Journal of Scientific Computing (2023) 96:44
https://doi.org/10.1007/s10915-023-02257-8

A Generalized Framework for Direct Discontinuous Galerkin Methods for Nonlinear Diffusion Equations

Mustafa Engin Danis[1,2] · **Jue Yan**[1]

Received: 19 October 2022 / Revised: 19 April 2023 / Accepted: 17 May 2023 /
Published online: 17 June 2023

Abstract
In this study, we propose a unified, general framework for the direct discontinuous Galerkin methods. In the new framework, the antiderivative of the nonlinear diffusion matrix is not needed. This allows a simple definition of the numerical flux, which can be used for general diffusion equations with no further modification. We also present the nonlinear stability analyses of the new direct discontinuous Galerkin methods and perform several numerical experiments to evaluate their performance. The numerical tests show that the symmetric and the interface correction versions of the method achieve optimal convergence and are superior to the nonsymmetric version, which demonstrates optimal convergence only for problems with diagonal diffusion matrices but loses order for even degree polynomials with a non-diagonal diffusion matrix. Singular or blow up solutions are also well captured with the new direct discontinuous Galerkin methods.

Keywords Discontinuous Galerkin method · Nonlinear diffusion equations · Stability · Convergence

Mathematics Subject Classification 65M12 · 65M60

1 Introduction

In this paper, we continue to study direct discontinuous Galerkin method [1] and other three versions of the direct discontinuous Galerkin (DDG) method [2–4] for solving the nonlinear diffusion equation

$$\frac{\partial U}{\partial t} = \nabla \cdot (A(U)\nabla U), \qquad (\mathbf{x}, t) \in \Omega \times (0, T), \tag{1}$$

✉ Jue Yan
jyan@iastate.edu

Mustafa Engin Danis
danis@iastate.edu

1 Department of Mathematics, Iowa State University, Ames 50011, USA

2 Department of Aerospace Engineering, Iowa State University, Ames 50011, USA

with the initial data $U_0(\mathbf{x}) = U(\mathbf{x}, 0)$. Nonlinear diffusion matrix $A(U) \in \mathbb{R}^{d\times d}$ can be anisotropic and none symmetric, but is assumed to have real and nonnegative eigenvalues. We adapt $\Omega \subset \mathbb{R}^d$ to denote the computational domain. In this study, we consider a 2-dimensional setting with $d = 2$. Our focus is to derive a generalized and unified DDG method for nonlinear diffusion Eq. (1) that can be easily extended and applied to system and multi dimensional cases.

The discontinuous Galerkin (DG) method was first introduced by Reed and Hill for neutron transport equations in 1973 [5]. However, it is after a series of papers by Cockburn, Shu et al. [6–9] that the DG method became the archetype of high order methods used in the scientific community. Essentially, the DG method is a finite element method, but with a discontinuous piecewise polynomial space defined for the numerical solution and test function in each element. Due to this property, the DG method has a smaller and more compact stencil compared to its continuous counterpart. Hence, the data structure required to implement DG methods is extremely local, which allows efficient parallel computing and hp-adaptation.

One of the key features of the DG method is that the communication between computational elements is established through a *numerical flux* defined at element interfaces. In this regard, the DG method bears a striking similarity to finite volume method, where a Riemann solver is employed to calculate the numerical flux. Therefore, the DG method enjoys the high-order polynomial approximations as a finite element method while benefiting from the characteristic decomposition of the wave propagation provided by Riemann solvers as a finite volume method. For this reason, the DG method has been successfully applied to hyperbolic problems, i.e. compressible Euler equations, in the last three decades, cf. [10–12].

On the other hand, for elliptic and parabolic problems, i.e. linear/nonlinear diffusion equations, the numerical flux must involve a proper definition for the solution gradient at element interfaces. It is, in fact, the variety of this definition that leads to several DG methods such as the interior penalty (IPDG) methods [13–15], the nonsymmetric interior penalty (NIPG) method [16, 17] and the symmetric interior penalty (SIPG) method [18–20]. Another important group of DG methods for solving diffusion problems include the method of Bassi and Rebay (BR) and its variations [21–24]; the local DG (LDG) method [25–27]; the method of Baumann and Oden (BO) [28, 29]; hybridized DG (HDG) method [30]. Recent works include the weakly over-penalized SIPG method [31]; weak DG [32] method and ultra weak DG method [33, 34]. For a review of these methods, we refer to [35, 36] and the references therein. Among the many DG methods mentioned, there is little discussion of nonlinear diffusion equations except Bassi and Rebay [21] and Local DG related methods [25, 30].

In addition to all the efforts mentioned above to devise a numerical flux for the diffusive terms in elliptic/parabolic equations, Liu and Yan [1] introduced Direct DG (DDG) method for nonlinear diffusion equations. Inspired by the exact trace formula corresponding to the solution of heat equation with a smooth initial data that contains a discontinuous point, Liu and Yan derived a simple formula for the numerical flux to compute the solution derivative at the element interface. Although DDG method is proven to converge to the exact solution optimally when measured in an energy norm, it suffers from an order loss in the L_2-norm when the solution space is approximated by even degree polynomials. In order to recover optimal convergence, Liu and Yan [2] developed the direct DG method with interface correction (DDGIC). In their subsequent studies, Yan and collaborators presented symmetric and nonsymmetric versions of the DDG method [3, 4]. Even though DDG methods degenerate to the IPDG method with piecewise constant and linear polynomial approximations, there exist a number of advantages with DDG methods for higher order approximations. For such advantages, we refer to the discussions on a third order bound preserving scheme in

[37], superconvergence to ∇U in [38, 39] and elliptic interface problems with different jump interface conditions in [40].

Despite the aforementioned favorable features, in the previous versions of DDG methods, the numerical flux definition is based on the antiderivative of the nonlinear diffusion matrix $A(U)$. However, this antiderivative might not exist if the diffusion matrix $A(U)$ is complicated enough. Therefore, the previous DDG methods are not applicable to nonlinear equations with such diffusion matrices. One important example where the diffusion matrix $A(U)$ cannot be integrated explicitly is the energy equation of compressible Navier–Stokes equations. A similar difficulty arises for the interface terms involving test function. This problem is addressed by defining a new direction vector on element interfaces, which depends on the nonlinear diffusion matrix $A(U)$ and geometric information of the interface. With the introduction of the nonlinear direction vector, the evaluation of the nonlinear numerical flux is greatly simplified. Interface terms can also be clearly defined with no ambiguity. Danis and Yan recently applied the method in [41] to solve compressible Navier–Stokes equations with DDGIC method. This treatment of a generic diffusion process opens up the possibility of a straightforward extension of all DDG versions to the complicated nonlinear diffusion equations, which motivates this study.

In this paper, the concept of the nonlinear direction vector is extended to all versions of DDG method in a generalized, unified framework. The new framework does not only address the problem of calculating the antiderivative of the diffusion matrix, but also provides an easy and practical recipe for using the DDG methods for general system of conservation laws. Moreover, interface terms of all versions of DDG methods are presented within a unified format that is clean and easy to be evaluated. Nonlinear stability analyses are presented for the new DDGIC, symmetric DDG and nonsymmetric DDG methods, and we investigate their performance in several numerical experiments. Since DDG methods degenerate to the IPDG method with low order approximations as mentioned, all numerical tests are conducted with high order polynomial approximations. In the numerical tests, optimal order of accuracy are obtained for DDGIC and symmetric DDG methods over uniform triangular meshes while a slight fraction of order loss is observed for nonsymmetric DDG method with even degree polynomial approximations. It is also shown that singular or blow up phenomena can be well captured under the new DDG framework.

Throughout the paper, we denote the exact solution of Eq. (1) by the uppercase U and the DG solution of Eq. (1) by the lowercase u. The rest of the paper is organized as follows. In Sect. 2, we briefly review the direct DG methods. In Sect. 3, the new methodology is described. In Sect. 4, nonlinear stability analysis are presented. Implementation details of the new methods are explained and several numerical examples are presented in Sect. 5. Finally, we draw our conclusions in Sect. 6.

2 A Review of Direct DG Methods

In this section, we will present a brief review of the original direct DG methods [1–4] and the required notation for later use.

We consider a shape regular triangular mesh partition $\mathcal{T}_h$ of Ω such that $\overline{\Omega} = \cup_{K\in\mathcal{T}_h} K$. For each element K, we denote the diameter of the inscribed circle by h_K. Furthermore, we define the numerical solution space as

$$\mathbb{V}_h^k := \{v \in L^2(\Omega) : v(x, y)\big|_K \in \mathbb{P}_k(K)\},$$

where $h = \max_K\{h_K\}$ and $\mathbb{P}_k(K)$ represents the space of polynomials of degree k in two dimensions. Note that this solution space is discontinuous across element interfaces. For this purpose, we adopt the following notation for the interface solution jump and average

$$[\![u]\!] = u^+ - u^-, \qquad \{\!\{u\}\!\} = \frac{u^+ + u^-}{2}, \quad \forall (x, y) \in \partial K,$$

where u^+ and u^- are the solution values calculated from the exterior and interior of the element K.

Next, for a test function $v \in \mathbb{V}_h^k$, we multiply Eq. (1) by v, integrate it by parts and apply the divergence theorem to obtain the weak form of Eq. (1). Along with the initial projection, the weak form is then given by

$$\int_K u_t v \, dxdy + \int_K A(u)\nabla u \cdot \nabla v \, dxdy - \int_{\partial K} \widehat{A(u)\nabla u} \cdot \mathbf{n} \, v \, ds = 0, \quad \forall v \in \mathbb{V}_h^k, \tag{2}$$

$$\int_K u(x, y, 0)v(x, y) \, dxdy = \int_K U_0(x, y)v(x, y) \, dxdy, \quad \forall v \in \mathbb{V}_h^k. \tag{3}$$

In Eq. (2), the volume integration is performed over individual elements K and the surface integral is performed over the element boundary ∂K. Here, $\mathbf{n}$ is the outward unit normal vector on ∂K. Furthermore, since $u(x, y, t)$ is discontinuous across the elements, $A(u)\nabla u$ is multi-valued on ∂K. For this reason, $A(u)\nabla u$ is written with a hat in the surface integral term in Eq. (2). In fact, $\widehat{A(u)\nabla u}$ is known as the numerical flux. The original DDG method [1] defines the numerical flux as

$$\widehat{a_{ij}(u)u_{x_j}} = \frac{\beta_0}{h_e}[\![b_{ij}(u)]\!]n_j + \{\!\{b_{ij}(u)_{x_j}\}\!\} + \beta_1 h_e [\![b_{ij}(u)_{x_1x_j}n_1 + b_{ij}(u)_{x_2x_j}n_2]\!], \tag{4}$$

where we denote by $a_{ij}(u)$ the ij component of the diffusion matrix $A(u)$. Here, $b_{ij}(u)$ are the components of the matrix $B(u)$. Basically, the components of $B(u)$ are the antiderivatives of $a_{ij}(u)$ and it is defined as $b_{ij}(u) = \int^u a_{ij}(s)ds$. In Eq. (4), h_e is the average of the element diameters sharing the edge ∂K, n_j are the components of the unit normal $\mathbf{n}$ for $j = 1, 2$, and the subscripts x_j denote the partial derivative with respect to the corresponding spatial coordinate axis for $j = 1, 2$. Furthermore, (β_0, β_1) is a pair of coefficients that affects the stability and optimal convergence of the DDG method. Along with Eqs. (2) and (4), the definition of the original direct DG method [1] is now completed.

It is well-known that the original DDG method loses an order for even degree polynomials [1]. This problem is fixed either by including a jump term for the test function or introducing a *numerical flux* for the test function. What determines the name of the corresponding DDG version is in fact how these additional terms are implemented.

2.1 The DDG Method with Interface Correction

The scheme formulation of the original DDGIC method [2] is given as

$$\int_K u_t v \, dxdy + \int_K A(u)\nabla u \cdot \nabla v \, dxdy - \int_{\partial K} \widehat{A(u)\nabla u} \cdot \mathbf{n} \, v \, ds + \int_{\partial K} [\![B(u)]\!]\{\!\{\nabla v\}\!\} \cdot \mathbf{n} \, ds = 0, \quad \forall v \in \mathbb{V}_h^k,$$

where the numerical flux $\widehat{A(u)\nabla u}$ is calculated using Eq. (4).

2.2 The Symmetric DDG Method

The scheme formulation of the original symmetric DDG method [3] is given by

$$\int_K u_t v \, dxdy + \int_K A(u)\nabla u \cdot \nabla v \, dxdy - \int_{\partial K} \widehat{A(u)\nabla u} \cdot \mathbf{n} \, v \, ds + \int_{\partial K} [\![B(u)]\!] \widehat{\nabla v} \cdot \mathbf{n} \, ds = 0, \quad \forall v \in \mathbb{V}_h^k.$$

As in the DDGIC method, the numerical flux $\widehat{A(u)\nabla u}$ is calculated by Eq. (4) while the numerical for the test function is given as

$$\begin{aligned} \widehat{v_x} &= \beta_0 \frac{[\![v]\!]}{h_e} n_1 + \{\{v_x\}\} + \beta_1 h_e [\![v_{xx} n_1 + v_{yx} n_2]\!], \\ \widehat{v_y} &= \beta_0 \frac{[\![v]\!]}{h_e} n_2 + \{\{v_y\}\} + \beta_1 h_e [\![v_{xy} n_1 + v_{yy} n_2]\!]. \end{aligned} \tag{5}$$

2.3 The Nonsymmetric DDG Method

The scheme formulation of the original nonsymmetric DDG method [4] is given by

$$\int_K u_t v \, dxdy + \int_K A(u)\nabla u \cdot \nabla v \, dxdy - \int_{\partial K} \widehat{A(u)\nabla u} \cdot \mathbf{n} \, v \, ds - \int_{\partial K} \widetilde{A(v)\nabla v} \cdot \mathbf{n} [\![u]\!] \, ds = 0, \quad \forall v \in \mathbb{V}_h^k.$$

Similar to the other DDG versions, the numerical flux $\widehat{A(u)\nabla u}$ is calculated by Eq. (4). The numerical flux for the test function is defined similarly but with a different penalty coefficient β_{0v}:

$$\widetilde{a_{ij}(v)v_{x_j}} = \frac{\beta_{0v}}{h_e} [\![b_{ij}(v)]\!] n_j + \{\{b_{ij}(v)_{x_j}\}\} + \beta_1 h_e [\![b_{ij}(v)_{x_1 x_j} n_1 + b_{ij}(v)_{x_2 x_j} n_2]\!]. \tag{6}$$

3 The New DDG Framework for Nonlinear Diffusion Equations

As can be seen in the previous section, the numerical flux definition of original DDG versions is based on calculating an antiderivative matrix $B(u)$ that is calculated according to

$$B(u) = \int^u A(s) ds.$$

A major drawback occurs when the components of the diffusion matrix $A(u)$ cannot be integrated explicitly. In such cases, none of the original DDG versions can be implemented. A striking example is the energy equation of compressible Navier–Stokes equations. This drawback limits the use of the original DDG versions only to simple applications where the antiderivative matrix $B(u)$ is available.

The new framework is based on the adjoint-property of inner product, which was used in the proof of a bound-preserving limiter with DDGIC method [37]. On the continuous level, the integrand of the surface integral in the weak form Eq. (2) can be rewritten as

$$A(u)\nabla u \cdot \mathbf{n} = \nabla u \cdot A(u)^T \mathbf{n}.$$

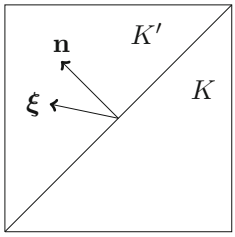


Fig. 1 The new direction vector $\boldsymbol{\xi}$

By applying the adjoint-property, we define a new direction vector $\boldsymbol{\xi}(u) = A(u)^T\mathbf{n}$. As shown in Fig. 1, the new direction vector is simply obtained by stretching/compressing and rotating the unit normal vector $\mathbf{n}$ through diffusion matrix $A(u)$. On the discrete level, the new direction vector can be calculated by

$$\boldsymbol{\xi}\left(\{\{u\}\}\right) = A(\{\{u\}\})^T\mathbf{n}. \tag{7}$$

The numerical flux can suitably be defined as

$$\widehat{A(u)\nabla u}\cdot\mathbf{n} = \widehat{\nabla u}\cdot\boldsymbol{\xi}\left(\{\{u\}\}\right),$$

where the numerical flux $\widehat{\nabla u} = (\widehat{u}_x, \widehat{u}_y)$ can be computed by the original DDG numerical flux formula for the heat equation [1]:

$$\begin{aligned}\widehat{u}_x &= \beta_0\frac{[\![u]\!]}{h_e}n_1 + \{\{u_x\}\} + \beta_1 h_e[\![u_{xx}n_1 + u_{yx}n_2]\!],\\ \widehat{u}_y &= \beta_0\frac{[\![u]\!]}{h_e}n_2 + \{\{u_y\}\} + \beta_1 h_e[\![u_{xy}n_1 + u_{yy}n_2]\!].\end{aligned} \tag{8}$$

Now, we reformulate all DDG versions for Eq. (1) according to the new framework: Find $u \in \mathbb{V}_h^k$ such that

$$\begin{aligned}&\int_K u_t v\,dxdy + \int_K A(u)\nabla u\cdot\nabla v\,dxdy\\ &\quad - \int_{\partial K}\widehat{\nabla u}\cdot\boldsymbol{\xi}(\{\{u\}\})v\,ds + \sigma\int_{\partial K}[\![u]\!]\widetilde{\nabla v}\cdot\boldsymbol{\xi}(\{\{u\}\})\,ds = 0, \quad \forall v\in\mathbb{V}_h^k,\end{aligned} \tag{9}$$

where $\sigma = 0$ for the basic DDG scheme, $\sigma = 1$ for DDGIC and symmetric DDG schemes, and $\sigma = -1$ for the nonsymmetric DDG scheme. Furthermore, we denote by $\widetilde{\nabla v} = (\widetilde{v}_x, \widetilde{v}_y)$ the numerical flux for the test function $v \in \mathbb{V}_h^k$. Along with the following definitions of the numerical flux for the test function, Eq. (9) defines the new DDG versions:
The baseline DDG scheme ($\sigma = 0$):

$$\widetilde{\nabla v} = \mathbf{0}. \tag{10}$$

The DDGIC scheme ($\sigma = 1$):

$$\widetilde{\nabla v} = \{\{\nabla v\}\}. \tag{11}$$

The symmetric DDG scheme ($\sigma = 1$):

$$\begin{aligned} \widetilde{v}_x &= \beta_0 \frac{[\![v]\!]}{h_e} n_1 + \{\{v_x\}\} + \beta_1 h_e [\![v_{xx} n_1 + v_{yx} n_2]\!], \\ \widetilde{v}_y &= \beta_0 \frac{[\![v]\!]}{h_e} n_2 + \{\{v_y\}\} + \beta_1 h_e [\![v_{xy} n_1 + v_{yy} n_2]\!]. \end{aligned} \tag{12}$$

The nonsymmetric DDG scheme ($\sigma = -1$):

$$\begin{aligned} \widetilde{v}_x &= \beta_{0v} \frac{[\![v]\!]}{h_e} n_1 + \{\{v_x\}\} + \beta_1 h_e [\![v_{xx} n_1 + v_{yx} n_2]\!], \\ \widetilde{v}_y &= \beta_{0v} \frac{[\![v]\!]}{h_e} n_2 + \{\{v_y\}\} + \beta_1 h_e [\![v_{xy} n_1 + v_{yy} n_2]\!]. \end{aligned} \tag{13}$$

In [37], the adjoint property of inner product was only used in the proof of positivity-preserving (Theorem 3.2). All numerical tests of [37] involving nonlinear diffusion equations were implemented with the original DDGIC scheme formulation [2].

3.1 Boundary Conditions

We briefly discuss the implementation of boundary conditions. For all edges falling on the domain boundaries, i.e. $\partial K \in \partial\Omega$, the second derivative term in the numerical flux Eq. (8) is dropped and only the solution jump and gradient average are explored to approximate the numerical flux on the domain boundary

$$\begin{aligned} \widehat{u}_x &= \beta_0 \frac{[\![u]\!]}{h_e} n_1 + \{\{u_x\}\}, \\ \widehat{u}_y &= \beta_0 \frac{[\![u]\!]}{h_e} n_2 + \{\{u_y\}\}. \end{aligned}$$

For periodic boundary conditions where the periodicity is given as $U(x, y) = U(x + L_x, y + L_y)$, the boundary condition for exterior " + " state is set from the interior " − " state of the periodic boundary such that

$$\begin{aligned} u^+(x, y) &= u^-(x + L_x, y + L_y), \\ \nabla u^+(x, y) &= \nabla u^-(x + L_x, y + L_y). \end{aligned}$$

For Dirichlet boundary condition where the solution on a part of the domain boundary is given, $U(x, y)|_{(x,y)\in\partial\Omega} = u_D(x, y)$, we simply set

$$\begin{aligned} u^+(x, y) &= u_D(x, y), \\ \nabla u^+(x, y) &= \nabla u^-(x, y). \end{aligned}$$

For Neumann boundary condition where the solution flux along the outward normal on a part of the domain boundary is given, i.e. $A(u)\nabla U(x, y) \cdot \mathbf{n}|_{(x,y)\in\partial\Omega} = u_N(x, y)$, we set

$$\widehat{A(u)\nabla u} \cdot \mathbf{n} = u_N(x, y).$$

We comment that there is no need to adapt direct DG method numerical flux formula (8) on the solution's gradient when Neumann type boundary condition is considered.

Remark 1 We summarize the main features and advantages of the generalized DDG methods (9) for nonlinear diffusion Eq. (1).

- Nonlinearity of the diffusion process goes into the corresponding new direction vector defined at the cell interface that greatly simplifies the implementation of the DDG method.
- Numerical flux for the solution's gradient ∇u can be approximated by the linear numerical flux formula of the original DDG. Since the solution's gradient is independent of the governing equation, this allows the code reuse for general nonlinear diffusion problems.
- The nonlinear direction vectors are further applied to define the nonlinear interface terms involving test function.

4 Nonlinear Stability of the New DDG Methods

In this section, we will discuss the nonlinear stability theory of the DDG methods developed in Sect. 3. To simplify the discussion, we will assume periodic boundary conditions on the domain boundaries. The important inequalities used in the proofs of the main theorems are discussed later in Appendix A.

We say that the DDG method is stable in L_2 sense if

$$\int_\Omega u^2(x, y, T)\, dxdy \le \int_\Omega U_0^2(x, y)\, dxdy, \quad \forall T \ge 0.$$

Note that the primal weak formulation of the new DDG methods is obtained by summing Eqs. (3) and (9) over all element $K \in \mathcal{T}_h$.

$$\int_\Omega u_t v\, dxdy + \mathbb{B}(u, v) = 0, \quad \forall v \in \mathbb{V}_h^k. \tag{14}$$

$$\int_\Omega u(x, y, 0)v(x, y)\, dxdy = \int_\Omega U_0(x, y)v(x, y)\, dxdy, \quad \forall v \in \mathbb{V}_h^k, \tag{15}$$

Here, $U_0(x, y)$ denotes the initial data and $\mathbb{B}(u, v)$ is given by

$$\mathbb{B}(u, v) = \int_\Omega A(u)\nabla u \cdot \nabla v\, dxdy + \sum_{e\in\mathcal{E}_h} \int_e \left([\![v]\!] \widehat{\nabla u} + \sigma [\![u]\!] \widetilde{\nabla v} \right) \cdot \boldsymbol{\xi}(\{\{u\}\})\, ds = 0. \tag{16}$$

where $\mathcal{E}_h = \cup_{K\in\mathcal{T}_h} \partial K$ represents the set of all element edges.

Theorem 2 (Stability of nonsymmetric DDG method) *Let the model parameter $\sigma = -1$ in the scheme formulation Eq. (9) that is equipped with the numerical flux for the gradient of the numerical solution Eq. (8) and the numerical flux for the gradient of the test function Eq. (13). If $\beta_0 \ge \beta_{0v}$, then we have*

$$\int_\Omega u^2(x, y, T)\, dxdy \le \int_\Omega U_0^2(x, y)\, dxdy, \quad \forall T \ge 0.$$

Proof By setting $u = v$, we integrate Eq. (14) with respect to time over $(0, T)$.

$$\frac{1}{2}\int_\Omega u^2(x, y, T)\, dxdy + \int_0^T \mathbb{B}(u, u)\, dt = \frac{1}{2}\int_\Omega u^2(x, y, 0)\, dxdy, \tag{17}$$

where

$$\int_0^T \mathbb{B}(u, u)\, dt = \int_0^T \int_\Omega A(u)\nabla u \cdot \nabla u\, dxdy\, dt + \sum_{e\in\mathcal{E}_h} \int_0^T \int_e [\![u]\!] (\widehat{\nabla u} - \widetilde{\nabla u}) \cdot \boldsymbol{\xi}(\{\{u\}\})\, ds\, dt. \tag{18}$$

Since $A(u)$ is positive definite, we have $A(u)\nabla u \cdot \nabla u = \nabla u^T A(u)\nabla u \geq 0$ and thus,

$$\int_0^T \int_\Omega A(u)\nabla u \cdot \nabla u \, dxdy \, dt \geq 0, \quad \forall T \geq 0. \tag{19}$$

Furthermore, we have that

$$\widehat{\nabla u} - \widetilde{\nabla u} = \frac{\beta_0^*}{h_e}[\![u]\!]\mathbf{n},$$

where $\beta_0^* = \beta_0 - \beta_{0v} \geq 0$ by the assumptions of the theorem. Recalling that $\boldsymbol{\xi}(\{\{u\}\}) = A(\{\{u\}\})^T \mathbf{n}$ and $A(u)$ is the positive definite, we can write

$$[\![u]\!](\widehat{\nabla u} - \widetilde{\nabla u}) \cdot \boldsymbol{\xi}(\{\{u\}\}) = \frac{\beta_0^*}{h_e}[\![u]\!]^2 \left(\mathbf{n} \cdot A(\{\{u\}\})^T \mathbf{n}\right) \geq 0.$$

Thus, we have

$$\int_0^T \int_e [\![u]\!](\widehat{\nabla u} - \widetilde{\nabla u}) \cdot \boldsymbol{\xi}(\{\{u\}\}) \, ds \, dt \geq 0, \quad \forall e \in \mathcal{E}_h, \ \forall T \geq 0. \tag{20}$$

Substituting Eqs. (19) and (20) into Eq. (18), and then Eq. (18) into Eq. (17), we obtain

$$\int_\Omega u^2(x, y, T) \, dxdy \leq \int_\Omega u^2(x, y, 0) \, dxdy, \quad \forall T \geq 0.$$

Finally, we apply the Schwarz inequality to the initial projection Eq. (15) with $v = u(x, y, 0)$ and obtain

$$\int_\Omega u^2(x, y, 0) \, dxdy \leq \int_\Omega U_0^2(x, y) \, dxdy,$$

which completes the proof. □

Theorem 3 (Stability of symmetric DDG method) *Assume that $A(u)$ is a positive definite matrix with positive eigenvalues and there exists $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{V}_h^k$. Furthermore, let the model parameter $\sigma = 1$ in the scheme formulation Eq. (9) that is equipped with the numerical flux for the gradient of the numerical solution Eq. (8) and the numerical flux for the gradient of the test function Eq. (12). If $\beta_0 \geq C(\beta_1)k^2 \left(\frac{\gamma^*}{\gamma}\right)^2 \beta_1^2$, then we have*

$$\frac{1}{2}\int_\Omega u^2(x, y, T) \, dxdy + \left(1 - C(\beta_1)k^2 \left(\frac{\gamma^*}{\gamma}\right)^2 \frac{\beta_1^2}{\beta_0}\right) \int_0^T \int_\Omega A(u)\nabla u \cdot \nabla u \, dxdy \, dt$$
$$\leq \frac{1}{2}\int_\Omega U_0^2(x, y) \, dxdy,$$

where $C(\beta_1) = C_1/2\beta_1^2 + 2C_2 > 0$ and $C_1, C_2 > 0$ are constants.

Proof By setting $u = v$, Eq. (14) becomes

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2 \, dxdy + \mathbb{B}(u, u) = 0, \tag{21}$$

where

$$\mathbb{B}(u, u) = \int_\Omega A(u)\nabla u \cdot \nabla u \, dxdy + 2\sum_{e \in \mathcal{E}_h} \int_e [\![u]\!]\widehat{\nabla u} \cdot \boldsymbol{\xi}(\{\{u\}\}) \, ds = 0. \tag{22}$$

Note that

$$[\![u]\!]\widehat{\nabla u}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) = \frac{\beta_0}{h}[\![u]\!]^2\mathbf{n}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) + [\![u]\!]\{\!\{\nabla u\}\!\}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) + \beta_1 h[\![u]\!][\![\nabla(\nabla u\cdot\mathbf{n})]\!]\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}).$$

Therefore, invoking Lemmas 10 to 12, we get

$$\begin{aligned}
&\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\widehat{\nabla u}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \\
&\quad\geq \sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{h}\left\|[\![u]\!]\right\|^2_{L^2(e)} \\
&\qquad-\sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)} - \sum_{K\in\mathcal{T}_h} C_1\frac{(\gamma^*k)^2}{4\gamma\beta_0}\|\nabla u\|^2_{L^2(K)} \\
&\qquad-\sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)} - \sum_{K\in\mathcal{T}_h} C_2\frac{(\gamma^*\beta_1k)^2}{\gamma\beta_0}\|\nabla u\|^2_{L^2(K)} \\
&\quad= -\sum_{K\in\mathcal{T}_h} C_1\frac{(\gamma^*k)^2}{4\gamma\beta_0}\|\nabla u\|^2_{L^2(K)} - \sum_{K\in\mathcal{T}_h} C_2\frac{(\gamma^*\beta_1k)^2}{\gamma\beta_0}\|\nabla u\|^2_{L^2(K)} \\
&\quad= -\sum_{K\in\mathcal{T}_h}\frac{1}{2}\left(\frac{C_1}{2\beta_1^2}+2C_2\right)\frac{(\gamma^*\beta_1k)^2}{\gamma\beta_0}\|\nabla u\|^2_{L^2(K)}.
\end{aligned}$$

This can be rewritten as

$$\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\widehat{\nabla u}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \geq -\frac{C(\beta_1)k^2}{2}\frac{(\gamma^*)^2}{\gamma}\frac{\beta_1^2}{\beta_0}\sum_K\|\nabla u\|^2_{L^2(K)}.$$

where $C(\beta_1) = C_1/2\beta_1^2 + 2C_2$. Next, we use the assumption on the eigenvalues of $A(u)$:

$$\sum_K\|\nabla u\|^2_{L^2(K)} \leq \frac{1}{\gamma}\sum_K\int_K \nabla u\cdot A(u)\nabla u\,dxdy,$$

and obtain

$$\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\widehat{\nabla u}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \geq -\frac{C(\beta_1)k^2}{2}\left(\frac{\gamma^*}{\gamma}\right)^2\frac{\beta_1^2}{\beta_0}\sum_K\int_K\nabla u\cdot A(u)\nabla u\,dxdy. \quad (23)$$

Substituting this Eq. (23) into Eq. (22), and then, Eq. (22) into Eq. (21) gives

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2\,dxdy + \left(1 - C(\beta_1)k^2\left(\frac{\gamma^*}{\gamma}\right)^2\frac{\beta_1^2}{\beta_0}\right)\int_\Omega A(u)\nabla u\cdot\nabla u\,dxdy \leq 0. \quad (24)$$

Lastly, we integrate Eq. (24) over $(0, T)$ and recall

$$\int_\Omega u^2(x,y,0)\,dxdy \leq \int_\Omega U_0^2(x,y)\,dxdy,$$

from the proof of Theorem 2. This completes the proof provided that we have $\beta_0 \geq C(\beta_1)k^2\left(\frac{\gamma^*}{\gamma}\right)^2\beta_1^2$. □

Theorem 4 (Stability of DDGIC method) *Assume that $A(u)$ is a positive definite matrix with positive eigenvalues and there exists $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{V}_h^k$. Furthermore, let the model parameter $\sigma = 1$ in the scheme formulation Eq. (9) that is equipped with the numerical flux for the gradient of the numerical solution Eq. (8) and the numerical flux for the gradient of the test function Eq. (11). If $\beta_0 \geq C(\beta_1)k^2 \left(\frac{\gamma^*}{\gamma}\right)^2 \beta_1^2$, then we have*

$$\frac{1}{2}\int_\Omega u^2(x,y,T)\,dxdy + \left(1 - C(\beta_1)k^2\left(\frac{\gamma^*}{\gamma}\right)^2\frac{\beta_1^2}{\beta_0}\right)\int_0^T\int_\Omega A(u)\nabla u\cdot\nabla u\,dxdy\,dt \leq \frac{1}{2}\int_\Omega U_0^2(x,y)\,dxdy.$$

where $C(\beta_1) = C_1/\beta_1^2 + C_2 > 0$ and $C_1, C_2 > 0$ are constants.

Proof The proof is very similar to the proof of Theorem 3. Therefore, we will only lay out the sketch of the proof. By setting $u = v$, Eq. (14) becomes

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2\,dxdy + \mathbb{B}(u,u) = 0,$$

where

$$\mathbb{B}(u,u) = \int_\Omega A(u)\nabla u\cdot\nabla u\,dxdy + \sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\left(\widehat{\nabla u} + \{\!\{\nabla u\}\!\}\right)\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds = 0.$$

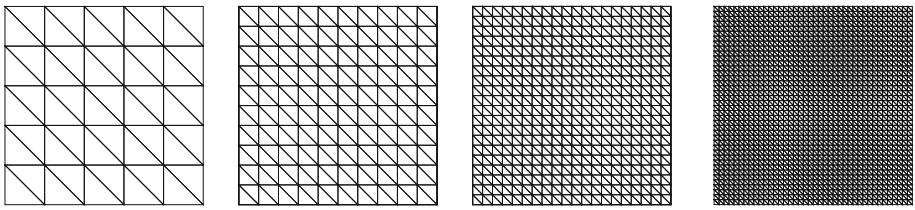
Note that

$$[\![u]\!]\left(\widehat{\nabla u} + \{\!\{\nabla u\}\!\}\right)\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) = \frac{\beta_0}{h}[\![u]\!]^2\mathbf{n}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) + 2[\![u]\!]\{\!\{\nabla u\}\!\}\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}) + \beta_1 h[\![u]\!][\![\nabla(\nabla u\cdot\mathbf{n})]\!]\cdot\boldsymbol{\xi}(\{\!\{u\}\!\}).$$

As in the proof of Theorem 3, we invoke Lemmas 10 to 12:

$$\begin{aligned}
&\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\left(\widehat{\nabla u} + \{\!\{\nabla u\}\!\}\right)\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \\
&\geq \sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{h}\left\|[\![u]\!]\right\|_{L^2(e)}^2 \\
&\quad - \sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|_{L^2(e)}^2 - \sum_{K\in\mathcal{T}_h} C_1\frac{(\gamma^* k)^2}{\gamma\beta_0}\|\nabla u\|_{L^2(K)}^2 \\
&\quad - \sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|_{L^2(e)}^2 - \sum_{K\in\mathcal{T}_h} C_2\frac{(\gamma^*\beta_1 k)^2}{\gamma\beta_0}\|\nabla u\|_{L^2(K)}^2 \\
&= -C(\beta_1)k^2\frac{(\gamma^*)^2}{\gamma}\frac{\beta_1^2}{\beta_0}\sum_K\|\nabla u\|_{L^2(K)}^2,
\end{aligned}$$

where $C(\beta_1) = C_1/\beta_1^2 + C_2$. Then, following the same lines of steps as in the proof of Theorem 3 will lead to the desired result. □



(a) h (b) $h/2$ (c) $h/4$ (d) $h/8$

Fig. 2 The set of structured meshes with $h = L/5$, where L is the square domain size

5 Numerical Examples

In this section, the results of several numerical examples are presented. A set of uniform triangular meshes is used to discretize the computational domain Ω as shown in Fig. 2. In all simulations, we set $\beta_0 = (k+1)^2$, $\beta_{0v} = \frac{\beta_0}{2}$ and $\beta_1 = \frac{1}{2k(k+1)}$ in the numerical flux definitions. The time integration is performed by a third order explicit strong stability-preserving (SSP) Runge–Kutta scheme [42]. Unless stated otherwise, the time step size Δt is determined by the following Courant-Friedrichs-Levy (CFL) rule

$$\mu \frac{\Delta t}{\min_K h_K^2} < \lambda,$$

where $\lambda = \frac{1}{2k+1}$ is the CFL number taken for P_k polynomials (refer to Table 2.2 in [43]). In this paper, unless stated, we choose a uniform CFL value of $\lambda = 0.1$ for all P_2, P_3 and P_4 approximations. Here μ is the diffusion constant. Since positivity-preserving limiter is used in Example 5.5 and Example 5.6, a more restrictive time step size is applied through all numerical tests

$$\mu \frac{\Delta t}{\min_K h_K^2} < \omega_k \lambda, \tag{25}$$

where ω_k depending on the degree k is the minimum quadrature weight to guarantee positivity [44]. We highlight that such polynomial point values used to evaluate positivity limiter are different to those in [37, 39]. Nevertheless, the quadrature points [44] applied to Example 5.5 and Example 5.6 for positivity all work well. The focus of the article is to verify the effectiveness of the new numerical flux formula and the accuracy of the new direct DG methods, we adapt a small time step size (25) to reduce the errors in time and leave spatial discretization errors being the dominant error.

The convergence rates are reported in the L_2 and L_∞ norms. To do that, we employ the $(k+1)th$ order quadrature rule to calculate the L_2-error while the L_∞-errors are measured using 361 points generated by the same quadrature rule in each element.

Example 5.1 In this example, we consider the heat equation

$$\frac{\partial U}{\partial t} = \mu \Delta U,$$

with periodic boundary conditions on $\Omega = [0,1] \times [0,1]$ where μ is a constant diffusion coefficient. Note that the diffusion matrix $A(u)$ is given as

$$A(u) = \mu \mathbb{I},$$

Table 1 L_2 errors for Example 5.1 at $T = 1$

	L_2 errors and orders for *the DDGIC method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.47E−03	3.10E−04	3.00	3.88E−05	3.00	4.85E−06	3.00
$k = 3$	1.87E−04	1.17E−05	4.00	7.29E−07	4.00	4.55E−08	4.00
$k = 4$	1.16E−05	3.64E−07	5.00	1.14E−08	5.00	3.55E−10	5.00
	L_2 errors and orders for *the symmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.82E−03	3.55E−04	2.99	4.45E−05	2.99	5.56E−06	3.00
$k = 3$	2.10E−04	1.30E−05	4.01	8.10E−07	4.00	5.06E−08	4.00
$k = 4$	1.28E−05	4.02E−07	4.99	1.26E−08	5.00	3.93E−10	5.00
	L_2 errors and orders for *the nonsymmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.29E−03	2.88E−04	2.99	3.73E−05	2.95	5.20E−06	2.84
$k = 3$	1.85E−04	1.15E−05	4.01	7.21E−07	4.00	4.50E−08	4.00
$k = 4$	1.13E−05	3.66E−07	4.95	1.26E−08	4.86	5.12E−10	4.62

where $\mathbb{I} \in \mathbb{R}^{2\times 2}$ is the identity matrix. The initial condition for this example is obtained from the following exact solution at $t = 0$

$$U(x, y, t) = e^{-8\pi^2 \mu t} \cos(2\pi(x + y)).$$

In this example, we set $\mu = 0.01$ and carry out direct DG approximations up to $T = 1$. In addition, all quadrature rules are exact up to polynomials of degree $2k + 1$. On the coarsest mesh h, time step values of $\Delta t = 9.21 \times 10^{-4}$, 2.77×10^{-4} and 1.0×10^{-4} for $k = 2, 3, 4$ are applied. We highlight that these small time step sizes are used in all accuracy tests for linear problems.

The L_2 and L_∞ errors are listed in Tables 1 and 2, respectively. We observe that optimal $(k + 1)th$ order convergence is obtained by the DDGIC and symmetric DDG methods. For nonsymmetric DDG method, a slight order loss is observed with even degree polynomial approximations, i.e. $k = 2$ and $k = 4$. Recall in [4] for one-dimensional problems and on uniform mesh, optimal convergence order is observed for the nonsymmetric DDG method.

Example 5.2 In this numerical test, we now consider an anisotropic diffusion equation with mixed derivatives

$$\frac{\partial U}{\partial t} = \mu \left(2U_{xx} + 3U_{xy} + 3U_{yy}\right),$$

with periodic boundary conditions on $\Omega = [0, 1] \times [0, 1]$. Note that this equation is still linear, i.e. diffusion matrix $A(u)$ is a constant coefficient matrix. Moreover, it can be written in a nonsymmetric form

$$A(u) = \mu \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}.$$

Table 2 L_∞ errors for Example 5.1 at $T = 1$

	L_∞ errors and orders for *the DDGIC method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	8.58E−03	1.03E−03	3.06	1.31E−04	2.97	1.65E−05	2.99
$k = 3$	7.77E−04	5.40E−05	3.85	3.43E−06	3.98	2.17E−07	3.98
$k = 4$	4.63E−05	1.35E−06	5.10	4.23E−08	4.99	1.32E−09	5.00
	L_∞ errors and orders for *the symmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	6.15E−03	7.01E−04	3.13	8.96E−05	2.97	1.13E−05	2.99
$k = 3$	6.18E−04	4.36E−05	3.83	2.83E−06	3.94	1.79E−07	3.98
$k = 4$	3.39E−05	1.01E−06	5.07	3.26E−08	4.96	1.03E−09	4.99
	L_∞ errors and orders for *the nonsymmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	1.14E−02	1.39E−03	3.03	1.83E−04	2.92	2.32E−05	2.98
$k = 3$	9.58E−04	6.56E−05	3.87	4.10E−06	4.00	2.59E−07	3.98
$k = 4$	6.04E−05	1.86E−06	5.02	6.08E−08	4.94	1.96E−09	4.96

A nonsymmetric diffusion matrix is chosen to see how the new DDG methods would behave in such a setting. The initial condition is set from the following exact solution at $t = 0$

$$U(x, y, t) = e^{-32\pi^2 \mu t} \cos(2\pi y) \cos(4\pi x - 2\pi y). \tag{26}$$

As in the previous example, we set $\mu = 0.01$ and output the approximations at $T = 1$, and all quadrature rules are exact up to polynomials of degree $2k + 1$.

The L_2 and L_∞ errors are listed in Tables 3 and 4, respectively. The DDGIC and symmetric DDG methods behave similarly in all cases with optimal convergence order obtained. The nonsymmetric DDG method demonstrates optimal $(k+1)$ convergence only for $k = 3$ while it loses an order for even degree polynomials, which is clearer compared to Example 5.1.

Remark 5 This problem has been also tested equivalently with the diffusion matrix

$$A(u) = \mu \begin{pmatrix} 2 & 1.5 \\ 1.5 & 3 \end{pmatrix},$$

which is symmetric-positive definite. We note that the performance of the new DDG methods does not change with this diffusion matrix and the same results in Tables 3 and 4 are obtained. Therefore, those results are not reported.

Example 5.3 In this example, we consider *the porous medium equation*

$$\frac{\partial U}{\partial t} = \mu \Delta (U^\gamma), \tag{27}$$

where γ is a model parameter. Note that this equation models a nonlinear diffusion process for $\gamma \neq 1$, i.e. coefficients of the diffusion matrix $A(u)$ are functions of u:

$$A(u) = \mu \begin{pmatrix} \gamma u^{\gamma-1} & 0 \\ 0 & \gamma u^{\gamma-1} \end{pmatrix} = \mu \gamma u^{\gamma-1} \mathbb{I},$$

Table 3 L_2 errors for Example 5.2 at $T = 1$

	L_2 errors and orders for *the DDGIC method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	1.34E−02	2.59E−03	2.37	2.14E−04	3.60	1.76E−05	3.61
$k = 3$	3.37E−03	1.16E−04	4.87	5.31E−06	4.45	3.12E−07	4.09
$k = 4$	4.62E−04	1.19E−05	5.29	3.73E−07	4.99	1.18E−08	4.99
	L_2 errors and orders for *the symmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	1.41E−02	3.01E−03	2.22	2.56E−04	3.56	1.96E−05	3.71
$k = 3$	3.68E−03	1.27E−04	4.86	5.51E−06	4.53	3.20E−07	4.11
$k = 4$	4.92E−04	1.22E−05	5.33	3.88E−07	4.98	1.22E−08	4.99
	L_2 errors and orders for *the nonsymmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	1.17E−02	2.47E−03	2.25	4.34E−04	2.51	9.21E−05	2.24
$k = 3$	3.67E−03	2.00E−04	4.20	1.07E−05	4.23	6.35E−07	4.07
$k = 4$	5.70E−04	1.15E−05	5.63	5.70E−07	4.34	3.47E−08	4.04

Table 4 L_∞ errors for Example 5.2 at $T = 1$

	L_∞ errors and orders for *the DDGIC method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.28E−02	4.33E−03	2.40	4.03E−04	3.43	4.04E−05	3.32
$k = 3$	7.81E−03	3.68E−04	4.41	2.18E−05	4.08	1.41E−06	3.96
$k = 4$	1.30E−03	3.48E−05	5.22	1.22E−06	4.84	3.97E−08	4.94
	L_∞ errors and orders for *the symmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.38E−02	4.96E−03	2.26	4.71E−04	3.40	4.54E−05	3.37
$k = 3$	8.33E−03	4.03E−04	4.37	2.03E−05	4.31	1.28E−06	3.99
$k = 4$	1.37E−03	3.61E−05	5.25	1.27E−06	4.82	4.17E−08	4.94
	L_∞ errors and orders for *the nonsymmetric DDG method*						
	h	$h/2$	Order	$h/4$	Order	$h/8$	Order
$k = 2$	2.13E−02	4.40E−03	2.28	8.01E−04	2.46	1.69E−04	2.25
$k = 3$	8.39E−03	5.87E−04	3.84	4.10E−05	3.84	2.66E−06	3.95
$k = 4$	1.49E−03	3.44E−05	5.43	1.50E−06	4.52	6.94E−08	4.44

where $\mathbb{I} \in \mathbb{R}^{2\times 2}$ is the identity matrix. The diffusion matrix is diagonal, but for $\gamma > 1$, Eq. (27) becomes highly nonlinear. In order to assess the performance of the new DDG methods on a highly nonlinear diffusion problem and measure the convergence rate, we solve Eq. (27) on $\Omega = [0, 1] \times [0, 1]$ with $\gamma = 3$ and employ the method of manufactured solutions by

Table 5 L_2 errors for Example 5.3 at $T = 1$

	L_2 errors and orders for *the DDGIC method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	3.19E−03	4.02E−04	2.99	5.00E−05	3.01
$k = 3$	2.46E−04	1.29E−05	4.26	7.48E−07	4.10
$k = 4$	1.74E−05	5.63E−07	4.95	1.54E−08	5.20
	L_2 errors and orders for *the symmetric DDG method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	3.26E−03	4.37E−04	2.90	5.64E−05	2.95
$k = 3$	2.87E−04	1.43E−05	4.33	8.33E−07	4.10
$k = 4$	1.85E−05	6.12E−07	4.91	1.69E−08	5.18
	L_2 errors and orders for *the nonsymmetric DDG method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	2.93E−03	4.41E−04	2.73	5.19E−05	3.09
$k = 3$	2.09E−04	1.31E−05	3.99	8.54E−07	3.94
$k = 4$	1.72E−05	5.23E−07	5.04	1.50E−08	5.13

enforcing the solution

$$U(x, y, t) = e^{-8\pi^2 \mu t} \sin(2\pi(x + y)).$$

The initial condition for this problem is obtained from this manufactured solution at $t = 0$. Also, note that the boundary conditions are periodic. Furthermore, we set $\mu = 0.01$ and conduct approximations to $T = 1$, and employ quadrature rules that are exact up to polynomials of degree $4k+1$. On the coarsest mesh h, time step values of $\Delta t = 10^{-4}$, 1.94×10^{-5} and 5.31×10^{-6} for $k = 2, 3, 4$ are applied, respectively.

The L_2 and L_∞ errors are listed in Tables 5 and 6. We observe that all DDG methods converge with optimal $(k+1)th$ order accuracy for all reported cases. Although this problem is nonlinear, the optimal convergence might be due to the diagonal structure of the nonlinear diffusion matrix $A(u)$.

Example 5.4 In this example, we consider the same equation as in Example 5.3 but with the model coefficient $\gamma = 2$ on $\Omega = [-10, 10] \times [-10, 10]$. Here, we investigate the performance of the new DDG methods for two initially disconnected, merging bumps which are defined by the following:

$$U_0(x, y) = \begin{cases} e^{\frac{-1}{6-(x-2)^2-(y+2)^2}}, & (x-2)^2 - (y+2)^2 < 6 \\ e^{\frac{-1}{6-(x+2)^2-(y-2)^2}}, & (x+2)^2 - (y-2)^2 < 6 \\ 0, & \text{otherwise.} \end{cases}$$

Note that $U(x, y, t) = 0$ on $\partial\Omega$ for $t \geq 0$. In this case, we solve the problem on $h/16$ mesh along with $\mu = 1$ and up to $T = 4$. In Fig. 3, we present a third order ($k = 2$) symmetric DDG solution. The solutions corresponding to other DDG versions are similar, hence they are not included. We observe that bumps are diffused quickly and merged with finite time. The results are in good agreement with those in literature [3, 4, 45].

Table 6 L_∞ errors for Example 5.3 at $T = 1$

	L_∞ errors and orders for *the DDGIC method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	2.42E−02	3.87E−03	2.65	5.33E−04	2.86
$k = 3$	2.29E−03	7.78E−05	4.88	3.45E−06	4.49
$k = 4$	2.60E−04	1.06E−05	4.62	3.50E−07	4.92
	L_∞ errors and orders for *the symmetric DDG method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	2.29E−02	3.54E−03	2.69	4.97E−04	2.83
$k = 3$	2.05E−03	5.61E−05	5.19	2.78E−06	4.33
$k = 4$	2.45E−04	1.02E−05	4.59	3.41E−07	4.90
	L_∞ errors and orders for *the nonsymmetric DDG method*				
	h	$h/2$	Order	$h/4$	Order
$k = 2$	2.69E−02	4.46E−03	2.59	6.06E−04	2.88
$k = 3$	2.45E−03	6.97E−05	5.14	4.32E−06	4.01
$k = 4$	2.79E−04	1.15E−05	4.61	3.70E−07	4.95

Example 5.5 In this example, we continue to study the same equation as in Example 5.3 but with the model coefficient $\gamma = 2$ on $\Omega = [-1, 1] \times [-1, 1]$. Zero boundary condition $U(x, y, t) = 0$ on $\partial\Omega$ for $t \geq 0$ is applied and the initial condition is defined as

$$U_0(x, y) = \begin{cases} 1, & (x, y) \in [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}] \\ 0, & \text{otherwise.} \end{cases}$$

We set $\mu = 1$, solve the problem on $h/16$ mesh and run the simulations to $T = 0.005$.

Since the initial condition is discontinuous, a maximum-principle-satisfying (MPS) limiter is employed to guarantee the stability of the numerical solutions obtained by new DDG methods. The limiter in [39], which is similar to the linear scaling limiter in [46], is implemented to keep the numerical solution in $0 \leq u(x, y, t) \leq 1$ for $t \geq 0$. Note that the CFL condition Eq. (25) is still in use, and the DDG parameters β_0 and β_1 are kept the same.

In Fig. 4, the numerical solution obtained by the DDGIC method for a third order approximation ($k = 2$) is shown. We observe that the numerical solution diffuses out smoothly in a stable manner and is in good agreement with Example 5.7 of [47].

Example 5.6 In this example, we consider

$$\frac{\partial U}{\partial t} = \mu \left(2U_{xx} + 3(U^{1.5})_{yy}\right) + U^2, \tag{28}$$

with the initial condition $U_0(x, y) = 200 \sin(\pi x) \sin(\pi y)$ on $\Omega = [0, 1] \times [0, 1]$. We apply the homogeneous Dirichlet boundary condition. Due to the last term on the right-hand side of Eq. (28), the solution eventually blows up. If a positivity-preserving limiter is not used, the numerical solution immediately becomes negative in the region near the initial bump. This leads to an execution error due to square-root of a negative number. Therefore, the linear scaling limiter of [39] (similar to the limiter in [46]) is employed to maintain the stability

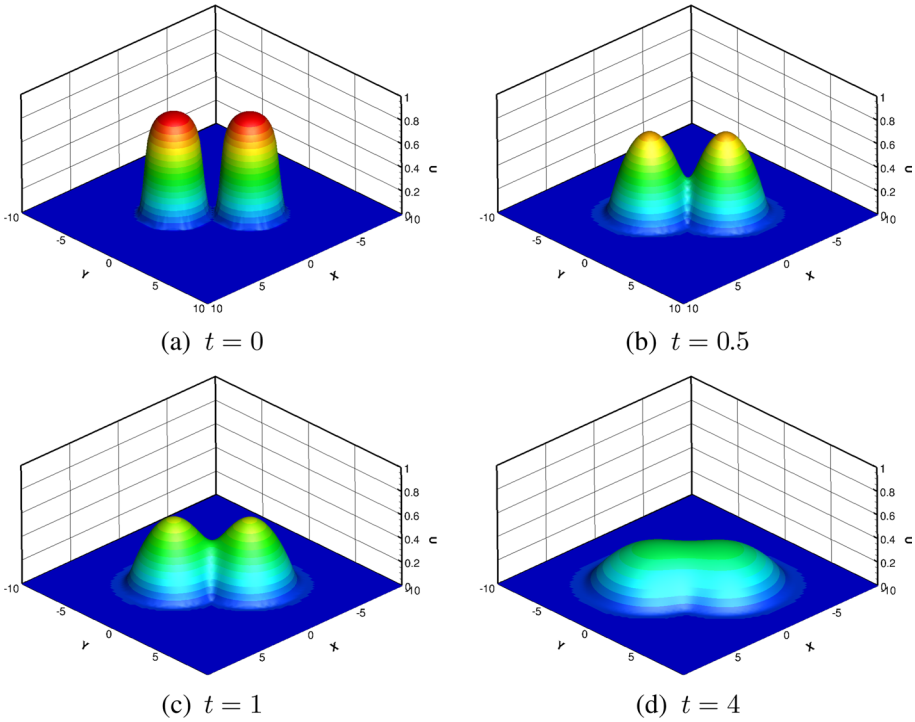


Fig. 3 The contour plot corresponding to the third ($k = 2$) order symmetric DDG solution for Example 5.4. 17 equally spaced contours are used between 0 and 0.8

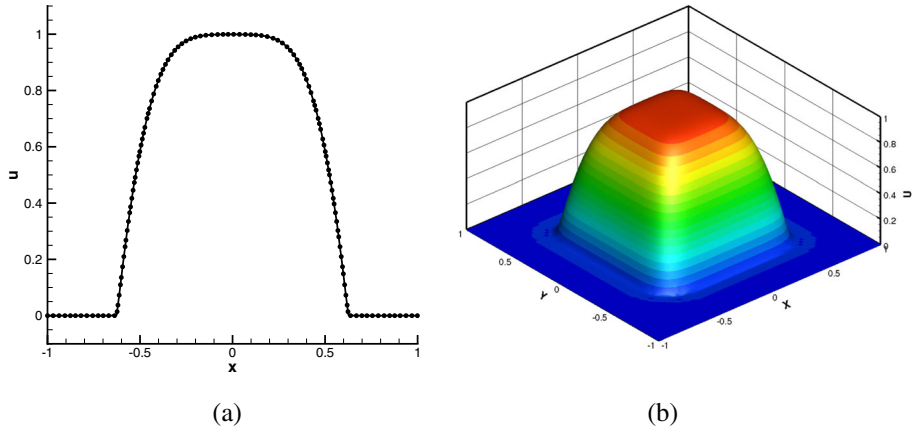


Fig. 4 The third ($k = 2$) order numerical solution obtained by DDGIC with MPS limiter for Example 5.4 at $T = 0.005$. **a** The solution along $y = 0$, **b** surface plot. 19 equally spaced contour levels are used between 0 and 1

of the numerical approximation. However, the limiter's lower bound is set to 0 to preserve the positivity of the numerical solution. Furthermore, since the solution does not have an upper bound due to the source term, the CFL condition in Eq. (25) is not able to maintain the

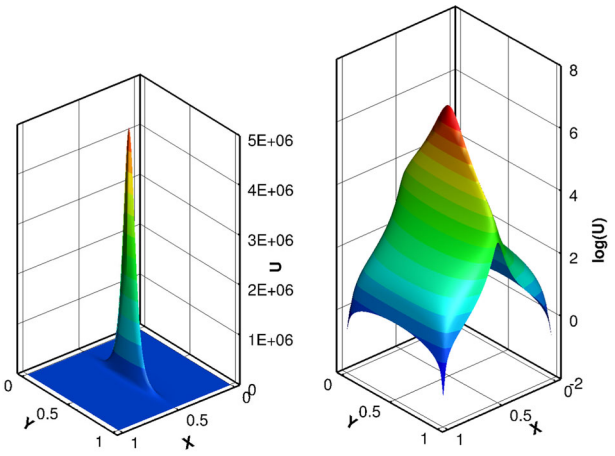


Fig. 5 The third ($k = 2$) order numerical solution obtained by DDGIC with MPS limiter for Example 5.6 at $T = 1.81552 \times 10^{-2}$

stability of the nonlinear diffusion process. Therefore, we follow [48] and modify Eq. (25) as

$$\mu \frac{\Delta t}{\min_K h_K^2} < \min\left(\omega_k \lambda, \frac{1}{\max_K u}\right).$$

The modified CFL condition adaptively decreases Δt while the solution u gets bigger. As in the previous example, the DDG parameters β_0, β_1 are kept the same. The $h/16$ mesh is used to solve this problem. We set the CFL value as $\lambda = 0.01$ and choose $\mu = 1$, and the quadrature rules are exact up to polynomials of degree $(2k + 1)th$. As for the stopping criterion, we follow [48] and take $\Delta t < 10^{-13}$. However, it is worth emphasizing that the focus of this paper is not to design a positivity-preserving limiter. Instead, we simply explore the performance of the new DDG methods in computationally challenging cases. We follow [49] and restart the Runge-Kutta time step with $\Delta t/2$ when it is no longer possible to maintain the positivity of the numerical solution in one or more cells.

For a third ($k = 2$) order solution, we observe that the restart algorithm is only activated at the last time-step at $t = 1.8155 \times 10^{-2}$ for all DDG versions. However, it turns out that the time step becomes smaller than the machine precision ϵ, i.e. $\Delta t < \epsilon$, during the restart procedure. This means that it is no longer possible to continue time-stepping without violating the positivity in one or more cells. Therefore, we denote $t = 1.81552 \times 10^{-2}$ as the blow-up time. Note that this value is slightly smaller than the blow-up time $t = 1.82378 \times 10^{-2}$ reported in [48]. Thus, we obtain slightly lower values for $\max_K u_K$. In Fig. 5, we show the simulation for a third order ($k = 2$) numerical solution obtained by the DDGIC method when the solution blows up at $t = 1.81552 \times 10^{-2}$. Since the results obtained by other DDG versions are similar, they are not shown. We observe that the numerical solution does not have oscillations and qualitatively similar to that in [48].

6 Concluding Remarks

In this paper, we have unified the new framework for direct discontinuous Galerkin (DDG) methods by extending the new DDGIC method [41] to the symmetric and the nonsymmetric DDG versions. Unlike their original counterparts, the new DDG methods do not require evaluating an antiderivative of the nonlinear diffusion matrix. By constructing a new direction vector at each element interface, a linear numerical flux is used regardless of the problem type. Furthermore, the nonlinear linear stability theory of the new methods has been developed and several numerical experiments have been conducted to perform the error analysis. In all numerical examples, the DDGIC and symmetric DDG methods demonstrated optimal $(k+1)th$ order convergence and their performances were assessed to be equivalent. On the other hand, the performance of the nonsymmetric DDG method varied in all numerical examples. The nonsymmetric DDG method achieved optimal $(k+1)th$ order convergence for odd degree polynomials in all examples while an order loss was observed with the even degree polynomials except for the cases with diagonal diffusion matrices. In short, a high-order accuracy is achieved by all new DDG methods. However, the DDGIC and symmetric DDG methods were found to be superior to the nonsymmetric version.

Funding Research work of the author is supported by National Science Foundation grant DMS-1620335 and Simons Foundation grant 637716.

A. Important Inequalities

In this section, we discuss important inequalities used in the proofs of the stability analysis for symmetric DDG and DDGIC methods.

Lemma 6 (Young's inequality) *Suppose that* $a, b \geq 0$, $1 < p, q, \infty$, *and that* $\frac{1}{p} + \frac{1}{q} = 1$. *Then, we have that*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

A corollary to Lemma 6 can be obtained by considering $ab = (a\epsilon^{1/p})(b/\epsilon^{1/p})$ for $\epsilon > 0$.

Corollary 7 *Suppose that* $a, b \geq 0$, $1 < p, q, \infty$, *and that* $\frac{1}{p} + \frac{1}{q} = 1$. *Furthermore, if* $\epsilon > 0$, *then*

$$ab \leq \frac{\epsilon a^p}{p} + \frac{b^q}{q\epsilon^{q/p}}.$$

We will also recall the following lemma due to Ern and Guermord [50]:

Lemma 8 *Let* K *be an element in the mesh partition* $\mathcal{T}_h$ *of the computational domain* $\Omega \subset \mathbb{R}^d$, $v \in \mathbb{V}_h^k$ *and* $l \in \mathbb{N}$. *There exists a constant* $C > 0$ *for any non-negative integer* $m \leq l$ *such that*

$$|v|_{W^{l,p}(K)} \leq C h_K^{m-l+d\left(\frac{1}{p}-\frac{1}{r}\right)} |v|_{W^{m,r}(K)}, \quad \forall p, r \in [1, \infty].$$

Proof See the proof of Lemma 12.1 in [50]. □

Next, we will derive a series of essential inequalities used in the proof of the stability results for symmetric DDG and DDGIC methods.

Lemma 9 *Assume that $A(u) \in \mathbb{R}^{2\times 2}$ is positive definite with positive eigenvalues and there exist $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{R}$. If $\boldsymbol{\xi}(u)$ is given as in Eq. (7), then $\forall \mathbf{x} \in \mathbb{R}^2$ there holds*

$$|\boldsymbol{\xi}(u) \cdot \mathbf{x}| \leq \gamma^* \|\mathbf{x}\|,$$

where we denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^2$.

Proof Using the Scharwz inequality, we have

$$|\boldsymbol{\xi}(u) \cdot \mathbf{x}| \leq \|\boldsymbol{\xi}(u)\| \|\mathbf{x}\|.$$

Let $\mathbf{e}_1, \mathbf{e}_2$ be the orthonormal eigenvectors corresponding to the eigenvalues γ and γ^*, respectively. Since $A(u)$ is positive-definite, its eigenvectors form a basis of $\mathbb{R}^2$. Then, the unit normal vector can be written as $\mathbf{n} = n_1\mathbf{e}_1 + n_2\mathbf{e}_2$, and then, it follows that

$$\begin{aligned}\|\boldsymbol{\xi}(u)\|^2 &= \left\| A(\{\{u\}\})^T \mathbf{n} \right\|^2 = \|\gamma_1 n_1 \mathbf{e}_1 + \gamma_2 n_2 \mathbf{e}_2\|^2 = \gamma_1^2 n_1^2 + \gamma_2^2 n_2^2 \\ &\leq (\gamma^*)^2 \left(n_1^2 + n_2^2\right) = (\gamma^*)^2\end{aligned}$$

and the conclusion holds. □

Lemma 10 *Suppose that $A(u) \in \mathbb{R}^{2\times 2}$ is positive definite with positive eigenvalues and there exist $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{R}$. If $\beta_0 \geq 0$, then we have that*

$$\sum_{e\in\mathcal{E}_h} \int_e \frac{\beta_0}{h_e} [\![u]\!]^2 \mathbf{n} \cdot \boldsymbol{\xi}(\{\{u\}\})\, ds \geq \sum_{e\in\mathcal{E}_h} \int_e \frac{\gamma\beta_0}{h_e} [\![u]\!]^2\, ds.$$

Proof Recall the definition of the new direction vector $\boldsymbol{\xi}\{\{u\}\} = A(\{\{u\}\})^T \mathbf{n}$. Then,

$$\frac{\beta_0}{h_e} [\![u]\!]^2 \mathbf{n} \cdot \boldsymbol{\xi}(\{\{u\}\}) = \frac{\beta_0}{h_e} [\![u]\!]^2 \mathbf{n} \cdot A(\{\{u\}\})^T \mathbf{n}.$$

Since $\mathbf{n} \cdot A(\{\{u\}\})^T \mathbf{n} \geq \gamma_1 \geq \gamma$, the conclusion follows. □

Lemma 11 *Suppose that $A(u) \in \mathbb{R}^{2\times 2}$ is positive definite with positive eigenvalues and there exist $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{V}_h^k$. If $\beta_0 \geq 0$, then there exists a constant $C > 0$ such that*

$$\sum_{e\in\mathcal{E}_h} \int_e [\![u]\!] \{\{\nabla u\}\} \cdot \boldsymbol{\xi}(\{\{u\}\})\, ds \geq -\sum_{e\in\mathcal{E}_h} \frac{\gamma\beta_0}{2h} \left\| [\![u]\!] \right\|_{L^2(e)}^2 - \sum_{K\in\mathcal{T}_h} C \frac{(\gamma^* k)^2}{4\gamma\beta_0} \|\nabla u\|_{L^2(K)}^2.$$

Furthermore, under the same assumptions, there also holds

$$\sum_{e\in\mathcal{E}_h} \int_e [\![u]\!] \{\{\nabla u\}\} \cdot \boldsymbol{\xi}(\{\{u\}\})\, ds \geq -\sum_{e\in\mathcal{E}_h} \frac{\gamma\beta_0}{4h} \left\| [\![u]\!] \right\|_{L^2(e)}^2 - \sum_{K\in\mathcal{T}_h} C \frac{(\gamma^* k)^2}{2\gamma\beta_0} \|\nabla u\|_{L^2(K)}^2.$$

Proof Note that $\{\{\nabla u\}\} = \frac{1}{2}(\nabla u)^+ + \frac{1}{2}(\nabla u)^-$. Then, we have

$$\begin{aligned}
&\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\{\{\nabla u\}\}\cdot\boldsymbol{\xi}(\{\{u\}\})\,ds \\
&= \frac{1}{2}\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\left((\nabla u)^+ + (\nabla u)^-\right)\cdot\boldsymbol{\xi}(\{\{u\}\})\,ds \\
&\geq -\frac{1}{2}\sum_{e\in\mathcal{E}_h}\int_e \left|[\![u]\!]\right|\left(\left|(\nabla u)^+\cdot\boldsymbol{\xi}(\{\{u\}\})\right| + \left|(\nabla u)^-\cdot\boldsymbol{\xi}(\{\{u\}\})\right|\right)\,ds \\
&\geq -\frac{\gamma^*}{2}\sum_{e\in\mathcal{E}_h}\int_e \left|[\![u]\!]\right|\left\|(\nabla u)^+\right\|\,ds - \frac{\gamma^*}{2}\sum_{e\in\mathcal{E}_h}\int_e \left|[\![u]\!]\right|\left\|(\nabla u)^-\right\|\,ds,
\end{aligned}$$

where we have invoked Lemma 9 in the last step. Moreover, by Corollary 7 with $\epsilon = \frac{\gamma\beta_0}{\gamma^* h}$, we obtain

$$\gamma^*\int_e \left|[\![u]\!]\right|\left\|(\nabla u)^\pm\right\|\,ds \leq \frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)} + \frac{(\gamma^*)^2 h}{2\gamma\beta_0}\left\|(\nabla u)^\pm\right\|^2_{L^2(e)}. \tag{29}$$

So that we have

$$\begin{aligned}
&\sum_{e\in\mathcal{E}_h}\int_e [\![u]\!]\{\{\nabla u\}\}\cdot\boldsymbol{\xi}(\{\{u\}\})\,ds \\
&\geq -\sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)} \\
&-\sum_{e\in\mathcal{E}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|(\nabla u)^+\right\|^2_{L^2(e)} - \sum_{e\in\mathcal{E}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|(\nabla u)^-\right\|^2_{L^2(e)}.
\end{aligned} \tag{30}$$

Note that the last two terms above are simply summations over individual edges in the triangulation $\mathcal{T}_h$, and each summation is responsible for accumulating $\left\|(\nabla u)^\pm\right\|^2_{L^2(e)}$ only from one side of the edge. In the global sense, these summations accumulate $\left\|(\nabla u)^{interior}\right\|^2_{L^2(\partial K)}$ for each cell K in the domain. Thus, they can be converted into a single summation over cells. That is,

$$\sum_{e\in\mathcal{E}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|(\nabla u)^+\right\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|(\nabla u)^-\right\|^2_{L^2(e)} = \sum_{K\in\mathcal{T}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|\nabla u\right\|^2_{L^2(\partial K)}. \tag{31}$$

This expression is useful since we can invoke the trace inequality

$$\sum_{K\in\mathcal{T}_h}\frac{(\gamma^*)^2 h}{4\gamma\beta_0}\left\|\nabla u\right\|^2_{L^2(\partial K)} \leq \sum_{K\in\mathcal{T}_h} C\frac{(\gamma^* k)^2}{4\gamma\beta_0}\left\|\nabla u\right\|^2_{L^2(K)}, \tag{32}$$

for some constant $C > 0$. Thus, substituting Eqs. (31) and (32) into Eq. (30) completes the first part of the proof. The second part follows after following the same steps but using $\epsilon = \frac{\gamma\beta_0}{2\gamma^* h}$ in Eq. (29) instead. □

Lemma 12 *Suppose that $A(u) \in \mathbb{R}^{2\times 2}$ is positive definite and there exist $\gamma, \gamma^* \in \mathbb{R}^+$ such that the eigenvalues $(\gamma_1(u), \gamma_2(u))$ of $A(u)$ lie between $[\gamma, \gamma^*]$ for $\forall u \in \mathbb{V}_h^k$. If $\beta_0 \geq 0$, then*

there exists a constant $C > 0$ such that

$$\sum_{e\in\mathcal{E}_h}\int_e \beta_1 h[\![u]\!][\![\nabla(\nabla u\cdot\mathbf{n})]\!]\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \geq -\sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)} - \sum_{K\in\mathcal{T}_h} C\frac{(\gamma^*\beta_1 k)^2}{\gamma\beta_0}\left\|\nabla u\right\|^2_{L^2(K)}.$$

Proof By convention, the outward unit normal vector $\mathbf{n}$ is understood as $\mathbf{n} = \mathbf{n}^+$. Also, it can be understood in terms of the inward unit normal vector as $\mathbf{n} = -\mathbf{n}^-$. Therefore, the jump term for the second derivatives can be rewritten as

$$[\![\nabla(\nabla u\cdot\mathbf{n})]\!] = (\nabla(\nabla u\cdot\mathbf{n}))^+ + (\nabla(\nabla u\cdot\mathbf{n}))^-.$$

With this understanding, we have that

$$\begin{aligned}
&\int_e \beta_1 h[\![u]\!][\![\nabla(\nabla u\cdot\mathbf{n})]\!]\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds\\
&= \int_e \beta_1 h[\![u]\!]\left((\nabla(\nabla u\cdot\mathbf{n}))^+ + (\nabla(\nabla u\cdot\mathbf{n}))^-\right)\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds\\
&\geq -\int_e \beta_1 h\left|[\![u]\!]\right|\left|\left((\nabla(\nabla u\cdot\mathbf{n}))^+ + (\nabla(\nabla u\cdot\mathbf{n}))^-\right)\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\right|\,ds\\
&\geq -\gamma^*\int_e \beta_1 h\left|[\![u]\!]\right|\left(\left\|(\nabla(\nabla u\cdot\mathbf{n}))^+\right\| + \left\|(\nabla(\nabla u\cdot\mathbf{n}))^-\right\|\right)\,ds.
\end{aligned}$$

Note that we have invoked Lemma 9 and triangle inequality in the last step. Furthermore, by Corollary 7 with $\epsilon = \frac{\gamma\beta_0}{2\gamma^*\beta_1 h^2}$, we obtain

$$\gamma^*\int_e \beta_1 h\left|[\![u]\!]\right|\left\|(\nabla(\nabla u\cdot\mathbf{n}))^\pm\right\|\,ds \leq \frac{\gamma\beta_0}{4h}\left\|[\![u]\!]\right\|^2_{L^2(e)} + \frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|(\nabla(\nabla u\cdot\mathbf{n}))^\pm\right\|^2_{L^2(e)}.$$

Thus,

$$\begin{aligned}
\sum_{e\in\mathcal{E}_h}\int_e \beta_1 h[\![u]\!][\![\nabla(\nabla u\cdot\mathbf{n})]\!]\cdot\boldsymbol{\xi}(\{\!\{u\}\!\})\,ds \geq &-\sum_{e\in\mathcal{E}_h}\frac{\gamma\beta_0}{2h}\left\|[\![u]\!]\right\|^2_{L^2(e)}\\
&-\sum_{e\in\mathcal{E}_h}\frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|(\nabla(\nabla u\cdot\mathbf{n}))^+\right\|^2_{L^2(e)}\\
&-\sum_{e\in\mathcal{E}_h}\frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|(\nabla(\nabla u\cdot\mathbf{n}))^-\right\|^2_{L^2(e)}.
\end{aligned} \tag{33}$$

As in the proof of Lemma 11, we convert the summations over edges to a summation over cells:

$$\begin{aligned}
&\sum_{e\in\mathcal{E}_h}\frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|(\nabla(\nabla u\cdot\mathbf{n}))^+\right\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}_h}\frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|(\nabla(\nabla u\cdot\mathbf{n}))^-\right\|^2_{L^2(e)}\\
&= \sum_{K\in\mathcal{T}_h}\frac{(\gamma^*\beta_1)^2h^3}{\gamma\beta_0}\left\|\nabla(\nabla u\cdot\mathbf{n})\right\|^2_{L^2(\partial K)}.
\end{aligned} \tag{34}$$

At this point, it might be tempting to invoke the trace theorem for the norm on the right-hand side of the above equation. However, a more useful inequality can be obtained by considering

the Euclidean norm $\|\nabla(\nabla u \cdot \mathbf{n})\|^2$. We first note that

$$\begin{aligned}\|\nabla(\nabla u \cdot \mathbf{n})\|^2 &= (u_{xx}n_1 + u_{yx}n_2)^2 + (u_{xy}n_1 + u_{yy}n_2)^2 \\ &= (u_{xx}^2 + u_{xy})^2 n_1^2 + (u_{yx}^2 + u_{yy})^2 n_2^2 + 2n_1n_2(u_{xx}u_{yx} + u_{xy}u_{yy}).\end{aligned}$$

For the cross-product term above, we invoke Lemma 6 with $p = q = 2$

$$\begin{aligned}2n_1n_2(u_{xx}u_{yx} + u_{xy}u_{yy}) &= 2\left((n_2u_{xx})(n_1u_{yx}) + (n_2u_{xy})(n_1u_{yy})\right) \\ &\le (u_{xx}^2 + u_{xy})^2 n_2^2 + (u_{yx}^2 + u_{yy})^2 n_1^2.\end{aligned}$$

Since $n_1^2 + n_2^2 = 1$, we obtain

$$\|\nabla(\nabla u \cdot \mathbf{n})\|^2 \le u_{xx}^2 + u_{xy}^2 + u_{yx}^2 + u_{yy}^2.$$

Using this and the trace inequality gives

$$\begin{aligned}\|\nabla(\nabla u \cdot \mathbf{n})\|_{L^2(\partial K)}^2 &= \int_{\partial K} \|\nabla(\nabla u \cdot \mathbf{n})\|^2 \, ds \\ &\le \int_{\partial K} \left(u_{xx}^2 + u_{xy}^2 + u_{yx}^2 + u_{yy}^2\right) ds \\ &\le C\frac{k^2}{h}\int_K \left(u_{xx}^2 + u_{xy}^2 + u_{yx}^2 + u_{yy}^2\right) dxdy = C\frac{k^2}{h}|u|_{H^2(K)}^2.\end{aligned}$$

By Lemma 8 with $d = l = p = r = 2$ and $m = 1$, we have

$$|u|_{H^2(K)} \le \frac{C}{h}|u|_{H^1(K)} = \frac{C}{h}\|\nabla u\|_{L^2(K)},$$

which leads to

$$\|\nabla(\nabla u \cdot \mathbf{n})\|_{L^2(\partial K)}^2 \le C\frac{k^2}{h^3}\|\nabla u\|_{L^2(K)}^2. \tag{35}$$

Finally, substituting Eqs. (34) and (35) in Eq. (33) leads to the desired result. □

References

1. Liu, Hailiang, Yan, Jue: The direct discontinuous galerkin (DDG) methods for diffusion problems. SIAM J. Numer. Anal. **47**(1), 675–698 (2008)
2. Liu, Hailiang, Yan, Jue: The direct discontinuous galerkin (DDG) method for diffusion with interface corrections. Commun. Comput. Phys. **8**(3), 541 (2010)
3. Vidden, Chad, Yan, Jue: A new direct discontinuous galerkin method with symmetric structure for nonlinear diffusion equations. J. Comput. Math. **31**(6), 638–662 (2013)
4. Yan, Jue: A new nonsymmetric discontinuous galerkin method for time dependent convection diffusion equations. J. Sci. Comput. **54**(2), 663–683 (2013)
5. Reed, William H., Hill, T.R.: Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, (1973)
6. Cockburn, Bernardo, Shu, Chi-Wang.: Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. General framework. Math. Comput. **52**(186), 411–435 (1989)
7. Cockburn, Bernardo, Lin, San-Yih., Shu, Chi-Wang.: 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)
8. Cockburn, Bernardo, Hou, Suchung, Shu, Chi-Wang.: The runge-kutta local projection discontinuous galerkin finite element method for conservation laws. iv. The multidimensional case. Math. Comput. **54**(190), 545–581 (1990)

9. Cockburn, Bernardo, Shu, Chi-Wang.: The runge-kutta discontinuous galerkin method for conservation laws v: multidimensional systems. J. Comput. Phys. **141**(2), 199–224 (1998)
10. Cockburn, B., Johnson, C., Shu, C.-W., Tadmor, E.: Advanced Numerical Approximation of Nonlinear Hyperbolic Equations. Lecture Notes in Mathematics, vol. 1697. Springer-Verlag, Berlin (1998)
11. Shu, C.-W.: Discontinuous Galerkin method for time-dependent problems: survey and recent developments. In *Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations*, vol. 157 of *IMA Vol. Math. Appl.*, pp 25–62. Springer, (2014)
12. Zhang, Xiangxiong, Shu, Chi-Wang.: Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. Proc. R. Soc. A Math. Phys. Eng. Sci. **467**(2134), 2752–2776 (2011)
13. Arnold, Douglas N.: An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Analy. **19**(4), 742–760 (1982)
14. Wheeler, Mary Fanett: An elliptic collocation-finite element method with interior penalties. SIAM J. Numer. Anal. **15**(1), 152–161 (1978)
15. Baker, Garth A.: Finite element methods for elliptic equations using nonconforming elements. Math. Comput. **31**(137), 45–59 (1977)
16. Rivière, Béatrice., Wheeler, Mary F., Girault, Vivette: Improved energy estimates for interior penalty, constrained and discontinuous galerkin methods for elliptic problems part. i. Comput. Geosci. **3**(3), 337–360 (1999)
17. Rivière, Béatrice., Wheeler, Mary F., Girault, Vivette: 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)
18. Hartmann, Ralf, Houston, Paul: Symmetric interior penalty dg methods for the compressible navier-stokes equations i: method formulation. Int. J. Numer. Anal. Model. **3**(1), 1–20 (2006)
19. Hartmann, Ralf, Houston, Paul: Symmetric interior penalty dg methods for the compressible navier-stokes equations ii: goal-oriented a posteriori error estimation. Int. J. Numer. Anal. Model. **3**(2), 141–162 (2006)
20. Hartmann, Ralf, Houston, Paul: An optimal order interior penalty discontinuous galerkin discretization of the compressible Navier–stokes equations. J. Comput. Phys. **227**(22), 9670–9685 (2008)
21. Bassi, F., Rebay, S.: A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier-stokes equations. J. Comput. Phys. **131**(2), 267–279 (1997)
22. Bassi, F., Rebay, S.: Gmres discontinuous galerkin solution of the compressible Navier–Stokes equations. In *Discontinuous Galerkin Methods*, pp 197–208. Springer, (2000)
23. Bassi, F., Rebay, S.: A high order discontinuous galerkin method for compressible turbulent flows. In: *Discontinuous Galerkin Methods*, pp 77–88. Springer, (2000)
24. Bassi, Francesco, Crivellini, Andrea, Rebay, Stefano, Savini, Marco: Discontinuous galerkin solution of the reynolds-averaged Navier–Stokes and k-ω turbulence model equations. Comput. Fluids **34**(4–5), 507–540 (2005)
25. Cockburn, Bernardo, Shu, Chi-Wang.: The local discontinuous galerkin method for time-dependent convection-diffusion systems. SIAM J. Numer. Anal. **35**(6), 2440–2463 (1998)
26. Cockburn, Bernardo, Dawson, Clint: Approximation of the velocity by coupling discontinuous galerkin and mixed finite element methods for flow problems. Comput. Geosci. **6**(3), 505–522 (2002)
27. Yan, Jue, Shu, Chi-Wang.: A local discontinuous galerkin method for kdv type equations. SIAM J. Numer. Anal. **40**(2), 769–791 (2002)
28. Baumann, Carlos Erik, Oden, J Tinsley: A discontinuous hp finite element method for convection-diffusion problems. Comput. Methods Appl. Mech. Eng. **175**(3–4), 311–341 (1999)
29. Baumann, Carlos Erik, Oden, J Tinsley: A discontinuous hp finite element method for the euler and Navier–Stokes equations. Int. J. Numer. Methods Fluids **31**(1), 79–95 (1999)
30. Cockburn, Bernardo, Gopalakrishnan, Jayadeep, Lazarov, Raytcho: Unified hybridization of discontinuous galerkin, mixed, and continuous galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**(2), 1319–1365 (2009)
31. Brenner, Susanne C., Owens, Luke, Sung, Li-Yeng.: A weakly over-penalized symmetric interior penalty method. Electron. Trans. Numer. Anal. **30**, 107 (2008)
32. Lin, Guang, Liu, Jiangguo, Sadre-Marandi, Farrah: A comparative study on the weak galerkin, discontinuous galerkin, and mixed finite element methods. J. Comput. Appl. Math. **273**, 346–362 (2015)
33. Cheng, Yingda, Shu, Chi-Wang.: A discontinuous galerkin finite element method for time dependent partial differential equations with higher order derivatives. Math. Comput. **77**(262), 699–730 (2008)
34. Chen, Anqi, Li, Fengyan, Cheng, Yingda: An ultra-weak discontinuous galerkin method for schrödinger equation in one dimension. J. Sci. Comput. **78**(2), 772–815 (2018)
35. Arnold, Douglas N., Brezzi, Franco, Cockburn, Bernardo, Marini, L Donatella: Unified analysis of discontinuous galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**(5), 1749–1779 (2002)

36. Shu, Chi-Wang.: Discontinuous galerkin methods for time-dependent convection dominated problems: Basics, recent developments and comparison with other methods. In: Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Lecture Notes in Computational Science and Engineering, pp. 371–399. Springer International Publishing, Cham (2016)
37. Chen, Zheng, Huang, Hongying, Yan, Jue: Third order maximum-principle-satisfying direct discontinuous galerkin methods for time dependent convection diffusion equations on unstructured triangular meshes. J. Comput. Phys. **308**, 198–217 (2016)
38. Zhang, M., Yan, J.: Fourier type super convergence study on DDGIC and symmetric DDG methods. J. Sci. Comput. **73**(2–3), 1276–1289 (2017)
39. Qiu, C., Liu, Q., Yan, J.: Third order positivity-preserving direct discontinuous Galerkin method for chemotaxis keller-segel equation. J. Comput. Phys. **433**, 110191 (2020)
40. Huang, H., Li, J., Yan, J.: High order symmetric direct discontinuous Galerkin method for elliptic interface problems with fitted mesh. J. Comput. Phys. **409**, 109301 (2020)
41. Danis, Mustafa E.: Yan, Jue: a new direct discontinuous Galerkin method with interface correction for two-dimensional compressible Navier–Stokes equations. J. Comput. Phys. **452**, 110904 (2022)
42. Shu, Chi-Wang., Osher, Stanley: Efficient implementation of essentially non-oscillatory shock-capturing schemes. J. Comput. Phys. **77**(2), 439–471 (1988)
43. Cockburn, Bernardo, Shu, Chi-Wang.: Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. J. Sci. Comput. **16**(3), 173–261 (2001)
44. Zhang, Xiangxiong, Xia, Yinhua, Shu, Chi-Wang.: Maximum-principle-satisfying and positivity-preserving high order discontinuous galerkin schemes for conservation laws on triangular meshes. J. Sci. Comput. **50**(1), 29–62 (2012)
45. Liu, Yuanyuan, Shu, Chi-Wang., Zhang, Mengping: High order finite difference weno schemes for nonlinear degenerate parabolic equations. SIAM J. Sci. Comput. **33**(2), 939–965 (2011)
46. Zhang, Xiangxiong, Shu, Chi-Wang.: On maximum-principle-satisfying high order schemes for scalar conservation laws. J. Comput. Phys. **229**(9), 3091–3120 (2010)
47. Jie, Du., Yang, Yang: Maximum-principle-preserving third-order local discontinuous galerkin method for convection-diffusion equations on overlapping meshes. J. Comput. Phys. **377**, 117–141 (2019)
48. Guo, Li., Yang, Yang: Positivity preserving high-order local discontinuous galerkin method for parabolic equations with blow-up solutions. J. Comput. Phys. **289**, 181–195 (2015)
49. Zhang, X.: On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations. J. Comput. Phys. **328**, 301–343 (2017)
50. Ern, Alexandre, Guermond, Jean-Luc.: Finite Elements I: Approximation and Interpolation, vol. 72. Springer Nature (2021)