The direct discontinuous Galerkin method with symmetric structure for diffusion problems

by

Chad Nathaniel Vidden

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Applied Mathematics (Numerical Analysis)

Program of Study Committee:

Jue Yan, Major Professor

L. Steven Hou

Hailiang Liu

Paul Sacks

Zhi Jian Wang

Iowa State University

Ames, Iowa

2012

DEDICATION

To my loving wife and daughter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank all who helped me with various aspects of this research and dissertation writing. First, I thank my committee members Drs. L. Steven Hou, Hailiang Liu, Paul Sacks, and Zhi Jian Wang for useful comments and suggestions. I thank Dr. Liu for many interesting conversations and valuable advice. Finally, I thank my advisor, Dr. Jue Yan, for her expert advice, patience, and support in achieving my goals.

ABSTRACT

In this thesis, a discontinuous Galerkin (DG) finite element method for nonlinear diffusion equations named the symmetric direct discontinuous Galerkin (DDG) method is studied. The scheme is first developed for the one dimensional heat equation using the DG approach. To define a numerical flux for the numerical solution derivative, the solution derivative trace formula of the heat equation with discontinuous initial data is used. A numerical flux for the test function is introduced in order to arrive at a symmetric scheme. Having a symmetric scheme is the key to proving an optimal $L^2(L^2)$ error estimate. In addition, stability results and an optimal energy error estimate are proven. In order to ensure stability of the scheme, a notion of flux admissibility is defined. Flux admissibility is analyzed resulting in explicit guidelines for choosing free coefficients in the numerical flux formula. The scheme is extended to one dimensional nonlinear diffusion, nonlinear convection diffusion, as well as two dimensional linear and nonlinear diffusion problems. Numerical examples are carried out to demonstrate the optimal $(k+1)$th order of accuracy for the method with degree k polynomial approximations for both linear and nonlinear problems, under one-dimensional and two-dimensional settings. In addition, admissibility analysis results are explored numerically.

CHAPTER 1. INTRODUCTION

1.1 General Introduction to Discontinuous Galerkin Methods

The discontinuous Galerkin (DG) method is a class of finite element methods first introduced by Reed and Hill [31] in 1973. A Galerkin finite element method has the characteristic of having the same function space for both the numerical solution and test functions. DG methods are named for their piecewise discontinuous function space, usually chosen to be polynomials, for both the numerical solution and test functions. There are many nice properties which have increased the popularity of these methods in the past years. Some features include allowance for arbitrary domain decompositions, high order accuracy of numerical solution, complete freedom in changing the polynomial degrees in each element independent of that in the neighbors, and extremely local data structure resulting in high parallel efficiency. These robust and accurate methods have quickly attracted the interest of the scientific community and are currently finding use in very diverse applications such as electromagnetism, magneto-hydrodynamics, meteorology, oceanography, semiconductor device simulation, weather forecasting, amd turbulent flows, among many others.

The original DG method was introduced by Reed and Hill [31] in 1973 for neutron transport equations, which are linear hyperbolic equations. The first analysis of this method was completed by LeSaint and Raviart [24] in 1974 for general mesh triangulations and rectangular grids, and this analysis was improved by Johnson and Pitkäranta [23] and Peterson [30]. It wasn't until the 1990s when DG methods became a very active area of research. A major development of the DG method known as the Runge-Kutta DG (RKDG) methods for nonlinear hyperbolic conservation laws was carried out by Cockburn, Shu, and collaborators ([14] [13] [12] [11] [15]) in the late 80s and early 90s. In 1992 Richter [32] extended the original DG

method to linear convection-diffusion equations. In 1997, Bassi and Rebay [4] extended DG methods further to compressible Navier-Stokes equations. In 1998, Bassi and Rebay's work was generalized to become the well known local DG (LDG) methods through work by Cockburn and Shu [9]. We refer to [7, 8, 10] for reviews and further references.

Working independently of the above, a class of penalty methods for elliptic problems was developed. This began in 1971 with Nitsche [29] where the first discontinuous basis function was used in approximations of solutions to elliptic equations involving a penalty term. This was applied to finite element methods by Babuska in 1973 [2]. This evolved to become the interior penalty (IP) methods which are still widely used today. The foundation of symmetric IP Galerkin (SIPG) methods was established in the late 70s and early 80s primarily by Baker [3], Wheeler [46], and Arnold [1].

1.2 DG Methods for Diffusion

Using DG methods for diffusion problems has been considered since the late 90s. This has been a challenging task because of the difficulty in properly defining the numerical solution derivative at cell interfaces. Because the numerical solution is allowed to be discontinuous across cell interfaces, appropriate numerical fluxes for diffusion terms need to be defined. In 2001 Shu [35] illustrated some problems resulting from using DG methods for diffusion problems. To illustrate this, he considered the simplest case, the regular heat equation

$$U_t - U_{xx} = 0, \tag{1.1}$$

in one dimension. A standard DG formulation formally leads to

$$\int_{I_j} u_t v \, dx + \int_{I_j} u_x v_x \, dx - (\widehat{u_x})_{j+1/2} \, v^-_{j+1/2} + (\widehat{u_x})_{j-1/2} \, v^+_{j-1/2} = 0, \tag{1.2}$$

for u, v piecewise polynomials on computational cell $I_j = [x_{j-1/2}, x_{j+1/2}]$. The main difference here when compared to the hyperbolic case is that the numerical flux $\widehat{u_x}$ for diffusion is for the solution derivative at the cell interface, not simply the function value as in the hyperbolic case. A naive choice of this numerical flux is a central flux $\widehat{u_x} = \frac{1}{2}(u_x^- + u_x^+)$. Here, because u is multivalued at cell interfaces, u^+ denotes the function value from the right and u^- denotes

the function value from the left. This seems like a natural choice, but one can see for the piecewise constant case the scheme doesn't make sense at all. One can see that both the numerical flux and second integration term are zero, implying that the time derivative of the numerical solution is zero. For linear and better cases, the DG solution proves to be stable but is completely inconsistent with the exact solution. That is, the method converges to a solution different from the exact solution. This illustrates some of the challenges when considering DG methods for diffusion.

There have been several DG methods suggested in literature for solving diffusion problems. One is the local DG (LDG) method of Cockburn and Shu [9]. The basic idea of LDG is to rewrite the heat equation (1.1) as the first order system

$$u_t - q_x = 0, \qquad q - u_x = 0,$$

then formally apply the DG technique to arrive at

$$\int_{I_j} u_t v \, dx + \int_{I_j} q v_x \, dx - \widehat{q}_{j+1/2} v^-_{j+1/2} + \widehat{q}_{j-1/2} v^+_{j-1/2} = 0,$$
$$\int_{I_j} q w \, dx + \int_{I_j} u w_x \, dx - \widehat{u}_{j+1/2} w^-_{j+1/2} + \widehat{u}_{j-1/2} w^+_{j-1/2} = 0.$$

This method was first considered by Bassi and Rebay [4] with central fluxes $\widehat{u} = \frac{1}{2}\left(u^+ + u^-\right)$ and $\widehat{q} = \frac{1}{2}\left(q^+ + q^-\right)$. This flux choice converges, though it proves to be suboptimal for odd order polynomial approximation. Cockburn and Shu instead propose the alternating flux $\widehat{u} = u^-$ and $\widehat{q} = q^+$, and they showed it converges at optimal rates. In some situations, especially diffusion dominated problems, computational cost of LDG methods are higher than others.

Another approach was given by Baumann and Oden in 1998. Their idea was to modify the inconsistent scheme (1.2) to arrive at the following.

$$\int_{I_j} u_t v \, dx + \int_{I_j} u_x v_x \, dx - \left(\widehat{u_x}\right)_{j+1/2} v^-_{j+1/2} + \left(\widehat{u_x}\right)_{j-1/2} v^+_{j-1/2}$$
$$-\frac{1}{2}\left(v_x\right)^-_{j+1/2}\left(u^+_{j+1/2} - u^-_{j+1/2}\right) - \frac{1}{2}\left(v_x\right)^-_{j-1/2}\left(u^+_{j-1/2} - u^-_{j-1/2}\right) = 0$$

Again, the central flux is used. The addition of these interface terms imply stability of this scheme. Just as above, this scheme only makes sense for linear or better polynomial approximations because otherwise numerical flux terms, interface terms, and the second space integration

all are zero. In addition, suboptimal rates of convergence are seen for even polynomial approximations.

The method of Baumann and Oden is closely related to a class of DG methods known as interior penalty (IP) methods, originally dating back to the 1970s. This class includes the original symmetric IP Galerkin method from the late 70s, the nonsymmetric IP Galerkin (NIPG) method [33] which came soon after the Baumann-Oden method, and the incomplete IP Galerkin (IIPG) method [16] in 2004. IP methods rely on penalty terms, typically proportional to the jump of the numerical solution across cell interfaces, to give stability to the scheme. For the heat equation (1.1), these methods take the form

$$\int_{I_j} u_t v \, dx \int_{I_j} u_x v_x \, dx - (\widehat{u_x})_{j+1/2} \, v^-_{j+1/2} + (\widehat{u_x})_{j-1/2} \, v^+_{j-1/2}$$
$$- \epsilon \frac{1}{2} (v_x)^-_{j+1/2} \left(u^+_{j+1/2} - u^-_{j+1/2} \right) - \epsilon \frac{1}{2} (v_x)^-_{j-1/2} \left(u^+_{j-1/2} - u^-_{j-1/2} \right) - \frac{\sigma}{\Delta x} [u][v] \Big|^{j+1/2}_{j-1/2} = 0,$$

where $\widehat{u_x}$ denotes the above central flux, $[u]$ is the jump of the numerical solution across an interface $[u] = u^+ - u^-$, and σ is a penalty parameter chosen large enough to ensure stability and convergence. Choosing $\epsilon = -1$ gives SIPG, $\epsilon = 1$ for NIPG, and $\epsilon = 0$ for IIPG.

There are more recent works including those by Van Leer and Nomura in 2005 in [41] and later in [43, 42] where a new recovery principle is used to create new DG schemes for diffusion. The basic idea here is to recover an underlying smooth solution from the discontinuous DG solution. Specifically, for each piecewise continuous polynomial of degree k defined on two adjacent cells, they construct a polynomial of degree $2k+1$ and then compute the polynomial derivative at the interface, thus giving a precise value for the numerical flux at cell interfaces. The main issue with this approach is the high complexity for higher order polynomial approximations. In 2006, Gassner et al. in [20] and later in [28, 21] used repeated integration by parts on the diffusion term in combination with a Riemann solver. That is, exact solutions for the diffusion equation with a step function Riemann initial data are used to provide a flux definition for the solution derivative resulting from the integration by parts. This idea is related to direct DG ideas as discussed below. In 2007, Cheng and Shu in [5] also used repeated integration by parts to design a suitable numerical flux for time dependent problems with high order spacial derivatives. This relies on suitably chosen numerical fluxes for all derivatives of the numerical

solution up to one lower than the order of the PDE. For nonlinear problems, choosing numerical fluxes becomes more difficult in order to ensure scheme stability.

1.3 Review of the Direct Discontinuous Galerkin Methods

Recently in [26], a direct discontinuous Galerkin (DDG) method was developed for solving diffusion equations. The scheme is based on the direct weak formulation of the heat equation, and a general numerical flux formula for the solution derivative was proposed. An optimal kth order error estimate in an energy norm was obtained for P^k polynomial approximations of linear diffusion equations. However, numerical experiments in [26] show that when measured under L^2 and L^∞ norms, the scheme accuracy is sensitive to the coefficient choices in the numerical flux formula. That is, for higher order P^k ($k \geq 4$) polynomial approximations it is difficult to identify suitable coefficients in the numerical flux formula to obtain optimal $(k+1)$th order of accuracy.

In [27], extra interface correction terms were introduced into the scheme formulation, and a refined version of the DDG method is obtained. A simpler numerical flux formula is used in [27] and numerically optimal $(k+1)th$ order of accuracy under L^2 and L^∞ norms are achieved in numerical experiments for any P^k polynomial approximations. The refined DDG method is not sensitive to the coefficients in the numerical flux formula. That is, numerical tests show a large class of admissible numerical fluxes can lead to the optimal convergence rates.

1.4 Summary of Major Work

There are two broad contributions of this thesis. The first is the development and analysis of a discontinuous Galerkin method for diffusion problems known as the symmetric direct discontinuous Galerkin method. This DG method is shown to have a number of desirable qualities including stability and optimal rates of convergence in energy and L^2 norms. The second contribution is the admissibility analysis which provides guidelines for choosing free coefficients within the numerical flux formula. This technique can be applied to a whole class of DG methods.

The symmetric DDG method was inspired by recent works known as the direct DG methods as described above. The DDG method [26] and the DDG method with interface corrections [27] are schemes which both lack symmetric properties. Without this symmetry, there is no known way to obtain L^2 error analysis and guarantee optimal rates of convergence. For the symmetric DDG scheme, a numerical flux for the test function derivative was introduced. With the same numerical flux formula for the solution derivative and the test function derivative, the bilinear form for the diffusion term thus obtained has a symmetry property. This symmetric structure is the key feature required to prove an optimal $L^2(L^2)$ error estimate for the numerical solution. The $L^2(L^2)$ error estimate was established using a parabolic lifting argument. This $L^2(L^2)$ error estimate is the primary motivation for the work in this thesis, and currently the only other method which can be shown to have this result for diffusion is the SIPG method. In addition, the scheme is shown to be consistent, stable, and have optimal convergence rates in the energy norm. One-dimensional and two-dimensional numerical examples are carried out to illustrate the capacity of this method. $(k+1)th$ optimal order of accuracy with piecewise P^k polynomial approximations are obtained for both linear and nonlinear diffusion problems in a number of settings. This part of the thesis has recently been submitted for publication to the SIAM Journal on Numerical Analysis [44].

Also, explicit guidelines for choosing admissible numerical fluxes for the one dimensional heat equation on a uniform mesh are given. The technique used to obtain this result is different when compared to related work. What makes it different is that the problem is solved using a combinatorial argument rather than a PDE argument. The idea is to begin with a notion of flux admissibility which is exactly the needed requirement for stability of the scheme. Then, the admissibility condition is analyzed directly. The resulting computation involves Hilbert matrices, and in the end, some combinatorics techniques. These guidelines are developed for the SDDG scheme, but results immediately follow for the SIPG method. Compared to the SIPG method [1], the penalty coefficient of the symmetric DDG method can be decreased from k^2 to $k^2/4$. In addition, the same analysis can be easily adapted to be applied to existing methods such as the original DDG method, the DDG method with interface corrections, as well as any interior penalty method. A continued study of numerical flux admissibility for

nonuniform mesh, nonlinear diffusion, and higher spacial dimensions is planned.

1.5 Outline of Thesis

In Chapter 2, the symmetric direct discontinuous Galerkin method is formulated and analyzed for the one dimensional heat equation. Formulation is done in the typical DG fashion. To define the numerical flux, the exact solution for the heat equation in an analogous setting is used as inspiration. In order to make the scheme symmetric, a test function numerical flux is defined which is of the same form as the numerical solution flux. Next, in order to choose suitable free coefficients in the numerical flux formula, a notion of flux admissibility is defined. This admissibility condition is the exact requirement needed to guarantee stability of the scheme. Then, this admissibility condition is analyzed completely resulting in an explicit formula used for choosing flux coefficients. Optimal error estimates are then shown for both the energy norm as well as the L^2 norm. Last, numerical results are given illustrating performance on uniform and nonuniform meshes as well as exploration of admissibility analysis results. In all cases, optimal rates of convergence are observed in L^∞ and L^2 norms.

For Chapter 3, the scheme is extended to one dimensional nonlinear diffusion equations and nonlinear convection diffusion equations. The resulting schemes are a direct extension of the formulation given in Chapter 2. A notion of admissibility for these cases is defined, again motivated by stability requirements. Last, numerical tests in a number of settings are given and show optimal rates of convergence in L^∞ and L^2 norms.

In Chapter 4, the symmetric DDG scheme is extended even further to handle two dimensional linear and nonlinear cases. Again, flux admissibility is defined and resulting stability is shown. For the linear case, L^2 error analysis can be shown using the exact strategy as in the one dimensional case with minor differences related to the increase in dimensionality. Last, numerical examples are given and discussed.

Chapter 5 gives concluding remarks as well as future directions of research. Last, appendices A, B, and C provided coding notes, Hilbert matrix results, as well as a derivation of the formula used as inspiration for the numerical flux definition.

CHAPTER 2. ONE DIMENSIONAL LINEAR DIFFUSION PROBLEMS

In this chapter, the symmetric direct discontinuous Galerkin method is developed and analyzed for the one dimensional heat equation

$$U_t - U_{xx} = 0, \quad \text{for } (x,t) \in \Omega \times (0,T) \tag{2.1}$$

with initial data $U(x,0) = U_0(x)$ for $x \in \Omega \subset \mathbb{R}$ and periodic boundary conditions. We begin with this one dimensional linear case for simplicity of presentation, and much of the same analysis is carried out for nonlinear and higher dimensional cases in later chapters. Also, it is for simplicity of presentation to consider periodic boundary conditions. The scheme can be easily applied to any well-posed boundary conditions.

The scheme formulation is provided first using a standard discontinuous Galerkin approach. The key point of interest in the scheme formulation is the definition of the numerical flux. To define the numerical flux, the exact solution for the heat equation in an analogous setting is used as inspiration. In order to make the scheme symmetric, a test function numerical flux is defined which is of the same form as the numerical solution flux. Next, in order to choose suitable free coefficients in the numerical flux formula, a notion of flux admissibility is defined. This admissibility condition is the exact requirement needed to guarantee stability of the scheme. Then, this admissibility condition is analyzed completely resulting in an explicit formula used for choosing flux coefficients. Optimal error estimates are then shown for both the energy norm as well as the L^2 norm. Last, numerical results are given illustrating performance on uniform and nonuniform meshes as well as exploration of admissibility analysis results. In all cases, optimal rates of convergence are observed in L^∞ and L^2 norms.

2.1 Scheme Formulation, Stability, and Admissibility

First, partition the domain Ω into computational cells $\Omega = \bigcup_{j=1}^{N} I_j$, where $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $j = 1, ..., N$. The center of cell I_j is denoted by $x_j = \frac{1}{2}\left(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}}\right)$ and the size of the cell by $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. Denote $\Delta x = \max_j \Delta x_j$. We seek numerical solution u in the piecewise polynomial space defined as

$$\mathbb{V}_{\Delta x}^k := \left\{ v \in L^2(\Omega) : \quad v|_{I_j} \in P^k(I_j), \quad j = 1, \cdots, N \right\},$$

where $P^k(I_j)$ denotes the space of polynomials in I_j of degree at most k. We are now ready to formulate the DG scheme.

Multiply the heat equation (2.1) by any smooth function $V \in H^1(\Omega)$, integrate over I_j, and perform integration by parts to formally obtain

$$\int_{I_j} U_t V \, dx - U_x V \Big|_{j-1/2}^{j+1/2} + \int_{I_j} U_x V_x \, dx = 0,$$

where

$$U_x V \Big|_{j-1/2}^{j+1/2} = (U_x)_{j+1/2} V_{j+1/2} - (U_x)_{j-1/2} V_{j-1/2}.$$

Here, $(U_x)_{j\pm1/2}$ and $V_{j\pm1/2}$ denote the values of U_x and V at $x = x_{j\pm1/2}$ respectively.

Then, replace smooth function V by a test function $v \in \mathbb{V}_{\Delta x}^k$ and exact solution U by approximate solution $u \in \mathbb{V}_{\Delta x}^k$. Thus, as in [26], we have the original Direct Discontinuous Galerkin (DDG) scheme defined as follows: find the unique approximate solution $u \in \mathbb{V}_{\Delta x}^k$ such that for all test functions $v \in \mathbb{V}_{\Delta x}^k$ and all $1 \le j \le N$ we have that

$$\int_{I_j} u_t v \, dx - \widehat{u_x} v \Big|_{j-1/2}^{j+1/2} + \int_{I_j} u_x v_x \, dx = 0, \tag{2.2}$$

$$\int_{I_j} u(x,0) v(x) \, dx = \int_{I_j} U_0(x) v(x) \, dx, \tag{2.3}$$

where

$$\widehat{u_x} v \Big|_{j-1/2}^{j+1/2} = (\widehat{u_x})_{j+1/2} v_{j+1/2}^- - (\widehat{u_x})_{j-1/2} v_{j-1/2}^+.$$

This scheme is well defined provided that numerical flux $\widehat{u_x}$ is given. Motivated by the solution derivative trace formula of the heat equation with discontinuous initial data, in [26], the numerical flux was introduced as taking the form

$$\widehat{u_x} = \beta_0 \frac{[u]}{\Delta x} + \overline{u_x} + \beta_1 \Delta x [u_{xx}] + \beta_2 (\Delta x)^3 [u_{xxxx}] + \cdots . \tag{2.4}$$

Note, here and below we adopt the following notation.

$$u^{\pm} = \lim_{\epsilon \to 0^{\pm}} u(x+\epsilon, t), \quad [u] = u^{+} - u^{-}, \quad \overline{u} = \frac{u^{+} + u^{-}}{2}$$

Also, $\Delta x = \frac{\Delta x_j + \Delta x_{j+1}}{2}$ if the numerical flux is evaluated at the cell interface $x_{j+1/2}$. The numerical flux $\widehat{u_x}$ approximates U_x and involves the average $\overline{u_x}$ and the jumps of even order derivatives of u across the cell interfaces $x_{j+1/2}$. The coefficients $\beta_0, \beta_1, \beta_2, \cdots$ are chosen to ensure the stability and convergence of the method. Note, it is easily seen that this numerical flux is consistent in the sense that $\widehat{u_x} = u_x$ for smooth u. In addition this flux is conservative, meaning $\widehat{u_x}\big|_{j+1/2}$ is single valued. The direct use of the weak formulation of the PDE with suitable defined numerical flux gives name to this method as the direct DG (DDG) method. The solution derivative trace formula result is given in Appendix C, and a complete proof is provided as well. The original DDG paper [26] listed this result, but a proof was absent. Note, some odd order derivative averages are excluded from this result when defining the numerical flux in order to ensure consistency of the numerical scheme.

The observation that the derivative of the test function across cell interfaces contributes to the interface flux motivated a refinement of the original DDG scheme in [27]. Here a truncated version of the numerical flux (2.4) with $\beta_j = 0, j \geq 2$ was considered and interface values of the test function derivative v_x was added. The resulting scheme is known as DDG with interface corrections.

In order to theoretically guarantee optimal rates of convergence in the L^2 norm, a numerical flux term $\widehat{v_x}$ for the test function v is added to the original DDG method (2.2) to create a symmetric scheme. This test function numerical flux is exactly of the same form as $\widehat{u_x}$ given above in (2.4). The resulting new scheme, known as the symmetric DDG scheme, is formally defined as follows: find the unique approximate solution $u \in \mathbb{V}_{\Delta x}^k$ such that for all test functions $v \in \mathbb{V}_{\Delta x}^k$ and all $1 \leq j \leq N$ we have that

$$\int_{I_j} u_t v dx - \widehat{u_x} v \Big|_{j-1/2}^{j+1/2} + \int_{I_j} u_x v_x dx + ([u]\widehat{v_x})_{j+1/2} + ([u]\widehat{v_x})_{j-1/2} = 0, \tag{2.5}$$

with the numerical flux terms defined as

$$\begin{cases} \widehat{u_x} = \beta_0 \frac{[u]}{\Delta x} + \overline{u_x} + \beta_1 \Delta x [u_{xx}] \\ \widehat{v_x} = \beta_0 \frac{[v]}{\Delta x} + \overline{v_x} + \beta_1 \Delta x [v_{xx}]. \end{cases} \tag{2.6}$$

Notice we drop higher order terms in (2.4) and take a simpler numerical flux formula for $\widehat{u_x}$. Dropping these high order terms in the flux formula makes analysis considerably easier, though the same results would hold if high order terms were included. An identical numerical flux formula is used for the test function derivative $\widehat{v_x}$. Numerically we take the test function v to be nonzero only inside the cell I_j, thus only half of the terms in (2.6) contribute to the computation of $\widehat{v_x}$. Summing the scheme (2.5) over all cells I_j and introducing the bilinear form

$$\mathbb{B}(u,v) = \sum_{j=1}^{N} \int_{I_j} u_x v_x \, dx + \sum_{j=1}^{N} (\widehat{u_x}[v])_{j+1/2} + \sum_{j=1}^{N} ([u]\widehat{v_x})_{j+1/2}, \tag{2.7}$$

we obtain the primal weak formulation as the following,

$$\int_{\Omega} u_t v dx + \mathbb{B}(u,v) = 0. \tag{2.8}$$

It can be seen that the bilinear form $\mathbb{B}(u,v)$ has an obvious symmetry. That is, $\mathbb{B}(u,v) = \mathbb{B}(v,u)$. It is this feature which gives name to this method, the DDG method with symmetric structure.

Remark 2.1.1. *The structure of symmetric DDG method given by (2.5) is similar to the well known SIPG method by Arnold in [1]. In fact, for the $\beta_1 = 0$ case, the symmetric DDG method reduces exactly to SIPG with the penalty parameter taken as $\sigma = 2\beta_0$. Despite the similarities, the motivation for each of these methods is completely different. For symmetric DDG, the numerical flux is directly drawn from the solution gradient for the heat equation. For SIPG, knowledge that the analytic solution to these diffusion problems is smooth motivates introduction of the jumps of numerical solutions across cell interfaces as a penalty term.*

The so called multipenalty method is an extension of SIPG which was proposed by Arnold in [1] as a way to compensate for interpolation errors when considering adaptive mesh refinement. This extension includes penalties of jump high order derivatives of the numerical solution. This multipenalty method hasn't been utilized to my knowledge since its introduction.

In addition to method motivation, the inclusion of second order derivative jumps in the numerical flux of the symmetric DDG method seems to play a significant role. Also, recently in [19] Feng and Wu also explore the importance of higher order normal derivative jump terms of DG solutions for Helmholtz equations.

Up to now, we have taken the method of lines approach and have left time variable t continuous. For time discretization, the explicit third order TVD Runge-Kutta method [37, 36] was used in order to match the accuracy in space.

As for any DG method, the guiding principle for the choice of numerical flux is the stability requirement. We adopt the following admissibility criterion.

Definition 2.1.1. *(Admissibility) We call numerical flux $\widehat{u_x}$ in (2.6) admissible if there exists $\gamma \in (0,1)$, $\alpha > 0$ such that for any $u \in \mathbb{V}_{\Delta x}^k$,*

$$\gamma \sum_{j=1}^{N} \int_{I_j} u_x^2(x,t)\ dx + 2\sum_{j=1}^{N} \widehat{u_x}[u]_{j+1/2} \geq \alpha \sum_{j=1}^{N} \frac{[u]_{j+1/2}^2}{\Delta x}. \tag{2.9}$$

This admissibility is motivated by and ensures the following stability of the symmetric DDG method.

Theorem 2.1.1. *(Stability) Consider the symmetric DDG scheme (2.5). If the numerical flux (2.6) is admissible as described in (2.9), then we have*

$$\frac{1}{2}\int_{\Omega} u^2(x,T)dx + (1-\gamma)\int_0^T \sum_{j=1}^{N} \int_{I_j} u_x^2(x,t)dxdt + \alpha \int_0^T \sum_{j=1}^{N} \frac{[u]_{j+1/2}^2}{\Delta x} dt \leq \frac{1}{2}\int_{\Omega} U_0^2(x)dx. \tag{2.10}$$

Proof. Starting with the primal weak formulation (2.8), choose the test function $v = u$. Then, we have that

$$\frac{d}{dt}\int_{\Omega} \frac{u^2}{2}dx + \sum_{j=1}^{N} \int_{I_j} (u_x)^2\ dx + 2\sum_{j=1}^{N} (\widehat{u_x}[u])_{j+1/2} = 0.$$

Then, using the admissibility assumption,

$$\frac{d}{dt}\int_{\Omega} \frac{u^2}{2}dx + (1-\gamma)\sum_{j=1}^{N} \int_{I_j} (u_x)^2\ dx + \alpha \sum_{j=1}^{N} \frac{[u]_{j+1/2}^2}{\Delta x} \leq 0.$$

Last, integrate in time from 0 to T to get the needed result. □

Note, Theorem 2.10 implies our solution u is L^2 stable in the sense that $\|u(\cdot,T)\|_2 \leq \|U_0\|_2$. Note, here and below we denote the usual L^2 norm as

$$\|f\|_2 = \|f\|_{L^2(\Omega)} = \left(\int_\Omega f^2 \, dx\right)^{1/2}.$$

2.2 Admissibility Analysis

Next, we need to develop some guidelines which ensure that our chosen numerical flux is indeed admissible. Here, a uniform mesh will be considered. The theorem below provides this.

Theorem 2.2.1. *(Admissibility) The numerical flux $\widehat{u_x}$ (2.6) is admissible provided that*

$$\beta_0 > 4\left(\beta_1^2\left(\frac{k^2(k^2-1)^2}{3}\right) - \beta_1\left(\frac{k^2(k^2-1)}{2}\right) + \frac{k^2}{4}\right), \quad \text{for } k \geq 1. \tag{2.11}$$

To minimize β_0, we have

$$\beta_1^* = \frac{3}{4(k^2-1)} \quad \text{and} \quad \beta_0^* = \frac{k^2}{4}. \tag{2.12}$$

Here k is the degree of the approximate polynomial space $\mathbb{V}_{\Delta x}^k$. For $k = 0$ we require $\beta_0 = \frac{1}{2}$ for consistency with the finite difference method.

The next Lemmas provided the needed results to prove Theorem 2.2.1.

Lemma 2.2.1. *The numerical flux $\widehat{u_x}$ is admissible if for $k = 0$, $2\beta_0 = 1$ and for $k > 0$,*

$$\beta_0 > 2 \max_{u \in \mathbb{V}_{\Delta x}^k} \frac{\Delta x(\overline{u_x} + \beta_1 \Delta x[u_{xx}])_{j+1/2}^2}{\int_{I_j} u_x^2 \, dx + \int_{I_{j+1}} u_x^2 \, dx}, \qquad \forall j.$$

Proof. For $k = 0$, we only need $2\beta_0 \geq \alpha$ for some $\alpha > 0$. So, for $0 < \alpha < 1$, $\beta_0 = \frac{1}{2}$ is admissible. For $k > 0$, consider the following. First, take α small and $\gamma = \frac{1}{2}$. Below, all jumps and averages are evaluated at $x_{j+1/2}$.

$$\begin{aligned}
&\sum_{j=1}^N \left(\Delta x(2\beta_0 - \alpha)\gamma \int_{I_j} u_x^2 \, dx + (2\beta_0 - \alpha)^2[u]^2 + 2\Delta x(2\beta_0 - \alpha)(\overline{u}_x + \Delta x\beta_1[u_{xx}])[u]\right) \\
&\quad = \sum_{j=1}^N \left(\Delta x(2\beta_0 - \alpha)\frac{\gamma}{2}\left(\int_{I_j} u_x^2 \, dx + \int_{I_{j+1}} u_x^2 \, dx\right) + (2\beta_0 - \alpha)^2[u]^2\right. \\
&\qquad \left. + 2\Delta x(2\beta_0 - \alpha)(\overline{u}_x + \Delta x\beta_1[u_{xx}])[u]\right) \\
&\quad \geq \sum_{j=1}^N \left(\Delta x^2\left(\overline{u}_x + \Delta x\beta_1[u_{xx}]\right)^2 + (2\beta_0 - \alpha)^2[u]^2 + 2\Delta x(2\beta_0 - \alpha)(\overline{u}_x + \Delta x\beta_1[u_{xx}])[u]\right)
\end{aligned}$$

$$= \sum_{j=1}^{N} \left(\Delta x(\overline{u}_x + \Delta x \beta_1 [u_{xx}]) + (2\beta_0 - \alpha)[u]\right)^2 \geq 0$$

So, we have admissibility of this flux if

$$(2\beta_0 - \alpha)\frac{\gamma}{2}\left(\int_{I_j} u_x^2 \, dx + \int_{I_{j+1}} u_x^2 \, dx\right) \geq \Delta x \left(\overline{u}_x + \Delta x \beta_1 [u_{xx}]\right)^2, \quad \forall j.$$

The assumption that

$$\beta_0 > 2 \max_{u \in \mathbb{V}^k_{\Delta x}} \frac{\Delta x (\overline{u_x} + \beta_1 \Delta x [u_{xx}])^2_{j+1/2}}{\int_{I_j} u_x^2 \, dx + \int_{I_{j+1}} u_x^2 \, dx}, \qquad \forall j$$

implies that this is true, guaranteeing admissibility. □

Lemma 2.2.2. *Further, $\widehat{u_x}$ is admissible for $k > 0$ if*

$$\beta_0 > 4\rho(H^{-1}O)$$

where $H = \left(\frac{1}{i+j-1}\right)_{i,j=1}^{k}$ is a Hilbert matrix and $O = (O_{ij})_{i,j=1}^{k}$ for $O_{ij} = \left(\frac{1}{2} - \beta_1(i-1)\right)\left(\frac{1}{2} - \beta_1(j-1)\right)$. Also, here $\rho(A)$ denotes the spectral radius of matrix A.

Proof. Note first that

$$\max_{u \in \mathbb{V}^k_{\Delta x}} \frac{\Delta x (\overline{u_x} + \beta_1 \Delta x [u_{xx}])^2_{j+1/2}}{\int_{I_j} u_x^2 \, dx + \int_{I_{j+1}} u_x^2 \, dx} = \max_{v \in \mathbb{V}^{k-1}_{\Delta x}} \frac{\Delta x (\overline{v} + \beta_1 \Delta x [v_x])^2_{j+1/2}}{\int_{I_j} v^2 \, dx + \int_{I_{j+1}} v^2 \, dx}.$$

Being that v is a piecewise polynomial, denote $v_L = v\big|_{I_j}$ and $v_R = v\big|_{I_{j+1}}$. To make things cleaner, use the change of variables $\xi = \frac{x - x_{j-1/2}}{\Delta x}$ for v_L and $\xi = \frac{x_{j+3/2} - x}{\Delta x}$ for v_R. Then we have that $\xi \in [0,1]$ in both cases. Then, we have the following.

$$\begin{aligned}
&\max_{v \in \mathbb{V}^{k-1}_{\Delta x}} \frac{\Delta x (\overline{v} + \beta_1 \Delta x [v_x])^2_{j+1/2}}{\int_{I_j} v^2 \, dx + \int_{I_{j+1}} v^2 \, dx} \\
&= \max_{\substack{v_L \in P^{k-1}(I_j) \\ v_R \in P^{k-1}(I_{j+1})}} \frac{\Delta x \left(\frac{v_L}{2} + \frac{v_R}{2} + \Delta x \beta_1 (\partial_x v_R - \partial_x v_L)\right)^2_{j+1/2}}{\int_{I_j} v_L^2 \, dx + \int_{I_{j+1}} v_R^2 \, dx} \\
&= \max_{v_L, v_R \in P^{k-1}[0,1]} \frac{\Delta x \left(\frac{v_L(1)}{2} + \frac{v_R(1)}{2} - \beta_1(\partial_\xi v_R(1) + \partial_\xi v_L(1))\right)^2}{\Delta x \int_0^1 v_L^2 \, d\xi + \Delta x \int_0^1 v_R^2 \, d\xi} \\
&= \max_{v_L, v_R \in P^{k-1}[0,1]} \frac{\left(\left(\frac{v_L(1)}{2} - \beta_1 \partial_\xi v_L(1)\right) + \left(\frac{v_R(1)}{2} - \beta_1 \partial_\xi v_R(1)\right)\right)^2}{\int_0^1 v_L^2 \, d\xi + \int_0^1 v_R^2 \, d\xi}
\end{aligned}$$

Then, consider $v_L = \sum_{m=1}^{k} a_m \xi^{m-1}$ and $v_R = \sum_{m=1}^{k} b_m \xi^{m-1}$. Note that

$$v_L(1) = \sum_{m=1}^{k} a_m, \qquad \partial_\xi v_L(1) = \sum_{m=1}^{k} (m-1) a_m, \qquad \int_0^1 v_L^2 \, d\xi = \sum_{\ell,m=1}^{k} a_\ell a_m \frac{1}{\ell+m-1}.$$

Similar computations hold for v_R. Continuing the above, we have the following.

$$\begin{aligned}
&\max_{v_L, v_R \in P^{k-1}[0,1]} \frac{\left(\left(\frac{v_L(1)}{2} - \beta_1 \partial_\xi v_L(1)\right)\right) + \left(\left(\frac{v_R(1)}{2} - \beta_1 \partial_\xi v_R(1)\right)\right)^2}{\int_0^1 v_L^2 \, d\xi + \int_0^1 v_R^2 \, d\xi} \\
&= \max_{\vec{a}, \vec{b} \in \mathbb{R}^k} \frac{\left(\sum_{\ell=1}^{k} a_\ell \left(\frac{1}{2} - \beta_1(\ell-1)\right) + \sum_{\ell=1}^{k} b_\ell \left(\frac{1}{2} - \beta_1(\ell-1)\right)\right)^2}{\sum_{\ell,m=1}^{k} a_\ell a_m \frac{1}{\ell+m-1} + \sum_{\ell,m=1}^{k} b_\ell b_m \frac{1}{\ell+m-1}} \\
&= \max_{\vec{a}, \vec{b} \in \mathbb{R}^k} \frac{\sum_{\ell,m=1}^{k} \left(a_\ell a_m + 2 a_\ell b_m + b_\ell b_m\right) \left(\frac{1}{2} - \beta_1(\ell-1)\right) \left(\frac{1}{2} - \beta_1(m-1)\right)}{\sum_{\ell,m=1}^{k} a_\ell a_m \frac{1}{\ell+m-1} + \sum_{\ell,m=1}^{k} b_\ell b_m \frac{1}{\ell+m-1}} \\
&= \max_{\vec{a}, \vec{b} \in \mathbb{R}^k} \frac{(\vec{a}+\vec{b})^T O (\vec{a}+\vec{b})}{\vec{a}^T H \vec{a} + \vec{b}^T H \vec{b}} \\
&= \max_{\vec{x}, \vec{y} \in \mathbb{R}^k} \frac{(\vec{x}+\vec{y})^T (H^{-1/2} O H^{-1/2})(\vec{x}+\vec{y})}{\|\vec{x}\|_2^2 + \|\vec{y}\|_2^2} \\
&\leq \max_{\vec{x}, \vec{y} \in \mathbb{R}^k} \frac{(\vec{x}+\vec{y})^T (H^{-1/2} O H^{-1/2})(\vec{x}+\vec{y})}{\frac{1}{2}\|\vec{x}+\vec{y}\|_2^2} \\
&= 2 \max_{\vec{z} \in \mathbb{R}^k} \frac{\vec{z}^T (H^{-1/2} O H^{-1/2}) \vec{z}}{\|\vec{z}\|_2^2} \\
&= 2\rho(H^{-1/2} O H^{-1/2}) \\
&= 2\rho(H^{-1} O)
\end{aligned}$$

Above, the transformation $\vec{x} = H^{1/2}\vec{a}$, $\vec{y} = H^{1/2}\vec{b}$ and the fact that the Hilbert matrix H is symmetric and positive definite was used. Combining this with Lemma 2.2.1, we get the needed result. □

Lemma 2.2.3. *(Hilbert Matrix Properties) For the $k \times k$ Hilbert matrix $H = (H_{ij})$, $H_{ij} = \frac{1}{i+j-1}$, we have the following.*

$$\left(H^{-1}\right)_{ij} = (-1)^{i+j}(i+j-1)\binom{k+i-1}{k-j}\binom{k+j-1}{k-i}\binom{i+j-2}{i-1}^2 \tag{2.13}$$

$$\sum_{i,j=1}^{k} \left(H^{-1}\right)_{ij} = k^2 \tag{2.14}$$

$$\sum_{j=1}^{k} j \left(H^{-1}\right)_{ij} = (k^2 - i + 1) \sum_{j=1}^{k} \left(H^{-1}\right)_{ij} \tag{2.15}$$

$$\sum_{i,j=1}^{k} i\left(H^{-1}\right)_{ij} = \sum_{i,j=1}^{k} j\left(H^{-1}\right)_{ij} = \frac{k^2(k^2+1)}{2} \tag{2.16}$$

$$\sum_{i,j=1}^{k} i^2\left(H^{-1}\right)_{ij} = \sum_{i,j=1}^{k} j^2\left(H^{-1}\right)_{ij} = \frac{k^2(k^4+4k^2+1)}{6} \tag{2.17}$$

For a proof of (2.13) and (2.14), we cite work done in [34] and [39]. Properties (2.15), (2.16), and (2.17) to my knowledge are new results and are shown in Appendix B.

Lemma 2.2.4.

$$\rho(H^{-1}O) = \beta_1^2\left(\frac{k^2(k^2-1)^2}{3}\right) - \beta_1\left(\frac{k^2(k^2-1)}{2}\right) + \frac{k^2}{4}$$

Proof. Multiplying term by term, we have that each entry of $H^{-1}O$ is given by

$$\begin{aligned}(H^{-1}O)_{ij} &= \sum_{m=1}^{k}(-1)^{i+m}(i+m-1)\binom{k+i-1}{k-m}\binom{k+m-1}{k-i}\binom{i+m-2}{i-1}^2 \\ &\qquad \left(\frac{1}{2}-\beta_1(m-1)\right)\left(\frac{1}{2}-\beta_1(j-1)\right) \\ &= c_i d_j\end{aligned}$$

where

$$\begin{aligned}c_i &= \sum_{m=1}^{k}(-1)^{i+m}(i+m-1)\binom{k+i-1}{k-m}\binom{k+m-1}{k-i}\binom{i+m-2}{i-1}^2\left(\frac{1}{2}-\beta_1(m-1)\right) \\ d_j &= \frac{1}{2}-\beta_1(j-1).\end{aligned}$$

Then, $H^{-1}O = \vec{c}\vec{d}^T$ where $\vec{c} = (c_1, c_2, \ldots, c_k)^T$ and $\vec{d} = (d_1, d_2, \ldots, d_k)^T$. Note that using (2.15), each c_i can be rewritten as

$$\begin{aligned}c_i &= \sum_{m=1}^{k}(-1)^{i+m}(i+m-1)\binom{k+i-1}{k-m}\binom{k+m-1}{k-i}\binom{i+m-2}{i-1}^2\left(\frac{1}{2}-\beta_1(m-1)\right) \\ &= \left(\frac{1}{2}+\beta_1\right)\sum_{m=1}^{k}\left(H^{-1}\right)_{im} - \beta_1\sum_{m=1}^{k} m\left(H^{-1}\right)_{im} \\ &= \left(\frac{1}{2}+\beta_1\right)\sum_{m=1}^{k}\left(H^{-1}\right)_{im} - \beta_1(k^2-i+1)\sum_{m=1}^{k}\left(H^{-1}\right)_{im} \\ &= \left(\frac{1}{2}-\beta_1(k^2-i)\right)\sum_{m=1}^{k}\left(H^{-1}\right)_{im}\end{aligned}$$

So, $H^{-1}O$ is a rank one matrix. It is known that rank one matrices have one non-zero eigenvalue given in this case by $\vec{d}^T\vec{c}$ with corresponding eigenvector $\vec{c}$ [22]. A quick justification of this nonzero eigenvalue is given below.

$$H^{-1}O\vec{c} = \left(\vec{c}\vec{d}^T\right)\vec{c} = \vec{c}\left(\vec{d}^T\vec{c}\right) = \left(\vec{d}^T\vec{c}\right)\vec{c} = \lambda\vec{c}$$

Then, the spectrum is given explicitly as

$$\begin{aligned}
\rho(H^{-1}O) = \vec{d}^T\vec{c} &= \sum_{\ell=1}^{k} c_\ell d_\ell \\
&= \sum_{\ell=1}^{k}\left(\frac{1}{2}-\beta_1(\ell-1)\right)\left(\left(\frac{1}{2}-\beta_1(k^2-\ell)\right)\sum_{m=1}^{k}\left(H^{-1}\right)_{\ell m}\right) \\
&= \sum_{\ell=1}^{k}\left(\left(\left(\frac{1}{4}-\frac{k^2}{2}\beta_1+\frac{1}{2}\beta_1-k^2\beta_1^2\right)+\left(\beta_1^2k^2+\beta_1^2\right)\ell-\beta_1^2\ell^2\right)\sum_{m=1}^{k}\left(H^{-1}\right)_{\ell m}\right) \\
&= \left(\frac{1}{4}-\frac{k^2}{2}\beta_1+\frac{1}{2}\beta_1-k^2\beta_1^2\right)\sum_{\ell,m=1}^{k}\left(H^{-1}\right)_{\ell m}+\left(\beta_1^2k^2+\beta_1^2\right)\sum_{\ell,m=1}^{k}\ell\left(H^{-1}\right)_{\ell m} \\
&\quad -\beta_1^2\sum_{\ell,m=1}^{k}\ell^2\left(H^{-1}\right)_{\ell m} \\
&= \left(\frac{1}{4}-\frac{k^2}{2}\beta_1+\frac{1}{2}\beta_1-k^2\beta_1^2\right)k^2+\left(\beta_1^2k^2+\beta_1^2\right)\frac{k^2(k^2+1)}{2}-\beta_1^2\frac{k^2(k^4+4k^2-1)}{6} \\
&= \beta_1^2\left(\frac{k^2(k^2-1)^2}{3}\right)-\beta_1\left(\frac{k^2(k^2-1)}{2}\right)+\frac{k^2}{4}.
\end{aligned}$$

Here, special sums (2.14), (2.16), and (2.17) were made use of. □

Proof. (of Theorem 2.2.1) Combining the results of Lemmas 2.2.1, 2.2.2, and 2.2.4, the first inequality is clear. For the minimized β_0^* and corresponding β_1^*, the function

$$g(\beta_1) = \beta_1^2\left(\frac{k^2(k^2-1)^2}{3}\right)-\beta_1\left(\frac{k^2(k^2-1)}{2}\right)+\frac{k^2}{4}$$

needs to be minimized. Function g is just a concave up parabola in β_1, so there is only one minimum β_1^* which can be computed by considering $g'(\beta_1^*)=0$. We have that $\beta_1^* = \frac{3}{4(k^2-1)}$ and $g(\beta_1^*) = \frac{k^2}{16}$. Last, note that this $\beta_1^* > 0$ for all $k > 1$ as needed. For $k = 1$, $g(\beta_1) = \frac{1}{4}$ and no minimization is needed. □

The admissibility Theorem 2.2.1 provides a way to choose suitable β_0 and β_1 in the numerical flux formula (2.6). From (2.11) we see any (β_0, β_1) pair that falls in the parabolic shaded regions

in Figure 2.1 leads to an admissible numerical flux. The minimized (β_0^*, β_1^*) pair of (2.12) is listed in Table 2.1 with $k = 0, \cdots, 10$. For numerical tests in §2.5, Table 2.1 is used for choosing (β_0, β_1) pairs.

Table 2.1 Minimized admissible (β_0^*, β_1^*) for one dimension heat equation

k	0	1	2	3	4	5	6	7	8	9	10
β_0^*	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{11}{4}$	$\frac{9}{2}$	$\frac{27}{4}$	$\frac{19}{2}$	$\frac{51}{4}$	$\frac{33}{2}$	$\frac{83}{4}$	$\frac{51}{2}$
β_1^*	0	0	$\frac{1}{4}$	$\frac{3}{32}$	$\frac{1}{20}$	$\frac{1}{32}$	$\frac{3}{140}$	$\frac{1}{64}$	$\frac{1}{84}$	$\frac{3}{320}$	$\frac{1}{132}$

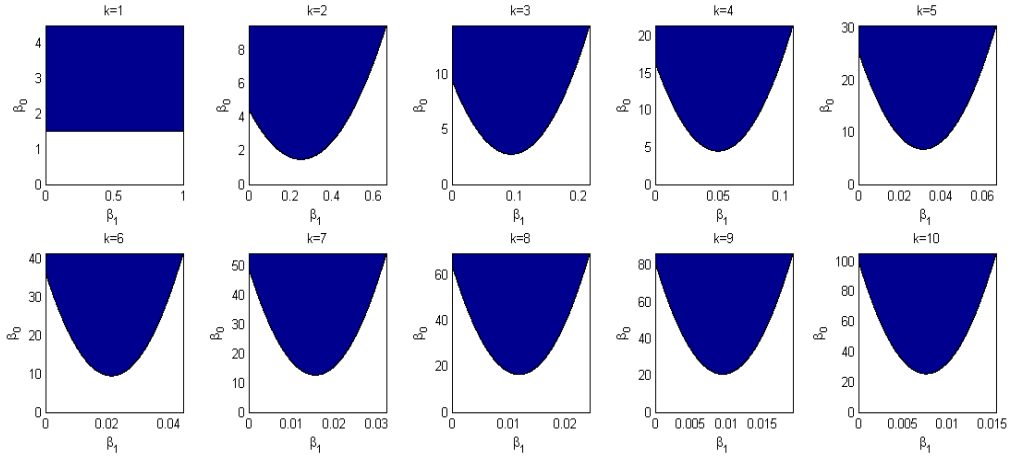

Figure 2.1 Admissibility regions for one dimensional heat equation

Note that in each polynomial degree $k \geq 2$ case, having β_1 nonzero allows for the smallest choice of β_0. Compared to the SIPG method [1], the penalty coefficient ($\sigma = 2\beta_0$ with $\beta_1 = 0$ case) can be decreased from k^2 to four times smaller as $\frac{k^2}{4}$ (here we take $\gamma = 1$ as for the elliptic case). If $\beta_1 = 0$, then a quick calculation shows that β_0 needs to be considerably larger in order to ensure admissibility. Table 2.2 shows admissible values of β_0 where $\beta_1 = 0$.

Table 2.2 Admissible β_0 for $\beta_1 = 0$ for one dimension heat equation

k	0	1	2	3	4	5	6	7	8	9	10
β_0	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{9}{2}$	$\frac{19}{2}$	$\frac{33}{2}$	$\frac{51}{2}$	$\frac{73}{2}$	$\frac{99}{2}$	$\frac{129}{2}$	$\frac{163}{2}$	$\frac{201}{2}$

Remark 2.2.1. *As mentioned above, the symmetric DDG scheme reduces exactly to the SIPG method when $\beta_1 = 0$. So this admissibility analysis provides guidelines for choosing penalty*

parameter $\sigma = 2\beta_0$ for SIPG as well. Estimates for the SIPG penalty parameter σ for the elliptic case have been obtained by Epshteyn and Riviere in 2007 [17]. This work has a very different approach than outlined above. The basic idea was to find the smallest necessary σ in order to ensure coercivity of the bilinear form within the numerical scheme. This relies on explicit constants depending on polynomial power and mesh information in trace inverse inequalities as given by Warburton and Hesthaven in 2003 [45].

Remark 2.2.2. *This same admissibility analysis applies to the original DDG [26] as well as DDG with interface correction [27] with minor modification.*

2.3 Energy Norm Error Estimate

Next, we will study the energy norm error estimate for the linear diffusion equation (2.1). Denote the energy norm associated with this scheme as being

$$|||u(\cdot,t)||| := \left(\int_\Omega u^2 \, dx + (1-\gamma) \int_0^t \sum_{j=1}^N \int_{I_j} u_x^2 \, dx d\tau + \alpha \int_0^t \sum_{j=1}^N \frac{[u]_{j+1/2}^2}{\Delta x} \, d\tau \right)^{1/2} \tag{2.18}$$

with $\gamma \in (0,1)$ and $\alpha > 0$ from (2.9). The form of this energy norm is inspired by the stability estimate (2.10). Before carrying on the error estimate, we first list the following two Lemmas as approximation properties of the finite element space $\mathbb{V}_{\Delta x}^k$ as in reference [6]. These results are used multiple times in both the energy estimate below as well as the $L^2(L^2)$ error analysis.

Before they are stated, we need the following notation. Let $W^{k,p}(\Omega)$ denote the Sobolev space, $\|f\|_{W^{k,p}(\Omega)}$ be the associated Sobolev norm and $|f|_{W^{k,p}(\Omega)}$ as the seminorm consisting of the L^p norms of the highest order derivatives. Also, denote $H^2(\Omega)$ as being the regular Hilbert space $W^{k,2}(\Omega)$.

Lemma 2.3.1. *(Approximation property [6]) Let $K \subset \mathbb{R}^n$ be any regular element in the sense that $\rho\Delta x \leq$ diam $(K) \leq \Delta x$ for some constant ρ. Let $U \in W^{k+1,p}(\Omega)$ and $\mathbb{P}(U)$ be the L^2 projection of U in $\mathbb{V}_{\Delta x}^k$. Then we have the following approximation property,*

$$||U - \mathbb{P}(U)||_{W^{m,q}(K)} \leq c_k (\Delta x)^{n/q - n/p} |U|_{W^{k+1,p}(K)} (\Delta x)^{k+1-m}. \tag{2.19}$$

Here $p, q \in [1,\infty]$, $m \geq 0$ and $k \geq 0$ are integers, and the constant c_k solely depends on k.

Here the L^2 projection of U, $\mathbb{P}(U)$, is defined as the unique function in $\mathbb{V}_{\Delta x}^k$ such that for all $v \in \mathbb{V}_{\Delta x}^k$ and all j

$$\int_{I_j} (U - \mathbb{P}(U))v \, dx = 0.$$

Lemma 2.3.2. *(Inverse inequality [6]) Given the finite dimensional piecewise polynomial space $\mathbb{V}_{\Delta x}^k$, and $1 \le p \le \infty, 1 \le q \le \infty$ and $0 \le m \le l$, and a regular element $K \subset \mathbb{R}^n$, there exists C independent of Δx such that for all $v \in \mathbb{V}_{\Delta x}^k$, we have*

$$\|v\|_{W^{m,q}(K)} \le C\Delta x^{l-m+n/q-n/p}\|v\|_{W^{l,p}(K)}. \tag{2.20}$$

Theorem 2.3.1. *(Energy norm error estimate) Let $e := u - U$ be the error between the exact solution U and the numerical solution u of the symmetric DDG method (2.5). If the numerical flux (2.6) is admissible as defined in (2.9), then the energy norm of the error satisfies the inequality*

$$|||e(\cdot, T)||| \le C|||\partial_x^{k+1}U(\cdot, T)|||(\Delta x)^k, \tag{2.21}$$

where $C = C(k, \gamma, \alpha)$ is a constant depending on k, γ, α but is independent of U and Δx.

Proof. First, rewrite the error as

$$e = u - U = u - \mathbb{P}(U) + \mathbb{P}(U) - U = \mathbb{P}(e) - (U - \mathbb{P}(U)). \tag{2.22}$$

Here $\mathbb{P}(U)$ denotes the L^2 projection of U into $\mathbb{V}_{\Delta x}^k$. Then, with (2.22) we have

$$|||e(\cdot, T)||| \le |||\mathbb{P}(e)(\cdot, T)||| + |||(U - \mathbb{P}(U))(\cdot, T)|||.$$

Using the standard polynomial projection estimate of $\mathbb{V}_{\Delta x}^k$ as in Lemma 2.3.1 and the definition of the energy norm (2.18), we have

$$|||(U - \mathbb{P}(U))(\cdot, T)||| \le C|||\partial_x^{k+1}U(\cdot, T)|||(\Delta x)^k. \tag{2.23}$$

Thus, we only need to find a bound for $|||\mathbb{P}(e)(\cdot, T)|||$. Notice here and below we use capital letter C to represent a generic constant. Define the bilinear form $\mathbb{C}(\cdot, \cdot)$ as

$$\mathbb{C}(w, v) = \int_0^T \int_\Omega w_t v \, dxdt + \int_0^T \sum_{j=1}^N \int_{I_j} w_x v_x \, dxdt + \Theta(T, w, v),$$

with

$$\Theta(T,w,v)=\int_0^T\sum_{j=1}^N(\widehat{w_x}[v])_{j+1/2}\,dt+\int_0^T\sum_{j=1}^N(\widehat{v_x}[w])_{j+1/2}\,dt.$$

From the scheme definition (2.5), we have $\mathbb{C}(u,v)=0$ and $\mathbb{C}(U,v)=0$ for any test function $v\in\mathbb{V}^k_{\Delta x}$. This implies $\mathbb{C}(e,v)=0$ for all such v as well. With (2.22), then we have that $\mathbb{C}(\mathbb{P}(e),v)=\mathbb{C}(U-\mathbb{P}(U),v)$ for all $v\in\mathbb{V}^k_{\Delta x}$. Taking $v=u-\mathbb{P}(U)=\mathbb{P}(e)$, we then have

$$\mathbb{C}(\mathbb{P}(e),\mathbb{P}(e))=\mathbb{C}(U-\mathbb{P}(U),\mathbb{P}(e)). \tag{2.24}$$

For the left hand side of (2.24), noting that $\mathbb{P}(e)(\cdot,0)=0$ in addition to the admissibility condition (2.9) for the numerical flux, we obtain

$$\begin{aligned}\mathbb{C}(\mathbb{P}(e),\mathbb{P}(e))&=\frac{1}{2}\|\mathbb{P}(e)(\cdot,T)\|^2+\int_0^T\sum_{j=1}^N\|\mathbb{P}(e)_x(\cdot,t)\|^2_{I_j}\,dt+\Theta(T,\mathbb{P}(e),\mathbb{P}(e))\\&\geq\frac{1}{2}\|\mathbb{P}(e)(\cdot,T)\|^2+\int_0^T\sum_{j=1}^N\|\mathbb{P}(e)_x(\cdot,t)\|^2_{I_j}\,dt+\alpha\int_0^T\sum_{j=1}^N\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}\,dt\\&\quad-\gamma\int_0^T\sum_{j=1}^N\|\mathbb{P}(e)_x(\cdot,t)\|^2_{I_j}\,dt\\&=|||\mathbb{P}(e)(\cdot,T)|||^2-\frac{1}{2}\|\mathbb{P}(e)(\cdot,T)\|^2\end{aligned}$$

So we have that

$$\mathbb{C}(\mathbb{P}(e),\mathbb{P}(e))\geq|||\mathbb{P}(e)(\cdot,T)|||^2-\frac{1}{2}||\mathbb{P}(e)(\cdot,T)||^2. \tag{2.25}$$

Now, turn the attention to the right hand side of (2.24),

$$\mathbb{C}(U-\mathbb{P}(U),\mathbb{P}(e))=\int_0^T\int_\Omega(U-\mathbb{P}(U))_t\mathbb{P}(e)\,dxdt+\int_0^T\sum_{j=1}^N\int_{I_j}(U-\mathbb{P}(U))_x\mathbb{P}(e)_x\,dxdt+\Theta(T,U-\mathbb{P}(U),\mathbb{P}(e)) \tag{2.26}$$

with

$$\begin{aligned}\Theta(T,U-\mathbb{P}(U),\mathbb{P}(e))&=\int_0^T\sum_{j=1}^N\left(\widehat{(U-\mathbb{P}(U))_x}[\mathbb{P}(e)]\right)_{j+1/2}\,dt\\&\quad+\int_0^T\sum_{j=1}^N\left([U-\mathbb{P}(U)]\widehat{\mathbb{P}(e)_x}\right)_{j+1/2}\,dt\\&=I_1+I_2.\end{aligned}$$

For the first term in (2.26), because $\mathbb{P}(e) \in \mathbb{V}_{\Delta x}^k$, by the projection definition we have

$$\int_0^T \int_\Omega (U - \mathbb{P}(U))_t \mathbb{P}(e)\ dxdt = 0.$$

For the second term of (2.26),

$$\int_0^T \sum_{j=1}^N \int_{I_j} (U - \mathbb{P}(U))_x \mathbb{P}(e)_x\ dxdt \le C|||\partial_x^{k+1} U(\cdot,T)|||\Delta x^k + \frac{1-\gamma}{4} \int_0^T \sum_{j=1}^N \|\mathbb{P}(e)_x\|_{I_j}^2 dt.$$

Here Cauchy's inequality and Lemma 2.3.1 was used. For the third term of $\Theta(T, U - \mathbb{P}(U), \mathbb{P}(e))$, we estimate I_1 and I_2 separately.

$$\begin{aligned} I_1 &= \int_0^T \sum_{j=1}^N \left(\widehat{(U - \mathbb{P}(U))}_x [\mathbb{P}(e)] \right)_{j+1/2}\ dt \\ &\le \epsilon_1 \int_0^T \sum_{j=1}^N \frac{[\mathbb{P}(e)]_{j+1/2}^2}{\Delta x}\ dt + C|||\partial_x^{k+1} U(\cdot,T)|||^2 (\Delta x)^{2k} \end{aligned}$$

Here, Cauchy's inequality along with approximation property (2.3.1) were used. Now, bound I_2.

$$\begin{aligned} I_2 &= \int_0^T \sum_{j=1}^N \left(\widetilde{\mathbb{P}(e)}_x [U - \mathbb{P}(U)] \right)_{j+1/2}\ dt \\ &\le \epsilon_2 \int_0^T \sum_{j=1}^N \|\mathbb{P}(e)_x\|_{I_j}^2\ dt + \epsilon_3 \int_0^T \sum_{j=1}^N \frac{[\mathbb{P}(e)]_{j+1/2}^2}{\Delta x}\ dt + C|||\partial_x^{k+1} U(\cdot,T)|||^2 (\Delta) x^{2k} \end{aligned}$$

Each term is bounded individually by first using Cauchy's inequality, then using appropriate inverse inequalities (2.3.2) and approximation property (2.3.1). Next, choose $\epsilon_1, \epsilon_2, \epsilon_3$ terms from Cauchy's inequality to match terms in the energy norm.

Now we have the right hand side of (2.26) as

$$\mathbb{C}(U - \mathbb{P}(U), \mathbb{P}(e))) \le \frac{1}{2}|||\mathbb{P}(e)(\cdot,T)|||^2 - \frac{1}{2}||\mathbb{P}(e)(\cdot,T)||^2 + C|||\partial_x^{k+1} U(\cdot,T)|||^2 (\Delta x)^{2k}.$$

With the left hand side estimate (2.25), we have that $|||\mathbb{P}(e)(\cdot,T)||| \le C|||\partial_x^{k+1} U(\cdot,T)|||(\Delta x)^k$. The needed result follows. □

2.4 $L^2(L^2)$ Error Estimate

In this section, we carry out the $L^2(L^2)$ *a priori* error analysis of the symmetric DDG method (2.5)-(2.6) for the 1-D model equation (2.1). Optimal $(k+1)$th order of accuracy

is obtained with P^k polynomial approximations. For simplicity of presentation, the uniform partition of the computational domain with mesh size Δx assumed. Also, the letter C is used to represent a generic constant. The $L^2(L^2)$ error estimate for two-dimensional linear diffusion equation is a straight forward extension and is provided as Theorem 4.1.3 in §4.1.

Remark 2.4.1. *The general technique used below is not new and has been applied in a number of situations for SIPG methods among others. The primary reference used is the thesis of Shuyu Sun [40], but this technique dates back to the early 80s.*

2.4.1 Symmetric DDG $L^2(L^2)$ Error Estimate

Theorem 2.4.1. *(Symmetric DDG $L^2(L^2)$ error estimate in one dimension) Consider the 1-D linear model equation (2.1). Let $e := u - U$ be the error between the exact solution U and the numerical solution u of the symmetric DDG method (2.5)-(2.6), then we have*

$$\|e\|_{L^2(0,T;L^2)} \leq C\Delta x^{k+1}\left(\|U\|_{L^\infty(0,T;H^{k+1})} + \|U\|_{L^2(0,T;H^{k+1})} + \Delta x\|U_t\|_{L^2(0,T;H^k)}\right).$$

Proof. We carry out the proof of the theorem in three steps. First, we apply the parabolic lift Theorem 2.4.2 below and obtain the following.

$$\begin{aligned}\|e\|_{L^2(0,T;L^2)} \leq & C\Delta x\left(\|e\|_{L^\infty(0,T;L^2)} + \|e_x\|_{L^2(0,T;L^2)} + \left(\int_0^T \sum_{j=1}^N \frac{[e]_{j+1/2}^2}{\Delta x}\,dt\right)^{1/2}\right) \\ & + C\Delta x^2\|e_t\|_{L^2(0,T;L^2)} + C\Delta x^{3/2}\left(\int_0^T \sum_{j=1}^N (\overline{e_x})_{j+1/2}^2\,dt\right)^{1/2} \\ & + C\Delta x^{5/2}\left(\int_0^T \sum_{j=1}^N [e_{xx}]_{j+1/2}^2\,dt\right)^{1/2}\end{aligned}$$

Second, we apply the time derivative estimate Theorem 2.4.3 and the interface error estimate Theorem 2.4.4 to bound the $\|e_t\|_{L^2(0,T;L^2)}$ term and the higher derivative interface terms and obtain

$$\begin{aligned}\|e\|_{L^2(0,T;L^2)} \leq & C\Delta x\left(\|e\|_{L^\infty(0,T;L^2)} + \|e_x\|_{L^2(0,T;L^2)} + \left(\int_0^T \sum_{j=1}^N \frac{[e]_{j+1/2}^2}{\Delta x}\,dt\right)^{1/2}\right) \\ & + C\Delta x^{k+1}\|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^{k+2}\|U_t\|_{L^2(0,T;H^k)} + C\Delta x^{k+1}\|U\|_{L^\infty(0,T;H^{k+1})}.\end{aligned}$$

Finally, we combine the energy estimate provided by Theorem 2.3.1 and the above estimate to complete the proof. Note we use the equivalency of ℓ^1 and ℓ^2 norms for the last step estimate. □

2.4.2 Parabolic lift for symmetric DDG method

The $L^2(L^2)$ error estimate is enhanced to optimal $(k+1)$th convergence rates through the following parabolic lift theorem.

Theorem 2.4.2. *(Parabolic lift) Let $e := u - U$ be the error of the symmetric DDG method (2.5)-(2.6). Assume $e(\cdot, t) \in H^k(I_j)$, for all j. We then have*

$$\begin{aligned}\|e\|_{L^2(0,T;L^2)} \le C\Delta x \left(\|e\|_{L^\infty(0,T;L^2)} + \|e_x\|_{L^2(0,T;L^2)}\right) + C\Delta x^2 \|e_t\|_{L^2(0,T;L^2)} \\ + C\Delta x^{1/2} \left(\int_0^T \sum_{j=1}^N [e]_{j+1/2}^2 \, dt\right)^{1/2} + C\Delta x^{3/2} \left(\int_0^T \sum_{j=1}^N (\overline{e_x})_{j+1/2}^2 \, dt\right)^{1/2} \\ + C\Delta x^{5/2} \left(\int_0^T \sum_{j=1}^N [e_{xx}]_{j+1/2}^2 \, dt\right)^{1/2}.\end{aligned}$$

Proof. Let's consider the following dual (backward) problem,

$$\begin{cases} -\Phi_t - \Phi_{xx} = e, & x \in \Omega, t \in [0,T] \\ \Phi = 0, & x \in \partial\Omega, t \in [0,T] \\ \Phi = 0, & x \in \Omega, t = T. \end{cases}$$

By dual regularity, there exists a unique solution Φ to this backward problem such that the following result holds,

$$\|\Phi\|_{L^\infty(0,T;H^1)} + \|\Phi\|_{L^2(0,T;H^2)} \le C\|e\|_{L^2(0,T;L^2)}. \tag{2.27}$$

As a reference for this result, see [18]. Then rewriting $\|e(\cdot,t)\|_{L^2}$ in terms of Φ, we have

$$\begin{aligned}
\|e(\cdot,t)\|_{L^2}^2 &= \sum_{j=1}^{N}\int_{I_j} e^2\ dx = \sum_{j=1}^{N}\int_{I_j} e\left(-\Phi_t - \Phi_{xx}\right)\ dx \\
&= -\frac{d}{dt}\sum_{j=1}^{N}\int_{I_j} e\Phi\ dx + \sum_{j=1}^{N}\int_{I_j} e_t\Phi\ dx + \sum_{j=1}^{N}\int_{I_j} e_x\Phi_x\ dx - \sum_{j=1}^{N} e\Phi_x\Big|_{j-1/2}^{j+1/2} \\
&= -\frac{d}{dt}\sum_{j=1}^{N}\int_{I_j} e\Phi\ dx + \sum_{j=1}^{N}\int_{I_j} e_t\Phi\ dx + \sum_{j=1}^{N}\int_{I_j} e_x\Phi_x\ dx + \sum_{j=1}^{N}\Phi_x[e]_{j+1/2} \\
&= -\frac{d}{dt}\sum_{j=1}^{N}\int_{I_j} e\Phi\ dx + \sum_{j=1}^{N}\int_{I_j} e_t\Phi\ dx + \mathbb{B}(e,\Phi) - \beta_1\Delta x\sum_{j=1}^{N}[\Phi_{xx}][e]_{j+1/2}.
\end{aligned}$$

These last two steps hold because $\Phi \in \mathbb{C}^{1,\alpha}(\Omega)$ from the regularity of the dual problem. Here, the bilinear form $\mathbb{B}(e,\Phi)$ is as defined in (2.7). Let $\mathbb{P}(\Phi)$ denote the L^2 projection of Φ into $\mathbb{V}_{\Delta x}^k$. Since $\mathbb{P}(\Phi) \in \mathbb{V}_{\Delta x}^k$, from the scheme primal formulation (2.8) we have $\langle e_t, \mathbb{P}(\Phi)\rangle + \mathbb{B}(e,\mathbb{P}(\Phi)) = 0$. Here, we use the notation $\langle w, v\rangle = \langle w, v\rangle_{L^2} = \int_\Omega wv\ dx$. Now we can formally rewrite $\|e(\cdot,t)\|_{L^2}^2$ as

$$\|e(\cdot,t)\|_{L^2}^2 = -\frac{d}{dt}\sum_{j=1}^{N}\langle e,\Phi\rangle_{I_j} + \sum_{j=1}^{N}\langle e_t, \Phi - \mathbb{P}(\Phi)\rangle_{I_j} + \mathbb{B}(e,\Phi-\mathbb{P}(\Phi)) - \beta_1\Delta x\sum_{j=1}^{N}[\Phi_{xx}][e]_{j+1/2}. \tag{2.28}$$

Next, estimate the terms on the right hand side of (2.28). We first bound (we assume $k \geq 1$ with the approximation polynomial space $\mathbb{V}_{\Delta x}^k$)

$$\langle e_t, \Phi - \mathbb{P}(\Phi)\rangle_\Omega \leq \|e_t\|_{L^2}\|\Phi - \mathbb{P}(\Phi)\|_{L^2} \leq C\Delta x^2\|e_t\|_{L^2}\|\Phi\|_{H^2}.$$

Then we have,

$$\begin{aligned}
\mathbb{B}(e,\Phi-\mathbb{P}(\Phi)) - \beta_1\Delta x\sum_{j=1}^{N}[\Phi_{xx}][e]_{j+1/2} &= \sum_{j=1}^{N}\int_{I_j} e_x(\Phi-\mathbb{P}(\Phi))_x\ dx + \sum_{j=1}^{N}\widehat{e}_x[\Phi-\mathbb{P}(\Phi)]_{j+1/2} \\
&\quad + \sum_{j=1}^{N}\widehat{(\Phi-\mathbb{P}(\Phi))}_x[e]_{j+1/2} - \beta_1\Delta x\sum_{j=1}^{N}[\Phi_{xx}][e]_{j+1/2} \\
&= I_1 + I_2 + I_3
\end{aligned}$$

with

$$I_1 = \sum_{j=1}^{N}\int_{I_j} e_x(\Phi-\mathbb{P}(\Phi))_x\ dx \leq \|e_x\|_{L^2}\|(\Phi-\mathbb{P}(\Phi))_x\|_{L^2} \leq C\Delta x\|e_x\|_{L^2}\|\Phi\|_{H^2},$$

and

$$
\begin{aligned}
I_2 &= \sum_{j=1}^{N} \widehat{e}_x[\Phi - \mathbb{P}(\Phi)]_{j+1/2} \\
&= \frac{\beta_0}{\Delta x}\sum_{j=1}^{N}[e][\Phi - \mathbb{P}(\Phi)]_{j+1/2} + \sum_{j=1}^{N}\overline{e}_x[\Phi - \mathbb{P}(\Phi)]_{j+1/2} + \beta_1\sum_{j=1}^{N}\Delta x[e_{xx}][\Phi - \mathbb{P}(\Phi)]_{j+1/2} \\
&\le \left(C\beta_0\Delta x^{1/2}\left(\sum_{j=1}^{N}[e]_{j+1/2}^2\right)^{1/2} + C\Delta x^{3/2}\left(\sum_{j=1}^{N}(\overline{e}_x)_{j+1/2}^2\right)^{1/2} \right. \\
&\qquad \left. + C\beta_1\Delta x^{5/2}\left(\sum_{j=1}^{N}[e_{xx}]_{j+1/2}^2\right)^{1/2}\right)\|\Phi\|_{H^2}.
\end{aligned}
$$

For the above I_1 and I_2 estimates, we need Cauchy-Schwarz inequality and the projection error estimate of $[\Phi - \mathbb{P}(\Phi)]_{j+1/2}$ (Lemma 2.3.1). Also, the following inequality was used to bound the $[\Phi - \mathbb{P}(\Phi)]_{j+1/2}^2$ terms.

$$
[w]^2 = (w^+ - w^-)^2 = (w^+)^2 - 2w^+w^- + (w^-)^2 \le (w^+)^2 + 2|w^+||w^-| + (w^-)^2 \le 2(w^+)^2 + 2(w^-)^2
$$

Similarly we can estimate the I_3 term as follows,

$$
\begin{aligned}
I_3 &= \sum_{j=1}^{N}\widehat{(\Phi - \mathbb{P}(\Phi))}_x[e]_{j+1/2} - \beta_1\Delta x\sum_{j=1}^{N}[\Phi_{xx}][e]_{j+1/2} \\
&= \frac{\beta_0}{\Delta x}\sum_{j=1}^{N}[\Phi - \mathbb{P}(\Phi)][e]_{j+1/2} + \sum_{j=1}^{N}\overline{(\Phi - \mathbb{P}(\Phi))_x}[e]_{j+1/2} - \beta_1\Delta x\sum_{j=1}^{N}[\mathbb{P}(\Phi)_{xx}][e]_{j+1/2} \\
&\le \left(C\beta_0\Delta x^{1/2} + C\Delta x^{1/2} + C\beta_1\Delta x^{1/2}\right)\left(\sum_{j=1}^{N}[e]_{j+1/2}^2\right)^{1/2}\|\Phi\|_{H^2}.
\end{aligned}
$$

Notice we need to use the inverse inequality (Lemma 2.3.2) for the last term in the above inequality. That is,

$$
[\mathbb{P}(\Phi)_{xx}]_{j+1/2} \le C\Delta x^{-1/2}|\mathbb{P}(\Phi)|_{H^2(I_j\cup I_{j+1})} \le C\Delta x^{-1/2}|\Phi|_{H^2(I_j\cup I_{j+1})}.
$$

Finally, apply these bounds to the right hand side of (2.28), and integrate in time from 0 to T we obtain,

$$
\|e(\cdot,t)\|_{L^2(0,T;L^2)}^2 \le \sum_{j=1}^{N}\int_{I_j} e(\cdot,0)\Phi(\cdot,0)\,dx + C\|\Phi\|_{L^2(0,T;H^2)}\Pi,
$$

where

$$\Pi = \Delta x^2 \|e_t\|_{L^2(0,T;L^2)} + \Delta x \|e_x\|_{L^2(0,T;L^2)} + [2\beta_0 + 1 + \beta_1]\, \Delta x^{1/2} \left(\int_0^T \sum_{j=1}^N [e]_{j+1/2}^2 \, dt \right)^{1/2} \tag{2.29}$$

$$+ \Delta x^{3/2} \left(\int_0^T \sum_{j=1}^N (\overline{e}_x)_{j+1/2}^2 \, dt \right)^{1/2} + \beta_1 \Delta x^{5/2} \left(\int_0^T \sum_{j=1}^N [e_{xx}]_{j+1/2}^2 \, dt \right)^{1/2}.$$

For the initial term we have,

$$\begin{aligned} \sum_{j=1}^N \int_{I_j} e(\cdot,0)\Phi(\cdot,0) \, dx &= \sum_{j=1}^N \int_{I_j} e(\cdot,0)(\Phi(\cdot,0) - \mathbb{P}(\Phi)(\cdot,0)) \, dx \\ &\le \|e(\cdot,0)\|_{L^2} \|\Phi(\cdot,0) - \mathbb{P}(\Phi)(\cdot,0)\|_{L^2} \\ &\le C\Delta x \|e(\cdot,0)\|_{L^2} \|\Phi(\cdot,0)\|_{H^1} \le C\Delta x \|e\|_{L^\infty(0,T;L^2)} \|\Phi\|_{L^\infty(0,T;H^1)}. \end{aligned}$$

This implies that

$$\|e\|_{L^2(0,T;L^2)}^2 \le C\Delta x \|e\|_{L^\infty(0,T;L^2)} \|\Phi\|_{L^\infty(0,T;H^1)} + C\|\Phi\|_{L^2(0,T;H^2)} \Pi.$$

With the dual regularity result (2.27) we finally obtain,

$$\|e\|_{L^2(0,T;L^2)} \le C\Delta x \|e\|_{L^\infty(0,T;L^2)} + C\Pi,$$

where Π is as given in (2.29). □

2.4.3 Time derivative error estimate and interface error estimate

To finish the $L^2(L^2)$ error estimate, we need to bound the time derivative error term $\|e_t\|_{L^2(0,T;L^2)} = \|u_t - U_t\|_{L^2(0,T;L^2)}$ as well as the two high order derivative interface terms. Theorem 2.4.3 provides this for the time derivative, and Theorem 2.4.4 considers the two interface terms.

Theorem 2.4.3. *(Time derivative $L^2(L^2)$ error estimate) Let $e := u - U$ be the error of the symmetric DDG method (2.5)-(2.6), then we have,*

$$\begin{aligned} &\|e_t\|_{L^2(0,T;L^2)} + \|e_x\|_{L^\infty(0,T;L^2)} \\ &\qquad \le C\Delta x^{k-1} \|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^k \|U\|_{L^\infty(0,T;H^{k+1})} + C\Delta x^k \|U_t\|_{L^2(0,T;H^k)}. \end{aligned}$$

Proof. As in the energy norm error estimate, we rewrite the error as $e = u - U = u - \mathbb{P}(U) + \mathbb{P}(U) - U = \mathbb{P}(e) - \xi$. For sake of presentation we introduce notation $\xi := U - \mathbb{P}(U)$. Then, again use $\langle v, w\rangle$ to denote the L^2 inner product. From the DDG scheme (2.5), we get that for all $v \in \mathbb{V}^k_{\Delta x}$,

$$\langle e_t, v\rangle + \mathbb{B}(e, v) = 0,$$

which implies that

$$\langle \mathbb{P}(e)_t, v\rangle + \mathbb{B}(\mathbb{P}(e), v) = \langle \xi_t, v\rangle + \mathbb{B}(\xi, v).$$

The bilinear form $\mathbb{B}(\cdot,\cdot)$ is as defined in (2.7). Choose $v = \mathbb{P}(e)_t \in \mathbb{V}^k_{\Delta x}$. We have,

$$\langle \mathbb{P}(e)_t, \mathbb{P}(e)_t\rangle + \mathbb{B}(\mathbb{P}(e), \mathbb{P}(e)_t) = \langle \xi_t, \mathbb{P}(e)_t\rangle + \mathbb{B}(\xi, \mathbb{P}(e)_t). \tag{2.30}$$

After integration in time, the goal will be to bound the left hand side of (2.30) below and the right hand side of (2.30) above to obtain the estimate of $\|\mathbb{P}(e)_t\|_{L^2(0,T;L^2)}$. Beginning with the left hand side, we have the terms $\langle \mathbb{P}(e)_t, \mathbb{P}(e)_t\rangle = \|\mathbb{P}(e)_t\|^2_{L^2}$ and

$$\begin{aligned}
\mathbb{B}(\mathbb{P}(e), \mathbb{P}(e)_t) &= \sum_{j=1}^{N} \int_{I_j} \mathbb{P}(e)_x (\mathbb{P}(e)_t)_x \, dx + \sum_{j=1}^{N} \widehat{\mathbb{P}(e)}_x [\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N} \widehat{(\mathbb{P}(e)_t)}_x [\mathbb{P}(e)]_{j+1/2} \\
&= \sum_{j=1}^{N} \int_{I_j} \mathbb{P}(e)_x (\mathbb{P}(e)_t)_x \, dx + \frac{2\beta_0}{\Delta x} \sum_{j=1}^{N} [\mathbb{P}(e)][\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N} \overline{\mathbb{P}(e)}_x [\mathbb{P}(e)_t]_{j+1/2} \\
&\quad + \beta_1 \Delta x \sum_{j=1}^{N} [\mathbb{P}(e)_{xx}][\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N} \overline{(\mathbb{P}(e)_t)}_x [\mathbb{P}(e)]_{j+1/2} \\
&\quad + \beta_1 \Delta x \sum_{j=1}^{N} [(\mathbb{P}(e)_t)_{xx}][\mathbb{P}(e)]_{j+1/2} \\
&= \frac{\partial}{\partial t}(T_1 + T_2 + T_3 + T_4),
\end{aligned}$$

where

$$T_1 = \frac{1}{2} \sum_{j=1}^{N} \int_{I_j} (\mathbb{P}(e)_x)^2 \, dx, \quad T_2 = \frac{\beta_0}{\Delta x} \sum_{j=1}^{N} [\mathbb{P}(e)]^2_{j+1/2},$$

$$T_3 = \sum_{j=1}^{N} \overline{\mathbb{P}(e)}_x [\mathbb{P}(e)]_{j+1/2}, \quad T_4 = \beta_1 \Delta x \sum_{j=1}^{N} [\mathbb{P}(e)_{xx}][\mathbb{P}(e)]_{j+1/2}.$$

Note, here the symmetry of the bilinear form $\mathbb{B}(\cdot,\cdot)$ is essential to obtain this complete time derivative. We then integrate in time (2.30) and have the left hand side as,

$$\int_0^t \langle \mathbb{P}(e)_t, \mathbb{P}(e)_t\rangle \, d\tau + \int_0^t \mathbb{B}(\mathbb{P}(e), \mathbb{P}(e)_t) \, d\tau = \int_0^t \|\mathbb{P}(e)_t\|_{L^2} \, d\tau + \sum_{i=1}^{4} T_i(t) - \sum_{i=1}^{4} T_i(0). \tag{2.31}$$

We leave $T_1(t)$ and $T_2(t)$ terms for now because they are positive. For $T_3(t)$ and $T_4(t)$ terms we have,

$$\Big|T_3(t)\Big| = \Big|\sum_{j=1}^{N} \overline{\mathbb{P}(e)}_x [\mathbb{P}(e)]_{j+1/2}\Big| \leq \epsilon_1 \Delta x \sum_{j=1}^{N} (\overline{\mathbb{P}(e)}_x)^2_{j+1/2} + \frac{1}{4\epsilon_1}\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}$$
$$\leq \epsilon_1 C\|\mathbb{P}(e)_x\|^2_{L^2} + \frac{1}{4\epsilon_1}\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}$$

and

$$\Big|T_4(t)\Big| = \Big|\beta_1 \Delta x \sum_{j=1}^{N} [\mathbb{P}(e)_{xx}][\mathbb{P}(e)]_{j+1/2}\Big| \leq \epsilon_2 \Delta x \sum_{j=1}^{N} \Delta x^2 [\mathbb{P}(e)_{xx}]^2_{j+1/2} + \frac{\beta_1^2}{4\epsilon_2}\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}$$
$$\leq \epsilon_2 C\|\mathbb{P}(e)_x\|^2_{L^2} + \frac{\beta_1^2}{4\epsilon_2}\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}.$$

Again, the inverse inequalities are used for the above T_3 and T_4 estimates. The constant C solely depends on the polynomial degree of the approximation space $\mathbb{V}^k_{\Delta x}$. Here we choose small enough ϵ_1 and ϵ_2 to guarantee $\frac{1}{2} - C(\epsilon_1 + \epsilon_2) > 0$. Since, by definition, $\mathbb{P}(e)(0) = u(0) - \mathbb{P}(U(0)) = 0$, thus we have $T_i(0) = 0$ for $i = 1, \cdots, 4$. Now we are ready to obtain a lower bound of (2.31) as follows,

$$\int_0^t \langle \mathbb{P}(e)_t, \mathbb{P}(e)_t\rangle \, d\tau + \int_0^t \mathbb{B}(\mathbb{P}(e), \mathbb{P}(e)_t) \, d\tau \geq \int_0^t \|\mathbb{P}(e)_t\|^2_{L^2} \, d\tau + \left[\frac{1}{2} - \epsilon_1 C - \epsilon_2 C\right]\|\mathbb{P}(e)_x\|^2_{L^2} \tag{2.32}$$
$$+ \left[\beta_0 - \frac{1}{4\epsilon_1} - \frac{\beta_1^2}{4\epsilon_2}\right]\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}$$
$$\geq \int_0^t \|\mathbb{P}(e)_t\|^2_{L^2} \, d\tau + C\|\mathbb{P}(e)_x\|^2_{L^2} + C\sum_{j=1}^{N}\frac{[\mathbb{P}(e)]^2_{j+1/2}}{\Delta x}.$$

Next, consider the right hand side of (2.30). First we have,

$$\langle \xi_t, \mathbb{P}(e)_t\rangle \leq \epsilon_3\|\mathbb{P}(e)_t\|^2_{L^2} + \frac{1}{4\epsilon_3}\|\xi_t\|^2_{L^2} \leq \epsilon_3\|\mathbb{P}(e)_t\|^2_{L^2} + C\frac{\Delta x^{2k}}{4\epsilon_3}\|U_t\|^2_{H^k}.$$

We require $U_t \in H^k(\Omega)$ in order for this projection estimate to hold. Also, we have the

following,

$$\begin{aligned}\mathbb{B}(\xi,\mathbb{P}(e)_t) &= \sum_{j=1}^{N}\int_{I_j}\xi_x(\mathbb{P}(e)_t)_x\ dx + \sum_{j=1}^{N}\widehat{\xi}_x[\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N}\widehat{(\mathbb{P}(e)_t)}_x[\xi]_{j+1/2}\\ &= \sum_{j=1}^{N}\int_{I_j}\xi_x(\mathbb{P}(e)_t)_x\ dx + \frac{2\beta_0}{\Delta x}\sum_{j=1}^{N}[\xi][\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N}\overline{\xi}_x[\mathbb{P}(e)_t]_{j+1/2}\\ &+ \Delta x\beta_1\sum_{j=1}^{N}[\xi_{xx}][\mathbb{P}(e)_t]_{j+1/2} + \sum_{j=1}^{N}\overline{(\mathbb{P}(e)_t)}_x[\xi]_{j+1/2} + \Delta x\beta_1\sum_{j=1}^{N}[(\mathbb{P}(e)_t)_{xx}][\xi]_{j+1/2}\\ &= \sum_{i=1}^{6}S_i.\end{aligned}$$

With the projection error estimate and the inverse inequalities, the estimate of these S_i's are obtained as below. Here we drop the subscripts $[\cdot]_{j+1/2}$ to save the space.

$$S_1 = \sum_{j=1}^{N}\int_{I_j}\xi_x(\mathbb{P}(e)_t)_x\ dx \le \Delta x^2\epsilon_4\|(\mathbb{P}(e)_t)_x\|_{L^2}^2 + \frac{1}{4\epsilon_4\Delta x^2}\|\xi_x\|_{L^2}^2 \le \epsilon_4 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C}{4\epsilon_4}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

$$S_2 = \frac{2\beta_0}{\Delta x}\sum_{j=1}^{N}[\xi][\mathbb{P}(e)_t] \le \Delta x\epsilon_5\sum_{j=1}^{N}[\mathbb{P}(e)_t]^2 + \frac{\beta_0^2}{4\epsilon_5\Delta x^3}\sum_{j=1}^{N}[\xi]^2 \le \epsilon_5 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C\beta_0^2}{4\epsilon_5}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

$$S_3 = \sum_{j=1}^{N}\overline{\xi}_x[\mathbb{P}(e)_t] \le \Delta x\epsilon_6\sum_{j=1}^{N}[\mathbb{P}(e)_t]^2 + \frac{1}{4\epsilon_6\Delta x}\sum_{j=1}^{N}(\overline{\xi}_x)^2 \le \epsilon_6 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C}{4\epsilon_6}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

$$S_4 = \Delta x\beta_1\sum_{j=1}^{N}[\xi_{xx}][\mathbb{P}(e)_t] \le \Delta x\epsilon_7\sum_{j=1}^{N}[\mathbb{P}(e)_t]^2 + \frac{\Delta x}{4\epsilon_7}\sum_{j=1}^{N}\beta_1^2[\xi_{xx}]^2 \le \epsilon_7 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C\beta_1^2}{4\epsilon_7}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

$$S_5 = \sum_{j=1}^{N}\overline{(\mathbb{P}(e)_t)}_x[\xi] \le \Delta x^3\epsilon_8\sum_{j=1}^{N}\overline{(\mathbb{P}(e)_t)}_x^2 + \frac{1}{4\epsilon_8\Delta x^3}\sum_{j=1}^{N}[\xi]^2 \le \epsilon_8 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C}{4\epsilon_8}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

$$S_6 = \Delta x\beta_1\sum_{j=1}^{N}[(\mathbb{P}(e)_t)_{xx}][\xi] \le \Delta x^5\epsilon_9\sum_{j=1}^{N}[(\mathbb{P}(e)_t)_{xx}]^2 + \frac{1}{4\epsilon_9\Delta x^3}\sum_{j=1}^{N}\beta_1^2[\xi]^2 \le \epsilon_9 C\|\mathbb{P}(e)_t\|_{L^2}^2 + \frac{C\beta_1^2}{4\epsilon_9}\Delta x^{2k-2}|U|_{H^{k+1}}^2$$

We choose small enough ϵ_i, $i = 3,\cdots,9$ to balance the left hand side term of $\|\mathbb{P}(e)_t\|_{L^2}^2$. Integrate in time, we obtain the upper bound of the right hand side of (2.30) as,

$$\int_0^t\langle\xi_t,\mathbb{P}(e)_t\rangle\ d\tau + \int_0^t\mathbb{B}(\xi,\mathbb{P}(e)_t)\ d\tau \le \epsilon\int_0^t\|\mathbb{P}(e)_t\|_{L^2}^2\ d\tau + C\Delta x^{2k}\|U_t\|_{L^2(0,T;H^k)}^2 + C\Delta x^{2k-2}\|U\|_{L^2(0,T;H^{k+1})}^2, \tag{2.33}$$

where $\epsilon = \epsilon_3 + C\sum_{i=4}^{9}\epsilon_i$. Then, combining (2.33) with (2.32) we obtain

$$\int_0^t\|\mathbb{P}(e)_t\|_{L^2}^2 + \sum_{j=1}^{N}\int_{I_j}(\mathbb{P}(e)_x)^2\ dx + \sum_{j=1}^{N}\frac{[\mathbb{P}(e)]_{j+1/2}^2}{\Delta x} \le C\Delta x^{2k-2}\|U\|_{L^2(0,T;H^{k+1})}^2 + C\Delta x^{2k}\|U_t\|_{L^2(0,T;H^k)}^2.$$

Recall that $e = \mathbb{P}(e) - \xi$ and make use of the triangle inequality and projection error estimates of $\xi = U - \mathbb{P}(U)$, finally we have

$$\|e_t\|_{L^2(0,T;L^2)} + \|e_x\|_{L^\infty(0,T;L^2)} \le C\Delta x^{k-1}\|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^k\|U\|_{L^\infty(0,T;H^{k+1})} + C\Delta x^k\|U_t\|_{L^2(0,T;H^k)}.$$

□

Theorem 2.4.4. *(Interface error estimates) Let $e := u - U$ be the error of the symmetric DDG method (2.5)-(2.6), we have*

$$\Delta x^{3/2}\left(\int_0^T \sum_{j=1}^{N} (\overline{e_x})^2_{j+1/2}\, dt\right)^{1/2} \leq C\Delta x^{k+1}\|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^{k+1}\|U\|_{L^\infty(0,T;H^{k+1})}$$

$$\Delta x^{5/2}\left(\int_0^T \sum_{j=1}^{N} [e_{xx}]^2_{j+1/2}\, dt\right)^{1/2} \leq C\Delta x^{k+1}\|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^{k+1}\|U\|_{L^\infty(0,T;H^{k+1})}.$$

Proof. These results can be obtained using similar techniques as in the proof of the above theorem. Again we use $\xi = U - \mathbb{P}(U)$ to represent the projection error. We have

$$\begin{aligned}
\Delta x^{3/2}\left(\int_0^T \sum_{j=1}^{N} (\overline{(U-u)_x})^2_{j+1/2}\, dt\right)^{1/2} &= \Delta x^{3/2}\left(\int_0^T \sum_{j=1}^{N} (\overline{(U-\mathbb{P}(U)+\mathbb{P}(U)-u)_x})^2_{j+1/2}\, dt\right)^{1/2} \\
&\leq \sqrt{2}\Delta x^{3/2}\left(\int_0^T \sum_{j=1}^{N} (\overline{\xi_x})^2_{j+1/2} + \sum_{j=1}^{N} (\overline{\mathbb{P}(e)_x})^2_{j+1/2}\, dt\right)^{1/2} \\
&\leq C\Delta x^{3/2}\left(\Delta x^{k-1/2}\|U\|_{L^2(0,T;H^{k+1})} + \Delta x^{-1/2}\|\mathbb{P}(e)_x\|_{L^2(0,T;L^2)}\right) \\
&\leq C\Delta x^{k+1}\|U\|_{L^2(0,T;H^{k+1})} + C\Delta x^{k+1}\|U\|_{L^\infty(0,T;H^{k+1})}.
\end{aligned}$$

In the last step, results from the energy norm estimate were made use of. Proof of the second interface estimate is similar. □

2.5 Numerical Examples

In this section, numerical examples are provided to illustrate the performance of the symmetric DDG method. Here, one dimensional linear problems are considered. Nonlinear and two dimensional problems are considered in upcoming chapters. For each computation, Table 2.1 is used as a guideline for the choice of numerical flux coefficients (β_0, β_1). Both uniform and nonuniform mesh computations are considered. Also, exploration of the admissibility result (2.2.1) is provided.

Example 2.5.1: 1-D linear diffusion equation

$$\begin{cases} U_t - U_{xx} = 0, & x \in [0, 2\pi] \\ U(x,0) = \sin(x), \end{cases} \tag{2.34}$$

with periodic boundary conditions.

We use this example to verify the optimal convergence of the symmetric DDG method using three tests. The first is carried out on a uniform mesh. Note, for $k = 0, 1$ in this linear case, the symmetric DDG scheme is the same as the SIPG method, thus we begin with quadratic polynomial approximations. Degree k polynomial approximations with $k = 2, \cdots, 6$ are tested, and optimal $(k+1)$th orders of convergence are achieved. See Table 2.3 for L^2 and L^∞ errors and orders of convergence. Note that in this and the remaining examples, the L^∞ error is obtained by evaluating 200 sample points per cell. Also, recall that N in Table 2.3 denotes the total number of computational cells. The second test addresses the numerical

Table 2.3 1-D linear diffusion equation (2.34), P^k polynomial approximations with uniform mesh. $T = 1$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 10$	$N = 20$		$N = 40$		$N = 80$	
β_0	3/2	$k = 2$	L^2	1.92E-03	2.36E-04	3.0	2.93E-05	3.0	3.66E-06	3.0
β_1	1/4		L^∞	3.64E-03	4.70E-04	3.0	5.92E-05	3.0	7.42E-06	3.0
β_0	11/4	$k = 3$	L^2	2.60E-05	1.58E-06	4.0	9.81E-08	4.0	6.12E-09	4.0
β_1	3/32		L^∞	5.87E-05	3.67E-06	4.0	2.32E-07	4.0	1.46E-08	4.0
β_0	9/2	$k = 4$	L^2	6.92E-07	2.07E-08	5.1	6.40E-10	5.0	1.99E-11	5.0
β_1	1/20		L^∞	1.68E-06	5.33E-08	5.0	1.67E-09	5.0	5.23E-11	5.0
				$N = 8$	$N = 12$		$N = 16$		$N = 20$	
β_0	27/4	$k = 5$	L^2	1.86E-07	1.67E-08	5.9	2.99E-09	6.0	7.87E-10	6.0
β_1	1/32		L^∞	3.25E-07	2.97E-08	5.9	5.37E-09	6.0	1.42E-09	6.0
β_0	19/2	$k = 6$	L^2	3.06E-09	1.32E-10	7.7	1.48E-11	7.6	2.81E-12	7.4
β_1	3/140		L^∞	4.84E-09	2.40E-10	7.4	2.97E-11	7.3	6.02E-12	7.2

flux admissibility provided by Theorem 2.2.1, which is explored computationally. There is a wide range of flux coefficients (β_0, β_1) which are admissible as defined by (2.9). Figure 2.1 illustrates the admissibility region for (β_0, β_1) pairs with $k = 1, 2, \ldots, 10$. In this test, different β_1 values are selected and the correspondingly smallest admissible β_0 is computed from (2.11) with $\alpha = 1, \gamma = \frac{1}{2}$. We list convergence results for quadratic approximations ($k = 2$) in Table 2.4, and numerically the optimal 3rd order of accuracy is observed for wide range of (β_0, β_1) pairs. Taking this further, a wide range of (β_0, β_1) pairs were considered. Specifically, $0 \le \beta_1 \le 0.7$ and $0 \le \beta_0 \le 8$ were each divided uniformly into 21 points, resulting in a total of 441 pairs. This same computation was carried out on a uniform mesh for both $N = 20$ and

$N = 40$. Then, resulting errors and orders of convergence where computed. Figures 2.2 and 2.3 give a contour plot of the orders of convergence for each of these (β_0, β_1) pairs in the L^2 and L^∞ norms. In addition, the parabola seen in Figure 2.1 as provided from Theorem 2.2.1 is plotted for comparison with theoretical results. If the run did not converge as was seen for the clearly non-admissible pairs, the order was simply set to zero. For the L^2 case, the plot is slightly deceiving. The lighter red region above the parabola graph include rates of convergence very near 3, while orders in the darker red region are slightly larger than 3. In order to explore this further, these contour plots were refined. That is, for all computed orders less than 3, the values in the plot were set to be three. This illustrates the regions of the contour plot which have orders of convergence greater than 3. Figures 2.4 and 2.5 illustrate this for these two norms. It is interesting to see in both cases, larger choices of β_1 result in higher orders of convergence. This is more apparent in the L^2 norm.

Table 2.4 1-D linear diffusion equation (2.34), uniform mesh, P^2 quadratic polynomial approximations. Admissibility test with different choices of (β_0, β_1) pair.

				Error	Error	Order	Error	Order	Error	Order
				$N = 10$	$N = 20$		$N = 40$		$N = 80$	
β_0	9/2	$k = 2$	L^2	1.68E-03	1.75E-04	3.3	2.07E-05	3.1	2.55E-06	3.0
β_1	1/2		L^∞	2.62E-03	3.13E-04	3.1	3.87E-05	3.0	4.83E-06	3.0
β_0	3/2	$k = 2$	L^2	1.92E-03	2.36E-04	3.0	2.93E-05	3.0	3.66E-06	3.0
β_1	1/4		L^∞	3.64E-03	4.70E-04	3.0	5.92E-05	3.0	7.42E-06	3.0
β_0	9/4	$k = 2$	L^2	5.65E-04	7.10E-05	3.0	8.90E-06	3.0	1.11E-06	3.0
β_1	1/8		L^∞	1.04E-03	1.33E-04	3.0	1.66E-05	3.0	2.08E-06	3.0
β_0	171/50	$k = 2$	L^2	2.90E-04	3.61E-05	3.0	4.50E-06	3.0	5.63E-07	3.0
β_1	1/20		L^∞	5.80E-04	7.31E-05	3.0	9.16E-06	3.0	1.15E-06	3.0
β_0	393/100	$k = 2$	L^2	2.59E-04	3.19E-05	3.0	3.97E-06	3.0	4.96E-07	3.0
β_1	1/40		L^∞	5.20E-04	6.50E-05	3.0	8.12E-06	3.0	1.01E-06	3.0

The third test is implemented on a nonuniform mesh. The nonuniform mesh is generated by repeating the pattern $\frac{\Delta x}{5}$, $\frac{3\Delta x}{10}$, and $\frac{\Delta x}{2}$, where $\Delta x = \frac{2\pi}{N}$. Table 2.5 provides the convergence results on such a nonuniform mesh. Note, for this nonuniform mesh it was observed computationally that a larger β_0 was needed for the numerical flux pair. Similar requirement on large β_0 is needed for the SIPG method.

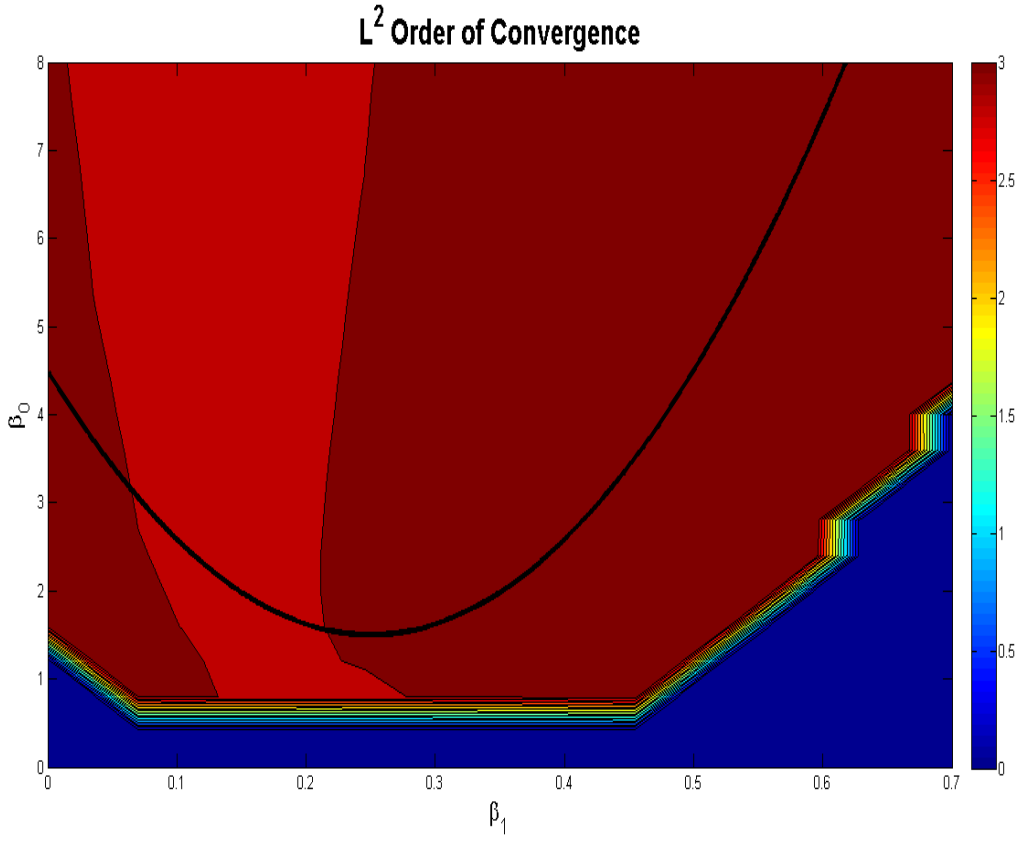

Figure 2.2 Computational admissibility region for $k = 2$ on uniform mesh.

Table 2.5 1-D linear diffusion equation (2.34), P^k approximations on nonuniform mesh. $T = 1$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 18$	$N = 36$		$N = 54$		$N = 72$	
β_0	20	$k = 2$	L^2	1.41E-04	1.79E-05	3.0	5.45E-06	2.9	2.39E-06	2.9
β_1	1/4		L^∞	4.18E-04	5.95E-05	2.8	1.80E-05	2.9	7.68E-06	3.0
β_0	25	$k = 3$	L^2	6.70E-06	4.23E-07	4.0	8.41E-08	4.0	2.68E-08	4.0
β_1	3/32		L^∞	2.22E-05	1.35E-06	4.0	2.75E-07	3.9	8.71E-08	4.0
β_0	25	$k = 4$	L^2	1.02E-07	3.02E-09	5.1	4.00E-10	5.0	9.67E-11	4.9
β_1	1/20		L^∞	3.52E-07	1.12E-08	5.0	1.50E-09	5.0	3.59E-10	5.0

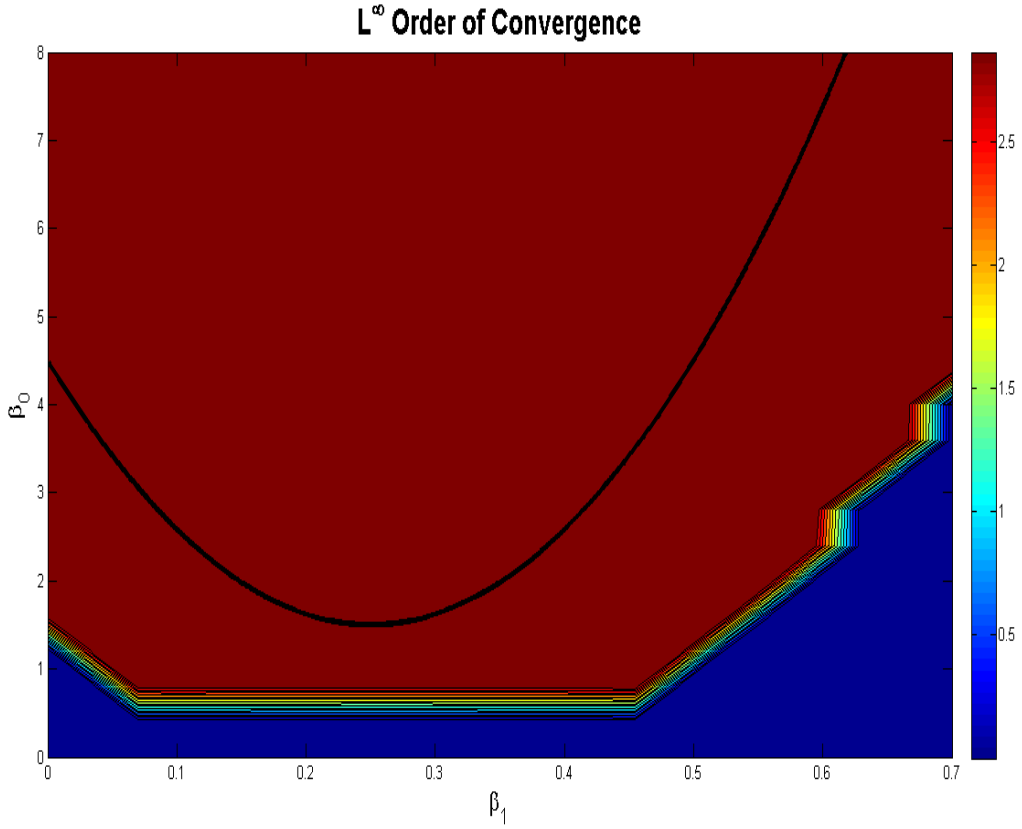

Figure 2.3 Computational admissibility region for $k = 2$ on uniform mesh.

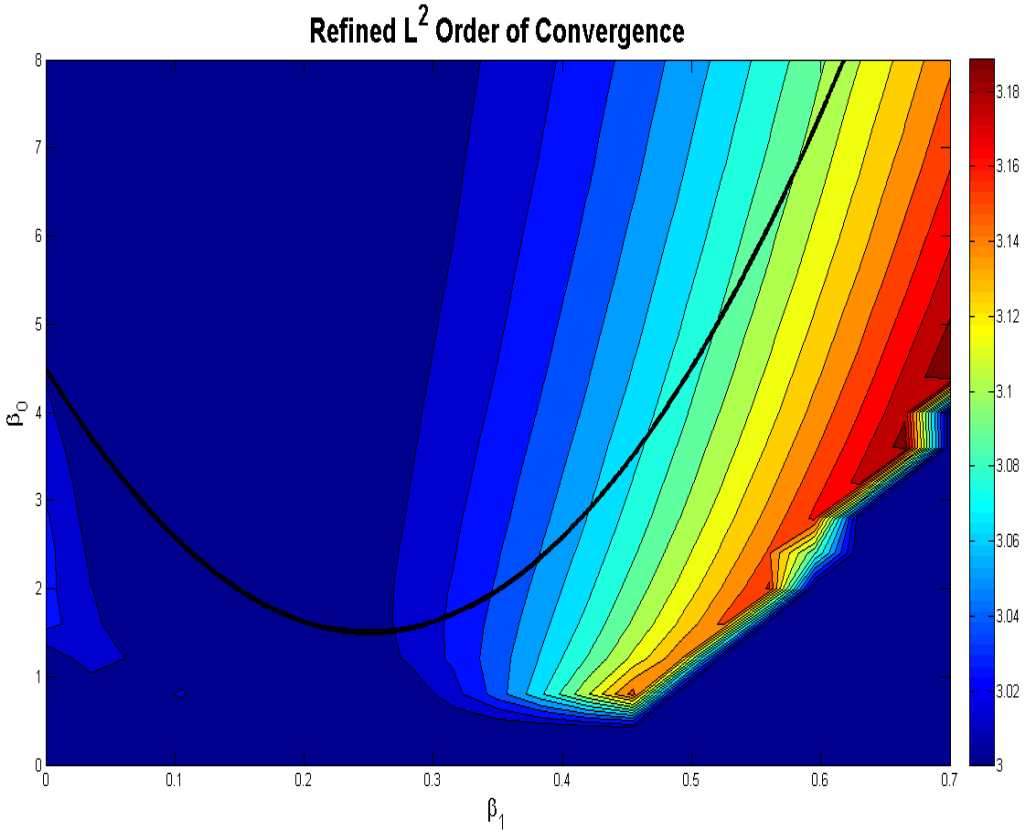

Figure 2.4 Refined computational admissibility region for $k = 2$ on uniform mesh.

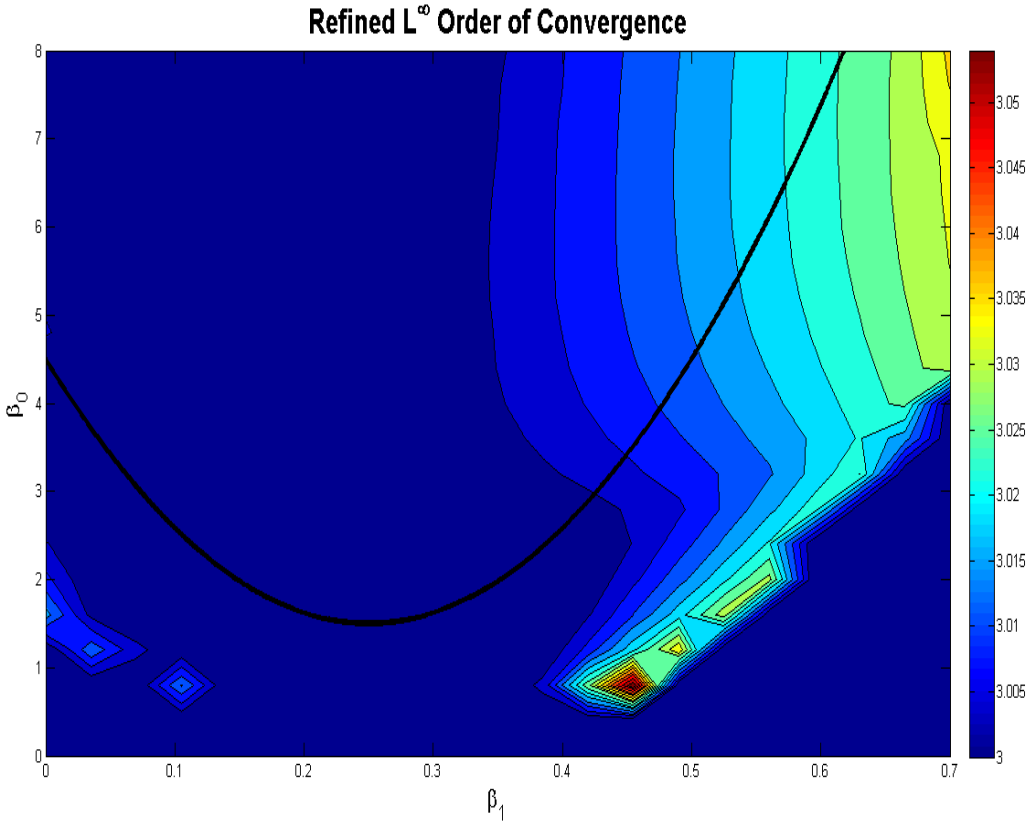

Figure 2.5 Refined computational admissibility region for $k = 2$ on uniform mesh.

CHAPTER 3. ONE DIMENSIONAL NONLINEAR CONVECTION DIFFUSION PROBLEMS

This chapter extends work done in Chapter 2 to one dimensional nonlinear diffusion and nonlinear convection diffusion problems. The scheme formulation for the one dimension linear case is directly made use of. A notion of flux admissibility is defined which ensures nonlinear stability of the scheme. For convection terms, the symmetric DDG method is coupled with existing methods designed for convection. Because of the nonlinearity, error analysis cannot be performed. Last, numerical examples are provided in a number of settings. The symmetric DDG method proves to handle problems of this type quite well.

3.1 Extension to One Dimensional Nonlinear Diffusion Problems

In this section, we extend the above symmetric DDG scheme to the one-dimensional nonlinear diffusion equation

$$U_t - (a(U)U_x)_x = 0, \quad \text{for } (x,t) \in \Omega \times (0,T) \tag{3.1}$$

with initial data $U(x,0) = U_0(x)$ and periodic boundary conditions. Here, we assume the diffusion coefficient $a(U) \geq 0$. Also, denote $b(u) = \int_0^u a(s)\, ds$. Then, $b(U)_x = a(U)U_x$.

Partition the domain $\Omega = \bigcup_{j=1}^N I_j$ and consider the solution space $\mathbb{V}_{\Delta x}^k$ as above. Then, taking inspiration from the linear case, we have the following scheme. Find the approximate solution $u \in \mathbb{V}_{\Delta x}^k$ of U in (3.1) such that for all test functions $v \in \mathbb{V}_{\Delta x}^k$ and on all I_j,

$$\int_{I_j} u_t v \, dx - \widehat{b(u)_x} v \Big|_{j-1/2}^{j+1/2} + \int_{I_j} b(u)_x v_x \, dx + ([b(u)]\widehat{v_x})_{j+1/2} + ([b(u)]\widehat{v_x})_{j-1/2} = 0,$$

with the numerical fluxes defined as

$$\begin{cases} \widehat{b(u)_x} = \beta_0 \frac{[b(u)]}{\Delta x} + \overline{b(u)_x} + \beta_1 \Delta x [b(u)_{xx}], \\ \\ \widehat{v_x} = \beta_0 \frac{[v]}{\Delta x} + \overline{v_x} + \beta_1 \Delta x [v_{xx}]. \end{cases} \tag{3.2}$$

Summing over all computational cells I_j provides the primal weak formulation,

$$\int_\Omega u_t v \; dx + \mathbb{B}(b(u), v) = 0,$$

where bilinear form $\mathbb{B}(b(u), v)$ is as given in (2.7). Note, symmetry of this bilinear form is maintained in the sense of $\mathbb{B}(b(u), v) = \mathbb{B}(v, b(u))$ for the nonlinear diffusion equation.

As in the linear case, we have a notion of admissibility of numerical flux terms and resulting stability.

Definition 3.1.1. *(Nonlinear Admissibility) We call numerical flux $\widehat{b(u)_x}$ in (3.6) admissible if there exists $\gamma \in (0,1)$, $\alpha > 0$ such that for any $u \in \mathbb{V}^k_{\Delta x}$,*

$$\gamma \sum_{j=1}^{N} \int_{I_j} a(u) u_x^2(x,t) \; dx + \sum_{j=1}^{N} \left(\widehat{b(u)_x}[u] + \widehat{u_x}[b(u)] \right)_{j+1/2} \geq \alpha \sum_{j=1}^{N} \frac{b(u^*)[u]^2_{j+1/2}}{\Delta x}. \tag{3.3}$$

Theorem 3.1.1. *(Nonlinear Stability) Consider the symmetric DDG scheme (3.1). If the numerical flux (3.6) is admissible as described in (3.3), then we have*

$$\frac{1}{2}\int_\Omega u^2(x,T)dx + (1-\gamma)\int_0^T \sum_{j=1}^{N} \int_{I_j} a(u) u_x^2(x,t) dx dt + \alpha \int_0^T \sum_{j=1}^{N} \frac{b(u^*)[u]^2_{j+1/2}}{\Delta x} dt \leq \frac{1}{2}\int_\Omega U_0^2(x) dx. \tag{3.4}$$

3.2 Extension to One Dimensional Nonlinear Convection Diffusion Problems

Next, we extend the symmetric DDG scheme further to the one-dimensional nonlinear convection diffusion equation

$$U_t + f(U)_x - (a(U)U_x)_x = 0, \quad \text{for} (x,t) \in \Omega \times (0,T) \tag{3.5}$$

with initial data $U(x,0) = U_0(x)$ and periodic boundary conditions. Again, we assume the diffusion coefficient $a(U) \geq 0$. Also, denote $b(s) = \int a(s) \; ds$. Then, $b(U)_x = a(U)U_x$.

Partition the domain $\Omega = \bigcup_{j=1}^{N} I_j$ and consider the solution space $\mathbb{V}_{\Delta x}^k$ as above. Then, as a direct extension of the nonlinear case, we have the following scheme. Find the approximate solution $u \in \mathbb{V}_{\Delta x}^k$ of U in (3.5) such that for all test functions $v \in \mathbb{V}_{\Delta x}^k$ and on all I_j,

$$\int_{I_j} u_t v \, dx + \left(\widetilde{f(u)} - \widehat{b(u)_x}\right) v \Big|_{j-1/2}^{j+1/2} - \int_{I_j} \left(f(u) - b(u)_x\right) v_x \, dx + ([b(u)]\widehat{v_x})_{j+1/2} + ([b(u)]\widehat{v_x})_{j-1/2} = 0,$$

with the numerical fluxes defined as

$$\begin{cases} \widehat{b(u)_x} = \beta_0 \frac{[b(u)]}{\Delta x} + \overline{b(u)_x} + \beta_1 \Delta x [b(u)_{xx}], \\ \widehat{v_x} = \beta_0 \frac{[v]}{\Delta x} + \overline{v_x} + \beta_1 \Delta x [v_{xx}], \\ \widetilde{f(u)} = \widetilde{f}(u^-, u^+) = \frac{1}{2}\left(f(u-) + f(u+) - \alpha(u^+ - u^-)\right) \end{cases} \tag{3.6}$$

Here, $\widetilde{f(u)}$ is the Lax-Friedrichs flux where $\alpha = \max_{u \in \{u^-, u^+\}} |f'(u)|$ (for reference, see [25]).

3.3 Numerical Examples

In this section, additional numerical examples are provided. Here, one dimensional nonlinear diffusion problems and nonlinear convection diffusion problems are considered.

Example 3.3.1: 1-D nonlinear porous medium equation

$$U_t - (2UU_x)_x = 0, \quad x \in [-12, 12]. \tag{3.7}$$

The exact solution is given by

$$U(x,t) = \begin{cases} (t+1)^{-1/3}\left(3 - \frac{x^2}{12(t+1)^{2/3}}\right), & |x| < 6(t+1)^{1/3} \\ 0, & |x| \geq 6(t+1)^{1/3}. \end{cases}$$

We see that the wave solution travels with finite speed. Accuracy tests are carried out and results are listed in Table 3.1 at final time $T = 1$. We obtain $(k+1)$th order of accuracy with P^k polynomial approximations. Note that errors and orders are computed within domain $[-6, 6]$ where the solution is smooth. In Figure 3.1 we illustrate the evolution of the symmetric DDG solution at times $T = 1, 2, 4$. The symmetric DDG solution resolves the two kinked corners well with no oscillations.

Example 3.3.2: General porous medium equation

$$U_t - (U^m)_{xx} = 0, \quad x \in [-6, 6]. \tag{3.8}$$

Table 3.1 1-D nonlinear porous medium equation (3.7), P^k polynomial approximations, $T = 1$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 40$	$N = 80$		$N = 160$		$N = 320$	
β_0	1/2	$k = 0$	L^2	3.54E-02	1.77E-02	1.0	8.84E-03	1.0	4.42E-03	1.0
β_1	0		L^∞	1.45E-01	7.36E-02	1.0	3.71E-02	1.0	1.87E-02	1.0
β_0	2	$k = 1$	L^2	1.29E-03	3.20E-04	2.0	8.02E-05	2.0	2.00E-05	2.0
β_1	1/80		L^∞	4.69E-03	1.15E-03	2.0	2.92E-04	2.0	7.29E-05	2.0
β_0	2	$k = 2$	L^2	3.25E-05	7.30E-08	8.8	1.49E-09	5.6	1.40E-10	3.4
β_1	1/80		L^∞	4.04E-04	9.48E-07	8.7	3.31E-09	8.2	3.03E-10	3.4

Here we consider zero boundary condition. Then, the Barrenblatt solution is given as

$$U(x,t) = \begin{cases} (t+1)^{-\frac{1}{m+1}} \left(3 - \frac{m-1}{2m(m+1)} \frac{x^2}{(t+1)^{2/(m+1)}}\right), & |x| \le \sqrt{\frac{6m(m+1)}{m-1}(t+1)^{\frac{1}{m+1}}}, \\ 0, & |x| \ge \sqrt{\frac{6m(m+1)}{m-1}(t+1)^{\frac{1}{m+1}}}. \end{cases} \tag{3.9}$$

For the first test, the solution was computed until end time $T = 1$ for quadratic approximations on a uniform mesh with $N = 80$. Here, Barrenblatt's solution was used as an initial condition. Figure 3.2 depicts runs for the $m = 2, 3, 5, 8$ cases. For the $m = 2, 3, 5$ cases $(\beta_0, \beta_1) = (2, \frac{1}{12})$. In the $m = 8$ case, $(\beta_0, \beta_1) = (4, \frac{1}{80})$.

Next, a two box collision test was considered. Here, equation (3.8) was considered with two different initial conditions. For the first, $m = 5$ was taken and the initial profile of U was set to be

$$U(x,0) = \begin{cases} 1, & x \in (-3.7, -0.7) \bigcup (0.7, 3.7) \\ 0, & \text{otherwise.} \end{cases} \tag{3.10}$$

The evolution for piecewise quadratic approximations with mesh partition $N = 160$ at times $t = 0, 0.2, 0.8, 1.5$ are listed in Figure 3.3. Here and below, $(\beta_0, \beta_1) = (2, \frac{1}{12})$. Next, the same box collision test was performed with boxes of differing height with $m = 6$.

$$U(x,0) = \begin{cases} 1, & x \in (-4, -1) \\ 2, & x \in (0, 3) \\ 0, & \text{otherwise.} \end{cases} \tag{3.11}$$

The evolution for piecewise quadratic approximations with mesh partition $N = 160$ at times $t = 0, 0.01, 0.04, 0.06, 0.12, 0.8$ are listed in Figure 3.4. In both of these test, the initial boxes

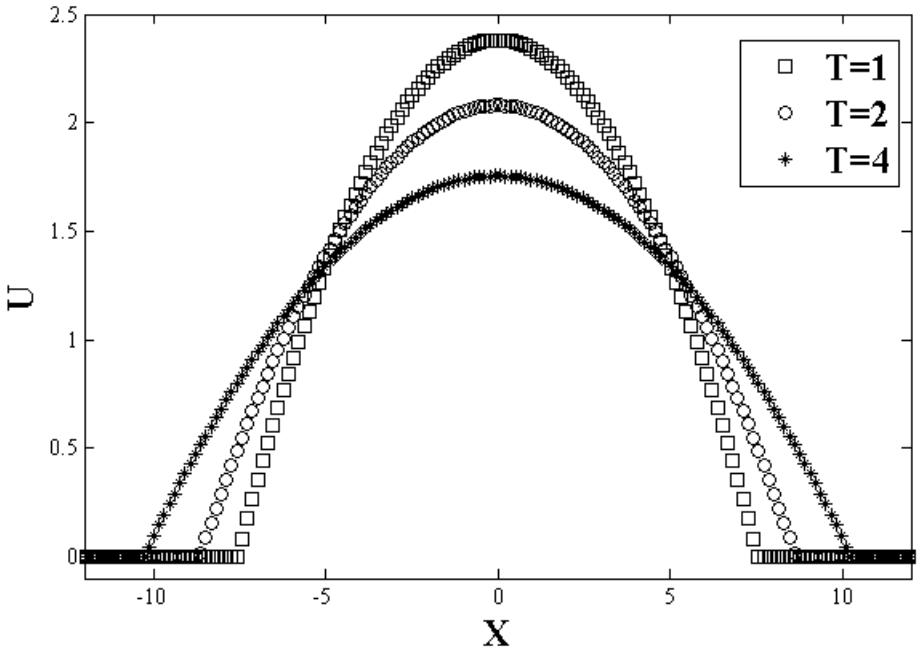

Figure 3.1 Evolution of the symmetric DDG solution to the 1-D nonlinear porous medium equation (3.7) at $T = 1, 2, 4$.

propagate with finite speed, and the solutions smooth and even out over time.

Example 3.3.3: Fisher-Kolmogorov equation

$$U_t - U_{xx} = U(1-U), \quad x \in [-20, 20]. \tag{3.12}$$

This is a nonlinear reaction diffusion equation, and the one solution is a traveling wave $U(x,t) = \left(1 + Ae^{\frac{x-ct}{\sqrt{6}}}\right)^{-2}$ where $c = \frac{5}{\sqrt{6}}$ and $A > 0$ is an arbitrary positive constant. For the first test, the solution was computed for end time $T = 2$. Table 3.2 gives errors and order of accuracy. The initial condition was set to match the exact solution. In each case, $(k+1)$ order accuracy is seen for degree k approximations. For the second test, discontinuous initial data

$$U(x,0) = \begin{cases} 1, & x \le 0, \\ 0, & x > 0, \end{cases}$$

was considered. The propagation of the solution was computed at times $t = 0, 1, 3, 5, 7$ and is displayed in Figure 3.5.

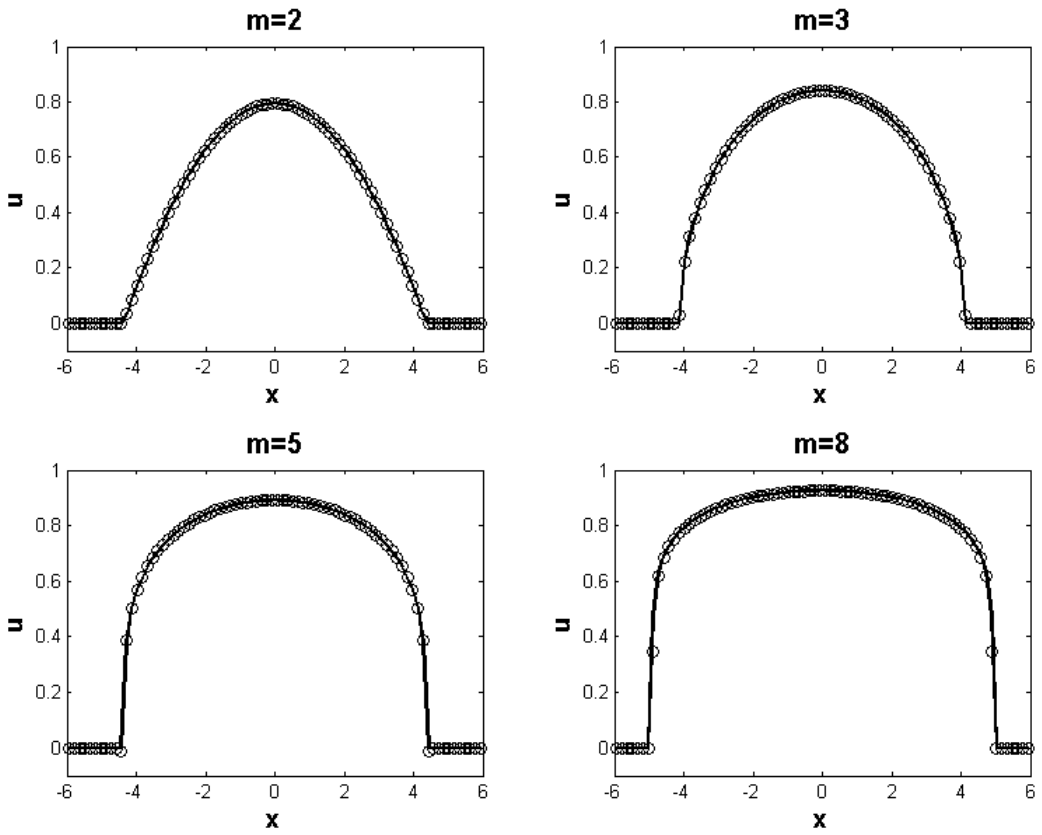

Figure 3.2 General porous medium equation (3.8) for $m = 2, 3, 5, 8$. Here, solid lines give the exact solution while circles give the quadratic symmetric DDG solution.

Example 3.3.4: Fully nonlinear convection diffusion equation

$$U_t + (U^2)_x - (U^2)_{xx} = f(x,t), \quad x \in [-3,3], \tag{3.13}$$

with the exact solution constructed to be $U(x,t) = e^{x-t}$. As a result, $f(x,t) = -e^{x-t} - 2e^{2(x-t)}$. For boundary conditions, the exact solution is used. In addition, the Lax-Friedrichs flux as mentioned above is used for the nonlinear convection term. Table 3.3 lists errors and orders of convergence for $k = 2, 3, 4$ cases at end time $t = 1.0$.

Example 3.3.5: Strongly degenerate nonlinear convection diffusion equation

$$U_t + (U^2)_x - \epsilon(\nu(U)U_x)_x = 0, \quad x \in [-2,2], \tag{3.14}$$

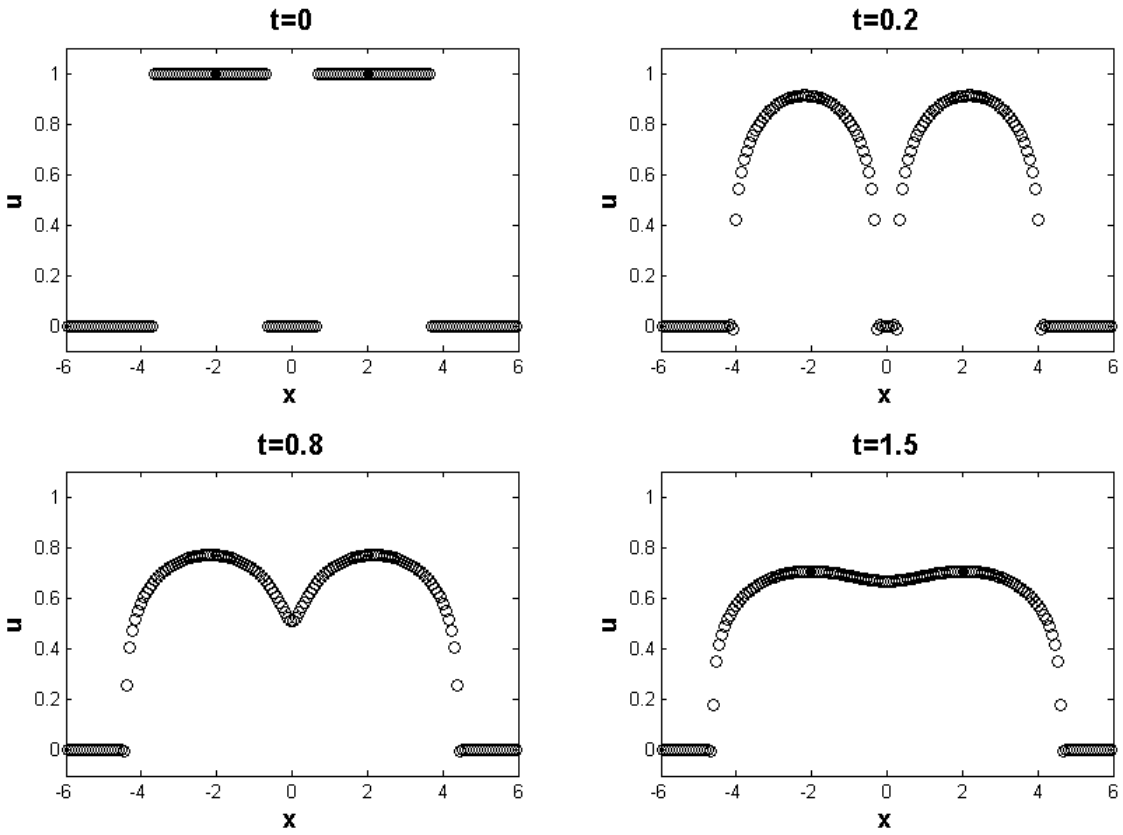

Figure 3.3 Quadratic approximation of general porous medium equation (3.8) for $m = 5$ with initial data (3.10), $N = 160$.

with parameter $\epsilon = 0.1$ and

$$\nu(U) = \begin{cases} 0, & |U| \le 0.25, \\ 1, & |U| > 0.25. \end{cases}$$

Then, this equation is hyperbolic if $|U(x,t)| \le 0.25$ and is parabolic otherwise. Here a zero boundary condition was used along with initial condition

$$U(x,0) = \begin{cases} 1, & -\frac{1}{\sqrt{2}} - 0.4 < x < -\frac{1}{\sqrt{2}} + 0.4, \\ -1, & \frac{1}{\sqrt{2}} - 0.4 < x < +\frac{1}{\sqrt{2}} + 0.4, \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

Here, a piecewise quadratic approximation was compute with the use of the minmod slope

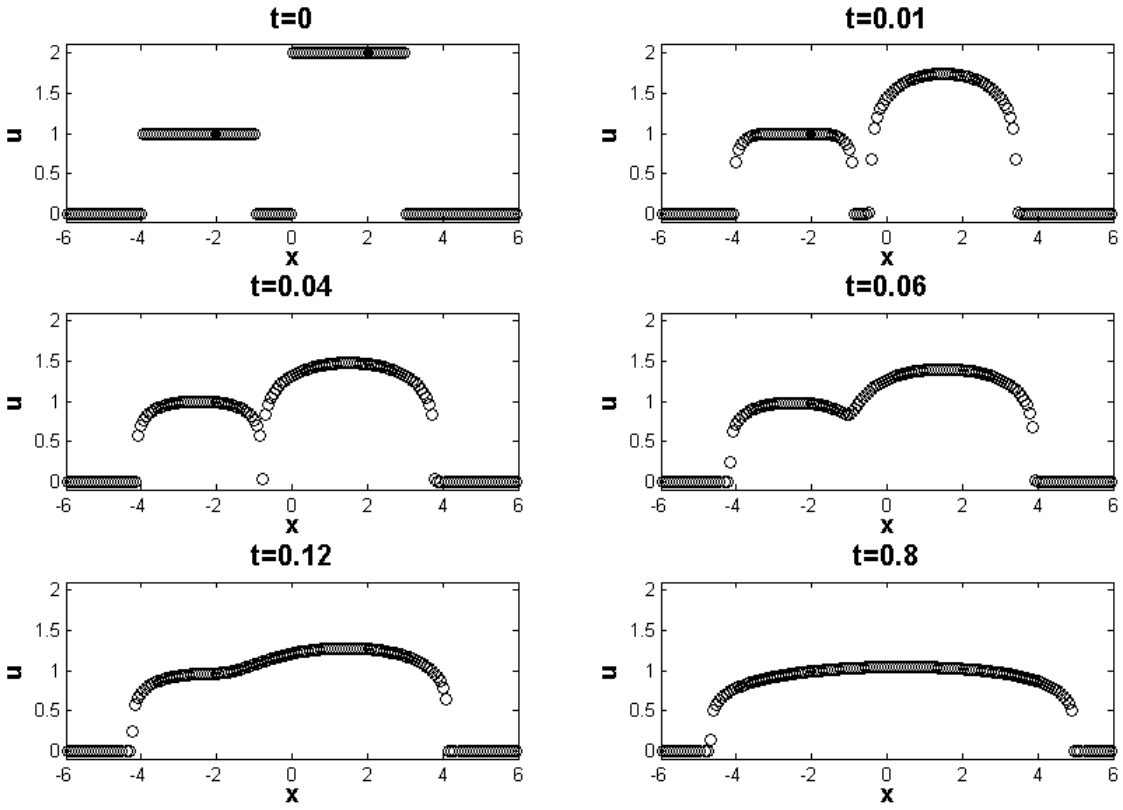

Figure 3.4 Quadratic approximation of general porous medium equation (3.8) for $m = 6$ with initial data (3.11), $N = 160$.

limiter as described in [38]. Figure 3.6 gives the exact solution at initial time $t = 0$ as well as symmetric DDG solution at time $t = 0.7$ for mesh size $N = 100$ and $N = 600$.

Table 3.2 1-D Fisher-Kolmogorov equation (3.12), P^k polynomial approximations, $T = 2$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 40$	$N = 80$		$N = 120$		$N = 160$	
β_0	2	$k = 2$	L^2	2.41E-05	2.92E-06	3.0	8.61E-07	3.0	3.63E-07	3.0
β_1	1/12		L^∞	1.14E-04	1.40E-05	3.0	4.10E-06	3.0	1.73E-06	3.0
β_0	2	$k = 3$	L^2	1.93E-06	9.33E-08	4.4	1.73E-08	4.2	5.33E-09	4.1
β_1	1/12		L^∞	1.01E-05	4.45E-07	4.5	7.90E-08	4.3	2.34E-08	4.2
β_0	4	$k = 4$	L^2	2.96E-08	8.36E-10	5.1	1.08E-10	5.1	2.54E-11	5.0
β_1	1/40		L^∞	1.53E-07	4.73E-09	5.0	6.17E-10	5.0	1.46E-10	5.0

Table 3.3 1-D nonlinear convection diffusion equation (3.13), P^k polynomial approximations, $T = 1$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 10$	$N = 20$		$N = 40$		$N = 80$	
β_0	2	$k = 2$	L^2	6.20E-04	6.54E-05	3.2	1.84E-05	3.1	7.57E-06	3.1
β_1	1/12		L^∞	1.53E-03	1.96E-04	3.0	5.85E-05	3.0	2.48E-05	3.0
β_0	2	$k = 3$	L^2	8.09E-06	5.09E-07	4.0	1.01E-07	4.0	3.19E-08	4.0
β_1	1/12		L^∞	2.72E-05	1.83E-06	3.9	3.71E-07	3.9	1.19E-07	4.0
β_0	4	$k = 4$	L^2	1.65E-07	4.21E-09	5.3	5.24E-10	5.1	1.15E-10	5.2
β_1	1/40		L^∞	3.85E-07	1.29E-08	4.9	1.73E-09	4.9	4.49E-10	4.7

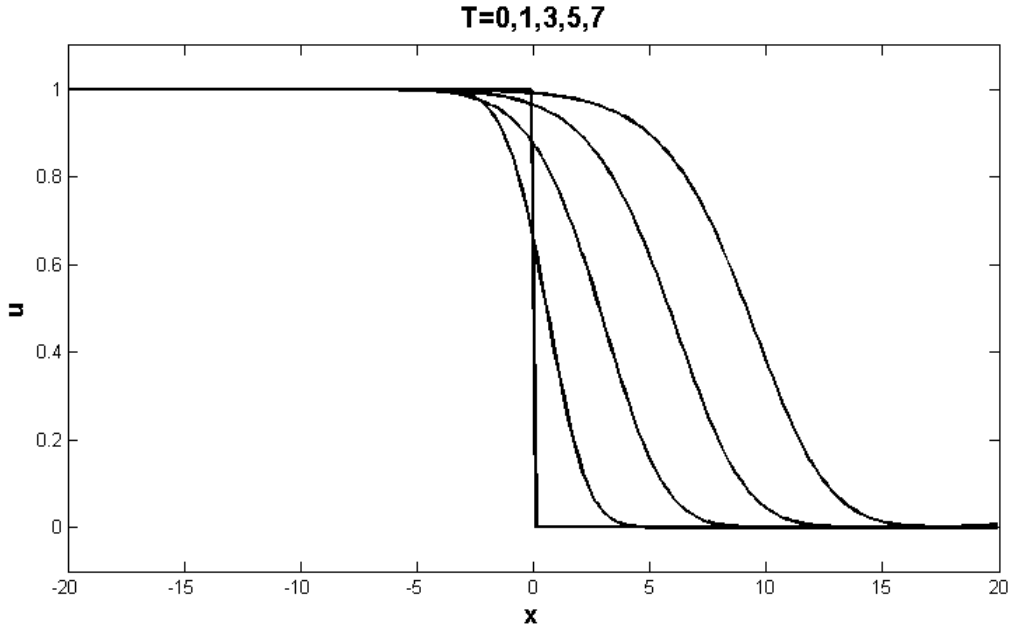

Figure 3.5 Evolution of the symmetric DDG solution to the 1-D Fisher-Kolmogorov equation (3.12) at $T = 0, 1, 3, 5, 7$.

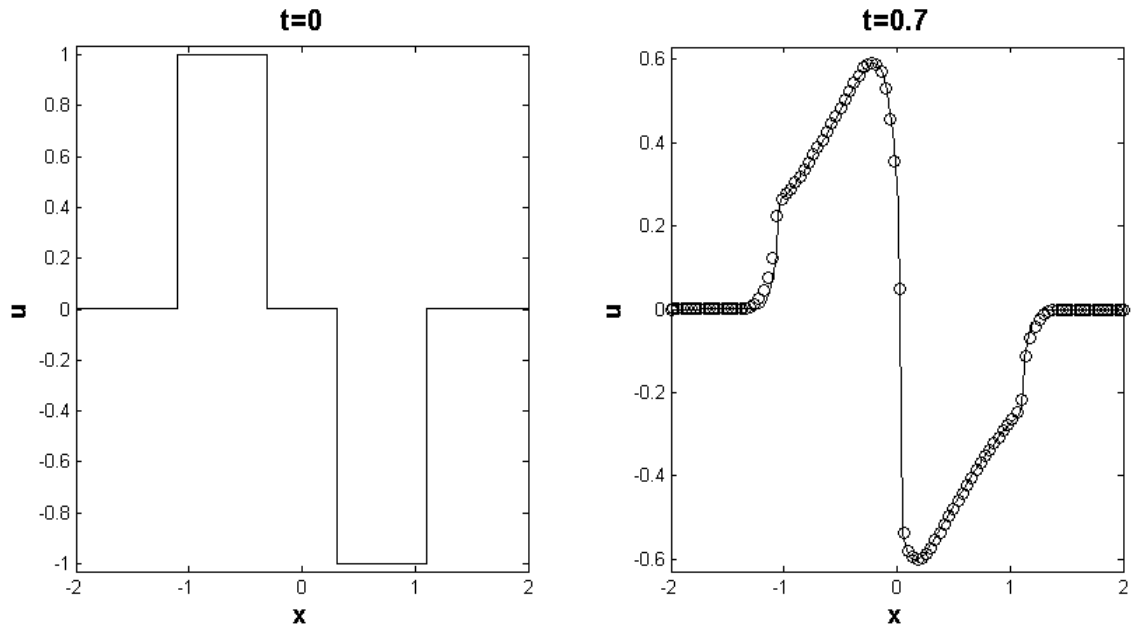

Figure 3.6 Strongly degenerate convection diffusion equation (3.12). Left: initial profile. Right: Symmetric diffusion approximation at time $T = 0.7$ for $N = 100$ (circles) and $N = 600$ (line).

CHAPTER 4. TWO DIMENSIONAL NONLINEAR DIFFUSION EQUATIONS

In this chapter we consider the two-dimensional nonlinear parabolic equation,

$$U_t - \nabla \cdot (A(U)\nabla U) = 0, \quad (\mathbf{x}, t) \in \Omega \times (0, T), \tag{4.1}$$

subject to initial data $U(\mathbf{x}, 0) = U_0(\mathbf{x})$ and periodic boundary conditions. The matrix $A(U) = (a_{ij}(U))$ is assumed symmetric positive definite and $\mathbf{x} = (x_1, x_2) \in \Omega \subset \mathbb{R}^2$. Similar to the one-dimensional case, we denote $b_{ij}(U) = \int a_{ij}(U)\, dU$, $i = 1, 2$, $j = 1, 2$.

First, the scheme is formulate for a model 2D heat equation on a triangular mesh. Flux admissibility is defined and ensures stability of the scheme. Further, an energy norm error estimate and L^2 error estimate can be shown using exact strategies as seen in Chapter 2 with some minor differences related to the change in dimension. Next, this linear case is extended to the nonlinear case in a straightforward manner. Last, numerical examples are provided.

Before we begin, some new concepts and notation is needed for higher dimension. Let $T_{\Delta \mathbf{x}} = \{K\}$ be a shape-regular partition of the domain Ω with elements K and denote $\Delta \mathbf{x} = \max_K \operatorname{diam}(K)$. As before, define $P^k(K)$ as the space of polynomials in the element K which are of degree at most k. Then, we have the piecewise polynomial numerical solution space as below,

$$\mathbb{V}^k_{\Delta \mathbf{x}} = \{v \in L^2(\Omega) : v\big|_K \in P^k(K), \forall K \in T_{\Delta \mathbf{x}}\}.$$

Along the element boundary ∂K, we use v^{int_K} to denote the value of v evaluated from inside the element K. Correspondingly we use v^{ext_K} to denote the value of v evaluated from outside the element K (inside the neighboring element). The average and jump of v on edge ∂K are defined as

$$\overline{v} = \frac{1}{2}\left(v^{ext_K} + v^{int_K}\right), \quad [v] = v^{ext_K} - v^{int_K}.$$

4.1 Scheme Formulation for Two Dimensional Model Equation

For sake of presentation, we first consider the case where $A(U) = I$ in (4.1). This gives us the below 2-D heat equation

$$U_t - \Delta U = 0, \quad (\mathbf{x}, t) \in \Omega \times (0, T). \tag{4.2}$$

As in the 1-D case, multiply the equation by test function, integrate over the computational cell K, perform integration by parts, add interface terms to symmetrize the scheme, and we have the following symmetric DDG scheme formulation. We seek the numerical solution $u \in \mathbb{V}^k_{\Delta\mathbf{x}}$ of U in (4.2) such that for all test functions $v \in \mathbb{V}^k_{\Delta\mathbf{x}}$ and on all elements K we have

$$\int_K u_t v \, d\mathbf{x} + \int_K \nabla u \cdot \nabla v \, d\mathbf{x} - \int_{\partial K} \widehat{u_{\mathbf{n}}} v^{int_K} \, ds + \int_{\partial K} \widehat{v_{\mathbf{n}}}[u] \, ds = 0, \tag{4.3}$$

where the numerical flux at the cell boundary ∂K is defined as

$$\widehat{w_{\mathbf{n}}} = \widehat{\nabla w \cdot \mathbf{n}} = \beta_0 \frac{[w]}{\Delta\mathbf{x}} + \overline{\frac{\partial w}{\partial \mathbf{n}}} + \beta_1 \Delta\mathbf{x}[w_{\mathbf{nn}}]. \tag{4.4}$$

Note, in the numerical flux definition, $\Delta\mathbf{x}$ is the average of the diameter of K and the diameter of its neighboring element. Here $\mathbf{n} = (n_1, n_2)$ is the outward unit normal along the element boundary ∂K. If the cell boundaries are straight lines, such as the triangular meshes, the numerical flux can be further simplified as

$$\widehat{w_{\mathbf{n}}} = \widehat{w_{x_1}} n_1 + \widehat{w_{x_2}} n_2,$$

with

$$\begin{cases} \widehat{w_{x_1}} = \beta_0 \frac{[w]}{\Delta\mathbf{x}} n_1 + \overline{w_{x_1}} + \beta_1 \Delta\mathbf{x}[w_{x_1x_1} n_1 + w_{x_2x_1} n_2] \\ \\ \widehat{w_{x_2}} = \beta_0 \frac{[w]}{\Delta\mathbf{x}} n_2 + \overline{w_{x_2}} + \beta_1 \Delta\mathbf{x}[w_{x_1x_2} n_1 + w_{x_2x_2} n_2]. \end{cases}$$

Again, the test function v is taken to be zero outside the element K, thus only one side (inside of K) contributes to the computation of $\widehat{v_{\mathbf{n}}}$ along the element boundary ∂K. Then, as in the 1-D case, we can define a notion of numerical flux admissibility in order to ensure L^2 stability.

Definition 4.1.1. *(Numerical flux admissibility) We call numerical flux* $\widehat{w_{\boldsymbol{n}}}$ *admissible if there exists* $\gamma \in (0,1)$, $\alpha > 0$ *such that for any* $w \in \mathbb{V}^k_{\Delta\mathbf{x}}$,

$$\gamma \sum_{K\in T_\Delta} \int_K |\nabla w|^2 \, d\mathbf{x} + 2 \sum_{K\in T_\Delta} \int_{\partial K} \widehat{w_{\boldsymbol{n}}}[w] \, ds \geq \alpha \sum_{K\in T_\Delta} \int_{\partial K} \frac{[w]^2}{\Delta\mathbf{x}} \, ds. \tag{4.5}$$

This admissibility ensures the following stability of the symmetric DDG method.

Theorem 4.1.1. *(Stability) Consider the symmetric DDG scheme (4.3)-(4.4). If the numerical flux is admissible as described in (4.5), then we have*

$$\frac{1}{2}\int_\Omega u^2(\mathbf{x},T)\, d\mathbf{x} + (1-\gamma)\int_0^T \sum_{K\in T_\Delta} \int_K |\nabla u|^2 \, d\mathbf{x}dt + \alpha \int_0^T \sum_{K\in T_\Delta} \int_{\partial K} \frac{[u]^2}{\Delta\mathbf{x}} \, dsdt \leq \frac{1}{2}\int_\Omega U_0^2(\mathbf{x}) \, d\mathbf{x}. \tag{4.6}$$

This can be proved directly by summation over all $K \in T_{\Delta\mathbf{x}}$ of (4.3) with $v = u$ and by using the admissibility condition (4.5). For the 2-D linear model equation (4.2), we have the following energy norm and $L^2(L^2)$ error estimates.

Theorem 4.1.2. *(Energy norm error estimate in two dimensions) Consider the 2-D linear model equation (4.2). Let* $e := u - U$ *be the error between the exact solution* U *and the numerical solution* u *of the symmetric DDG method (4.3)-(4.4), we have for multiindex* α, $|\alpha| = k+1$,

$$|||u - U||| \leq C|||D^\alpha U(\cdot,T)|||(\Delta\mathbf{x})^k.$$

Here, the energy norm is of the same form as the one dimension case.

Theorem 4.1.3. *($L^2(L^2)$ error estimate in two dimensions) Consider the 2-D linear model equation (4.2). Let* $e := u - U$ *be the error between the exact solution* U *and the numerical solution* u *of the symmetric DDG method (4.3)-(4.4), we have,*

$$\|u - U\|_{L^2(0,T;L^2)} \leq C(\Delta\mathbf{x})^{k+1} \left(\|U\|_{L^\infty(0,T;H^{k+1})} + \|U\|_{L^2(0,T;H^{k+1})} + \Delta\mathbf{x}\|U_t\|_{L^2(0,T;H^k)} \right).$$

Remark 4.1.1. *For the proof of Theorems 4.1.2 and 4.1.3, the same proof as the one dimension case can be used here with a minor differences. The primary difference is the increase in dimensionality from one to two. In the approximation properties provided by Lemmas 2.3.1 and 2.3.2, the space dimension* n *appears in the terms* $\frac{n}{q} - \frac{n}{p}$. *In many cases where these*

estimates are used, $p = q$ and there is no issue. In some cases though, specifically on the estimation of boundary terms, $q = \infty$ and $p = 2$. This case requires a little thought in higher dimension, but the result is the same in the end. To illustrate this, this calculation is performed for one term in the proof of the Parabolic Lift Theorem 2.4.2. All similar calculations can be handled in the same manner. Consider the first term in the estimation of I_2.

$$\begin{aligned}
\frac{\beta_0}{\Delta\mathbf{x}}\sum_K\int_{\partial K}[e][\Phi-\mathbb{P}(\Phi)]\,ds &\le \frac{\beta_0}{\Delta\mathbf{x}}\left(\sum_K\int_{\partial K}[e]^2\,ds\right)^{1/2}\left(\sum_K\int_{\partial K}[\Phi-\mathbb{P}(\Phi)]^2\,ds\right)^{1/2}\\
&\le \frac{\beta_0}{\Delta\mathbf{x}}\left(\sum_K\int_{\partial K}[e]^2\,ds\right)^{1/2}\|\Phi-\mathbb{P}(\Phi)\|_\infty\left(\sum_K\int_{\partial K}\,ds\right)^{1/2}\\
&\le \frac{\beta_0}{\Delta\mathbf{x}}\left(\sum_K\int_{\partial K}[e]^2\,ds\right)^{1/2}C\Delta\mathbf{x}^1\|\Phi\|_{H^2}\Delta\mathbf{x}^{1/2}\\
&\le C\beta_0\Delta\mathbf{x}^{1/2}\left(\sum_K\int_{\partial K}[e]^2\,ds\right)^{1/2}\|\Phi\|_{H^2}
\end{aligned}$$

4.2 Extension to Two Dimensional Nonlinear Diffusion Problems

We consider the fully nonlinear 2-D case as given in (4.1). The scheme formulation is given as follows. We seek approximation $u \in \mathbb{V}^k_{\Delta\mathbf{x}}$ of U in (4.1) such that the following scheme is satisfied for all test functions $v \in \mathbb{V}^k_{\Delta\mathbf{x}}$ on all elements K,

$$\int_K u_t v\,d\mathbf{x}+\int_K\sum_{i,j=1}^{2}b_{ij}(u)_{x_j}v_{x_i}\,d\mathbf{x}-\int_{\partial K}\sum_{i,j=1}^{2}\widehat{b_{ij}(u)_{x_j}}n_i v^{int_K}\,ds+\int_{\partial K}\sum_{i,j=1}^{2}\widehat{v_{x_j}}n_i[b_{ij}(u)]\,ds=0. \tag{4.7}$$

For $j = 1, 2$, the numerical flux terms are defined as

$$\widehat{b_{ij}(u)_{x_j}} = \beta_0\frac{[b_{ij}(u)]}{\Delta\mathbf{x}}n_j+\overline{b_{ij}(u)_{x_j}}+\beta_1\Delta\mathbf{x}[b_{ij}(u)_{x_1x_j}n_1+b_{ij}(u)_{x_2x_j}n_2]$$

$$\widehat{v_{x_j}} = \beta_0\frac{[v]}{\Delta\mathbf{x}}n_j+\overline{v_{x_j}}+\beta_1\Delta\mathbf{x}[v_{x_1x_j}n_1+v_{x_2x_j}n_2].$$

We should specify that the 2-D numerical examples in the following section are implemented on rectangular meshes.

4.3 Numerical Examples

Example 4.3.1: 2-D linear diffusion equation

$$\begin{cases} U_t - \epsilon \left(U_{xx} + U_{yy}\right) = 0, \quad (x, y) \in [0, 2\pi] \times [0, 2\pi] \\ U(x, 0) = \sin(x + y), \end{cases} \tag{4.8}$$

with periodic boundary conditions and $\epsilon = 0.01$. P^k polynomial approximations are carried out and errors and orders are listed in Table 4.1 with $k = 2, 3, 4$. Optimal convergence is obtained. Note the numerical flux coefficients are chosen according to the 1-D analysis.

Table 4.1 2-D linear diffusion equation (4.8), P^k approximations with $k = 2, 3, 4$. $T = 5$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 10$	$N = 20$		$N = 30$		$N = 40$	
β_0	3/2	$k = 2$	L^2	6.32E-03	9.13E-04	2.8	2.71E-04	3.0	1.14E-04	3.0
β_1	1/4		L^∞	1.24E-02	1.84E-03	2.7	5.47E-04	3.0	2.30E-04	3.0
β_0	11/4	$k = 3$	L^2	3.74E-04	2.51E-05	3.9	5.01E-06	4.0	1.61E-06	3.9
β_1	3/32		L^∞	2.32E-03	1.24E-04	4.2	2.48E-05	4.0	7.85E-06	4.0
β_0	9/2	$k = 4$	L^2	1.79E-05	5.07E-07	5.1	6.60E-08	5.0	1.56E-08	5.0
β_1	1/20		L^∞	1.01E-04	3.00E-06	5.1	3.93E-07	5.0	9.34E-08	5.0

Example 4.3.2: 2-D anisotropic linear diffusion equation

$$\begin{cases} U_t - \epsilon \left(U_{xx} + U_{xy} + U_{yy}\right) = 0, \quad (x, y) \in [0, 2\pi] \times [0, 2\pi] \\ U(x, 0) = \sin(x + y), \end{cases} \tag{4.9}$$

with $\epsilon = 0.01$. On the rectangular mesh $I_i \times I_j = [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$, the numerical flux for the mixed term should be taken as (according to (**??**)),

$$\widehat{u_x} = \overline{u_x} + \beta_1 \Delta \mathbf{x} [u_{yx}] \quad \text{at } y = y_{j \pm 1/2}.$$

Again, accuracy test is carried out with P^k approximations and errors and orders are listed in Table 4.2 with final time $T = 5$. Optimal $(k + 1)$th order of convergence is obtained.

Example 4.3.3: 2-D nonlinear porous medium equation

$$U_t - (U^2)_{xx} - (U^2)_{yy} = 0, \quad (x, y) \in [-10, 10] \times [-10, 10] \tag{4.10}$$

Table 4.2 2-D anisotropic diffusion equation (4.9), P^k polynomial approximations with $k = 2, 3, 4$.

				Error	Error	Order	Error	Order	Error	Order
				$N = 10$	$N = 20$		$N = 30$		$N = 40$	
β_0	5	$k = 2$	L^2	2.69E-03	3.29E-04	3.0	9.69E-05	3.0	4.08E-05	3.0
β_1	1/12		L^∞	1.87E-02	2.27E-03	3.0	6.68E-04	3.0	2.82E-04	3.0
β_0	5	$k = 3$	L^2	3.11E-04	2.03E-05	3.9	4.11E-06	3.9	1.33E-06	3.9
β_1	1/40		L^∞	2.03E-03	1.30E-04	4.0	2.61E-05	4.0	8.32E-06	4.0
β_0	30	$k = 4$	L^2	1.94E-05	5.54E-07	5.1	7.24E-08	5.0	1.77E-08	4.9
β_1	1/40		L^∞	8.51E-05	2.26E-06	5.2	2.90E-07	5.1	6.91E-08	5.0

with zero boundary conditions. The initial condition is given by two bumps as

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

Piecewise linear approximation with numerical flux coefficients $\beta_0 = \frac{3}{2}$ and $\beta_1 = \frac{1}{10}$ is implemented on a 80×80 rectangular mesh. Note, even though a linear approximation is used, second order jump terms are included within the numerical flux in the scheme formulation (4.7) because of the nonlinearity of the given problem. Results of this test are illustrated in Figure 3 with $T = 0, 0.5, 1.0$, and 4.0. The symmetric DDG solution effectively captures the evolution of the surface with sharp resolution.

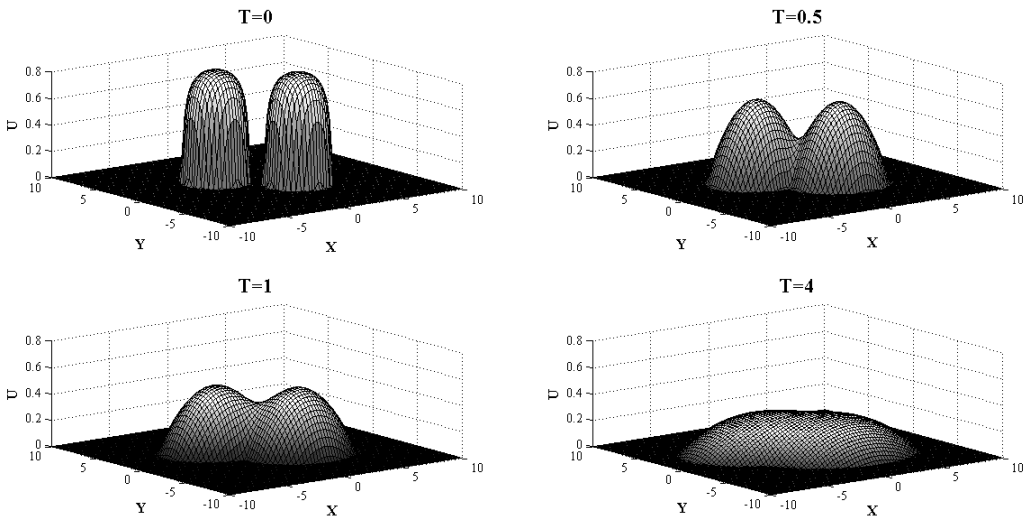

Figure 4.1 Evolution of the symmetric DDG solution to the 2-D nonlinear porous medium equation (4.10) at times $T = 0, 0.5, 1$, and 4.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

The symmetric direct discontinuous Galerkin method was formulated and analyzed for one dimensional linear diffusion, nonlinear diffusion, nonlinear convection diffusion problems as well as two dimensional linear and nonlinear problems. One characteristic of this scheme is that the numerical flux is defined using the exact solution for the heat equation in an analogous setting was used as inspiration. In order to make the scheme symmetric, a test function numerical flux was defined which is of the same form as the numerical solution flux. Next, in order to choose suitable free coefficients in the numerical flux formula, a notion of flux admissibility was defined and analyzed completely resulting in an explicit formula used for choosing flux coefficients. Optimal error estimates were then shown for both the energy norm as well as the L^2 norm for certain cases. Last, numerical results were given illustrating performance on uniform and nonuniform mesh as well as exploration of admissibility analysis results.

5.2 Future Work

In this section, outlines of future work are given. Many of the below topics have been considered, but the work is incomplete.

5.2.1 Further Admissibility Analysis

One project which is currently being considered is a complete admissibility analysis as in §2.2 for one dimensional nonuniform mesh and nonlinear cases as well as two dimensional cases.

The nonuniform mesh case has been considered in detail, but computation is considerably more difficult. For this case, the same definition of flux admissibility (2.9) as in the uniform

mesh is considered. Only now, $\Delta x = \Delta x_{j+1/2} = \frac{1}{2}(\Delta x_j + \Delta x_{j+1})$. The exact same procedure as in §2.2 for the uniform mesh case can be followed with needed modifications, only in the analog of Lemma 2.2.2, computation becomes difficult. Using some severe estimation, it can be shown that the numerical flux is admissible provided

$$\beta_0 > 4M^2 \left(\beta_1^2 \left(1 + M^4\right) \left(\frac{k^2(k^4 - 2k^2 + 2)}{3} \right) - \beta_1 \left(k^2(k^2 - 1)\right) + k^2 \right),$$

where,

$$M = \max_j \frac{\max\{\Delta x_j, \Delta x_{j+1}\}}{\min\{\Delta x_j, \Delta x_{j+1}\}}$$

for $k \geq 1$. After performing numerical tests as in Example 2.5.1, this estimate proves to be sufficient, though more restrictive than needed. The reason for this is the appearance of M^6 within the estimate. For a highly nonuniform mesh, this value becomes quite large and dominates the estimate. I believe this result can be significantly improved using this same strategy with new ideas and techniques.

The nonuniform diffusion admissibility analysis has also been considered carefully, though a practical estimate has not been obtained. The admissibility definition considered is given by (3.3). Again, the same strategy as in §2.2 was considered. The primary difficulty again is with the analog of Lemma 2.2.2. Now, the problem becomes finding a way to eliminate the nonlinearity. This is difficult when considering $\widehat{b(u)_x}$.

Two dimensional admissibility analysis is yet to be considered, though I feel especially for the linear uniform mesh case, something can be said.

5.2.2 Maximum Principle Satisfying Method

An important property of the heat equation (2.1) is that it satisfies a weak maximum principle. That is, if

$$M = \max_x U_0(x), \quad m = \max_x U_0(x), \tag{5.1}$$

then, we have that $U(x,t) \in [m, M]$ for all x and t. Similar things can be said for nonlinear and higher dimensional diffusion equations. Then, it is natural to desire that a numerical method designed for these problems maintain this important property because solutions outside the interval $[m, M]$ often have no physical meaning, such as negative density.

This problem of designing maximum principle satisfying methods proves to be quite challenging. Recently, work in this area for hyperbolic conservation laws has been done by Shu and Zhang in [49]. They were able to construct high order DG and finite volume schemes satisfying a weak maximum principle. All that was needed to do this was to apply an appropriate limiter for each time evolution. Then, by looking at the quadrature representation of the polynomial approximate solution, one can prove the numerical method satisfies a maximum principle.

This problem for diffusion differs from hyperbolic problems in a significant way. For the hyperbolic case, the numerical flux depends on u^+ and u^- on cell interfaces. As a result, the proof involves polynomial values at endpoints only, all other quadrature points remain the same. For the diffusion case, the numerical flux depends on u_x^+ and u_x^-. Changing these values at cell interfaces changes the polynomials within the entire computational cell. This makes analysis considerably more challenging and generalizing work for any polynomial degree is difficult.

Despite these challenges, DDG methods seem to maintain this positivity preserving property. Dr. Yan has some promising result regarding this topic, I would like to participate in this project or apply the technique to applications involving convection diffusion equations.

5.2.3 Extensions, Superconvergence, and Posteriori Error Estimates

Now that a foundation is established for the symmetric DDG methods, there are many potential applications to other classes of problems. In particular, compressible Navier-Stokes equations and nonlinear porous medium equations are topics of interest. Also, application to general elliptic type equations would provide interesting future work. It is expected that the resulting mass matrix of these DDG methods have a lower condition number in comparison with Interior Penalty methods. This is yet to be explored in detail. Also, detailed Fourier analysis has been performed for these DDG methods in [48]. This work has indicated superconvergence rates at quadrature points. It would be quite interesting to see how DDG methods compare to existing methods with regards to superconvergence. In addition, posteriori error estimates is something which I would like to see applied to DDG methods.

APPENDIX A. CODING NOTES

One Dimension

Below are some notes for coding this symmetric DDG scheme for the above heat equation (2.1). So, we seek approximation $u \in \mathbb{V}_{\Delta x}^k$ of U such that the following scheme is satisfied for all test functions $v \in \mathbb{V}_{\Delta x}^k$ and all cells I_j where $\Omega = \bigcup_j I_J$.

$$\int_{I_j} u_t v \, dx = \widehat{u_x} v \Big|_{j-1/2}^{j+1/2} - \int_{I_j} u_x v_x \, dx - \left(\widehat{v_x}[u] \Big|_{j+1/2} + \widehat{v_x}[u] \Big|_{j-1/2} \right) = \text{RES}$$

Choose the basis functions on I_j to be the cell centered polynomials

$$\phi_j^m = \left(\frac{x - x_j}{\Delta x_j} \right)^m, \quad m = 0, 1, \ldots, k.$$

The above scheme holding for all test functions $v \in \mathbb{V}_{\Delta x}^k$ equates to ensuring the scheme holds for each of these basis elements. Then, we can write u as follows.

$$u \Big|_{I_j} = \sum_{m=0}^{k} u_j^m(t) \phi_j^m$$

$$u_x \Big|_{I_j} = \sum_{m=1}^{k} \frac{m}{\Delta x_j} u_j^m(t) \phi_j^{m-1}$$

Then, define the mass matrix

$$M = \left(\int_{-1/2}^{1/2} x^\mu x^\nu \, dx \right)_{\mu,\nu=0}^{k}.$$

Change of variable for the basis functions is used here in order to simplify computation. In this case, mass matrix M is the same for any choice of stepsize and therefore can be precomputed and stored. Then, our scheme can now be written as the following system of ODEs for each cell I_j.

$$\Delta x_j M \left(\overrightarrow{u_j} \right)_t = \text{RES}$$

which implies

$$\Rightarrow (\overrightarrow{u_j})_t = \frac{1}{\Delta x_j} M^{-1}\mathrm{RES} = R(u,t)$$

Here,

$$\overrightarrow{u_j} = \begin{pmatrix} u_j^0(t) \\ u_j^1(t) \\ u_j^1(t) \\ \vdots \\ u_j^{mp}(t) \end{pmatrix}.$$

This system of ODEs is solved via a regular 3 step Runge-Kutta method.

$$\begin{aligned} u^{(1)} &= u^n + \Delta t R(u^n, t^n) \\ u^{(2)} &= \frac{3}{4}u^n + \frac{1}{4}u^{(1)} + \frac{1}{4}\Delta t R(u^{(1)}, t^n + \Delta t) \\ u^{n+1} &= \frac{1}{3}u^n + \frac{2}{3}u^{(2)} + \frac{2}{3}\Delta t R(u^{(2)}, t^n + \frac{1}{2}\Delta t) \end{aligned}$$

APPENDIX B. HILBERT MATRIX RESULTS

This appendix gives all Hilbert matrix results required for the admissibility analysis in Section 2.2. The primary combinatorial technique used is known as Wilf's Snake Oil Method [47]. This technique will be illustrated in proofs of below Lemmas.

The Hilbert matrix H is named for David Hilbert and was first discussed in 1894. It is defined as a square matrix with entries $H_{ij} = \frac{1}{i+j-1}$. In its original setting, this matrix came about naturally when considering integration of the product of two arbitrary polynomials. That is, $H_{ij} = \int_0^1 x^{i+j-2}\, dx$. This is the same setting which gave rise to Hilbert matrix computations in this thesis. It is easy to see that H is symmetric and positive definite. These two properties prove useful below and in Section 2.2. Here, we will consider the $k \times k$ Hilbert matrix and denote the dimension by H_k. Also, as mentioned in Section 2.2, the inverse of this matrix is known [34] and is listed below for reference.

$$\left(H_k^{-1}\right)_{ij} = (-1)^{i+j}(i+j-1)\binom{k+i-1}{k-j}\binom{k+j-1}{k-i}\binom{i+j-2}{i-1}^2$$

Theorem B.0.1.

$$\sum_{i,j=1}^{k} i\left(H_k^{-1}\right)_{ij} = \sum_{i,j=1}^{k} j\left(H_k^{-1}\right)_{ij} = \frac{k^2(k^2+1)}{2}$$

Proof. Proceed by means of induction on dimension k of the Hilbert matrix. The $k = 1, 2$ cases can be shown by hand. Then, assume the result holds for some k. Then, it remains to show

$$\sum_{i,j=1}^{k+1} j\left(H_{k+1}^{-1}\right)_{ij} - \sum_{i,j=1}^{k} j\left(H_k^{-1}\right)_{ij} = \frac{(k+1)^2((k+1)^2+1)}{2} - \frac{k^2(k^2+1)}{2} = (2k+1)(k^2+k+1).$$

First, not that for $1 \le i, j \le n$,

$$\begin{aligned}
\left(H_{k+1}^{-1}\right)_{ij} &= (-1)^{i+j}(i+j-1)\binom{(k+1)+i-1}{(k+1)-j}\binom{(k+1)+j-1}{(k+1)-i}\binom{i+j-2}{i-1}^2 \\
&= \frac{(k+i)(k+j)}{(k+1-i)(k+1-j)}(-1)^{i+j}(i+j-1)\binom{k+i-1}{k-j}\binom{k+j-1}{k-i}\binom{i+j-2}{i-1}^2 \\
&= \frac{(k+i)(k+j)}{(k+1-i)(k+1-j)}\left(H_k^{-1}\right)_{ij}.
\end{aligned}$$

Then, we have the following calculation.

$$\begin{aligned}
&\sum_{i,j=1}^{k+1} j\left(H_{k+1}^{-1}\right)_{ij} - \sum_{i,j=1}^{k} j\left(H_k^{-1}\right)_{ij} \\
&\qquad = \sum_{i,j=1}^{k} j\left(\frac{(k+i)(k+j)}{(k+1-i)(k+1-j)} - 1\right)\left(H_k^{-1}\right)_{ij} + \sum_{j=1}^{k} j\left(H_{k+1}^{-1}\right)_{k+1j} \\
&\qquad\quad + \sum_{i=1}^{k}(k+1)\left(H_{k+1}^{-1}\right)_{ik+1} + (n+1)\left(H_{k+1}^{-1}\right)_{k+1k+1} \\
&\qquad = \sum_{i,j=1}^{k} j\left(\frac{(2k+1)(i+j-1)}{(k+1-i)(k+1-j)}\right)\frac{(-1)^{i+j}}{i+j-1}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k-i)!(j-1)!(k-j)!} \\
&\qquad\quad + \sum_{j=1}^{k}(j+k+1)\frac{(-1)^{k+1+j}}{k+j}\frac{\prod_{n=0}^{k}(k+1+n)(j+n)}{k!(j-1)!(k+1-j)!} \\
&\qquad\quad + (k+1)\frac{(-1)^{2k+2}}{2k+1}\frac{\prod_{n=0}^{k}(k+1+n)(k+1+n)}{k!k!} \\
&\qquad = (2k+1)\left(\sum_{i,j=1}^{k} j\frac{(-1)^{i+j}}{(k+1-i)(k+1-j)}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k-i)!(j-1)!(k-j)!}\right. \\
&\qquad\quad + \sum_{j=1}^{k}(j+k+1)(-1)^{k+1+j}\frac{\prod_{n=0}^{k-1}(k+1+n)(j+n)}{k!(j-1)!(k+1-j)!} \\
&\qquad\quad \left. +(k+1)\frac{\prod_{n=0}^{k-1}(k+1+n)(k+1+n)}{k!k!}\right) \\
&\qquad = (2k+1)\left(\sum_{i,j=1}^{k+1} j(-1)^{i+j}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k+1-i)!(j-1)!(k+1-j)!}\right) \\
&\qquad = (2k+1)\left(\sum_{i,j=0}^{k}(j+1)(-1)^{i+j}\frac{\prod_{n=0}^{k-1}(i+1+n)(j+1+n)}{i!(k-i)!j!(k-j)!}\right) \\
&\qquad = (2k+1)\left(\sum_{i,j=0}^{k}(j+1)(-1)^{i+j}\binom{2i}{i}\binom{k+i}{2i}\binom{2j}{j}\binom{k+j}{2j}\right)
\end{aligned}$$

$$= (2k+1)\left(\sum_{j=0}^{k}(j+1)(-1)^{j+k}\binom{2j}{j}\binom{k+j}{2j}\left(\sum_{i=0}^{k}(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i}\right)\right)$$
$$= (2k+1)\left(\sum_{j=0}^{k}(j+1)(-1)^{j+k}\binom{2j}{j}\binom{k+j}{2j}\right)$$
$$= (2k+1)(k^2+k+1).$$

In the final two steps, Lemmas B.0.1 and B.0.2 were made use of. □

Lemma B.0.1.

$$\sum_{i=0}^{k}(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} = 1$$

Proof. Define function f as

$$f(k) := \sum_{i=0}^{k}(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i}$$

for $k \geq 1$. Then, define some function F with power series given by $F(x) = \sum_{k\geq 1} f(k)x^k$ for $|x| < 1$. The goal will be to compute this power series representation of F by interchanging order of summation and end up with a simple function. Then, by comparing coefficients of the power series of this simple function, we end up with a formula for $f(k)$. Pretty cool huh? Surprisingly, this strategy works for a number of combinatorial summation formulas. Then, compute as outlined.

$$\begin{aligned}
F(x) &= \sum_{k\geq 1} f(k)x^k = \sum_{k\geq 1} x^k \sum_{i=0}^{k}(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} \\
&= \sum_{i\geq 1}\sum_{k\geq i} x^k(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} + \sum_{k\geq 1}(-1)^k x^k \\
&= \sum_{i\geq 1}\binom{2i}{i}x^{-i}\sum_{k\geq i}\binom{k+i}{2i}(-x)^{k+i} + (-x)\sum_{k\geq 0}(-x)^k \\
&= \sum_{i\geq 1}\binom{2i}{i}x^{-i}\sum_{m\geq 2i}\binom{m}{2i}(-x)^m + (-x)\frac{1}{1+x} \\
&= \sum_{i\geq 1}\binom{2i}{i}x^{-i}\frac{(-x)^{2i}}{(1+x)^{2i+1}} - \frac{x}{1+x}, \quad \text{(reference [47])} \\
&= \sum_{i\geq 1}\binom{2i}{i}\frac{x^i}{(1+x)^{2i+1}} - \frac{x}{1+x}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1+x}\sum_{i\geq 1}\binom{2i}{i}y^i - \frac{x}{1+x}, \quad y = \frac{x}{(1+x)^2} \\
&= \frac{1}{1+x}\left(-1+\sum_{i\geq 0}\binom{2i}{i}y^i\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(-1+\frac{1}{\sqrt{1-4y}}\right) - \frac{x}{1+x}, \quad \text{(reference [47])} \\
&= \frac{1}{1+x}\left(-1+\frac{1+x}{1-x}\right) - \frac{x}{1+x} \\
&= \frac{x}{1-x} \\
&= x\sum_{k\geq 0}x^k \\
&= \sum_{k\geq 1}x^k
\end{aligned}
$$

Therefore, $f(k) = 1$ for all $k \geq 1$ as needed. □

Lemma B.0.2.

$$\sum_{i=0}^{k}(i+1)(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} = k^2+k+1$$

Proof. This proof will follow the same strategy as in Lemma B.0.1. Define function f as

$$f(k) := \sum_{i=0}^{k}(i+1)(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i}$$

and some function F with power series given by $F(x) = \sum_{k\geq 1} f(k)x^k$ for $|x| < 1$. Then, compute this power series as below.

$$
\begin{aligned}
F(x) = \sum_{k\geq 1}f(k)x^k &= \sum_{k\geq 1}x^k\sum_{i\geq 0}(i+1)(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} \\
&= \sum_{i\geq 1}\sum_{k\geq i}x^k(i+1)(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} + \sum_{k\geq 1}x^k(-1)^k \\
&= \sum_{i\geq 1}(i+1)\binom{2i}{i}x^{-i}\sum_{k\geq i}\binom{k+i}{2i}(-x)^{k+i} + (-x)\sum_{k\geq 0}x^k(-1)^k \\
&= \sum_{i\geq 1}(i+1)\binom{2i}{i}x^{-i}\sum_{m\geq 2i}\binom{m}{2i}(-x)^m + (-x)\frac{1}{1+x} \\
&= \sum_{i\geq 1}(i+1)\binom{2i}{i}x^{-i}\frac{(-x)^{2i}}{(1+x)^{2i+1}} - \frac{x}{1+x}, \quad \text{(reference [47])}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{1+x}\sum_{i\geq 1}(i+1)\binom{2i}{i}\left(\frac{x}{(1+x)^2}\right)^i - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(\sum_{i\geq 1}(i+1)\binom{2i}{i}y^i\right) - \frac{x}{1+x}, \quad y = \frac{x}{(1+x)^2} \\
&= \frac{1}{1+x}\left(\frac{d}{dy}y\sum_{i\geq 1}\binom{2i}{i}y^i\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(\frac{d}{dy}y\left(-1+\sum_{i\geq 0}\binom{2i}{i}y^i\right)\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(\frac{d}{dy}y\left(-1+\frac{1}{\sqrt{1-4y}}\right)\right) - \frac{x}{1+x}, \quad \text{(reference [47])} \\
&= \frac{1}{1+x}\left(-1+\frac{1}{\sqrt{1-4y}}+\frac{2y}{(1-4y)^{3/2}}\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(-1+\frac{1+x}{1-x}+\frac{2x(1+x)}{(1-x)^3}\right) - \frac{x}{1+x} \\
&= \frac{x}{1-x}+\frac{x}{(1-x)^2}+\frac{x(1+x)}{(1-x)^3} \\
&= x\sum_{k\geq 0}x^k + \sum_{k\geq 1}kx^k + \sum_{k\geq 1}k^2x^k \\
&= \sum_{k\geq 1}(k^2+k+1)x^k
\end{aligned}
$$

The final power series come from computing the first and second derivatives of the regular geometric series. Therefore, $f(k) = k^2 + k + 1$ as needed. □

Theorem B.0.2.

$$\sum_{i,j=1}^{k} i^2\left(H_k^{-1}\right)_{ij} = \sum_{i,j=1}^{k} j^2\left(H_k^{-1}\right)_{ij} = \frac{k^2(k^4+4k^2+1)}{6}$$

Proof. The strategy here is the same as in the proof of Theorem B.0.1. Again, utilize induction on dimension k. The $k = 1, 2$ cases can be shown by hand. Assume the result holds for some k. It remains to show

$$\sum_{i,j=1}^{k+1} j^2\left(H_{k+1}^{-1}\right)_{ij} - \sum_{i,j=1}^{k} j^2\left(H_k^{-1}\right)_{ij} = \frac{(2k+1)(k^4+2k^3+5k^2+4k+2)}{2}.$$

Then, performing the same calculation as in the proof of Theorem B.0.1, everything is identical

except the power of the j term.

$$\begin{aligned}
&\sum_{i,j=1}^{k+1} j^2 \left(H_{k+1}^{-1}\right)_{ij} - \sum_{i,j=1}^{k} j^2 \left(H_k^{-1}\right)_{ij} \\
&\qquad = (2k+1)\left(\sum_{j=0}^{k}(j+1)^2(-1)^{j+k}\binom{2j}{j}\binom{k+j}{2j}\right) \\
&\qquad = (2k+1)\frac{k^4+2k^3+5k^2+4k+2}{2}.
\end{aligned}$$

Again, Lemma B.0.1 was needed. For the last step, Lemma B.0.3 is made use of. □

Lemma B.0.3.

$$\sum_{i=0}^{k}(i+1)^2(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i} = \frac{k^4+2k^3+5k^2+4k+2}{2}$$

Proof. This proof will follow the same strategy as in Lemmas B.0.1 and B.0.2, so some details are left out. Define function f as

$$f(k) := \sum_{i=0}^{k}(i+1)^2(-1)^{i+k}\binom{2i}{i}\binom{k+i}{2i}$$

and some function F with power series given by $F(x) = \sum_{k\geq 1} f(k)x^k$ for $|x| < 1$. Then, compute this power series as below.

$$\begin{aligned}
F(x) &= \sum_{k\geq 1} f(k)x^k \\
&= \frac{1}{1+x}\left(\sum_{i\geq 1}(i+1)^2\binom{2i}{i}y^i\right) - \frac{x}{1+x}, \quad y = \frac{x}{(1+x)^2} \\
&= \frac{1}{1+x}\left(\frac{d}{dy}y\left(\frac{d}{dy}y\sum_{i\geq 1}\binom{2i}{i}y^i\right)\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(\left(\frac{d}{dy}y\right)^2\left(-1+\frac{1}{\sqrt{1-4y}}\right)\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(\frac{d}{dy}y\right)\left(-1+\frac{1}{\sqrt{1-4y}}+\frac{2y}{(1-4y)^{3/2}}\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(-1+\frac{1}{\sqrt{1-4y}}+\frac{6y}{(1-4y)^{3/2}}+\frac{12y^2}{(1-4y)^{5/2}}\right) - \frac{x}{1+x} \\
&= \frac{1}{1+x}\left(-1+\frac{1+x}{1-x}+\frac{6x(1+x)}{(1-x)^3}+\frac{12x^2(1+x)}{(1-x)^5}\right) - \frac{x}{1+x}
\end{aligned}$$

$$= \frac{x}{1-x} + 2\frac{x}{(1-x)^2} + \frac{5}{2}\frac{x(1+x)}{(1-x)^3} + \frac{x(x^2+4x+1)}{(1-x)^4} + \frac{1}{2}\frac{x(x^3+11x^2+11x+1)}{(1-x)^5}$$
$$= x\sum_{k\geq 0} x^k + 2\sum_{k\geq 1} kx^k + \frac{5}{2}\sum_{k\geq 1} k^2 + \sum_{k\geq 1} k^3x^k + \frac{1}{2}\sum_{k\geq 1} k^4x^k$$
$$= \sum_{k\geq 1} \frac{k^4+2k^3+5k^2+4k+2}{2} x^k$$

Therefore, $f(k) = \frac{k^4+2k^3+5k^2+4k+2}{2}$ as needed. □

Theorem B.0.3.

$$\sum_{j=1}^{k} j\left(H_k^{-1}\right)_{ij} = (k^2-i+1)\sum_{j=1}^{k}\left(H_k^{-1}\right)_{ij}$$

Proof. Proceed by means of induction on dimension k of the Hilbert matrix. The $k = 1, 2$ cases can be shown by hand. Then, assume the result holds for some k. Then, it remains to show

$$\sum_{j=1}^{k+1} j\left(H_{k+1}^{-1}\right)_{ij} - \sum_{j=1}^{k} j\left(H_k^{-1}\right)_{ij} = (2k+1)(k^2+k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(i+n)}{(i-1)!(k+1-i)!}.$$

Again, this is the same calculation as in Theorem B.0.1 with summation over j only, so many steps are much the same.

$$\sum_{j=1}^{k+1} j\left(H_{k+1}^{-1}\right)_{ij} - \sum_{j=1}^{k} j\left(H_k^{-1}\right)_{ij}$$
$$= \sum_{j=1}^{k} j\left(\frac{(k+i)(k+j)}{(k+1-i)(k+1-j)} - 1\right)\left(H_k^{-1}\right)_{ij} + (k+1)\left(H_{k+1}^{-1}\right)_{ik+1}$$
$$= \sum_{j=1}^{k} j\left(\frac{(2k+1)(i+j-1)}{(k+1-i)(k+1-j)}\right)\frac{(-1)^{i+j}}{i+j-1}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k-i)!(j-1)!(k-j)!}$$
$$+ (k+1)\frac{(-1)^{k+1+i}}{k+i}\frac{\prod_{n=0}^{k}(k+1+n)(i+n)}{k!(i-1)!(k+1-i)!}$$
$$= (2k+1)\left(\sum_{j=1}^{k} j\frac{(-1)^{i+j}}{(k+1-i)(k+1-j)}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k-i)!(j-1)!(k-j)!}\right.$$
$$\left. + (k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(k+1+n)(i+n)}{k!(i-1)!(k+1-i)!}\right)$$
$$= (2k+1)\left(\sum_{j=1}^{k+1} j(-1)^{i+j}\frac{\prod_{n=0}^{k-1}(i+n)(j+n)}{(i-1)!(k+1-i)!(j-1)!(k+1-j)!}\right)$$
$$= (2k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(i+n)}{(i-1)!(k+1-i)!}\left(\sum_{j=1}^{k} j(-1)^{j+k+1}\frac{\prod_{n=0}^{k-1}(j+n)}{(j-1)!(k+1-j)!}\right)$$

$$= (2k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(i+n)}{(i-1)!(k+1-i)!}\left(\sum_{j=0}^{k} j(-1)^{j+k}\frac{\prod_{n=0}^{k-1}(j+1+n)}{j!(k-j)!}\right)$$

$$= (2k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(i+n)}{(i-1)!(k+1-i)!}\left(\sum_{j=0}^{k} j(-1)^{j+k}\binom{2j}{j}\binom{k+j}{2j}\right)$$

$$= (2k+1)(k^2+k+1)(-1)^{k+1+i}\frac{\prod_{n=0}^{k-1}(i+n)}{(i-1)!(k-i)!}.$$

In the final step, Lemma B.0.2 was made use of. □

As an alternative proof, Theorem B.0.1 can be shown directly by making use of B.0.3 in addition to the identity $\sum_{i,j=1}^{k}\left(H_k^{-1}\right)_{ij} = n^2$ (2.14).

$$\begin{aligned}
\sum_{i,j=1}^{k} i\left(H_k^{-1}\right)_{ij} &= \sum_{j=1}^{k}\left(\sum_{i=1}^{k} i\left(H_k^{-1}\right)_{ij}\right)\\
&= \sum_{j=1}^{k}\left((k^2-j+1)\sum_{i=1}^{k}\left(H_k^{-1}\right)_{ij}\right)\\
&= (k^2+1)\sum_{i,j=1}^{k}\left(H_k^{-1}\right)_{ij} - \sum_{i,j=1}^{k} j\left(H_k^{-1}\right)_{ij}\\
&= (k^2+1)k^2 - \sum_{i,j=1}^{k} i\left(H_k^{-1}\right)_{ij}\\
&\Rightarrow \sum_{i,j=1}^{k} i\left(H_k^{-1}\right)_{ij} = \frac{k^2(k^2+1)}{2}
\end{aligned}$$

Here, the symmetry of H_k is made use of.

Remark B.0.1. *There are many such summation formulas which can be shown for Hilbert matrices using this same technique, many of them are quite interesting. Specifically, $\sum_{i,j=1}^{k} i^p\left(H_k^{-1}\right)_{ij}$ can be computed for any positive integer p. The difficulty for larger p is seen by looking at Lemma B.0.3. The difference seen will be p differentiations instead of only 2. This has been considered in general, but difficult computations arise. I feel that there is an elegant proof of such a general result, though I have not had the time or motivation to consider this. In addition, just as Theorem B.0.3 can be used to show Theorem B.0.1 directly, a similar result established through Lemma B.0.3 could be used to show Theorem B.0.2.*

APPENDIX C. SOLUTION GRADIENT FOR HEAT EQUATION

Below, the motivation of numerical flux choice (2.6) is provided. The main result was stated in [26], but no proof was given. A proof is provided here.

Theorem C.0.4. *(Solution gradient for the heat equation) Consider the the following heat equation*

$$\begin{cases} u_t - u_{xx} = 0 \\ u(x,0) = g(x) \end{cases}$$

where initial data g is smooth aside from one discontinuity at $x = 0$. Then, we have the following.

$$\begin{aligned} u_x(0,t) &= \sum_{m=0}^{\infty} \frac{2^m t^m}{\sqrt{4\pi t}(2m-1)!!}[D^{2m}g] + \sum_{m=0}^{\infty} \frac{2^m t^m}{(2m)!!}\overline{D^{2m+1}g} \\ &= \frac{1}{\sqrt{4\pi t}}[g] + \overline{\partial_x g} + \sqrt{\frac{t}{\pi}}[\partial_x^2 g] + t\overline{\partial_x^3 g} + \cdots \end{aligned}$$

Here, all jumps and averages of g are evaluated at $x = 0$ and the double factorial notation is defined as follows.

$$n!! = \begin{cases} n \cdot (n-2) \ldots 3 \cdot 1, & n > 0, \ \textit{odd} \\ n \cdot (n-2) \ldots 4 \cdot 2, & n > 0, \ \textit{even} \\ 0, & n = -1, 0 \end{cases}$$

Proof. Note first that the solution to the above heat equation is well known [18] and given as follows.

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/(4t)} g(y)\, dy$$

Then, because g is smooth on $(-\infty, 0)$ and $(0, \infty)$, then it has a convergent power series on each interval as given below.

$$g(y) = \sum_{i=0}^{\infty} \frac{1}{i!} D^i g(0 \pm \epsilon)(y \pm \epsilon)^i$$

on $(0, \pm\infty)$ for some $\epsilon > 0$. Then, taking limits as $\epsilon \to 0^+$, we have that

$$g(y) = \sum_{i=0}^{\infty} \frac{1}{i!} D^i g(0^{\pm}) y^i$$

on $(0, \pm\infty)$. Then, using the form of solution $u(x,t)$ as given above, we have the following.

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \left(\int_0^{\infty} e^{-(x-y)^2/(4t)} g(y)\ dy + \int_{-\infty}^{0} e^{-(x-y)^2/(4t)} g(y)\ dy \right)$$

This implies that

$$u_x(0,t) = \frac{1}{\sqrt{4\pi t}} \left(\int_0^{\infty} \frac{y}{2t} e^{-y^2/(4t)} g(y)\ dy + \int_{-\infty}^{0} \frac{y}{2t} e^{-y^2/(4t)} g(y)\ dy \right) = \frac{1}{\sqrt{4\pi t}} \left(I_1 + I_2 \right),$$

where,

$$\begin{aligned}
I_1 &= \int_0^{\infty} \frac{y}{2t} e^{-y^2/(4t)} g(y)\ dy \\
&= \int_0^{\infty} \frac{y}{2t} e^{-y^2/(4t)} \left(\sum_{i=0}^{\infty} \frac{1}{i!} D^i g(0^+) y^i \right)\ dy \\
&= \sum_{i=0}^{\infty} D^i g(0^+) \int_0^{\infty} \frac{y^{i+1}}{2ti!} e^{-y^2/(4t)}\ dy \\
&= \sum_{m=0}^{\infty} D^{2m} g(0^+) \int_0^{\infty} \frac{y^{2m+1}}{2t(2m)!} e^{-y^2/(4t)}\ dy + \sum_{m=0}^{\infty} D^{2m+1} g(0^+) \int_0^{\infty} \frac{y^{2m+2}}{2t(2m+1)!} e^{-y^2/(4t)}\ dy \\
&= \sum_{m=0}^{\infty} \frac{2^m}{(2m-1)!!} t^m D^{2m} g(0^+) + \sum_{m=0}^{\infty} \frac{2^m}{(2m)!!} t^m \sqrt{4\pi t} D^{2m+1} g(0^+).
\end{aligned}$$

Here, the result from Lemma C.0.4 was made use of. Similarly, we have the following calculation

for I_2.

$$\begin{aligned}
I_2 &= \int_{-\infty}^{0} \frac{y}{2t} e^{-y^2/(4t)} g(y)\ dy \\
&= \int_{-\infty}^{0} \frac{y}{2t} e^{-y^2/(4t)} \left(\sum_{i=0}^{\infty} \frac{1}{i!} D^i g(0^-) y^i \right)\ dy \\
&= \sum_{i=0}^{\infty} D^i g(0^-) \int_{-\infty}^{0} \frac{y^{i+1}}{2ti!} e^{-y^2/(4t)}\ dy \\
&= \sum_{i=0}^{\infty} D^i g(0^-) \int_{0}^{\infty} \frac{(-y)^{i+1}}{2ti!} e^{-y^2/(4t)}\ dy \\
&= \sum_{i=0}^{\infty} (-1)^{i+1} D^i g(0^-) \int_{0}^{\infty} \frac{y^{i+1}}{2ti!} e^{-y^2/(4t)}\ dy \\
&= -\sum_{m=0}^{\infty} \frac{2^m}{(2m-1)!!} t^m D^{2m} g(0^-) + \sum_{m=0}^{\infty} \frac{2^m}{(2m)!!} t^m \sqrt{4\pi t} D^{2m+1} g(0^-)
\end{aligned}$$

Therefore,

$$\begin{aligned}
u_x(0,t) &= \frac{1}{\sqrt{4\pi t}} (I_1 + I_2) \\
&= \sum_{m=0}^{\infty} \frac{2^m t^m}{\sqrt{4\pi t}(2m-1)!!} [D^{2m} g] + \sum_{m=0}^{\infty} \frac{2^m t^m}{(2m)!!} \overline{D^{2m+1} g}.
\end{aligned}$$

□

To obtain this result, the following Lemma is required.

Lemma C.0.4.

$$\int_{0}^{\infty} \frac{y^{i+1}}{2ti!} e^{-y^2/(4t)}\ dy = \begin{cases} \frac{2^m}{(2m-1)!!} t^m, & i = 2m, \quad m \geq 0 \\ \frac{2^m}{(2m)!!} t^m \sqrt{4\pi t}, & i = 2m+1, \quad m \geq 0 \end{cases}$$

Proof. Proceed by means of induction. Start with the following two bases cases. For $i = 0$, we have the following.

$$\int_{0}^{\infty} \frac{y}{2t} e^{-y^2/(4t)}\ dy = -e^{-y^2/4t} \Big|_0^{\infty} = -0 + 1 = 1$$

For $i = 1$, we have the following.

$$\begin{aligned}
\int_0^\infty \frac{y^2}{2t} e^{-y^2/(4t)}\ dy &= \int_0^\infty y\partial_y \left(-e^{-y^2/(4t)}\right)\ dy \\
&= -ye^{-y^2/(4t)}\Big|_0^\infty + \int_0^\infty e^{-y^2/(4t)}\ dy \\
&= -0 + 0 + \frac{1}{2}\sqrt{\frac{\pi}{1/(4t)}} \\
&= \sqrt{\pi t}
\end{aligned}$$

Then, assume the result holds for $i = 2m$ and $i = 2m+1$ for some $m \geq 0$. Then, for $i = 2(m+1)$, we have the following.

$$\begin{aligned}
\int_0^\infty \frac{y^{i+1}}{2ti} e^{-y^2/(4t)}\ dy &= \int_0^\infty \frac{y^i}{i!}\partial_y \left(-e^{-y^2/(4t)}\right)\ dy \\
&= -\frac{y^i}{i!}e^{-y^2/(4t)}\Big|_0^\infty + \int_0^\infty \frac{y^{i-1}}{(i-1)!}e^{-y^2/(4t)}\ dy \\
&= -0 + 0 + \int_0^\infty \frac{y^{2m+1}}{(2m+1)!}e^{-y^2/(4t)}\ dy \\
&= \frac{2t}{2m+1}\int_0^\infty \frac{y^{2m+1}}{2t(2m)!}e^{-y^2/(4t)}\ dy \\
&= \frac{2t}{2m+1}\frac{2^m}{(2m-1)!!}t^m \\
&= \frac{2^{m+1}}{(2m+1)!!}t^{m+1}
\end{aligned}$$

For $i = 2(m+1)+1$, we have the following.

$$\begin{aligned}
\int_0^\infty \frac{y^{i+1}}{2ti} e^{-y^2/(4t)}\ dy &= \int_0^\infty \frac{y^{i-1}}{(i-1)!}e^{-y^2/(4t)}\ dy \\
&= \int_0^\infty \frac{y^{(2m+1)+1}}{(2m+2)!}e^{-y^2/(4t)}\ dy \\
&= \frac{2t}{2m+2}\int_0^\infty \frac{y^{(2m+1)+1}}{2t(2m+1)!}e^{-y^2/(4t)}\ dy \\
&= \frac{2t}{2m+2}\frac{2^m}{(2m)!!}t^m\sqrt{4\pi t} \\
&= \frac{2^{m+1}}{(2(m+1))!!}t^{m+1}\sqrt{4\pi t}
\end{aligned}$$

Thus ends the proof. □

BIBLIOGRAPHY

[1] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[2] Ivo Babuška. The finite element method with penalty. *Math. Comp.*, 27:221–228, 1973.

[3] G. A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.*, 31:45–59, 1977.

[4] F. Bassi and S. Rebay. GMRES discontinuous Galerkin solution of the compressible Navier-Stokes equations. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 197–208. Springer, Berlin, 2000.

[5] Y. Cheng and C.-W. Shu. A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives. *Math. Comp.*, 77(262):699–730, 2008.

[6] Philippe G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[7] B. Cockburn, C. Johnson, C.-W. Shu, and E. Tadmor. *Advanced numerical approximation of nonlinear hyperbolic equations*, volume 1697 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998. Papers from the C.I.M.E. Summer School held in Cetraro, June 23–28, 1997, Edited by Alfio Quarteroni, Fondazione C.I.M.E.. [C.I.M.E. Foundation].

[8] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. The development of discontinuous Galerkin methods. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 3–50. Springer, Berlin, 2000.

[9] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463 (electronic), 1998.

[10] B. Cockburn and C.-W. Shu. Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16(3):173–261, 2001.

[11] Bernardo Cockburn, Suchung Hou, and Chi-Wang Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comp.*, 54(190):545–581, 1990.

[12] Bernardo Cockburn, San Yih Lin, and Chi-Wang Shu. 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.

[13] Bernardo Cockburn and Chi-Wang Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comp.*, 52(186):411–435, 1989.

[14] Bernardo Cockburn and Chi-Wang Shu. The Runge-Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 25(3):337–361, 1991.

[15] Bernardo Cockburn and Chi-Wang Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. *J. Comput. Phys.*, 141(2):199–224, 1998.

[16] C. Dawson, S. Sun, and M. F. Wheeler. Compatible algorithms for coupled flow and transport. *Comput. Methods Appl. Mech. Engrg.*, 193(23-26):2565–2580, 2004.

[17] Yekaterina Epshteyn and Béatrice Rivière. Estimation of penalty parameters for symmetric interior penalty Galerkin methods. *J. Comput. Appl. Math.*, 206(2):843–872, 2007.

[18] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[19] X. Feng and H. Wu. Discontinuous Galerkin methods for the Helmholtz equation with large wave number. *SIAM J. Numer. Anal.*, 47(4):2872–2896, 2009.

[20] G. Gassner, F. Lörcher, and C. D. Munz. A contribution to the construction of diffusion fluxes for finite volume and discontinuous Galerkin schemes. *J. Comput. Phys.*, 224(2):1049–1063, 2007.

[21] Gregor Gassner, Frieder Lörcher, and Claus-Dieter Munz. A numerical diffusion flux for finite volume and discontinuous Galerkin schemes based on the diffusive Riemann problem. In *Computational fluid dynamics review 2010*, pages 91–101. World Sci. Publ., Hackensack, NJ, 2010.

[22] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[23] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comp.*, 46(173):1–26, 1986.

[24] P. Lasaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation. In *Mathematical aspects of finite elements in partial differential equations (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974)*, pages 89–123. Publication No. 33. Math. Res. Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1974.

[25] Randall J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[26] H. Liu and J. Yan. The direct discontinuous Galerkin (ddg) methods for diffusion problems. *SIAM J. Numer. Anal.*, 47(1):475–698, 2009.

[27] H. Liu and J. Yan. The direct discontinuous Galerkin (DDG) method for diffusion with interface corrections. *Commun. Comput. Phys.*, 8(3):541–564, 2010.

[28] Frieder Lörcher, Gregor Gassner, and Claus-Dieter Munz. An explicit discontinuous Galerkin scheme with local time-stepping for general unsteady diffusion equations. *J. Comput. Phys.*, 227(11):5649–5670, 2008.

[29] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971. Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[30] Todd E. Peterson. A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.*, 28(1):133–140, 1991.

[31] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[32] Gerard R. Richter. The discontinuous Galerkin method with diffusion. *Math. Comp.*, 58(198):631–643, 1992.

[33] B. Rivière, M. F. Wheeler, and V. Girault. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM J. Numer. Anal.*, 39(3):902–931 (electronic), 2001.

[34] Richard Savage and Eugene Lukacs. Tables of inverses of finite segments of the Hilbert matrix. In *Contributions to the solution of systems of linear equations and the determination of eigenvalues*, National Bureau of Standards Applied Mathematics Series No. 39, pages 105–108. U. S. Government Printing Office, Washington, D. C., 1954.

[35] C.-W. Shu. Different formulations of the discontinuous galerkin method for the viscous terms. *Advances in Scientific Computing, Z.-C. Shi, M. Mu, W. Xue and J. Zou, editors,Science Press, China*, pages 144–155, 2001.

[36] C.-W. Shu and S. Osher. Efficient implementation of essentially nonoscillatory shock-capturing schemes. *J. Comput. Phys.*, 77(2):439–471, 1988.

[37] C.-W. Shu and S. Osher. Efficient implementation of essentially nonoscillatory shock-capturing schemes. II. *J. Comput. Phys.*, 83(1):32–78, 1989.

[38] Chi-Wang Shu. Discontinuous Galerkin methods: general approach and stability. In *Numerical solutions of partial differential equations*, Adv. Courses Math. CRM Barcelona, pages 149–201. Birkhäuser, Basel, 2009.

[39] Richard B. Smith. Two theorems on inverses of finite segments of the generalized Hilbert matrix. *Math. Tables Aids Comput.*, 13:41–43, 1959.

[40] Shuyu Sun. *Discontinuous Galerkin methods for reactive transport in porous media*. ProQuest LLC, Ann Arbor, MI, 2003. Thesis (Ph.D.)–The University of Texas at Austin.

[41] B. van Leer and S. Nomura. Discontinuous Galerkin for diffusion. *Proceedings of 17th AIAA Computational Fluid Dynamics Conference (June 6 2005)*, AIAA-2005-5108.

[42] Bram van Leer, Marcus Lo, Rita Gitik, and Shohei Nomura. A venerable family of discontinuous Galerkin schemes for diffusion revisited. In *Adaptive high-order methods in computational fluid dynamics*, volume 2 of *Adv. Comput. Fluid Dyn.*, pages 185–201. World Sci. Publ., Hackensack, NJ, 2011.

[43] Marc van Raalte and Bram van Leer. Bilinear forms for the recovery-based discontinuous Galerkin method for diffusion. *Commun. Comput. Phys.*, 5(2-4):683–693, 2009.

[44] C. Vidden and J. Yan. Direct discontinuous Galerkin method for diffusion equations with symmetric structure. *SIAM J. Numer. Anal*, under review.

[45] T. Warburton and J. S. Hesthaven. On the constants in *hp*-finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.*, 192(25):2765–2773, 2003.

[46] M. F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15:152–161, 1978.

[47] Herbert S. Wilf. *generatingfunctionology*. A K Peters Ltd., Wellesley, MA, third edition, 2006.

[48] Mengping Zhang and Jue Yan. Fourier type error analysis of the direct discontinuous galerkin method and its variations for diffusion equations. *Journal of Scientific Computing*, pages 1–18. 10.1007/s10915-011-9564-5.

[49] Xiangxiong Zhang and Chi-Wang Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comput. Phys.*, 229(9):3091–3120, 2010.