2016

Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations

Mahboub Baccouch
University of Nebraska at Omaha, mbaccouch@unomaha.edu

Applied Numerical Mathematics ••• (••••) •••–•••

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations

Mahboub Baccouch

Department of Mathematics, University of Nebraska at Omaha, Omaha, NE 68182, United States

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form 3 February 2016
Accepted 31 March 2016
Available online xxxx

I would like to dedicate this work to my Father, Ahmed Baccouch, who unfortunately passed away during the completion of this work

Keywords:
Discontinuous Galerkin method
Nonlinear ordinary differential equations
Superconvergence
A posteriori error estimation
Adaptive mesh refinement

ABSTRACT

We develop and analyze a new residual-based *a posteriori* error estimator for the discontinuous Galerkin (DG) method for nonlinear ordinary differential equations (ODEs). The *a posteriori* DG error estimator under investigation is computationally simple, efficient, and asymptotically exact. It is obtained by solving a local residual problem with no boundary condition on each element. We first prove that the DG solution exhibits an optimal $\mathcal{O}(h^{p+1})$ convergence rate in the L^2-norm when p-degree piecewise polynomials with $p \geq 1$ are used. We further prove that the DG solution is $\mathcal{O}(h^{2p+1})$ superconvergent at the downwind points. We use these results to prove that the p-degree DG solution is $\mathcal{O}(h^{p+2})$ super close to a particular projection of the exact solution. This superconvergence result allows us to show that the true error can be divided into a significant part and a less significant part. The significant part of the discretization error for the DG solution is proportional to the $(p+1)$-degree right Radau polynomial and the less significant part converges at $\mathcal{O}(h^{p+2})$ rate in the L^2-norm. Numerical experiments demonstrate that the theoretical rates are optimal. Based on the global superconvergent approximations, we construct asymptotically exact *a posteriori* error estimates and prove that they converge to the true errors in the L^2-norm under mesh refinement. The order of convergence is proved to be $p+2$. Finally, we prove that the global effectivity index in the L^2-norm converges to unity at $\mathcal{O}(h)$ rate. Several numerical examples are provided to illustrate the global superconvergence results and the convergence of the proposed estimator under mesh refinement. A local adaptive procedure that makes use of our local *a posteriori* error estimate is also presented.

1. Introduction

In this paper, we investigate the superconvergence properties and analyze a residual-based *a posteriori* error estimator of the discretization errors for the discontinuous Galerkin (DG) method applied to the following first-order initial-value problem (IVP) of nonlinear ordinary differential equation (ODE) on $[0, T]$:

$$\frac{d\vec{u}}{dt} = \vec{f}(t, \vec{u}), \quad t \in [0, T], \quad \vec{u}(0) = \vec{u}_0, \tag{1.1}$$

where $\vec{u} : [0, T] \to \mathbb{R}^n$, $\vec{u}_0 \in \mathbb{R}^n$, and $\vec{f} : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$. In our analysis, we assume that the solution exists and is unique. According to the ODE theory, the condition $\vec{f} \in C^1([0, T] \times \mathbb{R}^n)$ is sufficient to guarantee the existence and uniqueness of the

E-mail address: mbaccouch@unomaha.edu.

http://dx.doi.org/10.1016/j.apnum.2016.03.008
0168-9274/

solution to (1.1). We note that an IVP for higher-order ODE may be solved using the DG method for solving first-order system of the form (1.1) since one can transform higher-order equations into several coupled first-order equations by introducing new unknowns.

ODEs solvers are important tools for the computational solutions of higher-order ODEs and many partial differential equations (PDEs). For instance, the application of the standard finite element method or DG in space to solve time-dependent PDEs generates a coupled system of ODEs. Once the spatial discretization is constructed, one would then need to employ a suitable ODE solver for the time discretization. If the mesh size in space becomes small then often the ODE system becomes more and more stiff. It is very well-known that explicit schemes suffer from extremely small time step restriction for stability. Therefore, explicit time discretization techniques are not a suitable choice and implicit, at least A-stable methods are desirable. There are many A-stable higher-order time discretization schemes designed for different purposes in the literature, such as the implicit Runge–Kutta (IRK) methods and DG methods with higher polynomial order. Despite the popularity of high-order IRK methods for integrating systems of ODEs, we choose the DG method because of the following advantages: (i) it is A-stable, (ii) it is locally conservative, and (iii) it exhibits strong superconvergence that can be used to estimate the discretization error. Furthermore, it is very natural to construct high-order DG methods and, for future developments, we can apply the well-known adaptive DG techniques for changing the polynomial degree as well as the length of the time intervals. Finally, if we use space–time DG methods for the discretization of evolution PDEs, we would have a uniform variational approach in space and time which may provide a useful tool for the future analysis of the fully discrete problem and the construction of simultaneous space–time adaptive methods.

The DG method considered here is a class of finite element methods using completely discontinuous piecewise polynomials for the numerical solution and the test functions. DG method combines many attractive features of the classical finite element and finite volume methods. It is a powerful tool for approximating some differential equations which model problems in physics, especially in fluid dynamics or electrodynamics. Comparing with the standard finite element method, the DG method has a compact formulation, *i.e.*, the solution within each element is weakly connected to neighboring elements. DG method was initially introduced by Reed and Hill in 1973 as a technique to solve neutron transport problems [36]. In 1974, Lesaint and Raviart [33] presented the first numerical analysis of the method for a linear advection equation. Since then, DG methods have been used to solve ODEs [6,19,32,33], hyperbolic [15–18,23,24,35,29,31,20,45,34,2,3,12,5] and diffusion and convection–diffusion [13,14,43,25] PDEs. The proceedings of Cockburn et al. [22] and Shu [39] contain a more complete and current survey of the DG method and its applications.

In recent years, the study of superconvergence and *a posteriori error* estimates of DG methods has been an active research field in numerical analysis, see the monographs by Verfürth [41], Wahlbin [42], and Babuška and Strouboulis [9]. A knowledge of superconvergence properties can be used to (*i*) construct simple and asymptotically exact *a posteriori* estimates of discretization errors and (*ii*) help detect discontinuities to find elements needing limiting, stabilization and/or refinement. *A posteriori* error estimates play an essential role in assessing the reliability of numerical solutions and in developing efficient adaptive algorithms. Typically, *a posteriori* error estimators employ the known numerical solution to derive estimates of the actual solution errors. They are also used to steer adaptive schemes where either the mesh is locally refined (h-refinement) or the polynomial degree is raised (p-refinement). For an introduction to the subject of *a posteriori* error estimation see the monograph of Ainsworth and Oden [8]. Superconvergence properties for finite element and DG methods for ordinary differential equations have been studied in [6,7,27,33,44,40]. The first superconvergence result for standard DG solutions of hyperbolic PDEs appeared in Adjerid et al. [6]. The authors presented numerical results that show that standard DG solutions of one-dimensional linear and nonlinear hyperbolic problems using p-degree polynomial approximations exhibit an $\mathcal{O}(h^{p+2})$ superconvergence rate at the roots of $(p+1)$-degree Radau polynomial. They further established a strong $\mathcal{O}(h^{2p+1})$ superconvergence at the downwind end of every element.

Related theoretical results in the literature including superconvergence results and error estimates of the DG methods for ODEs are given in [33,27,26,32,30]. In particular, Lesaint and Raviart [33] studied the numerical solution of the initial value problem (1.1) by a DG method. Their scheme is equivalent to some implicit Runge–Kutta method, strongly A-stable one-step method. Delfour et al. [27] analyzed a class of Galerkin methods derived from discontinuous piecewise polynomial spaces. These schemes generalize the method proposed by Lesaint and Raviart [33]. In their DG method, the approximated solution at t_j, a point of discontinuity in the approximating polynomial u_h, is taken as an average across the jump: $\alpha_j u_h(t_j^-) + (1-\alpha_j)u_h(t_j^+)$. The cases $\alpha_j = 0, 0.5, 1$ correspond, respectively, to Euler's explicit, improved, and implicit schemes. Later, Delfour and Dubeau [26] studied the approximation of the solution of the nonlinear ODEs by discontinuous piecewise polynomials. They introduced a more general theory of one-step (such as implicit Runge–Kutta and Crank–Nicholson schemes), hybrid and multistep methods (such as Adams–Bashforth and Adams–Moulton schemes). Also, we mention the work of Johnson [32] in which *a priori* error estimates for a class of implicit one-step methods generated by the DG time discretization are proven. Estep [30] analyzed a finite element method for the integration of IVPs in ODEs. The author obtained quasi-optimal *a priori* and *a posterior* error bounds. They used these results to construct a rigorous and robust theory of global error control. The author also derived an asymptotic error estimate that is used in a discussion of the behavior of the error. Recently, Deng and Xiong [28] introduced and analyzed a DG finite element method with interpolated coefficients for an IVP of nonlinear ODE. They used the finite element projection for an auxiliary linear problem as comparison function and proved an optimal superconvergence results. Subsequently, Adjerid and Baccouch [4,11,10] investigated the global convergence of the implicit residual-based *a posteriori* error estimates of Adjerid et al. [6]. They proved that, for smooth solutions, these *a posteriori* error estimates at a fixed time t converge to the true spatial errors in the L^2-norm under

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

mesh refinement. More recently, the author [10] presented and analyzed new *a posteriori* error estimates for a DG formulation applied to nonlinear scalar conservation laws in one space dimension. We used the superconvergence result of Meng et al. [34] to prove that the DG discretization error estimates converge to the true spatial errors under mesh refinement at $\mathcal{O}(h^{p+3/2})$ rate. Finally, we proved that the global effectivity index in the L^2-norm converges to unity at $\mathcal{O}(h^{1/2})$ rate. Let us emphasize that these convergence rates are not optimal since the order of superconvergence proved in [34] is not optimal. Our computational results indicated that the observed numerical convergence rates are higher than the theoretical rates. However, there is no theoretical justification of these results so far.

In this work, we construct and analyze an implicit residual-based *a posteriori* error estimate of the discretization errors for the DG method for first-order initial-value problems. We prove that the p-degree DG solution exhibits an optimal $\mathcal{O}(h^{p+1})$ convergence rate in the L^2-norm for general nonlinear ODEs with a Lipschitz-continuous right-hand side. We further prove that the p-degree DG solution is $\mathcal{O}(h^{2p+1})$ superconvergent at the end of each step. We use these results to prove that the p-degree DG solution is $\mathcal{O}(h^{p+2})$ super close to a particular projection of the exact solution. Our numerical experiments demonstrate optimal rates of convergence. Superconvergence towards the particular projection of the exact solution allows us to show that the leading error term on each element is proportional to a $(p+1)$-degree right Radau polynomial. We use the superconvergence at Radau points to construct asymptotically exact *a posteriori* error estimates and prove that, for smooth solutions, these *a posteriori* DG error estimates converge to the true errors at $\mathcal{O}(h^{p+2})$ rate in the L^2-norm. Finally, we prove that the global effectivity indices in the L^2-norm converge to unity at $\mathcal{O}(h)$ rate. Our proofs are valid for arbitrary regular meshes and for P^p polynomials with $p \geq 1$. To the best knowledge of the author, this is the first proof of the *a posteriori* error estimates for the nonlinear ODEs. In [4], we used similar techniques to prove the global convergence of the residual-based *a posteriori* error estimates in the L^2-norm for the DG method applied to one-dimensional linear convection problems.

This paper is organized as follows: In section 2, we present the discrete DG method for solving the nonlinear ODE and we introduce some notation and definitions. We also present a few preliminary results which will be needed in our *a posteriori* error analysis. In section 3, we present the DG error analysis and prove our main superconvergence results. In section 4, we present the *a posteriori* error estimation procedure and prove that these error estimates converge to the true errors under mesh refinement in L^2-norm. In section 5, we present several numerical examples to validate our theoretical results. We also propose an adaptive algorithm based on the local *a posteriori* error estimates. We conclude and discuss our results in section 6.

2. The DG scheme

The error analysis of nonlinear scalar and vector IVPs having smooth solutions is similar. For this, we restrict our theoretical discussion to the scalar case; extension to systems follows the same lines. Without loss of generality, we consider the following nonlinear scalar IVP:

$$u' = f(t,u), \quad t \in [0,T], \quad u(0) = u_0, \tag{2.1}$$

where $f(t,u) : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $u_0 \in \mathbb{R}$ are given. In our analysis, we assume that $f(t,u)$ is sufficiently smooth with respect to the variables t and u. In particular, we always assume that $|f_u(t,u)| \leq M_1$ on the set $D = [0,T] \times \mathbb{R} \subset \mathbb{R}^2$, where M_1 is a positive constant. We note that the assumption $|f_u(t,u)| \leq M_1$ is sufficient to guarantee that the function $f(t,u)$ satisfies a Lipschitz condition in the variable u on the set D with Lipschitz constant M_1, *i.e.*,

$$|f(t,u) - f(t,v)| \leq M_1|u-v|, \quad \text{whenever } (t,u) \text{ and } (t,v) \in D. \tag{2.2}$$

First, we present the discrete DG scheme for the problem (2.1). We adopt the usual notation of the DG method. We divide the interval $\Omega = [0,T]$ into N subintervals $I_j = [t_{j-1}, t_j]$, $j = 1, \ldots, N$, where $0 = t_0 < t_1 < \cdots < t_N = T$. We denote the length of I_j by $h_j = t_j - t_{j-1}$. We also denote $h = \max\limits_{1 \leq j \leq N} h_j$ and $h_{min} = \min\limits_{1 \leq j \leq N} h_j$ as the length of the largest and smallest subinterval, respectively. Here, we consider regular meshes, that is $\frac{h}{h_{min}} \leq \lambda$, where $\lambda \geq 1$ is a constant (independent of h) during mesh refinement. If $\lambda = 1$, then the mesh is uniformly distributed. In this case, the nodes and mesh size are defined by $t_j = jh$, $j = 0, 1, \ldots, N$, $h = T/N$.

Throughout this paper, we define $v(t_j^-)$ and $v(t_j^+)$ to be the left limit and the right limit of the function v at the discontinuity point t_j, *i.e.*, $v(t_j^-) = \lim\limits_{s \to 0^-} v(t_j + s)$ and $v(t_j^+) = \lim\limits_{s \to 0^+} v(t_j + s)$. We also use $[v](t_j)$ to denote the jump of v at t_j, *i.e.*, $[v](t_j) = v(t_j^+) - v(t_j^-)$.

The weak DG formulation is obtained by multiplying (2.1) by a smooth test function v and integrating over an arbitrary element I_j. After integrating by parts we obtain the following weak formulation:

$$\int_{I_j} v' u dt + \int_{I_j} f(t,u) v dt - u(t_j) v(t_j) + u(t_{j-1}) v(t_{j-1}) = 0. \tag{2.3}$$

We define the piecewise-polynomial space V_h^p as the space of polynomials of degree at most p in each subinterval I_j, *i.e.*,

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

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

where $P^p(I_j)$ is the space of polynomials of degree at most p on I_j. Note that polynomials in the space V_h^p are allowed to have discontinuities across element boundaries.

Next, we approximate the exact solution $u(t)$ by a piecewise polynomial $u_h(t) \in V_h^p$. We note that u_h is not necessarily continuous at the endpoints of I_j. The discrete formulation consists of finding $u_h \in V_h^p$ such that: $\forall\ v \in V_h^p$ and $j = 1, \ldots, N$,

$$\int_{I_j} v' u_h dt + \int_{I_j} f(t, u_h) v dt - \hat{u}_h(t_j) v(t_j^-) + \hat{u}_h(t_{j-1}) v(t_{j-1}^+) = 0, \tag{2.4a}$$

where the numerical flux $\hat{u}_h(t_j)$ is the discrete approximation to the trace of u at the node $t = t_j$.

In order to complete the definition of the discrete DG method we need to select $\hat{u}_h$ on the boundaries of I_j. In this paper, we take the classical upwind numerical flux:

$$\hat{u}_h(t_0) = u_0, \quad \hat{u}_h(t_j) = u_h(t_j^-), \quad j = 1, \ldots, N. \tag{2.4b}$$

Implementation: We note that the approximate solution $u_h(t)$ can be efficiently computed in an element-by-element fashion. Specifically, we can obtain the DG solution $u_h(t)$ in the following ordering: first we can easily obtain $u_h(t)$ in I_1 using (2.4) with $j = 1$ since $u_h(t_0^-) = u_0$ is given. We then obtain $u_h(t)$ in I_2 since $u_h(t)$ in I_1 is already available. This process can be repeated to obtain $u_h(t)$ in $I_3, \ldots, I_N$. In practice, $u_h(t)$ can be computed locally for each I_j as follows: Expressing $u_h(t)$ as a linear combination of orthogonal basis $L_{i,j}(t)$, $i = 0, \ldots, p$, where $L_{i,j}$ is the ith-degree Legendre polynomial on I_j, *i.e.*, $u_h(t) = \sum_{i=0}^p c_{i,j} L_{i,j}(t)$, $t \in I_j$, where $\{L_{i,j}(t)\}_{i=0}^{i=p}$ is a local basis of $P^p(I_j)$, and choosing the test functions $v = L_{k,j}(t)$, $k = 0, \ldots, p$, we obtain a small system of nonlinear algebraic equations on each I_j. The resulting system can be solved for the unknown coefficients $c_{0,j}, \ldots, c_{p,j}$ using *e.g.*, the classical Newton–Raphson method. Once we compute the DG solution on each element I_j, $j = 1, \ldots, N$, we obtain an approximate solution (piecewise discontinuous polynomial functions of degree $\leq p$) to our original IVP (2.1). For complete details of the DG methods for ODEs as well as their properties, we refer the readers to [27,32,33].

Notation and definitions: We begin by defining some norms that will be used throughout the paper. We define the L^2 inner product of two integrable functions, u and v, on the interval $I_j = [t_{j-1}, t_j]$ as $(u, v)_{I_j} = \int_{I_j} u(t) v(t) dt$. Denote $\|u\|_{0,I_j} = (u, u)_{I_j}^{1/2}$ to be the standard L^2-norm of u on I_j. Moreover, the standard L^∞-norm of u on I_j is defined by $\|u\|_{\infty,I_j} = \sup_{t \in I_j} |u(t)|$. Let $H^s(I_j)$, where $s = 0, 1, \ldots$, denote the standard Sobolev space of square integrable functions on I_j with all derivatives $u^{(k)}$, $k = 0, 1, \ldots, s$ being square integrable on I_j, *i.e.*,

$$H^s(I_j) = \left\{ u \ :\ \int_{I_j} |u^{(k)}(t)|^2 dt < \infty,\ 0 \leq k \leq s \right\},$$

and equipped with the norm

$$\|u\|_{s,I_j} = \left(\sum_{k=0}^{s} \left\| u^{(k)} \right\|_{0,I_j}^2 \right)^{1/2}.$$

The $H^s(I_j)$-seminorm of a function u on I_j is given by $|u|_{s,I_j} = \left\| u^{(s)} \right\|_{0,I_j}$.

We also define the norms on the whole computational domain Ω as follows:

$$\|u\|_{0,\Omega} = \left(\sum_{j=1}^{N} \|u\|_{0,I_j}^2 \right)^{1/2}, \quad \|u\|_{\infty,\Omega} = \max_{1 \leq j \leq N} \|u\|_{\infty,I_j}, \quad \|u\|_{s,\Omega} = \left(\sum_{j=1}^{N} \|u\|_{s,I_j}^2 \right)^{1/2}.$$

The seminorm on the whole computational domain Ω is defined as follows

$$|u|_{s,\Omega} = \left(\sum_{j=1}^{N} |u|_{s,I_j}^2 \right)^{1/2}.$$

We note that if $u \in H^s(\Omega)$, $s = 1, 2, \ldots$, the norm $\|u\|_{s,\Omega}$ on the whole computational domain is the standard Sobolev norm $\left(\sum_{k=0}^{s} \left\| u^{(k)} \right\|_{0,\Omega}^2 \right)^{1/2}$. For convenience, we use $\|u\|_{I_j}$ and $\|u\|$ to denote $\|u\|_{0,I_j}$ and $\|u\|_{0,\Omega}$, respectively.

In our analysis we need the pth-degree Legendre polynomial which can be defined by Rodrigues formula [1]

$$\tilde{L}_p(\xi) = \frac{1}{2^p p!} \frac{d^p}{d\xi^p} \left((\xi^2 - 1)^p \right), \quad -1 \leq \xi \leq 1.$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

The Legendre polynomial satisfies the following important properties:

$$\tilde{L}_p(1)=1, \quad \tilde{L}_p(-1)=(-1)^p, \tag{2.5a}$$

$$\int_{-1}^{1}\tilde{L}_p(\xi)\tilde{L}_q(\xi)d\xi=\frac{2}{2p+1}\delta_{pq}, \quad \text{where } \delta_{pq} \text{ is the Kronecker symbol.} \tag{2.5b}$$

We note that the $(p+1)$-degree Legendre polynomial on $[-1,1]$ can be written as

$$\tilde{L}_{p+1}(\xi)=\frac{(2p+2)!}{2^{p+1}((p+1)!)^2}\xi^{p+1}+\tilde{q}_p(\xi), \quad \text{where } \tilde{q}_p\in P^p([-1,1]).$$

Next, we define the $(p+1)$-degree right Radau polynomial on $[-1,1]$ as

$$\tilde{R}_{p+1}(\xi)=\tilde{L}_{p+1}(\xi)-\tilde{L}_p(\xi), \ -1\le\xi\le 1, \tag{2.6}$$

which has $p+1$ real distinct roots, $-1<\xi_0<\cdots<\xi_p=1$.

Mapping the physical element I_j into a reference element $[-1,1]$ by the standard affine mapping

$$t=\frac{t_j+t_{j-1}}{2}+\frac{h_j}{2}\xi, \tag{2.7}$$

we obtain the shifted Legendre and Radau polynomials on I_j:

$$L_{p+1,j}(t)=\tilde{L}_{p+1}\left(\frac{2t-t_j-t_{j-1}}{h_j}\right), \quad R_{p+1,j}(t)=\tilde{R}_{p+1}\left(\frac{2t-t_j-t_{j-1}}{h_j}\right).$$

Next, we define the monic Radau polynomial, $\psi_{p+1,j}(t)$, on I_j as

$$\psi_{p+1,j}(t)=\frac{h_j^{p+1}[(p+1)!]^2}{(2p+2)!}R_{p+1,j}(t)=c_p h_j^{p+1}R_{p+1,j}(t), \quad \text{where } c_p=\frac{((p+1)!)^2}{(2p+2)!}. \tag{2.8}$$

Throughout this paper the roots of $R_{p+1,j}(t)$ are denoted by $t_{j,i}=\frac{t_j+t_{j-1}}{2}+\frac{h_j}{2}\xi_i$, $i=0,1,\ldots,p$.

In the next lemma, we state and prove the following results which will be needed in our *a posteriori* error analysis.

Lemma 2.1. *The polynomials $L_{p,j}$ and $\psi_{p+1,j}$ satisfy the following properties*

$$\left\|L_{p,j}\right\|_{I_j}^2=\frac{h_j}{2p+1}, \tag{2.9a}$$

$$\int_{I_j}\psi'_{p+1,j}\psi_{p+1,j}dt=-k_1h_j^{2p+2}, \tag{2.9b}$$

$$\left\|\psi_{p+1,j}\right\|_{I_j}^2=(2p+2)k_2h_j^{2p+3}, \tag{2.9c}$$

where $k_1=2c_p^2$, $k_2=\frac{k_1}{(2p+1)(2p+3)}$, and $c_p=\frac{((p+1)!)^2}{(2p+2)!}$.

Proof. In order to prove (2.9a), we use the mapping (2.7) and the orthogonality relation (2.5b) to obtain

$$\left\|L_{p,j}\right\|_{I_j}^2=\int_{I_j}L_{p,j}^2(t)dt=\frac{h_j}{2}\int_{-1}^{1}\tilde{L}_p^2(\xi)d\xi=\frac{h_j}{2}\frac{2}{2p+1}=\frac{h_j}{2p+1}. \tag{2.10}$$

Next, we will prove (2.9b). By the property of the Legendre polynomial (2.5a), we have

$$\tilde{R}_{p+1}(1)=R_{p+1,j}(t_j)=\psi_{p+1,j}(t_j)=0, \quad \tilde{R}_{p+1}(-1)=R_{p+1,j}(t_{j-1})=2(-1)^{p+1}. \tag{2.11}$$

Therefore,

$$\begin{aligned}\int_{I_j}\psi'_{p+1,j}\psi_{p+1,j}dt&=\frac{1}{2}\psi_{p+1,j}^2(t_j)-\frac{1}{2}\psi_{p+1,j}^2(t_{j-1})=-\frac{1}{2}\psi_{p+1,j}^2(t_{j-1})\\&=-\frac{1}{2}\left(c_ph_j^{p+1}R_{p+1,j}(t_{j-1})\right)^2=-2c_p^2h_j^{2p+2}=-k_1h_j^{2p+2}, \quad \text{where } k_1=2c_p^2.\end{aligned}$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Using (2.8) and the orthogonality relation (2.5b), we write

$$\left\|\psi_{p+1,j}\right\|_{I_j}^2 = c_p^2 h_j^{2p+2}\int_{I_j} R_{p+1,j}^2(t)dt = \frac{c_p^2 h_j^{2p+3}}{2}\int_{-1}^{1}\tilde{R}_{p+1}^2(\xi)d\xi = \frac{c_p^2 h_j^{2p+3}}{2}\int_{-1}^{1}\left(\tilde{L}_{p+1}(\xi)-\tilde{L}_p(\xi)\right)^2 d\xi$$

$$= \frac{c_p^2 h_j^{2p+3}}{2}\int_{-1}^{1}(\tilde{L}_{p+1}^2(\xi)+\tilde{L}_p^2(\xi))d\xi = (2p+2)k_2 h_j^{2p+3}, \quad \text{where } k_2 = \frac{k_1}{(2p+1)(2p+3)}. \qquad \square$$

For $p \geq 1$, we consider two special projection operators, $P_h^{\pm}$, which are defined as follows: For any smooth function u, the restrictions of $P_h^+ u$ and $P_h^- u$ to I_j are polynomials in $P^p(I_j)$ satisfying

$$\int_{I_j}(P_h^- u - u)v dt = 0, \ \forall\, v \in P^{p-1}(I_j), \quad \text{and} \quad (P_h^- u - u)(t_j^-) = 0, \tag{2.12a}$$

$$\int_{I_j}(P_h^+ u - u)v dt = 0, \ \forall\, v \in P^{p-1}(I_j), \quad \text{and} \quad (P_h^+ u - u)(t_{j-1}^+) = 0. \tag{2.12b}$$

These special projections are used in the error estimates of the DG methods to derive optimal L^2 error bounds in the literature, *e.g.*, in [20]. They are mainly used to eliminate the jump terms at the element boundaries in the error estimates in order to prove the optimal L^2 error estimates.

In our analysis, we need the following well-known projection results. Their proofs can be found in [21]: For any $u \in H^{p+1}(I_j)$ with $j = 1, \ldots, N$, there exists a constant C independent of the mesh size h such that

$$\left\|u - P_h^{\pm} u\right\|_{I_j} \leq C h_j^{p+1}\,|u|_{p+1,I_j}, \quad \left\|(u - P_h^{\pm} u)'\right\|_{I_j} \leq C h_j^{p}\,|u|_{p,I_j}. \tag{2.13}$$

Moreover, we recall the inverse properties of the finite element space V_h^p that will be used in our error analysis: For any $v_h \in V_h^p$, there exists a positive constant C independent of v_h and h, such that, $\forall\ j = 1, \ldots, N$,

$$\left\|v_h^{(k)}\right\|_{I_j} \leq C h_j^{-k}||v_h||_{I_j}, \quad k \geq 1, \tag{2.14a}$$

$$\left|v_h(t_{j-1}^+)\right| + \left|v_h(t_j^-)\right| \leq C h_j^{-1/2}||v_h||_{I_j}. \tag{2.14b}$$

From now on, the notation C, C_1, C_2, etc. will be used to denote positive constants that are independent of the discretization parameters, but which may depend upon the exact smooth solution of the ODE (2.1) and its derivatives. Furthermore, all the constants will be generic, *i.e.*, they may represent different constant quantities in different occurrences.

Throughout this paper, $e = u - u_h$ will be used to the error between the exact solution of (2.1) and the numerical solutions defined in (2.4). Let the projection error be defined as $\epsilon = u - P_h^- u$ and the error between the numerical solution and the projection of the exact solution be defined as $\bar{e} = P_h^- u - u_h$. We note that the true error can be split as $e = (u - P_h^- u) + (P_h^- u - u_h) = \epsilon + \bar{e}$.

3. Superconvergence error analysis

In this section, we investigate the superconvergence properties of the DG method. We first prove optimal L^2 error estimate for the DG solution. Then, we prove that the DG solution is $\mathcal{O}(h^{2p+1})$ superconvergent at the downwind point of each element. We use these results to prove that the DG solution is $\mathcal{O}(h^{p+2})$ super close to the special projection of the exact solution $P_h^- u$. Finally, we prove superconvergence at Radau points. In order to prove these results, we need to derive some error equations.

Subtracting (2.4a) from (2.3) with $v \in V_h^p$ and using the numerical flux (2.4b), we obtain the DG orthogonality condition for the error e on I_j:

$$\int_{I_j} v' e dt + \int_{I_j}(f(t,u) - f(t,u_h))v dt - e(t_j^-)v(t_j^-) + e(t_{j-1}^-)v(t_{j-1}^+) = 0, \quad \forall\, v \in V_h^p, \tag{3.1}$$

which, after using a simple integration by parts on the first term, is equivalent to

$$\int_{I_j}\left(e' - f(t,u) + f(t,u_h)\right)v dt + [e](t_{j-1})v(t_{j-1}^+) = 0. \tag{3.2}$$

Using the classical Taylor's series with integral remainder in the variable u, we write

$$f(t,u)-f(t,u_h)=\theta(u-u_h)=\theta e, \quad \text{where } \theta=\int_0^1 f_u(t,u+s(u_h-u))ds=\int_0^1 f_u(t,u-se)ds, \tag{3.3}$$

since $u-u_h=e$. Substituting (3.3) into (3.2) yields

$$\int_{I_j}\left(e'-\theta e\right)v dt+[e](t_{j-1})v(t_{j-1}^+)=0, \quad \forall\, v\in V_h^p. \tag{3.4}$$

For convenience, we define the operator $\mathcal{A}_j(e;V)$ as

$$\mathcal{A}_j(e;V)=\int_{I_j}(e'-\theta e)V dt+[e](t_{j-1})V(t_{j-1}^+). \tag{3.5}$$

We note that (3.4) can be written as

$$\mathcal{A}_j(e;v)=0, \quad \forall\, v\in V_h^p. \tag{3.6}$$

On the one hand, after a simple integration by parts, $\mathcal{A}_j(e;V)$ can be written as

$$\mathcal{A}_j(e;V)=\int_{I_j}(-V'-\theta V)e dt+e(t_j^-)V(t_j^-)-e(t_{j-1}^-)V(t_{j-1}^+). \tag{3.7}$$

On the other hand, adding and subtracting P_h^+V to V, we write (3.5) as

$$\mathcal{A}_j(e;V)=\mathcal{A}_j(e;V-P_h^+V)+\mathcal{A}_j(e;P_h^+V). \tag{3.8}$$

Combining (3.8) and (3.6) with $v=P_h^+V\in P^p(I_j)$ and using the property of the projection P_h^+, *i.e.*, $(V-P_h^+V)(t_{j-1}^+)=0$, we obtain

$$\mathcal{A}_j(e;V)=\int_{I_j}(e'-\theta e)(V-P_h^+V)dt+[e](t_{j-1})(V-P_h^+V)(t_{j-1}^+)=\int_{I_j}(e'-\theta e)(V-P_h^+V)dt. \tag{3.9}$$

By the property of the projection P_h^+, we have

$$\int_{I_j}v'(V-P_h^+V)dt=0, \quad \forall\, v\in P^p(I_j), \tag{3.10}$$

since v is a polynomial of degree at most p and thus v' is a polynomial of degree at most $p-1$.

Substituting $e=\epsilon+\bar{e}$ into (3.9) and using (3.10) with $v=\bar{e}$, we get

$$\mathcal{A}_j(e;V)=\int_{I_j}(\epsilon'-\theta e)(V-P_h^+V)dt+\int_{I_j}\bar{e}'(V-P_h^+V)dt=\int_{I_j}(\epsilon'-\theta e)(V-P_h^+V)dt. \tag{3.11}$$

Next, we state and prove optimal L^2 error estimate for $\|e\|$.

Theorem 3.1. *Let u be the exact solution of (2.1), which is assumed to be sufficiently smooth with bounded derivatives, i.e., $\|u\|_{p+1,\Omega}$ is bounded. We further assume that $|f_u(t,u)|\le M_1$ on the set $D=[0,T]\times\mathbb{R}$. Let $p\ge 0$ and u_h be the DG solution of (2.4), then, for sufficiently small h, there exists a positive constant C independent of h such that,*

$$\|e\|\le Ch^{p+1}. \tag{3.12}$$

Proof. The main idea behind the proof is to construct the following adjoint problem: find a function φ such that

$$-\varphi'-\theta\varphi=e, \quad t\in[0,T) \quad \text{subject to} \quad \varphi(T)=0, \tag{3.13}$$

where $\theta=\theta(t)=\int_0^1 f_u(t,u(t)-se(t))ds$. This problem has the following exact solution

$$\varphi(t)=\frac{1}{\Theta(t)}\int_t^T\Theta(y)e(y)dy, \quad \text{where} \quad \Theta(t)=\exp\left(-\int_t^T\theta(s)ds\right). \tag{3.14}$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

First, we prove some regular estimates needed in our error analysis. We note that the assumption $|f_u(t,u)| \leq M_1$ is sufficient to guarantee that the function $\theta(t)$, $t \in [0,T]$ is bounded by M_1 since

$$\big|\theta(t)\big| \leq \int_0^1 \big|f_u(t,u(t)-se(t))\big|ds \leq \int_0^1 M_1 ds = M_1, \quad \forall\, t \in [0,T]. \tag{3.15a}$$

From the definition of Θ and the estimate (3.15a), we have

$$0 \leq \Theta(t) \leq \exp\left(\int_0^T |\theta(s)|ds\right) \leq \exp\left(\int_0^T M_1 ds\right) = \exp(M_1 T) = C_1. \tag{3.15b}$$

Similarly, we estimate $\frac{1}{\Theta(t)}$ as follows

$$0 \leq \frac{1}{\Theta(t)} = \exp\left(\int_t^T \theta(s)ds\right) \leq \exp\left(\int_0^T |\theta(s)|ds\right) \leq \exp\left(\int_0^T M_1 ds\right) = \exp(M_1 T) = C_1. \tag{3.15c}$$

Using the definition of φ, applying the estimates (3.15b)–(3.15c) and the Cauchy–Schwarz inequality, we obtain

$$|\varphi(t)| \leq \frac{1}{\Theta(t)} \int_t^T \Theta(y)|e(y)|dy \leq C_1 \int_t^T C_1|e(y)|dy \leq C_1^2 \int_0^T |e(y)|dy \leq C_1^2 T^{1/2} \|e\|, \quad t \in [0,T],$$

which, after squaring both sides and intergrading over Ω, yields

$$\|\varphi\|^2 \leq C_1^4 T^2 \|e\|^2 = C_2 \|e\|^2. \tag{3.16a}$$

We also need to estimate $|\varphi|_{1,\Omega}$. Using (3.13) and (3.15a), we get

$$\big|\varphi'\big| = \big|\theta\varphi + e\big| \leq M_1\big|\varphi\big| + \big|e\big|, \quad t \in [0,T].$$

Squaring both sides, using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, intergrading over Ω, and applying (3.16a), we arrive at

$$|\varphi|_{1,\Omega}^2 = \int_0^T \big|\varphi'\big|^2 dt \leq 2M_1^2 \|\varphi\|^2 + 2\|e\|^2 \leq (2M_1^2 C_2 + 2)\|e\|^2 = C_3 \|e\|^2. \tag{3.16b}$$

Finally, using the well-known projection result and the estimate (3.16b), we obtain

$$\big\|\varphi - P_h^+ \varphi\big\| \leq C_4 h\, |\varphi|_{1,\Omega} \leq C_5 h \|e\|. \tag{3.16c}$$

Now, we are ready to prove (3.12). On the one hand, using (3.7) with $V = \varphi$ and (3.13), we have

$$\mathcal{A}_j(e;\varphi) = \int_{I_j} (-\varphi' - \theta\varphi)e dt + e(t_j^-)\varphi(t_j) - e(t_{j-1}^-)\varphi(t_{j-1}) = \int_{I_j} e^2 dt + e(t_j^-)\varphi(t_j) - e(t_{j-1}^-)\varphi(t_{j-1}),$$

which, after summing over the elements and using the fact that $\varphi(T) = e(t_0^-) = 0$, gives

$$\sum_{j=1}^N \mathcal{A}_j(e;\varphi) = \|e\|^2 + e(T^-)\varphi(T^-) - e(t_0^-)\varphi(t_0^-) = \|e\|^2. \tag{3.17}$$

On the other hand, taking $V = \varphi$ in (3.11), we obtain

$$\mathcal{A}_j(e;\varphi) = \int_{I_j} (\epsilon' - \theta e)(\varphi - P_h^+\varphi)dt. \tag{3.18}$$

Summing over all elements and applying the Cauchy–Schwarz inequality yields

$$\sum_{j=1}^N \mathcal{A}_j(e;\varphi) \leq (\big\|\epsilon'\big\| + M_1\|e\|)\big\|\varphi - P_h^+\varphi\big\|.$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Applying the estimate (3.16c), we get

$$\sum_{j=1}^{N} \mathcal{A}_j(e;\varphi) \leq (C_0 h^p |u|_{p+1,\Omega} + M_1 \|e\|) C_1 h \|e\| \leq C(h^{p+1} + h\|e\|)\|e\|. \tag{3.19}$$

Combining the two formulas (3.17) and (3.19) yields

$$\|e\| \leq Ch^{p+1} + Ch\|e\|. \tag{3.20}$$

Consequently, $(1-Ch)\|e\| \leq Ch^{p+1}$, where C is a constant independent of h. Hence, for sufficiently small h, *e.g.*, $h \leq \frac{1}{2C}$, we have $\frac{1}{2}\|e\| \leq (1-Ch)\|e\| \leq Ch^{p+1}$, which gives $\|e\| \leq 2Ch^{p+1}$ for small h. This completes the proof of (3.12). □

In the next theorem, we prove that the DG solution is $\mathcal{O}(h^{2p+1})$ superconvergent at the downwind points. We further prove that the DG solution is $\mathcal{O}(h^{p+2})$ super close to $P_h^- u$.

Theorem 3.2. *Assume that the assumptions of Theorem 3.1 are satisfied. In addition, we assume that f_u is sufficiently smooth with respect to the variables t and u (for example, the function $h(t) = f_u(t,u(t)) \in C^p([0,T])$ is enough). Then there exists a positive constant C such that*

$$\left|e(t_k^-)\right| \leq Ch^{2p+1}, \quad k = 1,\ldots,N. \tag{3.21}$$

$$\left|\bar{e}(t_k^-)\right| \leq Ch^{2p+1}, \quad k = 1,\ldots,N. \tag{3.22}$$

$$\left\|\bar{e}'\right\| \leq Ch^{p+1}. \tag{3.23}$$

$$\|\bar{e}\| \leq Ch^{p+2}. \tag{3.24}$$

Proof. In order to prove global superconvergence of DG solutions at the mesh points we proceed by the duality argument. Let W be the solution of the following auxiliary problem:

$$W' + \theta W = 0, \quad t \in [0,t_k) \quad \text{subject to} \quad W(t_k) = 1, \tag{3.25}$$

where $\theta = \theta(t) = \int_0^1 f_u(t,u(t) - se(t))ds$ and $1 \leq k \leq N$. The exact solution is given by $W(t) = \exp\left(\int_t^{t_k} \theta(s)ds\right)$, $t \in \Omega_k = [0,t_k]$. We note that the estimate (3.15a) and the assumption $h(t) = f_u(t,u(t)) \in C^p([0,T])$ indicated that there exists a constant C such that

$$\|W\|_{p+1,\Omega_k} \leq C. \tag{3.26}$$

By (3.7) and (3.25), we have

$$\mathcal{A}_j(e;W) = \int_{I_j} (-W' - \theta W)e dt + e(t_j^-)W(t_j) - e(t_{j-1}^-)W(t_{j-1}) = e(t_j^-)W(t_j) - e(t_{j-1}^-)W(t_{j-1}).$$

Summing over the elements I_j, $j = 1,\ldots,k$ and using $W(t_k) = 1$ and $e(t_0^-) = 0$, we get

$$\sum_{j=1}^{k} \mathcal{A}_j(e;W) = e(t_k^-)W(t_k) - e(t_0^-)W(t_0) = e(t_k^-). \tag{3.27}$$

On the other hand, choosing $V = W$ in (3.11), we get

$$\mathcal{A}_j(e;W) = \int_{I_j} (\epsilon' - \theta e)(W - P_h^+ W)dt.$$

Summing over the elements I_j, $j = 1,\ldots,k$ with $1 \leq k \leq N$ and using (3.27), we obtain

$$e(t_k^-) = \sum_{j=1}^{k} \int_{I_j} (\epsilon' - \theta e)(W - P_h^+ W)dt.$$

Applying (3.15a) and the Cauchy–Schwarz inequality, we obtain

$$\left|e(t_k^-)\right| \leq (\left\|\epsilon'\right\|_{0,\Omega_k} + M_1 \|e\|_{0,\Omega_k}) \left\|W - P_h^+ W\right\|_{0,\Omega_k} \leq (\left\|\epsilon'\right\| + M_1\|e\|) \left\|W - P_h^+ W\right\|_{0,\Omega_k}.$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Applying (3.12), (2.13), and (3.26), we get

$$\left|e(t_k^-)\right| \leq (C_0 h^p |u|_{p+1,\Omega} + M_1 C_1 h^{p+1}) C_2 h^{p+1} |W|_{p+1,\Omega_k} \leq C(h^p + h^{p+1}) h^{p+1} = \mathcal{O}(h^{2p+1}), \tag{3.28}$$

for all $k = 1, \ldots, N$, which completes the proof of (3.21).

In order to show (3.22) we use the relation $e = \bar{e} + \epsilon$, the property of the projection P_h^-, *i.e.*, $\epsilon(t_k^-) = 0$, and the estimate (3.21) to get

$$\left|\bar{e}(t_k^-)\right| = \left|e(t_k^-) - \epsilon(t_k^-)\right| = \left|e(t_k^-)\right| = \mathcal{O}(h^{2p+1}).$$

Next, we will estimate $\|\bar{e}'\|$. By the property of the projection P_h^-, we have

$$\int_{I_j} \epsilon v' dt = 0, \quad \forall\, v \in P^p(I_j), \quad \text{and} \quad \epsilon(t_j^-) = 0, \quad j = 1, \ldots, N, \tag{3.29}$$

since v is a polynomial of degree at most p and thus v' is a polynomial of degree at most $p-1$.

Substituting $e = \epsilon + \bar{e}$ into (3.1) and applying (3.29) yields

$$\int_{I_j} v' \bar{e} dt + \int_{I_j} (f(t,u) - f(t,u_h)) v dt - \bar{e}(t_j^-) v(t_j^-) + \bar{e}(t_{j-1}^-) v(t_{j-1}^+) = 0,$$

which, after using a simple integration by parts on the first term, is equivalent to

$$\int_{I_j} \left(\bar{e}' - f(t,u) + f(t,u_h)\right) v dt + [\bar{e}](t_{j-1}) v(t_{j-1}^+) = 0. \tag{3.30}$$

Taking $v(t) = \bar{e}'(t) - (-1)^p \bar{e}'(t_{j-1}^+) L_{p,j}(t) \in P^p(I_j)$ in (3.30), we have, by the property $\tilde{L}_p(-1) = (-1)^p$ and the orthogonality relation (2.5b), $v(t_{j-1}^+) = 0$ and

$$\begin{aligned}\int_{I_j} (\bar{e}')^2 dt &= (-1)^p \bar{e}'(t_{j-1}^+) \int_{I_j} L_{p,j} \bar{e}' dt + \int_{I_j} (f(t,u) - f(t,u_h))(\bar{e}' - (-1)^p \bar{e}'(t_{j-1}^+) L_{p,j}) dt \\ &= \int_{I_j} (f(t,u) - f(t,u_h))(\bar{e}' - (-1)^p \bar{e}'(t_{j-1}^+) L_{p,j}) dt. \end{aligned} \tag{3.31}$$

Using the Lipschitz condition (2.2) and applying the Cauchy–Schwarz inequality yields

$$\begin{aligned}\|\bar{e}'\|_{I_j}^2 &\leq \int_{I_j} |f(t,u) - f(t,u_h)| \left(|\bar{e}'| + \left|\bar{e}'(t_{j-1}^+)\right| |L_{p,j}|\right) dt \leq M_1 \int_{I_j} |e| \left(|\bar{e}'| + \left|\bar{e}'(t_{j-1}^+)\right| |L_{p,j}|\right) dt \\ &\leq M_1 \|e\|_{I_j} \left(\|\bar{e}'\|_{I_j} + \left|\bar{e}'(t_{j-1}^+)\right| \|L_{p,j}\|_{I_j}\right). \end{aligned} \tag{3.32}$$

Combining (3.32) with (2.14b) and (2.9a), we obtain

$$\|\bar{e}'\|_{I_j}^2 \leq M_1 \|e\|_{I_j} \left(\|\bar{e}'\|_{I_j} + C_1 h_j^{-1/2} \|\bar{e}'\|_{I_j} \frac{h_j^{1/2}}{(2p+1)^{1/2}}\right) \leq C \|e\|_{I_j} \|\bar{e}'\|_{I_j}.$$

Therefore, $\|\bar{e}'\|_{I_j} \leq C \|e\|_{I_j}$. Squaring both sides, summing over all elements, and applying (3.12), we get

$$\|\bar{e}'\|^2 \leq C \|e\|^2 \leq C h^{2p+2}. \tag{3.33}$$

Finally, we will estimate $\|\bar{e}\|$. By the Fundamental Theorem of Calculus, we have

$$\left|\bar{e}(t)\right| = \left|\bar{e}(t_j^-) + \int_{t_j}^t \bar{e}'(s) ds\right| \leq \left|\bar{e}(t_j^-)\right| + \int_{I_j} \left|\bar{e}'(s)\right| ds, \quad \forall\, t \in I_j.$$

Squaring both sides, using $(a+b)^2 \leq 2a^2 + 2b^2$ and applying the Cauchy–Schwartz inequality, we obtain

$$\left|\bar{e}(t)\right|^2 \leq 2\left|\bar{e}(t_j^-)\right|^2 + 2\left(\int_{I_j} \left|\bar{e}'(s)\right| ds\right)^2 \leq 2\left|\bar{e}(t_j^-)\right|^2 + 2h_j \int_{I_j} \left|\bar{e}'(s)\right|^2 ds = 2\left|\bar{e}(t_j^-)\right|^2 + 2h_j \|\bar{e}'\|_{I_j}^2.$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Integrating this inequality with respect to t and using the estimate (3.22), we get

$$\|\bar{e}\|_{I_j}^2 \leq 2h_j\left|\bar{e}(t_j^-)\right|^2 + 2h_j^2\left\|\bar{e}'\right\|_{I_j}^2 \leq 2Ch_j^{4p+3} + 2h_j^2\left\|\bar{e}'\right\|_{I_j}^2.$$

Summing over all elements and using the estimate (3.23) and the fact that $h = \max h_j$, we obtain

$$\|\bar{e}\|^2 \leq C_1 h^{4p+2} + 2h^2\left\|\bar{e}'\right\|^2 \leq C_1 h^{4p+2} + 2C_2 h^{2p+4} = \mathcal{O}(h^{2p+4}), \tag{3.34}$$

since $4p+2 \geq 2p+4$ for $p \geq 1$. Thus, we have completed the proof of the theorem. □

Thus, the numerical solution u_h is closer to $P_h^- u$ than to the exact solution u. Next, we use the results of Theorem 3.2 to show that the true error e can be divided into a significant part and a less significant part. The significant part is proportional to the $(p+1)$-degree right Radau polynomial and the less significant part converges at $\mathcal{O}(h^{p+2})$ rate in the L^2-norm. We first define two interpolation operators π and $\hat{\pi}$. The operator π is defined as follows: For any function $u = u(t)$, $\pi u\big|_{I_j} \in P^p(I_j)$ and interpolates u at the roots $t_{j,i}$, $i = 0, 1, \ldots, p$, of the $(p+1)$-degree right Radau polynomial shifted to I_j. Next, the operator $\hat{\pi}$ is such that $\hat{\pi} u\big|_{I_j} \in P^{p+1}(I_j)$ and is defined as follows: $\hat{\pi} u\big|_{I_j}$ interpolates u at $t_{j,i}$, $i = 0, 1, \ldots, p$, and at an additional point $\bar{t}_j$ in I_j with $\bar{t}_j \neq t_{j,i}$, $i = 0, 1, \ldots, p$.

For the sake of completeness, we include the following results from [4] which will be needed in our error analysis. In particular, we show that the interpolation error can be divided into significant and less significant parts.

Lemma 3.1. *Let $u \in H^{p+2}(I_j)$ and P_h^- and π as defined above. Then*

$$u - \pi u = \phi_j + \gamma_j, \quad \text{on } I_j, \tag{3.35a}$$

where

$$\phi_j(t) = \alpha_j \psi_{p+1,j}(t), \quad \psi_{p+1,j}(t) = \prod_{i=0}^{p}(t - t_{j,i}), \quad \gamma_j = u - \hat{\pi} u, \tag{3.35b}$$

and α_j is the coefficient of t^{p+1} in the $(p+1)$-degree polynomial $\hat{\pi} u$. Furthermore,

$$\left\|\phi_j\right\|_{s,I_j} \leq Ch_j^{p+1-s}\|u\|_{p+1,I_j}, \quad s = 0, \ldots, p, \tag{3.35c}$$

$$\left\|\gamma_j\right\|_{s,I_j} \leq Ch_j^{p+2-s}\|u\|_{p+2,I_j}, \quad s = 0, \ldots, p+1. \tag{3.35d}$$

Finally,

$$\left\|\pi u - P_h^- u\right\|_{I_j} \leq Ch_j^{p+2}\ \|u\|_{p+2,I_j}. \tag{3.36}$$

Proof. The proof of these results can be found in the paper by Adjerid and Baccouch [4]. More precisely in its Lemma 2.1. □

Now, we are ready to state and prove the main global superconvergence result.

Theorem 3.3. *Under the assumptions of Theorem 3.2, there exists a positive constant C independent of h such that*

$$\|u_h - \pi u\| \leq Ch^{p+2}. \tag{3.37}$$

Moreover, the true error can be divided into a significant part and a less significant part as

$$e(t) = \alpha_j \psi_{p+1,j}(t) + \omega_j(t), \quad \text{on } I_j, \tag{3.38a}$$

where

$$\omega_j = \gamma_j + \pi u - u_h, \tag{3.38b}$$

and

$$\sum_{j=1}^{N}\left\|\omega_j\right\|_{I_j}^2 \leq Ch^{2(p+2)}, \qquad \sum_{j=1}^{N}\left\|\omega_j'\right\|_{I_j}^2 \leq Ch^{2(p+1)}. \tag{3.38c}$$

Finally,

$$\sum_{j=1}^{N}\left\|e'\right\|_{I_j}^2 \leq Ch^{2p} \quad \text{and} \quad \|e\|_{1,\Omega} \leq Ch^p. \tag{3.39}$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Proof. Adding and subtracting $P_h^- u$ to $u_h - \pi u$, we write

$$u_h - \pi u = u_h - P_h^- u + P_h^- u - \pi u = -\bar{e} + P_h^- u - \pi u.$$

Taking the L^2-norm and applying the triangle inequality, we obtain

$$\|u_h - \pi u\| \leq \|\bar{e}\| + \left\| P_h^- u - \pi u \right\|.$$

Using the estimates (3.24) and (3.36), we establish (3.37).

Next, we add and subtract πu to e, we have

$$e = u - \pi u + \pi u - u_h. \tag{3.40}$$

Furthermore, one can split the interpolation errors $u - \pi u$ on I_j as in (3.35a) to obtain

$$e = \phi_j + \gamma_j + \pi u - u_h = \phi_j + \omega_j, \quad \text{where} \quad \omega_j = \gamma_j + \pi u - u_h. \tag{3.41}$$

Next, we use the Cauchy–Schwarz inequality and the inequality $|ab| \leq \frac{1}{2}(a^2 + b^2)$ to write

$$\begin{aligned} \left\|\omega_j\right\|_{I_j}^2 &= \left(\gamma_j + \pi u - u_h, \gamma_j + \pi u - u_h\right)_{I_j} = \left\|\gamma_j\right\|_{I_j}^2 + 2\left(\gamma_j, \pi u - u_h\right)_{I_j} + \|\pi u - u_h\|_{I_j}^2 \\ &\leq 2\left(\left\|\gamma_j\right\|_{I_j}^2 + \|\pi u - u_h\|_{I_j}^2\right). \end{aligned}$$

Summing over all elements and applying (3.35d) with $s = 0$ and (3.37) yields the first estimate in (3.38c).

Using the Cauchy–Schwarz inequality and the inequality $|ab| \leq \frac{1}{2}(a^2 + b^2)$, we write

$$\left\|\omega_j'\right\|_{I_j}^2 = \left(\gamma_j' + (\pi u - u_h)', \gamma_j' + (\pi u - u_h)'\right)_{I_j} \leq 2\left(\left\|\gamma_j'\right\|_{I_j}^2 + \left\|(\pi u - u_h)'\right\|_{I_j}^2\right). \tag{3.42}$$

Using the inverse inequality (2.14a), *i.e.*, $\left\|(\pi u - u_h)'\right\|_{I_j} \leq Ch^{-1} \|(\pi u - u_h)\|_{I_j}$, we obtain the estimate

$$\left\|\omega_j'\right\|_{I_j}^2 \leq C\left(\left\|\gamma_j'\right\|_{I_j}^2 + h^{-2} \|\pi u - u_h\|_{I_j}^2\right).$$

Summing over all elements and applying (3.37) and the estimate (3.35d) with $s = 1$, we establish the second estimate in (3.38c).

In order to show (3.39), we note that

$$\|e\|_{1,\Omega}^2 = \|e\|^2 + \sum_{j=1}^{N} \left\|e'\right\|_{I_j}^2. \tag{3.43}$$

Differentiating (3.41) with respect to t, taking the L^2-norm, and applying the Cauchy–Schwarz inequality and the inequality $|ab| \leq \frac{1}{2}(a^2 + b^2)$, we get

$$\left\|e'\right\|_{I_j}^2 = \left(\phi_j' + \omega_j', \phi_j' + \omega_j'\right)_{I_j} \leq 2\left(\left\|\phi_j'\right\|_{I_j}^2 + \left\|\omega_j'\right\|_{I_j}^2\right).$$

Summing over all elements and applying (3.35c) and (3.38c), we obtain

$$\sum_{j=1}^{N} \left\|e'\right\|_{I_j}^2 \leq Ch^{2p}. \tag{3.44}$$

Finally, substituting (3.12) and (3.44) into (3.43) establishes (3.39). □

4. *A posteriori* error estimation

In this section, we present a technique to compute asymptotically correct *a posteriori* estimates of the DG errors for the nonlinear IVP (2.1). These estimates are computed by solving a local problem with no boundary condition on each element. We further prove that the DG discretization error estimates converge to the true spatial errors in the L^2-norm as $h \to 0$. Next, we present the weak finite element formulation to compute *a posteriori* error estimate for the nonlinear IVP (2.1).

Multiplying (2.1) by arbitrary smooth function v and integrating over an arbitrary element I_j, we get

$$\int_{I_j} u' v dt = \int_{I_j} f(t, u) v dt. \tag{4.1}$$

Replacing u by $u_h + e$ and choosing $v = \psi_{p+1,j}(t)$, we obtain

$$\int_{I_j} e' \psi_{p+1,j} dt = \int_{I_j} \left(f(t, u_h + e) - u_h' \right) \psi_{p+1,j} dt. \tag{4.2}$$

Substituting (3.38a), *i.e.*, $e(t) = \alpha_j \psi_{p+1,j}(t) + \omega_j(t)$, into the left-hand side of (4.2) yields

$$\alpha_j \int_{I_j} \psi_{p+1,j}' \psi_{p+1,j} dt = \int_{I_j} \left(f(t, u_h + e) - u_h' - \omega_j' \right) \psi_{p+1,j} dt. \tag{4.3}$$

Using (2.9b) and solving for α_j, we obtain

$$\alpha_j = -\frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(f(t, u_h + e) - u_h' - \omega_j' \right) \psi_{p+1,j} dt. \tag{4.4}$$

Our error estimate procedure consists of approximating the true error on each element I_j by the leading term as

$$e(t) \approx E(t) = a_j \psi_{p+1,j}(t), \quad t \in I_j, \tag{4.5a}$$

where the coefficient of the leading term of the error, a_j, is obtained from the coefficient α_j defined in (4.4) by neglecting the terms ω_j and e, *i.e.*,

$$a_j = -\frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(f(t, u_h) - u_h' \right) \psi_{p+1,j} dt. \tag{4.5b}$$

We note that our error estimates are obtained by solving local problems with no boundary conditions.

An accepted efficiency measure of *a posteriori* error estimates is the effectivity index. In this paper, we use the global effectivity index

$$\sigma = \frac{\|E\|}{\|e\|},$$

and is used to appraise the accuracy of the error estimate. Ideally, the global effectivity index should stay close to one and should converge to one under mesh refinement.

Next, we will show that the error estimate E converges to the exact error e in the L^2-norm as $h \to 0$. Furthermore, we will prove the convergence to unity of the global effectivity index σ under mesh refinement.

Before stating our main result we state and prove the following preliminary results.

Theorem 4.1. *Suppose that the assumptions of Theorem 3.2 are satisfied. If α_j and a_j are given by (4.4) and (4.5b), respectively, then there exists a positive constant C independent of h such that*

$$\sum_{j=1}^{N} \left(a_j - \alpha_j \right)^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \leq C h^{2p+4}. \tag{4.6}$$

$$\sum_{j=1}^{N} \left(a_j^2 + \alpha_j^2 \right) \left\| \psi_{p+1,j} \right\|_{I_j}^2 \leq C h^{2p+2}. \tag{4.7}$$

$$\sum_{j=1}^{N} \left| a_j^2 - \alpha_j^2 \right| \left\| \psi_{p+1,j} \right\|_{I_j}^2 \leq C h^{2p+3}. \tag{4.8}$$

Proof. Subtracting (4.4) from (4.5b), we obtain

$$a_j - \alpha_j = \frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(f(t, u_h + e) - f(t, u_h) - \omega_j' \right) \psi_{p+1,j} dt. \tag{4.9}$$

Thus,

$$\left| a_j - \alpha_j \right| \leq \frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(\left| f(t, u_h + e) - f(t, u_h) \right| + \left| \omega_j' \right| \right) \left| \psi_{p+1,j} \right| dt. \tag{4.10}$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Using the Lipschitz condition (2.2) and applying the Cauchy–Schwarz inequality yields

$$|a_j-\alpha_j| \leq \frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(M_1 |e| + \left|\omega_j'\right|\right) |\psi_{p+1,j}|\, dt \leq \frac{\|\psi_{p+1,j}\|_{I_j}}{k_1 h_j^{2p+2}} \left(M_1 \|e\|_{I_j} + \left\|\omega_j'\right\|_{I_j}\right). \tag{4.11}$$

Applying the inequality $(a+b)^2 \leq 2(a^2+b^2)$, we obtain

$$(a_j-\alpha_j)^2 \leq \frac{2\|\psi_{p+1,j}\|_{I_j}^2}{k_1^2 h_j^{4p+4}} \left(M_1^2 \|e\|_{I_j}^2 + \left\|\omega_j'\right\|_{I_j}^2\right). \tag{4.12}$$

Multiplying by $\|\psi_{p+1,j}\|_{I_j}^2$ and using (2.9c), *i.e.*, $\|\psi_{p+1,j}\|_{I_j}^2 = (2p+2)k_2 h_j^{2p+3}$ yields

$$(a_j-\alpha_j)^2 \|\psi_{p+1,j}\|_{I_j}^2 \leq \frac{2\|\psi_{p+1,j}\|_{I_j}^4}{k_1^2 h_j^{4p+4}} \left(M_1^2 \|e\|_{I_j}^2 + \left\|\omega_j'\right\|_{I_j}^2\right) \leq k_3 h_j^2 \left(\|e\|_{I_j}^2 + \left\|\omega_j'\right\|_{I_j}^2\right), \tag{4.13a}$$

where k_3 is a constant independent of the mesh size given by

$$k_3 = \frac{2(2p+2)^2 k_2^2}{k_1^2} \max(M_1^2, 1). \tag{4.13b}$$

Summing over all elements and using $h = \max\limits_{1\leq j\leq N} h_j$, we arrive at

$$\sum_{j=1}^{N} (a_j-\alpha_j)^2 \|\psi_{p+1,j}\|_{I_j}^2 \leq k_3 h^2 \left[\|e\|^2 + \sum_{j=1}^{N} \left\|\omega_j'\right\|_{I_j}^2\right].$$

Combining this estimate with (3.12) and (3.38c), we establish (4.6).

Next, we will prove (4.7). Since $u_h = u - e$ and $u' = f(t,u)$, we have

$$a_j = -\frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(f(t,u-e) - f(t,u) + e'\right) \psi_{p+1,j} dt. \tag{4.14}$$

Using the Lipschitz condition (2.2) and applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned} |a_j| &\leq \frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(|f(t,u-e) - f(t,u)| + |e'|\right) |\psi_{p+1,j}|\, dt \\ &\leq \frac{1}{k_1 h_j^{2p+2}} \int_{I_j} \left(M_1 |e| + |e'|\right) |\psi_{p+1,j}|\, dt \leq \frac{\|\psi_{p+1,j}\|_{I_j}}{k_1 h_j^{2p+2}} \left(M_1 \|e\|_{I_j} + \|e'\|_{I_j}\right). \end{aligned} \tag{4.15}$$

Applying the inequality $(a+b)^2 \leq 2(a^2+b^2)$, we get

$$a_j^2 \leq \frac{2\|\psi_{p+1,j}\|_{I_j}^2}{k_1^2 h_j^{4p+4}} \left(M_1^2 \|e\|_{I_j}^2 + \|e'\|_{I_j}^2\right). \tag{4.16}$$

Multiplying both sides by $\|\psi_{p+1,j}\|_{I_j}^2 = (2p+2)k_2 h_j^{2p+3}$, we obtain

$$\begin{aligned} a_j^2 \|\psi_{p+1,j}\|_{I_j}^2 &\leq \frac{2\|\psi_{p+1,j}\|_{I_j}^4}{k_1^2 h_j^{4p+4}} \left(M_1^2 \|e\|_{I_j}^2 + \|e'\|_{I_j}^2\right) = \frac{2(2p+2)^2 k_2^2}{k_1^2} h_j^2 \left(M_1^2 \|e\|_{I_j}^2 + \|e'\|_{I_j}^2\right) \\ &\leq k_3 h_j^2 \left(\|e\|_{I_j}^2 + \|e'\|_{I_j}^2\right), \end{aligned}$$

where k_3 is defined in (4.13b). Summing over all elements and using $h = \max\limits_{1\leq j\leq N} h_j$ and the estimates (3.12) and (3.39), we obtain

$$\sum_{j=1}^{N} a_j^2 \|\psi_{p+1,j}\|_{I_j}^2 \leq k_3 h^2 \left[\|e\|^2 + \|e'\|^2\right] \leq C_1 h^{2p+2}. \tag{4.17}$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

On the other hand, taking the L^2 inner product of $\psi_{p+1,j}$ and $\phi_j = \alpha_j \psi_{p+1,j}$, defined in (3.35b), and applying the Cauchy–Schwarz inequality, we write

$$|\alpha_j| \left\| \psi_{p+1,j} \right\|_{I_j}^2 = \left| (\phi_j, \psi_{p+1,j})_{I_j} \right| \leq \left\| \psi_{p+1,j} \right\|_{I_j} \left\| \phi_j \right\|_{I_j}. \tag{4.18}$$

Hence, we have

$$\alpha_j^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \leq \left\| \phi_j \right\|_{I_j}^2.$$

Summing over all elements and applying (3.35c), we obtain

$$\sum_{j=1}^{N} \alpha_j^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \leq \sum_{j=1}^{N} \left\| \phi_j \right\|_{I_j}^2 \leq C_2 h^{2p+2}. \tag{4.19}$$

Adding (4.17) and (4.19) yields (4.7).

Finally, we will prove (4.8). Using the Cauchy–Schwarz inequality, the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, and applying the estimates (4.6) and (4.7), we obtain

$$\begin{aligned}
\sum_{j=1}^{N} |\alpha_j^2 - a_j^2| \left\| \psi_{p+1,j} \right\|_{I_j}^2 &= \sum_{j=1}^{N} \left(|\alpha_j - a_j| \left\| \psi_{p+1,j} \right\|_{I_j} \right) \left(|\alpha_j + a_j| \left\| \psi_{p+1,j} \right\|_{I_j} \right) \\
&\leq \left(\sum_{j=1}^{N} (\alpha_j - a_j)^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \right)^{1/2} \left(\sum_{j=1}^{N} (\alpha_j + a_j)^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \right)^{1/2} \\
&\leq \sqrt{2} \left(\sum_{j=1}^{N} (\alpha_j - a_j)^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 \right)^{1/2} \left(\sum_{j=1}^{N} (\alpha_j^2 + a_j^2) \left\| \psi_{p+1,j} \right\|_{I_j}^2 \right)^{1/2} \\
&\leq \sqrt{2} \left(C h^{2p+4} \right)^{1/2} \left(C h^{2p+2} \right)^{1/2} = \mathcal{O}(h^{2p+3}). \qquad \square
\end{aligned}$$

The main results of this section are stated in the following theorem. In particular, we state and prove asymptotic results of our *a posteriori* error estimates.

Theorem 4.2. *Suppose that the assumptions of Theorem 3.2 are satisfied. If $E(t) = a_j \psi_{p+1,j}(t)$, $t \in I_j$, where a_j, $j = 1, \ldots, N$, are given by (4.5b), then there exists a positive constant C independent of h such that*

$$\|e - E\|^2 \leq C h^{2p+4}. \tag{4.20}$$

As a consequence, the DG method combined with the a posteriori error estimation procedure yields $\mathcal{O}(h^{p+2})$ superconvergent solution i.e.,

$$\|u - (u_h + E)\|^2 = \sum_{j=1}^{N} \left\| u - (u_h + a_j \psi_{p+1,j}) \right\|_{I_j}^2 \leq C h^{2p+4}. \tag{4.21}$$

Furthermore,

$$\|e\|^2 = \|E\|^2 + \hat{\epsilon} = \sum_{j=1}^{N} a_j^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 + \hat{\epsilon}, \quad \text{where } |\hat{\epsilon}| \leq C h^{2p+3}. \tag{4.22}$$

Finally, if there exists a constant $C = C(u) > 0$ independent of h such that

$$\|e\| \geq C h^{p+1}, \tag{4.23}$$

then the global effectivity index in the L^2 norm, which is defined as $\sigma = \frac{\|E\|}{\|e\|}$, converges to unity at $\mathcal{O}(h)$ rate, i.e.,

$$\sigma = 1 + \mathcal{O}(h). \tag{4.24}$$

Proof. First, we will prove (4.20) and (4.21). Since $e = \alpha_j \psi_{p+1,j} + \omega_j$ and $E = a_j \psi_{p+1,j}$ on I_j, we have

$$\|e - E\|_{I_j}^2 = \left\| (\alpha_j - a_j) \psi_{p+1,j} + \omega_j \right\|_{I_j}^2 \leq 2(\alpha_j - a_j)^2 \left\| \psi_{p+1,j} \right\|_{I_j}^2 + 2 \left\| \omega_j \right\|_{I_j}^2,$$

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

where we used the inequality $(a+b)^2 \leq 2a^2+2b^2$. Summing over all elements and applying the estimates (3.38c) and (4.6) yields

$$\|e-E\|^2 = \sum_{j=1}^{N} \|e-E\|_{I_j}^2 \leq 2\sum_{j=1}^{N} (\alpha_j - a_j)^2 \left\|\psi_{p+1,j}\right\|_{I_j}^2 + 2\sum_{j=1}^{N} \left\|\omega_j\right\|_{I_j}^2 \leq 2C_1 h^{2p+4} + 2C_2 h^{2p+4} = Ch^{2p+4}.$$

Using the relation $e = u - u_h$ and the estimate (4.20), we obtain

$$\sum_{j=1}^{N} \left\|u - (u_h + a_j \psi_{p+1,j})\right\|_{I_j}^2 = \|u - (u_h + E)\|^2 = \|e-E\|^2 \leq Ch^{2p+4}.$$

Next, we will prove (4.22). From (3.38a) the DG error can be split as $e = \phi_j + \omega_j$, on I_j. Taking the L^2 norm of e and using (3.35b) we have

$$\|e\|_{I_j}^2 = \left\|\phi_j\right\|_{I_j}^2 + 2(\phi_j, \omega_j)_{I_j} + \left\|\omega_j\right\|_{I_j}^2 = \alpha_j^2 \left\|\psi_{p+1,j}\right\|_{I_j}^2 + \epsilon_j, \tag{4.25a}$$

where

$$\epsilon_j = 2(\phi_j, \omega_j)_{I_j} + \left\|\omega_j\right\|_{I_j}^2. \tag{4.25b}$$

Summing over all elements, we obtain

$$\|e\|^2 = \sum_{j=1}^{N} \alpha_j^2 \left\|\psi_{p+1,j}\right\|_{I_j}^2 + \tilde{\epsilon}, \quad \text{where} \quad \tilde{\epsilon} = \sum_{j=1}^{N} \epsilon_j. \tag{4.26}$$

Next, we write the DG error as

$$\|e\|^2 = \sum_{j=1}^{N} a_j^2 \left\|\psi_{p+1,j}\right\|_{I_j}^2 + \hat{\epsilon}, \quad \text{where} \quad \hat{\epsilon} = \sum_{j=1}^{N} (\alpha_j^2 - a_j^2) \left\|\psi_{p+1,j}\right\|_{I_j}^2 + \tilde{\epsilon}. \tag{4.27}$$

On the one hand, applying the estimate (4.8), we get

$$|\hat{\epsilon}| \leq \sum_{j=1}^{N} |\alpha_j^2 - a_j^2| \left\|\psi_{p+1,j}\right\|_{I_j}^2 + |\tilde{\epsilon}| \leq Ch^{2p+3} + |\tilde{\epsilon}|. \tag{4.28}$$

On the other hand, from (4.25b), we apply the Cauchy–Schwarz inequality to obtain

$$|\tilde{\epsilon}| \leq \sum_{j=1}^{N} |\epsilon_j| \leq 2\sum_{j=1}^{N} \left\|\phi_j\right\|_{I_j} \left\|\omega_j\right\|_{I_j} + \sum_{j=1}^{N} \left\|\omega_j\right\|_{I_j}^2. \tag{4.29}$$

Applying the Cauchy–Schwarz inequality and using the estimates (3.35c) and (3.38c), we get

$$|\tilde{\epsilon}| \leq 2\left(\sum_{j=1}^{N} \left\|\phi_j\right\|_{I_j}^2\right)^{1/2} \left(\sum_{j=1}^{N} \left\|\omega_j\right\|_{I_j}^2\right)^{1/2} + \sum_{j=1}^{N} \left\|\omega_j\right\|_{I_j}^2 \leq C_1 h^{2p+3}. \tag{4.30}$$

Finally, combining (4.28) and (4.30) we complete the proof of (4.22).

In order to show (4.24), we use the triangle inequality to have

$$\big|\, \|E\| - \|e\| \,\big| \leq \|E-e\|,$$

which, after dividing by $\|e\|$, yields

$$\left| \frac{\|E\|}{\|e\|} - 1 \right| \leq \frac{\|E-e\|}{\|e\|}. \tag{4.31}$$

Applying the estimate (4.20) and the inverse estimate (4.23), we arrive at

$$\left| \frac{\|E\|}{\|e\|} - 1 \right| \leq Ch.$$

Therefore, $\frac{\|E\|}{\|e\|} = 1 + \mathcal{O}(h)$, which establishes (4.24). □

In the previous theorem, we proved that the global *a posteriori* error estimates converge to the true spatial errors at $\mathcal{O}(h^{p+2})$ rate. We further proved that the global effectivity index in the L^2-norm converges to unity at $\mathcal{O}(h)$ rate.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

Remark 4.1. We note that the computable quantity $u_h + E$ converge to the exact solution u at $\mathcal{O}(h^{p+2})$ rate. This accuracy enhancement is simply achieved by adding the error estimate E to the approximate solution u_h. Finally, we note that our DG error estimates are obtained by solving a local problem with no boundary conditions on each element. This leads to very efficient computations of the post-processed approximation $u_h + E$.

Remark 4.2. The performance of an error estimator is commonly measured by the effectivity index which is the ratio of the estimated error to the actual error. In particular, we say that the error estimator is asymptotically exact if the effectivity index approaches unity as the mesh size goes to zero. Thus, (4.24) indicated that our *a posteriori* error estimator is asymptotically exact. We note that E is a computable quantity since it only depends on the numerical solution u_h and the f. It provides an asymptotically exact *a posteriori* estimator on the actual error $\|e\|$. We would like to emphasize that our DG error estimate is computationally simple which make it useful in adaptive computations.

Remark 4.3. We note that, from (3.34), if $p = 0$ then $\|\bar{e}\| = \mathcal{O}(h)$ which is the same as $\|e\|$. Thus, when $p = 0$, there is no superconvergence for global errors. Also, our error estimate procedure does not apply.

Remark 4.4. The assumption (4.23) implies that terms of order $\mathcal{O}(h^{p+1})$ are present in the error. If this were not the case, the error estimate E might not be such a good approximation of the error e. Even though the proof of (4.24) is valid under the assumption (4.23), our computational results given in the next section demonstrate that the global effectivity index in the L^2-norm converge to unity at $\mathcal{O}(h)$ rate. Thus, the proposed error estimation technique is an excellent measure of the error.

We note that the *a priori* estimate (3.12) is optimal in the sense that the exponent of h is the largest possible. In fact, one may show that provided that the $(p+1)$st-order derivatives of the exact solution u do not vanish identically over the domain ($u \notin V_h^p$), then an inverse estimate of the form (4.23) is valid [8,37,38] for some positive constant C which depends on u but not on h. Combining (3.12) with (4.23), we show that u_h approximates u to $\mathcal{O}(h^{p+1})$ in the L^2 norm.

5. Computational results

In this section, we numerically validate our theoretical results. We compute the maximum DG error at shifted roots of $(p+1)$-degree right-Radau polynomial on each element I_j and then take the maximum over all elements. We also compute the DG error at the downwind point of each element and then take the maximum over all elements I_j, $j = 1, \ldots, N$. For simplicity, we denote

$$\|e\|^* = \max_{1\le j\le N}\left(\max_{0\le i\le p}|e(t_{j,i}^-)|\right), \quad |e|^* = \max_{1\le j\le N}|e(t_j^-)|,$$

where $t_{j,i}$ are the roots of $R_{p+1,j}(t)$. We also define δe and $\delta\sigma$ as

$$\delta e = \big|\,\|e\| - \|E\|\,\big|, \quad \delta\sigma = \big|\sigma - 1\big|.$$

Example 5.1. We consider the following nonlinear IVP

$$u' = -u - u^2, \quad t \in [0,1], \quad u(0) = 1.$$

The exact solution is given by $u(t) = \frac{1}{2e^t - 1}$.

We solve this problem using the DG method on uniform meshes obtained by partitioning the domain $[0,1]$ into N subintervals with $N = 5, 10, 20, 30, 40, 50$ and using the spaces P^p with $p = 0$–4.

In Fig. 1 we plot, in log–log scale, the L^2 errors $||e||$ and $||\bar{e}||$ versus h. For each P^p space, we fit, in a least-squares sense, the data sets with a linear function and then calculate from the fitting result the slopes of the fitting lines. The slopes of the fitting lines are shown on the graph. We observe that $||e|| = \mathcal{O}(h^{p+1})$ and $||\bar{e}|| = \mathcal{O}(h^{p+2})$. We observe that the DG solution is $\mathcal{O}(h^{p+2})$ super close to the projection $P_h^- u$ except for $p = 0$, where there is no superconvergence for global errors. We note that the observed numerical convergence rate is optimal for $p \ge 1$.

We present the zero-level curves of the true error e in Fig. 2 using $N = 10$ and for p ranging from 1 to 4. The Radau points of degree $p+1$ are shown on each element as $\times$ signs. We observe that the errors vanish at points close to the Radau points for all solutions except the one with $p = 0$, where there is no superconvergence for global errors and, thus, the error estimation procedure does not apply. The maximum errors $\|e\|^*$, $|e|^*$, and their orders of convergence shown in Fig. 3 indicate that the DG error e is $\mathcal{O}(h^{p+2})$ superconvergent at the roots of $(p+1)$-degree right Radau polynomial and $\mathcal{O}(h^{2p+1})$ at the end of each step. Thus, globally, the DG solution converges at an $\mathcal{O}(h^{2p+1})$ rate at the downwind end of each subinterval and at an $\mathcal{O}(h^{p+2})$ rate at the remainder of the Radau points. This is in full agreement with the theory.

On each element, we apply the error estimation procedure (4.5) to compute the error estimate for the DG solution. In Fig. 4, we present the convergence rates for the global errors $||e - E||$ using the spaces P^p with $p = 1$–4. These results

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

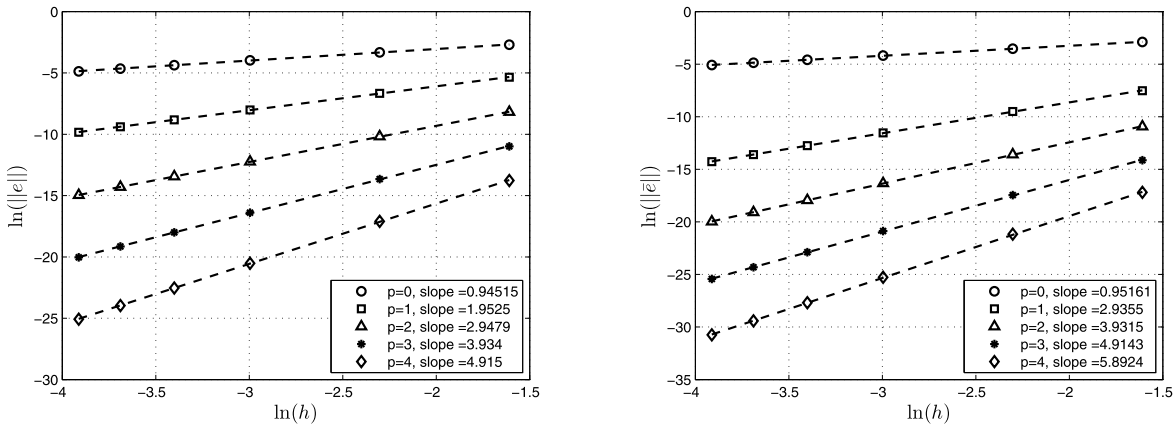

Fig. 1. Log–log plots of $||e||$ (left) and $||\bar{e}||$ (right) versus mesh sizes h for Example 5.1 on uniform meshes having $N = 5, 10, 20, 30, 40, 50$ elements using P^p, $p = 0$ to 4.

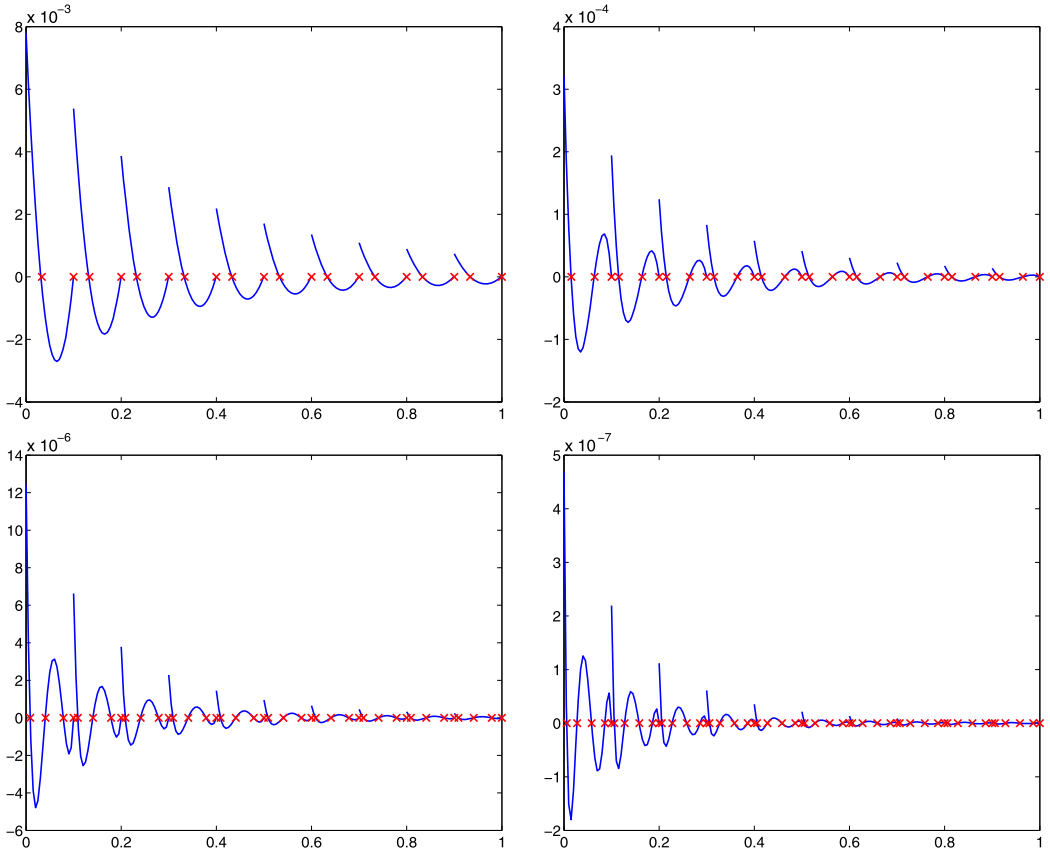

Fig. 2. Zero-level curves of e for Example 5.1 using P^p, $p = 1$–4 (upper left to lower right) on uniform meshes having $N = 10$ elements.

indicate that $||e - E|| = \mathcal{O}(h^{p+2})$. This is in full agreement with the theory. This example demonstrates that the convergence rate proved in this paper is optimal. We note that

$$||e - E|| = ||u - (u_h + E)|| = \mathcal{O}(h^{p+2}).$$

Thus, the computable quantities $u_h + E$ converge to the exact solution u at $\mathcal{O}(h^{p+2})$ rate. We note that this accuracy enhancement is achieved by adding the error estimate E to the DG solution only once at the end of the computation. This leads to a very efficient computation of the post-processed approximation $u_h + E$.

The true L^2 errors and the global effectivity indices shown in Table 1 indicate that our *a posteriori* error estimates converge to the true errors under both h- and p-refinements.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

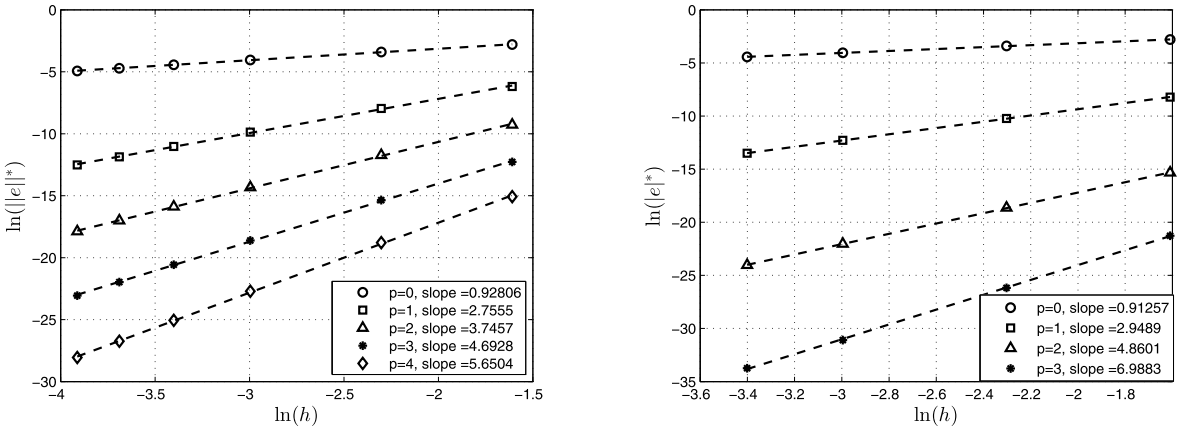

Fig. 3. (Left) Log–log plots of $||e||^*$ versus h for Example 5.1 using $N = 5, 10, 20, 30, 40, 50$ and P^p, $p = 0$ to 4. (Right) Log–log plots of $|e|^*$ versus h using $N = 5, 10, 20, 30$ elements using P^p, $p = 0$ to 3.

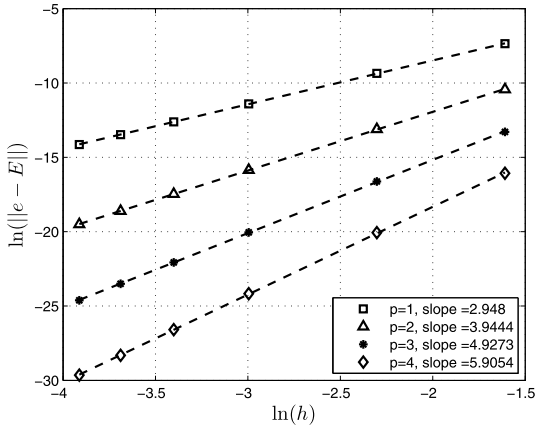

Fig. 4. Convergence rates of $||e - E||$ for Example 5.1 on uniform meshes having $N = 5, 10, 20, 30, 40, 50$ elements using P^p, $p = 1$ to 4.

Table 1
L^2 errors and global effectivity indices for Example 5.1 on uniform meshes having $N = 5, 10, 20, 30, 40, 50$ elements using P^p, $p = 1$ to 4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|\|e\|\|$	σ	$\|\|e\|\|$	σ	$\|\|e\|\|$	σ	$\|\|e\|\|$	σ
5	4.7637e−3	1.0362	2.7867e−4	1.0531	1.6847e−5	1.0637	1.0386e−6	1.0705
10	1.2750e−3	1.0179	3.7805e−5	1.0271	1.1742e−6	1.0326	3.7481e−8	1.0363
20	3.2849e−4	1.0089	4.8747e−6	1.0136	7.6227e−8	1.0164	1.2290e−9	1.0182
30	1.4736e−4	1.0059	1.4568e−6	1.0090	1.5201e−8	1.0109	1.6369e−10	1.0122
40	8.3262e−5	1.0044	6.1698e−7	1.0068	4.8296e−9	1.0082	3.9026e−11	1.0091
50	5.3429e−5	1.0035	3.1660e−7	1.0054	1.9827e−9	1.0066	1.2820e−11	1.0073

The results shown in Fig. 5 indicate that the numerical convergence rates for δe and $\delta\sigma$ are $\mathcal{O}(h^{p+2})$ and $\mathcal{O}(h)$, respectively. We note that the observed convergence rate for the global effectivity index is optimal. We note that the effectivity indices converge under h- and p-refinements. Numerical results further indicate that the error estimates converge to the true errors with decreasing mesh size and increasing polynomial degree p.

A posteriori error estimates are traditionally used to guide adaptive enrichment by h-refinement and to provide a measure of solution reliability. Adaptive methods based on *a posteriori* error estimates have become a common procedure for obtaining more accurate numerical solutions. In the next examples, we use our *a posteriori* errors estimates to construct efficient adaptive high-order DG method. Such adaptive DG method consists of successive loops of the cycle

SOLVE → ESTIMATE → MARK → REFINE.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

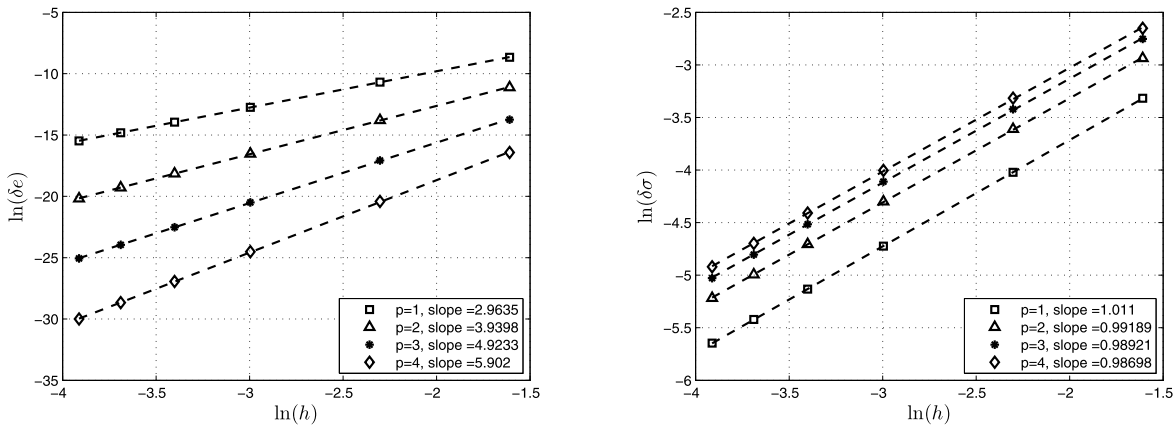

Fig. 5. Convergence rates for δe (left) and $\delta\sigma$ (right) for Example 5.1 on uniform meshes having $N = 5, 10, 20, 30, 40, 50$ elements using P^p, $p = 1$ to 4.

The local *a posteriori* errors estimate of section 4 is used to mark elements for refinement. The adaptive mesh refinement (AMR) refines the mesh according to the local *a posteriori* error indicator. Typically, an adaptive algorithm adds extra nodes to an existing mesh to refine the mesh where the error is large, for example near singularities or discontinuities.

In this paper, we propose a simple adaptive strategy which consists of solving the IVP (2.1) on a variable grid whose elements are denoted I_j, $j = 1, \ldots, N$. The local error estimate (4.5) is used to compute the local errors $\|E\|_{I_j}$, $j = 1, \ldots, N$ and the global error $\|E\| = \left(\sum_{j=1}^{N} \|E\|_{I_j}^2\right)^{1/2}$. Given a user tolerance $Tol > 0$, our adaptive refinement strategy consists of the following steps:

1. Set a tolerance Tol and a maximum bound on the number of nodes $N_{max} = 1000$. Set $\|E\| = 1$.
2. Create a coarse mesh with $N + 1$ nodes. Here we choose a uniform mesh with $N = 2$ elements.
3. While $N + 1 \leq N_{max}$ and $\|E\| \geq Tol$ do
   (a) Find the DG solution u_h as described in section 2.
   (b) Compute the L^2-norm of the local error estimates $\|E\|_{I_j}$, $j = 1, \ldots, N$ as described in section 4 and the global error estimate $\|E\| = \left(\sum_{j=1}^{N} \|E\|_{I_j}^2\right)^{1/2}$.
   (c) For all elements I_j
      i. if $\|E\|_{I_j} < \lambda \max_{j=1,\ldots,N} \|E\|_{I_j}$ then we accept the DG solution on I_j. Here $0 \leq \lambda \leq 1$ is a parameter which must be chosen by the user. Note that $\lambda = 0$ gives a uniform refinement, while $\lambda = 1$ gives no refinement at all.
      ii. else, we reject the DG solution on I_j and subdivide I_j into 2 uniform elements. Thus, we add the coordinate of the midpoint of the element I_j to the nodes.
4. endwhile

Remark 5.1. There are different possibilities for selecting the elements to be refined given the local errors $\|E\|_{I_j}$. In the above algorithm, we refine the element I_j if $\|E\|_{I_j} > \lambda \max_{j=1,\ldots,N} \|E\|_{I_j}$. Similarly, there are other stopping criteria which can be used to determine when the adaptive algorithm should stop.

In the next examples, we propose and test the above adaptive algorithm for the selection of the mesh.

Example 5.2. To test the adaptive algorithm, we consider the following linear problem

$$u' = \beta u, \quad t \in [0, 5], \quad u(0) = 1.$$

The exact solution is $u(t) = e^{\beta t}$. We apply our adaptive algorithm using $\beta = 1$ (unstable), $\beta = -1$ (stable), and $\beta = -20$ (stiff). In Figs. 6–8, we show the DG solutions and the sequence of meshes obtained by applying the above adaptive algorithm with $Tol = 10^{-2}$ for $p = 1$–4 using $\beta = 1$, $\beta = -1$, and $\beta = -20$, respectively. We observe that the adaptive procedure refines elements in which the *a posteriori* error estimator is large. From these results, we can see that, when λ is closer to 0, we get more uniform refinement near the portion with high approximation error. When λ is closer to 1, we get less uniform refinement near the portion with high approximation error. This example shows that our error estimates are efficient and accurate under adaptive mesh refinement.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

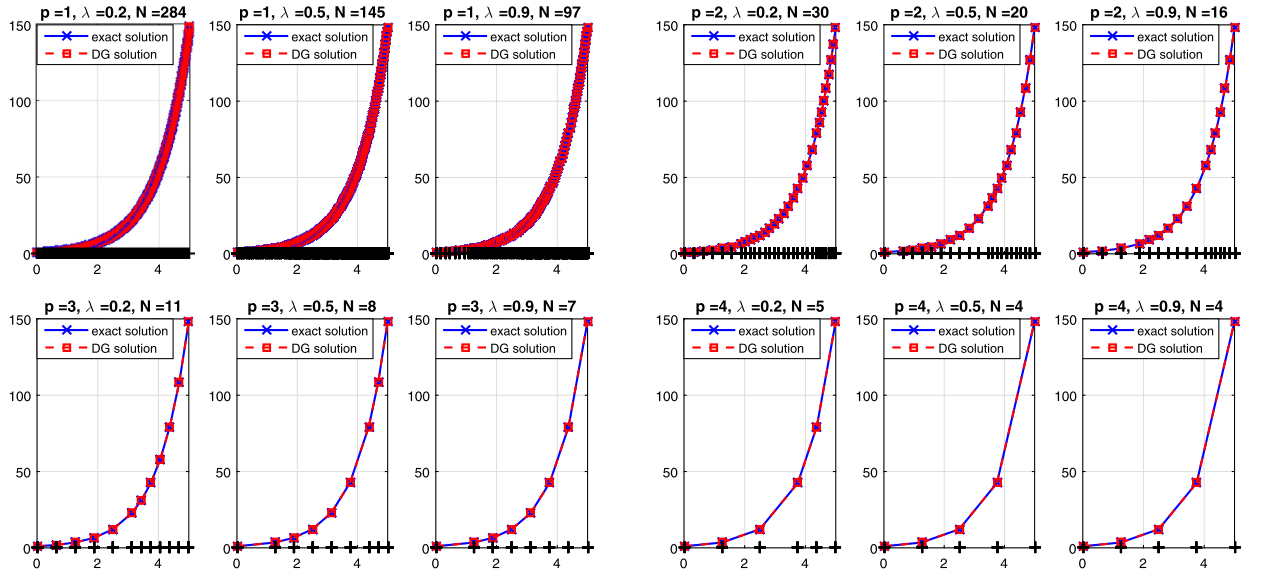

Fig. 6. DG solutions, exact solution, and final meshes for Example 5.2 with $\beta = 1$ using P^p, $p = 1$ to 4, and $Tol = 10^{-2}$.

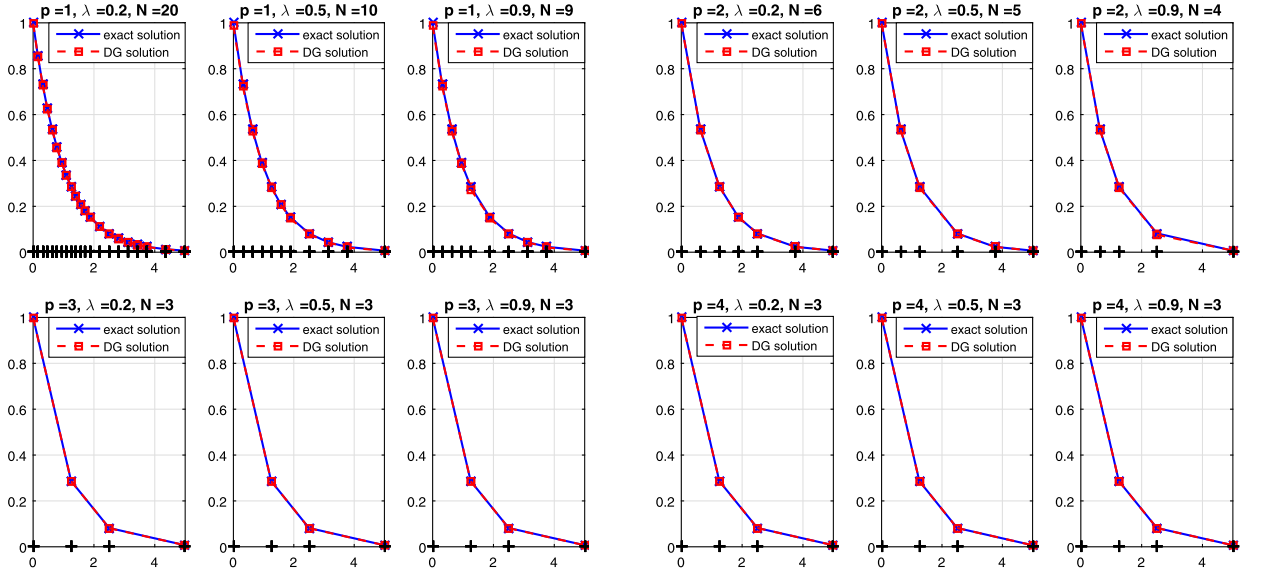

Fig. 7. DG solutions, exact solution, and final meshes for Example 5.2 with $\beta = -1$ using P^p, $p = 1$ to 4, and $Tol = 10^{-2}$.

Example 5.3. In this example, we apply our adaptive algorithm to the following linear problem

$$u' + u = f(t), \quad t \in [0, 1], \quad u(0) = u_0.$$

We select $f(t)$ and u_0 such that the exact solution is $u(t) = e^{-1000(x-0.5)^2} + 1$. In Fig. 9, we show the DG solutions and the sequence of meshes obtained by applying our adaptive algorithm with $Tol = 10^{-3}$ for $p = 1$–4. We observe that the adaptive procedure refines elements in which the *a posteriori* error estimator is large.

Example 5.4. As a final test, we consider the following IVP

$$u' + u = f(t), \quad t \in [0, 1], \quad u(0) = 1.$$

We select $f(t)$ such that the exact solution is $u(t) = 1 + 10^{-5}t(1-t)e^{15t}$, which is smooth function but it has a steep front and gradient near $t = 1$. We solve this problem using the DG method described in section 2. We start with a uniform meshes having $N = 2$ elements and using the space P^p with $p = 1$–4. Fig. 10 shows the DG solutions and the final mesh which is constructed through a sequence of successively refined meshes where elements for which the L^2 norm of the

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

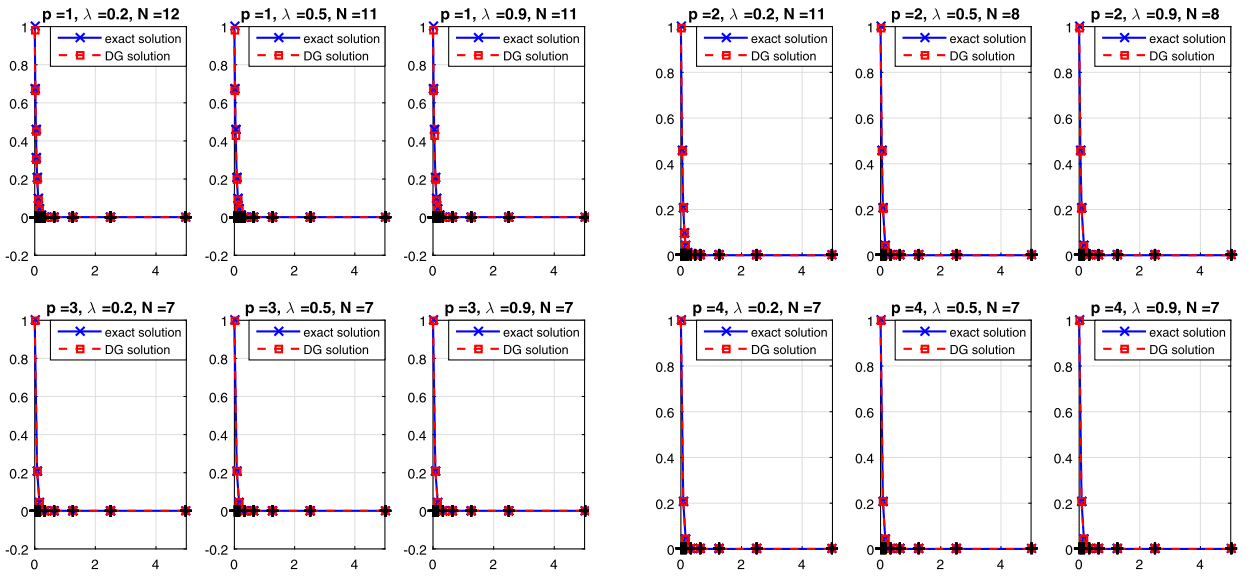

Fig. 8. DG solutions, exact solution, and final meshes for Example 5.2 with $\beta = -20$ using P^p, $p = 1$ to 4, and $Tol = 10^{-2}$.

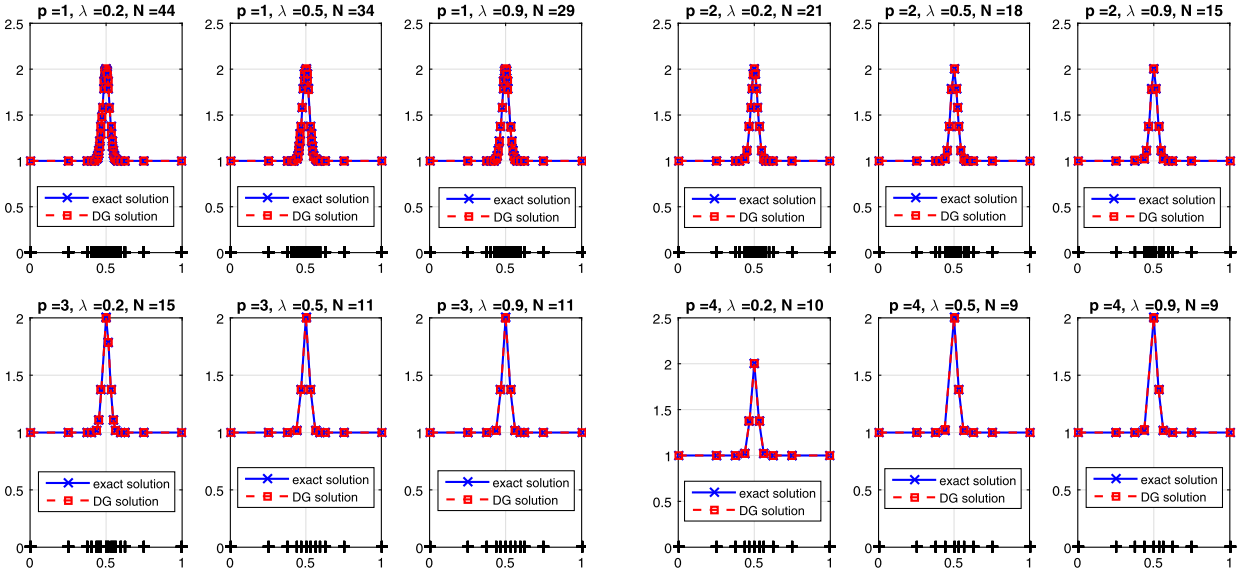

Fig. 9. DG solutions, exact solution, and final meshes for Example 5.3 using P^p, $p = 1$ to 4, and $Tol = 10^{-3}$.

local error estimate $\|E_u\|_{I_j}$ on each element is larger than $Tol = 10^{-3}$ are refined. We observe that the number of elements needed to achieve the error tolerance $Tol = 10^{-3}$ decreases under p-refinement. In Fig. 11, we show the DG solutions and the sequence of meshes obtained by applying our adaptive algorithm with $Tol = 10^{-5}$ for $p = 1$–4. We observe from our experiments that the local adaptive procedure is able to capture the steep front and gradient by refining the elements in the neighborhood of the point $t = 1$. These computational results indicate that our estimators are accurate on adaptively refined meshes and further suggest that they converge to the true error under adaptive mesh refinement.

6. Concluding remarks

In this paper, we analyzed the original discontinuous Galerkin (DG) finite element method for the approximation of initial-value problems for ordinary differential equations. We proved that the DG solution exhibits an optimal $\mathcal{O}(h^{p+1})$ convergence rate in the L^2-norm when p-degree piecewise polynomials are used. We further proved that the p-degree DG solution is $\mathcal{O}(h^{p+2})$ super close to a special projection of the exact solution. We used these results to show that the significant part of the discretization error for the DG solution is proportional to the $(p+1)$-degree right Radau polynomial.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

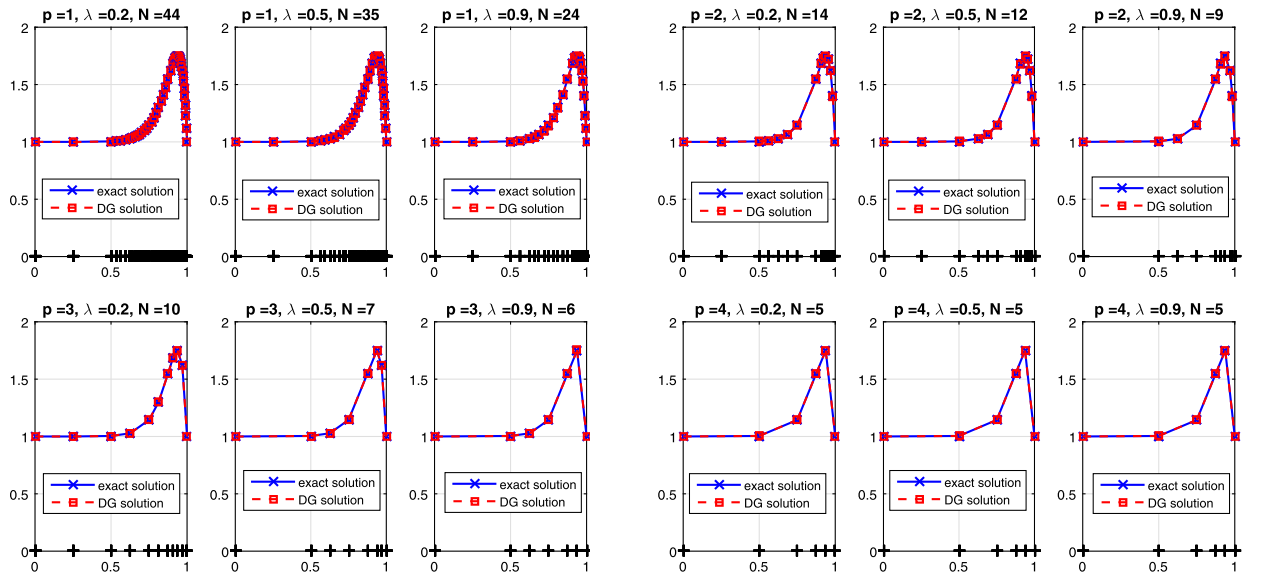

Fig. 10. DG solutions, exact solution, and final meshes for Example 5.4 using P^p, $p = 1$ to 4, and $Tol = 10^{-3}$.

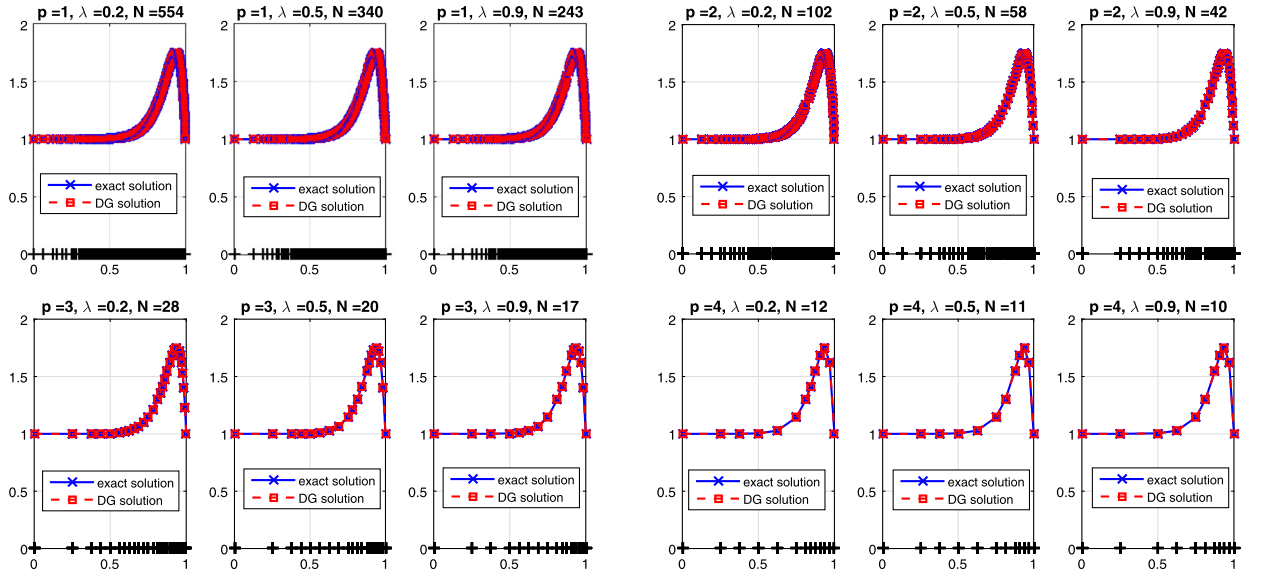

Fig. 11. DG solutions, exact solution, and final meshes for Example 5.4 using P^p, $p = 1$ to 4, and $Tol = 10^{-5}$.

These results are used to construct asymptotically correct *a posteriori* error estimates of spatial discretization errors. We proved that the DG discretization error estimates converge to the true spatial errors under mesh refinement in the L^2-norm. The order of convergence is proved to be of order $p + 2$. We further proved that the global effectivity index in the L^2-norm converges to unity at $\mathcal{O}(h)$ rate. Our proofs are valid for arbitrary regular meshes and for P^p polynomials with $p \geq 1$. The extension of our proofs for nonlinear vector IVPs is straight forward. We are currently investigating the superconvergence properties and the *a posteriori* error analysis of the DG method applied to boundary-value problems.

Acknowledgements

The authors would like to thank the two anonymous reviewers for the valuable comments and suggestions which improve the quality of the paper. This research was supported by the University Committee on Research and Creative Activity (UCRCA Proposal 2015-01-F) at the University of Nebraska at Omaha.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover, New York, 1965.
[2] S. Adjerid, M. Baccouch, The discontinuous Galerkin method for two-dimensional hyperbolic problems. Part I: superconvergence error analysis, J. Sci. Comput. 33 (2007) 75–113.
[3] S. Adjerid, M. Baccouch, The discontinuous Galerkin method for two-dimensional hyperbolic problems. Part II: *a posteriori* error estimation, J. Sci. Comput. 38 (2009) 15–49.
[4] S. Adjerid, M. Baccouch, Asymptotically exact *a posteriori* error estimates for a one-dimensional linear hyperbolic problem, Appl. Numer. Math. 60 (2010) 903–914.
[5] S. Adjerid, M. Baccouch, Adaptivity and error estimation for discontinuous Galerkin methods, in: X. Feng, O. Karakashian, Y. Xing (Eds.), Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations, in: The IMA Volumes in Mathematics and Its Applications, vol. 157, Springer International Publishing, Switzerland, 2014, pp. 63–96.
[6] S. Adjerid, K.D. Devine, J.E. Flaherty, L. Krivodonova, *A posteriori* error estimation for discontinuous Galerkin solutions of hyperbolic problems, Comput. Methods Appl. Mech. Eng. 191 (2002) 1097–1112.
[7] S. Adjerid, H. Temimi, A discontinuous Galerkin method for higher-order ordinary differential equations, Comput. Methods Appl. Mech. Eng. 197 (2007) 202–218.
[8] M. Ainsworth, J.T. Oden, *A Posteriori* Error Estimation in Finite Element Analysis, John Wiley, New York, 2000.
[9] I. Babuška, T. Strouboulis, The Finite Element Method and Its Reliability, Numerical Mathematics and Scientific Computation, Clarendon Press, 2001.
[10] M. Baccouch, *A posteriori* error estimates for a discontinuous Galerkin method applied to one-dimensional nonlinear scalar conservation laws, Appl. Numer. Math. 84 (2014) 1–21.
[11] M. Baccouch, Global convergence of *a posteriori* error estimates for a discontinuous Galerkin method for one-dimensional linear hyperbolic problems, Int. J. Numer. Anal. Model. 11 (2014) 172–192.
[12] M. Baccouch, S. Adjerid, Discontinuous Galerkin error estimation for hyperbolic problems on unstructured triangular meshes, Comput. Methods Appl. Mech. Eng. 200 (2010) 162–177.
[13] F. Bassi, S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations, J. Comput. Phys. 131 (1997) 267–279.
[14] C.E. Baumann, J.T. Oden, A discontinuous *hp* finite element method for convection–diffusion problems, Comput. Methods Appl. Mech. Eng. 175 (1999) 311–341.
[15] K.S. Bey, J.T. Oden, *hp*-version discontinuous Galerkin method for hyperbolic conservation laws, Comput. Methods Appl. Mech. Eng. 133 (1996) 259–286.
[16] K.S. Bey, J.T. Oden, A. Patra, *hp*-version discontinuous Galerkin method for hyperbolic conservation laws: a parallel strategy, Int. J. Numer. Methods Eng. 38 (1995) 3889–3908.
[17] K.S. Bey, J.T. Oden, A. Patra, A parallel *hp*-adaptive discontinuous Galerkin method for hyperbolic conservation laws, Appl. Numer. Math. 20 (1996) 321–386.
[18] R. Biswas, K. Devine, J.E. Flaherty, Parallel adaptive finite element methods for conservation laws, Appl. Numer. Math. 14 (1994) 255–284.
[19] K. Böttcher, R. Rannacher, Adaptive error control in solving ordinary differential equations by the discontinuous Galerkin method, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, IWR, 1996.
[20] Y. Cheng, C.-W. Shu, Superconvergence of discontinuous Galerkin and local discontinuous Galerkin schemes for linear hyperbolic and convection–diffusion equations in one space dimension, SIAM J. Numer. Anal. 47 (2010) 4044–4072.
[21] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland Pub. Co., Amsterdam–New York–Oxford, 1978.
[22] B. Cockburn, G.E. Karniadakis, C.W. Shu, Discontinuous Galerkin Methods Theory, Computation and Applications, Lecture Notes in Computational Science and Engineering, vol. 11, Springer, Berlin, 2000.
[23] B. Cockburn, S.Y. Lin, C.W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin methods of scalar conservation laws III: one dimensional systems, J. Comput. Phys. 84 (1989) 90–113.
[24] B. Cockburn, C.W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin methods for scalar conservation laws II: general framework, Math. Comput. 52 (1989) 411–435.
[25] B. Cockburn, C.W. Shu, The local discontinuous Galerkin method for time-dependent convection–diffusion systems, SIAM J. Numer. Anal. 35 (1998) 2440–2463.
[26] M. Delfour, F. Dubeau, Discontinuous polynomial approximations in the theory of one-step, hybrid and multistep methods for nonlinear ordinary differential equations, Math. Comput. 47 (1986) 169–189.
[27] M. Delfour, W. Hager, F. Trochu, Discontinuous Galerkin methods for ordinary differential equation, Math. Comput. 154 (1981) 455–473.
[28] K. Deng, Z. Xiong, Superconvergence of a discontinuous finite element method for a nonlinear ordinary differential equation, Appl. Math. Comput. 217 (2010) 3511–3515.
[29] K.D. Devine, J.E. Flaherty, Parallel adaptive *hp*-refinement techniques for conservation laws, Comput. Methods Appl. Mech. Eng. 20 (1996) 367–386.
[30] D. Estep, *A posteriori* error bounds and global error control for approximation of ordinary differential equations, SIAM J. Numer. Anal. 32 (1995) 1–48.
[31] J.E. Flaherty, R. Loy, M.S. Shephard, B.K. Szymanski, J.D. Teresco, L.H. Ziantz, Adaptive local refinement with octree load-balancing for the parallel solution of three-dimensional conservation laws, J. Parallel Distrib. Comput. 47 (1997) 139–152.
[32] C. Johnson, Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations, SIAM J. Numer. Anal. 25 (1988) 908–926.
[33] P. Lesaint, P. Raviart, On a finite element method for solving the neutron transport equations, in: C. de Boor (Ed.), Mathematical Aspects of Finite Elements in Partial Differential Equations, Academic Press, New York, 1974, pp. 89–123.
[34] X. Meng, C.-W. Shu, Q. Zhang, B. Wu, Superconvergence of discontinuous Galerkin methods for scalar nonlinear conservation laws in one space dimension, SIAM J. Numer. Anal. 50 (5) (2012) 2336–2356.
[35] T.E. Peterson, A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation, SIAM J. Numer. Anal. 28 (1) (1991) 133–140.
[36] W.H. Reed, T.R. Hill, Triangular mesh methods for the neutron transport equation, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, 1973.
[37] L. Schumaker, Spline Functions: Basic Theory, Cambridge University Press, Cambridge, New York, 2007.
[38] K. Segeth, A posteriori error estimation with the finite element method of lines for a nonlinear parabolic equation in one space dimension, Numer. Math. 83 (3) (1999) 455–475.
[39] C.-W. Shu, Discontinuous Galerkin method for time-dependent problems: survey and recent developments, in: X. Feng, O. Karakashian, Y. Xing (Eds.), Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations, in: The IMA Volumes in Mathematics and Its Applications, vol. 157, Springer International Publishing, 2014, pp. 25–62.
[40] H. Temimi, S. Adjerid, Error analysis of a discontinuous Galerkin method for systems of higher-order differential equations, Appl. Math. Comput. 219 (2013) 4503–4525.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008

[41] R. Verfürth, *A posteriori* error estimation and adaptive mesh-refinement techniques, Comput. Appl. Math. 50 (1994) 67–83.
[42] L. Wahlbin, Superconvergence in Galerkin Finite Element Methods, Lecture Notes in Mathematics, Springer, 1995.
[43] M.F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM J. Numer. Anal. 15 (1978) 152–161.
[44] Z. Xionga, C. Chenb, Superconvergence of rectangular finite element with interpolated coefficients for semilinear elliptic problem, Appl. Math. Comput. 181 (2006) 1577–1584.
[45] Y. Yang, C.-W. Shu, Analysis of optimal superconvergence of discontinuous Galerkin method for linear hyperbolic equations, SIAM J. Numer. Anal. 50 (2012) 3110–3133.

Please cite this article in press as: M. Baccouch, Analysis of *a posteriori* error estimates of the discontinuous Galerkin method for nonlinear ordinary differential equations, Appl. Numer. Math. (2016), http://dx.doi.org/10.1016/j.apnum.2016.03.008