*IMA Journal of Numerical Analysis* (2014) **34**, 1447–1488
doi:10.1093/imanum/drt040
Advance Access publication on November 14, 2013

# A local discontinuous Galerkin approximation for systems with *p*-structure

Lars Diening*
*Institute of Mathematics, LMU Munich, Theresienstr. 39, D-80333 Munich, Germany*
*Corresponding author: diening@math.lmu.de

and

Dietmar Kröner, Michael Růžička and Ioannis Toulopoulos
*Department of Applied Mathematics, University of Freiburg, Eckerstr. 1, D-79104 Freiburg, Germany*
dietmar@mathematik.uni-freiburg.de rose@mathematik.uni-freiburg.de
ioannis.toulopoulos@mathematik.uni-freiburg.de

[Received on 15 September 2012; revised on 29 July 2013]

In the present paper, we propose a local discontinuous Galerkin (LDG) approximation for systems with *p*-structure. We discuss appropriate DG spaces and numerical fluxes for the problem. Based on the primal formulation, we show stability (*a priori* estimates) of the method. We derive error estimates, which are optimal for linear ansatz functions. Moreover, we prove new functional analytic tools for the DG setting.

*Keywords*: discontinuous Galerkin; error bounds; degenerate elliptic equations; DG spaces.

## 1. The problem

We consider the numerical approximation of a vectorial system of *p*-Laplace type

$$\begin{aligned} -\mathrm{div}\boldsymbol{\mathcal{A}}(\nabla\mathbf{u}) &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D \quad \text{on } \Gamma_D, \\ \boldsymbol{\mathcal{A}}(\nabla\mathbf{u})\mathbf{n} &= \mathbf{a}_N \quad \text{on } \Gamma_N, \end{aligned} \tag{1.1}$$

by means of local discontinuous Galerkin (LDG) approximations. For given data $\mathbf{f}$, $\mathbf{u}_D$ and $\mathbf{a}_N$ we seek the unknown vector field $\mathbf{u} = (u_1, \ldots, u_d)^{\mathrm{T}}$ defined on $\Omega$. Here, $\Omega \subset \mathbb{R}^n$, $n \geqslant 2$, is a polyhedral, bounded domain with Lipschitz boundary $\partial\Omega$ which is decomposed into $\Gamma_D$ and $\Gamma_N$ satisfying $|\Gamma_D| \neq 0$, $\Gamma_D \cup \Gamma_N = \partial\Omega$, $\Gamma_D \cap \Gamma_N = \emptyset$, and $\mathbf{n}$ is the outer unit normal to $\partial\Omega$. Throughout the paper we assume that $\boldsymbol{\mathcal{A}}\colon \mathbb{R}^{d\times n} \to \mathbb{R}^{d\times n}$ has $(p,\delta)$-structure (cf. Assumption 2.1) and the relevant example which falls into this class is

$$\boldsymbol{\mathcal{A}}(\nabla\mathbf{u}) = (\delta + |\nabla\mathbf{u}|)^{p-2}\nabla\mathbf{u},$$

with $p \in (1,\infty)$ and $\delta \geqslant 0$. Introducing the additional unknowns $\mathbf{L}$ and $\mathbf{A}$ the system (1.1) can be re-written as a 'first-order' system for $\mathbf{u}$, $\mathbf{L}$ and $\mathbf{A}$

$$\begin{aligned} \mathbf{L} = \nabla\mathbf{u}, \quad \mathbf{A} = \boldsymbol{\mathcal{A}}(\mathbf{L}), \quad -\mathrm{div}\mathbf{A} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D \quad \text{on } \Gamma_D, \\ \mathbf{A}\mathbf{n} &= \mathbf{a}_N \quad \text{on } \Gamma_N. \end{aligned} \tag{1.2}$$

DG methods for elliptic problems have been introduced in the late 1990s. In the last decade they have received an extensive attention and by now they are well understood and rigorously analysed in the context of linear elliptic problems (cf. Arnold *et al.*, 2001/2002 for the Poisson problem). In contrast to this, second-order elliptic problems with nonlinear leading terms are not yet completely understood, even if their natural formulation is in the Hilbert space $W^{1,2}(\Omega)$. In fact, only recently appropriate functional analytical tools for the analysis of such problems have been developed (cf. Lasis & Süli, 2003; Bustinza & Gatica, 2004; Houston *et al.*, 2005; Di Pietro & Ern, 2010, 2012). To our knowledge only the papers of Buffa & Ortner (2009) and Burman & Ern (2008) treat the $p$-Laplace problem with DG methods. Both papers analyse minimizer of a scalar, variational problem with $(p,0)$-structure. This leads to an equation of the form (1.1) with $\mathcal{A}(\nabla u) = |\nabla u|^{p-2}\nabla u$, see Remark 2.15 for more details. It is shown that the sequence of approximate DG solutions converges to the original minimizer, but no rate of convergence is provided. On the other hand, it is known from Ebmeyer & Liu (2005) and Diening & Růžička (2007) that finite element solutions of (1.1) converge (in terms of the quasi-norm, see below) at least with a linear rate to the exact solution at least in the case $\Gamma_D = \partial\Omega$ and $\mathbf{u}_D = \mathbf{0}$. Similar results have been obtained in Belenki *et al.* (2012a) for the $p$-Stokes system, where an additional pressure term appears. Moreover, is has been shown in Diening & Kreuzer (2008) and Belenki *et al.* (2012b) that the adaptive finite element algorithm for (1.1) with piecewise linear ansatz functions converges with an optimal rate to the solution. These results and our interest in fluid–structure interaction problems have been the motivation for our work. We believe that the flexibility in the choice of the local polynomial spaces and the mesh geometry are a big advantage if one wants to treat problems for generalized Newtonian fluids having $p$-structure in moving domains, which are in fact to be determined by the problem itself (cf. Kröner *et al.*, 2013 for LDG methods applied to generalized Newtonian fluids in the case of a fixed domain). In this paper, we extend the analysis and techniques developed for finite element methods to a rigorous analysis of LDG methods, which turns out to be much more involved. In particular, we include more general boundary conditions in our analysis. We show in this paper that LDG-solutions also converge with a linear rate to the exact solution in some cases.

The paper is organized as follows: In the next section we introduce the notation, the appropriate function spaces, in particular appropriate spaces for DG functions, the basic assumption on the nonlinear operator and its basic consequences, discrete gradients and our numerical fluxes. Moreover, we derive the flux and the primal formulation of our problem. In Section 3, we prove stability of the method, i.e., *a priori* estimates (cf. Theorem 3.2). In Section 4, we prove the best approximation result (cf. Proposition 4.1) and error estimates for our problem (cf. Theorem 4.3, Theorem 4.8, Corollary 4.10). These are the first convergence rates for problems with $(p,\delta)$-structure, like the $p$-Laplace equation. Note that for linear ansatz functions the convergence rates are optimal. In the appendix, we prove various properties of the $L^2$-projection, the jump functionals, the discrete gradient and the extended Scott–Zhang projection for DG functions. Finally, we prove Poincaré inequalities and trace theorems in the Orlicz space setting for DG and Sobolev functions.

## 2. The LDG formulation: notation and main results

In this section we introduce our LDG formulation of problem (1.1). Before that we introduce the notation we will use, state the precise assumptions on the nonlinear elliptic operator and discuss its basic consequences.

## 2.1 *Function spaces*

We use $c$ and $C$ to denote generic constants, which may change from line to line, but not depending on the crucial quantities. Moreover, we write $f \sim g$ if and only if there exists constants $c, C > 0$ such that $cf \leqslant g \leqslant Cf$.

We will use the customary Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $W^{k,p}(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a bounded, polyhedral domain with Lipschitz continuous boundary $\partial\Omega$ which is decomposed into $\Gamma_D$ and $\Gamma_N$ satisfying $|\Gamma_D| \neq 0$, $\Gamma_D \cup \Gamma_N = \partial\Omega$ and $\Gamma_D \cap \Gamma_N = \emptyset$. We will denote by $\|\cdot\|_p$ the norm in $L^p(\Omega)$ and by $\|\cdot\|_{k,p}$ the norm in $W^{k,p}(\Omega)$. The space $W^{1,p}_{\Gamma_D}(\Omega)$ is the closure of the compactly supported, smooth functions $C_0^\infty(\Omega \cup \Gamma_N)$ in $W^{1,p}(\Omega)$. We equip $W^{1,p}_{\Gamma_D}(\Omega)$ (based on Poincaré's inequality) with the gradient norm $\|\nabla\cdot\|_p$. To deal with the discrete analogue of $W^{1,p}_{\Gamma_D}(\Omega)$ we need the following construction: Let $\Omega' \supsetneq \Omega$, be a polyhedral, bounded domain with Lipschitz continuous boundary such that

$$\partial\Omega \setminus \partial\Omega' = \Gamma_D, \quad \partial\Omega \cap \partial\Omega' = \Gamma_N.$$

For a normed space $X$, we denote its topological dual space by $X^*$. We do not distinguish between function spaces for scalar-, vector- or tensor-valued functions. However, we will denote vector-valued functions by boldface letters and tensor-valued functions by capital boldface letters. The scalar product between two vectors $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$. The scalar product between two tensors $\mathbf{P}$ and $\mathbf{Q}$ is denoted by $\mathbf{P} : \mathbf{Q}$, and we use the notation $|\mathbf{P}|^2 = \mathbf{P} : \mathbf{P}^{\mathrm{T}}$. We denote by $|M|$ the $n$-dimensional or $(n-1)$-dimensional Lebesgue measure of a measurable set $M$. The mean value of a locally integrable function $f$ over a measurable set $M \subset \Omega$ is denoted by $\langle f \rangle_M := \fint_M f \,\mathrm{d}x = (1/|M|) \int_M f \,\mathrm{d}x$. Moreover, we use the notation $(f, g) := \int_\Omega fg \,\mathrm{d}x$, whenever the right-hand side is well defined.

We will also use Orlicz and Sobolev–Orlicz spaces (cf. Rao & Ren, 1991). A real convex function $\psi : \mathbb{R}^{\geqslant 0} \to \mathbb{R}^{\geqslant 0}$ is said to be an *N-function*,[1] if $\psi(0) = 0$, $\psi(t) > 0$ for $t > 0$, $\lim_{t\to 0} \psi(t)/t = 0$, as well as $\lim_{t\to\infty} \psi(t)/t = \infty$. As a consequence, there exists $\psi'$, the right derivative of $\psi$, which is nondecreasing and satisfies $\psi'(0) = 0$, $\psi'(t) > 0$ for $t > 0$, and $\lim_{t\to\infty} \psi'(t) = \infty$. We define the *conjugate N-function* $\psi^*$ by $\psi^*(t) := \sup_{s\geqslant 0}(st - \psi(s))$ for all $t \geqslant 0$. If $\psi'$ is strictly increasing and therefore invertible, then $(\psi^*)' = (\psi')^{-1}$. A given N-function $\psi$ satisfies the $\Delta_2$-condition (in short, $\varphi \in \Delta_2$), if there exists $K > 0$ such that for all $t \geqslant 0$, $\psi(2t) \leqslant K\psi(t)$ holds. We denote the smallest such constant by $\Delta_2(\psi)$. We say that $\psi$ satisfies the $\nabla_2$-condition (in short, $\psi \in \nabla_2$), if $\psi^*$ satisfies the $\Delta_2$-condition. In the following we always assume that $\psi$ satisfies the $\Delta_2$- and the $\nabla_2$-condition. Under this condition we have

$$\psi^*(\psi'(t)) \sim \psi(t). \tag{2.1}$$

We denote by $L^\psi(\Omega)$ and $W^{1,\psi}(\Omega)$ the classical Orlicz and Sobolev–Orlicz spaces, i.e., $f \in L^\psi(\Omega)$ if the *modular* $\rho_\psi(f) = \rho_{\psi,\Omega}(f) := \int_\Omega \psi(|f|) \,\mathrm{d}x$ is finite and $f \in W^{1,\psi}(\Omega)$ if $f, \nabla f \in L^\psi(\Omega)$. Equipped with the induced Luxembourg norm $\|f\|_\psi := \inf\{\lambda > 0 : \int_\Omega \psi(|f|/\lambda) \,\mathrm{d}x \leqslant 1\}$ the space $L^\psi(\Omega)$ becomes a Banach space. The same holds for the space $W^{1,\psi}(\Omega)$ if it is equipped with the norm $\|\cdot\|_\psi + \|\nabla\cdot\|_\psi$. Note that the dual space $(L^\psi(\Omega))^*$ can be identified with the space $L^{\psi^*}(\Omega)$. By $W^{1,\psi}_{\Gamma_D}(\Omega)$, we denote the closure of $C_0^\infty(\Omega \cup \Gamma_N)$ in $W^{1,\psi}(\Omega)$ and equip it with the gradient norm $\|\nabla\cdot\|_\psi$.

[1] N stands for 'nice'.

We need the following refined version of the Young inequality: for all $\varepsilon > 0$ there exists $c_\varepsilon > 0$, depending only on $\Delta_2(\psi), \Delta_2(\psi^*) < \infty$, such that for all $s, t \geqslant 0$

$$\begin{aligned} ts &\leqslant \varepsilon \psi(t) + c_\varepsilon \psi^*(s), \\ t\psi'(s) + \psi'(t)s &\leqslant \varepsilon \psi(t) + c_\varepsilon \psi(s) \end{aligned} \tag{2.2}$$

hold.

### 2.2 *Basic properties of the nonlinear operator*

In the whole paper, we assume that the nonlinear operator $\boldsymbol{\mathcal{A}}(\cdot)$ has $(p,\delta)$-structure, which will be defined now. A detailed discussion and full proofs can be found in Diening & Ettwein (2008) and Růžička & Diening (2007).

ASSUMPTION 2.1 (nonlinear operator) We assume that the nonlinear operator $\boldsymbol{\mathcal{A}}\colon \mathbb{R}^{d\times n} \to \mathbb{R}^{d\times n}$ belongs to $C^0(\mathbb{R}^{d\times n}, \mathbb{R}^{d\times n}) \cap C^1(\mathbb{R}^{d\times n} \setminus \{\mathbf{0}\}, \mathbb{R}^{d\times n})$, satisfies $\boldsymbol{\mathcal{A}}(\mathbf{0}) = \mathbf{0}$. Moreover, we assume that the operator $\boldsymbol{\mathcal{A}}$ has $(p,\delta)$-structure, i.e., there exist $p \in (1,\infty)$, $\delta \in [0,\infty)$, and constants $C_0, C_1 > 0$ such that

$$\sum_{j,l=1}^{n} \sum_{i,k=1}^{d} \partial_{kl} \mathcal{A}_{ij}(\mathbf{P}) Q_{ij} Q_{kl} \geqslant C_0 (\delta + |\mathbf{P}|)^{p-2} |\mathbf{Q}|^2, \tag{2.3a}$$

$$|\partial_{kl} \mathcal{A}_{ij}(\mathbf{P})| \leqslant C_1 (\delta + |\mathbf{P}|)^{p-2} \tag{2.3b}$$

are satisfied for all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{d\times n}$ with $\mathbf{P} \neq \mathbf{0}$ and all $i,k = 1,\ldots,d$, $j,l = 1,\ldots,n$. The constants $C_0$, $C_1$ and $p$ are called the *characteristics* of $\boldsymbol{\mathcal{A}}$.

REMARK 2.2 We emphasize that, if not otherwise stated, the constants in the paper depend only on the characteristics of $\boldsymbol{\mathcal{A}}$, but are independent of $\delta \geqslant 0$.

REMARK 2.3 We would like to point out that most of the results proved in the present paper also hold for systems with Orlicz-structure: Given an N-function $\psi$ with $\psi''(t)t \sim \psi'(t)$ (which implies also $\psi \in \Delta_2 \cap \nabla_2$), we can replace the $(\delta + |\mathbf{P}|)^{p-2}$ terms in (2.3a) and (2.3b) by $\psi''(|\mathbf{P}|)$. Such systems arise, for example, as the Euler–Lagrange equation of the local minimizers of the energy $\int_\Omega \psi(|\nabla \mathbf{u}|)\,\mathrm{d}x$, where $\boldsymbol{\mathcal{A}}(\nabla \mathbf{u}) = (\psi'(|\nabla \mathbf{u}|)/|\nabla \mathbf{u}|)\nabla \mathbf{u}$. We refer to Diening *et al.* (2009) and the references therein.

Defining for $t \geqslant 0$ a special N-function $\varphi = \varphi_{p,\delta}$ by

$$\varphi(t) := \int_0^t \alpha(s)\,\mathrm{d}s \quad \text{with } \alpha(t) := (\delta + t)^{p-2} t, \tag{2.4}$$

we can replace on the right-hand side of (2.3) $C_i(\delta + |\mathbf{P}|)^{p-2}$ by $\tilde{C}_i \varphi''(|\mathbf{P}|)$, $i = 0, 1$. The function $\varphi$ satisfies uniformly in $t$ the important equivalence

$$\varphi''(t)t \sim \varphi'(t) \tag{2.5}$$

since $\min\{1, p-1\}(\delta + t)^{p-2} \leqslant \varphi''(t) \leqslant \max\{1, p-1\}(\delta + t)^{p-2}$. Moreover, $\varphi$ satisfies the $\Delta_2$-condition with $\Delta_2(\varphi) \leqslant c2^{\max\{2,p\}}$ (hence independent of $\delta$). This implies that, uniformly in $t$, we have

$$\varphi'(t)t \sim \varphi(t). \tag{2.6}$$

The conjugate function $\varphi^*$ satisfies $\varphi^*(t) \sim (\delta^{p-1}+t)^{p'-2}t^2$ with $1 = 1/p + 1/p'$. Also, $\varphi^*$ satisfies the $\Delta_2$-condition with $\Delta_2(\varphi^*) \leqslant c2^{\max\{2,p'\}}$.

REMARK 2.4 An important example of a nonlinear operator $\boldsymbol{\mathcal{A}}$ satisfying Assumption 2.1 is given by $\boldsymbol{\mathcal{A}}(\mathbf{P}) = \varphi'(|\mathbf{P}|)|\mathbf{P}|^{-1}\mathbf{P}$. In this case, the characteristics of $\boldsymbol{\mathcal{A}}$, namely $C_0$, $C_1$ and $p$, depend only on $p$ and are independent of $\delta \geqslant 0$.

REMARK 2.5 Note that the spaces $L^p(\Omega)$ and $L^\varphi(\Omega)$, as well as $W^{1,p}(\Omega)$ and $W^{1,\varphi}(\Omega)$, are isomorphic. The equivalence of the corresponding norms depends only on $\delta$ and $p$.

Closely related to the nonlinear operator $\boldsymbol{\mathcal{A}}$ with $(p,\delta)$-structure are the functions $\mathbf{F}\colon \mathbb{R}^{d\times n} \to \mathbb{R}^{d\times n}$, $\mathbf{F}^*\colon \mathbb{R}^{d\times n} \to \mathbb{R}^{d\times n}$ defined through

$$\begin{aligned} \mathbf{F}(\mathbf{P}) &:= (\delta + |\mathbf{P}|)^{(p-2)/2}\mathbf{P}, \\ \mathbf{F}^*(\mathbf{P}) &:= (\delta^{p-1} + |\mathbf{P}|)^{(p'-2)/2}\mathbf{P}, \end{aligned} \tag{2.7}$$

where $1 = 1/p + 1/p'$.

Other important tools are the *shifted N-functions* (cf. Růžička & Diening, 2007; Diening & Ettwein, 2008; Diening & Kreuzer, 2008; Belenki *et al.*, 2012b; Růžička, 2013). For an N-function $\psi$, we define the family of shifted N-functions $\{\psi_a\}_{a\geqslant 0}$ for $t \geqslant 0$ by

$$\psi_a(t) := \int_0^t \psi_a'(s)\,\mathrm{d}s \quad \text{with} \;\; \psi_a'(t) := \psi'(a+t)\frac{t}{a+t}. \tag{2.8}$$

It follows from Lemma 26 of Diening & Ettwein (2008) that the conjugate function of the shifted N-function satisfies, for all $t \geqslant 0$,

$$(\psi_a)^*(t) \sim (\psi^*)_{\psi'(a)}(t), \tag{2.9}$$

with constants depending only on $\Delta_2(\psi)$, $\Delta_2(\psi^*)$.

In the special case of $\varphi$ defined in (2.4) related to the $(p,\delta)$-structure, we have that $\varphi_a(t) \sim (\delta + a + t)^{p-2}t^2$ and also $(\varphi_a)^*(t) \sim ((\delta+a)^{p-1}+t)^{p'-2}t^2$. The family $\{\varphi_a\}_{a\geqslant 0}$ satisfies the $\Delta_2$-condition uniformly in $a \geqslant 0$, with $\Delta_2(\varphi_a) \leqslant c2^{\max\{2,p\}}$ and $\Delta_2((\varphi_a)^*) \leqslant c2^{\max\{2,p\}}$.

Moreover, we have (cf. Růžička & Diening, 2007; Diening & Ettwein, 2008; Diening & Kreuzer, 2008) the following lemma:

LEMMA 2.6 (Change of shift) Let $\psi$ be an N-function such that $\psi$ and $\psi^*$ satisfy the $\Delta_2$-condition. Then, for all $\delta \in (0,1)$, there exists $c_\delta$ such that for all $\mathbf{P},\mathbf{Q} \in \mathbb{R}^{d\times n}$, and for all $t \geqslant 0$

$$\begin{aligned} \psi_{|\mathbf{P}|}(t) &\leqslant c_\delta \psi_{|\mathbf{Q}|}(t) + \delta\psi_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|), \\ (\psi_{|\mathbf{P}|})^*(t) &\leqslant c_\delta (\psi_{|\mathbf{Q}|})^*(t) + \delta\psi_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|). \end{aligned} \tag{2.10}$$

Moreover, we have $\psi_{|\mathbf{Q}|}(|\mathbf{P}-\mathbf{Q}|) \sim \psi_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|)$.

The connection between $\boldsymbol{\mathcal{A}}$, $\mathbf{F}$, $\mathbf{F}^*$, $\{\varphi_a\}_{a\geqslant 0}$, and $\{(\varphi^*)_a\}_{a\geqslant 0}$ is best explained by the following proposition (cf. Růžička & Diening, 2007; Diening & Ettwein, 2008).

PROPOSITION 2.7 Let $\mathcal{A}$ satisfy Assumption 2.1, let $\varphi$ be defined in (2.4), and let $\mathbf{F}$ and $\mathbf{F}^*$ be defined in (2.7). Then

$$(\mathcal{A}(\mathbf{P})-\mathcal{A}(\mathbf{Q})):(\mathbf{P}-\mathbf{Q})\sim|\mathbf{F}(\mathbf{P})-\mathbf{F}(\mathbf{Q})|^2 \tag{2.11a}$$

$$\sim\varphi_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|), \tag{2.11b}$$

$$|\mathbf{F}^*(\mathbf{P})-\mathbf{F}^*(\mathbf{Q})|^2\sim(\varphi^*)_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|), \tag{2.11c}$$

uniformly in $\mathbf{P},\mathbf{Q}\in\mathbb{R}^{d\times n}$. Moreover, uniformly in $\mathbf{Q}\in\mathbb{R}^{d\times n}$,

$$\mathcal{A}(\mathbf{Q})\cdot\mathbf{Q}\sim|\mathbf{F}(\mathbf{Q})|^2\sim\varphi(|\mathbf{Q}|). \tag{2.11d}$$

The constants depend only on the characteristics of $\mathcal{A}$.

There also holds

$$|\mathcal{A}(\mathbf{P})-\mathcal{A}(\mathbf{Q})|\sim\varphi'_{|\mathbf{P}|}(|\mathbf{P}-\mathbf{Q}|)\quad\forall\mathbf{P},\mathbf{Q}\in\mathbb{R}^{d\times n}, \tag{2.12}$$

$$|\mathcal{A}(\mathbf{P})|\sim\varphi'(|\mathbf{P}|)\quad\forall\mathbf{P}\in\mathbb{R}^{d\times n}. \tag{2.13}$$

Using this, (2.11c), (2.13), (2.1) and (2.11d) we also get

$$|\mathbf{F}^*(\mathcal{A}(\mathbf{Q}))|^2\sim\varphi^*(|\mathcal{A}(\mathbf{Q})|)\sim|\mathbf{F}(\mathbf{Q})|^2\sim\varphi(|\mathbf{Q}|)\quad\forall\mathbf{Q}\in\mathbb{R}^{d\times n}. \tag{2.14}$$

LEMMA 2.8 Let $\mathcal{A}$ satisfy Assumption 2.1. For all $t\geqslant 0$ and for all $\mathbf{Q},\mathbf{P}\in\mathbb{R}^{d\times n}$ we have

$$(\varphi^*)_{|\mathcal{A}(\mathbf{P})|}(t)\sim(\varphi_{|\mathbf{P}|})^*(t), \tag{2.15}$$

$$(\varphi^*)_{|\mathcal{A}(\mathbf{P})|}(|\mathcal{A}(\mathbf{Q})-\mathcal{A}(\mathbf{P})|)\sim\varphi_{|\mathbf{P}|}(|\mathbf{Q}-\mathbf{P}|), \tag{2.16}$$

$$|\mathbf{F}^*(\mathcal{A}(\mathbf{Q}))-\mathbf{F}^*(\mathcal{A}(\mathbf{P}))|^2\sim|\mathbf{F}(\mathbf{Q})-\mathbf{F}(\mathbf{P})|^2. \tag{2.17}$$

*Proof.* From (2.9) we know that $(\varphi^*)_{|\varphi'(a)|}(t)\sim(\varphi_a)^*(t)$. Now (2.13) and the fact that $\varphi_b(t)\sim\varphi_e(t)$ for $b\sim e$ imply (2.15). Using (2.15), (2.12) and (2.1), we obtain

$$(\varphi^*)_{|\mathcal{A}(\mathbf{P})|}(|\mathcal{A}(\mathbf{Q})-\mathcal{A}(\mathbf{P})|)\sim(\varphi_{|\mathbf{P}|})^*(\varphi'_{|\mathbf{P}|}(|\mathbf{Q}-\mathbf{P}|))\sim\varphi_{|\mathbf{P}|}(|\mathbf{Q}-\mathbf{P}|).$$

This proves (2.16). The second equivalence follows from (2.11). □

From Lemma 2.6–2.8 we immediately obtain the following shift-change result.

COROLLARY 2.9 For all $\mathbf{P},\mathbf{Q}\in\mathbb{R}^{d\times n}$ and for all $t\geqslant 0$, the following

$$(\varphi_{|\mathbf{P}|})(t)\leqslant c(\varphi)_{|\mathbf{Q}|}(t)+c|\mathbf{F}(\mathbf{P})-\mathbf{F}(\mathbf{Q})|^2,$$

$$(\varphi_{|\mathbf{P}|})^*(t)\leqslant c(\varphi_{|\mathbf{Q}|})^*(t)+c|\mathbf{F}(\mathbf{P})-\mathbf{F}(\mathbf{Q})|^2,$$

$$(\varphi^*)_{|\mathcal{A}(\mathbf{P})|}(t)\leqslant c(\varphi^*)_{|\mathcal{A}(\mathbf{Q})|}(t)+c|\mathbf{F}(\mathbf{P})-\mathbf{F}(\mathbf{Q})|^2,$$

$$\varphi^*)_{|\mathbf{P}|}(t)\leqslant c(\varphi^*)_{|\mathbf{Q}|}(t)+c|\mathbf{F}^*(\mathbf{P})-\mathbf{F}^*(\mathbf{Q})|^2$$

hold.

The following lemma shows that $\nabla\mathbf{F}(\mathbf{L})$ also controls $\nabla\mathbf{F}^*(\mathbf{A})$.

LEMMA 2.10 Let $\mathcal{A}$ satisfy Assumption 2.1. Then we have

$$|\nabla \mathbf{F}^*(\mathcal{A}(\mathbf{P}))| \sim |\nabla \mathbf{F}(\mathbf{P})|$$

for every function $\mathbf{P} : \Omega \to \mathbb{R}^{d\times n}$ such that one of the two sides is well defined.

*Proof.* Let $(\tau_h f)(x) := |h|^{-1}(f(x+h) - f(x))$ denote the difference quotient. Then Lemma 2.8 implies $|\tau_h \mathbf{F}^*(\mathcal{A}(\mathbf{L}))|^2 \sim |\tau_h \mathbf{F}(\mathbf{L})|^2$. Passing to the limit $|h| \to 0$ proves the claim. □

REMARK 2.11 (Natural energy spaces) If $\mathcal{A}$ satisfies Assumption 2.1, we see from (2.11d) and (2.14) that the natural function spaces for $\mathbf{u}$, $\mathbf{L}$ and $\mathbf{A} = \mathcal{A}(\mathbf{L})$ are $W^{1,\varphi}(\Omega)$, $L^{\varphi}(\Omega)$ and $L^{\varphi^*}(\Omega)$, respectively.

REMARK 2.12 (Natural distance) In view of the previous lemma, we have, for all $\mathbf{u}, \mathbf{w} \in W^{1,\varphi}(\Omega)$,

$$(\mathcal{A}(\nabla \mathbf{u}) - \mathcal{A}(\nabla \mathbf{w}), \nabla \mathbf{u} - \nabla \mathbf{w}) \sim \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla \mathbf{w})\|_2^2 \sim \int_\Omega \varphi_{|\nabla \mathbf{u}|}(|\nabla \mathbf{u} - \nabla \mathbf{w}|)\,\mathrm{d}x.$$

The constants depend only on the characteristics of $\mathcal{A}$. The last expression equals the quasi-norm introduced in Barrett & Liu (1994) raised to the power $\rho = \max\{p, 2\}$. This ensures that our results can also be expressed in terms of the quasi-norm. We refer to all three equivalent quantities as the *natural distance*.

In view of Proposition 2.7, one can deduce many useful properties of the natural distance and of the quantities $\mathbf{F}$, $\mathbf{F}^*$ and $\mathcal{A}$ from the corresponding properties of the shifted N-functions $\{\varphi_a\}$. For example, the following important estimate follows directly from (2.12), Young's inequality (2.2) and (2.11).

LEMMA 2.13 For all $\varepsilon > 0$, there exists a constant $c_\varepsilon > 0$ depending only on $\varepsilon > 0$ and the characteristics of $\mathcal{A}$ such that for all vector fields $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W^{1,\varphi}(\Omega)$

$$(\mathcal{A}(\nabla \mathbf{u}) - \mathcal{A}(\nabla \mathbf{v}), \nabla \mathbf{w} - \nabla \mathbf{v}) \leqslant \varepsilon \|\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla \mathbf{v})\|_2^2 + c_\varepsilon \|\mathbf{F}(\nabla \mathbf{w}) - \mathbf{F}(\nabla \mathbf{v})\|_2^2.$$

## 2.3 *DG spaces, jumps and averages*

Let $\mathcal{T}_h$ be a family of shape regular triangulations of our domain $\Omega$ consisting of $n$-dimensional simplices $K$ with diameter $h_K$ less than $h$. For simplicity, we assume in the paper that always $h \leqslant 1$. For a simplex $K \in \mathcal{T}_h$, we denote by $\rho_K$ the supremum of the diameters of inscribed balls. We assume that there exists a constant $\omega_0$ independent on $h$ and $K \in \mathcal{T}_h$ such that $h_K \rho_K^{-1} \leqslant \omega_0$. The smallest such constant $\omega_0$ is called the chunkiness of $\mathcal{T}_h$. Note that in the following all constants may depend on the chunkiness $\omega_0$, but are independent of $h$. Let $S_K$ denote the neighbourhood of $K$, i.e., the patch $S_K$ is the union of all simplices of $\mathcal{T}_h$ touching $K$. We further assume for our triangulation that the interior of each $S_K$ is connected. One easily sees that under these assumptions we get that $|K| \sim |S_K|$ and that the number of simplices in $S_K$ and the number of patches to which a simplex belongs to are uniformly bounded with respect to $h > 0$ and $K \in \mathcal{T}_h$. We define the faces of $\mathcal{T}_h$ as follows: an interior face of $\mathcal{T}_h$ is the nonempty interior of $\partial K \cap \partial K'$, where $K$ and $K'$ are two adjacent elements of $\mathcal{T}_h$. For the face $\gamma := \partial K \cap \partial K'$, we use the notation $S_\gamma := K \cup K'$. A boundary face of $\mathcal{T}_h$ is the nonempty interior of $\partial K \cap \partial \Omega$, where $K$ is a boundary element of $\mathcal{T}_h$. For the face $\gamma := \partial K \cap \partial \Omega$, we use the notation $S_\gamma := K$. By $\Gamma_I$, $\Gamma_D$ and $\Gamma_N$ we denote the interior, the Dirichlet and the Neumann faces, respectively, and put $\Gamma := \Gamma_I \cup \Gamma_D \cup \Gamma_N$. We assume that each $K \in \mathcal{T}_h$ has at most one face from $\Gamma_D \cup \Gamma_N$. We introduce the following scalar products on the boundaries through

$$\langle f, g\rangle_\Gamma := \sum_{\gamma\in\Gamma} \int_\gamma fg \,\mathrm{d}s,$$

whenever the right-hand side is well defined. Analogously, we define $\langle\cdot,\cdot\rangle_{\Gamma_D}$, $\langle\cdot,\cdot\rangle_{\Gamma_N}$ and $\langle\cdot,\cdot\rangle_{\Gamma_I}$. We also extend the notation for modulars to the boundaries by setting $\rho_{\psi,B}(\mathbf{G}) := \int_B \psi(|\mathbf{G}|)\,\mathrm{d}s$ for $\mathbf{G} \in L^\psi(B)$, where $B = \Gamma_I$ or $B = \Gamma_D$ or $B = \Gamma_I \cup \Gamma_D$.

We denote by $\mathcal{P}_m(K)$, with $m \in \mathbb{N}_0$, the space of scalar, vector-valued or tensor-valued continuous functions, which are polynomials of degree at most $m$ on a simplex $K \in \mathcal{T}_h$. Given a triangulation of $\Omega$ with the above properties, given an N-function $\psi$, given $q \in [1,\infty]$ and given $k \in \mathbb{N}$, we define

$$\begin{aligned} X_h^k &= X_h^k(\Omega) := \{\mathbf{G} \in L^1(\Omega) | \mathbf{G}|_K \in \mathcal{P}_k(K) \text{ for all } K \in \mathcal{T}_h\},\\ V_h^k &= V_h^k(\Omega) := \{\mathbf{g} \in L^1(\Omega) | \mathbf{g}|_K \in \mathcal{P}_k(K) \text{ for all } K \in \mathcal{T}_h\},\\ W_{\mathrm{DG}}^{1,\psi} &= W_{\mathrm{DG}}^{1,\psi}(\Omega) := \{\mathbf{g} \in L^1(\Omega) | \mathbf{g}|_K \in W^{1,\psi}(K) \text{ for all } K \in \mathcal{T}_h\},\\ W_{\mathrm{DG}}^{1,q} &= WDG^{1,q}(\Omega) := \{\mathbf{g} \in L^1(\Omega) | \mathbf{g}|_K \in W^{1,q}(K) \text{ for all } K \in \mathcal{T}_h\}. \end{aligned} \tag{2.18}$$

Note that both $W^{1,\psi}(\Omega) \subset W_{\mathrm{DG}}^{1,\psi}(\Omega)$ and $V_h^k(\Omega) \subset W_{\mathrm{DG}}^{1,\psi}(\Omega)$. In a slight abuse of notation, we will also use $V_h^k$ and $W_{\mathrm{DG}}^{1,\psi}$ to denote the corresponding function spaces of scalar- or tensor-valued functions. For $\mathbf{g} \in W_{\mathrm{DG}}^{1,\psi}$, we denote by $\nabla_h \mathbf{g}$ the local distributional gradient, and note that for each $K \in \mathcal{T}_h$ the interior trace of $\mathbf{g}$ on $\partial K$ is well defined. For each face $\gamma$ of $K$, we define this interior trace by $\mathrm{tr}_\gamma^K(\mathbf{g})$.

Let $\Pi_{\mathrm{DG}}\colon L^1(\Omega) \to V_h^k(\Omega)$ denote the (local) $L^2$-projection onto $V_h^k(\Omega)$, i.e.,

$$(\Pi_{\mathrm{DG}}\mathbf{g}, \mathbf{z}_h) = (\mathbf{g}, \mathbf{z}_h) \quad \forall \mathbf{z}_h \in V_h^k. \tag{2.19}$$

The analogous definition is used for $\Pi_{\mathrm{DG}} : L^1(\Omega) \to X_h^k(\Omega)$. This projection plays an important role in our analysis and we collect the relevant properties in Appendix A.1.

Before we define a norm on the spaces $V_h^k$ and $W_{\mathrm{DG}}^{1,\psi}$, we introduce jump and averaging operators, and derive a relationship between the global and the local gradient. Let $g, \mathbf{g}, \mathbf{G} \in W_{\mathrm{DG}}^{1,\psi}$. For interior faces $\gamma$, we denote by $[\![g\mathbf{n}]\!]_\gamma$, $[\![\mathbf{g}\otimes\mathbf{n}]\!]_\gamma$ and $[\![\mathbf{G}\mathbf{n}]\!]_\gamma$ the normal jump, i.e., the jump of $g\mathbf{n}$, $\mathbf{g}\otimes\mathbf{n}$ and $\mathbf{G}\mathbf{n}$, respectively. For example, $[\![\mathbf{G}\mathbf{n}]\!]_\gamma$ is defined on an interior face $\gamma \in \Gamma_I$ shared by the adjacent elements $K^-$ and $K^+ \in \mathcal{T}_h$ with outer normals $\mathbf{n}^-$ and $\mathbf{n}^+$, respectively, by

$$[\![\mathbf{G}\mathbf{n}]\!]_\gamma := \mathrm{tr}_\gamma^{K^+}(\mathbf{G})\mathbf{n}^+ + \mathrm{tr}_\gamma^{K^-}(\mathbf{G})\mathbf{n}^-.$$

Note that we write the normal explicitly. This differs from the usual notation, but ensures always a well-defined expression.

For all interior faces we denote by $\{\cdot\}$ the trace average. For example, $\{\mathbf{g}\}$ is defined on an interior face $\gamma \in \Gamma_I$ shared by the adjacent elements $K^-, K^+ \in \mathcal{T}_h$ by

$$\{\mathbf{g}\}_\gamma := \tfrac{1}{2}(\mathrm{tr}_\gamma^{K^+}(\mathbf{g}) + \mathrm{tr}_\gamma^{K^-}(\mathbf{g})).$$

We omit the index $\gamma$ for jumps and averages if there is no danger of confusion.

Now we want to deal with the Dirichlet boundary data on $\Gamma_D$. Recall that $\Omega' \supsetneq \Omega$ is a polyhedral, bounded domain with Lipschitz continuous boundary such that $\partial\Omega \setminus \partial\Omega' = \Gamma_D$, $\partial\Omega \cap \partial\Omega' = \Gamma_N$. Let $\mathcal{T}_h'$ denote an extension of the triangulation $\mathcal{T}_h$ to $\Omega'$, having the same properties as $\mathcal{T}_h$ (in particular with a similar chunkiness). We extend our notation to this setting by adding a superposed 'prime' to it. In particular, we denote by $\Gamma_I'$, $S_K'$ and $S_\gamma'$ the interior faces, the neighbourhood of $K$ and $\gamma$, respectively, of $\mathcal{T}_h'$. We define

$$W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega) := \{\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega') | \mathbf{g}|_{\Omega'\setminus\Omega} = \mathbf{0}\}.$$

So, functions from $W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ are elements of $W^{1,\psi}_{\mathrm{DG}}(\Omega)$, which are (virtually) extended by zero to $\Omega' \setminus \Omega$. Therefore, it is very natural to define the jumps of $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ on $\Gamma_D$ by

$$[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma := \mathrm{tr}^\Omega_\gamma(\mathbf{g}) \otimes \mathbf{n} \quad \text{for } \gamma \in \Gamma_D. \tag{2.20}$$

If $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$, then for all $\mathbf{X} \in C_0^\infty(\Omega')$ we have

$$\sum_{K\in\mathcal{T}_h'} \int_{\partial K} \mathbf{g} \cdot \mathbf{X}\mathbf{n}\, \mathrm{d}x = \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \mathbf{X}\rangle_{\Gamma_I \cup \Gamma_D}.$$

Hence, in the sense of distributions $\mathcal{D}'(\Omega')$, we have

$$\nabla \mathbf{g} = \nabla_h \mathbf{g} - \sum_{\gamma\in\Gamma_I\cup\Gamma_D} \mu_\gamma [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma, \tag{2.21}$$

where $\mu_\gamma$ denotes the $(n-1)$-dimensional Hausdorff measure on $\gamma$ and where the jumps on $\Gamma_D$ are defined by (2.20). In other words, for $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ and $\mathbf{X} \in C_0^\infty(\Omega')$

$$\langle \nabla \mathbf{g}, \mathbf{X}\rangle_{\mathcal{D}'(\Omega')\times\mathcal{D}(\Omega')} = (\nabla_h \mathbf{g}, \mathbf{X}) - \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \mathbf{X}\rangle_{\Gamma_I\cup\Gamma_D}.$$

It is natural to extend this functional to DG functions $\mathbf{X}_h \in X_h^k$ by

$$\mathbf{X}_h \mapsto (\nabla_h \mathbf{g}, \mathbf{X}_h) - \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \{\mathbf{X}_h\}\rangle_{\Gamma_I\cup\Gamma_D}, \tag{2.22}$$

where we used for a compact and unified notation the convention that

$$\{\mathbf{X}_h\}_\gamma := \mathrm{tr}^\Omega_\gamma(\mathbf{X}_h) \quad \text{for } \gamma \in \Gamma_D.$$

It is easy to see (cf. Appendix A.2) that this functional is continuous on $X_h^k$ and can be represented by a unique Riesz representative form $X_h^k$, which we denote by $\nabla^h_{\mathrm{DG}}\mathbf{g}$. In particular, for every $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$, we define the *discrete gradient* $\nabla^h_{\mathrm{DG}}\mathbf{g} \in X_h^k$ (via Riesz representation) by the relation

$$(\nabla^h_{\mathrm{DG}}\mathbf{g}, \mathbf{X}_h) := (\nabla_h \mathbf{g}, \mathbf{X}_h) - \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \{\mathbf{X}_h\}\rangle_{\Gamma_I\cup\Gamma_D} \tag{2.23}$$

for all $\mathbf{X}_h \in X_h^k$. In Section A.2, we discuss the discrete gradient and the jump functionals (defined below) in more detail. In particular, it is shown that these objects are well defined and proper estimates are presented.

It is useful to split the second part of the functional in (2.23) into smaller pieces. For $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ and $\gamma \in \Gamma_I \cup \Gamma_D$, we define the jump functionals $\mathbf{R}^{\gamma}_h \mathbf{g} \in X^k_h$ (via Riesz representation) by the formula

$$(\mathbf{R}^{\gamma}_h \mathbf{g}, \mathbf{X}_h) := \langle [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma, \{\mathbf{X}_h\} \rangle_\gamma, \tag{2.24}$$

for all $\mathbf{X}_h \in X^k_h$. Moreover, let $\mathbf{R}_h := \sum_{\gamma \in \Gamma_I \cup \Gamma_D} \mathbf{R}^{\gamma}_h$, i.e.,

$$(\mathbf{R}_h \mathbf{g}, \mathbf{X}_h) := \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \{\mathbf{X}_h\} \rangle_{\Gamma_I \cup \Gamma_D}. \tag{2.25}$$

Since every $\mathbf{g}_h \in V^k_h$ can also be seen as an element of $W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$, we use the definitions in (2.23), (2.24) and (2.25) also for all $\mathbf{g}_h \in V^k_h$. For this we extend our convention (2.20) to $\mathbf{g}_h \in V^k_h$, i.e.,

$$[\![\mathbf{g}_h \otimes \mathbf{n}]\!]_\gamma := \mathrm{tr}^{\Omega}_{\gamma}(\mathbf{g}_h) \otimes \mathbf{n} \quad \text{for } \gamma \in \Gamma_D. \tag{2.26}$$

The discrete gradient $\nabla^h_{\mathrm{DG}} \mathbf{g}_h$ and the jump functionals $\mathbf{R}^{\gamma}_h \mathbf{g}_h$ and $\mathbf{R}_h \mathbf{g}_h$ have been first introduced in Di Pietro & Ern (2010) and Bustinza & Gatica (2004) for $\mathbf{g}_h \in V^k_h$.

With these definitions, we have the following pointwise identities for $\mathbf{g}_h \in V^k_h$

$$\nabla^h_{\mathrm{DG}} \mathbf{g}_h = \nabla_h \mathbf{g}_h - \mathbf{R}_h \mathbf{g}_h \tag{2.27}$$

and for $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$

$$\nabla^h_{\mathrm{DG}} \mathbf{g} = \Pi_{\mathrm{DG}} \nabla_h \mathbf{g} - \mathbf{R}_h \mathbf{g}. \tag{2.28}$$

Further properties and estimates for $\nabla^h_{\mathrm{DG}}$, $\mathbf{R}^{\gamma}_h$ and $\mathbf{R}_h$ are presented in Appendix A.2.

We define the semi-modulars $m_{\psi,h}$ and $M_{\psi,h}$ for $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ by

$$\begin{aligned}
m^I_{\psi,h}(\mathbf{g}) &:= h\rho_{\psi,\Gamma_I}(h^{-1}[\![\mathbf{g} \otimes \mathbf{n}]\!]),\\
m^D_{\psi,h}(\mathbf{g}) &:= h\rho_{\psi,\Gamma_D}(h^{-1}[\![\mathbf{g} \otimes \mathbf{n}]\!]) = h\rho_{\psi,\Gamma_D}(h^{-1}\mathbf{g} \otimes \mathbf{n}),\\
m_{\psi,h}(\mathbf{g}) &:= m^I_{\psi,h}(\mathbf{g}) + m^D_{\psi,h}(\mathbf{g}),\\
M_{\psi,h}(\mathbf{g}) &:= \rho_{\psi,\Omega}(\nabla_h \mathbf{g}) + m_{\psi,h}(\mathbf{g}).
\end{aligned} \tag{2.29}$$

The semi-modular $M_{\psi,h}$ on $W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$ is equivalent to $\rho_{\psi,\Omega}(\nabla\cdot)$ on $W^{1,\psi}_{\mathrm{DG}}(\Omega)$. The induced Luxembourg semi-norm and norm are $|\cdot|_{m_{\psi,h}}$ and $\|\cdot\|_{M_{\psi,h}}$, respectively. Note that for every $\mathbf{g} \in W^{1,\psi}_{\Gamma_D}(\Omega)$ we have $m_{\psi,h}(\mathbf{g}) = 0$ and $M_{\psi,h}(\mathbf{g}) = \rho_{\psi,\Omega}(\nabla \mathbf{g})$, so $M_{\psi,h}(\cdot)$ is an extension of the modular $\rho_{\psi,\Omega}(\nabla\cdot)$ on $W^{1,\psi}_{\Gamma_D}$ to the DG setting.

We will see later that due to the zero boundary values of $W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$, the semi-modular $M_{\psi,\Omega}$ is in fact a modular, see Lemma A.9. This is in complete analogy to the case $W^{1,\psi}_{\Gamma_D}(\Omega)$.

REMARK 2.14 In the special case $\psi = \varphi$, we have due to (2.11d)

$$\begin{aligned}
m_{\varphi,h}(\mathbf{g}) &\sim h\|\mathbf{F}(h^{-1}[\![\mathbf{g} \otimes \mathbf{n}]\!])\|^2_{2,\Gamma_I \cup \Gamma_D},\\
M_{\varphi,h}(\mathbf{g}) &\sim \|\mathbf{F}(\nabla_h \mathbf{g})\|^2_{2,\Omega} + h\|\mathbf{F}(h^{-1}[\![\mathbf{g} \otimes \mathbf{n}]\!])\|^2_{2,\Gamma_I \cup \Gamma_D}.
\end{aligned} \tag{2.30}$$

### 2.4 *Fluxes and LDG formulation*

In order to obtain the LDG formulation of (1.1), we multiply the equations in $(1.2)_1$ by $\mathbf{X}_h \in X_h^k$, $\mathbf{Y}_h \in X_h^k$ and $\mathbf{z}_h \in V_h^k$, respectively, and obtain, for all $K \in \mathcal{T}_h$,

$$
\begin{aligned}
\int_K \mathbf{L} : \mathbf{X}_h \,\mathrm{d}x &= -\int_K \mathbf{u} \cdot \operatorname{div}_h \mathbf{X}_h \,\mathrm{d}x + \int_{\partial K} \mathbf{u} \cdot (\mathbf{X}_h \mathbf{n}) \,\mathrm{d}s,\\
\int_K \mathbf{A} : \mathbf{Y}_h \,\mathrm{d}x &= \int_K \boldsymbol{\mathcal{A}}(\mathbf{L}) : \mathbf{Y}_h \,\mathrm{d}x,\\
\int_K \mathbf{A} : \nabla_h \mathbf{z}_h \,\mathrm{d}x &= \int_K \mathbf{f} \cdot \mathbf{z}_h \,\mathrm{d}x + \int_{\partial K} \mathbf{z}_h \cdot \mathbf{A}\mathbf{n} \,\mathrm{d}s,
\end{aligned}
$$

where we used $\mathbf{A} : (\mathbf{z}_h \otimes \mathbf{n}) = \mathbf{z}_h \cdot (\mathbf{A}\mathbf{n})$. Replacing in the volume integrals the fields $\mathbf{u}$, $\mathbf{L}$ and $\mathbf{A}$ by their discrete counterparts $\mathbf{u}_h$, $\mathbf{L}_h$ and $\mathbf{A}_h$, respectively, and in the surface integrals $\mathbf{u}$ and $\mathbf{A}$ by the numerical fluxes $\hat{\mathbf{u}}_h := \hat{\mathbf{u}}(\mathbf{u}_h)$ and $\hat{\mathbf{A}}_h := \hat{\mathbf{A}}(\mathbf{u}_h, \mathbf{A}_h)$, we obtain

$$
\begin{aligned}
\int_K \mathbf{L}_h : \mathbf{X}_h \,\mathrm{d}x &= -\int_K \mathbf{u}_h \cdot \operatorname{div}_h \mathbf{X}_h \,\mathrm{d}x + \int_{\partial K} \hat{\mathbf{u}}_h \cdot (\mathbf{X}_h \mathbf{n}) \,\mathrm{d}s,\\
\int_K \mathbf{A}_h : \mathbf{Y}_h \,\mathrm{d}x &= \int_K \boldsymbol{\mathcal{A}}(\mathbf{L}_h) : \mathbf{Y}_h \,\mathrm{d}x,\\
\int_K \mathbf{A}_h : \nabla_h \mathbf{z}_h \,\mathrm{d}x &= \int_K \mathbf{f} \cdot \mathbf{z}_h \,\mathrm{d}x + \int_{\partial K} \mathbf{z}_h \cdot (\hat{\mathbf{A}}_h \mathbf{n}) \,\mathrm{d}s.
\end{aligned}
\tag{2.31}
$$

The numerical fluxes are defined by

$$
\hat{\mathbf{u}}(\mathbf{u}_h) := \begin{cases} \{\mathbf{u}_h\} & \text{on } \Gamma_I,\\ \mathbf{u}_D^* & \text{on } \Gamma_D,\\ \mathbf{u}_h & \text{on } \Gamma_N, \end{cases}
\tag{2.32}
$$

and[2]

$$
\hat{\mathbf{A}}(\mathbf{u}_h, \mathbf{A}_h) := \begin{cases} \{\mathbf{A}_h\} - \alpha \boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h \otimes \mathbf{n}]\!]) & \text{on } \Gamma_I,\\ \mathbf{A}_h - \alpha \boldsymbol{\mathcal{A}}(h^{-1}((\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n})) & \text{on } \Gamma_D,\\ \mathbf{a}_N \otimes \mathbf{n} & \text{on } \Gamma_N, \end{cases}
\tag{2.33}
$$

where $\alpha > 0$ is some constant, and $\mathbf{u}_D$ and $\mathbf{a}_N$ are the given boundary data. We assume that $\mathbf{u}_D^* \in W^{1,\varphi}(\Omega)$ is some approximation of $\mathbf{u}_D$. In fact, for $\mathbf{u}_D^*$ we have in mind $\mathbf{u}_D^* = \mathbf{u}$ or $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$. The fluxes are consistent, since $\hat{\mathbf{u}}(\mathbf{u}) = \mathbf{u}$ and $\hat{\mathbf{A}}(\mathbf{u}, \mathbf{A}) = \mathbf{A}$ (if $\mathbf{u}_D^* = \mathbf{u}$) for regular functions $\mathbf{u}$ and $\mathbf{A}$ satisfying the boundary conditions $\mathbf{u} = \mathbf{u}_D^*$ on $\Gamma_D$ and $\boldsymbol{\mathcal{A}}(\nabla \mathbf{u})\mathbf{n} = \mathbf{a}_N$ on $\Gamma_N$ (cf. $(1.2)_{2,3}$). The fluxes are also conservative since they are single-valued.

There is a special reason for our flux $\hat{\mathbf{u}}(\mathbf{u}_h)$ from (2.32). We will see later in (2.39) that this choice implies $\mathbf{L}_h = \nabla_{\mathrm{DG}}^h \mathbf{u}_h$ if $\mathbf{u}_D = \mathbf{0}$. As explained above, $\nabla_{\mathrm{DG}}^h$ is the natural generalization of the distributional gradient. So $\mathbf{L}_h = \nabla_{\mathrm{DG}}^h \mathbf{u}_h$ is the natural DG equivalent of $\mathbf{L} = \nabla \mathbf{u}$. The flux $\hat{\mathbf{A}}(\mathbf{u}_h, \mathbf{A}_h)$ is a natural generalization of the corresponding fluxes for the Laplace problem (cf. Arnold *et al.*, 2001/2002) and the $p$-Laplace problem (cf. Burman & Ern, 2008), taking into account that for $\delta > 0$, the operator $\boldsymbol{\mathcal{A}}$ is not homogeneous, i.e., $\boldsymbol{\mathcal{A}}(\lambda \mathbf{P}) \neq \lambda^p \boldsymbol{\mathcal{A}}(\mathbf{P})$.

[2] In the definition of the numerical flux $\hat{\mathbf{A}}(\mathbf{u}_h, \mathbf{A}_h)$, one could also use $\boldsymbol{\mathcal{A}}(\mathbf{P}) = \varphi'(|\mathbf{P}|)|\mathbf{P}|^{-1}\mathbf{P}$ from Remark 2.4 instead of the general $\boldsymbol{\mathcal{A}}$ with $(p, \delta)$-structure.

A straightforward computation shows that for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega)$ and $\mathbf{X} \in W^{1,\psi^*}_{\mathrm{DG}}(\Omega)$, we have

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\int_{\partial K}\mathbf{g}\cdot\mathbf{X}\mathbf{n}\,\mathrm{d}x &= \langle\{\mathbf{g}\},[\![\mathbf{X}\mathbf{n}]\!]\rangle_{\Gamma_I}+\langle[\![\mathbf{g}\otimes\mathbf{n}]\!],\{\mathbf{X}\}\rangle_{\Gamma_I}+\langle\mathbf{g}\otimes\mathbf{n},\mathbf{X}\rangle_{\Gamma_D\cup\Gamma_N}\\ &= \langle\{\mathbf{g}\},[\![\mathbf{X}\mathbf{n}]\!]\rangle_{\Gamma_I}+\langle[\![\mathbf{g}\otimes\mathbf{n}]\!],\{\mathbf{X}\}\rangle_{\Gamma_I\cup\Gamma_D}+\langle\mathbf{g}\otimes\mathbf{n},\mathbf{X}\rangle_{\Gamma_N}.\end{aligned} \tag{2.34}$$

Inserting (2.32) and (2.33) into (2.31), summing the result over $K\in\mathcal{T}_h$ and using (2.34) we arrive at

$$\begin{aligned}(\mathbf{L}_h,\mathbf{X}_h) &= -(\mathbf{u}_h,\mathrm{div}_h\mathbf{X}_h)+\langle\{\mathbf{u}_h\},[\![\mathbf{X}_h\mathbf{n}]\!]\rangle_{\Gamma_I}+\langle\mathbf{u}_D^*,\mathbf{X}_h\mathbf{n}\rangle_{\Gamma_D}+\langle\mathbf{u}_h,\mathbf{X}_h\mathbf{n}\rangle_{\Gamma_N},\\ (\mathbf{A}_h,\mathbf{Y}_h) &= (\boldsymbol{\mathcal{A}}(\mathbf{L}_h),\mathbf{Y}_h),\\ (\mathbf{A}_h,\nabla_h\mathbf{z}_h) &= (\mathbf{f},\mathbf{z}_h)+\langle\{\mathbf{A}_h\},[\![\mathbf{z}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}+\langle\mathbf{A}_h,\mathbf{z}_h\otimes\mathbf{n}\rangle_{\Gamma_D}+\langle\mathbf{a}_N,\mathbf{z}_h\rangle_{\Gamma_N}\\ &\quad -\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h\otimes\mathbf{n}]\!]),[\![\mathbf{z}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}-\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}),\mathbf{z}_h\otimes\mathbf{n}\rangle_{\Gamma_D}.\end{aligned}$$

Using partial integration together with (2.34) in the first equation, we obtain

$$\begin{aligned}(\mathbf{L}_h,\mathbf{X}_h) &= (\nabla_h\mathbf{u}_h,\mathbf{X}_h)-\langle[\![\mathbf{u}_h\otimes\mathbf{n}]\!],\{\mathbf{X}_h\}\rangle_{\Gamma_I}-\langle(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n},\mathbf{X}_h\rangle_{\Gamma_D},\\ (\mathbf{A}_h,\mathbf{Y}_h) &= (\boldsymbol{\mathcal{A}}(\mathbf{L}_h),\mathbf{Y}_h),\\ (\mathbf{A}_h,\nabla_h\mathbf{z}_h) &= (\mathbf{f},\mathbf{z}_h)+\langle\{\mathbf{A}_h\},[\![\mathbf{z}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}+\langle\mathbf{A}_h,\mathbf{z}_h\otimes\mathbf{n}\rangle_{\Gamma_D}+\langle\mathbf{a}_N,\mathbf{z}_h\rangle_{\Gamma_N}\\ &\quad -\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h\otimes\mathbf{n}]\!]),[\![\mathbf{z}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}-\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}),\mathbf{z}_h\otimes\mathbf{n}\rangle_{\Gamma_D}.\end{aligned} \tag{2.35}$$

Using the notation of the discrete gradient $\nabla^h_{\mathrm{DG}}$ and the definition of $\mathbf{R}_h$ in (2.25), we get the *flux formulation* of (1.1): For given data $\mathbf{u}_D \in W^{1-1/p,p}(\Gamma_D)$,[3] $\mathbf{f}\in L^{p'}(\Omega)$ and $\mathbf{a}_N \in L^{p'}(\Gamma_N)$ find $(\mathbf{u}_h,\mathbf{L}_h,\mathbf{A}_h)\in V_h^k\times X_h^k\times X_h^k$ such that for all $(\mathbf{X}_h,\mathbf{Y}_h,\mathbf{z}_h)\in X_h^k\times X_h^k\times V_h^k$

$$\begin{aligned}(\mathbf{L}_h,\mathbf{X}_h) &= (\nabla^h_{\mathrm{DG}}\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*,\mathbf{X}_h),\\ (\mathbf{A}_h,\mathbf{Y}_h) &= (\boldsymbol{\mathcal{A}}(\mathbf{L}_h),\mathbf{Y}_h),\\ (\mathbf{A}_h,\nabla^h_{\mathrm{DG}}\mathbf{z}_h) &= (\mathbf{f},\mathbf{z}_h)+\langle\mathbf{a}_N,\mathbf{z}_h\rangle_{\Gamma_N}-\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h\otimes\mathbf{n}]\!]),[\![\mathbf{z}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}\\ &\quad -\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}),\mathbf{z}_h\otimes\mathbf{n}\rangle_{\Gamma_D}.\end{aligned} \tag{2.36}$$

Let us now see whether the weak solution of our original problem (1.1) satisfies a similar system. First of all note that one can easily show, using the theory of monotone operators, that for all data $\mathbf{u}_D\in W^{1-1/p,p}(\Gamma_D)$, $\mathbf{f}\in L^{p'}(\Omega)$ and $\mathbf{a}_N\in L^{p'}(\Gamma_N)$ there exists a weak solution $\mathbf{u}\in W^{1,\varphi}(\Omega)$, i.e., $\mathbf{u}-\mathbf{u}_D\in W^{1,\varphi}_{\Gamma_D}(\Omega)$ and

$$\int_\Omega \boldsymbol{\mathcal{A}}(\nabla\mathbf{u}):\nabla\boldsymbol{\phi}\,\mathrm{d}x=\int_\Omega \mathbf{f}\cdot\boldsymbol{\phi}\,\mathrm{d}x+\int_{\Gamma_N}\mathbf{a}_N\cdot\boldsymbol{\phi}\,\mathrm{d}s$$

is satisfied for all $\boldsymbol{\phi}\in W^{1,\varphi}_{\Gamma_D}(\Omega)$. Moreover, using modular trace and Poincaré inequalities, we obtain the *a priori* estimate

$$\rho_{\varphi,\Omega}(\nabla\mathbf{u})\leqslant c(\rho_{\varphi,\Omega}(\nabla\mathbf{u}_D)+\rho_{\varphi^*,\Omega}(\mathbf{f})+\rho_{\varphi^*,\Gamma_N}(\mathbf{a}_N)). \tag{2.37}$$

[3] Note that the boundary data $\mathbf{u}_D\in W^{1-1/p,p}(\Gamma_D)$ can be extended to $\Omega$. The extension, which is denoted again by $\mathbf{u}_D$, belongs to $W^{1,\varphi}(\Omega)$.

Using the notation $\mathbf{L} = \nabla \mathbf{u}$, $\mathbf{A} = \boldsymbol{\mathcal{A}}(\mathbf{L})$ we see that $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^{\varphi}(\Omega) \times L^{\varphi^*}(\Omega)$. If additionally $\mathbf{A} \in W^{1,1}(\Omega)$, we get, using partial integration and $\mathbf{A}\mathbf{n} = \mathbf{a}_N$ on $\Gamma_N$, for all $\mathbf{z}_h \in V_h^k$

$$
\begin{aligned}
(\mathbf{f}, \mathbf{z}_h) &= (-\operatorname{div}\mathbf{A}, \mathbf{z}_h)\\
&= (\mathbf{A}, \nabla_h \mathbf{z}_h) - \langle \mathbf{A}, [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} - \langle \mathbf{A}, \mathbf{z}_h \otimes \mathbf{n}\rangle_{\Gamma_D} - \langle \mathbf{A}, \mathbf{z}_h \otimes \mathbf{n}\rangle_{\Gamma_N}\\
&= (\mathbf{A}, \nabla_h \mathbf{z}_h) - \langle \{\mathbf{A}\}, [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I \cup \Gamma_D} - \langle \mathbf{a}_N, \mathbf{z}_h\rangle_{\Gamma_N}.
\end{aligned}
$$

Using the definition of $\Pi_{\mathrm{DG}}$, (2.27) and the definition of the jump functional (2.25), we obtain

$$
\begin{aligned}
(\mathbf{A}, \nabla_h \mathbf{z}_h) &= (\Pi_{\mathrm{DG}}\mathbf{A}, \nabla_h \mathbf{z}_h)\\
&= (\Pi_{\mathrm{DG}}\mathbf{A}, \nabla_{\mathrm{DG}}^h \mathbf{z}_h) + (\Pi_{\mathrm{DG}}\mathbf{A}, \mathbf{R}_h \mathbf{z}_h)\\
&= (\mathbf{A}, \nabla_{\mathrm{DG}}^h \mathbf{z}_h) + \langle \{\Pi_{\mathrm{DG}}\mathbf{A}\}, [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I \cup \Gamma_D}.
\end{aligned}
$$

Thus, the original solution $(\mathbf{u}, \mathbf{L}, \mathbf{A})$ satisfies a system similar to (2.36), namely

$$
\begin{aligned}
(\mathbf{L}, \mathbf{X}_h) &= (\nabla \mathbf{u}, \mathbf{X}_h),\\
(\mathbf{A}, \mathbf{Y}_h) &= (\boldsymbol{\mathcal{A}}(\mathbf{L}), \mathbf{Y}_h),\\
(\mathbf{A}, \nabla_{\mathrm{DG}}^h \mathbf{z}_h) &= (\mathbf{f}, \mathbf{z}_h) + \langle \mathbf{a}_N, \mathbf{z}_h\rangle_{\Gamma_N} + \langle \{\mathbf{A}\} - \{\Pi_{\mathrm{DG}}\mathbf{A}\}, [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I \cup \Gamma_D}
\end{aligned}
\tag{2.38}
$$

for all $(\mathbf{X}_h, \mathbf{Y}_h, \mathbf{z}_h) \in X_h^k \times X_h^k \times V_h^k$. Note that $[\![\mathbf{u} \otimes \mathbf{n}]\!] = \mathbf{0}$ on $\Gamma_I$ and $\mathbf{u} - \mathbf{u}_D = \mathbf{0}$ on $\Gamma_D$, which implies that the last two terms in $(2.36)_3$ for $\mathbf{u}_h$ replaced by $\mathbf{u}$ could be added to $(2.38)_3$ if $\mathbf{u}_D^* = \mathbf{u}_D$. Consequently, the difference between (2.38) and (2.36) lies in the appearance of different gradients in the first equations and in the additional appearance of the projection error in $(2.38)_3$.

## 2.5 *Primal formulation*

We will now eliminate in our LDG system (2.36) the variables $\mathbf{L}_h$ and $\mathbf{A}_h$ to derive a system expressed only in terms of $\mathbf{u}_h$.

To this end, first note that from (2.36) follows

$$
\mathbf{L}_h = \nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*, \tag{2.39}
$$

$$
\mathbf{A}_h = \Pi_{\mathrm{DG}} \boldsymbol{\mathcal{A}}(\mathbf{L}_h). \tag{2.40}
$$

We can plug this into the last equation of (2.36) to obtain

$$
\begin{aligned}
(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*), \nabla_{\mathrm{DG}}^h \mathbf{z}_h) &= (\mathbf{f}, \mathbf{z}_h) + \langle \mathbf{a}_N, \mathbf{z}_h\rangle_{\Gamma_N} - \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h \otimes \mathbf{n}]\!]), [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I}\\
&\quad - \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}), \mathbf{z}_h \otimes \mathbf{n}\rangle_{\Gamma_D}
\end{aligned}
\tag{2.41}
$$

for all $\mathbf{z}_h \in V_h^k$. This is the *primal formulation* of our system.

By standard methods, one can easily prove the existence of a solution $\mathbf{u}_h$ of (2.41). In turn, we also obtain the existence of $\mathbf{L}_h$ and $\mathbf{A}_h$ from (2.39) and (2.40), respectively. This also shows the solvability of (2.36).

REMARK 2.15 Let us consider the situation, where $\boldsymbol{\mathcal{A}}(\mathbf{P}) = \varphi'(|\mathbf{P}|)|\mathbf{P}|^{-1}$. For every $\mathbf{w}_h \in V_h^k$, we define the energy

$$\mathcal{J}_h(\mathbf{w}_h) := \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{w}_h + \mathbf{R}_h \mathbf{u}_D^*) + \alpha m_{\varphi,h}(\mathbf{w}_h - \mathbf{u}_D^*) - (\mathbf{f}, \mathbf{w}_h) - \langle \mathbf{a}_N, \mathbf{w}_h \rangle_{\Gamma_N}.$$

Then the solution $\mathbf{u}_h$ of (2.41) is the unique minimizer of $\mathcal{J}_h$ on $V_h^k$. Indeed, (2.41) just says that the first variation vanishes at $\mathbf{u}_h$, i.e., $(\delta J)(\mathbf{u}_h)(\mathbf{z}_h) = 0$ for all $\mathbf{z}_h \in V_h^k$. Moreover, $\mathcal{J}_h$ is strictly convex, since $\varphi$ is strictly convex due to $p \in (1, \infty)$. The super linearity of $\varphi$ (due to $p > 1$) implies that $\mathcal{J}_h$ is coercive with respect to the norm induced by the modular $\rho_{\varphi,\Omega}(|\nabla_{\mathrm{DG}}^h \cdot|) + m_{\varphi,h}(\cdot)$.

This energy functional was already proposed by Buffa & Ortner (2009). In that article the authors also show the convergence of minimizers $\mathbf{u}_h$ to the solution $\mathbf{u}$ for $h \to 0$. The special case $\Gamma_N = \emptyset$, $\mathbf{u}_D = \mathbf{0}$ and $\varphi(t) = (1/p)t^p$ has also been studied by Burman & Ern (2008). In this situation, $\mathcal{J}_h(\mathbf{w}_h) = \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{w}_h) + \alpha m_{\varphi,h}(\mathbf{w}_h) - (\mathbf{f}, \mathbf{w}_h)$.

## 3. *A priori* estimates

In this section, we derive *a priori* estimates for our DG solution. Let us start showing that for our choices of $\mathbf{u}_D^*$ we can control the jump terms.

LEMMA 3.1 Let $\mathbf{u}_D^* = \mathbf{u}$ or $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$. Then we have

$$m_{\varphi,h}(\mathbf{u}_D^* - \mathbf{u}) \leqslant c\rho_{\varphi,\Omega}(\nabla \mathbf{u}).$$

*Proof.* For $\mathbf{u}_D^* = \mathbf{u}$ the assertion is trivial, while for $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$ it follows from (A.42). □

Now we can formulate the main result of this section.

THEOREM 3.2 Let $\mathbf{u}_D^* = \mathbf{u}$ or $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$ and let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V_h^k \times X_h^k \times X_h^k$ be a solution of (2.36) for some $\alpha > 0$. Then this solution satisfies the following *a priori* estimate

$$\begin{aligned}
&\int_\Omega \varphi(|\mathbf{L}_h|)\,\mathrm{d}x + \alpha h \int_{\Gamma_I} \varphi(h^{-1}|[\![\mathbf{u}_h \otimes \mathbf{n}]\!]|)\,\mathrm{d}s + \alpha h \int_{\Gamma_D} \varphi(h^{-1}|[\![(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}]\!]|)\,\mathrm{d}s \\
&\quad + \int_\Omega \varphi^*(|\mathbf{A}_h|)\,\mathrm{d}x + \min\{1,\alpha\} \int_\Omega \varphi(|\nabla_h \mathbf{u}_h|)\,\mathrm{d}x + \min\{1,\alpha\} \int_\Omega \varphi(|\mathbf{u}_h - \mathbf{u}|)\,\mathrm{d}x \\
&\quad \leqslant c \int_\Omega \varphi(|\nabla \mathbf{u}|)\,\mathrm{d}x + c \int_\Omega \varphi^*(|\mathbf{f}|)\,\mathrm{d}x + c \int_{\Gamma_N} \varphi^*(|\mathbf{a}_N|)\,\mathrm{d}s,
\end{aligned}$$

with $c$ depending only on $\alpha$, the characteristics of $\boldsymbol{\mathcal{A}}$, and the chunkiness $\omega_0$.

*Proof.* To prove the assertion we use in the primal formulation (2.41) the test function $\mathbf{z}_h = \mathbf{u}_h - \Pi_{\mathrm{DG}}\mathbf{u}$. From (2.27),

$$\nabla_{\mathrm{DG}}^h(\mathbf{u}_h - \Pi_{\mathrm{DG}}\mathbf{u}) = (\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - (\mathbf{R}_h \mathbf{u}_D^* - \mathbf{R}_h \Pi_{\mathrm{DG}}\mathbf{u}) - \nabla_h \Pi_{\mathrm{DG}}\mathbf{u}$$

and (2.39) we obtain

$$
\begin{aligned}
&\tfrac{1}{2}(\boldsymbol{\mathcal{A}}(\mathbf{L}_h),\mathbf{L}_h)+\tfrac{1}{2}(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*),\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*)\\
&\quad+\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h\otimes\mathbf{n}]\!]),[\![\mathbf{u}_h\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}+\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}),(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}\rangle_{\Gamma_D}\\
&\qquad=(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*),\nabla_h\Pi_{\mathrm{DG}}\mathbf{u})+(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*),\mathbf{R}_h(\mathbf{u}_D^*-\Pi_{\mathrm{DG}}\mathbf{u}))\\
&\qquad\quad+(\mathbf{f},\mathbf{u}_h-\mathbf{u}+\mathbf{u}-\Pi_{\mathrm{DG}}\mathbf{u})+\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h\otimes\mathbf{n}]\!]),[\![\Pi_{\mathrm{DG}}\mathbf{u}\otimes\mathbf{n}]\!]\rangle_{\Gamma_I}\\
&\qquad\quad-\alpha\langle\boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h-\mathbf{u}_D^*)\otimes\mathbf{n}),(\mathbf{u}_D^*-\Pi_{\mathrm{DG}}\mathbf{u})\otimes\mathbf{n}\rangle_{\Gamma_D}+\langle\mathbf{a}_N,\mathbf{u}_h-\Pi_{\mathrm{DG}}\mathbf{u}\rangle_{\Gamma_N}\\
&\qquad=:I_1+I_2+I_3+I_4+I_5+I_6.
\end{aligned}
\tag{3.1}
$$

Before we estimate the terms $I_i$, $i=1,\ldots,6$, we collect the information we can obtain from the left-hand side. From Lemma 2.11 follows:

$$
\begin{aligned}
\boldsymbol{\mathcal{A}}(\mathbf{L}_h):\mathbf{L}_h&\sim\varphi(|\mathbf{L}_h|),\\
\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*):(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*)&\sim\varphi(|\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*|),\\
\alpha\boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{v}\otimes\mathbf{n}]\!]):[\![\mathbf{v}\otimes\mathbf{n}]\!]&\sim\alpha h\varphi(h^{-1}|[\![\mathbf{v}\otimes\mathbf{n}]\!]|)
\end{aligned}
$$

and consequently the left-hand side of (3.1) is equivalent to

$$
\begin{aligned}
&\rho_{\varphi,\Omega}(\mathbf{L}_h)+\rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*)+\alpha m_{\varphi,h}^I(\mathbf{u}_h)+\alpha m_{\varphi,h}^D(\mathbf{u}_h-\mathbf{u}_D^*)\\
&\quad=\rho_{\varphi,\Omega}(\mathbf{L}_h)+\rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h\mathbf{u}_h+\mathbf{R}_h\mathbf{u}_D^*)+\alpha m_{\varphi,h}(\mathbf{u}_h-\mathbf{u}_D^*),
\end{aligned}
\tag{3.2}
$$

since $[\![\mathbf{u}_D^*\otimes\mathbf{n}]\!]=\mathbf{0}$ on $\Gamma_I$. From (3.2) we can also obtain information about $\rho_{\varphi^*,\Omega}(\mathbf{A}_h)$, $\rho_{\varphi,\Omega}(\nabla_h\mathbf{u}_h)$ and $\rho_{\varphi,\Omega}(\mathbf{u}_h-\mathbf{u})$ as we show in the following.

From (2.40) we know that $\mathbf{A}_h=\Pi_{\mathrm{DG}}(\boldsymbol{\mathcal{A}}(\mathbf{L}_h))$, so the $L^\varphi$-stability of $\Pi_{\mathrm{DG}}$ (cf. (A.12)), (2.13) and (2.1) imply

$$
\rho_{\varphi^*,\Omega}(\mathbf{A}_h)\leqslant c\rho_{\varphi^*,\Omega}(\boldsymbol{\mathcal{A}}(\mathbf{L}_h))\leqslant c\rho_{\varphi^*,\Omega}(\varphi'(|\mathbf{L}_h|))\leqslant c\rho_{\varphi,\Omega}(\mathbf{L}_h).
\tag{3.3}
$$

From (2.39) and (2.27) we obtain

$$
\mathbf{L}_h=\nabla_h\mathbf{u}_h-\mathbf{R}_h(\mathbf{u}_h-\mathbf{u}_D^*).
\tag{3.4}
$$

Therefore, it follows from (A.25) that

$$
\rho_{\varphi,\Omega}(\nabla_h\mathbf{u}_h)\leqslant c\rho_{\varphi,\Omega}(\mathbf{L}_h)+c\rho_{\varphi,\Omega}(\mathbf{R}_h(\mathbf{u}_h-\mathbf{u}_D^*))\leqslant c\rho_{\varphi,\Omega}(\mathbf{L}_h)+cm_{\varphi,h}(\mathbf{u}_h-\mathbf{u}_D^*).
$$

This and Lemma 3.1 yield

$$
\begin{aligned}
M_{\varphi,h}(\mathbf{u}_h-\mathbf{u})&=\rho_{\varphi,\Omega}(\nabla_h(\mathbf{u}_h-\mathbf{u}))+m_{\varphi,h}(\mathbf{u}_h-\mathbf{u})\\
&\leqslant c\rho_{\varphi,\Omega}(\nabla_h\mathbf{u}_h)+cm_{\varphi,h}(\mathbf{u}_h-\mathbf{u}_D^*)+cm_{\varphi,h}(\mathbf{u}_D^*-\mathbf{u})+c\rho_{\varphi,\Omega}(\nabla\mathbf{u})\\
&\leqslant c\rho_{\varphi,\Omega}(\nabla_h\mathbf{u}_h)+cm_{\varphi,h}(\mathbf{u}_h-\mathbf{u}_D^*)+c\rho_{\varphi,\Omega}(\nabla\mathbf{u}).
\end{aligned}
\tag{3.5}
$$

Putting everything together and using Lemma A.9, we obtain that the left-hand side of (3.1) is larger than

$$c(\rho_{\varphi,\Omega}(\mathbf{L}_h) + \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) + \rho_{\varphi^*,\Omega}(\mathbf{A}_h) + \alpha m_{\varphi,h}(\mathbf{u}_h - \mathbf{u}_D^*)$$
$$+ \min\{1,\alpha\}(\rho_{\varphi,\Omega}(\mathbf{u}_h - \mathbf{u}) + \rho_{\varphi,\Omega}(\nabla_h \mathbf{u}_h) + M_{\varphi,h}(\mathbf{u}_h - \mathbf{u}) - c\rho_{\varphi,\Omega}(\nabla \mathbf{u}))).$$

Now we can estimate the terms on the right-hand side of (3.1). We have, using (2.13), Young's inequality, (2.1) and the $L^\varphi$-gradient stability of $\Pi_{\mathrm{DG}}$ (cf. (A.12)),

$$\begin{aligned}
|I_1| &\leqslant \varepsilon \rho_{\varphi^*,\Omega}(\varphi'(|\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*|)) + c_\varepsilon \rho_{\varphi,\Omega}(\nabla_h \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) + c_\varepsilon \rho_{\varphi,\Omega}(\nabla \mathbf{u}),
\end{aligned}$$

and in addition with (A.25), Lemma 3.1 and (A.16)

$$\begin{aligned}
|I_2| &\leqslant \varepsilon \rho_{\varphi^*,\Omega}(\varphi'(|\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*|)) + c_\varepsilon \rho_{\varphi,\Omega}(\mathbf{R}_h(\mathbf{u}_D^* - \Pi_{\mathrm{DG}} \mathbf{u})) \\
&\leqslant \varepsilon \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) + c_\varepsilon m_{\varphi,h}(\mathbf{u}_D^* - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) + c_\varepsilon m_{\varphi,h}(\mathbf{u}_D^* - \mathbf{u}) + c_\varepsilon m_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon \rho_{\varphi,\Omega}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) + c_\varepsilon \rho_{\varphi,\Omega}(\nabla \mathbf{u}).
\end{aligned}$$

Next, we estimate with Young's inequality, and (A.9)

$$\begin{aligned}
|I_3| &\leqslant \varepsilon \rho_{\varphi,\Omega}(\mathbf{u}_h - \mathbf{u}) + c_\varepsilon \rho_{\varphi^*,\Omega}(\mathbf{f}) + c\rho_{\varphi,\Omega}(\mathbf{u} - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon \rho_{\varphi,\Omega}(\mathbf{u}_h - \mathbf{u}) + c_\varepsilon \rho_{\varphi^*,\Omega}(\mathbf{f}) + c\rho_{\varphi,\Omega}(h\nabla \mathbf{u}).
\end{aligned}$$

Using $[\![\mathbf{u}_D^* \otimes \mathbf{n}]\!] = \mathbf{0}$ on $\Gamma_I$, (2.13), Young's inequality, (2.1), Lemma 3.1 and (A.16), we obtain

$$\begin{aligned}
|I_4 + I_5| &= \alpha h |\langle \boldsymbol{\mathcal{A}}(h^{-1}[\![(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}]\!]), h^{-1}[\![(\mathbf{u}_D^* - \Pi_{\mathrm{DG}} \mathbf{u}) \otimes \mathbf{n}]\!]\rangle_{\Gamma_I \cup \Gamma_D}| \\
&\leqslant \varepsilon h \alpha \rho_{\varphi^*,\Gamma_I \cup \Gamma_D}(\varphi'(|h^{-1}[\![(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}]\!]|)) + c_\varepsilon \alpha m_{\varphi,h}(\mathbf{u}_D^* - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon m_{\varphi,h}(\mathbf{u}_h - \mathbf{u}_D^*) + c_\varepsilon \alpha m_{\varphi,h}(\mathbf{u}_D^* - \mathbf{u}) + c_\varepsilon \alpha m_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant \varepsilon \alpha m_{\varphi,h}(\mathbf{u}_h - \mathbf{u}_D^*) + c_\varepsilon \alpha \rho_{\varphi,\Omega}(\nabla \mathbf{u}).
\end{aligned}$$

Finally, we estimate with Lemma A.10 and (A.17)

$$\begin{aligned}
|I_6| &\leqslant c_\varepsilon \rho_{\varphi^*,\Gamma_N}(\mathbf{a}_N) + \varepsilon \rho_{\varphi,\Gamma_N}(\mathbf{u}_h - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant c_\varepsilon \rho_{\varphi^*,\Gamma_N}(\mathbf{a}_N) + \varepsilon c \rho_{\varphi,\Gamma_N}(\mathbf{u}_h - \mathbf{u}) + \varepsilon c \rho_{\varphi,\Gamma_N}(\mathbf{u} - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant c_\varepsilon \rho_{\varphi^*,\Gamma_N}(\mathbf{a}_N) + \varepsilon c M_{\varphi,h}(\mathbf{u}_h - \mathbf{u}) + \varepsilon c M_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{DG}} \mathbf{u}) \\
&\leqslant c_\varepsilon \rho_{\varphi^*,\Gamma_N}(\mathbf{a}_N) + \varepsilon c M_{\varphi,h}(\mathbf{u}_h - \mathbf{u}) + \varepsilon c \rho_{\varphi,\Omega}(\nabla \mathbf{u}).
\end{aligned}$$

Choosing $\varepsilon$ small enough, we then obtain the assertion of Theorem 3.2. □

## 4. Error estimates

Using (2.41) and (2.38), we obtain our error equation

$$
\begin{aligned}
&(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \boldsymbol{\mathcal{A}}(\nabla \mathbf{u}), \nabla_{\mathrm{DG}}^h \mathbf{z}_h) + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h \otimes \mathbf{n}]\!]), [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} \\
&\quad + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}((\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n})), \mathbf{z}_h \otimes \mathbf{n}\rangle_{\Gamma_D} \\
&\qquad = \langle \{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\}, [\![\mathbf{z}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I \cup \Gamma_D}.
\end{aligned} \tag{4.1}
$$

We start with a Strang-type error estimate.

Proposition 4.1 Let $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^{\varphi}(\Omega) \times W^{1,\varphi^*}(\Omega)$ be a solution of (1.2). Further, let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V_h^k \times X_h^k \times X_h^k$ be a solution of (2.36) for some $\alpha > 0$ and some approximation $\mathbf{u}_D^* \in W^{1,\varphi}(\Omega)$ of $\mathbf{u}_D$. Then we have for all $\mathbf{w}_h \in V_h^k$

$$
\begin{aligned}
&\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \mathbf{u}_D^*) \\
&\quad \leqslant c\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{w}_h + \mathbf{R}_h \mathbf{u}_D^*) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + c\alpha m_{\varphi,h}^I(\mathbf{w}_h) + c\alpha m_{\varphi,h}^D(\mathbf{w}_h - \mathbf{u}_D^*) \\
&\quad\quad + c_\alpha h \rho_{\varphi^*, \Gamma_I \cup \Gamma_D}(\{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\})
\end{aligned} \tag{4.2}
$$

[4]with a constant $c$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$ and a constant $c_\alpha$ depending on $p$ and $\alpha^{-1}$. Moreover, note that $\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 = \|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2$.

*Proof.* Let $\mathbf{z}_h := \mathbf{u}_h - \mathbf{w}_h$, then

$$
\nabla_{\mathrm{DG}}^h \mathbf{z}_h = (\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^* - \nabla \mathbf{u}) - (\nabla_{\mathrm{DG}}^h \mathbf{w}_h + \mathbf{R}_h \mathbf{u}_D^* - \nabla \mathbf{u}),
$$

so the error equation (4.1) gives

$$
\begin{aligned}
&(\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \boldsymbol{\mathcal{A}}(\nabla \mathbf{u}), (\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \nabla \mathbf{u}) \\
&\quad + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h \otimes \mathbf{n}]\!]), [\![\mathbf{u}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}), (\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}\rangle_{\Gamma_D} \\
&\quad = (\boldsymbol{\mathcal{A}}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \boldsymbol{\mathcal{A}}(\nabla \mathbf{u}), (\nabla_{\mathrm{DG}}^h \mathbf{w}_h + \mathbf{R}_h \mathbf{u}_D^*) - \nabla \mathbf{u}) \\
&\qquad + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}[\![\mathbf{u}_h \otimes \mathbf{n}]\!]), [\![\mathbf{w}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} + \alpha \langle \boldsymbol{\mathcal{A}}(h^{-1}(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}), (\mathbf{w}_h - \mathbf{u}_D^*) \otimes \mathbf{n}\rangle_{\Gamma_D} \\
&\qquad - \langle \{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\}, [\![\mathbf{w}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} - \langle \{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\}, [\![(\mathbf{w}_h - \mathbf{u}_D^*) \otimes \mathbf{n}]\!]\rangle_{\Gamma_D} \\
&\qquad + \langle \{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\}, [\![\mathbf{u}_h \otimes \mathbf{n}]\!]\rangle_{\Gamma_I} + \langle \{\Pi_{\mathrm{DG}} \mathbf{A}\} - \{\mathbf{A}\}, [\![(\mathbf{u}_h - \mathbf{u}_D^*) \otimes \mathbf{n}]\!]\rangle_{\Gamma_D}.
\end{aligned}
$$

Now Young's inequality and Lemma 2.7 prove (4.2). Note that for the treatment of the last two terms, one has to choose an appropiate $\varepsilon$ to absorb the terms with $\mathbf{u}_h$ on the left-hand side. The second part follows from (2.39). □

From this Strang-type error estimate we will derive our error estimates by choosing $\mathbf{w}_h = \Pi_{\mathrm{SZ}} \mathbf{u}$. Since $\mathbf{u}$ is a standard Sobolev function, our extended Scott–Zhang operator applied to $\mathbf{u}$ is just the

---

[4] Note that $m_{\varphi,h}^I(\mathbf{u}_h) = m_{\varphi,h}^I(\mathbf{u}_h - \mathbf{u})$ since $\mathbf{u} \in W^{1,\varphi}(\Omega)$.

standard Scott–Zhang operator. Therefore, we can use the following result from Diening & Růžička (2007, Corollary 5.4).

PROPOSITION 4.2 Let $k \geqslant 1$ and $\mathbf{F}(\nabla\mathbf{u}) \in W^{1,2}(\Omega)$, then

$$\|\mathbf{F}(\nabla_h \Pi_{\mathrm{SZ}}\mathbf{u}) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 \leqslant ch^2 \|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2$$

with $c$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$.

Let us begin with the simple situation $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$.

THEOREM 4.3 Let $k \geqslant 1$. Let $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^{\varphi}(\Omega) \times W^{1,\varphi^*}(\Omega)$ be a solution of (1.2) and let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V_h^k \times X_h^k \times X_h^k$ be a solution of (2.36) for some $\alpha > 0$ and $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$. If $\mathbf{F}(\nabla\mathbf{u}) \in W^{1,2}(\Omega)$, then

$$\begin{aligned}
&\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h\mathbf{u}_D^*) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \Pi_{\mathrm{SZ}}\mathbf{u}) \\
&\quad \leqslant ch^2 \|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2 + c_\alpha h^{\min(2,p')} \rho_{\varphi^*,\Omega}(\nabla\mathbf{A})
\end{aligned}$$

with $c$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$ and the chunkiness $\omega_0$ and a constant $c_\alpha$ depending on $p, \alpha^{-1}$ and the chunkiness $\omega_0$. Moreover, note that $\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h\mathbf{u}_D^*) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 = \|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2$.

*Proof.* Since $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u} \in W^{1,\varphi}(\Omega)$ we have

$$\begin{aligned}
m_{\varphi,h}^I(\Pi_{\mathrm{SZ}}\mathbf{u}) &= m_{\varphi,h}^D(\Pi_{\mathrm{SZ}}\mathbf{u} - \mathbf{u}_D^*) = 0, \\
\nabla_{\mathrm{DG}}^h \Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h\mathbf{u}_D^* &= \nabla_h \Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h(\mathbf{u}_D^* - \Pi_{\mathrm{SZ}}\mathbf{u}) = \nabla_h \Pi_{\mathrm{SZ}}\mathbf{u}.
\end{aligned}$$

Therefore, the choice $\mathbf{w}_h := \Pi_{\mathrm{SZ}}\mathbf{u}$ in Proposition 4.1 implies

$$\begin{aligned}
&\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h\mathbf{u}_D^*) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \Pi_{\mathrm{SZ}}\mathbf{u}) \\
&\quad \leqslant c\|\mathbf{F}(\nabla_h \Pi_{\mathrm{SZ}}\mathbf{u}) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 + c_\alpha h \rho_{\varphi^*,\Gamma_I\cup\Gamma_D}(\{\Pi_{\mathrm{DG}}\mathbf{A}\} - \{\mathbf{A}\}).
\end{aligned}$$

Now the first term on the right-hand side is treated by Proposition 4.2, while for the second one we proceed as follows:

$$\begin{aligned}
h\int_{\Gamma_I\cup\Gamma_D} \varphi^*(\{\Pi_{\mathrm{DG}}\mathbf{A}\} - \{\mathbf{A}\})\,\mathrm{d}s &\leqslant c\sum_{K\in\mathcal{T}_h} h\int_{\partial K\cap\operatorname{int}\Omega'} \varphi^*(|\Pi_{\mathrm{DG}}\mathbf{A} - \mathbf{A}|)\,\mathrm{d}s \\
&\leqslant c\sum_{K\in\mathcal{T}_h} \int_K \varphi^*(h|\nabla\mathbf{A}|)\,\mathrm{d}x \\
&\leqslant c\sum_{K\in\mathcal{T}_h} h^{\min(2,p')} \int_K \varphi^*(|\nabla\mathbf{A}|)\,\mathrm{d}x,
\end{aligned}$$

where we used (A.15) and

$$\varphi^*(ht) \leqslant ch^{\min(2,p')}\varphi^*(t),$$

which holds uniformly with respect to $t, \delta \geqslant 0$ and $h \in (0, 1]$. This completes the proof. □

In the case $p \leqslant 2$, we can improve the previous result. For this we need the following lemma on averages, which is useful to interchange averages and also to get later an error estimate in terms of $\mathbf{A}$ and $\mathbf{A}_h$, which is best expressed by $\|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_{2,\Omega}^2$.

LEMMA 4.4 For all $K \in \mathcal{T}_h$ and all functions $\mathbf{P}\colon \Omega \to \mathbb{R}^{d\times n}$

$$\fint_K |\mathbf{F}(\mathbf{P}) - \langle \mathbf{F}(\mathbf{P})\rangle_K|^2 \,\mathrm{d}x \sim \fint_K |\mathbf{F}(\mathbf{P}) - \mathbf{F}(\langle \mathbf{P}\rangle_K)|^2 \,\mathrm{d}x$$
$$\sim \fint_K |\mathbf{F}^*(\boldsymbol{\mathcal{A}}(\mathbf{P})) - \mathbf{F}^*(\langle \boldsymbol{\mathcal{A}}(\mathbf{P})\rangle_K)|^2 \,\mathrm{d}x.$$

*Proof.* The case $\boldsymbol{\mathcal{A}}(\mathbf{P}) = \varphi'(|\mathbf{P}|)(\mathbf{P}/|\mathbf{P}|)$ is just (Diening *et al.*, 2012, Lemma 6.2). The proof in the general case is word by word the same and we therefore omit it. It is based on Lemma 2.7. □

THEOREM 4.5 Under the assumptions of Theorem 4.3, we have in the case $p \leqslant 2$

$$\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \Pi_{\mathrm{SZ}}\mathbf{u}) \leqslant c_\alpha h^2 \|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2$$

with $c_\alpha$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$, the chunkiness $\omega_0$ and on $\alpha^{-1}$. Moreover, note that $\|\mathbf{F}(\nabla_{\mathrm{DG}}^h \mathbf{u}_h + \mathbf{R}_h \mathbf{u}_D^*) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 = \|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2$.

*Proof.* In order to prove the statement, we just treat the projection error in a different way. Starting from

$$h\int_{\Gamma_I\cup\Gamma_D} \varphi^*(\{\Pi_{\mathrm{DG}}\mathbf{A}\} - \{\mathbf{A}\})\,\mathrm{d}s \leqslant c \sum_{K\in\mathcal{T}_h} h \int_{\partial K\cap \operatorname{int}\Omega'} \varphi^*(|\Pi_{\mathrm{DG}}\mathbf{A} - \mathbf{A}|)\,\mathrm{d}s$$

we obtain for any $\gamma \in \Gamma_I \cup \Gamma_D$ being a face of some $K \in \mathcal{T}_h$

$$\begin{aligned} h\int_\gamma \varphi^*(|\Pi_{\mathrm{DG}}\mathbf{A} - \mathbf{A}|)\,\mathrm{d}s \leqslant ch \int_\gamma & \varphi^*(|\Pi_{\mathrm{DG}}(\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K))|) \\ & + \varphi^*(|\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K)|)\,\mathrm{d}s =: I_1 + I_2, \end{aligned} \tag{4.3}$$

where we used that $\mathbf{A} = \boldsymbol{\mathcal{A}}(\mathbf{L}) = \boldsymbol{\mathcal{A}}(\nabla\mathbf{u})$. To treat $I_1$, we use (A.14) and the continuity of $\Pi_{\mathrm{DG}}$ (cf. (A.6)) to obtain

$$\begin{aligned} I_1 &\leqslant c\int_K \varphi^*(|\Pi_{\mathrm{DG}}(\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K))|)\,\mathrm{d}x \\ &\leqslant c\int_K \varphi^*(|\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K)|)\,\mathrm{d}x =: I_3. \end{aligned} \tag{4.4}$$

Using

$$\varphi^*(t) \leqslant (\varphi^*)_a(t), \tag{4.5}$$

which holds for $p \leqslant 2$ uniformly with respect to $a, t \geqslant 0$, we obtain

$$\begin{aligned} I_2 + I_3 &\leqslant ch \int_\gamma (\varphi^*)_{|\boldsymbol{\mathcal{A}}(\langle \mathbf{L}\rangle_K)|}(|\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle \mathbf{L}\rangle_K)|)\,\mathrm{d}s + c\int_K (\varphi^*)_{|\boldsymbol{\mathcal{A}}(\langle \mathbf{L}\rangle_K)|}(|\boldsymbol{\mathcal{A}}(\mathbf{L}) - \boldsymbol{\mathcal{A}}(\langle \mathbf{L}\rangle_K)|)\,\mathrm{d}x \\ &\leqslant ch \int_\gamma |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\langle \nabla \mathbf{u}\rangle_K)|^2\,\mathrm{d}s + c\int_K |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\langle \nabla \mathbf{u}\rangle_K)|^2\,\mathrm{d}x \\ &\leqslant c\int_K |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\langle \nabla \mathbf{u}\rangle_K)|^2\,\mathrm{d}s + ch^2\int_K |\nabla \mathbf{F}(\nabla \mathbf{u})|^2\,\mathrm{d}x \\ &\leqslant ch^2 \int_K |\nabla \mathbf{F}(\nabla \mathbf{u})|^2\,\mathrm{d}x, \end{aligned} \tag{4.6}$$

where we also used Lemma 2.8, Proposition 2.7, $\mathbf{L} = \nabla \mathbf{u}$; Lemma A.1 for $\psi(t) = t^2$, Poincaré's lemma on $W^{1,2}(K)$ and finally Lemma 4.4 and again Poincaré's lemma on $W^{1,2}(K)$. Putting together the estimates (4.3–4.6), we obtained the desired estimate of the projection error. This finishes the proof. □

For the jump projection error on the boundary $\Gamma_D$ we have:

Lemma 4.6 Let $k \geqslant 1$ and let $\mathbf{F}(\nabla \mathbf{u}) \in W^{1,2}(\Omega)$, then

$$m^D_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) \leqslant m_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) \leqslant ch^{\min\{2,p\}}(\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla \mathbf{u})).$$

*Proof.* Clearly, we have

$$m^D_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) \leqslant m_{\varphi,h}(\mathbf{u}_h - \Pi_{\mathrm{SZ}}\mathbf{u}).$$

To estimate the last term, we distinguish the cases $p \leqslant 2$ and $p \geqslant 2$.

*The case $p \leqslant 2$:* Using

$$\varphi(ht) \leqslant ch^{\min\{2,p\}}\varphi(t),$$

which holds uniformly with respect to $h \in (0,1)$, $t \geqslant 0$ and $\delta \geqslant 0$, and (A.43), we obtain

$$m_{\varphi,h}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) \leqslant c\sum_{K\in\mathcal{T}_h}\int_{S_K} \varphi(|h\nabla^2\mathbf{u}|)\,\mathrm{d}x \leqslant ch^p\int_\Omega \varphi(|\nabla^2\mathbf{u}|)\,\mathrm{d}x \leqslant ch^p.$$

In the last step we also used that for $p \leqslant 2$ we have (cf. Berselli *et al.*, 2010, Lemma 4.4)

$$\rho_{\varphi,\Omega}(\nabla^2\mathbf{u}) \leqslant c\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + c\rho_{\varphi,\Omega}(\nabla \mathbf{u}). \tag{4.7}$$

*The case $p \geqslant 2$:* In this case, we proceed slightly differently, since we do not have $\mathbf{u} \in W^{2,\varphi}(\Omega)$ for $p \geqslant 2$. Since

$$\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u} = (\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) - \Pi_{\mathrm{SZ}}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}), \tag{4.8}$$

we obtain from (A.40) for $\mathbf{g} = h^{-1}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})$, the estimate

$$\varphi(t) \leqslant \varphi_a(t) \tag{4.9}$$

which holds for $p \geqslant 2$ uniformly with respect to $a, t \geqslant 0$ and (2.11) that

$$
\begin{aligned}
h \int_{\gamma} \varphi(h^{-1}|\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}|)\,\mathrm{d}s &\leqslant c \int_{S_K} \varphi(|\nabla \mathbf{u} - \nabla \Pi_{\mathrm{SZ}}\mathbf{u}|)\,\mathrm{d}x \\
&\leqslant c \int_{S_K} \varphi_{|\nabla \mathbf{u}|}(|\nabla \mathbf{u} - \nabla \Pi_{\mathrm{SZ}}\mathbf{u})|)\,\mathrm{d}x \\
&\leqslant c \int_{S_K} |\mathbf{F}(\nabla \mathbf{u}) - \mathbf{F}(\nabla \Pi_{\mathrm{SZ}}\mathbf{u})|^2\,\mathrm{d}x.
\end{aligned}
$$

After summation over the traces we treat the last term by Proposition 4.2. This proves the claim. □

From the previous lemma, Theorems 4.3, 4.5 and the convexity of $\varphi$, we obtain the following corollary:

Corollary 4.7 Under the assumptions of Theorem 4.3, we have for the jump error on the boundary $\Gamma_D$

$$
m^D_{\varphi,h}(\mathbf{u} - \mathbf{u}_h) \leqslant c_\alpha \begin{cases} h^p(\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla \mathbf{u})) & \text{if } p \leqslant 2, \\ h^{p'}(\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla \mathbf{u}) + \rho_{\varphi^*,\Omega}(\nabla \mathbf{A})) & \text{if } p \geqslant 2 \end{cases}
$$

with $c_\alpha$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$, the chunkiness $\omega_0$ and on $\alpha^{-1}$.

Let us now consider the other case $\mathbf{u}^*_D = \mathbf{u}$.

Theorem 4.8 Let $k \geqslant 1$. Let $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^\varphi(\Omega) \times W^{1,\varphi^*}(\Omega)$ be a solution of (1.2) and let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V^k_h \times X^k_h \times X^k_h$ be a solution of (2.36) for some $\alpha > 0$ and $\mathbf{u}^*_D = \mathbf{u}$. If $\mathbf{F}(\nabla \mathbf{u}) \in W^{1,2}(\Omega)$, then

$$
\begin{aligned}
&\|\mathbf{F}(\nabla^h_{\mathrm{DG}}\mathbf{u}_h + \mathbf{R}_h\mathbf{u}) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \alpha m^I_{\varphi,h}(\mathbf{u}_h) + \alpha m^D_{\varphi,h}(\mathbf{u}_h - \mathbf{u}) \\
&\quad\leqslant c_\alpha \begin{cases} h^p(\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla \mathbf{u})) & \text{if } p \leqslant 2, \\ h^{p'}(\|\nabla \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla \mathbf{u}) + \rho_{\varphi^*,\Omega}(\nabla \mathbf{A})) & \text{if } p \geqslant 2 \end{cases}
\end{aligned}
$$

with $c_\alpha$ depending only on the characteristics of $\boldsymbol{\mathcal{A}}$, the chunkiness $\omega_0$ and $\alpha^{-1}$. Moreover, note that $\|\mathbf{F}(\nabla^h_{\mathrm{DG}}\mathbf{u}_h + \mathbf{R}_h\mathbf{u}) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 = \|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2$.

*Proof.* We use $\mathbf{u}^*_D = \mathbf{u}$, $\mathbf{w}_h := \Pi_{\mathrm{SZ}}\mathbf{u}$ in Proposition 4.1, the definition of $m_{\varphi,h}$ and $[\![\Pi_{\mathrm{SZ}}\mathbf{u} \otimes \mathbf{n}]\!] = [\![\mathbf{u} \otimes \mathbf{n}]\!] = \mathbf{0}$ on $\Gamma_I$ to obtain

$$
\begin{aligned}
&\|\mathbf{F}(\nabla^h_{\mathrm{DG}}\mathbf{u}_h + \mathbf{R}_h\mathbf{u}) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + \alpha m^I_{\varphi,h}(\mathbf{u}_h) + \alpha m^D_{\varphi,h}(\mathbf{u}_h - \mathbf{u}) \\
&\quad\leqslant c\|\mathbf{F}(\nabla^h_{\mathrm{DG}}\Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h\mathbf{u}) - \mathbf{F}(\nabla \mathbf{u})\|_2^2 + c\alpha m_{\varphi,h}(\Pi_{\mathrm{SZ}}\mathbf{u} - \mathbf{u}) \\
&\qquad + c_\alpha h \rho_{\varphi^*,\Gamma_I \cup \Gamma_D}(\{\Pi_{\mathrm{DG}}\mathbf{A}\} - \{\mathbf{A}\}) =: I_1 + I_2 + I_3.
\end{aligned}
$$

The term $I_3$ is treated as in Theorem 4.3 for $p \geqslant 2$ or as in Theorem 4.5 for $p \leqslant 2$. The term $I_2$ is handled by Lemma 4.6. Thus, we concentrate in the remaining proof on the treatment of $I_1$. To treat $I_1$ we use

$$
\nabla^h_{\mathrm{DG}}\Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h\mathbf{u} = \nabla_h \Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}).
$$

Thus, we obtain by Proposition 2.7, triangle inequality and again Proposition 2.7

$$
\begin{aligned}
I_1 &\leqslant \int_\Omega \varphi_{|\nabla\mathbf{u}|}(|\nabla_h \Pi_{\mathrm{SZ}}\mathbf{u} + \mathbf{R}_h(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}) - \nabla\mathbf{u}|)\,\mathrm{d}x \\
&\leqslant c\int_\Omega \varphi_{|\nabla\mathbf{u}|}(|\nabla_h \Pi_{\mathrm{SZ}}\mathbf{u} - \nabla\mathbf{u}|)\,\mathrm{d}x + c\int_\Omega \varphi_{|\nabla\mathbf{u}|}(|\mathbf{R}_h(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})|)\,\mathrm{d}x \\
&\leqslant c\|\mathbf{F}(\nabla_h \Pi_{\mathrm{SZ}}\mathbf{u}) - \mathbf{F}(\nabla\mathbf{u})\|_2^2 + c\int_\Omega \varphi_{|\nabla\mathbf{u}|}(|\mathbf{R}_h(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})|)\,\mathrm{d}x \\
&=: I_{1,1} + I_{1,2}.
\end{aligned}
$$

We can estimate $I_{1,1} \leqslant ch^2\|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2$ by Proposition 4.2. Using the properties of $\mathbf{R}_h$, namely $\mathbf{R}_h\mathbf{g} = \sum_{\gamma\in\Gamma_D}\mathbf{R}_h^\gamma\mathbf{g}$ for $\mathbf{g}\in W^{1,1}(\Omega)$ and $\operatorname{supp}\mathbf{R}_h^\gamma \subset S_\gamma$ are pairwise disjoint for $\gamma\in\Gamma_D$ we get[5]

$$
I_{1,2} = \sum_{\gamma\in\Gamma_D}\int_{S_\gamma} \varphi_{|\nabla\mathbf{u}|}(|\mathbf{R}_h^\gamma(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})|)\,\mathrm{d}x.
$$

Moreover, we have using Corollary 2.9, (A.23) with the definition of the jumps on $\Gamma_D$, (2.11) and Poincaré's lemma on $W^{1,2}(S_{S_\gamma})$

$$
\begin{aligned}
I_{1,2} &\leqslant c\sum_{\gamma\in\Gamma_D}\int_{S_\gamma} \varphi_{|\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}}|}(|\mathbf{R}_h^\gamma(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})|) + |\mathbf{F}(\nabla\mathbf{u}) - \mathbf{F}(\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}})|^2\,\mathrm{d}x \\
&\leqslant c\sum_{\gamma\in\Gamma_D}\left(h_\gamma\int_\gamma \varphi_{|\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}}|}(h_\gamma^{-1}|(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})\otimes\mathbf{n}|)\,\mathrm{d}s + \int_{S_{S_\gamma}} |\mathbf{F}(\nabla\mathbf{u}) - \mathbf{F}(\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}})|^2\,\mathrm{d}x\right) \\
&\leqslant c\sum_{\gamma\in\Gamma_D} h_\gamma\int_\gamma \varphi_{|\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}}|}(h_\gamma^{-1}|\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}|)\,\mathrm{d}s + ch^2\|\nabla\mathbf{F}(\nabla\mathbf{u})\|_{2,\Omega}^2.
\end{aligned}
$$

We obtain from (4.8), (A.40) for $\mathbf{g} = h^{-1}(\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u})$, (2.11), Corollary 2.9 and Poincaré's lemma on $W^{1,2}(S_{S_\gamma})$ that

$$
\begin{aligned}
&h\int_\gamma \varphi_{|\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}}|}(h^{-1}|\mathbf{u} - \Pi_{\mathrm{SZ}}\mathbf{u}|)\,\mathrm{d}s \\
&\quad\leqslant c\int_{S_{S_\gamma}} \varphi_{|\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}}|}(|\nabla\mathbf{u} - \nabla\Pi_{\mathrm{SZ}}\mathbf{u}|)\,\mathrm{d}x \\
&\quad\leqslant c\int_{S_{S_\gamma}} \varphi_{|\nabla\mathbf{u}|}(|\nabla\mathbf{u} - \nabla\Pi_{\mathrm{SZ}}\mathbf{u})|)\,\mathrm{d}x + c\int_{S_{S_\gamma}} |\mathbf{F}(\nabla\mathbf{u}) - \mathbf{F}(\langle\nabla\mathbf{u}\rangle_{S_{S_\gamma}})|^2\,\mathrm{d}x \\
&\quad\leqslant c\int_{S_{S_\gamma}} |\mathbf{F}(\nabla\mathbf{u}) - \mathbf{F}(\nabla\Pi_{\mathrm{SZ}}\mathbf{u})|^2\,\mathrm{d}x + ch_\gamma^2\int_{S_{S_\gamma}} |\nabla\mathbf{F}(\nabla\mathbf{u})|^2\,\mathrm{d}x.
\end{aligned}
$$

[5] Recall that for $\gamma\in\Gamma_D$ we denoted by $S_\gamma$ the only simplex in $\mathcal{T}_h$ that has $\gamma$ as a face.

Collecting the above estimates and using Proposition 4.2, we obtain

$$I_1 \leqslant c\|\mathbf{F}(\nabla\mathbf{u}) - \mathbf{F}(\nabla\Pi_{\mathrm{SZ}}\mathbf{u})\|_2^2 + ch^2\|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2 \leqslant ch^2\|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2.$$

This proves the claim. □

Now we want to derive an estimate for the error between $\mathbf{A}$ and $\mathbf{A}_h$, which is best expressed by $\|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2^2$. Recall that $\mathbf{A} = \boldsymbol{\mathcal{A}}(\mathbf{L})$ and $\mathbf{A}_h = \Pi_{\mathrm{DG}}(\boldsymbol{\mathcal{A}}(\mathbf{L}_h))$ by (2.40).

PROPOSITION 4.9 Let $k \geqslant 1$. Let $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^{\varphi}(\Omega) \times L^{\varphi^*}(\Omega)$ be a solution of (1.2) and let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V_h^k \times X_h^k \times X_h^k$ be a solution of (2.36) for some $\alpha > 0$. Then

$$\|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2^2 \leqslant c\|\mathbf{F}(\mathbf{L}) - \mathbf{F}(\mathbf{L}_h)\|_2^2 + ch^2\|\nabla\mathbf{F}(\nabla\mathbf{u})\|_2^2,$$

with constants depending only on the characteristics of $\boldsymbol{\mathcal{A}}$ and the chunkiness $\omega_0$.

*Proof.* From (2.11c) it follows that

$$\|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2^2 \sim \sum_{K\in\mathcal{T}_h} \int_K (\varphi^*)_{|\mathbf{A}|}(|\mathbf{A} - \mathbf{A}_h|). \tag{4.10}$$

Now by the triangle inequality, the uniform $\Delta_2$-constant for the family $(\varphi^*)_a$ for $a \geqslant 0$, and $\mathbf{A}_h = \Pi_{\mathrm{DG}}\boldsymbol{\mathcal{A}}(\mathbf{L}_h)$ we get for $K \in \mathcal{T}_h$

$$\begin{aligned}\int_K (\varphi^*)_{|\mathbf{A}|}(|\mathbf{A} - \mathbf{A}_h|)\,\mathrm{d}x &\leqslant \int_K (\varphi^*)_{|\mathbf{A}|}(|\mathbf{A} - \Pi_{\mathrm{DG}}\mathbf{A}|)\,\mathrm{d}x + \int_K (\varphi^*)_{|\mathbf{A}|}(|\Pi_{\mathrm{DG}}\mathbf{A} - \mathbf{A}_h|)\,\mathrm{d}x\\ &= \int_K (\varphi^*)_{|\mathbf{A}|}(|\mathbf{A} - \Pi_{\mathrm{DG}}\mathbf{A}|)\,\mathrm{d}x + \int_K (\varphi^*)_{|\mathbf{A}|}(|\Pi_{\mathrm{DG}}(\mathbf{A} - \boldsymbol{\mathcal{A}}(\mathbf{L}_h))|)\,\mathrm{d}x\\ &=: T_1 + T_2.\end{aligned}$$

We estimate with Corollary 2.9 (applied to $\varphi^*$ and $\mathbf{F}^*$)

$$\begin{aligned}T_1 &\leqslant c\int_K (\varphi^*)_{|\langle\mathbf{A}\rangle_K|}(|\mathbf{A} - \Pi_{\mathrm{DG}}\mathbf{A}|)\,\mathrm{d}x + c\int_K |\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\langle\mathbf{A}\rangle_K)|^2\,\mathrm{d}x\\ &=: T_{1,1} + T_{1,2}.\end{aligned}$$

Since $\Pi_{\mathrm{DG}}\boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K) = \boldsymbol{\mathcal{A}}(\langle\mathbf{L}\rangle_K)$, we estimate with (A.6) and Lemma 2.7

$$\begin{aligned}T_{1,1} &\leqslant c\int_K (\varphi^*)_{|\langle\mathbf{A}\rangle_K|}(|\mathbf{A} - \langle\mathbf{A}\rangle_K|)\,\mathrm{d}x + c\int_K (\varphi^*)_{|\langle\mathbf{A}\rangle_K|}(|\Pi_{\mathrm{DG}}(\mathbf{A} - \langle\mathbf{A}\rangle_K)|)\,\mathrm{d}x\\ &\leqslant c\int_K (\varphi^*)_{|\langle\mathbf{A}\rangle_K|}(|\mathbf{A} - \langle\mathbf{A}\rangle_K|)\,\mathrm{d}x\\ &\leqslant c\int_K |\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\langle\mathbf{A}\rangle_K)|^2\,\mathrm{d}x\\ &= cT_{1,2}.\end{aligned}$$

Similarly, we estimate with Corollary 2.9, (A.6), again Corollary 2.9, and Lemma 2.7

$$
\begin{aligned}
T_2 &\leqslant c \int_K (\varphi^*)_{|\langle \mathbf{A}\rangle_K|}(|\Pi_{\mathrm{DG}}(\mathbf{A} - \boldsymbol{\mathcal{A}}(\mathbf{L}_h))|)\,\mathrm{d}x + c\int_K |\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\langle \mathbf{A}\rangle_K)|^2\,\mathrm{d}x \\
&\leqslant c \int_K (\varphi^*)_{|\langle \mathbf{A}\rangle_K|}(|\mathbf{A} - \boldsymbol{\mathcal{A}}(\mathbf{L}_h)|)\,\mathrm{d}x + c\int_K |\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\langle \mathbf{A}\rangle_K)|^2\,\mathrm{d}x \\
&\leqslant c \int_K (\varphi^*)_{|\mathbf{A}|}(|\mathbf{A} - \boldsymbol{\mathcal{A}}(\mathbf{L}_h)|)\,\mathrm{d}x + c\int_K |\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\langle \mathbf{A}\rangle_K)|^2\,\mathrm{d}x \\
&\leqslant c \int_K |\mathbf{F}(\mathbf{L}) - \mathbf{F}(\mathbf{L}_h)|^2\,\mathrm{d}x + cT_{1,2}.
\end{aligned}
$$

Moreover, it follows from Lemma 4.4 and Poincaré's inequality applied to $\mathbf{F}(\nabla\mathbf{u})$ on $L^2(K)$ that

$$
\begin{aligned}
T_{1,2} &= \int_K |\mathbf{F}^*(\boldsymbol{\mathcal{A}}(\mathbf{L})) - \mathbf{F}^*(\langle \boldsymbol{\mathcal{A}}(\mathbf{L})\rangle_K)|^2\,\mathrm{d}x \leqslant c\int_K |\mathbf{F}(\mathbf{L}) - \langle \mathbf{F}(\mathbf{L})\rangle_K|^2\,\mathrm{d}x \\
&\leqslant c h_K^2 \int_K |\nabla \mathbf{F}(\mathbf{L})|^2\,\mathrm{d}x.
\end{aligned}
$$

The above estimates prove the claim. □

Proposition 4.9, Theorems 4.3, 4.5 and 4.8 immediately yield the following corollary.

Corollary 4.10 Let $k \geqslant 1$. Let $(\mathbf{u}, \mathbf{L}, \mathbf{A}) \in W^{1,\varphi}(\Omega) \times L^{\varphi}(\Omega) \times W^{1,\varphi^*}(\Omega)$ be a solution of (1.2) with $\mathbf{F}(\nabla\mathbf{u}) \in W^{1,2}(\Omega)$ and let $(\mathbf{u}_h, \mathbf{L}_h, \mathbf{A}_h) \in V_h^k \times X_h^k \times X_h^k$ be a solution of (2.36) for some $\alpha > 0$.

(i) If $\mathbf{u}_D^* = \Pi_{\mathrm{SZ}}\mathbf{u}$, then

$$
\begin{aligned}
&\|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2 + \|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \Pi_{\mathrm{SZ}}\mathbf{u}) \\
&\quad\leqslant c_\alpha \begin{cases} h^2 \|\nabla \mathbf{F}(\nabla\mathbf{u})\|_2^2 & \text{if } p \leqslant 2, \\ h^{p'}\big(\|\nabla \mathbf{F}(\nabla\mathbf{u})\|_2^2 + \rho_{\varphi^*,\Omega}(\nabla\mathbf{A})\big) & \text{if } p \geqslant 2. \end{cases}
\end{aligned}
$$

(ii) If $\mathbf{u}_D^* = \mathbf{u}$, then

$$
\begin{aligned}
&\|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2^2 + \|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2^2 + \alpha m_{\varphi,h}^I(\mathbf{u}_h) + \alpha m_{\varphi,h}^D(\mathbf{u}_h - \mathbf{u}) \\
&\quad\leqslant c_\alpha \begin{cases} h^{p}(\|\nabla \mathbf{F}(\nabla\mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla\mathbf{u})) & \text{if } p \leqslant 2, \\ h^{p'}(\|\nabla \mathbf{F}(\nabla\mathbf{u})\|_2^2 + \rho_{\varphi,\Omega}(\nabla\mathbf{u}) + \rho_{\varphi^*,\Omega}(\nabla\mathbf{A})) & \text{if } p \geqslant 2. \end{cases}
\end{aligned}
$$

The constants depend only on the characteristics of $\boldsymbol{\mathcal{A}}$, the chunkiness $\omega_0$ and $\alpha^{-1}$.

## 5. Numerical experiments

In this section, we apply the LDG method, described above, to solve numerically systems with $p$-structure. We are going to approximate the solution of (1.1) by the stationary solution of the following

TABLE 1 *Choice of the stabilization parameter* $\alpha$

| | $p$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| – | 1.25 | 4/3 | 1.5 | 5/3 | 1.8 | 2 | 2.25 | 2.5 | 3 | 4 |
| $\alpha$ | 0.06 | 0.1 | 0.2 | 0.5 | 1.0 | 2.0 | 2.0 | 2.5 | 2.5 | 2.5 |

initial boundary value problem in $\Omega := [-2,2] \times [-2,2]$:

$$\begin{aligned} \mathbf{u}_t - \operatorname{div}\boldsymbol{\mathcal{A}}(\nabla\mathbf{u}) &= \mathbf{f} \quad \text{in } \Omega \times (0,T], \\ \mathbf{u}(\cdot,0) &= \mathbf{u}_0 \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D \quad \text{on } \partial\Omega \times (0,T], \end{aligned} \tag{5.1}$$

i.e., by $\lim_{t\to\infty} \mathbf{u}(\cdot,t)$. The parameter $\delta$ in the operator $\boldsymbol{\mathcal{A}}$ is chosen as $\delta := 10^{-3}$. Again we write the system (5.1) as a first-order system (cf. (1.2)), use the weak formulation similar as in (2.36). For the numerical flux function as defined in (2.32) and (2.33) we choose the parameter $\alpha$ according to Table 1 as a function of $p$.

We denote by $\mathbf{S}(t) = [\mathbf{L}(t), \mathbf{A}(t)]^{\mathrm{T}}$ the vector with the degrees of freedom of $\mathbf{L}_h$ and $\mathbf{A}_h$, and by $\mathbf{U}(t)$ the vector of degrees of freedom of $\mathbf{u}_h$. The LDG discretization leads to a nonlinear system of ODEs of the following form:

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{M}\dfrac{\mathrm{d}\mathbf{U}(t)}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1(t,\mathbf{S}(t),\mathbf{U}(t)) \\ \mathbf{R}_2(t,\mathbf{S}(t),\mathbf{U}(t)) \end{pmatrix}, \tag{5.2}$$

where $\mathbf{M}$ is the mass matrix of $\mathbf{u}_h$, and the components of $\mathbf{R}_1$ and $\mathbf{R}_2$ are derived by the weak formulation of the system (5.1) similar to (2.36), where all terms except the one coming from the time derivative are brought to the right-hand side. We discretize (5.2) by an s-stage diagonally implicit Runge–Kutta (sDIRK) method (cf. Alexander, 1977; Diehl, 2007).

On the basis of this method, we have considered several test cases for different values of $p$ to confirm the theoretical results in Corollary 4.10. We compute numerical solutions of the initial boundary value problem (5.1) in case where the function

$$\mathbf{u}(\mathbf{x},t) = \begin{cases} u(\mathbf{x},t) = B(t)|\mathbf{x}|^{a(p)}x_2, \\ v(\mathbf{x},t) = B(t)|\mathbf{x}|^{a(p)}(-x_1), \end{cases} \tag{5.3}$$

with $B(t) := 1 + \exp(-100t)$ and $a(p) = 0.01$, is the exact solution of (5.1). The corresponding initial data, boundary data and the source term are determined by (5.3). The parameter $a(p) = 0.01$ is chosen such that $\mathbf{F}(\nabla\mathbf{u}(\cdot,t)) \in W^{1,2}(\Omega)$. We construct a starting triangulation $\mathcal{T}_{h_0}$ by subdividing a rectangular cartesian grid into regular triangles with different orientations. Finer grids are obtained by regular subdivision of the previous grid: Each triangle is subdivided into four equal triangles by connecting the midpoints of the edges. We use the ansatz functions as defined in (2.18) with $k = 1$. The convergence rates for

$$\|\mathbf{F}(\mathbf{L}_h) - \mathbf{F}(\mathbf{L})\|_2, \quad \|\mathbf{F}^*(\mathbf{A}) - \mathbf{F}^*(\mathbf{A}_h)\|_2 \quad \text{and} \quad (\alpha m^I_{\varphi,h}(\mathbf{u}_h) + \alpha m^D_{\varphi,h}(\mathbf{u}_h - \mathbf{u}^*_D))^{\frac{1}{2}}$$

TABLE 2 *Convergence rates for* $\|\mathbf{F}(\mathbf{L})-\mathbf{F}(\mathbf{L}_h)\|_2$ *in the case of* $\mathcal{P}_1(K)$ *as local space and* $\mathbf{u}_D^*=\mathbf{u}$ *on* $\partial\Omega$, $\mathbf{F}(\nabla\mathbf{u})\in W^{1,2}(\Omega)$

| $\frac{h_0}{2^m}$ \ $p$ | 1.25 | 4/3 | 1.5 | 5/3 | 1.8 | 2 | 2.25 | 2.5 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|
| $m=1$ | 1.01 | 1.3 | 1.01 | 0.98 | 0.95 | 0.90 | 0.83 | 0.8 | 0.77 | 0.72 |
| $m=2$ | 1.01 | 1.03 | 0.98 | 0.98 | 0.96 | 0.95 | 0.94 | 0.96 | 0.97 | 1.00 |
| $m=3$ | 0.90 | 0.86 | 0.96 | 0.96 | 0.95 | 0.94 | 0.94 | 0.94 | 0.95 | 0.96 |
| $m=4$ | 0.98 | 0.95 | 0.94 | 1.00 | 0.96 | 0.90 | 0.95 | 0.95 | 0.98 | 0.97 |
| $m=5$ | 0.98 | 0.94 | 0.95 | 1.00 | 0.96 | 0.99 | 0.95 | 0.94 | 0.99 | 1.0 |

TABLE 3 *Convergence rates for* $\|\mathbf{F}(\mathbf{L})-\mathbf{F}(\mathbf{L}_h)\|_2$ *in the case of* $\mathcal{P}_1(K)$ *as local space and* $\mathbf{u}_D^*=\Pi_{SZ}\mathbf{u}$ *on* $\partial\Omega$, $\mathbf{F}(\nabla\mathbf{u})\in W^{1,2}(\Omega)$

| $\frac{h_0}{2^m}$ \ $p$ | 1.25 | 4/3 | 1.5 | 5/3 | 1.8 | 2 | 2.25 | 2.5 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|
| $m=1$ | 1.01 | 1.01 | 0.95 | 0.92 | 0.9 | 0.95 | 0.80 | 0.8 | 0.78 | 0.70 |
| $m=2$ | 0.99 | 1.00 | 0.97 | 0.96 | 0.96 | 0.95 | 0.93 | 0.95 | 0.98 | 0.98 |
| $m=3$ | 0.81 | 0.96 | 0.95 | 0.94 | 0.92 | 0.92 | 0.93 | 0.94 | 0.95 | 0.95 |
| $m=4$ | 0.95 | 0.93 | 0.96 | 0.98 | 0.96 | 0.97 | 0.93 | 0.92 | 0.96 | 0.99 |
| $m=5$ | 0.96 | 0.93 | 0.96 | 0.97 | 1.02 | 0.97 | 0.95 | 0.95 | 0.97 | 1.0 |

TABLE 4 *Convergence rates for* $\|\mathbf{F}^*(\mathbf{A})-\mathbf{F}^*(\mathbf{A}_h)\|_2$ *in the case of* $\mathcal{P}_1(K)$ *as local space and* $\mathbf{u}_D^*=\mathbf{u}$ *on* $\partial\Omega$, $\mathbf{F}(\nabla\mathbf{u})\in W^{1,2}(\Omega)$

| $\frac{h_0}{2^m}$ \ $p$ | 1.25 | 4/3 | 1.5 | 5/3 | 1.8 | 2 | 2.25 | 2.5 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|
| $m=1$ | 1.22 | 1.4 | 1.06 | 1.02 | 0.97 | 0.95 | 0.82 | 0.8 | 0.75 | 0.71 |
| $m=2$ | 1.05 | 1.04 | 1.01 | 0.98 | 0.97 | 0.95 | 0.93 | 0.95 | 0.99 | 0.98 |
| $m=3$ | 0.95 | 0.94 | 0.98 | 0.97 | 0.96 | 0.92 | 0.93 | 0.93 | 0.95 | 0.95 |
| $m=4$ | 0.99 | 0.95 | 0.91 | 1.00 | 0.96 | 0.97 | 0.94 | 0.95 | 0.97 | 0.97 |
| $m=5$ | 1.00 | 0.93 | 0.95 | 0.99 | 0.95 | 0.97 | 0.95 | 0.95 | 0.95 | 1.0 |

for different values of $p$ and for a series of triangulations, obtained by regular, global refinement as described above with $h_0=1$, Tables 2–5, respectively. The convergence rates are based on the evaluation of the (time-dependent) solution at time $T=0.25$. The numerical experiments clearly support the theoretical results from Corollary 4.10. In fact, the experiments indicate that the sub-optimal theoretical results, which are mainly due to the sub-optimal treatment of the terms on $\Gamma_D$, are of a purely technical nature and are not reflected in the numerical experiments. Moreover, Tables 2 and 3 show that the numerical results are almost the same for both choices of the boundary values, namely for $\mathbf{u}_D^*=\mathbf{u}$ on $\partial\Omega$ and for $\mathbf{u}_D^*=\Pi_{SZ}\mathbf{u}$ on $\partial\Omega$.

TABLE 5 *Convergence rates for* $(\alpha m^I_{\varphi,h}(\mathbf{u}_h) + \alpha m^D_{\varphi,h}(\mathbf{u}_h - \mathbf{u}^*_D))^{1/2}$ *in the case of* $\mathcal{P}_1(K)$ *as local space and* $\mathbf{u}^*_D = \mathbf{u}$ *on* $\partial\Omega$, $\mathbf{F}(\nabla\mathbf{u}) \in W^{1,2}(\Omega)$

| | $p$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{h_0}{2^m}$ | 1.25 | 4/3 | 1.5 | 5/3 | 1.8 | 2 | 2.25 | 2.5 | 3 | 4 |
| $m = 1$ | 0.93 | 0.91 | 0.86 | 0.83 | 0.80 | 0.84 | 0.75 | 0.80 | 0.75 | 0.72 |
| $m = 2$ | 0.92 | 0.94 | 0.88 | 0.87 | 0.86 | 0.85 | 0.85 | 0.90 | 0.98 | 1.00 |
| $m = 3$ | 0.86 | 0.81 | 0.90 | 0.89 | 0.90 | 0.87 | 0.88 | 0.90 | 0.97 | 1.02 |
| $m = 4$ | 0.92 | 0.91 | 0.92 | 0.90 | 0.90 | 0.92 | 0.91 | 0.93 | 1.00 | 1.07 |
| $m = 5$ | 0.92 | 0.92 | 0.92 | 0.90 | 0.90 | 0.93 | 0.93 | 0.92 | 0.99 | 1.02 |


## Acknowledgements

The authors wish to thank Alexandre Ern for pointing out a problem in an earlier version of this paper.

## Funding

This research was supported by the Project C.2 'Fluid structure interactions' of DFG SFB/TR 71 'Geometric Partial Differential Equations' (to L.D., D.K., M.R. and I.T.)


## REFERENCES


ALEXANDER, R. (1977) Diagonally implicit Runge–Kutta methods for stiff o.d.e.'s. *SIAM J. Numer. Anal.*, **14**, 1006–1021.

ARNOLD, D., BREZZI, F., COCKBURN, B. & MARINI, L. (2001/2002) Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, **39**, 1749–1779.

BARRETT, J. W. & LIU, W. B. (1994) Quasi-norm error bounds for the finite element approximation of a non-Newtonian flow. *Numer. Math.*, **68**, 437–456.

BELENKI, L., BERSELLI, L. C., DIENING, L. & RŮŽIČKA, M. (2012a) On the finite element approximation of $p$-Stokes systems. *SIAM J. Numer. Anal.*, **50**, 373–397. doi:10.1137/10080436X.

BELENKI, L., DIENING, L. & KREUZER, C. (2012b) Optimality of an adaptive finite element method for the $p$-Laplacian equation. *IMA J. Numer. Anal.*, **32**, 484–510.

BERSELLI, L. C., DIENING, L. & RŮŽIČKA, M. (2010) Existence of strong solutions for incompressible fluids with shear dependent viscosities. *J. Math. Fluid Mech.*, **12**, 101–132. Online First, doi:10.1007/s00021-008-0277-y.

BRENNER, S. C. (2003) Poincaré–Friedrichs inequalities for piecewise $H^1$ functions. *SIAM J. Numer. Anal.*, **41**, 306–324.

BUFFA, A. & ORTNER, C. (2009) Compact embeddings of broken Sobolev spaces and applications. *IMA J. Numer. Anal.*, **29**, 827–855.

BURMAN, E. & ERN, A. (2008) Discontinuous Galerkin approximation with discrete variational principle for the nonlinear Laplacian. *C. R. Math. Acad. Sci. Paris*, **346**, 1013–1016.

BUSTINZA, R. & GATICA, G. (2004) A local discontinuous Galerkin method for nonlinear diffusion problems with mixed boundary conditions. *SIAM J. Sci. Comput.*, **26**, 152–177.

DI PIETRO, D. & ERN, A. (2010) Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier–Stokes equations. *Math. Comp.*, **79**, 1303–1330.

DI PIETRO, D. & ERN, A. (2012) *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques & Applications, vol. 69. Berlin: Springer.

Diehl, D. (2007) High-order schemes for simulation of compressible liquid–vapor flows with phase change. *Dissertation*, Universität Freiburg.

Diening, L. & Ettwein, F. (2008) Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Mathematicum*, **20**, 523–556.

Diening, L., Harjulehto, P., Hästö, P. & Růžička, M. (2011) *Lebesgue and Sobolev Spaces with Variable Exponents*. Lecture Notes in Mathematics, vol. 2017. Heidelberg: Springer, pp. x+509.

Diening, L., Kaplický, P. & Schwarzacher, S. (2012) BMO estimates for the $p$-Laplacian. *Nonlinear Anal.*, **75**, 637–650.

Diening, L. & Kreuzer, C. (2008) Linear convergence of an adaptive finite element method for the $p$-Laplacian equation. *SIAM J. Numer. Anal.*, **46**, 614–638.

Diening, L. & Růžička, M. (2007) Interpolation operators in Orlicz–Sobolev spaces. *Numer. Math.*, **107**, 107–129. doi:10.1007/s00211-007-0079-9.

Diening, L., Růžička, M. & Schumacher, K. (2010) A decomposition technique for John domains. *Ann. Acad. Sci. Fenn. Ser. A. I. Math.*, **35**, 87–114.

Diening, L., Stroffolini, B. & Verde, A. (2009) Everywhere regularity of functionals with $\varphi$-growth. *Manuscripta Math.*, **129**, 449–481.

Ebmeyer, C. & Liu, W. B. (2005) Quasi-norm interpolation error estimates for finite element approximations of problems with $p$–structure. *Numer. Math.*, **100**, 233–258.

Houston, P., Robson, J. & Süli, E. (2005) Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems. I. The scalar case. *IMA J. Numer. Anal.*, **25**, 726–749.

Kokilashvili, V. & Krbec, M. (1991) *Weighted Inequalities in Lorentz and Orlicz Spaces*. River Edge, NJ: World Scientific, pp. xii+233.

Kröner, D., Růžička, M. & Toulopoulos, I. (2013) Local discontinuous Galerkin numerical solutions of non-Newtonian incompressible flows modeled by $p$-Navier–Stokes equations. *J. Comp. Phys.*, submitted.

Kufner, A., John, O. & Fučík, S. (1977) *Function Spaces*. Praha: Academia.

Lasis, A. & Süli, E. (2003) *Poincare-type Inequalities for Broken Sobolev Spaces*. Oxford: Oxford University Computing Laboratory, pp. 1–23.

Perugia, I. & Schötzau, D. (2001) On the coupling of local discontinuous Galerkin and conforming finite element methods. *J. Sci. Comput.*, **16**, 411–433 (2002).

Rao, M. M. & Ren, Z. D. (1991) *Theory of Orlicz Spaces*. Monographs and Textbooks in Pure and Applied Mathematics, vol. 146. New York: Marcel Dekker Inc.

Růžička, M. (2013) Theoretical and numerical analyis of steady problems for generalized Newtonian fluids. *Topics in Mathematical Fluid-Mechanics*. Lecture Notes in Mathematics, vol. 2073., pp. 199–238. Proceedings of CIME Summer School, Cetraro 2010. Heidelberg: Springer.

Růžička, M. & Diening, L. (2007) Non-Newtonian fluids and function spaces. *Nonlinear Analysis, Function Spaces and Applications, Proceedings of NAFSA 2006 Prague*, vol. 8., pp. 95–144. Prague.

Scott, L. & Zhang, S. (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, **54**, 483–493.

## A. Appendix

In this appendix, we derive several useful tools for the numerical analysis of DG systems. Subdivided into subsections, we study the local $L^2$-projection, the discrete gradient, an extended Scott–Zhang interpolation operator and embedding theorems of Poincaré and trace type.

The estimates derived below are presented mostly in the context of modular estimates rather than norm estimates. Note that modular estimates are stronger than norm estimates, since the first always implies the latter (but not the other way round). For example, a modular estimate like

$$\rho_{\psi,\Omega}(Tf) \leqslant K_1 \rho_{\psi,\Omega}(f) \tag{A.1}$$

or

$$\rho_{\psi,\Omega}(Tf) \leqslant \rho_{\psi,\Omega}(K_1 f) \tag{A.2}$$

for some operator $T$ and all $f \in L^{\psi}(\Omega)$ implies immediately the corresponding norm estimate

$$\|Tf\|_{L^{\psi}(\Omega)} \leqslant K_2 \|f\|_{L^{\psi}(\Omega)} \tag{A.3}$$

for all $f \in L^{\psi}(\Omega)$. Note that (A.2) implies (A.1) if $K_1 \geqslant 1$.

Although it is often useful to work with N-functions, the two limit cases $L^1$ and $L^{\infty}$ are unfortunately excluded in this setting. Therefore, we will work in the following in the context of $\Phi$-functions: A convex, left-continuous function $\varphi\colon [0,\infty) \to [0,\infty]$ with $\varphi(0)=0$ $\lim_{t\to 0^+}\varphi(t)=0$ and $\lim_{t\to\infty}\varphi(t)=\infty$ is called a *$\Phi$-function*. Note that every N-function is a $\Phi$-function, so the concept is slightly more general. The definition of the Orlicz spaces $L^{\psi}$, the Sobolev–Orlicz spaces $W^{1,\psi}$, the modular $\rho_{\psi}$, the $\Delta_2$ and the $\nabla_2$-condition from Section 2.1 are also used for $\Phi$-functions.

The $\Phi$-functions $\varphi_p(t) := (1/p)t^p$ for $1 \leqslant p < \infty$ and $\varphi_{\infty}(t) := \infty\chi_{(1,\infty)}(t)$ generate the spaces $L^p(\Omega)$, $1 \leqslant p \leqslant \infty$. In particular, the results below include the situation of $L^p$-spaces (both modular and norm estimates).

In the previous sections we always assumed the validity of the $\Delta_2$-condition. Most of the estimates presented below do not need this assumption. However, in the absence of the $\Delta_2$-condition (for example in the $L^{\infty}$-context) one has to work with modular estimates of the form (A.2) rather than (A.1). For the sake of presentation, we will assume in the following that

$$\varphi \text{ is a } \Phi\text{-function satisfying the } \Delta_2\text{-condition.} \tag{A.4}$$

This ensures that we can put all constants as usual in front of the modular as in (A.1). However, note that only minimal changes have to be made to include the non-$\Delta_2$-case into our treatment.

Moreover, note that the constants in the following subsections depend also on chunkiness $\omega_0$ of $\mathcal{T}_h$.

## A.1 *Local $L^2$-projection*

In this section we examine the properties of the projection $\Pi_{\mathrm{DG}}\colon L^1(\Omega) \to V_h^k(\Omega)$ from (2.19), which was defined by

$$(\Pi_{\mathrm{DG}}\mathbf{g}, \mathbf{z}_h) = (\mathbf{g}, \mathbf{z}_h) \quad \forall \mathbf{z}_h \in V_h^k.$$

The projection is completely local and we could define $(\Pi_{\mathrm{DG}}\mathbf{g})|_K$ for $K \in \mathcal{T}_h$ by

$$(\Pi_{\mathrm{DG}}\mathbf{g}, \mathbf{z}_h)_K = (\mathbf{g}, \mathbf{z}_h)_K \quad \forall \mathbf{z}_h \in \mathcal{P}_k(K).$$

Since $\Pi_{\mathrm{DG}}$ is a local $L^2$-projection, we have obviously for $\mathbf{g} \in L^2(K)$

$$\fint_K |\Pi_{\mathrm{DG}}\mathbf{g}|^2 \,\mathrm{d}x \leqslant \fint_K |\mathbf{g}|^2 \,\mathrm{d}x. \tag{A.5}$$

Since $\mathcal{P}_k(K)$ is finite-dimensional, it follows by a standard scaling argument that for $K \in \mathcal{T}_h$ and $\mathbf{g} \in L^1(\Omega)$ we have

$$\|\Pi_{\mathrm{DG}}\mathbf{g}\|_{L^{\infty}(K)} \leqslant c \fint_K |\Pi_{\mathrm{DG}}\mathbf{g}| \,\mathrm{d}x = c \sup_{\substack{\mathbf{z}_h \in \mathcal{P}_k(K)\\ \|\mathbf{z}_h\|_{\infty} \leqslant 1}} \frac{1}{|K|} |(\Pi_{\mathrm{DG}}\mathbf{g}, \mathbf{z}_h)_K| \leqslant c \fint_K |\mathbf{g}| \,\mathrm{d}x.$$

Clearly, this inequality is stronger than (A.5). It follows by Jensen's inequality that for every $\varPhi$-function $\psi$ we have

$$\fint_K \psi(|\varPi_{\mathrm{DG}}\mathbf{g}|)\,\mathrm{d}x \leqslant c\psi(\|\varPi_{\mathrm{DG}}\mathbf{g}\|_{L^\infty(K)}) \leqslant c\psi\left(\fint_K |\mathbf{g}|\,\mathrm{d}x\right) \leqslant c\fint_K \psi(|\mathbf{g}|)\,\mathrm{d}x. \tag{A.6}$$

In the following let $0 \leqslant j \leqslant l \leqslant k+1$. Since $\varPi_{\mathrm{DG}}\mathbf{p} = \mathbf{p}$ for all $\mathbf{p} \in \mathcal{P}_k(K)$, we have

$$\begin{aligned}&\fint_K \psi(h_K^j|\nabla_h^j(\mathbf{g}-\varPi_{\mathrm{DG}}\mathbf{g})|)\,\mathrm{d}x\\&\quad\leqslant c \inf_{\mathbf{p}\in\mathcal{P}_k(T)}\left(\fint_K \psi(h_K^j|\nabla_h^j(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x + \fint_K \psi(h_K^j|\nabla_h^j\varPi_{\mathrm{DG}}(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x\right).\end{aligned}$$

To estimate the last term we use $\|\mathbf{p}\|_{L^\infty(K)} \sim \fint_K |\mathbf{p}|\,\mathrm{d}x$, an inverse estimate for polynomials, Jensen's inequality and the $L^\psi$-stability in (A.6) to obtain

$$\begin{aligned}\fint_K \psi(h_K^j|\nabla_h^j\varPi_{\mathrm{DG}}(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x &\leqslant \fint_K \psi(ch_K^j|\fint_K \nabla_h^j\varPi_{\mathrm{DG}}(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x\\&\leqslant \fint_K \psi\left(c\fint_K |\varPi_{\mathrm{DG}}(\mathbf{g}-\mathbf{p})|\right)\mathrm{d}x\\&\leqslant c\fint_K \psi(|\varPi_{\mathrm{DG}}(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x\\&\leqslant c\fint_K \psi(|\mathbf{g}-\mathbf{p}|)\,\mathrm{d}x.\end{aligned}$$

Thus we have

$$\fint_K \psi(h_K^j|\nabla_h^j(\mathbf{g}-\varPi_{\mathrm{DG}}\mathbf{g})|)\,\mathrm{d}x \leqslant c\inf_{\mathbf{p}\in\mathcal{P}_k(K)}\left(\fint_K \psi(h_K^j|\nabla_h^j(\mathbf{g}-\mathbf{p})|)\,\mathrm{d}x + \fint_K \psi(|\mathbf{g}-\mathbf{p}|)\,\mathrm{d}x\right).$$

We choose $\mathbf{p}$ as the averaged Taylor polynomial of $\mathbf{g}$ and get by the classical Sobolev–Poincaré estimates for $\varPhi$-functions (cf. Diening & Růžička, 2007, Corollary 3.3) that

$$\fint_K \psi(h_K^j|\nabla_h^j(\mathbf{g}-\varPi_{\mathrm{DG}}\mathbf{g})|)\,\mathrm{d}x \leqslant c\fint_K \psi(h_K^l|\nabla^l\mathbf{g}|)\,\mathrm{d}x \tag{A.7}$$

for all $K \in \mathcal{T}_h$ and $\mathbf{g} \in W^{l,\psi}(K)$ with $0 \leqslant j \leqslant l \leqslant k+1$.

Summation of (A.7) over $K \in \mathcal{T}_h$ with $l = 0, 1$ gives

$$\rho_{\psi,\varOmega}(\mathbf{g}-\varPi_{\mathrm{DG}}\mathbf{g}) \leqslant c\varphi_{\psi,\varOmega}(\mathbf{g}), \tag{A.8}$$

$$\rho_{\psi,\varOmega}(\mathbf{g}-\varPi_{\mathrm{DG}}\mathbf{g}) \leqslant c\varphi_{\psi,\varOmega}(h\nabla_h\mathbf{g}), \tag{A.9}$$

$$\rho_{\psi,\varOmega}(\nabla_h\mathbf{g}-\nabla_h\varPi_{\mathrm{DG}}\mathbf{g}) \leqslant c\varphi_{\psi,\varOmega}(\nabla_h\mathbf{g}) \tag{A.10}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega)$. In particular, this implies

$$\rho_{\psi,\Omega}(\Pi_{\mathrm{DG}}\mathbf{g}) \leqslant c\varphi_{\psi,\Omega}(\mathbf{g}), \tag{A.11}$$

$$\rho_{\psi,\Omega}(\nabla_h \Pi_{\mathrm{DG}}\mathbf{g}) \leqslant c\varphi_{\psi,\Omega}(\nabla_h \mathbf{g}) \tag{A.12}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega)$.

In the following we also want to derive estimate of $M_{\psi,h}(\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g})$. We need to follow trace theorem and a proper scaling argument.

Lemma A.1 Let $K \in \mathcal{T}_h$ and $\gamma$ be a face of $K$. Then for all $\mathbf{g} \in W^{1,\psi}(K)$

$$\fint_\gamma \psi(|\mathbf{g}|)\,\mathrm{d}s \leqslant c \fint_K \psi(|\mathbf{g}|)\,\mathrm{d}x + c \fint_K \psi(|h_\gamma \nabla \mathbf{g}|)\,\mathrm{d}x \tag{A.13}$$

with a constant independent of $h$.

*Proof.* Let $\gamma$ be a face of $K \in \mathcal{T}_h$. From the embedding $W^{1,\psi}(K) \hookrightarrow L^\psi(\gamma)$ (cf. Kufner *et al.*, 1977) and a standard contradiction argument one can show the modular inequality

$$\int_\gamma \psi(|\mathbf{g}|)\,\mathrm{d}s \leqslant c \int_K \psi(|\mathbf{g}|)\,\mathrm{d}x + c \int_K \psi(|\nabla \mathbf{g}|)\,\mathrm{d}x.$$

From this and standard scaling properties we easily obtain the claim. □

Remark A.2 Note that for all $q \in \mathcal{P}_m(K)$, $m \in \mathbb{N}_0$, we have

$$\fint_\gamma \psi(|\mathbf{g}|)\,\mathrm{d}s \leqslant c \fint_K \psi(|\mathbf{g}|)\,\mathrm{d}x, \tag{A.14}$$

where $\gamma$ is a face of some $K \in \mathcal{T}_h$. This follows immediately from the equivalence of norms in the finite-dimensional space $\mathcal{P}_m(K)$ and standard scaling properties.

Let $\mathbf{g} \in W^{1,\psi}(K)$, $K \in \mathcal{T}_h$ and $\gamma$ be a face of $K$. Then Lemma A.1 and (A.7) imply

$$\begin{aligned} h_\gamma &\int_\gamma \psi(h_\gamma^{-1}|\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g}|)\,\mathrm{d}s \\ &\leqslant c \int_K \psi(h_\gamma^{-1}|\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g}|)\,\mathrm{d}x + c \int_K \psi(|\nabla(\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g})|)\,\mathrm{d}x \\ &\leqslant c \int_K \psi(|\nabla \mathbf{g}|)\,\mathrm{d}x. \end{aligned} \tag{A.15}$$

This immediately implies by triangle inequality for the traces and by summation that

$$m_{\psi,h}(\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g}) \leqslant c\,\rho_{\psi,\Omega}(\nabla_h \mathbf{g}) \tag{A.16}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega)$. Combined with (A.10) we get

$$M_{\psi,h}(\mathbf{g} - \Pi_{\mathrm{DG}}\mathbf{g}) \leqslant c\rho_{\psi,\Omega}(\nabla_h \mathbf{g}) \tag{A.17}$$

and consequently also

$$M_{\psi,h}(\Pi_{\mathrm{DG}}\mathbf{g}) \leqslant cM_{\psi,h}(\mathbf{g}) \tag{A.18}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG}}(\Omega)$. If $\mathbf{g} \in W^{1,\psi}_{\Gamma_D}(\Omega)$, then it follows that

$$M_{\psi,h}(\Pi_{\mathrm{DG}}\mathbf{g}) \leqslant c\rho_{\psi,\Omega}(\nabla\mathbf{g}), \tag{A.19}$$

since $M_{\psi,h}(\mathbf{g}) = \rho_{\psi,\Omega}(\nabla\mathbf{g})$ for $\mathbf{g} \in W^{1,\psi}_{\Gamma_D}(\Omega)$.

### A.2 *Discrete gradient estimates*

In this section we present estimates related to the discrete gradient, see (2.23) and the jump functionals $\mathbf{R}_h^\gamma$ and $\mathbf{R}_h$, see (2.24) and (2.25). We also explain why these objects are well defined.

We start with the jump functionals. Let $\mathbf{g} \in W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$, which means in particular that $\mathbf{g}$ is zero in $\Omega' \setminus \Omega$. Let $\gamma \in \Gamma_I \cup \Gamma_D$. Then by the classical $W^{1,1}$-trace theorem, the two-sided traces of $\mathbf{g}$ on $\gamma$ are well defined and the mapping $\mathbf{g} \mapsto [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma$ is continuous from $W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega) \to L^1(\gamma)$. As a consequence, the functional

$$X_h^k \to \mathbb{R}; \quad \mathbf{X}_h \mapsto \langle [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma, \{\mathbf{X}_h\}\rangle_\gamma$$

is continuous (where the finite-dimensional $X_h^k$ can be equipped with any norm). Therefore, by the Riesz representation theorem there exists an element $\mathbf{R}_h^\gamma\mathbf{g} \in X_h^k$ such that

$$(\mathbf{R}_h^\gamma\mathbf{g}, \mathbf{X}_h) := \langle [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma, \{\mathbf{X}_h\}\rangle_\gamma \quad \text{for all } \mathbf{X}_h \in X_h^k, \tag{A.20}$$

where $\gamma \in \Gamma_I \cup \Gamma_D$. This is consistent with our definition in (2.24). Both the functional and its Riesz representative $\mathbf{R}_h^\gamma$ are called *jump functionals*. For $\mathbf{g} \in W^{1,2}_{\mathrm{DG},\Gamma_D}(\Omega)$ they are well-known under the name *lifting operator* (cf. Perugia & Schötzau, 2001; Bustinza & Gatica, 2004; Di Pietro & Ern, 2010).

Recall that $S_\gamma$, $\gamma \in \Gamma$, is the union of at most two adjacent elements $K \in \mathcal{T}_h$ to $\gamma$. Since the functional $\mathbf{X}_h \mapsto \langle [\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma, \{\mathbf{X}_h\}\rangle_\gamma$ depends only on the values of $\mathbf{X}_h$ on $S_\gamma$, the same must hold for $(\mathbf{R}_h^\gamma\mathbf{g}, \mathbf{X}_h)$. In particular, $\mathbf{R}_h^\gamma$ is zero outside of $S_\gamma$ and $\mathbf{R}_h^\gamma\mathbf{g} = \chi_{S_\gamma}\mathbf{R}_h^\gamma\mathbf{g}$. Thus, it would have also been possible to define $\mathbf{R}_h^\gamma\mathbf{g} \in X_h^k(S_\gamma)$ by (A.20) with $\mathbf{X}_h \in X_h^k(S_\gamma)$ and then extend it outside of $S_\gamma$ by zero.

Since $X_h^k(S_\gamma)$ is finite-dimensional, it follows by a standard scaling[6] argument that

$$\|\mathbf{R}_h^\gamma\mathbf{g}\|_{L^\infty(S_\gamma)} \leqslant c \fint_{S_\gamma} |\mathbf{R}_h^\gamma\mathbf{g}|\,\mathrm{d}x = c \sup_{\substack{\mathbf{X}_h \in X_h^k(S_\gamma) \\ \|\mathbf{X}_h\|_\infty \leqslant 1}} \frac{1}{|S_\gamma|}\mathbf{s}(\mathbf{R}_h^\gamma\mathbf{g}, \mathbf{X}_h),$$

where $c$ is independent of $\gamma$. Thus the estimate

$$\mathbf{s}(\mathbf{R}_h^\gamma\mathbf{g}, \mathbf{X}_h) \leqslant \int_\gamma |[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma|\,\mathrm{d}s \big\|\mathbf{X}_h\big\|_{L^\infty(S_\gamma)}$$

implies

$$\|\mathbf{R}_h^\gamma\mathbf{g}\|_{L^\infty(S_\gamma)} \leqslant c \fint_\gamma h_\gamma^{-1}|[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma|\,\mathrm{d}s,$$

[6] The two adjacent elements of $S_\gamma$ should be scaled separately.

where we used that $|S_\gamma| \sim h_\gamma |\gamma|$. Overall, we get the point-wise estimate

$$|\mathbf{R}_h^\gamma \mathbf{g}| \leqslant \chi_{S_\gamma} c \fint_\gamma h_\gamma^{-1} |[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma| \,\mathrm{d}s. \tag{A.21}$$

Thus, for any $\Phi$-function $\psi$ and any $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$, we get by Jensen's inequality

$$\psi(|\mathbf{R}_h^\gamma \mathbf{g}|) \leqslant \chi_{S_\gamma} c \fint_\gamma \psi(h_\gamma^{-1} |[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma|) \,\mathrm{d}s. \tag{A.22}$$

In particular, using again $|S_\gamma| \sim h_\gamma |\gamma|$, we obtain

$$\int_{S_\gamma} \psi(|\mathbf{R}_h^\gamma \mathbf{g}|) \,\mathrm{d}x \leqslant c h_\gamma \int_\gamma \psi(h_\gamma^{-1} |[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma|) \,\mathrm{d}s. \tag{A.23}$$

These estimates are valid for all $\gamma \in \Gamma_I \cup \Gamma_D$. We just have to keep in mind that for $\gamma \in \Gamma_D$ we have $[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma = \mathrm{tr}_\gamma^\Omega \mathbf{g} \otimes \mathbf{n}$ for $\mathbf{g} \in W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$.

Consistent with (2.25) we define $\mathbf{R}_h := \sum_{\gamma \in \Gamma_I \cup \Gamma_D} \mathbf{R}_h^\gamma$ as the sum of our jump functionals, i.e.,

$$(\mathbf{R}_h \mathbf{g}, \mathbf{X}_h) = \langle [\![\mathbf{g} \otimes \mathbf{n}]\!], \{\mathbf{X}_h\} \rangle_{\Gamma_I \cup \Gamma_D} \quad \text{for all } \mathbf{X}_h \in X_h^k. \tag{A.24}$$

Since $\mathbf{R}_h^\gamma$ is supported on $S_\gamma$ and the sets $S_\gamma$ are locally finite, we have

$$\rho_{\psi,\Omega}(\mathbf{R}_h \mathbf{g}) \leqslant c \sum_{\gamma \in \Gamma_I \cup \Gamma_D} h_\gamma \int_\gamma \psi(h_\gamma^{-1} |[\![\mathbf{g} \otimes \mathbf{n}]\!]_\gamma|) \,\mathrm{d}s = c m_{\psi,h}(\mathbf{g}) \tag{A.25}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$.

Consistent with (2.28) we define for every $\mathbf{g} \in W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$ the *discrete gradient* by $\nabla^h_{\mathrm{DG}} \mathbf{g} = \Pi_{\mathrm{DG}} \nabla_h \mathbf{g} - \mathbf{R}_h \mathbf{g}$. It immediately follows by (A.25) and (A.12) that

$$\rho_{\psi,\Omega}(\nabla^h_{\mathrm{DG}} \mathbf{g}) \leqslant c(\rho_{\psi,\Omega}(\nabla_h \mathbf{g}) + m_{\psi,h}(\mathbf{g})) = c M_{\psi,h}(\mathbf{g}). \tag{A.26}$$

Using also (A.19), we obtain

$$\rho_{\psi,\Omega}(\nabla_h \mathbf{g}) \leqslant c(\rho_{\psi,\Omega}(\nabla^h_{\mathrm{DG}} \mathbf{g}) + m_{\psi,h}(\mathbf{g})), \tag{A.27}$$

$$M_{\psi,h}(\mathbf{g}) = \rho_{\psi,\Omega}(\nabla_h \mathbf{g}) + m_{\psi,h}(\mathbf{g}) \leqslant c(\rho_{\psi,\Omega}(\nabla^h_{\mathrm{DG}} \mathbf{g}) + m_{\psi,h}(\mathbf{g})) \tag{A.28}$$

for all $\mathbf{g} \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$.

### A.3 *Scott–Zhang interpolation*

In this section we extend the classical Scott–Zhang interpolation operator to the DG setting. We restrict ourselves to the scalar situation, since the vectorial case follows by component-wise application. Let

$$V^k_{h,\mathrm{cont}}(\Omega) := V^k_h(\Omega) \cap W^{1,1}(\Omega),$$
$$V^k_{h,\mathrm{cont}}(\Omega') := V^k_h(\Omega') \cap W^{1,1}(\Omega').$$

Our goal is to construct a linear projection $\Pi_{\mathrm{SZ}} : W^{1,1}_{\mathrm{DG}}(\Omega') \to V^k_{h,\mathrm{cont}}(\Omega')$. In particular, DG functions are mapped to classical Sobolev functions (which have no jumps). Additionally, we want that $\Pi_{\mathrm{SZ}}$ maps functions from $W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$ (extended by zero to $\Omega' \setminus \Omega$) to $V^k_{h,\mathrm{cont}}(\Omega) \cap W^{1,1}_{\Gamma_D}(\Omega)$. To achieve this, it is easier for us to start with a linear projection from $W^{1,1}_{\mathrm{DG}}(\Omega')$ and restrict this operator later to suitable subspaces.

Let $\mathcal{N}'$ denote the standard set of nodal points of $\mathcal{T}'_h$ and $\Gamma'_I$ denote the set of interior faces of $\mathcal{T}'_h$. Further, let $\{\phi_a\}_{a\in\mathcal{N}'}$ denote the standard Lagrange basis of $V^k_{h,\mathrm{cont}}(\Omega')$. To each $a \in \mathcal{N}'$ we associate either an $n$-dimensional simplex $K = K_a$ (if $a \in \mathrm{int}K$) or an $(n-1)$-dimensional face $\gamma = \gamma_a$ (if $a \in \overline{\gamma}$), which we denote in both cases by $\sigma_a$.

For each $\sigma_a$, we denote by $\{\phi_{a,i}\}_i$ the local basis of $\{\phi_b|_{\sigma_a} | b \in \mathcal{N}'\}$ with a numbering such that $\phi_{a,1} = \phi_a|_{\sigma_a}$. Let $\{\beta_{a,i}\}_i$ denote the local dual basis with respect to the standard scalar product $(\cdot,\cdot)_{\sigma_a}$ on $\sigma_a$, i.e., $(\phi_{a,i}, \beta_{a,j}) = \delta_{i,j}$. Then for each $g \in W^{1,1}(\Omega)$ the standard Scott–Zhang operator is defined by

$$\Pi_{\mathrm{SZ}} g := \sum_{a\in\mathcal{N}'} (g, \beta_{a,1})_{\sigma_a} \phi_a.$$

In the DG setting we have to be more precise. In particular, if $\sigma_a$ is a face, and $g \in W^{1,1}_{\mathrm{DG}}(\Omega)$, then $g$ usually has two traces at $\sigma_a$. Therefore, $\sigma_a$ will be assigned to precisely one of the (at most) two simplices of $\mathcal{T}'_h$, which have $\sigma_a$ as a face. We denote this simplex by $K_a$. The scalar product $(g, \beta_{a,1})_{\sigma_a}$ is understood in such a way that the trace $g$ is evaluated at the side of the assigned simplex $K_a$. For a clear notation we write $(g|_{K_a}, \beta_{a,1})_{\sigma_a}$. If $\sigma_a$ is already an $n$-dimensional simplex, then we set $K_a := \sigma_a$.

With these conventions, we can extend the standard Scott–Zhang interpolation operator to functions from $W^{1,1}_{\mathrm{DG}}(\Omega')$ by

$$\Pi_{\mathrm{SZ}} g := \sum_{a\in\mathcal{N}'} (g|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a. \tag{A.29}$$

Let us point out that $\Pi_{\mathrm{SZ}}$ is just a classical Scott–Zhang interpolation operator if restricted to $W^{1,1}(\Omega')$.

It is clear that $\Pi_{\mathrm{SZ}}$ is a linear mapping from $W^{1,1}_{\mathrm{DG}}(\Omega')$ to $V^k_{h,\mathrm{cont}}(\Omega')$. We want to show that it is also a projection. As in the case of the standard Scott–Zhang operator, we have

$$(\phi_b|_{K_a}, \beta_{a,1})_{\sigma_a} = \delta_{a,b} \tag{A.30}$$

for all $a, b \in \mathcal{N}'$. Indeed, if $a = b$, then $\phi_a = \phi_{a,1}$ and $(\phi_{a,1}|_{K_a}, \beta_{a,1})_{\sigma_a} = 1$. If, on the other hand, $a \neq b$, then $\phi_a|_{K_a} = \phi_{a,j}$ for some $j \neq 1$ and $(\phi_{a,j}, \beta_{a,1})_{\sigma_a} = 0$. Now (A.30) implies that for every $b \in \mathcal{N}'$ we have

$$\Pi_{\mathrm{SZ}} \phi_b = \sum_{a\in\mathcal{N}'} (\phi_b|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a = \sum_{a\in\mathcal{N}'} \delta_{a,b}\phi_a = \phi_b.$$

This and the linearity of $\Pi_{\mathrm{SZ}}$ implies that $\Pi_{\mathrm{SZ}}$ is a linear projection from $W^{1,1}_{\mathrm{DG}}(\Omega')$ to $V^k_{h,\mathrm{cont}}(\Omega')$.

As the classical Scott–Zhang interpolation operator, also our extension has local nature. For $K \in \mathcal{T}'_h$, let $S'_K$ denote the neighbourhood of $K \in \mathcal{T}'_h$, i.e., the patch $S'_K$ is the union of all simplices of $\mathcal{T}'_h$ touching $K$. Indeed, if $K \in \mathcal{T}'_h$, then $\phi_b|_K = 0$, whenever $K \not\subset S'_{K_b}$. This implies that $(\Pi_{\mathrm{SZ}} g)|_K$ only depends on the values of $g$ on $S'_K$, i.e.,

$$(\Pi_{\mathrm{SZ}} g)|_K = \Pi_{\mathrm{SZ}}(g\chi_{S'_K}) \tag{A.31}$$

for all $g \in W^{1,1}_{\mathrm{DG}}(\Omega')$ and all $K \in \mathcal{T}'_h$.

In order to ensure that the subspace $W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$ is mapped to $V_h^k(\Omega') \cap W^{1,1}_{\Gamma_D}(\Omega')$, where $W^{1,1}_{\Gamma_D}(\Omega') := \{g \in W^{1,1}(\Omega') | g|_{\Omega' \setminus \Omega} = 0\}$, we have to take special care of the choice of the $\sigma_a$ and $K_a$ for $a \in \bar{\Gamma}_D$. If $a \in \bar{\Gamma}_D$, then $\sigma_a$ should be an $(n-1)$-dimensional face contained in $\Gamma_D$ and $K_a$ should be the element of $\mathcal{T}_h'$, which lies outside of $\Omega$ and has $\sigma_a$ as a face. If $a \notin \bar{\Gamma}_D$, then $\sigma_a$ should be an element of $\mathcal{T}_h'$ and $K_a := \sigma_a$. Note that this automatically ensures that for $a \in \bar{\Omega} \setminus \bar{\Gamma}_D$ we have $K_a \in \mathcal{T}_h$.

Let us explain now why $\Pi_{\mathrm{SZ}}$ maps the subspace $W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$ to $V_h^k(\Omega') \cap W^{1,1}_{\Gamma_D}(\Omega')$. To this end let $g \in W^{1,1}_{\mathrm{DG},\Gamma_D}(\Omega)$ and recall that $g$ is equal to zero on $\Omega' \setminus \Omega$. Then for every $a \in \mathcal{N}'$ with $a \in (\overline{\Omega'} \setminus \bar{\Omega}) \cup \bar{\Gamma}_D$, we have $K_a \subset \Omega' \setminus \Omega$, which implies $(g|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a = 0$. This implies that $\Pi_{\mathrm{SZ}} g = 0$ on $\Omega' \setminus \Omega$ and consequently we get $\Pi_{\mathrm{SZ}} g \in W^{1,1}_{\Gamma_D}(\Omega')$.

It follows from (A.31) that $(\Pi_{\mathrm{SZ}} g)|_\Omega$ only depends on the values of $g$ on $\bigcup_{K \subset \mathcal{T}_h} S_K'$ for $g \in W^{1,1}_{\mathrm{DG}}(\Omega')$, which is unfortunately slightly larger than $\Omega$. However, for $g \in W^{1,1}(\Omega')$ we can sharpen this to the following restriction property. Indeed, we show below that

$$g|_\Omega = h|_\Omega \;\Rightarrow\; (\Pi_{\mathrm{SZ}} g)|_\Omega = (\Pi_{\mathrm{SZ}} g)|_\Omega, \tag{A.32}$$

$$g|_{\Gamma_D} = h|_{\Gamma_D} \;\Rightarrow\; (\Pi_{\mathrm{SZ}} g)|_{\Gamma_D} = (\Pi_{\mathrm{SZ}} h)|_{\Gamma_D}. \tag{A.33}$$

We start with (A.32). Let $g \in W^{1,1}(\Omega')$. If $a \in \mathcal{N}$ with $a \in \bar{\Omega} \setminus \bar{\Gamma}_D$, then from the above discussion follows that $K_a \subset \Omega$. Thus $(g|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a$ depends only on $g|_\Omega$. If $a \in \mathcal{N}$ with $a \in \bar{\Gamma}_D$, then $\sigma_a \subset \Gamma_D$ and $(g|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a$ depends only on outer trace of $g$ on $\Gamma_D$. Since $g$ is a standard Sobolev function, this trace is equal to the inner trace of $g$ on $\Gamma_D$ and is uniquely determined by the values of $g$ on $\Omega$. If $a \in \mathcal{N}'$ with $a \in \Omega' \setminus \bar{\Omega}$, then $\phi_a = 0$ on $\Omega$, so these nodes do not contribute to $(\Pi_{\mathrm{SZ}} g)|_\Omega$. Overall, we have proved (A.32). Let us continue with (A.33). If $a \in \mathcal{N}'$ with $a \notin \bar{\Gamma}_D$, then $\phi_a|_{\Gamma_D} = 0$. If $a \in \mathcal{N}'$ with $a \in \bar{\Gamma}_D$, then $\sigma_a \subset \Gamma_D$ and $(g|_{K_a}, \beta_{a,1})_{\sigma_a} \phi_a$ only depends on $g$ restricted to $\Gamma_D$. This proves (A.33).

Because of (A.32) and (A.33), it is possible to consider $\Pi_{\mathrm{SZ}}$ as an operator from $W^{1,1}(\Omega)$ to $V^k_{h,\mathrm{cont}}(\Omega)$ and from $\mathrm{tr}_{\Gamma_D} W^{1,1}(\Omega)$ to $\mathrm{tr} V^k_{h,\mathrm{cont}}(\Gamma_D)$, where $\mathrm{tr}_{\Gamma_D} W^{1,1}(\Omega)$ is the $\Gamma_D$-trace space of $W^{1,1}(\Omega)$ and $V^k_{h,\mathrm{cont}}(\Gamma_D)$ is the space of $C^0(\bar{\Gamma}_D)$-function, which are $\mathcal{P}_k$ polynomials on each face $\gamma \subset \Gamma_D$.

Let us now show various estimates for $\Pi_{\mathrm{SZ}}$. It follows as in Scott & Zhang (1990) that for all $a \in \mathcal{N}'$ and all $i$

$$\|\phi_{a,i}\|_\infty \leqslant 1, \quad \|\beta_{a,i}\|_\infty \leqslant \frac{c}{|\sigma_a|}, \tag{A.34}$$

where $|\sigma_a|$ is either the $n$ or the $(n-1)$-dimensional measure of $\sigma_a$.

Lemma A.3 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2$, $K \in \mathcal{T}_h$ and $\gamma$ be a face of $K$. Then

$$\psi(|\langle f\rangle_\gamma - \langle f\rangle_K|) + \fint_K \psi(|f - \langle f\rangle_\gamma|)\,\mathrm{d}x + c\fint_K \psi(|f - \langle f\rangle_K|)\,\mathrm{d}x \leqslant c\fint_K \psi(h_K|\nabla f|)\,\mathrm{d}x$$

for all $f \in W^{1,\psi}(K)$.

*Proof.* The estimate $\fint_K \psi(|f - \langle f\rangle_K|)\,\mathrm{d}x \leqslant c\fint_K \psi(h_K|\nabla f|)\,\mathrm{d}x$ is a classical result using additionally a scaling argument (cf. Diening & Ettwein, 2008, Theorem 7, Diening *et al.*, 2010, Theorem 6.5).

By Jensen's inequality, Lemma A.1 and the $W^{1,1}$-Poincaré inequality we estimate

$$
\begin{aligned}
|\langle f\rangle_\gamma - \langle f\rangle_K| &\leqslant \fint_\gamma |f-\langle f\rangle_K|\,\mathrm{d}s \leqslant c\fint_K |f-\langle f\rangle_K|\,\mathrm{d}x + c\fint_K h_K|\nabla f|\,\mathrm{d}x\\
&\leqslant c\fint_K h_K|\nabla f|\,\mathrm{d}x.
\end{aligned}
$$

Thus, with Jensen's inequality we have

$$
\psi(|\langle f\rangle_\gamma - \langle f\rangle_K|) \leqslant c\psi\left(c\fint_K h_K|\nabla f|\,\mathrm{d}x\right) \leqslant c\fint_K \psi(h_K|\nabla f|)\,\mathrm{d}x.
$$

The remaining estimate for $\fint_K \psi(|f-\langle f\rangle_\gamma|)\,\mathrm{d}x$ follows by triangle inequality. □

Lemma A.4 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2 \cap \nabla_2$, $K \in \mathcal{T}_h$ and $\gamma$ be a face of $K$. Then

$$
\fint_\gamma \psi(|f-\langle f\rangle_\gamma|)\,\mathrm{d}x + \fint_\gamma \psi(|f-\langle f\rangle_K|)\,\mathrm{d}x \leqslant c\fint_K \psi(h_K|\nabla f|)\,\mathrm{d}x
$$

for all $f \in W^{1,\psi}(K)$.

*Proof.* Because of Lemma A.3 and the triangle inequality it suffices to estimate one of the terms, namely $\fint_\gamma \psi(|f-\langle f\rangle_K|)\,\mathrm{d}x$. By Diening *et al.* (2011, Lemma 8.2.1 b) it follows that $|f-\langle f\rangle_K| \leqslant cI_1(\chi_K\nabla f)$. Then as in the proof of Lemma A.8, we obtain

$$
\fint_\gamma \psi(|f-\langle f\rangle_K|)\,\mathrm{d}x \leqslant c\fint_K \psi(h_K|\nabla f|)\,\mathrm{d}x.
$$

Note that as in Lemma A.8 we assume $\psi \in \nabla_2$ only for simplicity of the proof. □

Note that in the case of vector-valued functions $\mathbf{g}$, one has to replace here and in the following the jump $[\![g\mathbf{n}]\!]$ by $[\![\mathbf{g}\otimes\mathbf{n}]\!]$.

Lemma A.5 Let $\gamma \in \Gamma_I'$ with $S_\gamma' = K_\gamma^1 \cup K_\gamma^2$. Then

$$
|\langle g\rangle_{K_\gamma^1} - \langle g\rangle_{K_\gamma^2}| \leqslant c\fint_{S_\gamma'} h_\gamma|\nabla_h g|\,\mathrm{d}x + c\fint_\gamma |[\![g\mathbf{n}]\!]_\gamma|\,\mathrm{d}s,
$$

for all $g \in W^{1,1}_{\mathrm{DG}}(S_\gamma')$.

*Proof.* We estimate by the properties of the mean value integral, the fact that $|a| = |a\mathbf{n}|$, since $|\mathbf{n}| = 1$, and Poincaré's inequality (Lemma A.3 with $\psi(t) = t$)

$$
\begin{aligned}
|\langle g\rangle_{K_\gamma^1} - \langle g\rangle_{K_\gamma^2}| &\leqslant |\langle g\rangle_{K_\gamma^1} - \langle g|_{K_\gamma^1}\rangle_\gamma| + \langle|[\![g\mathbf{n}]\!]_\gamma|\rangle_\gamma + |\langle g\rangle_{K_\gamma^2} - \langle g|_{K_\gamma^2}\rangle_\gamma|\\
&\leqslant \fint_{K_\gamma^1} |g-\langle g|_{K_\gamma^1}\rangle_\gamma|\,\mathrm{d}x + \fint_\gamma |[\![g\mathbf{n}]\!]_\gamma|\,\mathrm{d}s + \fint_{K_\gamma^2} |g-\langle g|_{K_\gamma^2}\rangle_\gamma|\,\mathrm{d}x.\\
&\leqslant c\fint_{K_\gamma^1} |\nabla g|\,\mathrm{d}x + \fint_\gamma |[\![g\mathbf{n}]\!]_\gamma|\,\mathrm{d}s + c\fint_{K_\gamma^2} |\nabla g|\,\mathrm{d}x.
\end{aligned}
$$

This proves the claim. □

The following lemma provides the basic estimates for our interpolation operator $\Pi_{\mathrm{SZ}}$.

LEMMA A.6 Let $K \in \mathcal{T}_h'$ and $g \in W^{1,1}_{\mathrm{DG}}(\Omega')$. Then

$$\|\nabla \Pi_{\mathrm{SZ}} g\|_{L^\infty(K)} \leqslant c h_K^{-1} \|(\Pi_{\mathrm{SZ}} g) - \langle g\rangle_K\|_{L^\infty(K)} \leqslant c h_K^{-1} \fint_K |\Pi_{\mathrm{SZ}} g - \langle g\rangle_K| \,\mathrm{d}x,$$

$$\|(\Pi_{\mathrm{SZ}} g) - \langle g\rangle_K\|_{L^\infty(K)} \leqslant c \fint_{S_K'} h_K |\nabla_h g| \,\mathrm{d}x + c \sum_{\gamma \in \Gamma_I(S_K')} \fint_\gamma |[\![ g\mathbf{n} ]\!]_\gamma| \,\mathrm{d}s,$$

where $\Gamma_I(S_K')$ denotes the faces of $\mathcal{T}_h'$ in the interior of $S_K'$.

*Proof.* By inverse estimates for polynomials and $\nabla \Pi_{\mathrm{SZ}} g = \nabla(\Pi_{\mathrm{SZ}} g - \langle g\rangle_K)$, we obtain

$$\begin{aligned}
\|\nabla \Pi_{\mathrm{SZ}} g\|_{L^\infty(K)} &= \|\nabla(\Pi_{\mathrm{SZ}} g - \langle g\rangle_K)\|_{L^\infty(K)} \\
&\leqslant c h_K^{-1} \|\Pi_{\mathrm{SZ}} g - \langle g\rangle_K\|_{L^\infty(K)} \\
&\leqslant c h_K^{-1} \fint_K |\Pi_{\mathrm{SZ}} g - \langle g\rangle_K| \,\mathrm{d}x.
\end{aligned}$$

This proves the first estimate of the claim.

Let $K \in \mathcal{T}_h'$. Then by (A.31) and (A.34) we estimate

$$\|(\Pi_{\mathrm{SZ}} g) - \langle g\rangle_K\|_{L^\infty(K)} = \|\Pi_{\mathrm{SZ}}(g - \langle g\rangle_K)\|_{L^\infty(K)} \leqslant c \sum_{\substack{a \in \mathcal{N}: \\ K_a \subset S_K'}} \langle |g|_{K_a} - \langle g\rangle_K| \rangle_{\sigma_a}.$$

We can estimate the summands by

$$\langle |g|_{K_a} - \langle g\rangle_K| \rangle_{\sigma_a} \leqslant c \fint_{K_a} |g - \langle g\rangle_K| \,\mathrm{d}x + c \fint_{K_a} h_{K_a} |\nabla_h g| \,\mathrm{d}x$$

where we used Lemma A.1 if $\sigma_a$ is a face of $K_a$. Since the interior of $S_K'$ is connected, we find for every $K_a$ a finite sequence $K_{a,j}$ (whose length only depends on the degeneracy of $\mathcal{T}_h$) starting at $K_a$ and ending in $K$ such that $K_{a,j}$ and $K_{a,j+1}$ share a face. This and Lemma A.5 allows one to replace the mean value $\langle g\rangle_K$ on $K_a$ by $\langle g\rangle_{K_a}$. In particular, we have

$$\begin{aligned}
\fint_{K_a} |g|_{K_a} - \langle g\rangle_K| \,\mathrm{d}x &\leqslant c \fint_{K_a} |g - \langle g\rangle_{K_a}| \,\mathrm{d}x + c \sum_{\gamma \in \Gamma_I(S_K')} \left( \fint_\gamma |[\![ g\mathbf{n} ]\!]| \,\mathrm{d}s + \fint_{S_\gamma} h_\gamma |\nabla_h g| \,\mathrm{d}x \right) \\
&\leqslant c \fint_{S_K'} h_K |\nabla_h g| \,\mathrm{d}x + c \sum_{\gamma \in \Gamma_I(S_K')} \fint_\gamma |[\![ g\mathbf{n} ]\!]| \,\mathrm{d}s,
\end{aligned}$$

where we also used Poincaré's inequality on $K_a$. This and the penultimate estimate yield the second estimate of the claim, since the number of elements in $S_K'$ is bounded by a constant depending only on the chunkiness of $\mathcal{T}_h'$. □

It follows immediately by Lemma A.6 and Jensen's inequality that

$$\begin{aligned}\int_K \psi(|\nabla \Pi_{SZ} g|)\,dx &+ \int_K \psi(h_K^{-1}|(\Pi_{SZ} g) - \langle g\rangle_K|)\,dx \\ &\leqslant c\int_{S'_K} \psi(|\nabla_h g|)\,dx + c\sum_{\gamma\in\Gamma_I(S'_K)} h_\gamma \int_\gamma \psi(h_\gamma^{-1}|[\![g\mathbf{n}]\!]_\gamma|)\,ds,\end{aligned} \tag{A.35}$$

where we used that $|K| \sim h_\gamma|\gamma|$ for the second term on the right-hand side. Together with Lemma A.3 we get for $g \in W^{1,\psi}_{DG}(\Omega')$

$$\begin{aligned}\int_K \psi(h_K^{-1}|g - \Pi_{SZ} g|)\,dx &\leqslant c\int_K \psi(h_K^{-1}|(\Pi_{SZ} g) - \langle g\rangle_K|)\,dx + c\int_K \psi(h_K^{-1}|g - \langle g\rangle_K|)\,dx \\ &\leqslant c\int_{S'_K} \psi(|\nabla_h g|)\,dx + c\sum_{\gamma\in\Gamma_I(S'_K)} h_\gamma \int_\gamma \psi(h_\gamma^{-1}|[\![g\mathbf{n}]\!]_\gamma|)\,ds.\end{aligned} \tag{A.36}$$

Summation of (A.35) and of (A.36) over $K \in \mathcal{T}'_h$ implies

$$\rho_{\psi,\Omega'}(h_K^{-1}(g - \Pi_{SZ} g)) + \rho_{\psi,\Omega'}(\nabla \Pi_{SZ} g) \leqslant cM_{\psi,h,\Omega'}(g) \tag{A.37}$$

for all $g \in W^{1,\psi}_{DG}(\Omega')$, where

$$m_{\psi,h,\Omega'}(g) := \sum_{\gamma\in\Gamma_I(\Omega')} h_\gamma \int_\gamma \psi(h_\gamma^{-1}|[\![g\mathbf{n}]\!]_\gamma|)\,ds,$$

$$M_{\psi,h,\Omega'}(g) := \rho_{\psi,\Omega'}(|\nabla_h g|) + m_{\psi,h,\Omega'}(g).$$

Let us explain the situation for $g \in W^{1,\psi}_{DG,\Gamma_D}(\Omega)$ (extended by zero on $\Omega' \setminus \Omega$). First of all,

$$m_{\psi,h,\Omega'}(g) = m_{\psi,h}(g),$$

$$M_{\psi,h,\Omega'}(g) = M_{\psi,h}(g)$$

for all $g \in W^{1,\psi}_{DG,\Gamma_D}(\Omega)$. Moreover, for all $g \in W^{1,\psi}_{DG,\Gamma_D}(\Omega)$, we have $g = \Pi_{SZ} g = 0$ on $\Omega' \setminus \Omega$.

In particular, (A.37) implies that

$$\rho_{\psi,\Omega}(h_K^{-1}(g - \Pi_{SZ} g)) + \rho_{\psi,\Omega}(\nabla \Pi_{SZ} g) \leqslant cM_{\psi,h}(g) \tag{A.38}$$

for all $g \in W^{1,\psi}_{DG,\Gamma_D}(\Omega)$.

On the other hand, our Scott–Zhang operator $\Pi_{SZ}$ coincides for functions from $W^{l,\psi}(\Omega)$ with the standard Scott–Zhang operator defined on $\Omega$ (cf. (A.32)). Thus we have (cf. Diening & Růžička, 2007, Theorem 4.3)

$$⨍_K \psi(h_K^j|\nabla_h^j(g - \Pi_{SZ} g)|)\,dx \leqslant c ⨍_{S'_K} \psi(h_K^l|\nabla^l g|)\,dx \tag{A.39}$$

for all $K \in \mathcal{T}_h$ and $g \in W^{l,\psi}(K)$ with $0 \leqslant j \leqslant l \leqslant k+1$, $k \geqslant 1$. In particular, this and Lemma A.1 imply for all $K \in \mathcal{T}_h$ and all faces $\gamma$ of $K$

$$\begin{aligned} h_\gamma \int_\gamma \psi(h_\gamma^{-1}|g - \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}s &\leqslant c \int_K \psi(h_\gamma^{-1}|g - \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}x + c \int_K \psi(|\nabla_h(g - \Pi_{\mathrm{SZ}} g)|)\,\mathrm{d}x \\ &\leqslant c \int_{S'_K} \psi(|\nabla g|)\,\mathrm{d}x, \end{aligned} \tag{A.40}$$

$$\begin{aligned} h_\gamma \int_\gamma \psi(h_\gamma^{-1}|g - \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}s &\leqslant c \int_K \psi(h_\gamma^{-1}|g - \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}x + c \int_K \psi(|\nabla_h(g - \Pi_{\mathrm{SZ}} g)|)\,\mathrm{d}x \\ &\leqslant c \int_{S'_K} \psi(|h\nabla^2 g|)\,\mathrm{d}x. \end{aligned} \tag{A.41}$$

Summation of (A.40) and of (A.41) over $K \in \mathcal{T}_h$ immediately implies

$$m_{\psi,h}(g - \Pi_{\mathrm{SZ}} g) + \rho_{\psi,\Omega}(h_K^{-1}(g - \Pi_{\mathrm{SZ}} g)) + \rho_{\psi,\Omega}(\nabla_h \Pi_{\mathrm{SZ}} g) \leqslant c\rho_{\psi,\Omega}(\nabla g), \tag{A.42}$$

$$m_{\psi,h}(g - \Pi_{\mathrm{SZ}} g) + \rho_{\psi,\Omega}(h_K^{-1}(g - \Pi_{\mathrm{SZ}} g)) + \rho_{\psi,\Omega}(\nabla g - \nabla_h \Pi_{\mathrm{SZ}} g) \leqslant c\rho_{\psi,\Omega}(h\nabla^2 g), \tag{A.43}$$

for all $g \in W^{1,\psi}(\Omega)$ and $g \in W^{2,\psi}(\Omega)$, respectively.

### A.4 *Poincaré and trace theorems*

The extended Scott–Zhang operator can be used to extend embedding and trace theorems for classical Sobolev spaces to the DG setting. A similar approach has been used in Brenner (2003), Lasis & Süli (2003) and Buffa & Ortner (2009) to prove similar, but slightly different versions of Poincaré and Sobolev-type embedding. In these works, the extended Scott–Zhang operator is replaced by an interpolation operator (reconstruction operator based on certain averages), which also maps from the DG space to the space of conformal discrete functions. Our use of the extended Scott–Zhang operator allows a little bit more flexibility, since it is in addition a projection and it preserves boundary values quite naturally.

We start with the classical setting.

Lemma A.7 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2$. For all $g \in W^{1,\psi}_{\Gamma_D}(\Omega)$ holds

$$\rho_{\psi,\Omega}(g) \leqslant c\rho_{\psi,\Omega}(\mathrm{diam}(\Omega)\nabla g),$$

where $c$ only depends on $\Omega$, $\Omega'$ and $\Delta_2(\psi)$.

*Proof.* We extend $g$ by zero on $\Omega' \setminus \Omega$, then $g \in W^{1,\psi}(\Omega')$. Let $B$ be a ball in $\Omega' \setminus \Omega$. Since $\Omega'$ is a domain with Lipschitz boundary, it is in particular a John domain (see Diening *et al.*, 2011, Section 7.4). Hence, by Diening *et al.* (2011, Lemma 8.2.1 b) it holds that[7]

$$|g(x)| \leqslant cI_1(\chi_\Omega \nabla g)(x) := c \int_\Omega \frac{|\nabla g(y)|}{|x-y|^{n-1}}\,\mathrm{d}y \tag{A.44}$$

[7] Note that in the definition of a John domain, with the help of the emanating chain condition, one can always choose the central John ball $B$ such that $B \subset \Omega' \setminus \Omega$.

for all $x \in \Omega'$. Note that $I_1$ is the standard Riesz potential operator. Set $R := \operatorname{diam}(\Omega)$. We can estimate $\int_\Omega R^{-1}|x-y|^{1-n}\,\mathrm{d}y \leqslant c_0$ with $c_0$ independent of $x \in \Omega$. Thus, we can apply Jensen's inequality with respect to the measure $\chi_\Omega c_0^{-1} R^{-1}|x-y|^{1-n}\,\mathrm{d}y$ to get

$$\begin{aligned}\rho_{\psi,\Omega}(g) &\leqslant \int_\Omega \psi\left(c\int_\Omega \frac{|\nabla g(y)|}{|x-y|^{n-1}}\,\mathrm{d}y\right)\mathrm{d}x\\ &\leqslant c\int_\Omega\int_\Omega \psi(R|\nabla g(y)|)R^{-1}|x-y|^{1-n}\,\mathrm{d}y\,\mathrm{d}x.\end{aligned}$$

Now the claim follows from $\int_\Omega |x-y|^{1-n}\,\mathrm{d}x \leqslant cR$. □

LEMMA A.8 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2 \cap \nabla_2$. For all $g \in W^{1,\psi}_{\Gamma_D}(\Omega)$ holds

$$\rho_{\psi,\Gamma_N}(g) \leqslant c\operatorname{diam}(\Omega)^{-1}\rho_{\psi,\Omega}(\operatorname{diam}(\Omega)\nabla g),$$

where $c$ only depends on $\Omega$, $\Omega'$, $\Delta_2(\psi)$ and $\Delta_2(\psi^*)$.

*Proof.* Let $R := \operatorname{diam}(\Omega)$. We proceed similar to the proof of Lemma A.7 and start with the estimate (A.44). Since $\psi \in \nabla_2$, it follows from Kokilashvili & Krbec (1991, Lemma 1.2.2 & Lemma 1.2.3) that there exist a $\Phi$-function $\rho$ and $\theta \in (0,1)$ with $\psi(\rho^{-1}(t)) \sim t^{1/\theta}$. Thus, Jensen's inequality with respect to the convex function $\rho$ and the measure $\chi_\Omega c_0^{-1}R^{-1}|x-y|^{1-n}\,\mathrm{d}y$ implies that

$$\begin{aligned}\rho_{\psi,\Gamma_N}(g) &\leqslant \int_{\Gamma_N} \psi\left(c\int_\Omega \frac{|\nabla g(y)|}{|x-y|^{n-1}}\,\mathrm{d}y\right)\mathrm{d}s(x)\\ &\leqslant c\int_{\Gamma_N} \psi\circ\rho^{-1}\left(\int_\Omega \rho(R|\nabla g(y)|)R^{-1}|x-y|^{1-n}\,\mathrm{d}y\right)\mathrm{d}s(x)\\ &\leqslant c\int_{\Gamma_N}\left(\int_\Omega \rho(R|\nabla g(y)|)R^{-1}|x-y|^{1-n}\,\mathrm{d}y\right)^{1/\theta}\mathrm{d}s(x).\end{aligned}$$

Let us denote the integrand of the outer integral by $I$. Then Hölder's inequality with exponent $1/\theta$ and $1/(1-\theta)$ with respect to the measure $|x-y|^{1-n}\,\mathrm{d}y$ implies for $\alpha > 0$

$$I \leqslant R^{-1/\theta}\int_\Omega \rho^{1/\theta}(R|\nabla g(y)|)|x-y|^{\alpha/\theta+1-n}\,\mathrm{d}y\left(\int_\Omega |x-y|^{-\alpha/(1-\theta)+1-n}\,\mathrm{d}y\right)^{(1-\theta)/\theta}.$$

We fix $\alpha > 0$ such that $-\alpha/(1-\theta)+1-n > -n$. This and $\rho^{1/\theta}(t) \sim \psi(t)$ gives

$$I \leqslant R^{-1/\theta}\int_\Omega \psi(R|\nabla g(y)|)|x-y|^{\alpha/\theta+1-n}\,\mathrm{d}y(R^{-\alpha/(1-\theta)+1})^{(1-\theta)/\theta}.$$

Together with the previous estimate and Fubini's theorem we have

$$\begin{aligned}\rho_{\psi,\Gamma_N}(g) &\leqslant cR^{-1-\alpha/\theta}\int_{\Gamma_N}\int_\Omega \psi(R|\nabla g(y)|)|x-y|^{\alpha/\theta+1-n}\,\mathrm{d}y\,\mathrm{d}s(x)\\ &\leqslant cR^{-1-\alpha/\theta}\int_\Omega \psi(R|\nabla g(y)|)\int_{\Gamma_N}|x-y|^{\alpha/\theta+1-n}\,\mathrm{d}s(x)\,\mathrm{d}y.\end{aligned}$$

Since $\alpha/\theta + 1 - n > 1 - n$ and $\Gamma_N$ is an $(n-1)$-dimensional Lipschitz surface, we have

$$\int_{\Gamma_N} |x-y|^{\alpha/\theta+1-n}\,\mathrm{d}s(x) \leqslant cR^{\alpha/\theta}$$

independent of $y \in \Omega$. Thus

$$\rho_{\psi,\Gamma_N}(g) \leqslant cR^{-1}\int_{\Omega} \psi(R|\nabla g(y)|)\,\mathrm{d}y$$

as desired. □

Lemma A.8 actually does not require that $\psi \in \nabla_2$. The $\nabla_2$-condition is only needed to simplify the proof.

Let us now turn to the DG setting.

LEMMA A.9 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2 \cap \nabla_2$. For all $g \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$

$$\rho_{\psi,\Omega}(g) \leqslant cM_{\psi,h}(\mathrm{diam}(\Omega)g)$$

holds, where $c$ only depends on $\Omega$, $\Omega'$, $\Delta_2(\psi)$ and $\Delta_2(\psi^*)$.

*Proof.* Let $g \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$. Then $\Pi_{\mathrm{SZ}}g \in W^{1,\psi}_{\Gamma_D}(\Omega)$. This and Lemma A.7 imply

$$\begin{aligned}\rho_{\psi,\Omega}(g) &\leqslant c\rho_{\psi,\Omega}(g-\Pi_{\mathrm{SZ}}g) + c\rho_{\psi,\Omega}(\Pi_{\mathrm{SZ}}g)\\ &\leqslant c\rho_{\psi,\Omega}(g-\Pi_{\mathrm{SZ}}g) + c\rho_{\psi,\Omega}(\mathrm{diam}(\Omega)\nabla\Pi_{\mathrm{SZ}}g).\end{aligned}$$

Now with (A.38) and $h_K \leqslant \mathrm{diam}(\Omega)$ it follows that

$$\rho_{\psi,\Omega}(g) \leqslant cM_{\psi,h}(\mathrm{diam}(\Omega)g).$$

This proves the claim. □

LEMMA A.10 Let $\psi$ be a $\Phi$-function with $\psi \in \Delta_2$. Then for all $g \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$

$$\rho_{\psi,\Gamma_N}(g) \leqslant c\,\mathrm{diam}(\Omega)^{-1}M_{\psi,h}(\mathrm{diam}(\Omega)g),$$

holds, where $c$ depends only on $\Omega$, $\Omega'$ and $\Delta_2(\psi)$.

*Proof.* Let $g \in W^{1,\psi}_{\mathrm{DG},\Gamma_D}(\Omega)$. Then $\Pi_{\mathrm{SZ}}g \in W^{1,\psi}_{\Gamma_D}(\Omega)$. Thus, with Lemma A.8 and (A.38), we obtain

$$\begin{aligned}\rho_{\psi,\Gamma_N}(g) &\leqslant c\rho_{\psi,\Gamma_N}(g-\Pi_{\mathrm{SZ}}g) + c\rho_{\psi,\Gamma_N}(\Pi_{\mathrm{SZ}}g)\\ &\leqslant c\rho_{\psi,\Gamma_N}(g-\Pi_{\mathrm{SZ}}g) + c\,\mathrm{diam}(\Omega)^{-1}\rho_{\psi,\Omega}(\mathrm{diam}(\Omega)\nabla\Pi_{\mathrm{SZ}}g)\\ &\leqslant c\rho_{\psi,\Gamma_N}(g-\Pi_{\mathrm{SZ}}g) + c\,\mathrm{diam}(\Omega)^{-1}M_{\psi,h}(\mathrm{diam}(\Omega)g).\end{aligned}$$

We estimate with Lemma A.4, Lemma A.1 and (A.35)

$$\begin{aligned}
\rho_{\psi,\Gamma_N}(g - \Pi_{\mathrm{SZ}} g) &\leqslant \sum_{\gamma \subset \Gamma_N} \int_\gamma \psi(|g - \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}s \\
&\leqslant c \sum_{\gamma \subset \Gamma_N} \int_\gamma \psi(|g - \langle g\rangle_K|) + \psi(|\langle g\rangle_K - \Pi_{\mathrm{SZ}} g)|)\,\mathrm{d}s \\
&\leqslant c \sum_{K \in \mathcal{T}_h\,:\,\partial K \cap \Gamma_N \neq \emptyset} h_K^{-1} \int_K \psi(h_K |\nabla_h g|)\,\mathrm{d}x \\
&\quad + c \sum_{K \in \mathcal{T}_h\,:\,\partial K \cap \Gamma_N \neq \emptyset} h_K^{-1} \int_K \psi(|\langle g\rangle_K - \Pi_{\mathrm{SZ}} g|) + \psi(h_K |\nabla \Pi_{\mathrm{SZ}} g|)\,\mathrm{d}x \\
&\leqslant c h_K^{-1} \int_{S'_K} \psi(h_K |\nabla_h g|)\,\mathrm{d}x + c \sum_{\gamma \in \Gamma_I(S'_K)} \int_\gamma \psi(|[\![ g\mathbf{n} ]\!]_\gamma|)\,\mathrm{d}s.
\end{aligned}$$

From $h_K \leqslant \mathrm{diam}(\Omega)$ and the convexity of $\psi$, it follows that

$$\rho_{\psi,\Gamma_N}(g - \Pi_{\mathrm{SZ}} g) \leqslant c\,\mathrm{diam}(\Omega)^{-1} M_{\psi,h}(\mathrm{diam}(\Omega) g).$$

And the claim follows. $\square$