ERROR CONTROL FOR A CLASS OF RUNGE KUTTA DISCONTINUOUS GALERKIN METHODS FOR NONLINEAR CONSERVATION LAWS

ANDREAS DEDNER, CHARALAMBOS MAKRIDAKIS, AND MARIO OHLBERGER

ABSTRACT. We propose an a–posteriori error estimate for the Runge Kutta Discontinuous Galerkin method (RK–DG) of arbitrary order in arbitrary space dimensions. For stabilization of the scheme a general framework of projections is introduced. Finally it is demonstrated numerically how the a posteriori error estimate is used to design both an efficient grid adaption and gradient limiting strategy. Numerical experiments show the stability of the scheme and the gain in efficiency in comparison with computations on uniform grids.

1. INTRODUCTION

In this paper we study a generalized version of the Runge Kutta Discontinuous Galerkin approximation of Cockburn and Shu (see [8, 6, 9]) for non linear scalar conservation laws in several space dimensions. As a prototype conservation law, consider the Cauchy initial value problem

$$\partial_t u + \nabla \cdot f(u) = 0 \quad \text{in } \mathbb{R}^d \times \mathbb{R}^+ , \tag{1a}$$

$$u(x,0) = u_0(x) \quad \text{in } \mathbb{R}^d. \tag{1b}$$

Here $u : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ denotes the dependent solution variable, $f \in C^1(\mathbb{R})$ denotes the flux function, and $u_0 \in \mathrm{BV}(\mathbb{R}^d) \cap \mathrm{L}^\infty(\mathbb{R}^d)$ the initial data with $u_0 \in [A, B]$ a. e. It is well known, see for example [12, 11], that (1a)-(1b) admits an unique entropy weak solution in the class of functions of bounded variation (BV). For later use let us briefly recall that an entropy weak solution is a weak solution of (1a)-(1b) which satisfies for all entropy pairs (S, F_S)

$$-\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} \left(S(u)\partial_t\phi + F_S(u)\cdot\nabla\phi\right)\, dt\, dx - \int_{\mathbb{R}^d} S(u_0)\phi(x,0)\, dx \leq 0 \quad \text{for all } \phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)\,. \tag{2}$$

Recall that (S, F_S) is called an entropy - entropy flux pair or more simply an entropy pair for the equation (1a), iff S is convex and

$$F_S' = S'f'. \tag{3}$$

Numerical methods for nonlinear hyperbolic conservation laws are usually rather complicated since they need to approximate a partial differential equation with non-standard stability behaviour. It turns out that in many computational successful

Date: February 11, 2005.

Key words and phrases. discontinuous Galerkin, higher order, adaptive methods, error estimate, finite element.

methods the theoretical backup is very limited. This is partly because when constructing high order methods stabilization terms have to be added so that in the limit the solution satisfying (2) is computed. These mechanisms include shock capturing terms or limiters that result in complicated and highly nonlinear schemes, see [7] for a comprehensive review of high-order finite difference, finite volume, and finite element methods for hyperbolic conservation laws. The Runge Kutta Discontinuous Galerkin approximation of Cockburn and Shu, [9], is a very successful method that combines many desirable properties. It is based on totally discontinuous finite element spaces for the space discretization while the time discretization is based on appropriate Runge-Kutta schemes. The available theory for RK-DG methods for nonlinear problems is limited to certain stability and TVD properties proved in [10, 9] and to error estimates for one dimensional smooth solutions (see [30]). The problem of showing convergence towards the unique entropy solution for the high-order version of these methods seems rather difficult.

In this paper we consider a generalized version of RK-DG methods designed for use with dynamic mesh modification. We are interested in the following question: "is it possible to establish a rigorous error control for RK-DG methods in mesh adaptive computations?" An answer to this question will be based on certain a posteriori estimates and does not necessarily depend on available a priori convergence results for the method. One of the consequences is that it provides a (nonstandard) way of theoretical backup for a method with no available convergence theory, compare with [13] where this was done for MUSCL finite difference schemes.

First we prove a posteriori error estimates for generalized DG methods. We then use these estimates to provide an adaptive $h-p$ algorithm that is used together with a rigorous error control. The computational performance of the resulting methods and algorithms is tested in one dimensional examples.

The literature on a posteriori error control and adaptive solution algorithms for Discontinuous Galerkin approximations is rare. We refer for instance to [15, 23] where duality techniques were used for designing adaptive schemes, and [1] where asymptotically correct a posteriori estimates of spatial discretization errors for smooth solutions were derived in one space dimension.

2. Formulation of the generalized DG methods and main results

Let $\mathcal{T}$ denote an element decomposition of $\mathbb{R}^d$ with control volumes $T_j \in \mathcal{T}, j \in J$ such that $\cup_{T\in\mathcal{T}}\overline{T} = \mathbb{R}^d$. Let h_T denote a length scale associated with each control volume T, e.g. $h_T \equiv \text{diam}(T)$. For two distinct control volumes T_i and T_j in $\mathcal{T}$, the intersection is either an oriented edge (2-D) or face (3-D) S_{ij} with oriented normal ν_{ij} or else a set of measure at most $d-2$. The set $N(j)$ denotes the index set of neighboring control volumes to T_j and the index set of the oriented edges or faces of the grid is given by $\mathcal{E} = \{(j,l) | T_j \in \mathcal{T}, l \in N(j), j > l\}$. The set of edges or faces of the element decomposition $\mathcal{T}$ will be denoted by Γ.

On $\mathcal{T}$ we define the space of (possible) discontinuous piecewise polynomials of degree p by

$$V_h^p := \{v_h \in BV(\mathbb{R}^d) | \, v_T := v_h|_T \in \mathbb{P}_p \text{ for all } T \in \mathcal{T}\}. \tag{4}$$

Let us denote by $\Pi_{V_h^p}$ the L^2-projection into V_h^p. Furthermore, following standard notation, $[v_h]|_{S_{ij}} := (v_j|_{S_{ij}} - v_i|_{S_{ij}})\nu_{ij}$ is the jump of v_h on the edge S_{ij}, and $\{v_h\}_{S_{ij}} := 1/2(v_j|_{S_{ij}} + v_i|_{S_{ij}})$ denotes the mean of v_h at an interface.

The semi-discrete DG-finite element scheme is the basis of the definition of RK-DG methods.

Definition 2.1 (Space-discrete DG approximation). *$u_h \in C^1(0,T;V_h^p)$ is called a semi-discrete DG approximation of (1a)-(1b), iff*

$$u_h(0) = \Pi_{V_h^p}(u_0), \tag{5a}$$

$$\frac{d}{dt}(u_h(t), v_h) - (f(u_h(t)), \nabla v_h) + (f_h(u_h(t)), [v_h])_\Gamma = 0 \quad \text{for all } v_h \in V_h^p. \tag{5b}$$

Here $(\cdot,\cdot)$ denotes the L^2 inner product, $(\cdot,\cdot)_\Gamma$ denotes the L^2 inner product on the set of interfaces in Γ, and f_h denotes a given numerical flux function that is uniquely defined on the interfaces of the element decomposition. Detailed assumptions on f_h will be stated below.

Note that due to the fact that both u_h and the test space V_h^p are completely discontinuous, the global definition of the scheme (5b) is equivalent to the following local definition.

$$\frac{d}{dt}(u_j(t), v_j)_{T_j} - (f(u_j(t)), \nabla v_j)_{T_j} + \sum_{l \in N(j)} (f_{jl}(u_j(t), u_l(t)), v_j)_{S_{jl}} = 0, \tag{6}$$
$$\text{for all } v_j \in \mathbb{P}_p, T_j \in \mathcal{T}.$$

Here $(\cdot,\cdot)_{T_j}$, $(\cdot,\cdot)_{S_{jl}}$ denote the local inner product on T_j, S_{jl} respectively, and $f_{jl}(u_j(t), u_l(t))$ is the restriction of $f_h(u_h)$ to S_{jl}. Note that the numerical fluxes $f_{jl}(u_j(t), u_l(t))$ are usually defined via a standard finite difference "upwind - type" one dimensional flux, and it is the only source of "artificial viscosity" in the scheme (5b). We make the following standard assumptions on the numerical flux function.

Assumption 2.2 (Numerical flux function). *The numerical fluxes are supposed to be functions $f_{jl} \in C^1(\mathbb{R}^2, \mathbb{R})$ which satisfies for all $u, v, u', v' \in [A, B]$ the following conditions (respectively: monotony, conservation, regularity, consistency):*

$$\partial_u f_{jl}(u,v) \geq 0, \qquad \partial_v f_{jl}(u,v) \leq 0, \qquad f_{jl}(u,v) = -f_{lj}(v,u), \tag{7a}$$

$$f_{jl}(u,u) = n_{jl}|S_{jl}|\mathbf{f}(u), \qquad |f_{jl}(u,v) - f_{jl}(u',v')| \leq LS_{jl}|(|u-u'| + |v-v'|). \tag{7b}$$

In the literature of DG methods the stabilization due to the "upwinding" of the discrete fluxes is usually accompanied by extra artificial "shock capturing" terms as in [16, 17, 4] or limiting projections as in [9]. (Error estimates for the shock-capturing DG method were obtained in [4]). The RK-DG methods introduced by Cockburn and Shu are based on a combination of limiting projections and Runge Kutta discretization of the ode (5b). Therefore, in the next step we are going to introduce limiting projections in the discretization that will be chosen later in Section 5.

2.1. **Generalized semi-discrete DG approximation.** We introduce a hybrid scheme that incorporates all the characteristics of a RK-DG scheme used with mesh modification with time, but assumes that the ode in time is solved exactly in each time step. To this end we introduce a partition of the time interval $(0,T)$, $\{0 = t^0, ..., t^N = T\}$, and we define the time step $\Delta t^n := t^{n+1} - t^n$. With each time interval $(t^n, t^{n+1}]$ we associate a (possibly different) finite element space V_h^p denoted by

$$V_{h,n}^p := \{v_h \in BV(\mathbb{R}^d) \,|\, v_h|_T \in \mathbb{P}_p \text{ for all } T \in \mathcal{T}_n\}. \tag{8}$$

The associated index set of $\mathcal{T}_n$ is denoted by J^n. In the sequel, when this will not be a source of confusion we might drop the index n in objects related to the finite element space.

To define a local projection operator we proceed as follows: We define $\overline{v_h}$ through $\overline{v_j} := \Pi_{V_h^0}(v)|_{T_j}$ for any $v \in L^2(\Omega)$, i.e., $\overline{v_h}$ is the element wise average of v. Furthermore, with each n we associate projections $\Lambda_h^{n,t}$ with the following properties.

Assumption 2.3 (Projection operator). *The projection $\Lambda_h^{n,t}$ is supposed to be a continuous function with respect to t on the interval $[t^n, t^{n+1}]$. If $t \in (t^n, t^{n+1}]$, the operators act $\Lambda_h^{n,t} : V_{h,n}^p \to V_{h,n}^p$ and satisfy*

$$\overline{\Lambda_h^{n,t}(v_h(\cdot,t))} = \overline{v_h(t)}, \quad t \in (t^n, t^{n+1}]\,. \tag{9}$$

In addition $\Lambda_h^{n,t^n} : V_{h,n-1}^p \to V_{h,n}^p$ is a projection to the new mesh, still with the property

$$\overline{\Lambda_h^{n,t^n}(v_h(\cdot,t^n))} = \overline{v_h(t^n)}\,. \tag{10}$$

In the last equation the element wise average is taken in the new mesh, i.e., corresponds to the projection $\Pi_{V_{h,n}^0}$. At t^n the two operators Λ_h^{n,t^n} and Λ_h^{n-1,t^n} satisfy

$$\|\Lambda_h^{n,t^n}(u_h) - \overline{u_h}\|_\infty \le \|\Lambda_h^{n-1,t^n}(u_h) - \overline{u_h}\|_\infty \tag{11}$$

Properties (9), (10) lead to a conservation of mass, whereas assumption (11) guarantees that the gradients in the discrete solution are not increased between time steps. Note that $\Lambda_h^{n,t}$ accounts for both limiting projections and projections to the new spaces. We define the restriction of $\Lambda_h^{n,t}$ on the element T_j by $\Lambda_j^{n,t}$:

$$\Lambda_j^{n,t} \equiv \Lambda_h^{n,t} \qquad \text{in } T_j \times [t^n, t^{n+1}],\ j \in J^n.$$

We now define the generalized semi-discrete DG approximation.

Definition 2.4 (Generalized semi-discrete DG approximation). *Let us suppose that a projection $\Lambda_h^{n,t}$ with the above properties is given. In addition assume that the discrete fluxes f_{ij} are monotone. The function u_h is called a generalized semi-discrete DG approximation of (1a)-(1b), if for $u_h^{-1} := \Pi_{V_{h,0}^p}(u_0)$ u_h satisfies:*

For $n = 0, ..., N-1$, $u_h^n|_{[t^n,t^{n+1}]} \in C^1(t^n, t^{n+1}; V_{h,n}^p)$ is defined through

$$u_h^n(t^n) \;:=\; \Lambda_h^{n,t^n}(u_h^{n-1}(t^n))\,, \tag{12a}$$

$$\frac{d}{dt}(u_j^n(t), v_j)_{T_j} \;=\; -\sum_{l \in N(j)} (f_{jl}(\Lambda_j^{n,t}(u_h^n(t)), \Lambda_l^{n,t}(u_h^n(t))), v_j)_{S_{jl}} + (f(\Lambda_j^{n,t}(u_h^n(t))), \nabla v_j)_{T_j}, \text{ for all } v_j \in \mathbb{P}_p, j \in J^n, t \in (t^n, t^{n+1})\,. \tag{12b}$$

The global approximation $u_h \in L^\infty(0,T; V_{h,n}^p)$ is defined through $u_h(0) := u_h^{-1}$, and $u_h|_{(t^n,t^{n+1}]} := u_h^n|_{(t^n,t^{n+1}]}$.

In Section 5 we combine the above method with Runge-Kutta time discretizations of the ode (12b) to obtain the generalized class of fully discrete RK-DG methods. This class includes the method of Cockburn and Shu but we consider also alternative choices for the limiting projections motivated by the a posteriori result for (12a)-(12b) proved in the sequel.

2.2. **A–posteriori error estimate for the semi-discrete DG method.** We will show a posteriori estimates for the error $\|(u-u_h)(T)\|_{L^1}$. To do that we compare u and u_h with $\widetilde{u}_h$ defined as:

$$\widetilde{u}_h(t) = \Lambda_h^{n,t}(u_h(t)) \quad \text{for } t \in (t^n, t^{n+1}], \quad n = 0, \dots, N-1 \ . \tag{13}$$

Then $\|(\widetilde{u}_h-u_h)(T)\|_{L^1}$ is an a posteriori quantity and the control of $\|(u-\widetilde{u}_h)(T)\|_{L^1}$ will be obtained in the sequel by employing Kruzkov estimates.

Note that by definition $\widetilde{u}_h$ might be discontinuous at the time nodes t^n. This will be the case either when the spatial mesh is modified at this node, or when we decide to use different projections on $(t^{n-1}, t^n]$ and $(t^n, t^{n+1}]$. In fact due to the definitions of u_h and the projections we have

$$\widetilde{u}_h(t^{n+}) - \widetilde{u}_h(t^n) = \Lambda_h^{n,t^n} u_h(t^n) - \Lambda_h^{n-1,t^n} u_h(t^n) = (\Lambda_h^{n,t^n} - \Lambda_h^{n-1,t^n}) u_h(t^n)\,. \tag{14}$$

Before stating our main result we introduce the following notation:

$$\widetilde{u}_j = \widetilde{u}_h \quad \text{in } T_j, \qquad \widetilde{u}^n = \widetilde{u}_h \quad \text{in } (t^n, t^{n+1}], \qquad \widetilde{u}^n(t^n) = \widetilde{u}_h(t^{n+}),$$

with the obvious extension for combined indexes.

Theorem 2.5 (A–posteriori error estimate for the semi-discrete DG method)**.** *Let u_h be given by the semi-discrete generalized DG method (12a)-(12b). For $\widetilde{u}_h$ given by* (13) *we have the following a–posteriori error estimate*

$$\begin{aligned} ||(u-u_h)(T)||_{L^1(B_R(x_0))} &\le ||(\widetilde{u}_h - u_h)(T)||_{L^1(B_R(x_0))} + ||(u-\widetilde{u}_h)(T)||_{L^1(B_R(x_0))} \\ &\le ||(\widetilde{u}_h - u_h)(T)||_{L^1(B_R(x_0))} + \eta_h \end{aligned}$$

where $\eta_h := \eta_0 + \sqrt{K_1\eta_1} + \sqrt{K_2\eta_2}$, $\eta_0 := \sum_{j\in J^0} \eta_{0,j}$, $\eta_i := \sum_n \sum_{j\in J^n} \eta_{i,j}^n, i = 1, 2,$ *and the local contributions* $\eta_{i,j}^n$ *are given as*

$$\eta_{0,j} := \int_{T_j} |u_0 - \widetilde{u}_j^0(0)|, \tag{15a}$$

$$\begin{aligned} \eta_{1,j}^n := & \int_{t^n}^{t^{n+1}} \int_{T_j} h_j \left|\partial_t \widetilde{u}_j + \nabla\cdot f(\widetilde{u}_j)\right| + \frac{1}{2}\int_{t^n}^{t^{n+1}} \sum_{l\in N(j)} h_{jl} \int_{S_{jl}} Q_{jl}(\widetilde{u}_j, \widetilde{u}_l)|\widetilde{u}_j - \widetilde{u}_l| \\ & + \int_{T_j} h_j |\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n)|, \end{aligned} \tag{15b}$$

$$\begin{aligned} \eta_{2,j}^n := & \int_{t^n}^{t^{n+1}} ||\overline{\widetilde{u}_j^n} - \widetilde{u}_j^n||_{L^\infty(T_j)} \int_{T_j} |\partial_t \widetilde{u}_j^n + \nabla\cdot f(\widetilde{u}_j^n)| \\ & + \frac{1}{2}\int_{t^n}^{t^{n+1}} \sum_{l\in N(j)} \max_{k\in\{j,l\}} ||\overline{\widetilde{u}_k^n} - \widetilde{u}_k^n||_{L^\infty(S_{jl})} \int_{S_{jl}} Q_{jl}(\widetilde{u}_j, \widetilde{u}_l)|\widetilde{u}_j - \widetilde{u}_l| \\ & + ||\overline{\widetilde{u}^{n-1}}(t^n) - \widetilde{u}^{n-1}(t^n)||_{L^\infty(T_j)} \int_{T_j} |\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n)|. \end{aligned} \tag{15c}$$

Here, we used the notation

$$Q_{jl}(u,v) := \frac{2f_{jl}(u,v) - f_{jl}(u,u) - f_{jl}(v,v)}{u-v}, \quad h_{jl} := \mathrm{diam}(T_j \cup T_l).$$

The error estimator in Theorem 2.5 is composed of the two parts η_1, η_2. The first part corresponds to the standard estimates known for first order schemes [5, 20, 26] and the second part of the estimate corresponds to error terms which are only present in higher order approximations. In the following Corollary 2.6 we have rearranged these terms so that the estimate is more suitable for designing an adaptive scheme. The estimate is localized on the control volumes $(t^n, t^{n+1}) \times T_j$. Here the error is split into three terms $R^n_{T,j}, R^n_{S,j}, R^n_{\Lambda,j}$ that account for the different sources of error in the Discontinues Galerkin method. The first term is the element residual found in most error estimates for finite element methods; the second term takes into account the diffusion in the numerical flux function; the third term combines two sources of error between successive time steps, on the one hand side the error due to limiting and on the other hand the error due to grid changes.

Corollary 2.6. *With the assumptions and notations of Theorem 2.5 it follows that*

$$\eta_h \leq \eta_0 + R_h, \quad \text{with} \quad R_h^2 = 2\sum_n \sum_{j\in J^n} \rho_j^n \left(R^n_{T,j} + R^n_{S,j} + R^n_{\Lambda,j} \right) \tag{16}$$

and

$$\rho_j^n := K_1 h_j + K_2 \max_{k\in\{j,\, l\in N(j)\}} ||\overline{\widetilde{u}_k^n} - \widetilde{u}_k^n||_{L^\infty((t^n,t^{n+1})\times T_k)}, \tag{17a}$$

$$R^n_{T,j} := \int_{t^n}^{t^{n+1}} \int_{T_j} \left| \partial_t \widetilde{u}_j + \nabla \cdot f(\widetilde{u}_j) \right|, \tag{17b}$$

$$R^n_{S,j} := \int_{t^n}^{t^{n+1}} \sum_{l\in N(j)} \int_{S_{jl}} Q_{jl}(\widetilde{u}_j, \widetilde{u}_l) |\widetilde{u}_j - \widetilde{u}_l|, \tag{17c}$$

$$R^n_{\Lambda,j} := \int_{T_j} |\widetilde{u}^n(t^{n+1}) - \widetilde{u}^{n+1}(t^{n+1})|. \tag{17d}$$

Proof. The estimate $\eta_h \leq \eta_0 + R_h$ is a direct consequence of Theorem 2.5 if we estimate $\sqrt{K_1\eta_1} + \sqrt{K_2\eta_2} \leq \sqrt{2K_1\eta_1 + 2K_2\eta_2}$. By rearranging terms in $K_1\eta_1 + K_2\eta_2$ and using $h_{jl} \leq h_j + h_l$, equation (16) follows. □

We are now going to discuss some aspects of the error estimate:

Semi-discrete vs. fully discrete estimates. The above a posteriori result is extended in a straightforward manner when the ode (5a) is discretized by Euler's method. RK-DG methods though use high order Runge Kutta schemes for time discretization. We have decided to present the result in the above hybrid semi-discrete scheme since we expect that an a posteriori result for high-order RK-DG methods will be of the form of Theorem 2.5 plus additional terms for the error due to time discretization. The proof of a result for high order Runge-Kutta schemes requires new ideas and is the subject of future work.

The error bound of Theorem 2.5 is used to design our adaptive algorithm in Section 5. To do that we need to introduce in Section 4 the fully discrete generalized

RK-DG methods and to express it as an ode for each time slab $[t^n, t^{n+1}]$. This is done in Section 4 by using the "continuous extension" for Runge Kutta schemes introduced by Zennaro [29].

First order vs. high order estimates. In the case where $p = 0$ the DG method reduces to a standard finite volume scheme that allows mesh modification with n. Then the first term in $\eta^n_{1,j}$ and the whole $\eta^n_{2,j}$ will be zero. The last term in $\eta^n_{1,j}$ will account for coarsening errors due to mesh modification. Such terms were not included in the previous a posteriori estimates for finite volume schemes, [5, 20, 25]. Another implication due to higher order polynomials used in the finite element spaces is the appearance of $\overline{\widetilde{u}^n_j} - \widetilde{u}^n_j$ in various norms in the term $\eta^n_{2,j}$. A comparison with $\eta^n_{1,j}$ leads to the conclusion that it would be desirable to have $\overline{\widetilde{u}^n_j} - \widetilde{u}^n_j = O(h_j)$. In general this is not guaranteed unless $\widetilde{u}^n_j$ is a result of certain limiting projections which restrict gradients or/and polynomial degrees of u_h. This observation is one of the main motivations for the choice of the limiting projections and the design of the adaptive algorithm in Section 5.

Computational "convergence" of the estimators. Theorem 2.5 is a rather general result that covers any projection $\Lambda^{n,t}_h$ with the properties (9) and (10). In addition due to the generality of Kruzkov's estimates used in the proof the above a posteriori bound can be seen as a "worst case scenario" upper bound. It is clear that if in the computational runs the estimators converge to zero then the error will do the same. On the other hand in certain test cases examined in Section 6 it happens also that the estimators computationally do not converge to zero while the error does. In this sense Theorem 2.5 allows for the design of error control algorithms based on upper bound estimates. This issue is discussed in detail in Sections 6 and 5. At this point we would like only to note that, among many other choices presented in Section 5, $h-p$ versions of RK-DG methods seem to allow error control algorithms based on the estimators of Theorem 2.5. On the other hand for the test problems discussed in this paper it seems that for the RK-DG methods with limiters from [8, 6] the estimator of Theorem 2.5 computationally does not converge to zero while the error does. Whether this is a weakness of the bound in Theorem 2.5 or indeed reflects the fact that this method does not converge to the entropy solution in all cases is not clear. The modified limiters based on gradient restrictions suggested in Section 5 address this issue. Concluding, we are able to provide adaptive error control based algorithms for both $h-p$ and gradient restriction generalized versions of DG methods.

The rest of the paper is organized as follows: In Section 3 we prove Theorem 2.5. The proof is based on an abstract Kruzkov estimate for approximations of the entropy solution of the conservation law (Theorem 4.2) and on a weak cell entropy inequality for the method (Lemma 4.3). In Section 4 we present the fully discrete generalized RK-DG method and their continuous in time form with the help of the "continuous extension" for Runge Kutta schemes. In Section 5 we present the limiting projections and the adaptive error control based algorithms for the corresponding DG methods. In the last Section 6 we discuss the computational performance of the various methods in several test cases.

3. Proof of the a–posteriori error estimate

3.1. **Abstract error estimate.** In this subsection we establish an error estimate for approximations of conservation laws. It is an extension to smooth entropies of the corresponding results in [19, 18, 3]. The notation and the form of the result follows [18, Lemma 4.1]. We start with the definition of the entropy residual.

Definition 3.1 (Entropy Residual R_S). *Let $\tilde{u} \in L^\infty(\mathbb{R}^d \times \mathbb{R}^+)$ be an arbitrary function. Then, corresponding to the definition of an entropy weak solution, we define the entropy residual R_S by*

$$\langle R_S(\tilde{u}), \phi\rangle := \iint\limits_{\mathbb{R}^d\times\mathbb{R}^+} S(\tilde{u})\partial_t\phi + F_S(\tilde{u})\cdot\nabla\phi + \int\limits_{\mathbb{R}^d} S(u_0)\phi(\cdot,0). \tag{18}$$

For our a–posteriori error estimate we require a regularization of the standard Kruzkov entropy $S(v) = S(v-k) = |v-k|$.

Definition 3.2 ($\delta-$regularized Kruzkov entropy). *Let $\bar{S} \in C^2(\mathbb{R}, \mathbb{R}^+)$ be given as*

$$\bar{S}(v) = (6v^2 - v^4)/8 \textit{ if } |v| \le 1 \quad \textit{and } \bar{S}(v) = |v| - 3/8 \textit{ otherwise.}$$

For any $\delta > 0, v \in \mathbb{R}$ let us define $S_\delta : \mathbb{R} \to \mathbb{R}^+$ by $S_\delta(v) := \delta\bar{S}(\frac{v}{\delta})$. Furthermore, define $F_{S,\delta} : \mathbb{R}^2 \to \mathbb{R}$ for any $v, k \in \mathbb{R}$ by $F_{S,\delta}(v,k) := \int_\kappa^v f'(w)S'_\delta(w-k)dw$.

In the following result, u_h stands for any approximation problem (1a)-(1b).

Theorem 3.3 (Abstract Kruzkov estimate). *Let $u_h, u \in L^\infty_{loc}([0,\infty), L^1_{loc}(\mathbb{R}^d))$ be right continuous in t, with values in $L^1_{loc}(\mathbb{R}^d)$. Assume that u is the entropy solution of a given conservation law, i.e., it satisfies (1a)-(1b). Let $S(v) = S(v-k) = S_\delta(v-k)$ be the $\delta-$regularized Kruzkov entropy and by $F_S(v) = F_S(v\,,\,k) = F_{S,\delta}(v\,,\,k)$ the corresponding entropy flux. Let Ψ a nonnegative test function $\Psi \in C_c^\infty((0,\infty)\times\mathbb{R}^d)$ and assume that u_h satisfies,*

$$-\langle R_S(u_h), \Psi\rangle = -\iint_{(0,\infty)\times\mathbb{R}^d} \big(S(u_h - k)\partial_t\Psi + F_S(u_h\,,\,k)\cdot\nabla_x\Psi\big)dtdx$$
$$\le \iint_{(0,\infty)\times\mathbb{R}^d} \Big(\beta_O B_O(\Psi) + \alpha_G|\partial_t\Psi| + \sum_j \beta_H^j B_H^j(\frac{\partial\Psi}{\partial x_j})\Big)dxdt, \quad \textit{for all } k \in \mathbb{R},$$

*where α_G, β_O, β_H^j, are nonnegative k-*independent but possibly δ dependent *functions in $L^1_{loc}([0,\infty)\times\mathbb{R}^d)$ and $\alpha_G \in L^\infty_{loc}([0,\infty), L^1_{loc}(\mathbb{R}^d))$.*
For fixed $\Delta, \delta > 0$, let $\mathcal{T}_h = \{K\}$ be a given element decomposition of $[0,\infty)\times\mathbb{R}^d$ into elements K, such that $\text{diam}(K_t) \le \Delta$ *in the case where B_O or B_H^j, is not identically zero; here $K_t = \{x : (t,x) \in K\}$.*

If in addition, for all $(t,x) \in K$, $1 \le i,j \le d$,

$$|B_O(\Psi)(t,x)| \le C \sup_{x'\in K_t} |\Psi(t,x')|\,, \quad |B_H^j(\frac{\partial\Psi}{\partial x_j})(t,x)| \le C \sup_{x'\in K_t} |\frac{\partial\Psi}{\partial x_j}(t,x')|\,, \tag{19}$$

where C is a uniform constant independent of Ψ and the element decomposition $\mathcal{T}_h$, then the following estimate holds: for any $T \ge 0$, $x_0 \in \mathbb{R}^d$, $R > 0$, $\rho > 0$ with $M = Lip(f)$, we have:

$$\int_{|x-x_0|<R} |u_h(T,x) - u(T,x)|dx \le \int_{B_0} |u_h(0,x) - u(0,x)|dx + C\,(M\,+1)TV(u^0)\;\Delta$$
$$+ C\{k_1 TV(u^0) + k_0\,\chi_{\mathrm{supp}(u_h-u)(T)}\,(R+\Delta)^d\,\}\,\delta$$
$$+ C(1+\tfrac{T(1+M)}{\Delta})\sup_{0\le t\le T+\rho}\int_{B_t}\alpha_G(t,x)dx$$
$$+ C\iint_{0\le t\le T\; x\in B_t^\Delta}\Big(\beta_O(t,x) + \frac{1}{\Delta}\sum_{j=1}^d \beta_H^j(t,x)\Big)dxdt\,,$$

where $B_t = B(x_0, R + M(T-t) + \Delta)$*,* $B_t^\Delta = B(x_0, R + M(T-t) + 2\Delta)$*, and* χ_D *denotes the characteristic function of the set* D*.*

Proof. The proof follows [3, 19, 24]: we seek two nonnegative functions $\Phi, \zeta \in C_c^\infty((0,\infty)\times\mathbb{R}^d)$ that will be specified later and we set

$$\phi(t,x,s,y) = \Phi(t,x)\zeta(t-s, x-y)\,.$$

Then the approximate inequality for $\Psi = \phi(\cdot,\cdot,s,y)$ with fixed $(s,y)\in(0,\infty)\times\mathbb{R}^d$ and $k = u(s,y)$ yields

$$-\iiiint\Big[S(u_h(t,x) - u(s,y)\,)\partial_t\phi(t,x,s,y) + F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\phi(t,x,s,y)\Big]dsdtdydx \le R^\beta\,, \tag{20}$$

where

$$R^\beta = \iiiint \beta_O(t,x)B_O(\phi(t,x,s,y)) + \alpha_G(t,x)|\partial_t\phi(t,x,s,y)| + \sum_j \beta_H^j(t,x)B_H^j(\partial_{x_j}\phi(t,x,s,y))\,dsdtdydx\,. \tag{21}$$

Notice that B_O, B_H^j act in (t,x). Next, using the fact $\partial_t\zeta = -\partial_s\zeta$, $\nabla_x\zeta = -\nabla_y\zeta$, we obtain

$$\begin{aligned}
-\Big[&S(u_h(t,x) - u(s,y)\,)\partial_t\phi(t,x,s,y) + F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\phi(t,x,s,y)\Big]\\
&= -\Big[S(u_h(t,x) - u(s,y)\,)\partial_t\Phi(t,x)\\
&\qquad + F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\Phi(t,x)\Big]\zeta(t-s,x-y)\\
&\quad -\Big[S(u_h(t,x) - u(s,y)\,)\partial_t\zeta(t-s,x-y)\\
&\qquad + F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\zeta(t-s,x-y)\Big]\Phi(t,x)\\
&= -\Big[S(u_h(t,x) - u(s,y)\,)\partial_t\Phi(t,x)\\
&\qquad + F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\Phi(t,x)\Big]\zeta(t-s,x-y)\\
&\quad +\Big[F_S(u_h(t,x)\,,\,u(s,y)) - F_S(u(s,y)\,,\,u_h(t,x))\cdot\nabla_y\zeta(t-s,x-y)\Big]\Phi(t,x)\\
&\quad +\Big[S(u(s,y) - u_h(t,x)\,)\partial_s\zeta(t-s,x-y)
\end{aligned}$$

$$+F_S(u(s,y)\,,\,u_h(t,x))\cdot\nabla_y\zeta(t-s,x-y)\Big]\Phi(t,x).$$

We emphasize the presence of the second term in the last equality which is due to the lack of symmetry of the regularized entropy flux $F_S(v\,,\,w)$. Integrating with respect to all variables and noticing then that the last term is positive by the fact that u is the entropy solution of the conservation law, we get

(22)

$$-\iiiint\Big[S(u_h(t,x)-u(s,y)\,)\partial_t\Phi(t,x)\\ +F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\Phi(t,x)\Big]\zeta(t-s,x-y)dsdtdydx\le R^\beta+R^S\,,$$

where

$$R^S=-\iiiint\Big[F_S(u_h(t,x)\,,\,u(s,y))\\ -F_S(u(s,y)\,,\,u_h(t,x))\Big]\cdot\nabla_y\zeta(t-s,x-y)\Phi(t,x)dsdtdydx\,. \tag{23}$$

We proceed with a specific choice of $\Phi(t,x)$. Let $\theta>0$, we define $Y_\theta(t)$ so that $Y_\theta(-\infty)=0$ and $Y'_\theta(t)=\frac{1}{\theta}Y'(\frac{t}{\theta})$, where $Y'\in C_c^\infty((0,1))$, $Y'\ge 0$ and $\int Y'=1$. For a new parameter $\varepsilon>0$ we set

$$\begin{aligned} &X(t)=Y_\varepsilon(t)-Y_\varepsilon(t-T)\ \in C_c^\infty((0,T+\varepsilon)),\ X\ge 0\\ &X_0(t)=Y_\varepsilon(t),\ X_T(t)=Y_\varepsilon(t-T),\ X'_0,X'_T\ \ge 0\,. \end{aligned} \tag{24}$$

Finally we define $\psi(t,x)=1-Y_\theta\big(|x-x_0|-R-\Delta/2-M(T-t)\big)\ge 0$. We now set

$$\Phi(t,x)=X(t)\psi(t,x)\,.$$

Notice that $\Phi\in C^\infty$ as long as $M\varepsilon\le R+\Delta/2$.

In the sequel it will be crucial to note that the entropy flux satisfies

$$|F_S(v\,,\,w)\,|\le M\,S(v-w)\,, \tag{25}$$

where M is the Lipschitz constant of f. Indeed, in the case where $v\ge w$,

$$|F_S(v\,,\,w)\,|=|\int_w^v f'(s)S'(s-w)\,ds\,|\le M\int_w^v S'(s-w)\,ds=M\,S(v-w)\,. \tag{26}$$

Similarly for $v\le w$. Using the specific choice of Φ, the Lipschitz condition on f, and (25) we have

$$\begin{aligned} S(u_h(t,x)-u(s,y))\partial_t\Phi(t,x)&+F_S(u_h(t,x)\,,\,u(s,y))\cdot\nabla_x\Phi(t,x)\\ &=S(u_h(t,x)-u(s,y))X'(t)\psi(t,x)\\ &\quad-X(t)Y'_\theta\Big\{MS(u_h(t,x)-u(s,y))-F_S(u_h(t,x)\,,\,u(s,y))\cdot\tfrac{x-x_0}{|x-x_0|}\Big\}\\ &\le S(u_h(t,x)-u(s,y))\,. \end{aligned}$$

Therefore

$$-\iiiint S(u_h(t,x)-u(s,y))X'(t)\psi(t,x)\zeta(t-s,x-y)dsdtdydx\le R^\beta+R^S\,.$$

We add and subtract $S(u_h(t,x)-u(t,y)), S(u_h(t,x)-u(t,x))$. Using the fact that the Lipschitz constant of S is 1 we finally get

$$-\iiiint S(u_h(t,x)-u(t,x))X'(t)\psi(t,x)\zeta(t-s,x-y)dsdtdydx\le R^t+R^x+R^\beta+R^S\,,$$

where

$$R^x = \iiiint |u(t,x) - u(t,y)|\, |X'(t)| \psi(t,x)\zeta(t-s, x-y) dsdtdydx\,,$$
$$R^t = \iiiint |u(t,y) - u(s,y)|\, |X'(t)| \psi(t,x)\zeta(t-s, x-y) dsdtdydx\,.$$

Then notice that in view of the definitions of the regularized entropy S and of X we have

$$\begin{aligned}
-S(&u_h(t,x) - u(t,x))X'(t)\psi(t,x)\\
&= -S(u_h(t,x) - u(t,x))X_0'(t)\psi(t,x) + S(u_h(t,x) - u(t,x))X_T'(t)\psi(t,x)\\
&\geq -|u_h(t,x) - u(t,x)|X_0'(t)\psi(t,x) + |u_h(t,x) - u(t,x)|X_T'(t)\psi(t,x)\\
&\quad - \delta\, k_0 \chi_{\mathrm{supp}(u_h-u)(t,x)}\, X_T'(t)\,.
\end{aligned}$$

Note that for θ fixed and ε small,

$$1_{B(x_0, R+M(T-t)+\Delta/2)} \leq \psi(t,x) \leq 1_{B(x_0, R+M(T-t)+\Delta/2+\theta}\,.$$

We pass to the $\varepsilon \to 0$ limit with the choice $\theta = \Delta/4$. As in [21, 3] we have

$$0 \leq \limsup_{\varepsilon\to 0} \Big\{ \iiiint S(u_h(t,x) - u(t,x))X'(t)\psi(t,x)\zeta(t-s,x-y)dsdtdydx$$
$$+R^t + R^x + R^\beta + R^S \Big\}$$

where

$$\begin{aligned}
\limsup_{\varepsilon\to 0} &\iiiint S(u_h(t,x) - u(t,x))X'(t)\psi(t,x)\zeta(t-s,x-y)dsdtdydx\\
&\leq \int_{|x-x_0|<R+MT+\Delta/2+\theta} |u_h(0,x) - u(0,x)|dx + \int_{|x-x_0|<R+\Delta/2} \delta\, k_0 \chi_{\mathrm{supp}(u_h-u)(T,x)} dx\\
&- \int_{|x-x_0|<R+\Delta/2} |u_h(T,x) - u(T,x)|dx,
\end{aligned}$$

and

$$\limsup_{\varepsilon\to 0} R^t \leq 2E^t, \quad \limsup_{\varepsilon\to 0} R^x \leq 2E^x\,.$$

The term R^β is bounded as in [19, 3]. The last term R^S is handled as in [5, 24]. Indeed integrating by parts with respect to y and using the properties of the $\delta-$regularized entropy flux we get

$$|R^S| \leq C\|F''\|_\infty\, k_1 TV(u^0)\,\delta.$$

The proof is therefore complete. □

3.2. **Estimate on the entropy residual.** To apply the abstract theorem of the previous subsection we need to estimate

$$\langle R_S(\tilde{u}_h), \phi\rangle$$

for ϕ being a test function and $\tilde{u}_h$ defined in (13). This will be done in the following lemmas. We start with a *weak* discrete local entropy inequality:

Lemma 3.4 (Weak cell entropy inequality for the semi discrete DG approximation)**.** *Let* (S, F_S) *denote a smooth entropy pair. Then the following cell entropy inequality holds for* $\widetilde{u}_h$*.*

$$I_j^n := I_{1,j}^n + I_{2,j}^n + I_{3,j}^n + I_{4,j}^n = -D_j^n \leq 0, \tag{27}$$

where for $\phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)$,

$$\begin{aligned}
I_{1,j}^n &= (\partial_t S(\widetilde{u}_j^n) + \nabla \cdot F_S(\widetilde{u}_j^n), \overline{\phi_j})_{T_j}, \\
I_{2,j}^n &= \sum_{l \in N(j)} \Big(F_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - F_S(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}, \overline{\phi_j} \Big)_{S_{jl}}, \\
I_{3,j}^n &= (\partial_t \widetilde{u}_j^n + \nabla \cdot f(\widetilde{u}_j^n), (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j})_{T_j}, \\
I_{4,j}^n &= \sum_{l \in N(j)} \Big(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}, (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j} \Big)_{S_{jl}}, \\
D_j^n &= \sum_{l \in N(j)} \Big(\int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) \int_w^{\widetilde{u}_j^n} S''(s)\, ds dw, \overline{\phi_j} \Big)_{S_{jl}}.
\end{aligned}$$

Here $F_{jl}(\alpha, \beta) = \int_\alpha^\beta \partial_s f_{jl}(\alpha, s) S'(s) ds + F_S(\alpha)$ *is a discrete entropy flux that is consistent with* F_S*.*

Proof. Let $p \geq 0$. We start by choosing $v_h = S'(\overline{\widetilde{u}_h^n})\overline{\phi_h}$ in the local form the scheme (12b). This yields

$$(\partial_t u_j^n, S'(\overline{\widetilde{u}_j^n})\overline{\phi_j})_{T_j} + \sum_{l \in N(j)} (f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n), S'(\overline{\widetilde{u}_j^n})\overline{\phi_j})_{S_{jl}} = 0.$$

Next, (9) implies $(\partial_t u_j^n, v_j)_{T_j} = (\partial_t \widetilde{u}_j^n, v_j)_{T_j}$ for all $v_h \in V_{h,n}^0$. Therefore

$$(\partial_t \widetilde{u}_j^n, S'(\overline{\widetilde{u}_j^n})\overline{\phi_j})_{T_j} + \sum_{l \in N(j)} (f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n), S'(\overline{\widetilde{u}_j^n})\overline{\phi_j})_{S_{jl}} = 0.$$

We insert zeros to get

$$\begin{aligned}
0 = &(\partial_t \widetilde{u}_j^n + \nabla \cdot f(\widetilde{u}_j^n), S'(\widetilde{u}_j^n)\overline{\phi_j})_{T_j} \\
&+ \sum_{l \in N(j)} \Big(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}, S'(\widetilde{u}_j^n)\overline{\phi_j} \Big)_{S_{jl}} \\
&+ (\partial_t \widetilde{u}_j^n + \nabla \cdot f(\widetilde{u}_j^n), (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j})_{T_j} \\
&+ \sum_{l \in N(j)} \Big(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}, (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j} \Big)_{S_{jl}}.
\end{aligned}$$

To complete the proof, it remains to show that

$$\begin{aligned}
&(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}) S'(\widetilde{u}_j^n) \\
= \ &F_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - F_S(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl} + \int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) \int_w^{\widetilde{u}_j^n} S''(s) \mathrm{dsdw}.
\end{aligned}$$

Indeed,

$$\begin{aligned}
&(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}) S'(\widetilde{u}_j^n) = (f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f_{jl}(\widetilde{u}_j^n, \widetilde{u}_j^n)) S'(\widetilde{u}_j^n) \\
= &\int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) \mathrm{dw} S'(\widetilde{u}_j^n)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) S'(w) dw + \int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) \big(S'(\widetilde{u}_j^n) - S'(w)\big) \mathrm{dw} \\
&= F_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - F_{jl}(\widetilde{u}_j^n, \widetilde{u}_j^n) + \int_{\widetilde{u}_j^n}^{\widetilde{u}_l^n} \partial_w f_{jl}(\widetilde{u}_j^n, w) \int_w^{\widetilde{u}_j^n} S''(s) \mathrm{dsdw}\,.
\end{aligned}
$$

□

Remark 3.5. *1) Note that the dissipation term D_j^n in the cell entropy inequality (27) is positive because of the monotonicity of the numerical flux and the convexity of S. 2) If $p = 0$, we have $I_{3,j}^n, I_{4,j}^n = 0$. 3) Also note that, as expected for a high order scheme, the weak cell entropy like inequality is, in general, not a real cell entropy inequality in the classical sense. But its use is important to conclude the following estimates, compare to* [13].

Lemma 3.6 (Entropy residual for the semi-discrete DG approximation). *Let u_h, $\widetilde{u}_h$ as before. Then, the following identity holds true for all $\phi \in C_0^1(\mathbb{R}^d \times (0,T), \mathbb{R}^+)$*

$$
\langle R_S(u_h), \phi \rangle \geq T_1 + T_2 + T_3 + T_4 + T_5 + T_6, \tag{28}
$$

where

$$
\begin{aligned}
T_1 &:= \iint\limits_{\mathbb{R}^d \times \mathbb{R}^+} \Big(\partial_t S(\widetilde{u}_h) + \nabla \cdot F_S(\widetilde{u}_h)\Big)\Big(\overline{\phi_h} - \phi\Big), \\
T_2 &:= \int_0^T \sum_{j \in J^n} \sum_{l \in N(j)} \Big(F_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - F_S(\widetilde{u}_j^n), \overline{\phi_h} - \phi\Big)_{S_{jl}}, \\
T_3 &:= \sum_n \int\limits_{t^n}^{t^{n+1}} \sum_{j \in J^n} \Big(\partial_t \widetilde{u}_j^n + \nabla \cdot f(\widetilde{u}_j^n), (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j}\Big)_{T_j}, \\
T_4 &:= \sum_n \int\limits_{t^n}^{t^{n+1}} \sum_{j \in J^n} \sum_{l \in N(j)} \Big(f_{jl}(\widetilde{u}_j^n, \widetilde{u}_l^n) - f(\widetilde{u}_j^n) \cdot \mathbf{n}_{jl}, (S'(\overline{\widetilde{u}_j^n}) - S'(\widetilde{u}_j^n))\overline{\phi_j}\Big)_{S_{jl}}, \\
T_5 &:= \sum_n \sum_{j \in J^n} \Big(\widetilde{u}_j^n(t^n) - \widetilde{u}_j^{n-1}(t^n), \int_0^1 S'(u_j(\theta)) d\theta \; (\overline{\phi_h}(t^n) - \phi(t^n))\Big)_{T_j}, \\
T_6 &:= \sum_n \sum_{j \in J^n} \Big(\widetilde{u}_j^n(t^n) - \widetilde{u}_j^{n-1}(t^n), (S'(\overline{\widetilde{u}_j^{n-1}(t^n)}) - \int_0^1 S'(v^n(\theta)) d\theta) \; \overline{\phi_h}(t^n)\Big)_{T_j}
\end{aligned}
$$

where in the definition of T_5 and T_6 we use the abbreviation

$$
v^n(\theta) := \widetilde{u}^{n-1}(t^n) + \theta(\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n)).
$$

Note that T_1 is the element residual and T_2 is the jump residual in space. T_3 and T_4 are to be seen as a kind of stability errors coming from the higher order approximation and T_5, T_6 account for possible discontinuities in time of the projected function $\widetilde{u}_h$.

Proof. A summation of the cell entropy like inequality (27) on all elements $T_j \in \mathcal{T}_n$ and an integration in time leads to

$$
I_h \;:=\; \sum_n \int_{t^n}^{t^{n+1}} \sum_{j \in J^n} \Big(I_{1,j}^n + I_{2,j}^n + I_{3,j}^n + I_{4,j}^n\Big) \leq 0.
$$

Next, let us look at the entropy residual. Using integration by parts in time and locally in space we get

$$
\begin{aligned}
\langle R_S(\widetilde{u}_h),\phi\rangle \;=\; & -\sum_n \int_{t^n}^{t^{n+1}} \sum_{j\in J^n} \int_{T_j} \partial_t S(\widetilde{u}_j^n)\phi + \nabla\cdot F_S(\widetilde{u}_j^n)\phi \\
& +\sum_n \int_{t^n}^{t^{n+1}} \sum_{j\in J^n} \sum_{l\in N(j)} \int_{S_{jl}} F(\widetilde{u}_j^n)\cdot \mathbf{n}_{jl}\phi \\
& +\sum_n \sum_{j\in J^n} \int_{T_j} \left(S(\widetilde{u}_j^n)(t^{n+1})\phi(t^{n+1}) - S(\widetilde{u}_j^n)(t^n)\right)\phi(t^n)\right).
\end{aligned}
$$

Noting that due to the conservation property of the numerical flux and since ϕ is continuous we have $\int_0^T \sum\limits_{j\in J^n} \sum\limits_{l\in N(j)} \int_{S_{jl}} F_{jl}(\widetilde{u}_j^n,\widetilde{u}_l^n)\phi = 0$, and by rearranging the summation we get.

$$
\begin{aligned}
\langle R_S(\widetilde{u}_h),\phi\rangle \;=\; & -\sum_n \int_{t^n}^{t^{n+1}} \sum_{j\in J^n} \Big(\partial_t S(\widetilde{u}_j^n) + \nabla\cdot F_S(\widetilde{u}_j^n),\phi\Big)_{T_j} \\
& +\sum_n \int_{t^n}^{t^{n+1}} \sum_{j\in J^n} \sum_{l\in N(j)} \Big(F_{jl}(\widetilde{u}_j^n,\widetilde{u}_l^n) - F(\widetilde{u}_j^n)\cdot\mathbf{n}_{jl},\phi\Big)_{S_{jl}} \\
& -\sum_n \sum_{j\in J^n} \Big(S(\widetilde{u}^n)(t^n) - S(\widetilde{u}^{n-1})(t^n),\phi(t^n)\Big)_{T_j}. \qquad (29)
\end{aligned}
$$

Note that, regarding the last term, since the above sums are reduced to integrals over the spatial domain, we have decided to split them in sums over $T_j \in \mathcal{T}_n$ although $\widetilde{u}^{n-1}(t^n) \in V_{h,n-1}^p$. Next, using the property (10) of the projections we obtain

$$
\sum_n \sum_{j\in J^n} \Big(\widetilde{u}_j^n(t^n) - \widetilde{u}_j^{n-1}(t^n), S'(\overline{\widetilde{u}_j^{n-1}(t^n)})\overline{\phi_h}(t^n)\Big)_{T_j} = 0.
$$

Defining $v^n(\theta) := \widetilde{u}^{n-1}(t^n) + \theta(\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n))$ we can rewrite the last summand (29) of the residual as follows

$$
\begin{aligned}
& -\sum_n \sum_{j\in J^n} \Big(S(\widetilde{u}^n)(t^n) - S(\widetilde{u}^{n-1})(t^n),\phi(t^n)\Big)_{T_j} \\
&= -\sum_n \sum_{j\in J^n} \Big(S(\widetilde{u}^n)(t^n) - S(\widetilde{u}^{n-1})(t^n),\phi(t^n)\Big)_{T_j} \\
&\quad +\sum_n \sum_{j\in J^n} \Big((\widetilde{u}^n - \widetilde{u}^{n-1})(t^n), S'(\overline{\widetilde{u}^{n-1}(t^n)})\overline{\phi_h}(t^n)\Big)_{T_j} \\
&= \sum_n \sum_{j\in J^n} \Big((\widetilde{u}^n - \widetilde{u}^{n-1})(t^n), S'(\overline{\widetilde{u}^{n-1}(t^n)})\overline{\phi_h}(t^n) - \int_0^1 S'(v^n(\theta))\mathrm{d}\theta\phi(t^n)\Big)_{T_j} \\
&= \sum_n \sum_{j\in J^n} \Big((\widetilde{u}^n - \widetilde{u}^{n-1})(t^n), \int_0^1 S'(v^n(\theta))\mathrm{d}\theta\; (\overline{\phi_h}(t^n) - \phi(t^n))\Big)_{T_j}
\end{aligned}
$$

$$+\sum_n \sum_{j\in J^n} \Big((\widetilde{u}^n - \widetilde{u}^{n-1})(t^n), (S'(\overline{\widetilde{u}^{n-1}(t^n)}) - \int_0^1 S'(v^n(\theta))\mathrm{d}\theta)\ \overline{\phi_h}(t^n) \Big)_{T_j},$$

Finally by comparing the Residual with I_h we arrive at the final result noticing

$$\begin{aligned} \langle R_S(u_h), \phi\rangle &\geq \langle R_S(u_h), \phi\rangle + I_h \\ &= T_1 + T_2 + T_3 + T_4 + T_5 + T_6. \end{aligned}$$

□

We conclude by further estimating the T_i terms in (28).

Lemma 3.7 (Estimate on the residual). *The following estimates on the contributions to the residual hold true*

$$\begin{aligned} |T_1 + T_2 + T_5| &\leq ||S'||_{L^\infty} ||\nabla\phi||_{L^\infty} \sum_n \sum_{j\in J^n} \eta^n_{1,j}, \\ |T_3 + T_4 + T_6| &\leq ||S''||_{L^\infty} ||\phi||_{L^\infty} \sum_n \sum_{j\in J^n} \eta^n_{2,j}, \end{aligned}$$

where the local error indicators $\eta^n_{i,j}$ are defined in Theorem 2.5 above.

Proof. The goal of estimating the terms T_i is to get some power of the mesh size h from the differences in the test functions. On the other hand, every power of h that we might gain must be paid for by a higher derivative of either ϕ or S. Since we will later choose ϕ to approximate certain δ-functions and S to approximate the non-smooth Kruzkov entropies, derivatives of ϕ and S' will blow up with a certain rate, depending on the corresponding approximation parameters. The goal is therefore, to restrict to first derivatives of ϕ and second derivatives of S.

Estimate on T_1: To estimate this term, we need some local estimate on the difference of the test functions. As $\Pi_{V_h^0}$ is exact on polynomials of degree $p = 0$, we get

$$|(\Pi_{V_h^0}(\phi)(x) - \phi(x))|_{T_j}| \leq h_j ||\nabla\phi||_{L^\infty(T_j)}$$

which finally leads to the estimate on T_1.

Estimate on T_2: We get by rearranging the summation in space:

$$\begin{aligned} T_2 &= \int_0^T \sum_{(j,l)\in\mathcal{E}^n} \int_{S_{jl}} (F_{jl}(\widetilde{u}_j, \widetilde{u}_l) - F_{jl}(\widetilde{u}_j, \widetilde{u}_j)(\overline{\phi_j} - \phi) \\ &\quad - \int_0^T \sum_{(j,l)\in\mathcal{E}^n} \int_{S_{jl}} (F_{jl}(\widetilde{u}_j, \widetilde{u}_l) - F_{jl}(\widetilde{u}_l, \widetilde{u}_l)(\overline{\phi_l} - \phi). \end{aligned}$$

From here we get the estimate

$$\begin{aligned} |T_2| &\leq ||S'||_{L^\infty} \sum_n \int_{t^n}^{t^{n+1}} \sum_{(j,l)\in\mathcal{E}^n} h_{jl} ||\nabla\phi||_{L^\infty(T_j\cup T_l)} \int_{S_{jl}} Q_{jl}(\widetilde{u}_j, \widetilde{u}_l) |\widetilde{u}_j - \widetilde{u}_l| \\ &\leq ||S'||_{L^\infty} ||\nabla\phi||_{L^\infty} \sum_n \sum_{j\in J^n} \frac{1}{2} \int_{t^n}^{t^{n+1}} \sum_{l\in N(j)} h_{jl} \int_{S_{jl}} Q_{jl}(\widetilde{u}_j, \widetilde{u}_l) |\widetilde{u}_j - \widetilde{u}_l| \end{aligned}$$

where we made use of the monotonicity of the numerical fluxes f_{jl}.

Analog to the estimates for T_1, T_2 we derive for T_3 and T_4:

$$
\begin{aligned}
|T_3| &\leq ||S''||_{L^\infty}||\phi||_{L^\infty}\sum_n\sum_{j\in J^n}\int_{t^n}^{t^{n+1}} ||\overline{\widetilde{u}_j^n} - \widetilde{u}_j^n||_{L^\infty(T_j)}\int_{T_j}|\partial_t\widetilde{u}_h + \nabla\cdot f(\widetilde{u}_h)|,\\
|T_4| &\leq ||S''||_{L^\infty}||\phi||_{L^\infty}\sum_n\sum_{j\in J^n}\frac{1}{2}\int_{t^n}^{t^{n+1}}\sum_{l\in N(j)}\max_{k\in\{j,l\}}||\overline{\widetilde{u}_k^n} - \widetilde{u}_k^n||_{L^\infty(S_{jl})}\\
&\qquad \int_{S_{jl}} Q_{jl}(\widetilde{u}_j,\widetilde{u}_l)|\widetilde{u}_j - \widetilde{u}_l|.
\end{aligned}
$$

Estimate on T_5 and T_6:

$$
\begin{aligned}
|T_5| &\leq ||S'||_{L^\infty}||\nabla\phi(t^n)||_{L^\infty(T_j)}\sum_n\sum_{j\in J^n} h_j\int_{T_j}|\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n)|,\\
|T_6| &\leq ||S''||_{L^\infty}||\phi(t^{n+1})||_{L^\infty(T_j)}\\
&\qquad \sum_n\sum_{j\in J^n}||\overline{\widetilde{u}_j^{n-1}}(t^n) - \widetilde{u}_j^{n-1}(t^n)||_{L^\infty(T_j)}\int_{T_j}|\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n)|.
\end{aligned}
$$

For the last estimate we used

$$
\begin{aligned}
&|S'(\overline{\widetilde{u}_j^{n-1}(t^n)}) - \int_0^1 S'(v^n(\theta))\mathrm{d}\theta|\\
&\qquad \leq ||S''||_{L^\infty}\int_0^1|\overline{\widetilde{u}^{n-1}(t^n)} - \widetilde{u}^{n-1}(t^n) + \theta(\widetilde{u}^n(t^n) - \widetilde{u}^{n-1}(t^n))|\mathrm{d}\theta
\end{aligned}
$$

which gives us the bound on T_6 since the function under the integral is monotone decreasing in θ due to (11).

The estimate of the Theorem now follows by introducing the notation from Theorem 2.5. □

We are ready now to complete the proof of Theorem 2.5.

Proof of Theorem 2.5. Lemma 3.7 shows that $\widetilde{u}_h$ satisfies the assumption of Theorem 3.3 with $\alpha_G := 0$ and $\beta_O, \sum_k \beta_H^k$ given by the following local contributions

$$
\begin{aligned}
\beta_O|_{T_j\times[t^n,t^{n+1})} &:= \frac{1}{\Delta t\ |T_j|}||S''||_{L^\infty}\eta_{2,j}^n,\\
\sum_k \beta_H^k|_{T_j\times[t^n,t^{n+1})} &:= \frac{1}{\Delta t\ |T_j|}\ \eta_{1,j}^n.
\end{aligned}
$$

Using the definition of the entropy $S = S_\delta$ (see Def.3.2) we estimate $||S''||_{L^\infty} \leq K_S\frac{1}{\delta}$. Theorem 2.5 now follows from Theorem 3.3 by choosing the regularization parameters Δ, δ as

$$
\Delta := \sqrt{\frac{\sum_n\sum_{j\in J^n}\eta_{1,j}^n}{K_1}}, \qquad \delta := \sqrt{\frac{\sum_n\sum_{j\in J^n}\eta_{2,j}^n}{K_2}} \tag{30}
$$

where $K_1 := (M+1)TV(u^0), K_2 := K_S^{-1}(k_1 TV(u^0) + k_0\,\chi_{\mathrm{supp}(u_h-u)(T)}\,(R+1)^d)$.
□

4. Fully discrete RK-DG method and continuous in time extension

In this section we will briefly present the generalized class of RK-DG methods that result from the full discretization of the hybrid semi-discrete method of Definition 2.4 analyzed in the previous sections. This is a variation of the RK-DG methods of Cockburn and Shu [9].

For the time discretization, the known class of strongly stability preserving Runge Kutta methods is used, [9]. Thus let us suppose that we can write the semi-discrete DG method as a system of ODEs for a vector valued function $U : (0,T) \to \mathbb{R}^N$, where N corresponds to the degrees of freedom of u_h in the space discretization. Then (5b) can be written in the general form:

$$\frac{\mathrm{d}}{\mathrm{dt}} U(t) = L(U(t),t). \tag{31}$$

A general explicit m stage Runge-Kutta method for integrating (31) in time can be algorithmically represented as

$$W^l := U^n + \Delta t \sum_{k=1}^{l-1} a_{lk} L^k \;, \quad l = 1, ..., m, \tag{32a}$$

$$L^l := L(W^l, t^n + c_l \Delta t) \;, \quad l = 1, ..., m, \tag{32b}$$

$$U^{n+1} := U^n + \Delta t \sum_{k=1}^{m} b_k L^k \;. \tag{32c}$$

To ensure consistency, the additional constraints $\sum_{k=1}^{m} b_k = 1$, $c_l = \sum_{k=1}^{l-1} a_{lk} \in [0,1]$ have do be imposed. The scheme is characterized by the values b_k, $k = 1, ..., m$ and a lower triangular matrix a_{lk}, $l = 2, ..., m$, $k < l$.

For particular strongly stability preserving (SSP) explicit Runge Kutta methods we refer to [28] and to the review articles [14, 27].

The Runge-Kutta method, as presented above, gives only approximations at the discrete time steps t^n. In order to obtain a continuous approximation in time, we seek for a polynomial approximation U_h in time, such that in each interval $[t^n, t^{n+1}]$ the Runge-Kutta scheme can be written in the form

$$\frac{\mathrm{d}}{\mathrm{dt}} U_h(t) = L_h(U_h(t),t).$$

A way to construct such polynomials is given by the so called natural continuous extension (NCE) of Runge-Kutta methods, introduced by Zennaro [29]. Since this construction is used in the numerical examples we present it below.

The main result of [29] can be summarized as follows. Each m-stage Runge Kutta method of order $\tilde{m}$ has a natural continuous extension U_h of polynomial degree $\tilde{p}$ with $\frac{\tilde{m}+1}{2} \le \tilde{p} \le \min\{m^*, \tilde{m}\}$, where m^* is the number of distinct values of the coefficients c_l, in the sense that there exists m polynomials $b_l \in \mathbb{P}^{\tilde{p}}(0,1)$, $l = 1, ..., m$ such that

$$U_h(t^n) = U^n, \quad U_h(t^{n+1}) = U^{n+1},$$

$$U_h(t^n + s\Delta t) := U^n + \Delta t \sum_{k=1}^{m} b_k(s) L^k, \quad 0 \le s \le 1. \tag{33}$$

Let us suppose that U_h is given by the m-stage NCE Runge Kutta scheme (for an explicit construction see [29]). From (33) we get

$$U_h(t^n) = U^n \ , \tag{34}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}U_h(t) = \sum_{k=1}^{m} b'_k(\frac{t-t^n}{\Delta t})L^k, \quad t \in [t^n, t^{n+1}]. \tag{35}$$

Defining the discrete Operator L_h as

$$L_h(U_h, t) := \sum_{k=1}^{m} b'_k(\frac{t-t^n}{\Delta t})L^k,$$

we have reached our desired goal. The Runge Kutta method now writes

$$U_h(t^n) = U^n \ , \qquad \frac{\mathrm{d}}{\mathrm{dt}}U_h(t) = L_h(U_h, t). \tag{36}$$

Let us further define the fully-discrete RKDG-method in a form equivalent to (36). Starting from Definition 2.4 we first define the local operators L_j^n by

$$\langle L_j^n(u_h(t)), v_j\rangle|_{T_j} := (f(u_j^n(t)), \nabla v_j)_{T_j} - \sum_{l\in N(j)} (f_{jl}(u_j^n(t), u_l^n(t)), v_j)_{S_{jl}}, \tag{37}$$

$$\text{for all } T_j \in \mathcal{T}_h, n = 0, ..., N, v_h \in V_h^p.$$

and L_h^n through $\langle L^n(u_h(t)), v_h\rangle := \sum_{j\in J^n} \langle L_j^n(u_h(t)), v_j\rangle|_{T_j}$

Using the notation from above, the fully-discrete generalized Runge Kutta–DG method reads.

Definition 4.1 (Fully-discrete generalized RK–DG approximation)**.** *Let an m-stage Runge Kutta method be given according to (32a) (32b), (32c) and let us suppose that a projection $\Lambda_h^{n,t}$ with the properties (9), (10) is given. Furthermore, let the natural continuous extension of highest possible degree $\tilde{p}$ be given according to (33). Let us denote $\Lambda_h^{n,k} := \Lambda_h^{n,t^n+c_k\Delta t}$ for $k = 1, ..., m$.*

The function U_h is called a generalized fully-discrete RK–DG approximation of (1a)-(1b), if for $U_h^{-1} := \Lambda_h^{0,0}(u_0)$ it satisfies:

For $n = 0, ..., N-1$, $U_h^n := U_h|_{(t^n,t^{n+1}]} \in C^1(t^n, t^{n+1}; V_{h,n}^p)$ is defined through

$$U_h^n(t^n) := \Lambda_h^{n,t^n}(U_h^{n-1}(t^n)), \tag{38a}$$

$$(W_j^{n,l}, v_j) = (U_j^n(t^n), v_j) + \Delta t \sum_{k=1}^{l-1} a_{lk}\langle L_j^n(\Lambda_h^{n,k}(W_h^{n,k})), v_j\rangle, \tag{38b}$$

$$(U_j^n(t), v_j) = (U_j^n(t^n), v_j) + \Delta t \sum_{k=1}^{m} b_k(\frac{t-t^n}{\Delta t})\langle L_j^n(\Lambda_h^{n,k}(W_h^{n,k})), v_j\rangle, \tag{38c}$$

$$\text{for all } v_j \in \mathbb{P}_p, j \in J^n, t \in [t^n, t^{n+1}].$$

Thus, with the definition of L_h^n, the fully-discrete generalized RK–DG approximation satisfies on each time slab (t^n, t^{n+1}) the ordinary partial differential equation

$$(\partial_t U_h^n(t), v_h) = \sum_{k=1}^{m} b'_k(\frac{t-t^n}{\Delta t})\langle L_h^n(\Lambda_h^{n,k}(W_h^{n,k})), v_h\rangle, \ \forall \ v_h \in V_{h,n}^p, \tag{39}$$

Remark 4.2. *In our numerical experiments we used polynomial degree $p = 1, 2, 3$ for the space discretization combined with the NCE-Runge-Kutta method of the same degree. In* [29] *extensions of Runge-Kutta methods are constructed with optimal order up to $p = 4$ but for $p = 3$ and $p = 4$ it is necessary to include a stage reuse procedure to obtain the desired order. In our examples we have included stage reuse since this does not increase the computation cost of the scheme.*

5. Choice of the projections and adaptive strategy

In Definition 2.1 we introduced a class of semi-discrete DG–methods for arbitrary limiting projections $\Lambda_h^{n,t}$ and computational grids. In this subsection we are going to describe specific choices of projection operators that are motivated by our a–posteriori error estimate in Theorem 2.5. Furthermore, we give an adaptive strategy for local mesh refinement based on the error estimate.

The evolution of the solution in the interval $(t^n, t^{n+1}]$ is described in the following algorithm. We start with an initial guess $\tilde{\mathcal{T}}^n$ for the grid and $\tilde{\Lambda}_h^{n,t}$ for the limiting projection ($t \in (t^n, t^{n+1}]$).

- **Given:** Grid $\tilde{\mathcal{T}}^n$, projection $\tilde{\Lambda}_h^{n,t}$ for $t \in [t^n, t^{n+1}]$ and $u^n(t^n, x)$
- **do**
  (1) Let $\mathcal{T}^n = \tilde{\mathcal{T}}^n$ and $\Lambda_h^{n,t} = \tilde{\Lambda}_h^{n,t}$ for $t \in [t^n, t^{n+1}]$
  (2) Compute $u^n(t, x)$ for $t \in (t^n, t^{n+1}]$ on $\mathcal{T}^n$ using $\Lambda_h^{n,t}$
  (3) Compute indicators and new limiting projection $\tilde{\Lambda}_h^{n,t}$ on $\mathcal{T}^n$
  For $j \in J^n$ compute
    - $\rho_j^n, R_{T,j}^n, R_{S,j}^n$ (cf. Corollary 2.6)
    - $\tilde{\Lambda}_h^{n,t}$ for $t \in (t^n, t^{n+1}]$ (cf. Sec. 5.1)
    - $\tilde{R}_{\Lambda,j}^n := \int_{T_j} |\overline{\tilde{u}_k^{n+1}} - \tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1})|$
  (4) Compute error indicator for interval $(t^n, t^{n+1}]$ on $\mathcal{T}^n$:
  $R^n := 2\sum_{j\in J^n} \rho_j^n (R_{T,j}^n + R_{S,j}^n + \tilde{R}_{\Lambda,j}^n)$
  (5) Refine grid: $\mathcal{T}^n \to \tilde{\mathcal{T}}^n$ and project $\tilde{\Lambda}_h^{n,t}$ for $t \in (t^n, t^{n+1}]$ onto $\tilde{\mathcal{T}}^n$
- **while** $R^n > \mathrm{TOL}^n$
- define $\tilde{\mathcal{T}}^{n+1}$ by coarsening $\tilde{\mathcal{T}}^n$ so that
  $R^n + 2\sum_{j\in\tilde{J}^n} \rho_j^n (\int_{T_j} |\tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1}) - \Pi^{n\to n+1}\tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1})|) < \mathrm{TOL}^n$
- define $\tilde{\Lambda}_h^{n+1,t^{n+1}} := \Pi^{n\to n+1}\tilde{\Lambda}_h^{n,t^{n+1}}$ and $\tilde{\Lambda}_h^{n+1,t}$ for $t \in (t^{n+1}, t^{n+2}]$ on $\tilde{\mathcal{T}}^{n+1}$ using $\tilde{\Lambda}_h^{n,t^n}$

The algorithm is based on the assumption that

$$K_1 h_j + K_2 \max_{k\in\{j,l\in N(j)\}} ||\overline{\tilde{\Lambda}_h^{n,t} u_k^n} - \tilde{\Lambda}_h^{n,t} u_k^n||_{L^\infty((t^n,t^{n+1})\times T_k)} \leq$$
$$K_1 h_j + K_2 \max_{k\in\{j,l\in N(j)\}} ||\overline{\Lambda_h^{n,t} u_k^n} - \Lambda_h^{n,t} u_k^n||_{L^\infty((t^n,t^{n+1})\times T_k)}.$$

With the restrictive choice $\tilde{\Lambda}_h^{n,t}(v) = \overline{v}$ we have $\max_k ||\overline{\tilde{\Lambda}_h^{n,t} u_k^n} - \tilde{\Lambda}_h^{n,t} u_k^n||_{L^\infty} = 0$ and therefore $\rho_j^n = K_1 h_j$ as in the first order case. In fact with this limiting operator the DG scheme reduces to the first order finite volume scheme for which the convergence of the error indicator can be shown rigorously for $h \to 0$. Therefore, with a suitable choice of $\mathcal{T}^n$ the iteration (1)–(5) always terminates and in practice our scheme requires hardly any iterations.

Note that from our strategy it follows that

$$\sum_{j\in J^n} \rho_j^n \left(R^n_{T,j} + R^n_{S,j} + R^n_{\Lambda,j}\right) \le \mathrm{TOL}^n.$$

Thus, the error $||(u-\tilde{u}_h)(T)||_{L^1(B_R(x_0))}$ is bounded by some prescribed tolerance TOL which satisfies $\sum_n \mathrm{TOL}^n \le \mathrm{TOL}$. This is summarized in the following Lemma.

Lemma 5.1. *Let η_h denote the global error estimator from Theorem 2.5 and let a prescribed tolerance TOL be given. If the computational mesh is adapted due to the strategy described above using the methods described in the following subsections then it follows that*

$$\eta_h \le \mathrm{TOL}. \tag{39}$$

Thus, the adaptive strategy together with Theorem 2.5 yields a rigorous control on the error $||(u-\tilde{u}_h)(T)||_{L^1(B_R(x_0))}$.

5.1. **Choice of the projection operator.** Our algorithm is based on an initial guess for the projection operator which we denote with $\tilde{\Lambda}_h^{n,t}$. This is used to define the final projection operators $\Lambda_h^{n,t}$ in the generalized RK-DG method 2.4 or 4.1. We are now going to introduce two different approaches for constructing $\tilde{\Lambda}_h^{n,t}$. The first approach is based on a restriction of the gradients of the approximate solution based on the error estimate in Corollary 2.6. The second approach is a p-adaptive projection where the local polynomial degree of the approximate solution is chosen in accordance with the error indicators in Corollary 2.6. Together with the local mesh adaption strategy that we will discuss in the next subsection, both methods are then used in an hp-adaptive manner as discussed in the introduction to this section. The operator $\tilde{\Lambda}_h^{n,t}$ is always constructed on a fixed mesh $\mathcal{T}^n$ and then prolonged/restricted onto a modified mesh in such a way that refinement of cells does not change the projected function.

The goal of the choice of the projection $\tilde{\Lambda}_h^{n,t}$ is twofold. On the one hand side we need a projection or limiting of the solution in order to stabilize the scheme at least in the case of non-linear conservation laws in the vicinity of discontinuities. On the other hand, the factor ρ_j^n should be in $O(h_j)$ — in other words $||\overline{\Lambda_j^{n,t} u_h} - \Lambda_j^{n,t} u_h||_{L^\infty((t^n,t^{n+1})\times T_j)}$ should at least be of the order of h_j. Together with reasonable assumption on the boundedness of the residual terms $R^n_{T,j}, R^n_{S,j}, R^n_{\Lambda,j}$ this requirement guarantees the convergence of the errors estimate η_h for $h \to 0$. We can expect that in regions where the solution u is smooth the stated requirement is met even if we choose $\Lambda_j^{n,t} = id$, whereas near discontinuities the term $||\overline{u_h} - u_h||_{L^\infty}$ grows without bound. The projection should therefore only be active on mesh cells near discontinuities and not in smooth regions with steep gradients. Thus, we suggest to define a projection parameter λ_h as

$$\lambda_j^n(t) := \tilde{\lambda}_j^n + \tfrac{t-t^n}{\Delta t^n}\tilde{\lambda}_j^{n+1}, \qquad \tilde{\lambda}_j^n := \frac{h_j}{\left(h_j + \frac{1}{\Delta t^n}\left(R^n_{T,j}+R^n_{S,j}\right)\right)^{\frac{p+2}{p+1}}} \tag{40}$$

and to ensure that our projection operators yield a solution with the property

$$||\overline{\tilde{u}_j^n(\cdot,t)} - \tilde{u}_j^n(\cdot,t)||_{L^\infty(T_j)} \le \lambda_j^n(t). \tag{41}$$

We expect that the upper bound λ_j^n is of order h_j near discontinuities, whereas it is of order $h_j^{-\frac{1}{p+1}}$ in smooth regions. As the error $||\overline{u_j^n(\cdot,t)} - u_j^n(\cdot,t)||_{L^\infty(T_j)}$

is expected to converge with order h_j in smooth regions and to remain constant near discontinuities, the upper restriction leads to a projection of the solution near discontinuities and at the same time $||\overline{\widetilde{u}_j^n(\cdot,t)} - \widetilde{u}_j^n(\cdot,t)||_{L^\infty(T_j)}$ would be at least of order $\mathcal{O}(h_j)$. The bound (41) dictates the manner in which we construct the operator $\tilde{\Lambda}_j^{n,t}$ from a given projection $\Lambda_j^{n,t}$.

In the sequel we propose two possible choices for the projection $\tilde{\Lambda}_h^{n,t}$ in one space dimension, which both satisfy the upper bound (41) for given limiter function λ_h. In order to define the methods, let $\varphi_l, l = 0, ..., p$ denote the orthogonal basis of Legendre polynomials on the cell $T_j := (x_{j-1/2}, x_{j+1/2})$, such that $\varphi_l \in \mathbb{P}_l(T_j)$. We then have the local expansion

$$u_j^n(x,t) = \sum_{l=0}^{p} u_{j,l}^n(t)\varphi_l(x), \tag{42}$$

where $\varphi_0 = 1$ and thus $\overline{u_j^n(\cdot,t)} = u_{j,0}^n(t)$.

5.1.1. *P–adaptive method in 1D.* Let $1 \le l^* \le p$ denote the maximal index such that

$$\sum_{l=1}^{l^*} u_{j,l}^n(t^n)\varphi_l(x) \le \lambda_j^n(t^n), \qquad \text{for all } x \in T_j.$$

Then, the p-adaptive projection on the cell T_j is defined through

$$\tilde{\Lambda}_j^{n,t}(u_h(\cdot,t)) := \sum_{l=0}^{l^*} u_{j,l}^n(t)\varphi_l(x). \tag{43}$$

5.1.2. *Derivatives restriction method in 1D.* For fixed $t \in [t^n, t^{n+1}]$ let $1 \le l^* \le p$ denote the maximal index such that

$$\sum_{l=1}^{l^*} u_{j,l}^n(t)\varphi_l(x) \le \lambda_j^n(t), \qquad \text{for all } x \in T_j.$$

In contrast to the p-adaptive strategy, we allow here that the derivative of degree l^*+1 is not switched of completely but is reduced in such a way that the bound (41) still holds. In particular, we define

$$\tilde{\Lambda}_j^{n,t}(u_h(\cdot,t)) \ := \ \sum_{l=0}^{l^*} u_{j,l}^n(t)\varphi_l(x) + \widetilde{u}_{j,l^*+1}^n(t)\varphi_{l^*+1}(x) \tag{44}$$

where $\widetilde{u}_{j,l^*+1}^n(t)$ is given as

$$\widetilde{u}_{j,l^*+1}^n(t) := \operatorname{sgn}(u_{j,l^*+1}^n(t)) \ \min\{|u_{j,l^*+1}^n(t)|, \lambda_j^n(t) - ||\sum_{l=0}^{l^*} u_{j,l}^n(t)\varphi_l||_{L^\infty(T_j)}\}.$$

After the refinement of a cell T_j or the coarsening of a set of cells $(T_{j_k})_{k=1}^l$ the operator $\tilde{\Lambda}_h^{n,t}$ has to be modified to operate on the new grid. Both of the choices described above require the definition of λ_j^n on the new grid cells:

- **refinement** $(T_j \to (T_{j_k})_{k=1}^l)$: let $\lambda_{j_k}^n = \frac{1}{l}\lambda_j^n$
- **coarsening** $((T_{j_k})_{k=1}^l \to T_j)$: let $\lambda_j^n = \sum_{k=0}^l \lambda_{j_k}^n$

5.2. **Adaptive strategy for local mesh refinement.** In this subsection we describe an adaptive strategy for local mesh refinement or coarsening that is based on an equal distribution strategy of the error indicator η_h of Theorem 2.5. However, there are two significant modifications of the equal distribution strategy when compared with the strategy presented in [20] or [24]. The first modification is that we only distribute the error among those elements that significantly contribute to the error and secondly we also incorporate the projection error from mesh coarsening into the adaptive strategy. These modifications are of minor importance for smooth solution but result in quite different adaptive convergence behavior for problems with discontinuities. In detail, the new adaptive strategy is given as follows.

Using the notation of Corollary 2.6, let us define for a prescribed tolerance TOL the local error indicators η_j^n for some given $\Theta \in (0,1)$ as

$$\begin{aligned}\eta_j^0(M) &:= \frac{M}{(1-\Theta)\ \mathrm{TOL}}\eta_{0,j},\\ \eta_j^n(M) &:= \frac{2\ T\ M}{\Delta t^n\ (\Theta\ \mathrm{TOL}^n)^2}\rho_j^n(R_{T,j}^n + R_{S,j}^n + \tilde{R}_{\Lambda,j}^n)\end{aligned}$$

where we again have used the abbreviation $\tilde{R}_{\Lambda,j}^n := \int_{T_j} |\overline{\tilde{u}_j^{n+1}}(t^{n+1}) - \tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1})|$. The operator $\tilde{\Lambda}_h^{n,t^{n+1}}$ is again a suitable projection operator defined on the mesh $\mathcal{T}_n$ used to stabilize the scheme and to guarantee that ρ_j^n converges for $h_j \to 0$.

The adaptive strategy at the time t^{n+1} is then given as follows. For $\alpha \in (0, 0.5)$, $M \in \mathbb{N}$ let us define the set of significant elements as

$$I_s^n(M) := \{j \in I^n |\ \eta_j^n(M) \geq \alpha\},$$

and let M^n implicitly be defined through $M^n = |I_s^n(M^n)|$ where $|\cdot|$ denotes the cardinality of the set. We define ε^n as

$$\varepsilon^n := \sum_{j \in I^n \setminus I_s^n(M^n)} \eta_j^n(1)$$

and suppose that α is chosen small enough to ensure $\varepsilon^n \in (0, 0.5)$.

We then define for given $\beta \in (0,1)$ the sets

$$I_r := \{T_j |\ \eta_j^n(M^n) \geq (1-\varepsilon^n)\}, \qquad \tilde{I}_c := \{T_j |\ \eta_j^n(M^n) \leq \beta(1-\varepsilon^n)\}$$

and mark all elements of the set I_r for refinement and those in the set $\tilde{I}_c$ as candidates for coarsening. Coarsening of the mesh leads to an additional projection error of the approximate solution that contributes to the indicator $R_{\Lambda,j}^n$. We split this error into two parts according to $R_{\Lambda,j}^n \leq \tilde{R}_{\Lambda,j}^n + R_{c,j}^n$ with $R_{c,j}^n := \int_{T_j} |\tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1}) - \Pi^{n\to n+1}\tilde{\Lambda}_h^{n,t^{n+1}} u_j^n(t^{n+1})|$ using the operator $\Pi^{n \to n+1}$ to denote the L^2 projection from one grid to another. We calculate the error terms $\eta_{c,j}^n(M^n) := \eta_j^n(M^n) + \frac{2\ T\ M}{\Delta t^n\ (\Theta\ \mathrm{TOL}^n)^2}\rho_j^n R_{c,j}^n$ for all $T_j \in \tilde{I}_c$ and define the updated set I_c as

$$I_c := \{T_j \in \tilde{I}_c |\ \eta_{c,j}^n(M^n) \leq \beta(1-\varepsilon^n)\}$$

and mark all elements of the set I_c for coarsening. Finally, all elements in the set I_r are refined, until the refinement set I_r is an empty set. Then, all elements of the set I_c are coarsened.

5.3. **Evaluation of the semi-discrete error indicators.** In Theorem 2.5 we give an a–posteriori error estimate for the semi-discrete DG-method from Definition 2.4. Conventionally, Runge Kutta time discretizations that are used in practice only provide values for the approximate solution at the discrete time steps t^n. Thus, our error indicators could not be evaluated continuously in time. In order to give a suitable interpretation of the fully discrete Runge-Kutta solution (see Definition 4.1) we use the natural continuous extension as defined through equation (33). Thus, the approximate fully discrete solution is continuous in time on each time slab $[t^n, t^{n+1}]$ and all contributions of the error indicators $\eta_{i,j}^n$ of Theorem 2.5 are computable.

6. Adaptive numerical experiments in one space dimension

In this section we examine numerically the RK–DG–methods defined through Definition 4.1 together with the projections from Subsection 5.1 and the local adaptive grid refinement from Subsection 5.2. We study the convergence behavior of the estimator η_h from Theorem 2.5, as well as the convergence of the error between the RK–DG–approximation and the exact solution itself. As test problems, we look at a linear transport problem with smooth and discontinuous regions in the solution. This example is a scalar prototype for contact discontinuities. As a second very challenging example we chose the Buckley–Leverett equation. Here, the flux function is non-convex and thus the solution consists of compound waves. For such fluxes there exist several weak solutions that are compatible with a single entropy, but only one of those solutions is the unique entropy solution in the Kruzkov sense. It is well known that higher order numerical schemes may have difficulties in selecting this unique Kruzkov entropy solution (see also [2] and [22]).

In order to compare the efficiency of the selected RK–DG–methods we are going to plot the error estimators and errors against the overall number of grid cells $M_{\text{tot}}(\mathcal{T}_h) := \sum_{n=1}^N \sum_{T_j \in \mathcal{T}_h^n} 1$. As M_{tot} is available for uniform refined grids, as well as for adaptively refined grids, and as M_{tot} is proportional to the degrees of freedom for fixed polynomial degree p, this is a good way of comparing our adaptive method with standard approaches on uniform grids. Furthermore, let us define the experimental order of convergence of a grid dependent quantity e_h as

$$EOC(e_{H\to h}) := \log\Big(\frac{e(\mathcal{T}_H)}{e(\mathcal{T}_h)}\Big)\log^{-1}\Big(\frac{M_{\text{tot}}(\mathcal{T}_h)}{M_{\text{tot}}(\mathcal{T}_H)}\Big). \tag{45}$$

Note that a convergence rate $\mathcal{O}(h^p)$ on uniform grids in one space dimension corresponds to $EOC(e_{2h\to h}) = \frac{p}{2}$, as a refinement from grids with cells of size $2h$ to grids with cells of size h leads to two times the number of grid cells per time step and two times the number of time steps. This yields $M_{\text{tot}}(\mathcal{T}_h) = 4M_{\text{tot}}(\mathcal{T}_{2h})$.

6.1. **Linear transport equation.** As a first numerical example we look at the linear transport equation

$$\begin{aligned} \partial_t u + a\partial_x u &= 0 , \\ u(\cdot, 0) &= u_0(\cdot) \end{aligned}$$

with the constant transport velocity $a = 2$. For fixed initial data the solution u is then given by

$$u(x,t) = u_0(x - at) .$$

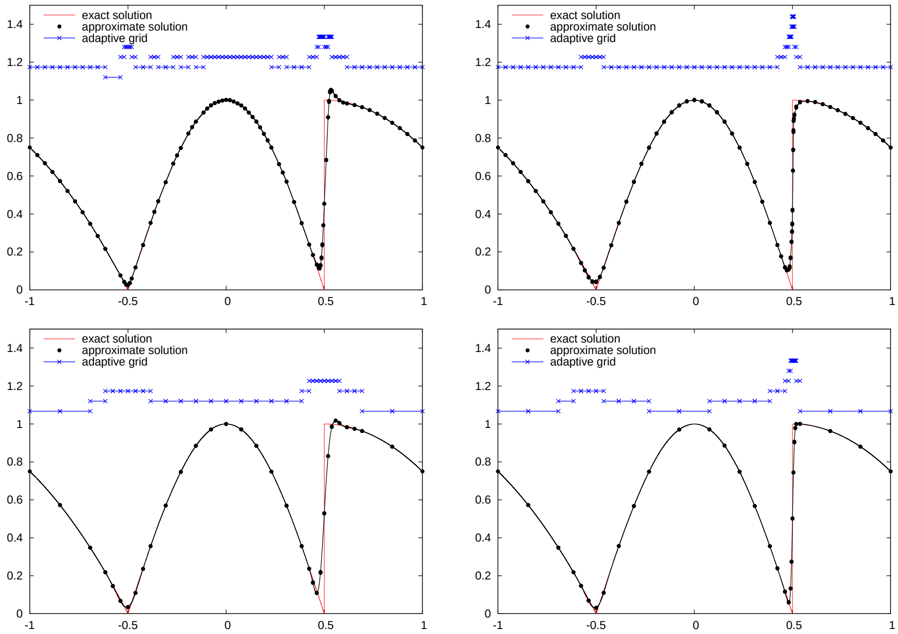

FIGURE 1. Comparison of the approximate solutions obtained with the *p–adaptive* method (left hand side) and the *derivatives restriction* method (right hand side) on adaptively refined grids with $p = 1$ (top) and $p = 2$ (bottom). For both computations we used the prescribed tolerance TOL $= 0.5$. The approximate solutions are compared with the exact solution at $T = 2.0$.

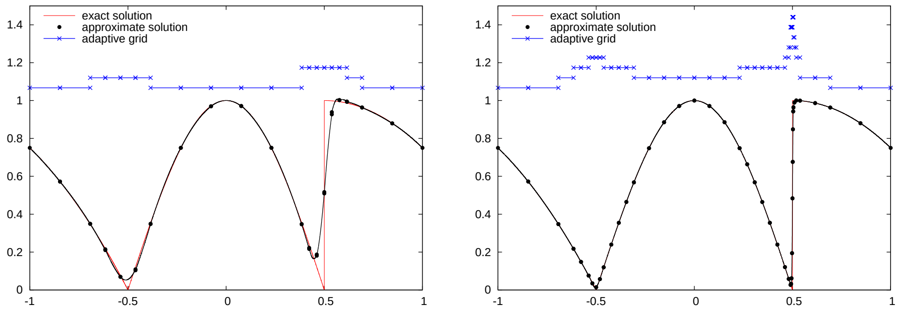

FIGURE 2. Comparison of the approximate solutions obtained with the *derivatives restriction* method for $p = 2$ with TOL $= 1$ (right hand side) and TOL $= 0.25$ (left hand side) on adaptively refined grids. The approximate solutions are compared with the exact solution (solid line) at $T = 2.0$.

We study the setting on $[-1, 1] \times [0, 2]$ with periodic boundary conditions for the following non-smooth initial data

$$u_0(x) := \begin{cases} 1 - (x + 1.5)^2, & \text{for } x < -0.5, \\ \sin((x + 0.5)\pi), & \text{for } -0.5 \le x < 0.5, \\ 1 - (x - 0.5)^2, & \text{for } 0.5 \le x, \end{cases}$$

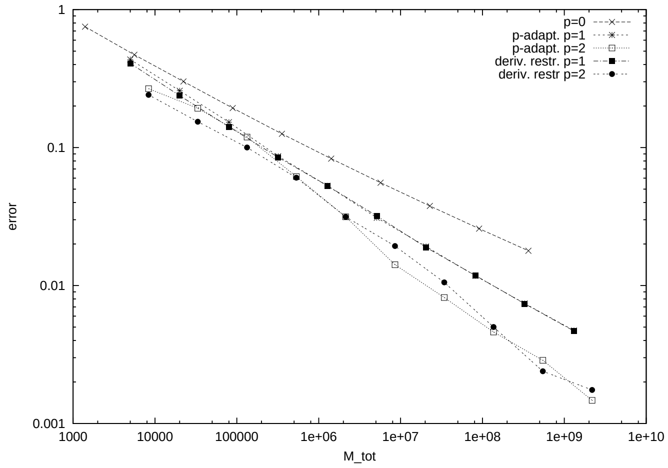

FIGURE 3. Convergence study for our new schemes on uniformly refined grids.

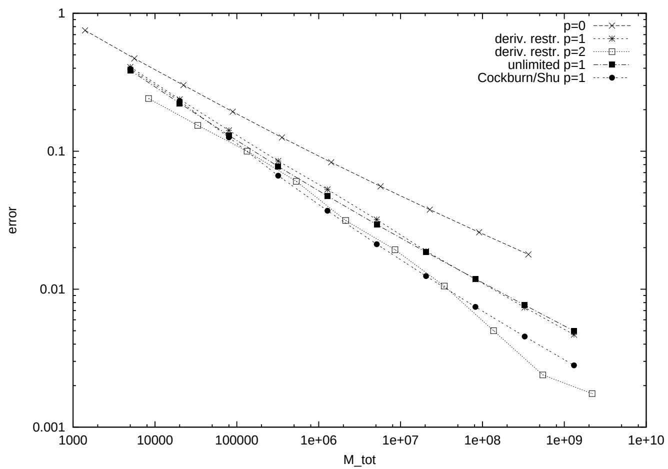

FIGURE 4. Comparison of the *derivatives restriction* method with DG–approximation of Cockburn and Shu and with DG–approximation without limiting projections.

Since $x - 2a = x - 4$ is equal to x on a $[-1, 1]$ periodic domain it follows that $u(x, 2) = u_0(x)$.

We first compare the two projection methods described in Subsection 5.1 for $p = 1$ and $p = 2$ (ref. Fig. 1). All results are computed with the adaptation strategy from Subsection 5.2 with TOL $= 0.5$. In Fig. 1 both the exact solution and the approximate solution are shown together with the grid density function.

The comparison of the projection methods for fixed polynomial degree shows that both methods lead to a good resolution of the smooth region as well as of the discontinuity. The *p-adaptive* method (Fig. 1(left)) produces slight overshoots in

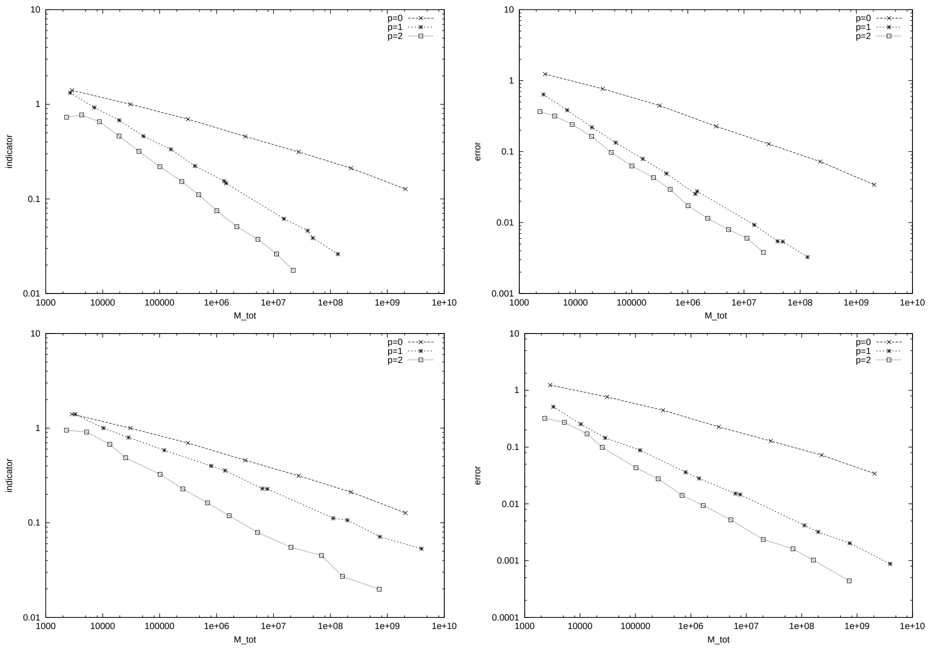

FIGURE 5. Convergence study for the new *p–adaptive* method (top) and the *derivatives restriction* method (bottom) on adaptively refined grids for $p = 0, 1, 2$. The value of the indicator η_h are plotted on the left hand side, while the true error is plotted on the right hand side.

front of the discontinuity but these decrease on finer grids. The refinement strategy together with the *derivatives restriction* method produces a slightly finer grid in the region of the discontinuity whereas the grid is coarser in the smooth regions.

A comparison of different polynomial degrees for fixed projection scheme shows the clear advantage of the quadratic polynomials in the smooth region. The finest grid resolution in the region of the discontinuity is 4 times larger for $p = 2$ and the grid is also considerably coarser in the smooth region. For example the calculation with the *derivatives restriction* method and $p = 1$ produces a final grid with 70 elements and an error of 0.028 whereas the final grid with $p = 2$ has only 30 cells while the error is about the same. As a consequence the overall number of grid cells $M_{\text{tot}}(\mathcal{T}_h)$ is more than five times larger for $p = 1$.

Results with $p = 2$ and the *derivatives restriction* method for different values of TOL are shown in Fig. 2. It can be clearly seen that the grid is hardly refined in the smooth regions of the solution whereas the fineness in the region of the discontinuity and also around the kink increases for smaller tolerance values. The coarsest grid level corresponds to a grid with 13 cells. With TOL $= 1$ only 7 cells are added to the final grid — two in the region of the kink and five in the shock region. With TOL $= 0.25$, 30 cells are added — about 50% of which are located in the shock region.

A comparison of the efficiency of our new method on uniform grids for $p = 0, 1, 2$ is shown in Fig. 3. The increase in efficiency due to an increase of the polynomial

TABLE 1. Experimental order of convergence for the error and the estimator of the new *p–adaptive* method and the *derivatives restriction* method on adaptively refined grids. The underlying computational data coincide with those of Fig. 5.

	derivatives restriction method		*p–adaptive* method	
p	$EOC(e_{H\to h})$	$EOC(\eta_{H\to h})$	$EOC(e_{H\to h})$	$EOC(\eta_{H\to h})$
0	0.292	0.193	0.292	0.193
1	0.431	0.228	0.476	0.368
2	0.544	0.342	0.515	0.450

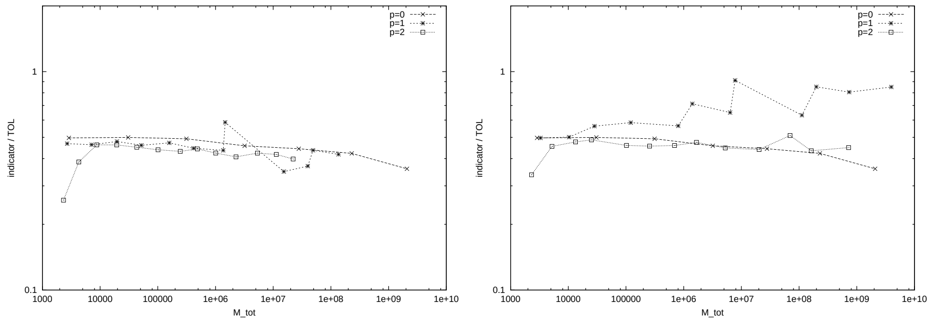

FIGURE 6. Efficiency study for the new *p–adaptive* method (left hand side) and the *derivatives restriction* method (right hand side) on adaptively refined grids for $p = 0, 1, 2$. The efficiency indexes η_h/TOL are plotted versus M_{tot}.

degree can be clearly seen. Furthermore the difference between our two projection methods is hardly significant. For $p = 1$ there is hardly any difference and also for $p = 2$ there is no clear indication of which method is the more efficient. In Fig. 4 we included the DG–approximation of Cockburn and Shu and the DG–approximation without any limiting projections for $p = 1$. For $p = 1$ our projection scheme is comparable to the DG-scheme without any limiting (but note that the scheme without projection is only stable for linear problems). Compared to the Cockburn and Shu version of the DG-method our scheme shows a lower efficiency; but we designed our scheme to be used in combination with h-adaptivity based on a rigorous a-posteriori analysis. The Cockburn and Shu scheme can not be used with our adaptive strategy since the indicators do not converge for this scheme and to our knowledge no rigorous error control is available for it. An error control for hp–adaptive DG–methods is the goal of this paper, not an increase in efficiency on uniform grids. Since we cannot compare our method with the Cockburn and Shu scheme on locally refined grids using the strategy presented in Subsection 5.2, we refrain from further comparisons with this method.

Next we compare our adaptive schemes for $p = 0, 1, 2$. In Fig. 5 the error and the error indicator η_h are shown, while in Tab. 1 the convergence rates as defined in equation (45) are given for the error $EOC(e_{H\to h})$ and the estimator $EOC(\eta_{H\to h})$. The error and the estimator show better convergence rates for higher polynomial degree. The convergence rates of the error are even better than what we expect to

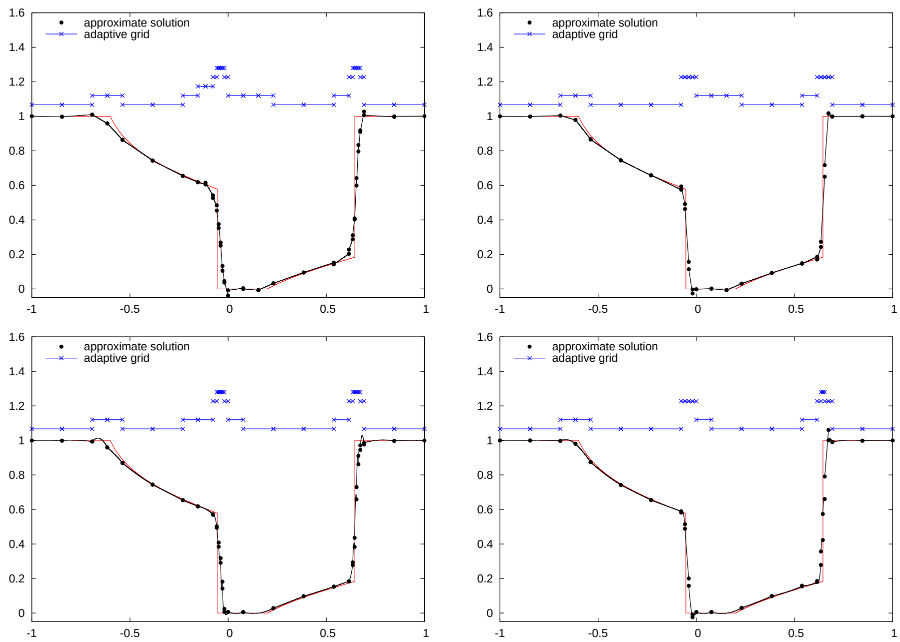

FIGURE 7. Comparison of the approximate solutions obtained with the p–*adaptive* method (left hand side) and the *derivatives restriction* method (right hand side) on adaptively refined grids with $p = 1$ (top row) and $p = 2$ (bottom row). For both computations we used the prescribed tolerance TOL $= 0.5$. The approximate solutions are compared with the exact solution at $T = 0.4$.

be optimal for discontinuous solutions on uniform computational grids. On uniform grids with mesh size h the optimal rate is supposed to be $h^{\frac{p+1}{p+2}}$ which corresponds to $EOC(e_{H\to h}) = \frac{1}{2}\frac{p+1}{p+2}$ (i.e. 0.250, 0.333, 0.375 for p=0,1,2). Although the convergence rate of the indicator differs from the convergence rate of the error the ratio between the prescribed tolerance and the indicator is about constant. In the optimal case this ratio should be close to one. Our adaptive strategy leads to an efficiency index of about $0.5 - 0.8$ (cf. Fig. 6).

6.2. **Buckley–Leverett problem.** As a second example we look at the Buckley–Leverett equation which is a one dimensional model for two phase flow in porous media where capillary pressure effects are neglected. The unknown variable $u : (-1, 1) \times (0, 0.4) \to \mathbb{R}$ is the saturation of the wetting phase within a two phase mixture. It satisfies the non-linear conservation law

$$\begin{aligned} u_t + \partial_x f(u) &= 0, \quad \text{on } (-1,1) \times (0, 0.4), \\ u(\cdot, 0) &= u_0, \quad \text{on } (-1, 1), \end{aligned}$$

where the fractional flow rate f is given as

$$f(s) = \frac{u^2}{u^2 + \frac{1}{2}(1-u)^2}.$$

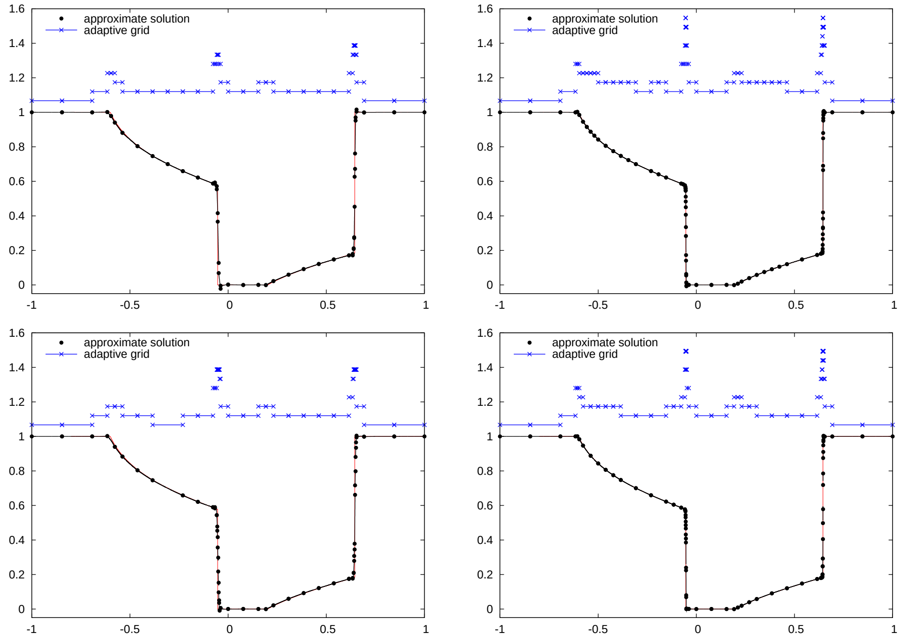

FIGURE 8. Comparison of the approximate solutions obtained with the *derivatives restriction* method for $p = 2$ with TOL $= 0.25$ (right hand side) and TOL $= 0.125$ (left hand side) on adaptively refined grids. The approximate solutions are compared with the exact solution (solid line) at $T = 0.4$.

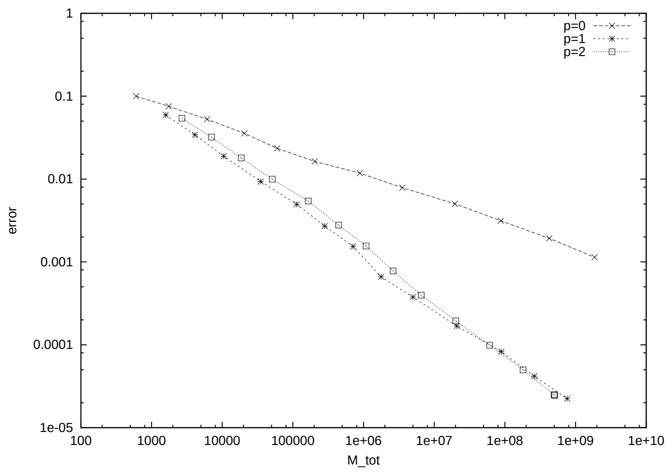

FIGURE 9. Convergence study for our new schemes on adaptively refined grids.

We look at this problem for the following initial data

$$u_0(x) := \begin{cases} 1, & \text{for } x < -0.6, \\ 0, & \text{for } -0.6 \leq x < 0.2, \\ 1, & \text{for } 0.2 \leq x. \end{cases}$$

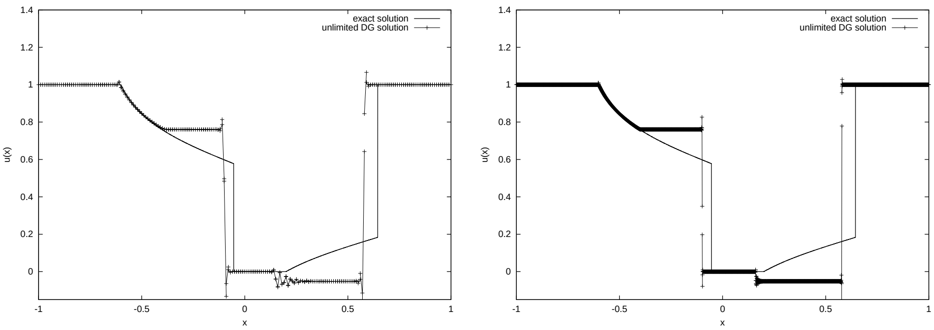

FIGURE 10. Solution of the unlimited DG–scheme for $p = 1$ on uniform grids with 200 (left) and 1600 (right) grid cells. The figure demonstrates that the unlimited DG–scheme converges against the wrong solution for the Buckley–Leverett problem.

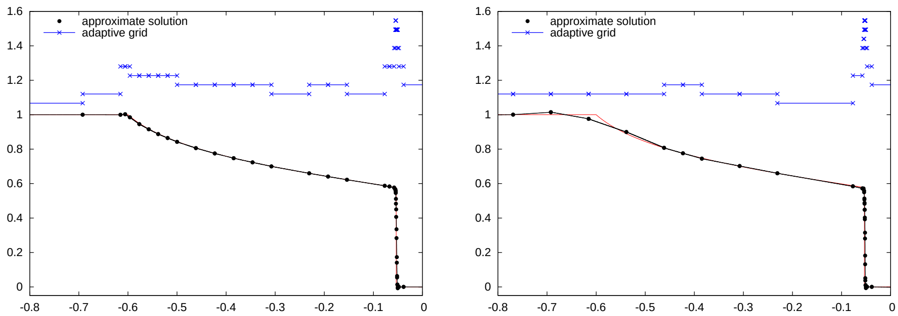

FIGURE 11. Comparison of the new adaptive *derivatives restriction* scheme with and without incorporated coarsening projection error. Only the part of the solution is plotted corresponding to the left Riemann problem at $T = 0.4$ with TOL $= 0.125$ in both cases.

Thus, the solution of our Buckley–Leverett problem consists of the solution of two distinct Riemann problems for t smaller than some critical time $T^* > 0.4$. The solution of each Riemann problem is a composed wave consisting of a rarefaction wave and an attached shock and the exact solution is known up to solving an ODE for the rarefaction waves.

In Fig. 7 we plot the exact solution together with the approximation using our adaptive strategy for $p = 1, 2$. Since the structure of the solution away from the discontinuities is far simpler than in the advection problem studied above, the advantage of the quadratic Ansatz functions is not evident. The grid density function hardly depends on the polynomial degree since almost all grid points are located in the shock regions. Only the kinks at the beginning of the rarefaction waves lead to additional slight refinement. Through a decrease of the tolerance value for the refinement indicator the difference in the resolution of the smooth regions can be demonstrated (cf. Fig. 7). Since the highest grid resolution produced by our refinement strategy is the same for $p = 1$ and $p = 2$ and the approximation error

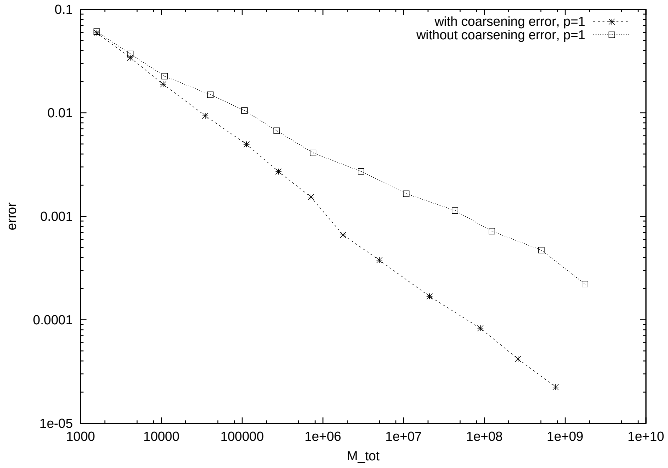

FIGURE 12. The figure shows the influence of the incorporated coarsening projection error on the efficiency of the adaptive algorithm. The error is plotted versus M_{tot} for the *derivatives restriction* scheme with and without incorporated coarsening projection error in the mesh adaption strategy.

is dominated by the shocks, the $p = 2$ version of the DG method does not lead to a more efficient scheme as can be seen from Fig. 9. This must be attributed to the smaller CFL stability restriction required in the higher order schemes and the resulting smaller time-steps. A more complicated structure of the solution — as can only be found in systems in higher space dimension — is required to demonstrate the advantage of a hp-adaptive strategy for non-linear conservation laws with discontinuous solutions.

The results so far show that our adaptive strategy and our projection methods both based on the error estimate from Theorem 2.5 lead to good schemes both for linear and non-linear test problems. In Fig. 10 we demonstrate the necessity of using a projection mechanism in the case of non-linear conservation laws. Without any projection mechanism the DG method clearly does not converges to the Kruzkov entropy solution but to some other solution where the shocks are detached from the rarefaction wave. Both our projection methods, although not monotone, lead to DG schemes which converge to the right entropy solution. Compare with [2] where computational convergence to the entropy solution is achieved through alternative adaptive methods.

We conclude our numerical experiments with results demonstrating the advantage of including the coarsening error in the indicator. Results computed with and without using this "jump" indicator are shown in Figs. 11 and 12. Including the error due to coarsening leads to a higher grid resolution around the kink at the left side of the rarefaction waves. Without this indicator the grid is coarsened to such a degree that the rarefaction wave is not sufficiently resolved and the convergence rate of the adaptive scheme is severely reduced.

7. Conclusion

We have proved an a posteriori error estimate for a class of semi-discrete Discontinuous Galerkin methods on adaptively refined computational meshes (see Theorem 2.5 and Section 3). The estimate provides a rigorous error control and is used for the design of stabilizing limiting projection operators (see Section 5.1) as well as for the design of a local grid adaptation strategy (see Section 5.2). Numerical examples in one space dimension demonstrate that the resulting adaptive schemes converge with higher order compared with the standard first order method with piecewise constant Ansatz functions. In addition it was shown that also the error estimator η_h from the a posteriori Theorem 2.5 converges with higher order for higher order methods. Finally, the principle ideas of the design of projection operators and the adaption strategy is not limited to the one dimensional case and gives rise to promising results also in multi space dimensions. The multi dimensional case and a study of the more involved fully discrete case is left for further studies.

8. Acknowledgment

This work was supported by the European RTN–network HYKE under contract number HPRN-CT-2002-00282. We thank Eitan Tadmor for the hint to the existence of the work on NCE Runge Kutta schemes.

References

[1] S. Adjerid, K. D. Devine, J. E. Flaherty, and L. Krivodonova. A posteriori error estimation for discontinuous Galerkin solutions of hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 191(11-12):1097–1112, 2002.

[2] C. Arvanitis, C. Makridakis, and Tzavaras A. E. Stability and convergence of a class of finite element schemes for hyperbolic systems of conservation laws. *SIAM J. Numer. Anal.*, 42(4):1357–1393, 2004.

[3] F. Bouchut and B. Perthame. Kruzkov's estimates for scalar conservation laws revisited. *Trans. Am. Math. Soc.*, 350(7):2847–2870, 1998.

[4] B. Cockburn and P.A. Gremaud. Error estimates for finite element methods for scalar conservation laws. *SIAM J. Num. Anal.*, 33:522–554, 1996.

[5] B. Cockburn and P.A. Gremaud. A priori error estimates for numerical methods for scalar conservation laws. Part I: The general approach. *Math. Comput.*, 65:533–573, 1996.

[6] B. Cockburn, S. Hou, and C.-W. 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.

[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 and C.-W. 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.

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

[10] 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.

[11] Constantine M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2000.

[12] T. Gallouët and R. Herbin. A uniqueness result for measure-valued solutions of nonlinear hyperbolic equations. *Differential Integral Equations*, 6(6):1383–1394, 1993.

[13] L. Gosse and C. Makridakis. Two a posteriori error estimates for one-dimensional scalar conservation laws. *SIAM J. Numer. Anal.*, 38(3):964–988, 2000.

[14] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.*, 43(1):89–112, 2001.

[15] R. Hartmann and P. Houston. Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws. *SIAM J. Sci. Comput.*, 24(3):979–1004, 2002.

[16] J. Jaffré, C. Johnson, and A. Szepessy. Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *Math. Models Methods Appl. Sci.*, 5(3):367–386, 1995.

[17] C. Johnson and A. Szepessy. Adaptive finite element methods for conservation laws based on a posteriori error estimates. *Commun. Pure Appl. Math.*, 48:199–234, 1995.

[18] T. Katsaounis and C. Makridakis. Finite volume relaxation schemes for multidimensional conservation laws. *Math. Comput.*, 70(234):533–553, 2001.

[19] M.A. Katsoulakis, G. Kossioris, and Ch. Makridakis. Convergence and error estimates of relaxation schemes for multidimensional conservation laws. *Commun. Partial Differ. Equations*, 24(3-4):395–424, 1999.

[20] D. Kröner and M. Ohlberger. A-posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multi dimensions. *Math. Comput.*, 69:25–39, 2000.

[21] S. N. Kruzkov. First order quasilinear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970.

[22] A. Kurganov, G. Petrova, and B. Popov. Adaptive semi-discrete central-upwind schemes for nonconvex hyperbolic conservation laws. Submitted to SIAM J. Sci. Comput., 2004.

[23] Mats G. Larson and Timothy J. Barth. A posteriori error estimation for adaptive discontinuous Galerkin approximations of hyperbolic systems. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 363–368. Springer, Berlin, 2000.

[24] M. Ohlberger. A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations. *M2AN Math. Model. Numer. Anal.*, 35(2):355–387, 2001.

[25] M. Ohlberger. Higher order finite volume methods on selfadaptive grids for convection dominated reactive transport problems in porous media. *Comput. Visual. Sci.*, 7(1):41–51, 2004.

[26] M. Ohlberger and J. Vovelle. Error estimate for the approximation of non-linear conservation laws on bounded domains by the finite volume method. Preprint 03-32, Mathematische Fakultät, Universität Freiburg, 2003.

[27] C.-W. Shu. A survey of strong stability preserving high order time discretizations. Scientific Computing Report Series (2001-18), Brown University, 2001.

[28] C.-W. Shu and S. Osher. Efficent implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Comp. Phys.*, 77:439–471, 1988.

[29] M. Zennaro. Natural continuous extensions of Runge-Kutta methods. *Math. Comput.*, 46:119–133, 1986.

[30] Q. Zhang and C.-W. Shu. Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws. *SIAM J. Numer. Anal.*, 42(2):641–666 (electronic), 2004.

Andreas Dedner, Abteilung für Angewandte Mathematik, Universität Freiburg, Hermann-Herder-Str. 10, D-79104 Freiburg

E-mail address: dedner@mathematik.uni-freiburg.de

Charalambos Makridakis, Department of Applied Mathematics, University of Crete, GR-71409 Heraklion

E-mail address: makr@tem.uoc.gr

Mario Ohlberger, Abteilung für Angewandte Mathematik, Universität Freiburg, Hermann-Herder-Str. 10, D-79104 Freiburg,

E-mail address: mario@mathematik.uni-freiburg.de