



**HAL**  
open science

# Discontinuous Skeletal Gradient Discretisation Methods on polytopal meshes

Daniele Di Pietro, Jérôme Droniou, Gianmarco Manzini

► **To cite this version:**

Daniele Di Pietro, Jérôme Droniou, Gianmarco Manzini. Discontinuous Skeletal Gradient Discretisation Methods on polytopal meshes. 2017. hal-01564598v1

**HAL Id: hal-01564598**

**<https://hal.science/hal-01564598v1>**

Preprint submitted on 19 Jul 2017 (v1), last revised 6 Dec 2017 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Discontinuous Skeletal Gradient Discretisation Methods on polytopal meshes

Daniele A. Di Pietro<sup>1</sup>, Jérôme Droniou<sup>2</sup>, and Gianmarco Manzini<sup>3</sup>

<sup>1</sup>Institut Montpellierain Alexander Grothendieck, CNRS, Univ. Montpellier (France)

<sup>2</sup>School of Mathematical Sciences, Monash University, Melbourne (Australia)

<sup>3</sup>T-5 Applied Mathematics and Plasma Physics Group, Los Alamos National Laboratory, Los Alamos, New Mexico (USA)

July 19, 2017

## Abstract

In this work we develop arbitrary-order Discontinuous Skeletal Gradient Discretisations (DSGD) on general polytopal meshes. Discontinuous Skeletal refers to the fact that the globally coupled unknowns are broken polynomial on the mesh skeleton. The key ingredient is a high-order gradient reconstruction composed of two terms: (i) a consistent contribution obtained mimicking an integration by parts formula inside each element and (ii) a stabilising term for which sufficient design conditions are provided. An example of stabilisation that satisfies the design conditions is proposed based on a local lifting of high-order residuals on a Raviart–Thomas–Nédélec subspace. We prove that the novel DSGDs satisfy coercivity, consistency, limit-conformity, and compactness requirements that ensure convergence for a variety of elliptic and parabolic problems. Links with Hybrid High-Order, non-conforming Mimetic Finite Difference and non-conforming Virtual Element methods are also studied. Numerical examples complete the exposition.

**MSC2010:** 65N08, 65N30, 65N12

**Keywords:** Gradient discretisation methods, gradient schemes, high-order Mimetic Finite Difference methods, Hybrid High-Order methods, Virtual Element methods, non-linear problems

## 1 Introduction

The numerical resolution of (linear or non-linear) partial differential equations (PDEs) is nowadays ubiquitous in the engineering practice. In this context, the design of convergent numerical schemes is a very active research topic. The Gradient Discretisation Method (GDM) is a recently introduced framework which identifies key design properties to obtain convergent schemes for a variety of linear and non-linear elliptic and parabolic problems. Several models of current use in fluid mechanics fall into the latter categories including, e.g., porous media flows governed by Darcy’s law, phase change problems governed by the Stefan problem [34], as well as simplified models of the viscous terms in power-law fluids corresponding the Leray–Lions elliptic operators. The latter also appear in the modelling of glacier motion [36], of incompressible turbulent flows in porous media [25], and in airfoil design [35].

A Gradient Discretisation (GD) is defined by a finite-dimensional space encoding the discrete unknowns, as well as two linear operators acting on the latter, and corresponding to reconstructions of scalar functions and of their gradient. For a given PDE problem, convergent GDs are characterised by four properties, which can also serve as guidelines for the design of new schemes: *coercivity*, which corresponds to a discrete Poincaré inequality; *GD-consistency*, which expresses the ability of the scalar and gradient reconstructions to approximate functions in the space where the continuous problem is set; *limit-conformity*, linking the two reconstructions through an approximate integration by parts formula; *compactness*, corresponding to a discrete counterpart of the Rellich theorem.

In the recent monograph [27], several classical discretisation methods have been interpreted in the GDM framework. These include: arbitrary order conforming, nonconforming, and mixed Finite Elements (FE) on standard meshes; various lowest-order Finite Volume methods on specific grids; lowest-order methods belonging to the Hybrid Mixed Mimetic family (see the unified presentation in [28] of the methods originally proposed in [9, 26, 33]) as well as nodal Mimetic Finite Differences (MFD) [10] on arbitrary polyhedral meshes; see also [5].

In this paper we present an important addition to the GDM framework: arbitrary-order methods supporting general polyhedral meshes. Our focus is on Discontinuous Skeletal methods [18], characterised by globally coupled unknowns that are broken polynomials on the mesh skeleton. Specifically, the primary source of inspiration are the recently introduced Hybrid High-Order (HHO) methods for linear [22, 20] and non-linear [16, 17] diffusion problems, and the high-order non-conforming MFD (ncMFD) method of [39]; see also [2] for an interpretation in the Virtual Element framework and [4] for an introduction to the latter. We also cite here the Hybridizable Discontinuous Galerkin methods of [14], whose link with the former methods has been studied in [13]; see also [7] for a unified formulation.

Let a polynomial degree  $k \geq 0$  be given. The Discontinuous Skeletal Gradient Discretisations (DSGD) studied here hinge on face unknowns that ensure the global coupling and that correspond to broken polynomials of total degree up to  $k$  on the mesh skeleton, as well as locally coupled element-based unknowns that correspond to broken polynomials of total degree up to  $l \in \{k - 1, k, k + 1\}$  on the mesh itself. The reconstruction of scalar functions is defined in a straightforward manner through the latter if  $l \geq 0$ , or by a suitable combination of face-based unknowns if  $l = -1$ . The gradient reconstruction, on the other hand, requires a more careful design. The seminal ideas to devise high-order gradient reconstructions on general meshes are already present, among others, in HHO methods (see, e.g., [22, Eq. (13)] and [16, Eq. (4.3)]) as well as in ncMFD methods (see [39, Eq. (21)]). These gradient reconstructions, however, are not suitable to define a convergent DSGD because they fail to satisfy the coercivity requirement. In the context of discontinuous Galerkin methods, a stable discrete gradient based on a variation of the scheme originally proposed in [3] for the treatment of viscous terms in the incompressible Navier–Stokes equations has been recently studied in [1]. In addition, when considering non-linear problems, the codomain of the gradient reconstruction has to be carefully selected in order for the GD-consistency requirement to be satisfied with optimal scaling in the meshsize for  $k \geq 1$  (this point was already partially recognised in [16]).

The main novelty of this work is the introduction of a gradient reconstruction that meets all the requirements to define a convergent GD, and which satisfies the limit-conformity requirement with an error that scales optimally in the meshsize. This gradient reconstruction is composed of two terms: a consistent contribution closely inspired by [16, Eq. (4.3)]

and a stabilisation term. Two design conditions are identified for the stabilisation term: (i) local stability and boundedness with respect to a suitable boundary seminorm and (ii)  $L^2$ -orthogonality to vector-valued polynomials of degree up to  $k$ . When considering problems posed in a non-Hilbertian setting, an additional condition is added stipulating that the stabilisation is built on a piecewise polynomial space. An example of stabilisation term that meets all of the above requirements is proposed based on a Raviart–Thomas–Nédélec space on a submesh.

The rest of the paper is organised as follows. In Section 2 we recall the basics of the GDM and give a few examples of linear and non-linear problems for which GDs are convergent under the coercivity, GD-consistency, limit-conformity, and compactness properties discussed above. The construction of arbitrary-order DSGD is presented in Section 3, the main results are stated in Section 3.5, and numerical examples are collected in Section 3.7. The links with HHO and ncMFD schemes are studied in detail in Section 4. Appendix A contains the proofs of the main results. The material is organised so that multiple levels of reading are possible: readers mainly interested in the numerical recipe and results can primarily focus on Sections 2–3; readers also interested in the relations with other polytopal methods can consult Section 4.

## 2 The Gradient Discretisation Method

We give here a brief presentation of the Gradient Discretisation Method (GDM) in the context of homogeneous Dirichlet boundary conditions, and we refer to the monograph [27] for more details and other boundary conditions.

### 2.1 Gradient Discretisations and Gradient Schemes

Let  $\Omega$  be a bounded polytopal domain in  $\mathbb{R}^d$ , where  $d \geq 1$  is the space dimension. We consider elliptic or parabolic problems whose weak formulation is set in  $W_0^{1,p}(\Omega)$ , where  $p \in (1, +\infty)$  denotes a Sobolev exponent which we assume fixed in what follows.

A *Gradient Discretisation (GD)* is a triplet  $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$  where:

- (i)  $X_{\mathcal{D},0}$  is a finite dimensional vector space on  $\mathbb{R}$  encoding the *discrete unknowns*, and accounting for the homogeneous Dirichlet boundary condition;
- (ii)  $\Pi_{\mathcal{D}} : X_{\mathcal{D},0} \rightarrow L^p(\Omega)$  is a linear mapping that reconstructs scalar functions in  $L^p(\Omega)$  from the discrete unknowns in  $X_{\mathcal{D},0}$ ;
- (iii)  $\nabla_{\mathcal{D}} : X_{\mathcal{D},0} \rightarrow L^p(\Omega)^d$  is a linear mapping that reconstructs the gradient of scalar functions in  $L^p(\Omega)^d$  from the unknowns in  $X_{\mathcal{D},0}$ . This reconstruction must be defined such that  $\|\nabla_{\mathcal{D}} \cdot\|_{L^p(\Omega)^d}$  is a norm on  $X_{\mathcal{D},0}$ .

In a nutshell, the GDM consists in selecting a GD and in replacing, in the weak formulation of the PDE, the continuous space and operators by the discrete ones provided by the GD. The scheme thus obtained is called a *gradient scheme (GS)*. To illustrate this procedure, consider the simple linear problem: Find  $u : \Omega \rightarrow \mathbb{R}$  such that

$$\begin{aligned} -\nabla \cdot (\mathbf{A} \nabla u) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{1}$$

with diffusion tensor  $\mathbf{\Lambda}$  bounded and uniformly coercive, and source term  $f \in L^2(\Omega)$ . The weak formulation of (1) is

$$\text{Find } u \in H_0^1(\Omega) \text{ such that, for all } v \in H_0^1(\Omega), \int_{\Omega} \mathbf{\Lambda} \nabla u \cdot \nabla v = \int_{\Omega} f v. \quad (2)$$

Given a gradient discretisation  $\mathcal{D}$ , the gradient scheme for (2) is then

$$\text{Find } u_{\mathcal{D}} \in X_{\mathcal{D},0} \text{ such that, for all } v_{\mathcal{D}} \in X_{\mathcal{D},0}, \int_{\Omega} \mathbf{\Lambda} \nabla_{\mathcal{D}} u_{\mathcal{D}} \cdot \nabla_{\mathcal{D}} v_{\mathcal{D}} = \int_{\Omega} f \Pi_{\mathcal{D}} v_{\mathcal{D}}. \quad (3)$$

The same procedure applies to non-linear problems. Consider, e.g., the following generalisation of (1) that corresponds to Leray–Lions operators: Find  $u : \Omega \rightarrow \mathbb{R}$  such that

$$\begin{aligned} -\nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, u, \nabla u) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \quad (4)$$

where the flux function  $\boldsymbol{\sigma} : \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies the requirements detailed in [27, Section 3.3.2]. A paradigmatic example of this class of problems is the  $p$ -Laplace equation which, for a fixed  $p \in (1, +\infty)$ , corresponds to the flux function

$$\boldsymbol{\sigma}(\mathbf{x}, u, \nabla u) = |\nabla u|^{p-2} \nabla u. \quad (5)$$

Assuming  $f \in L^{p'}(\Omega)$  with  $p' := \frac{p}{p-1}$ , the Problem (4) admits the following weak formulation:

$$\text{Find } u \in W_0^{1,p}(\Omega) \text{ such that, for all } v \in W_0^{1,p}(\Omega), \int_{\Omega} \boldsymbol{\sigma}(\mathbf{x}, u, \nabla u) \cdot \nabla v = \int_{\Omega} f v. \quad (6)$$

Given a gradient discretisation  $\mathcal{D}$ , the gradient scheme for (6) is then

$$\text{Find } u_{\mathcal{D}} \in X_{\mathcal{D},0} \text{ such that, for all } v_{\mathcal{D}} \in X_{\mathcal{D},0}, \int_{\Omega} \boldsymbol{\sigma}(\mathbf{x}, \Pi_{\mathcal{D}} u, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \cdot \nabla_{\mathcal{D}} v_{\mathcal{D}} = \int_{\Omega} f \Pi_{\mathcal{D}} v_{\mathcal{D}}. \quad (7)$$

## 2.2 Convergent Gradient Schemes

The accuracy and convergence of GSs, for linear and non-linear problems, can be assessed by a few properties of the underlying GDs. In many situations, and in all cases considered in this paper, GDs are obtained starting from a mesh of the domain. We consider here *polytopal meshes* corresponding to couples  $\mathcal{M}_h := (\mathcal{T}_h, \mathcal{F}_h)$ , where  $\mathcal{T}_h$  is a finite collection of polytopal elements  $T$  of maximum diameter equal to  $h > 0$ , while  $\mathcal{F}_h$  is a finite collection of hyperplanar faces  $F$ . It is assumed henceforth that the mesh  $\mathcal{M}_h$  matches the weak geometrical requirements detailed in [27, Definition 7.2]; see also [24, Section 2]. Our focus is on the so-called  $h$ -convergence analysis, where we consider a sequence of refined meshes  $(\mathcal{M}_h)_{h \in \mathcal{H}}$  whose sizes are collected in a countable set  $\mathcal{H} \subset \mathbb{R}_*^+$  having 0 as its unique accumulation point. We further assume that the polytopal mesh sequences that we deal with are *regular* in the sense of [24, Definition 3], and we denote by  $\varrho > 0$  the corresponding regularity parameter.

The following properties allow us to single out sequences  $(\mathcal{D}_h)_{h \in \mathcal{H}} = (X_{\mathcal{D}_h,0}, \Pi_{\mathcal{D}_h}, \nabla_{\mathcal{D}_h})_{h \in \mathcal{H}}$  of GDs that lead to gradient schemes that converge, for both linear and non-linear problems:

**(GD1) Coercivity.** Consider, for all  $h \in \mathcal{H}$ , the norm of the linear mapping  $\Pi_{\mathcal{D}_h}$  defined by:

$$C_{\mathcal{D}_h} := \max_{v \in X_{\mathcal{D}_h,0} \setminus \{0\}} \frac{\|\Pi_{\mathcal{D}_h} v\|_{L^p(\Omega)}}{\|\nabla_{\mathcal{D}_h} v\|_{L^p(\Omega)^d}}.$$

Then, there exists a real number  $C_P > 0$  such that  $C_{\mathcal{D}_h} \leq C_P$  for all  $h \in \mathcal{H}$ .

**(GD2) GD-Consistency.** For all  $h \in \mathcal{H}$ , let  $S_{\mathcal{D}_h} : W_0^{1,p}(\Omega) \rightarrow [0, +\infty)$  be such that

$$S_{\mathcal{D}_h}(\phi) := \min_{v \in X_{\mathcal{D}_h,0}} \left( \|\Pi_{\mathcal{D}_h} v - \phi\|_{L^p(\Omega)} + \|\nabla_{\mathcal{D}_h} v - \nabla \phi\|_{L^p(\Omega)^d} \right) \quad \forall \phi \in W_0^{1,p}(\Omega).$$

Then, it holds that

$$\lim_{h \rightarrow 0} S_{\mathcal{D}_h}(\phi) = 0 \quad \forall \phi \in W_0^{1,p}(\Omega). \quad (8)$$

**(GD3) Limit-conformity.** Let  $p' := \frac{p}{p-1}$  denote the dual exponent of  $p$ , and set  $\mathbf{W}^{p'}(\text{div}; \Omega) := \{\boldsymbol{\psi} \in L^{p'}(\Omega)^d : \nabla \cdot \boldsymbol{\psi} \in L^{p'}(\Omega)\}$ . For all  $h \in \mathcal{H}$ , let  $W_{\mathcal{D}_h} : \mathbf{W}^{p'}(\text{div}; \Omega) \rightarrow [0, +\infty)$  be such that, for all  $\boldsymbol{\psi} \in \mathbf{W}^{p'}(\text{div}; \Omega)$ ,

$$W_{\mathcal{D}_h}(\boldsymbol{\psi}) := \sup_{v \in X_{\mathcal{D}_h,0} \setminus \{0\}} \frac{1}{\|\nabla_{\mathcal{D}_h} v\|_{L^p(\Omega)^d}} \left| \int_{\Omega} \left( \nabla_{\mathcal{D}_h} v(\mathbf{x}) \cdot \boldsymbol{\psi}(\mathbf{x}) + \Pi_{\mathcal{D}_h} v(\mathbf{x}) \nabla \cdot \boldsymbol{\psi}(\mathbf{x}) \right) d\mathbf{x} \right|.$$

Then, it holds that

$$\lim_{h \rightarrow 0} W_{\mathcal{D}_h}(\boldsymbol{\psi}) = 0 \quad \forall \boldsymbol{\psi} \in \mathbf{W}^{p'}(\text{div}; \Omega). \quad (9)$$

**(GD4) Compactness.** For any  $v_h \in X_{\mathcal{D}_h,0}$  such that  $(\|\nabla_{\mathcal{D}_h} v_h\|_{L^p(\Omega)^d})_{h \in \mathcal{H}}$  is bounded, the sequence  $(\Pi_{\mathcal{D}_h} v_h)_{h \in \mathcal{H}}$  is relatively compact in  $L^p(\Omega)$ .

A few comments are of order. Property **(GD1)** is linked to the stability of the method, and stipulates that the  $L^p$ -norm of the reconstruction of scalar functions is uniformly controlled by the  $L^p$ -norm of the reconstruction of their gradient. It readily implies the uniform Poincaré inequality  $\|\Pi_{\mathcal{D}_h} v_h\|_{L^p(\Omega)} \leq C_P \|\nabla_{\mathcal{D}_h} v_h\|_{L^p(\Omega)^d}$  valid for all  $h \in \mathcal{H}$  and all  $v_h \in X_{\mathcal{D}_h,0}$ .

Properties **(GD2)** and **(GD3)** are linked to the consistency of the method. More specifically, property **(GD2)** states that the reconstructions  $\Pi_{\mathcal{D}_h}$  of scalar functions and  $\nabla_{\mathcal{D}_h}$  of their gradients are able to approximate functions that lie in the space  $W_0^{1,p}(\Omega)$  where the continuous problem is set. In the context of the FE convergence analysis, this property is an attribute of the underlying discrete space, and is usually called *approximability*; see, e.g., [32, Definition 2.14]. Property **(GD3)**, on the other hand, establishes a link between  $\Pi_{\mathcal{D}_h}$  and  $\nabla_{\mathcal{D}_h}$  in the form of a discrete integration by parts formula. Its counterpart in the context of the FE convergence analysis for linear problems is *asymptotic consistency*; see, e.g., [32, Definition 2.15]. Notice, however, that the formulation in **(GD3)** is in a sense more general, as it is not linked to a specific underlying problem and is in particular readily applicable to non-linear problems (whereas [32, Definition 2.15] is restricted to linear problems).

Finally, property **(GD4)** is a discrete Rellich compactness theorem, and can be regarded as the key ingredient to obtain strong convergence results by compactness techniques.

The above properties are sufficient to carry out a convergence analysis, either by error estimates (when the model is amenable to these) or by compactness, for a variety of linear

and non-linear elliptic or parabolic models. An example of such convergence results for gradient discretisations of the Leray–Lions problem (6) is provided by Theorems 1 and 2 below; see [27] for a comprehensive collection of convergence results for various linear and non-linear elliptic and parabolic problems.

**Theorem 1** (Convergence). *We assume that  $\sigma$  satisfies the classical properties of Leray–Lions operators (see [27, Section 3.3.2]). Let  $(\mathcal{D}_h)_{h \in \mathcal{H}}$  denote a sequence of GDs satisfying (GD1)–(GD4). Then, for all  $h \in \mathcal{H}$ , there exists at least one  $u_{\mathcal{D}_h} \in X_{\mathcal{D}_h,0}$  solution to (7) and, along a subsequence as  $h \rightarrow 0$ , (i)  $\Pi_{\mathcal{D}_h} u_{\mathcal{D}_h}$  converges strongly in  $L^p(\Omega)$  to a solution  $u$  of (6); (ii)  $\nabla_{\mathcal{D}_h} u_{\mathcal{D}_h}$  converges strongly in  $L^p(\Omega)^d$  to  $\nabla u$ .*

*Proof.* This is a special case of [27, Theorem 3.34].  $\square$

**Theorem 2** (Error estimates). *Let  $\mathcal{D}$  be a gradient discretisation, and let  $\sigma$  be the Leray–Lions operator corresponding to the  $p$ -Laplace equation (see (5)). Then there exists a unique  $u_{\mathcal{D}}$  solution to (7) and, if  $u$  is the solution to (6), then there exists  $C$  depending only on  $p$ ,  $f$  and an upper bound of  $C_{\mathcal{D}}$  such that*

- If  $1 < p \leq 2$ ,

$$\|u - \Pi_{\mathcal{D}} u_{\mathcal{D}}\|_{L^p(\Omega)} + \|\nabla u - \nabla_{\mathcal{D}} u_{\mathcal{D}}\|_{L^p(\Omega)} \leq C [S_{\mathcal{D}}(u) + S_{\mathcal{D}}(u)^{p-1} + W_{\mathcal{D}}(\sigma(\nabla u))].$$

- If  $2 \leq p$ ,

$$\|u - \Pi_{\mathcal{D}} u_{\mathcal{D}}\|_{L^p(\Omega)} + \|\nabla u - \nabla_{\mathcal{D}} u_{\mathcal{D}}\|_{L^p(\Omega)} \leq C \left[ S_{\mathcal{D}}(u) + S_{\mathcal{D}}(u)^{\frac{1}{p-1}} + W_{\mathcal{D}}(\sigma(\nabla u))^{\frac{1}{p-1}} \right].$$

*Proof.* These error estimates are simplified forms of the estimates in [27, Theorem 3.28].  $\square$

### 3 Discontinuous Skeletal Gradient Discretisations

In this section, we construct a family of Discontinuous Skeletal Gradient Discretisations (DSGD). The notation is closely inspired by HHO methods; see, e.g., [24].

#### 3.1 Local polynomial spaces and projectors

Local polynomial spaces on mesh elements or faces and projectors thereon play a crucial role in the design and analysis of DSGD methods.

For any  $X \subset \bar{\Omega}$ , we denote by  $(\cdot, \cdot)_X$  the standard  $L^2(X)$ - or  $L^2(X)^d$ -products. This notation is used in place of integrals when dealing with quantities that are inherently  $L^2$ -based. Let now  $X$  be a mesh element or face. For an integer  $\ell \geq 0$ ,  $\mathbb{P}^\ell(X)$  denotes the space spanned by the restriction to  $X$  of scalar-valued,  $d$ -variate (if  $X$  is a mesh element) or  $(d-1)$ -variate (if  $X$  is a face) polynomials of total degree  $\ell$  or less, and conventionally set  $\mathbb{P}^{-1}(X) := \{0\}$ .

Let again  $X$  denote a mesh element or face. The  $L^2$ -orthogonal projector  $\pi_X^{0,\ell} : L^1(X) \rightarrow \mathbb{P}^\ell(X)$  is defined as follows: For all  $v \in L^1(X)$ ,  $\pi_X^{0,\ell}$  is the unique polynomial in  $\mathbb{P}^\ell(X)$  such that

$$(\pi_X^{0,\ell} v - v, w)_X = 0 \quad \forall w \in \mathbb{P}^\ell(X). \quad (10)$$



In the vector case, the  $L^2$ -projector is defined component-wise and denoted by  $\pi_X^{0,\ell}$ .

For any mesh element  $T \in \mathcal{T}_h$ , we also define the elliptic projector  $\pi_T^{1,\ell} : W^{1,1}(T) \rightarrow \mathbb{P}^\ell(T)$  as follows: For all  $v \in W^{1,1}(T)$ ,  $\pi_T^{1,\ell} v$  is the unique polynomial in  $\mathbb{P}^\ell(T)$  that satisfies

$$(\nabla(\pi_T^{1,\ell} v - v), \nabla w)_T = 0 \text{ for all } w \in \mathbb{P}^\ell(T) \text{ and } (\pi_T^{1,\ell} v - v, 1)_T = 0.$$

On regular polytopal mesh sequences, both  $\pi_T^{0,\ell}$  and  $\pi_T^{1,\ell}$  have optimal approximation properties in  $\mathbb{P}^\ell(T)$  (see Theorem 1.1, Theorem 1.2, and Lemma 3.1 in [17]): For any  $\alpha \in \{0, 1\}$  and  $s \in \{\alpha, \dots, \ell + 1\}$ , there exists a real number  $C > 0$  independent of  $h$ , but possibly depending only on  $d, p, \varrho, \ell, \alpha$ , and  $s$ , such that, for all  $T \in \mathcal{T}_h$ , and all  $v \in W^{s,p}(T)$ ,

$$|v - \pi_T^{\alpha,\ell} v|_{W^{r,p}(T)} \leq Ch_T^{s-r} |v|_{W^{s,p}(T)} \quad \forall r \in \{0, \dots, s\}, \quad (11a)$$

and, if  $s \geq 1$ ,

$$h_T^{\frac{1}{p}} |v - \pi_T^{\alpha,\ell} v|_{W^{r,p}(\mathcal{F}_T)} \leq Ch_T^{s-r} |v|_{W^{s,p}(T)} \quad \forall r \in \{0, \dots, s-1\}, \quad (11b)$$

where  $W^{r,p}(\mathcal{F}_T) := \{v \in L^p(\partial T) : v|_F \in W^{r,p}(F) \text{ for all } F \in \mathcal{F}_T\}$  and  $h_T$  denotes the diameter of the element  $T$ .

### 3.2 Computing gradient projections from projections of scalar functions

We continue our discussion with a crucial remark concerning the computation of the  $L^2$ -orthogonal projection of the gradient from  $L^2$ -orthogonal projections of a scalar function and its traces. This remark will inspire the choice of the discrete unknowns as well as the definition of the gradient reconstruction in DSGD methods. In what follows, we work on a fixed mesh element  $T \in \mathcal{T}_h$ , denote by  $\mathcal{F}_T$  the set of mesh faces that lie on the boundary of  $T$  and, for all  $F \in \mathcal{F}_T$ , by  $\mathbf{n}_{TF}$  the normal vector to  $F$  pointing out of  $T$ .

Consider a function  $v \in W^{1,1}(T)$ . We note the following integration by parts formula, valid for all  $\phi \in C^\infty(\overline{T})^d$ :

$$(\nabla v, \phi)_T = -(v, \nabla \cdot \phi)_T + \sum_{F \in \mathcal{F}_T} (v, \phi \cdot \mathbf{n}_{TF})_F. \quad (12)$$

Let now an integer  $k \geq 0$  be fixed. Specialising (12) to  $\phi \in \mathbb{P}^k(T)^d$ , we obtain

$$(\pi_T^{0,k} \nabla v, \phi)_T = -(\pi_T^{0,k-1} v, \nabla \cdot \phi)_T + \sum_{F \in \mathcal{F}_T} (\pi_F^{0,k} v, \phi \cdot \mathbf{n}_{TF})_F, \quad (13)$$

where we have used (10) to insert  $\pi_T^{0,k}$  into the left-hand side, and  $\pi_T^{0,k-1}$  and  $\pi_F^{0,k}$  into the right-hand side after observing that  $\nabla \cdot \phi \in \mathbb{P}^{k-1}(T)$  and, since we are considering planar faces,  $\phi|_F \cdot \mathbf{n}_{TF} \in \mathbb{P}^k(F)$  for all  $F \in \mathcal{F}_T$ . The relation (13) show that computing the  $L^2$ -orthogonal projection of  $\nabla v$  on  $\mathbb{P}^k(T)^d$  does not require a full knowledge of the function  $v$ . All that is required is

- (i)  $\pi_T^{0,k-1} v$ , the  $L^2$ -projection of  $v$  on  $\mathbb{P}^{k-1}(T)$ . Other possible choices are  $\pi_T^{0,k} v$  or  $\pi_T^{0,k+1} v$  (in fact, any polynomial degree larger then or equal to  $k-1$  will do);
- (ii) for all  $F \in \mathcal{F}_T$ ,  $\pi_F^{0,k} v$ , the  $L^2$ -projection on  $\mathbb{P}^k(F)$  of the trace of  $v$  on  $F$ .



### 3.3 Space of discrete unknowns and reconstruction of scalar functions

Inspired by the previous remark, for two given integers  $k \geq 0$  and  $l \in \{k-1, k, k+1\}$  we consider the following set of discrete unknowns:

$$\underline{U}_h^{k,l} := \left( \times_{T \in \mathcal{T}_h} \mathbb{P}^l(T) \right) \times \left( \times_{F \in \mathcal{F}_h} \mathbb{P}^k(F) \right).$$

The choice  $l = k-1$  can be traced back to the ncMFD of [39], the choice  $l = k$  to the Hybrid High-Order method of [22], and the choice  $l = k+1$  to the Hybridizable Discontinuous Galerkin method of [38, Remark 1.2.4]. Notice that, for  $k = 0$  and  $l = k-1$ , element-based unknowns are not present.

For a generic element of  $\underline{U}_h^{k,l}$ , we use the standard HHO underlined notation

$$\underline{v}_h = ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}),$$

and we define the interpolator  $\underline{I}_h^{k,l} : W^{1,1}(\Omega) \rightarrow \underline{U}_h^{k,l}$  such that, for all  $v \in W^{1,1}(\Omega)$ ,

$$\underline{I}_h^{k,l} v := \left( (\pi_T^{0,l} v)_{T \in \mathcal{T}_h}, (\pi_F^{0,k} v|_F)_{F \in \mathcal{F}_h} \right).$$

To account for Dirichlet boundary conditions strongly, we introduce the subspace

$$\underline{U}_{h,0}^{k,l} := \left\{ \underline{v}_h \in \underline{U}_h^{k,l} : v_F \equiv 0 \text{ for all } F \in \mathcal{F}_h^b \right\},$$

where  $\mathcal{F}_h^b$  is the set collecting the mesh faces that lie on the boundary of  $\Omega$ .

The restrictions of  $\underline{U}_h^{k,l}$ ,  $\underline{I}_h^{k,l}$  and  $\underline{v}_h \in \underline{U}_h^{k,l}$  to a generic mesh element  $T \in \mathcal{T}_h$  are denoted by  $\underline{U}_T^{k,l}$ ,  $\underline{I}_T^{k,l}$ , and  $v_T$ , respectively. Moreover, we adopt the convention that, for all  $T \in \mathcal{T}_h$ ,

$$v_T := \sum_{F \in \mathcal{F}_T} \omega_{TF} v_F \text{ if } l < 0, \quad (14)$$

where, following [39, Appendix A], the weights  $\{\omega_{TF}\}_{F \in \mathcal{F}_T}$  are defined in such a way that  $\sum_{F \in \mathcal{F}_T} \omega_{TF} (q, 1)_F = (q, 1)_T$  for all  $q \in \mathbb{P}^1(T)$  (this condition is required in the above reference to obtain  $L^2$ -superconvergence, not treated in this work). For all  $\underline{v}_h \in \underline{U}_h^{k,l}$ , we also define the broken polynomial field  $v_h$  such that

$$v_h|_T := v_T \quad \forall T \in \mathcal{T}_h. \quad (15)$$

The space of discrete unknowns and the reconstruction of the scalar variable for a DSDG are given by, respectively,

$$X_{\mathcal{D}_h,0} := \underline{U}_{h,0}^{k,l} \quad \text{and} \quad \Pi_{\mathcal{D}_h} \underline{v}_h := v_h \text{ for all } \underline{v}_h \in \underline{U}_{h,0}^{k,l}. \quad (16)$$

### 3.4 Reconstruction of the gradient

To complete the definition of a DSGD, it remains to identify a reconstruction of the gradient, which makes the object of this section.

### 3.4.1 A consistent and limit-conforming high-order gradient

Let a mesh element  $T \in \mathcal{T}_h$  be fixed. Taking inspiration from the integration by parts formula (13), we define the gradient reconstruction  $\mathbf{G}_T^k : \underline{U}_T^{k,l} \rightarrow \mathbb{P}^k(T)^d$  such that, for any  $\underline{v}_T := (v_T, (v_F)_{F \in \mathcal{F}_T}) \in \underline{U}_T^{k,l}$ ,  $\mathbf{G}_T^k \underline{v}_T$  satisfies, for all  $\phi \in \mathbb{P}^k(T)^d$ ,

$$(\mathbf{G}_T^k \underline{v}_T, \phi)_T = -(v_T, \nabla \cdot \phi)_T + \sum_{F \in \mathcal{F}_T} (v_F, \phi \cdot \mathbf{n}_{TF})_F. \quad (17)$$

By construction, it holds for all  $v \in W^{1,1}(T)$ ,

$$\mathbf{G}_T^k \underline{I}_T^{k,l} v = \pi_T^{0,k}(\nabla v). \quad (18)$$

Recalling the estimates (11) on  $\pi_T^{0,k}$ , this implies that  $\mathbf{G}_T^k \underline{I}_T^{k,l} v$  optimally approximates  $\nabla v$  in  $\mathbb{P}^k(T)^d$  when  $v$  is smooth enough.

A reconstruction of the gradient that meets both the consistency and limit-conformity requirements expressed by (GD2) and (GD3), respectively, can be obtained at this point letting  $\nabla_{\mathcal{D}_h}$  be such that, for all  $\underline{v}_h \in \underline{U}_{h,0}^{k,l}$ ,

$$(\nabla_{\mathcal{D}_h} \underline{v}_h)|_T = \mathbf{G}_T^k \underline{v}_T \quad \forall T \in \mathcal{T}_h. \quad (19)$$

However, for general element shapes, the  $L^p$ -norm of this gradient reconstruction is not a norm on the space  $X_{\mathcal{D}_h,0} = \underline{U}_{h,0}^{k,l}$ , hence the coercivity requirement expressed by (GD1) cannot be met. This initial choice of reconstructed gradient therefore has to be stabilised by accounting for jumps between element and face unknowns. These jumps can be controlled in turn via a discrete counterpart of the  $W^{1,p}$ -seminorm, which gives us an emulated Sobolev structure on  $\underline{U}_h^{k,l}$ .

*Remark 3* (Crouzeix–Raviart discretisations). If  $T$  is a  $d$ -simplex (i.e., a triangle if  $d = 2$ , a tetrahedron if  $d = 3$ , etc.) and we take  $k = 0$  and  $l = -1$ , the gradient reconstruction  $\mathbf{G}_T^k \underline{v}_T$  defined by (17) coincides with the gradient of the unique Crouzeix–Raviart [15] function  $\varphi$  such that  $|F|^{-1} \int_T \varphi(\mathbf{x}_F) = v_F$  for all  $F \in \mathcal{F}_T$ . In this case, the  $L^p$ -norm of the global gradient defined by (19) defines a norm on the space of discrete unknowns, and stabilisation is not needed. To recover Crouzeix–Raviart discretisations, (16) has to be modified setting  $\Pi_{\mathcal{D}_h} \underline{v}_h|_T := \mathbf{r}_T^1 \underline{v}_T$  for all  $T \in \mathcal{T}_h$  and all  $\underline{v}_T \in \underline{U}_T^{0,-1}$ , where  $\mathbf{r}_T^1$  is the high-order reconstruction of scalar function defined in the following section.

### 3.4.2 High-order reconstruction of scalar functions and discrete $W^{1,p}$ -seminorm

Let  $\underline{v}_T \in \underline{U}_T^{k,l}$ . Recalling the convention (14),  $v_T$  defines a reconstruction of scalar functions inside  $T$  of degree  $\max(0, l)$ . However, taking again inspiration from the integration by parts formula (13), this time specialised to  $\phi = \nabla w$  with  $w \in \mathbb{P}^{k+1}(T)$ , one can define a higher-order reconstruction  $\mathbf{r}_T^{k+1} : \underline{U}_T^{k,l} \rightarrow \mathbb{P}^{k+1}(T)$  such that, for all  $\underline{v}_T \in \underline{U}_T^{k,l}$ ,  $\mathbf{r}_T^{k+1} \underline{v}_T$  satisfies, for all  $w \in \mathbb{P}^{k+1}(T)$ ,

$$(\nabla \mathbf{r}_T^{k+1} \underline{v}_T, \nabla w)_T = -(v_T, \Delta w)_T + \sum_{F \in \mathcal{F}_T} (v_F, \nabla w \cdot \mathbf{n}_{TF})_F. \quad (20a)$$

Equation (20a) defines  $\mathbf{r}_T^{k+1} \underline{v}_T$  up to an additive constant, which we fix by imposing

$$(\mathbf{r}_T^{k+1} \underline{v}_T - v_T, 1)_T = 0. \quad (20b)$$

*Remark 4* (Optimal approximation properties of  $\mathbf{r}_T^{k+1} \circ \underline{\mathbf{I}}_T^{k,l}$ ). When  $l \geq 0$ , following the reasoning of [22, Lemma 3], it can be proved that  $\mathbf{r}_T^{k+1} \circ \underline{\mathbf{I}}_T^{k,l} = \pi_T^{1,k+1}$ , and optimal approximation properties in  $\mathbb{P}^{k+1}(T)$  follow from (11) with  $\alpha = 1$  and  $\ell = k + 1$ . The case  $l < 0$ , on the other hand, can only occur when  $k = 0$ . Owing to the specific choice for the reconstruction of a (constant) element value in (14), optimal approximation properties analogous to (11) with  $\alpha = 1$  and  $\ell = 1$  can be proved also in this case.

To define the discrete Sobolev seminorm on  $\underline{U}_h^{k,l}$ , for all  $T \in \mathcal{T}_h$  we introduce the difference operators  $\delta_T^l : \underline{U}_T^{k,l} \rightarrow \mathbb{P}^l(T)$  and, for all  $F \in \mathcal{F}_T$ ,  $\delta_{TF}^k : \underline{U}_T^{k,l} \rightarrow \mathbb{P}^k(F)$  such that, for all  $\underline{v}_T \in \underline{U}_T^{k,l}$ ,

$$\delta_T^l \underline{v}_T := \pi_T^{0,l}(\mathbf{r}_T^{k+1} \underline{v}_T - v_T), \quad \delta_{TF}^k \underline{v}_T := \pi_F^{0,k}(\mathbf{r}_T^{k+1} \underline{v}_T - v_F) \quad \forall F \in \mathcal{F}_T. \quad (21)$$

The role of these difference operators in the context of HHO methods has been highlighted in [24, Section 3.1.4]. We also note here the following relation holds:

$$(\delta_T^l \underline{v}_T, (\delta_{TF}^k \underline{v}_T)_{F \in \mathcal{F}_T}) = \underline{\mathbf{I}}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T - \underline{v}_T, \quad (22)$$

which will be exploited in Section 4.5 and Lemma 20 below.

The discrete  $W^{1,p}$ -seminorm is defined setting

$$\|\underline{v}_h\|_{1,p,h}^p := \sum_{T \in \mathcal{T}_h} \|\underline{v}_T\|_{1,p,T}^p,$$

where, for all  $T \in \mathcal{T}_h$ , the local seminorm is such that, denoting by  $h_F$  the diameter of the face  $F$ ,

$$\|\underline{v}_T\|_{1,p,T}^p := \|\mathbf{G}_T^k \underline{v}_T\|_{L^p(T)^d}^p + |\underline{v}_T|_{p,\partial T}^p, \quad |\underline{v}_T|_{p,\partial T}^p := \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^p(F)}^p. \quad (23)$$

As a result of Proposition 19 below,  $\|\cdot\|_{1,p,h}$  is a norm on the subspace  $\underline{U}_{h,0}^{k,l}$ .

### 3.4.3 A stabilised reconstruction of the gradient

We can now describe the general form of the gradient  $\nabla_{\mathcal{D}_h} : \underline{U}_h^{k,l} \rightarrow L^2(\Omega)^d$ , built inside each mesh element from the consistent and limit-conforming part  $\mathbf{G}_T^k$  and a stabilising contribution:

$$(\nabla_{\mathcal{D}_h} \underline{v}_h)|_T = \mathcal{G}_T \underline{v}_T := \mathbf{G}_T^k \underline{v}_T + \mathbf{S}_T \underline{v}_T \quad \forall \underline{v}_h \in \underline{U}_h^{k,l}, \quad \forall T \in \mathcal{T}_h, \quad (24)$$

where  $\mathbf{S}_T : \underline{U}_T^{k,l} \rightarrow L^2(T)^d$  satisfies the following design conditions:

**(S1)**  *$L^2$ -stability and boundedness.* For all  $T \in \mathcal{T}_h$  and all  $\underline{v}_T \in \underline{U}_T^{k,l}$ , it holds that

$$\|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d} \simeq |\underline{v}_T|_{2,\partial T}, \quad (25)$$

where  $a \simeq b$  means  $Ca \leq b \leq C^{-1}a$  with real number  $C > 0$  independent of  $h$  and of  $T$ , but possibly depending on  $d$  and on discretisation parameters including  $\rho$ ,  $k$ , and  $l$ .

**(S2)** *Orthogonality.* For all  $\underline{v}_T \in \underline{U}_T^{k,l}$  and all  $\phi \in \mathbb{P}^k(T)^d$ , it holds

$$(\mathbf{S}_T \underline{v}_T, \phi)_T = 0. \quad (26)$$

**(S3) Image.** If  $p \neq 2$ , there exists  $k_S \in \mathbb{N}$  such that the image of  $\mathbf{S}_T$  is contained in  $\mathbb{P}^{k_S}(\mathcal{P}_T)$ , the space of broken polynomials of total degree up to  $k_S$  on a regular polytopal partition  $\mathcal{P}_T$  of  $T$ . Here, regular means that, for all  $P \in \mathcal{P}_T$ , denoting by  $r_P$  and  $h_P$  the inradius and diameter of  $P$ , respectively, it holds that

$$\varrho h_P \leq r_P, \quad \varrho h_T \leq h_P. \quad (27)$$

*Remark 5* ( $L^2$ -based stabilising contribution). Property **(S2)**, which is crucial to ensure the stabilising properties of  $\mathbf{S}_T$ , requires to work with an inner product space. In our case, a natural choice is  $L^2(T)^d$ . The role of orthogonality properties analogous to **(S2)** has been previously recognised in the context of specific stabilised method, see for example [8, Proposition 7] for the lowest-order Compatible Discretization Operator methods, [27, Proposition 12.5 and Theorem 13.5] for the Hybrid Mimetic Mixed methods and the nodal Mimetic Mixed Methods, and [31, Section 4.2] for numerical methods for elasticity models.

*Remark 6* ( $L^p$ -stability of  $\mathbf{S}_T$ ). Property **(S3)** is required to extend the stability properties expressed by **(S1)** to  $L^p$ ; see the proof of point (i) in Proposition 18 for further details.

The above construction of a DSGD is summarised in the following

**Definition 7** (Discontinuous Skeletal Gradient Discretisation). Given a polytopal mesh  $\mathcal{M}_h$ , a Discontinuous Skeletal Gradient Discretisation (DSGD) is given by  $\mathcal{D}_h = (X_{\mathcal{D}_h,0}, \Pi_{\mathcal{D}_h}, \nabla_{\mathcal{D}_h})$  where  $X_{\mathcal{D}_h,0}$  and  $\Pi_{\mathcal{D}_h}$  are defined by (16), and  $\nabla_{\mathcal{D}_h}$  is given by (24) with a family of stabilisations  $\{\mathbf{S}_T : T \in \mathcal{M}_h\}$  satisfying properties **(S1)**–**(S3)**.

### 3.5 Main results

The construction detailed above yields a GD that meets properties **(GD1)**–**(GD4)** identified in Section 2, as summarised in the following

**Theorem 8** (Properties of DSGD). *If  $(\mathcal{M}_h)_{h \in \mathcal{H}}$  is a regular sequence of polytopal meshes, then the sequence of the corresponding DSGDs  $(\mathcal{D}_h)_{h \in \mathcal{H}}$  given by Definition 7 satisfies properties **(GD1)**–**(GD4)**.*

*Proof.* See Appendix A.3.1. □

Since we are dealing with arbitrary-order methods, given the error estimates in Theorem 2, a relevant point consists in estimating the convergence rates of the quantities  $S_{\mathcal{D}_h}(\phi)$  (see **(GD2)**) and  $W_{\mathcal{D}}(\psi)$  (see **(GD3)**) when they arguments exhibit further regularity. This makes the object of the following

**Proposition 9** (Estimates on  $S_{\mathcal{D}}$  and  $W_{\mathcal{D}}$ ). *Let  $\mathcal{M}_h$  be a polytopal mesh and  $\mathcal{D}_h$  be a DSGD as in Definition 7. Then, denoting by  $a \lesssim b$  the inequality  $a \leq Cb$  with real number  $C > 0$  not depending on  $h$ , but possibly depending on  $d, p, \varrho, k, l$ , and  $k_S$ , it holds with  $l^+ := \max(l, 0)$ ,*

$$\forall \phi \in W_0^{1,p}(\Omega) \cap W^{l^++1,p}(\mathcal{T}_h), \quad \|\Pi_{\mathcal{D}_h} \underline{I}_h^{k,l} \phi - \phi\|_{L^p(\Omega)} \lesssim h^{l^++1} |\phi|_{W^{l^++1,p}(\mathcal{T}_h)}, \quad (28a)$$

$$\forall \phi \in W_0^{1,p}(\Omega) \cap W^{k+2,p}(\mathcal{T}_h), \quad \|\nabla_{\mathcal{D}_h} \underline{I}_h^{k,l} \phi - \nabla \phi\|_{L^p(\Omega)^d} \lesssim h^{k+1} |\phi|_{W^{k+2,p}(\mathcal{T}_h)}. \quad (28b)$$

As a consequence,

$$\forall \phi \in W_0^{1,p}(\Omega) \cap W^{\min(k,l^+)+2,p}(\mathcal{T}_h), \quad S_{\mathcal{D}_h}(\phi) \lesssim h^{\min(k,l^+)+1} \|\phi\|_{W^{\min(k,l^+)+2,p}(\mathcal{T}_h)}. \quad (29)$$

Moreover,

$$\forall \boldsymbol{\psi} \in \mathbf{W}^{p'}(\operatorname{div}; \Omega) \cap W^{k+1, p'}(\mathcal{T}_h)^d, \quad W_{\mathcal{D}_h}(\boldsymbol{\psi}) \lesssim h^{k+1} \|\boldsymbol{\psi}\|_{W^{k+1, p'}(\mathcal{T}_h)^d}. \quad (30)$$

Here, for an integer  $s \geq 0$  and a real number  $q \in [1, +\infty)$ ,  $W^{s, q}(\mathcal{T}_h) := \{v \in L^q(\Omega) : v|_T \in W^{s, q}(T) \forall T \in \mathcal{T}_h\}$  is the broken space on  $\mathcal{T}_h$  constructed on  $W^{s, q}$  and endowed with the norm

$$\|v\|_{W^{s, q}(\mathcal{T}_h)} := \left( \sum_{T \in \mathcal{T}_h} \|v|_T\|_{W^{s, q}(T)}^q \right)^{\frac{1}{q}}.$$

*Proof.* See Appendix A.3.2. □

*Remark 10* (Order of  $S_{\mathcal{D}_h}$ ). In the case  $l \geq k$ , (29) yields the optimal order  $\mathcal{O}(h^{k+1})$  for interpolations of smooth enough functions. If  $l = k - 1$ , one order is lost and (29) gives an  $\mathcal{O}(h^k)$  estimate (but, as shown by (28), this loss is only perceptible on the approximations of the functions, not of their gradients).

### 3.6 Local stabilising contribution based on a Raviart–Thomas–Nédélec subspace

We construct in this section a stabilising contribution that fulfills the requirements expressed by (S1)–(S3).

#### 3.6.1 An inspiring remark

Let, for the sake of brevity,

$$\boldsymbol{\delta}_{\nabla, T}^k := \nabla \mathbf{r}_T^{k+1} - \mathbf{G}_T^k.$$

We start by observing that it holds, for all  $\underline{v}_T \in \underline{U}_T^{k, l}$  and all  $\boldsymbol{\phi} \in \mathbb{P}^k(T)^d$ ,

$$\begin{aligned} (\boldsymbol{\delta}_{\nabla, T}^k \underline{v}_T, \boldsymbol{\phi})_T &= (\nabla \mathbf{r}_T^{k+1} \underline{v}_T, \boldsymbol{\phi})_T - (\mathbf{G}_T^k \underline{v}_T, \boldsymbol{\phi})_T \\ &= (v_T - \mathbf{r}_T^{k+1} \underline{v}_T, \nabla \cdot \boldsymbol{\phi})_T + \sum_{F \in \mathcal{F}_T} (\mathbf{r}_T^{k+1} \underline{v}_T - v_F, \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F \\ &= (\pi_T^{0, l} (v_T - \mathbf{r}_T^{k+1} \underline{v}_T), \nabla \cdot \boldsymbol{\phi})_T + \sum_{F \in \mathcal{F}_T} (\pi_F^{0, k} (\mathbf{r}_T^{k+1} \underline{v}_T - v_F), \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F \\ &= -(\delta_T^l \underline{v}_T, \nabla \cdot \boldsymbol{\phi})_T + \sum_{F \in \mathcal{F}_T} (\delta_{TF}^k \underline{v}_T, \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F \\ &= (\nabla \delta_T^l \underline{v}_T, \boldsymbol{\phi})_T + \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F, \end{aligned}$$

where we have used the definition of  $\boldsymbol{\delta}_{\nabla, T}^k$  in the first line, an integration by parts together with the definition (17) of  $\mathbf{G}_T^k$  in the second line, (10) together with the fact that  $\nabla \cdot \boldsymbol{\phi} \in \mathbb{P}^{k-1}(T) \subset \mathbb{P}^l(T)$  since  $l \geq k - 1$  and  $\boldsymbol{\phi}|_F \cdot \mathbf{n}_{TF} \in \mathbb{P}^k(F)$  for all  $F \in \mathcal{F}_T$  to introduce the projectors in the third line, the definition (21) of  $\delta_T^l$  and  $\delta_{TF}^k$  in the fourth line, and an integration by parts to conclude. Rearranging the terms, we arrive at

$$((\boldsymbol{\delta}_{\nabla, T}^k - \nabla \delta_T^l) \underline{v}_T, \boldsymbol{\phi})_T = \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F. \quad (31)$$

A few remarks are of order to illustrate the consequences of the above relation.

*Remark 11* (Control of the element-based difference through face-based differences). A first notable consequence is that the element-based difference  $(\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T$  can be controlled in terms of the face-based differences  $\{(\delta_{TF}^k - \delta_T^l)\underline{v}_T : F \in \mathcal{F}_T\}$ : For all  $T \in \mathcal{T}_h$  and all  $\underline{v}_T \in \underline{U}_T^{k,l}$ , it holds

$$\|(\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T\|_{L^2(T)^d} \lesssim |\underline{v}_T|_{2,\partial T}, \quad (32)$$

where  $a \lesssim b$  means  $a \leq Cb$  with real number  $C > 0$  independent of  $h$  and of  $T$ , but possibly depending on  $d, \varrho, k$ , and  $l$ . To prove (32), it suffices to observe that

$$\begin{aligned} \|(\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T\|_{L^2(T)^d} &= \sup_{\phi \in \mathbb{P}^k(T)^d, \|\phi\|_{L^2(T)^d}=1} ((\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T, \phi)_T \\ &= \sup_{\phi \in \mathbb{P}^k(T)^d, \|\phi\|_{L^2(T)^d}=1} \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l)\underline{v}_T, \phi \cdot \mathbf{n}_{TF})_F \\ &\leq \sup_{\phi \in \mathbb{P}^k(T)^d, \|\phi\|_{L^2(T)^d}=1} |\underline{v}_T|_{2,\partial T} h_T^{\frac{1}{2}} \|\phi \cdot \mathbf{n}_T\|_{L^2(\partial T)} \lesssim |\underline{v}_T|_{2,\partial T}, \end{aligned}$$

where we have used the fact that  $(\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T \in \mathbb{P}^k(T)^d$  in the first line, (31) in the second line, the Cauchy–Schwarz inequality in the third line, and the discrete trace inequality (71) below with  $p = 2$  to infer  $h_T^{\frac{1}{2}} \|\phi \cdot \mathbf{n}_T\|_{L^2(\partial T)} \lesssim \|\phi\|_{L^2(T)^d}$  and conclude.

*Remark 12* (Stabilisation based on a lifting of face-based differences). Let now  $\underline{v}_T \in \underline{U}_T^{k,l}$  be fixed. Relation (31) no longer holds true in general if we replace  $\phi$  by a function  $\eta$  belonging a space  $\mathfrak{S}_T$  larger than  $\mathbb{P}^k(T)^d$ . It then makes sense to define the nontrivial residual linear form  $\mathcal{R}_T(\underline{v}_T; \cdot) : \mathfrak{S}_T \rightarrow \mathbb{R}$  such that

$$\mathcal{R}_T(\underline{v}_T; \eta) := -((\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T, \eta)_T + \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l)\underline{v}_T, \eta \cdot \mathbf{n}_{TF})_F.$$

Assume now  $\mathfrak{S}_T$  large enough for the  $L^2(T)$ -norm of the Riesz representation  $\mathbf{L}_T \underline{v}_T \in \mathfrak{S}_T$  of  $\mathcal{R}_T(\underline{v}_T; \cdot)$  to control  $|\underline{v}_T|_{2,\partial T}$  (hence also  $\|(\delta_{\nabla,T}^k - \nabla\delta_T^l)\underline{v}_T\|_{L^2(T)^d}$  by (32)). Property (S1) is then fulfilled letting the stabilising contribution in (24) be such that  $\mathbf{S}_T \underline{v}_T = \mathbf{L}_T \underline{v}_T$  for all  $\underline{v}_T$ . This choice also satisfies (S2) by construction since  $\mathbb{P}^k(T)^d \subset \mathfrak{S}_T$  and  $\mathcal{R}_T(\underline{v}_T; \cdot)$  vanishes on  $\mathbb{P}^k(T)^d$  for any  $\underline{v}_T \in \underline{U}_T^{k,l}$  owing to (31). Finally, property (S3) is satisfied provided that  $\mathfrak{S}_T$  is a piecewise polynomial space on a regular polytopal partition of  $T$ .

The above procedure can be interpreted as a lifting on  $\mathfrak{S}_T$  of the face-based differences  $\{(\delta_{TF}^k - \delta_T^l)\underline{v}_T : F \in \mathcal{F}_T\}$  realised by means of the operator  $\mathbf{L}_T$ . This interpretation justifies the terminology employed in Section 3.6.3 below.

### 3.6.2 A Raviart–Thomas–Nédélec subspace

In this section we define a good candidate to play the role of the space  $\mathfrak{S}_T$  in Remark 12. From this point on, we work on a fixed mesh element  $T \in \mathcal{T}_h$  and assume, for the sake of simplicity, that (i) the faces of  $T$  are  $(d-1)$ -simplices and that (ii)  $T$  is star-shaped with respect to a point  $\mathbf{x}_T$  whose orthogonal distance  $d_{TF}$  from each face  $F \in \mathcal{F}_T$  satisfies

$$d_{TF} \geq \varrho h_T. \quad (33)$$

These assumptions can be relaxed using a simplicial submesh of  $T$  and at the price of a heavier notation. For all  $F \in \mathcal{F}_T$ , we denote by  $P_{TF}$  the  $d$ -simplex of base  $F$  and apex  $\mathbf{x}_T$ ,

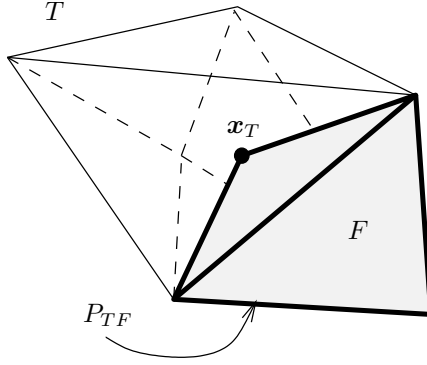


Figure 1: Illustration of  $P_{TF}$ .

and by  $\mathcal{F}_{TF}$  the set of  $(d-1)$ -simplicial faces of  $P_{TF}$ , see Figure 1. In what follows, we work on the face-based simplicial partition  $\mathcal{P}_T := \{P_{TF} : F \in \mathcal{F}_T\}$ .

For an integer  $m \geq 0$  and a face  $F \in \mathcal{F}_T$ , we let  $\mathbb{RT}^m(P_{TF}) := \mathbb{P}^m(P_{TF})^d + \mathbf{x}\mathbb{P}^m(P_{TF})$  denote the Raviart–Thomas–Nédélec space [41, 40] of degree  $m$  on the simplex  $P_{TF}$ . Each function  $\boldsymbol{\eta} \in \mathbb{RT}^m(P_{TF})$  is uniquely identified by the following degrees of freedom (see, e.g., [6, Proposition 2.3.4]):

$$\begin{aligned} (\boldsymbol{\eta} \cdot \mathbf{n}_\sigma, q)_\sigma & \quad \forall \sigma \in \mathcal{F}_{TF}, \forall q \in \mathbb{P}^m(\sigma), \\ (\boldsymbol{\eta}, \boldsymbol{\chi})_{P_{TF}} & \quad \forall \boldsymbol{\chi} \in \mathbb{P}^{m-1}(P_{TF})^d, \end{aligned}$$

where, for all  $\sigma \in \mathcal{F}_{TF}$ , the normal  $\mathbf{n}_\sigma$  points out of  $P_{TF}$ . Additionally, we note the following relation, valid for all  $\boldsymbol{\eta} \in \mathbb{RT}^m(P_{TF})$ :

$$\|\boldsymbol{\eta}\|_{L^2(P_{TF})^d}^2 \simeq \|\boldsymbol{\pi}_{P_{TF}}^{0,m-1} \boldsymbol{\eta}\|_{L^2(P_{TF})^d}^2 + \sum_{\sigma \in \mathcal{F}_{TF}} h_F \|\boldsymbol{\eta} \cdot \mathbf{n}_\sigma\|_{L^2(\sigma)}^2, \quad (34)$$

where  $\simeq$  means  $Ca \leq b \leq C^{-1}a$  with real number  $C > 0$  independent of  $h$  and of  $P_{TF}$ , but possibly depending on  $\varrho$  and  $m$ .

The candidate to play the role of the space  $\mathfrak{S}_T$  in Remark 12 is  $\mathbb{RT}^{k+1}(\mathcal{P}_T)$ , the broken Raviart–Thomas–Nédélec space of degree  $(k+1)$  on the submesh  $\mathcal{P}_T$ .

### 3.6.3 Lifting of face-based differences

We are now ready to construct the lifting of face-based differences. Owing to the specific choice of  $\mathfrak{S}_T$ , we can proceed face by face. Specifically, for all  $F \in \mathcal{F}_T$ , we define the lifting operator  $\mathbf{L}_{TF}^{k+1} : \underline{U}_T^{k,l} \rightarrow \mathbb{RT}^{k+1}(P_{TF})$  such that, for all  $\underline{v}_T \in \underline{U}_T^{k,l}$ ,  $\mathbf{L}_{TF}^{k+1} \underline{v}_T$  satisfies for all  $\boldsymbol{\eta} \in \mathbb{RT}^{k+1}(P_{TF})$

$$(\mathbf{L}_{TF}^{k+1} \underline{v}_T, \boldsymbol{\eta})_{P_{TF}} = -((\delta_{\nabla, T}^k - \nabla \delta_T^l) \underline{v}_T, \boldsymbol{\eta})_{P_{TF}} + ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F. \quad (35)$$

In what follows, we extend by zero  $\mathbf{L}_{TF}^{k+1} \underline{v}_T$  outside  $P_{TF}$ .

**Proposition 13** (Stabilisation based on a Raviart–Thomas–Nédélec subspace). *The following stabilising contribution satisfies properties (S1)–(S3):*

$$\mathbf{S}_T := \sum_{F \in \mathcal{F}_T} \mathbf{L}_{TF}^{k+1}. \quad (36)$$



*Proof.* (i) *Proof of (S1).* We abridge by  $a \lesssim b$  the inequality  $a \leq Cb$  with real number  $C$  independent of both  $h$  and  $T$ , but possibly depending on  $d$ ,  $\varrho$ ,  $k$ , and  $l$ . We start by proving that

$$|\underline{v}_T|_{2,\partial T} \lesssim \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}. \quad (37)$$

Let  $\boldsymbol{\eta} \in \mathbb{RT}^{k+1}(P_{TF})$  be such that

$$\begin{aligned} (\boldsymbol{\eta} \cdot \mathbf{n}_{TF}, q)_F &= h_F^{-1} ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, q)_F & \forall q \in \mathbb{P}^{k+1}(F), \\ (\boldsymbol{\eta} \cdot \mathbf{n}_\sigma, q)_\sigma &= 0 & \forall \sigma \in \mathcal{F}_{TF} \setminus \{F\}, \forall q \in \mathbb{P}^{k+1}(\sigma), \\ (\boldsymbol{\eta}, \boldsymbol{\chi})_{P_{TF}} &= 0 & \forall \boldsymbol{\chi} \in \mathbb{P}^k(P_{TF})^d. \end{aligned} \quad (38)$$

Plugging this definition into (35) and using the Cauchy–Schwarz inequality, we infer that

$$\begin{aligned} h_F^{-1} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^2(F)}^2 &\leq \|\mathbf{L}_{TF}^{k+1} \underline{v}_T\|_{L^2(P_{TF})^d} \|\boldsymbol{\eta}\|_{L^2(P_{TF})^d} \\ &\lesssim \|\mathbf{L}_{TF}^{k+1} \underline{v}_T\|_{L^2(P_{TF})^d} h_F^{-\frac{1}{2}} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^2(F)}, \end{aligned}$$

where we have used (34) and (38) with  $q = \boldsymbol{\eta} \cdot \mathbf{n}_{TF} \in \mathbb{P}^{k+1}(F)$  or  $q = \boldsymbol{\eta} \cdot \mathbf{n}_\sigma \in \mathbb{P}^{k+1}(\sigma)$  to estimate the  $L^2$ -norm of  $\boldsymbol{\eta}$ . Dividing by  $h_F^{-\frac{1}{2}} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^2(F)}$ , squaring, summing over  $F \in \mathcal{F}_T$ , and taking the square root of the resulting inequality proves (37).

Let us now prove that

$$\|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d} \lesssim |\underline{v}_T|_{2,\partial T}. \quad (39)$$

Letting in (35)  $\boldsymbol{\eta} = \mathbf{L}_{TF}^{k+1} \underline{v}_T$ , summing over  $F \in \mathcal{F}_T$ , and using multiple times the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}^2 &= \sum_{F \in \mathcal{F}_T} \|\mathbf{L}_{TF}^{k+1} \underline{v}_T\|_{L^2(P_{TF})^d}^2 \\ &\leq \left( \|(\delta_{\nabla,T}^k - \nabla \delta_T^l) \underline{v}_T\|_{L^2(T)^d}^2 + \sum_{F \in \mathcal{F}_T} h_F^{-1} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^2(F)}^2 \right)^{\frac{1}{2}} \\ &\quad \times \left( \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}^2 + h_T \|\mathbf{S}_T \underline{v}_T \cdot \mathbf{n}_T\|_{L^2(\partial T)}^2 \right)^{\frac{1}{2}} \\ &\lesssim |\underline{v}_T|_{2,\partial T} \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}, \end{aligned}$$

where we have used (32) and the discrete trace inequality (71) below with  $p = 2$  to conclude. Combining (37) and (39) gives (S1).

(ii) *Proof of (S2).* Let  $\boldsymbol{\phi} \in \mathbb{P}^k(T)^d$ , set  $\boldsymbol{\eta} = \boldsymbol{\phi}|_{P_{TF}} \in \mathbb{P}^k(P_{TF})^d \subset \mathbb{RT}^{k+1}(P_{TF})$  in (35), and sum over  $F \in \mathcal{F}_T$  to obtain

$$\begin{aligned} (\mathbf{S}_T \underline{v}_T, \boldsymbol{\phi})_T &= \sum_{F \in \mathcal{F}_T} (\mathbf{L}_{TF}^{k+1} \underline{v}_T, \boldsymbol{\phi})_{P_{TF}} \\ &= -((\delta_{\nabla,T}^k - \nabla \delta_T^l) \underline{v}_T, \boldsymbol{\phi})_T + \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, \boldsymbol{\phi} \cdot \mathbf{n}_{TF})_F = 0, \end{aligned}$$

where we have used (31) to conclude.

(iii) *Proof of (S3).* The regularity of the face-based partition  $\mathcal{P}_T$  follows from (33) in view of [23, Lemma 3]. Property (S3) is then satisfied with  $k_S := k + 2$ .  $\square$

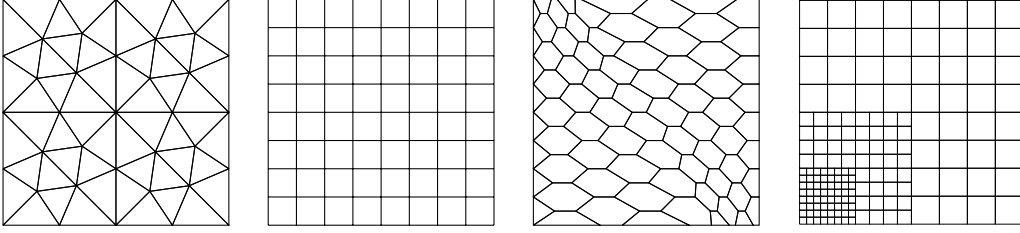


Figure 2: Triangular, Cartesian, hexagonal and locally refined meshes for the numerical examples of Section 3.7

*Remark 14* (The lowest-order case). When considering  $k = 0$  and  $l \in \{-1, 0\}$ , the first term in the right-hand side of (35) vanishes. As a matter of fact, (i)  $\delta_{\nabla, T}^k \underline{v}_T = \mathbf{0}$  since  $\mathbb{P}^0(T)^d = \nabla \mathbb{P}^1(T)$ , so that  $\mathbf{G}_T^0 = \nabla \mathbf{r}_T^1$ ; (ii)  $\delta_T^{-1} = \delta_T^0 = 0$  by (20b). A consequence of this remark is that, for all  $F \in \mathcal{F}_T$ , it suffices to lift the boundary term in the right-hand side of (35) on  $\mathbb{P}^0(P_{TF})^d$  instead of  $\mathbb{RT}^1(P_{TF})$ , since  $\boldsymbol{\eta}|_F \cdot \mathbf{n}_{TF} \in \mathbb{P}^0(F)$  whenever  $\boldsymbol{\eta} \in \mathbb{RT}^1(P_{TF})$ . With this choice, for all  $F \in \mathcal{F}_T$  we have the following explicit formula for the face-based lifting  $\mathbf{L}_{TF}^{\text{hmm}} : \underline{U}_T^{0,l} \rightarrow \mathbb{P}^0(P_{TF})^d$ : For all  $\underline{v}_T \in \underline{U}_T^{0,l}$ ,

$$\mathbf{L}_{TF}^{\text{hmm}} \underline{v}_T = \frac{|F|}{|P_{TF}|} \delta_{TF}^0 \underline{v}_T \mathbf{n}_{TF} = \frac{d}{d_{TF}} \delta_{TF}^0 \underline{v}_T \mathbf{n}_{TF}.$$

This lifting correspond to one of the stabilisations used for the Hybrid-Mixed-Mimetic methods of [28], obtained taking  $k = l = 0$ .

### 3.7 Numerical examples

In this section we provide numerical evidence to support our theoretical results.

#### 3.7.1 Trigonometric solution for $p \geq 2$

We solve inside the two-dimensional unit square  $\Omega = (0, 1)^2$  the  $p$ -Laplace problem (6) corresponding to the exact solution

$$u(\mathbf{x}) = \sin(\pi x_1) \sin(\pi x_2), \quad (40)$$

with  $p \in \{2, 3, 4\}$  and source term inferred from  $u$ . We consider the matching triangular, Cartesian, (predominantly) hexagonal, and locally refined mesh families depicted in Figure 2 and polynomial degrees ranging from 0 to 4. The first, second, and fourth mesh families originate from the FVCA5 benchmark [37], whereas the third from [23]. The local refinement in the third mesh family has no specific meaning here: its purpose is to demonstrate the seamless treatment of nonconforming interfaces.

We report in Figure 3 the error  $\|\nabla_{\mathcal{D}_h} (\underline{L}_h^{k,l} u - \underline{u}_h)\|_{L^p(\Omega)^d}$  versus the meshsize  $h$ . The reference slopes correspond to the convergence rates resulting from Theorem 2 together with Proposition 9 (more precisely, the order corresponds to the dominating term). The theoretical orders of convergence are perfectly matched for  $p = 2$ . Similar considerations hold for  $p = 3$  and  $k < 2$  whereas, for  $k \geq 2$ , the order of convergence is limited by the regularity of the function  $\boldsymbol{\psi} \mapsto |\boldsymbol{\psi}|^{p-2} \boldsymbol{\psi}$ , which impacts on the regularity of  $|\nabla u|^{p-2} \nabla u$ . Finally, for  $p = 4$

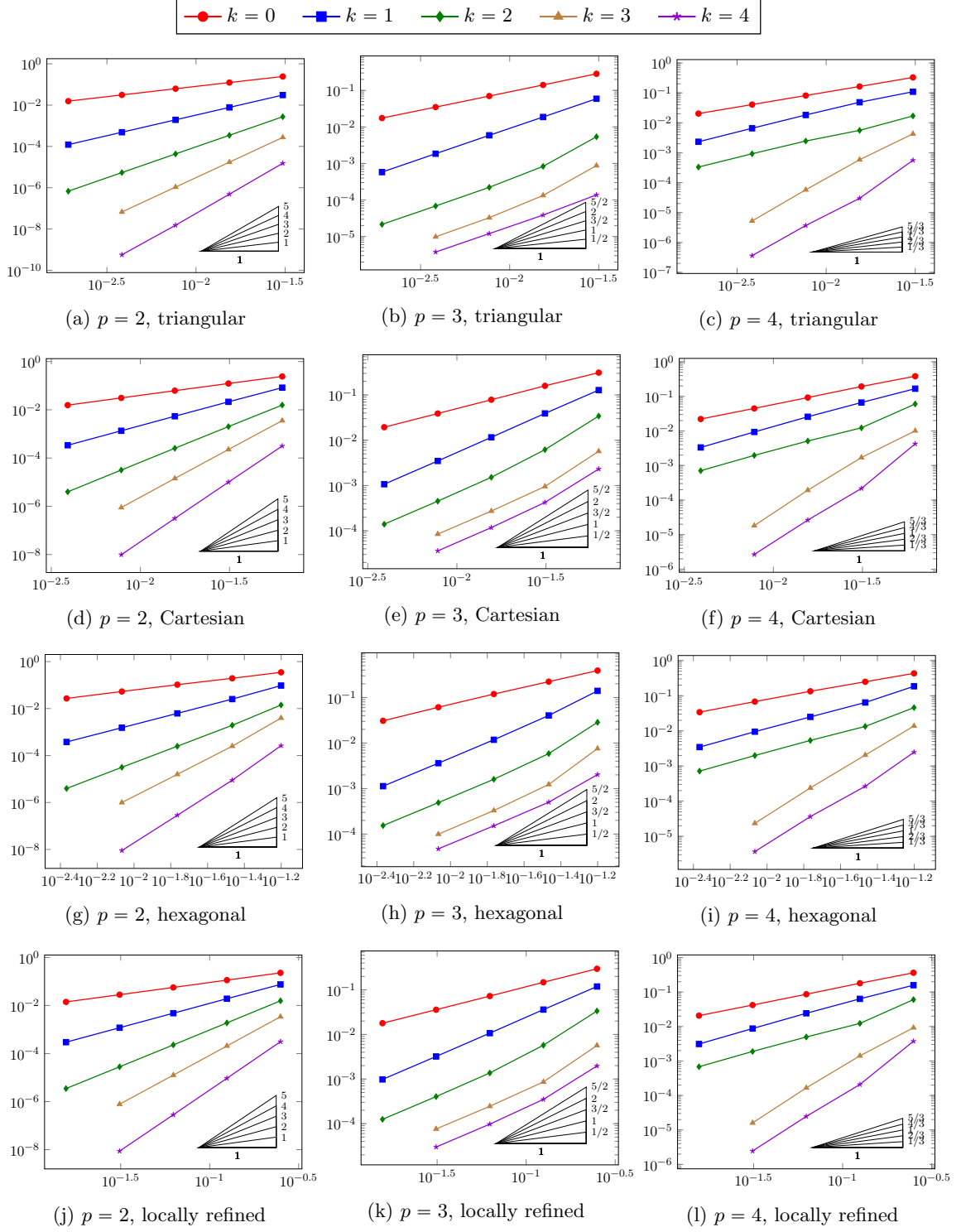


Figure 3:  $\|\nabla_{\mathcal{D}_h}(I_h^{k,l}u - \underline{u}_h)\|_{L^p(\Omega)^d}$  v.  $h$ . Trigonometric test case,  $p \in \{2, 3, 4\}$ , DSGD.

the theoretical orders of convergence are matched for  $k \in \{0, 1, 2\}$ , while faster convergence than predicted by the error estimates is observed for  $k \in \{3, 4\}$ . This phenomenon will be further investigated in the future. For a comparison with the HHO method of [16], see Figure 5 below.

### 3.7.2 Exponential solution for $p < 2$

As pointed out in [17], the trigonometric solution (40) does not have the required regularity to assess the convergence order of the DSGD method when  $1 < p < 2$ . For this reason, we consider instead the exponential solution

$$u(\mathbf{x}) = \exp(x_1 + \pi x_2),$$

and solve the  $p$ -Laplace problem with  $p = \frac{3}{4}$  with Dirichlet boundary conditions on  $\partial\Omega$  and right-hand side  $f$  inferred from the expression of  $u$ . The error  $\|\nabla_{\mathcal{D}_h}(\underline{L}_h^{k,l}u - \underline{u}_h)\|_{L^p(\Omega)^d}$  versus the meshsize  $h$  is plotted in Figure 4 for the mesh families illustrated in Figure 2 and polynomial degrees ranging from 0 to 4. For the sake of completeness, a comparison with the HHO method (47) is also included. We observe that both methods yield similar orders of convergence, but the corresponding errors differ in some cases by more than one order of magnitude. This difference is mainly related to the stabilisation contribution: when the error is measured using only the consistent part of the gradient, the results are much closer.

## 4 Alternative gradient and links with other methods

In this section we discuss an alternative to the gradient reconstruction  $\mathcal{G}_T$  defined by (24), and links with other methods.

### 4.1 A consistent gradient based on $\mathbb{r}_T^{k+1}$

An alternative to using  $\mathbf{G}_T^k$  in (24) would be to use  $\nabla \mathbb{r}_T^{k+1}$ . This would lead to a local gradient defined by

$$(\tilde{\nabla}_{\mathcal{D}_h} \underline{v}_h)|_T = \tilde{\mathcal{G}}_T \underline{v}_T := \nabla \mathbb{r}_T^{k+1} \underline{v}_T + \tilde{\mathbf{S}}_T \underline{v}_T \quad \forall \underline{v}_h \in \underline{U}_h^{k,l}, \forall T \in \mathcal{T}_h. \quad (41)$$

The stabilisation term  $\tilde{\mathbf{S}}_T$  is still required to satisfy the  $L^2$ -stability and boundedness property (S1) and the image property (S3). As shown in (78) below, the property (S2) is required on  $\mathbf{S}_T \underline{v}_T$  to ensure that it is orthogonal to the consistent part  $\mathbf{G}_T^k \underline{v}_T$  of  $\mathcal{G}_T \underline{v}_T$ . For  $\tilde{\mathcal{G}}_T \underline{v}_T$ , the consistent part is  $\nabla \mathbb{r}_T^{k+1} \underline{v}_T \in \nabla \mathbb{P}^{k+1}(T)$ , and the orthogonality property on  $\tilde{\mathbf{S}}_T \underline{v}_T$  can therefore be relaxed into

$$(\tilde{\mathbf{S}}_2) \text{ For all } \underline{v}_T \in \underline{U}_T^{k,l} \text{ and all } \phi \in \nabla \mathbb{P}^{k+1}(T), (\tilde{\mathbf{S}}_T \underline{v}_T, \phi)_T = 0,$$

where, with respect to (S2), the space for  $\phi$  is  $\nabla \mathbb{P}^{k+1}(T)$  instead of  $\mathbb{P}^k(T)^d$ .

Performing integration-by-parts and introducing projection operators in (20a) leads to

$$-(\nabla \delta_T^l \underline{v}_T, \nabla w)_T = \sum_{F \in \mathcal{F}_T} ((\delta_{TF}^k - \delta_T^l) \underline{v}_T, \nabla w \cdot \mathbf{n}_{TF})_T, \quad \forall w \in \mathbb{P}^{k+1}(T). \quad (42)$$

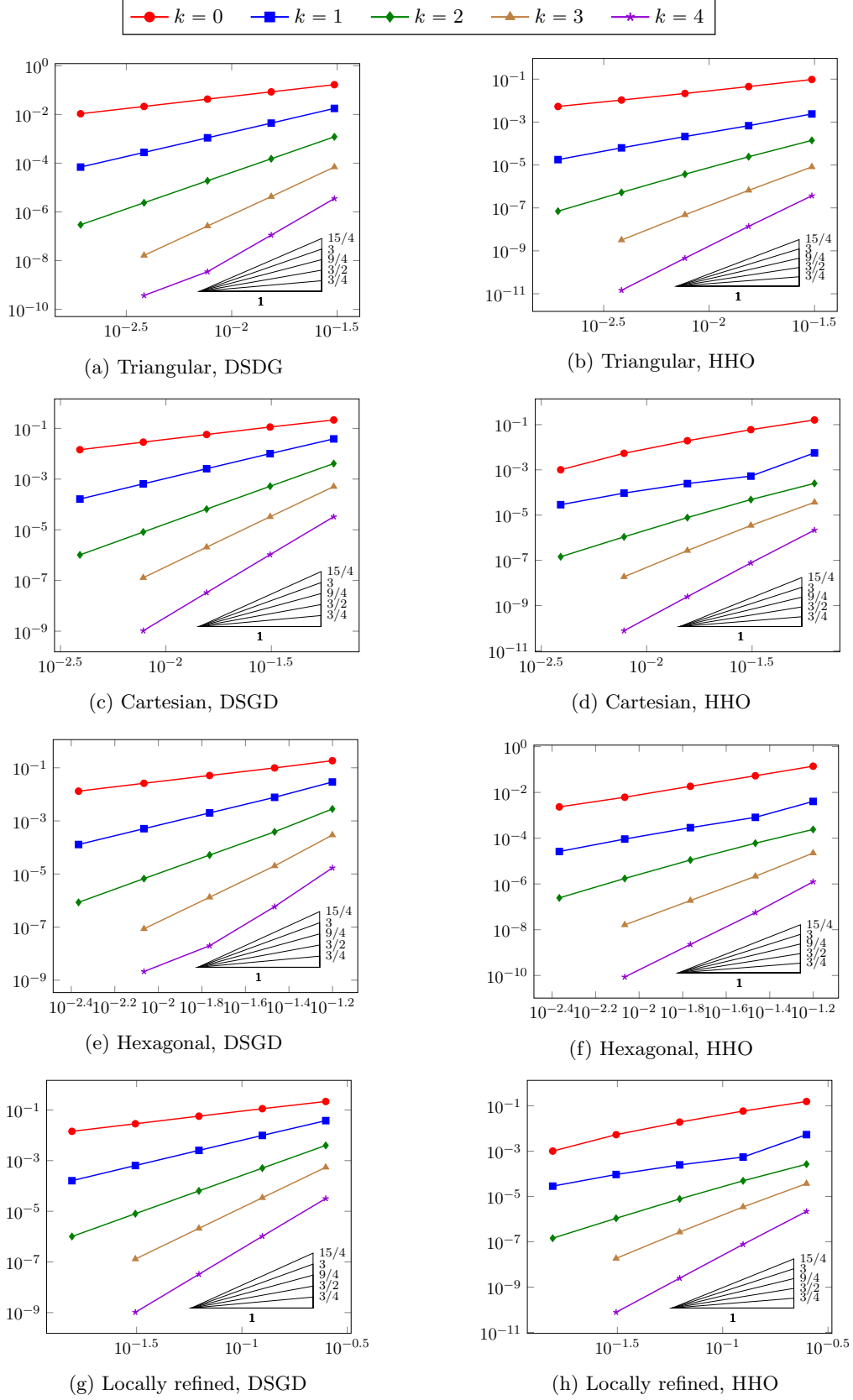


Figure 4: Error v.  $h$ , exponential test case,  $p = \frac{3}{4}$ . For the DSGD method (left) the error is measured by  $\|\nabla_{\mathcal{D}_h}(I_h^{k,l}u - \underline{u}_h)\|_{L^p(\Omega)^d}$ , for the HHO method (right) by  $\|I_h^{k,l}u - \underline{u}_h\|_{1,p,h}$ .

Comparing with (31), we see that the volumetric term  $\delta_{\nabla, T}^k v_T$  has disappeared, and the only remaining volumetric term  $\nabla \delta_T^l v_T$  belongs to  $\nabla \mathbb{P}^l(T)$ . This is a gain with respect to the situation in Section 3.6, in which the degree of the volumic term was constrained by  $\delta_{\nabla, T}^k v_T$  to be  $k$ , whatever the choice of  $l$ . As a consequence, the construction of  $\tilde{\mathbf{S}}_T$  can be done in a (possibly) smaller space than  $\mathbb{RT}^{k+1}(P_{TF})$ , as detailed in what follows.

First, considering  $w = -\delta_T^l v_T$  in (42) and using the trace inequality (71) with  $p = 2$  on  $\nabla \delta_T^l v_T$  yields (compare with Remark 11)

$$\|\nabla \delta_T^l v_T\|_{L^2(T)^d} \lesssim |v_T|_{2, \partial T}. \quad (43)$$

Then, following the ideas of Section 3.6.3, we construct the lifting  $\tilde{\mathbf{L}}_{TF}^k : U_T^{k,l} \rightarrow \mathbb{RT}^{\max(l,k)}(P_{TF})$  such that, for all  $\boldsymbol{\eta} \in \mathbb{RT}^{\max(l,k)}(P_{TF})$ ,

$$(\tilde{\mathbf{L}}_{TF}^k v_T, \boldsymbol{\eta})_{P_{TF}} = (\nabla \delta_T^l v_T, \boldsymbol{\eta})_{P_{TF}} + ((\delta_{TF}^k - \delta_T^l) v_T, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F. \quad (44)$$

Since  $\nabla \delta_T^l v_T \in \nabla \mathbb{P}^l(T) \subset \mathbb{P}^{l-1}(T)^d$  and  $(\delta_{TF}^k - \delta_T^l) v_T \in \mathbb{P}^{\max(l,k)}(F)$ , following the proof of (S1) in Proposition 13 shows that, with this choice of  $\tilde{\mathbf{L}}_{TF}^k$ , we only need (38) to hold for  $q \in \mathbb{P}^{\max(l,k)}(F)$  and  $\boldsymbol{\chi} \in \mathbb{P}^{l-1}(T)^d$ . The space  $\mathbb{RT}^{\max(l,k)}(P_{TF})$  enables these choices of  $q$  and  $\boldsymbol{\chi}$ , and the  $L^2$ -stability of  $\tilde{\mathbf{S}}_T := \sum_{F \in \mathcal{F}_T} \tilde{\mathbf{L}}_{TF}^k$  therefore follows. When  $l \in \{k-1, k\}$ , this stabilisation term is constructed on a piecewise  $\mathbb{RT}^k$  space, instead of a piecewise  $\mathbb{RT}^{k+1}$  space for  $\mathbf{S}_T$  in Section 3.6.3.

Although the choice (41) leads to coercive, consistent, limit-conforming, and compact families of gradient discretisations, it can turn out to be far from optimal for general problems. More precisely, its limit-conformity properties are much worse than those of (24). Indeed, the full orthogonality property (S2) is essential to establish the  $\mathcal{O}(h^{k+1})$  estimate (30) on  $W_{\mathcal{D}_h}$  (see Remark 23). In general, we cannot establish more than an  $\mathcal{O}(h)$  estimate on  $W_{\mathcal{D}_h}$ , irrespective of  $k$ , if the gradient is reconstructed via (41). For anisotropic linear problems, a modification of  $r_T^{k+1}$  embedding a dependence on the diffusion coefficient can be constructed to recover optimal rates of convergence [21]; for fully non-linear models, though, the only option to recover a truly high-order method seems to be using the gradient  $\mathbf{G}_T^k v_T$  in the full polynomial space  $\mathbb{P}^k(T)^d$ .

To illustrate numerically the loss of convergence experienced when using the gradient reconstruction (41) in the context of fully non-linear problems, we solve the problem described in Section 3.7.1 using two numerical methods: the HHO method of [16] (see (47) below), and the method obtained from the latter replacing  $\mathbf{G}_T^k$  by  $\nabla r_T^{k+1}$ . We report in Figure 5 the error  $\|\underline{I}_h^{k,l} u - u_h\|_{1,p,h}$  versus the meshsize  $h$ , with reference slopes corresponding to the estimates of convergence rates derived in [17, Theorem 3.2]. The leftmost column, corresponding to the Poisson problem with  $p = 2$ , shows that both  $\mathbf{G}_T^k$  and  $\nabla r_T^{k+1}$  can be used when  $W_{\mathcal{D}_h}$  is applied to  $\boldsymbol{\psi} = \nabla u$ . For  $p \in \{3, 4\}$ , on the other hand, a significant loss in the convergence rate is observed for  $k > 1$  (the dashed line corresponding to  $\nabla r_T^{k+1}$  departs from the solid line corresponding to  $\mathbf{G}_T^k$ ). Notice that, for  $k \geq 3$  and  $p = 3$ , the order of convergence is again limited by the regularity of the function  $\boldsymbol{\psi} \mapsto |\boldsymbol{\psi}|^{p-2} \boldsymbol{\psi}$ . The results presented here replace the ones of [17, Figure 3], which were affected by a bug in one of the libraries used in our code. From these new tests, the error estimates of [17, Theorem 3.2] appear to be sharp also for  $p > 2$ .

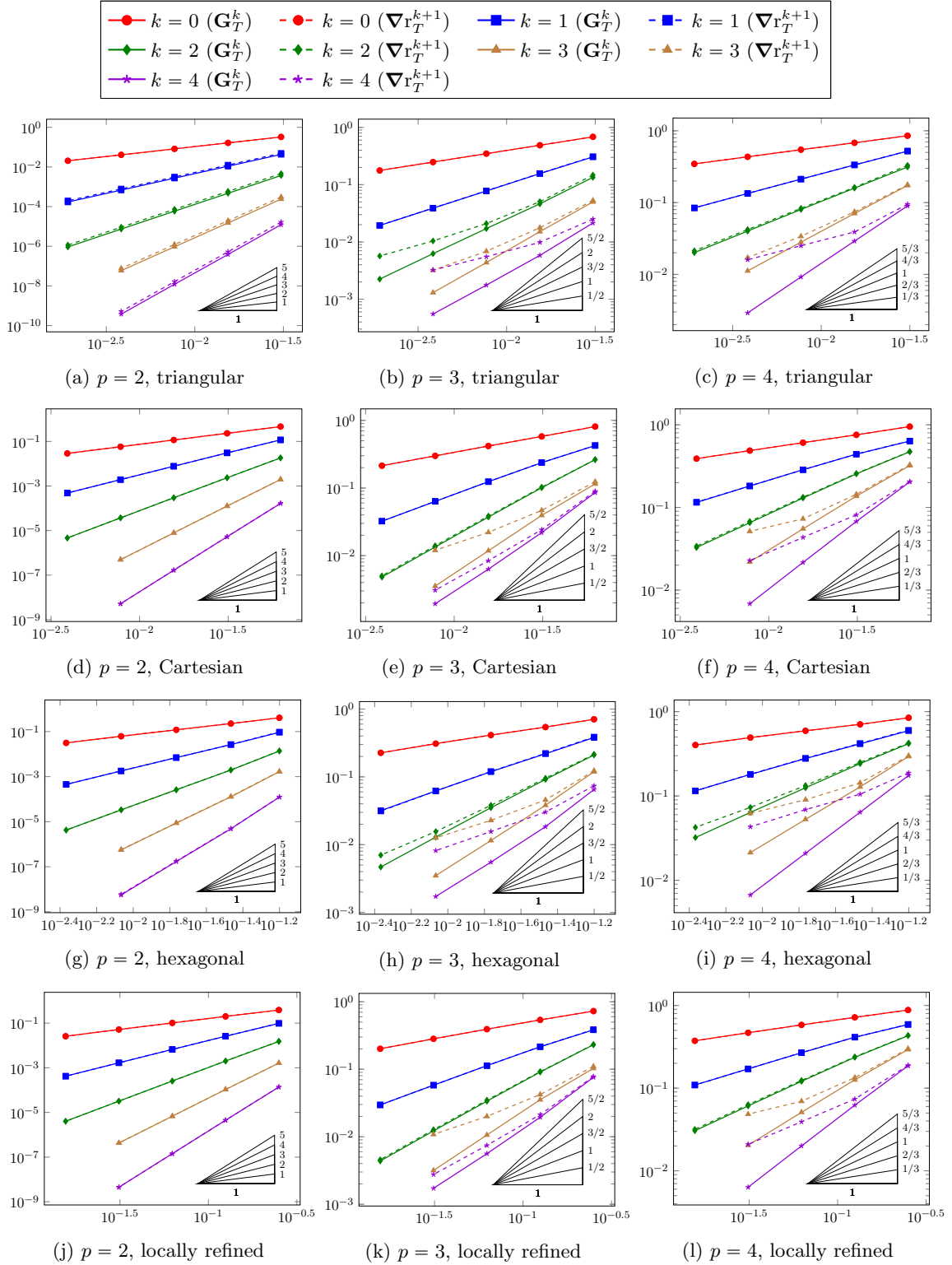


Figure 5:  $\|L_h^{k,l} u - \underline{u}_h\|_{1,p,h}$  v.  $h$ . Trigonometric test case,  $p \in \{2, 3, 4\}$ , HHO.



## 4.2 Hybrid High-Order methods

The HHO method proposed in [22] for problem (2) with  $\mathbf{\Lambda} = \mathbf{I}_d$  reads

$$\text{Find } \underline{u}_h \in \underline{U}_{h,0}^{k,l} \text{ such that, for all } \underline{v}_h \in \underline{U}_{h,0}^{k,l}, a_h^{\text{hho}}(\underline{u}_h, \underline{v}_h) = (f, v_h), \quad (45)$$

where the broken polynomial function  $v_h$  is defined by (15), and the bilinear form  $a_h^{\text{hho}} : \underline{U}_h^{k,l} \times \underline{U}_h^{k,l} \rightarrow \mathbb{R}$  is assembled from the elementary contributions

$$a_T^{\text{hho}}(\underline{u}_T, \underline{v}_T) := (\nabla \mathbf{r}_T^{k+1} \underline{u}_T, \nabla \mathbf{r}_T^{k+1} \underline{v}_T)_T + \sum_{F \in \mathcal{F}_T} h_F^{-1} ((\delta_{TF}^k - \delta_T^l) \underline{u}_T, (\delta_{TF}^k - \delta_T^l) \underline{v}_T)_F. \quad (46)$$

Here, the consistent gradient is  $\nabla \mathbf{r}_T^{k+1}$ , as in (41), and the stabilisation is not incorporated in the gradient reconstruction, but rather added as a separate term in the bilinear form.

For the non-linear problem (6), on the other hand, the HHO method considered in [16, 17] reads

$$\text{Find } \underline{u}_h \in \underline{U}_{h,0}^{k,l} \text{ such that, for all } \underline{v}_h \in \underline{U}_{h,0}^{k,l}, A_h^{\text{hho}}(\underline{u}_h, \underline{v}_h) = (f, v_h),$$

with function  $A_h^{\text{hho}} : \underline{U}_h^{k,l} \times \underline{U}_h^{k,l} \rightarrow \mathbb{R}$  assembled from the elementary contributions

$$A_T^{\text{hho}}(\underline{u}_T, \underline{v}_T) := \int_T \boldsymbol{\sigma}(\mathbf{G}_T^k \underline{u}_T) \cdot \mathbf{G}_T^k \underline{v}_T + \sum_{F \in \mathcal{F}_T} h_F^{p-1} \int_F |(\delta_{TF}^k - \delta_T^l) \underline{u}_T|^{p-2} (\delta_{TF}^k - \delta_T^l) \underline{u}_T (\delta_{TF}^k - \delta_T^l) \underline{v}_T. \quad (47)$$

Also in this case, stability is achieved by a separate term, and only the consistent (but not stable) gradient reconstruction appears in the consistency term. Notice that, unlike the linear case, the gradient reconstructed in the full polynomial space  $\mathbb{P}^k(T)^d$  is present here; see comments at the end of Section 4.1.

## 4.3 High-order non-conforming Mimetic Finite Difference

The ncMFD method of [39] hinges on degrees of freedom (DOFs) that are the polynomial moments of degree up to  $(k-1)$  inside the mesh elements and the polynomial moments of degree up to  $k$  on the mesh faces.  $X_h$  is the space of vectors gathering such degrees of freedom. Given a function  $v \in H^1(\Omega)$ , we denote by  $v^I$  its interpolation in  $X_h$ , that is, the vector collecting its moments in the elements and on the faces. If  $\mathbf{v} \in X_h$ , we denote by  $v_T \in \mathbb{P}^{k-1}(T)$ ,  $T \in \mathcal{T}_h$ , and  $v_F \in \mathbb{P}^k(F)$ ,  $F \in \mathcal{F}_h$ , the polynomials reconstructed from the moments represented by  $\mathbf{v}$ . This defines an isomorphism

$$\mathbf{v} \in X_h \mapsto \underline{v}_h = ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}) \in \underline{U}_h^{k,k-1}. \quad (48)$$

Denoting by  $\mathbf{v}_{|T}$  the sub-vector made of the DOFs of  $\mathbf{v} \in X_h$  in the mesh element  $T \in \mathcal{T}_h$  and on the mesh faces in  $\mathcal{F}_T$ , the ncMFD method for problem (2) with  $\mathbf{\Lambda} = \mathbf{I}_d$  reads

$$\text{Find } \mathbf{u} \in X_h \text{ such that, for all } \mathbf{v} \in X_h, \sum_{T \in \mathcal{T}_h} \mathbf{u}_{|T}^t \mathbf{M}_T \mathbf{v}_{|T} = \sum_{T \in \mathcal{T}_h} L_f(\mathbf{v}_{|T}), \quad (49)$$

where  $L_f(\mathbf{v}_{|T})$  is a discretisation of  $\int_T f v$ , and the matrix  $\mathbf{M}_T$  is positive semi-definite with suitable consistency and stability properties. Setting  $N_{d,k} = \dim(\mathbb{P}^{k-1}(T))$  and selecting a

basis  $(q_i)_{i=0,\dots,N_{d,k}-1}$  of  $\mathbb{P}^{k-1}(T)$  with  $q_0 = 1$ , the required consistency and stability properties on  $M_T$  lead to the following decomposition (see [39, Eq. (35)]):

$$M_T = M_T^0 + M_T^1 = \widehat{R}_T(\widehat{R}_T^t \widehat{N})^{-1} \widehat{R}_T^t + M_T^1$$

where  $A^t$  is the transpose of  $A$ ,  $\ker(M_T^1) = \{(v^I)|_T : v \in \mathbb{P}^{k-1}(T)\}$ ,  $\widehat{N}$  has columns  $\widehat{N}_i = (q_i)^I$  for  $i = 1, \dots, N_{d,k} - 1$ , and  $\widehat{R}_T$  is the matrix with columns  $(\widehat{R}_{T,i})_{i=1,\dots,N_{d,k}}$  defined by

$$\forall v \in X_h, (\widehat{R}_{T,i})^t v|_T = -(v_T, \Delta q_i) + \sum_{F \in \mathcal{F}_T} (v_F, \nabla q_i \cdot \mathbf{n}_{TF})_F. \quad (50)$$

The stabilising matrix  $M_T^1$  does not have any impact on the consistency of the method; its sole role is to stabilise the matrix  $M$  so that its kernel is  $\mathbb{R}1^I$  (which is expected: a matrix  $M_T$  representing the bilinear form  $\int_T \nabla u \cdot \nabla v$  should vanish on interpolants of constant functions). The matrix  $M_T^0$ , on the other hand, contains all the consistency properties of the method. The analysis of the stabilisation part is made in Section 4.5, alongside the analysis of the stabilisation in the HHO and ncVEM methods.

Let us analyse here the consistent part  $M_T^0$  of  $M_T$ . Take  $v \in X_h$ , and let  $v_h \in U_h^{k,k-1}$  corresponding to  $v$  through the isomorphism (48). Comparing (20a) and (50) shows that

$$(\nabla r_T^{k+1} v_T, \nabla q_i)_T = (\widehat{R}_{T,i})^t v|_T = (\widehat{R}_T^t v|_T)_i, \quad \forall i = 1, \dots, N_{d,k} - 1.$$

Hence, with obvious notations,

$$\begin{aligned} v|_T^t M_T^0 w|_T &= (\widehat{R}_T^t v|_T)^t (\widehat{R}_T^t \widehat{N})^{-1} (\widehat{R}_T^t w|_T) \\ &= [(\nabla r_T^{k+1} v_T, \nabla q_i)_T]_{i=1,\dots,N_{d,k}-1}^t (\widehat{R}_T^t \widehat{N})^{-1} [(\nabla r_T^{k+1} w_T, \nabla q_i)_T]_{i=1,\dots,N_{d,k}-1}. \end{aligned} \quad (51)$$

Applying (50) to  $v|_T = (q_j)^I$  and integrating by parts shows that  $(\widehat{R}_T^t \widehat{N})_{ij} = (\nabla q_i, \nabla q_j)_T$ , that is,  $\widehat{R}_T^t \widehat{N}$  is the Gram matrix of  $(\nabla q_i)_{i=1,\dots,N_{d,k}-1}$  in  $L^2(T)^d$ . Equation (51) can therefore be re-written

$$v|_T^t M_T^0 w|_T = (\nabla r_T^{k+1} v_T, \nabla r_T^{k+1} w_T)_T.$$

Thus, recalling the isomorphism (48) between  $X_h$  and  $U_h^{k,k-1}$ , the ncMFD method (49) is equivalent to

$$\begin{aligned} \text{Find } u_h \in U_h^{k,k-1} \text{ such that, for all } v_h \in U_h^{k,k-1}, \\ \sum_{T \in \mathcal{T}_h} (\nabla r_T^{k+1} u_T, \nabla r_T^{k+1} v_T)_T + s_T^{\text{mfd}}(u, v) = \sum_{T \in \mathcal{T}_h} L_f(v|_T), \end{aligned}$$

where  $s_T^{\text{mfd}}$  is the stabilising bilinear form assembled from the local matrices  $(M_T^1)_{T \in \mathcal{T}_h}$ . Here, the consistent part is constructed from  $\nabla r_T^{k+1}$ , as in (41), but the stabilisation is external to the gradient. In the ncMFD method, the loading term is discretised (for  $k \geq 2$ ) by

$$L_f(v|_T) = \int_T \pi_T^{0,k-1}(f) v_T = \int_T f v_T. \quad (52)$$

This discretisation has a lower order than the right-hand side of (3) with  $\Pi_{\mathcal{D}}$  defined by (16), but this does not impact the rates of convergence of ncMFD. A modification of (52) is however necessary for the low orders  $k = 0, 1$  – see [39, Section 2.7].

#### 4.4 Non-conforming Virtual Element Method

For a given mesh element  $T \in \mathcal{T}_h$ , we define the local non-conforming virtual element space as follows [2]:

$$V_k^{h,nc}(T) := \left\{ \mathbf{v}_h \in H^1(T) : \mathbf{n}_{TF} \cdot \nabla \mathbf{v}_h|_F \in \mathbb{P}^k(F) \quad \forall F \in \mathcal{F}_T, \quad \Delta \mathbf{v}_h \in \mathbb{P}^{k-1}(T) \right\}. \quad (53)$$

This space is finite dimensional and its functions are described by their face moments against the polynomials of degree up to  $k$  on each face  $F \in \mathcal{F}_T$  and the cell moments against polynomials of degree up to  $k - 1$  inside cell  $T$ . The unisolvency of these degrees of freedom has been proved in [2].

Let now  $H^1(\mathcal{T}_h)$  be the broken  $H^1$  space as in Proposition 9. The *global non-conforming virtual element space of degree  $k$*  is given by:

$$V_k^{h,nc}(\Omega) := \left\{ \mathbf{v}_h \in L^2(\Omega) \text{ such that } \int_F \llbracket v \rrbracket q = 0 \quad \forall F \in \mathcal{F}_T, \quad \forall q \in \mathbb{P}^k(F) \right. \\ \left. \text{and } \mathbf{v}_h|_T \in V_k^{h,nc}(T) \quad \forall T \in \mathcal{T}_h \right\}, \quad (54)$$

where  $\llbracket v \rrbracket$  denotes the *jump operator* with the usual definition at interfaces (the sign is not relevant), and extended to boundary faces setting  $\llbracket v \rrbracket := v$ .

The polynomials of degree up to  $(k + 1)$  are a subspace of  $V_k^{h,nc}(T)$ . The jump condition on the elements of  $V_k^{h,nc}(\Omega)$  ensure that  $\underline{I}_h^{k,k-1} : V_k^{h,nc}(\Omega) \rightarrow \underline{U}_h^{k,k-1}$  is well defined (there is only one interpolant on each face). Moreover, the unisolvent property of the set of DOFs shows that  $\underline{I}_h^{k,k-1}$  is an isomorphism between  $V_k^{h,nc}(\Omega)$  and  $\underline{U}_h^{k,k-1}$ .

The virtual element discretization of problem (2) with  $\mathbf{\Lambda} = \text{Id}$  reads as:

$$\text{Find } \mathbf{u}_h \in V_{k,0}^{h,nc}(\Omega) \text{ such that, for all } \mathbf{v}_h \in V_{k,0}^{h,nc}(\Omega), a_h(\mathbf{u}_h, \mathbf{v}_h) = (f_h, \mathbf{v}_h) \quad (55)$$

where the bilinear form  $a_h(\mathbf{u}_h, \mathbf{v}_h)$  approximates the left-hand side of (2),  $f_h|_T = \pi_T^{0,k-1} f$ , and  $V_{k,0}^{h,nc}(\Omega)$  is the subspace of  $V_k^{h,nc}(\Omega)$  whose functions are zero on the boundary of  $\Omega$ .

Mimicking the additivity of integrals, we assume that the virtual element bilinear form is the summation of local elemental terms

$$a_h(u_h, v_h) = \sum_{T \in \mathcal{T}_h} a_{h,T}(u_h, v_h). \quad (56)$$

Two different formulations of the local bilinear form  $a_{h,T}$  can be found in the literature, both including a consistency and a stability term:

– *first formulation* [2]: for every  $\mathbf{u}_h, \mathbf{v}_h \in V_k^{h,nc}(T)$ :

$$a_{h,T}(\mathbf{u}_h, \mathbf{v}_h) := (\nabla \pi_T^{1,k+1} \mathbf{u}_h, \nabla \pi_T^{1,k+1} \mathbf{v}_h)_T + \mathcal{S}((\text{Id} - \pi_T^{1,k+1}) \mathbf{u}_h, (\text{Id} - \pi_T^{1,k+1}) \mathbf{v}_h); \quad (57)$$

– *second formulation* [12, 11]: for every  $\mathbf{u}_h, \mathbf{v}_h \in V_k^{h,nc}(T)$ :

$$a_{h,T}(\mathbf{u}_h, \mathbf{v}_h) := (\pi_T^{0,k} \nabla \mathbf{u}_h, \pi_T^{0,k} \nabla \mathbf{v}_h)_T + \mathcal{S}((\text{Id} - \pi_T^{1,k+1}) \mathbf{u}_h, (\text{Id} - \pi_T^{1,k+1}) \mathbf{v}_h). \quad (58)$$

In these definitions, the first term on the right is the consistency term designed to provide the exactness of the integration whenever at the least one of the entries  $\mathbf{u}_h$  or  $\mathbf{v}_h$  is a polynomial of degree up to  $(k+1)$ . The second term is a stabilization, and  $\mathcal{S}$  can be any symmetric and positive definite bilinear form for which there exist two positive constants  $c_*$  and  $c^*$  such that

$$c_* \|\nabla \mathbf{v}_h\|_{L^2(T)^d}^2 \leq \mathcal{S}(\mathbf{v}_h, \mathbf{v}_h) \leq c^* \|\nabla \mathbf{v}_h\|_{L^2(T)^d}^2 \quad \forall \mathbf{v}_h \in V_k^{h,nc}(T) \text{ such that } \pi_T^{1,k+1} \mathbf{v}_h = 0. \quad (59)$$

*Remark 15.* The connection between formulation (57) and the ncMFD method of [39] reviewed in Section 4.3 has been established in [2].

#### 4.5 HHO, ncMFD and ncVEM are gradient discretisation methods

In the previous sections, we showed that the consistent gradient in the HHO, ncMFD and ncVEM methods (second formulation) is, for the Poisson problem, identical to the consistent gradient in (41). We show here that the stabilisations used in these methods can actually be represented by well-chosen stabilisation terms  $\mathbf{S}_T$  satisfying (S1)–(S3), and thus that these methods are gradient discretisation methods.

Following the discussion in the previous sections, for the Poisson problem the HHO, ncMFD and ncVEM (second formulation) can be written, upon an isomorphism of the space of discrete unknowns and with the proper choice of  $l \in \{k-1, k, k+1\}$ : Find  $\underline{u}_h \in \underline{U}_{h,0}^{k,l}$  such that, for all  $\underline{v}_h \in \underline{U}_{h,0}^{k,l}$ ,

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} (\nabla \mathbf{r}_T^{k+1} \underline{u}_T, \nabla \mathbf{r}_T^{k+1} \underline{v}_T)_T + \sum_{T \in \mathcal{T}_h} \text{Stab}_T(\underline{u}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{u}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T) \\ = \sum_{T \in \mathcal{T}_h} L_{T,f}(\underline{v}_T), \quad (60) \end{aligned}$$

where  $L_{T,f} : \underline{U}_T^{k,l} \rightarrow \mathbb{R}$  is a specific linear form and  $\text{Stab}_T$  is a symmetric bilinear form on  $\underline{U}_T^{k,l}$  that is coercive and stable on  $\text{Im}(\text{Id} - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1})$ , that is,

$$\forall \underline{v}_T \in \underline{U}_T^{k,l}, \quad \text{Stab}_T(\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T) \simeq |\underline{v}_T|_{2,\partial T}^2 \quad (61)$$

where  $a \simeq b$  means  $Ca \leq b \leq C^{-1}a$  with real number  $C > 0$  independent of  $h$  and of  $T$ .

*Remark 16* (Stabilisation term). For the HHO scheme (46), by (22) the vector  $((\delta_{TF}^k - \delta_T^l) \underline{v}_T)_{F \in \mathcal{F}_T}$  is the difference of the face- and cell-unknowns of  $\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T$ . Hence, the definition

$$\text{Stab}_T(\underline{u}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{u}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T) := \sum_{F \in \mathcal{F}_T} h_F^{-1} ((\delta_{TF}^k - \delta_T^l) \underline{u}_T, (\delta_{TF}^k - \delta_T^l) \underline{v}_T)_F$$

is valid and ensures that the stabilisation terms in (46) and (60) coincide. This also shows that  $\text{Stab}_T(\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T) = |\underline{v}_T|_{2,\partial T}^2$ , and thus that (61) holds.

For the ncVEM version (any formulation), comparing the stabilisations in (58)–(57) and in (60) leads to defining, for all  $\mathbf{u}_h, \mathbf{v}_h \in V_k^{h,nc}(T)$  and setting  $\underline{u}_T = \underline{I}_T^{k,l} \mathbf{u}_h$  and  $\underline{v}_T = \underline{I}_T^{k,l} \mathbf{v}_h$ ,

$$\text{Stab}_T(\underline{u}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{u}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T) := \mathcal{S}((\text{Id} - \pi_T^{1,k+1}) \mathbf{u}_h, (\text{Id} - \pi_T^{1,k+1}) \mathbf{v}_h).$$

Estimate (61) then follows from (59) and from (82) in Lemma 21.

Due to the orthogonality condition **(S2)**, the GS **(3)** for Problem **(1)** with  $\mathbf{\Lambda} = \mathbf{I}_d$ , based on the gradient reconstructions **(41)** (with  $\tilde{\mathbf{S}}_T = \mathbf{S}_T$ ) and some function reconstruction  $\tilde{\Pi}_{\mathcal{D}_h}$ , is given by

$$\sum_{T \in \mathcal{T}_h} (\nabla_{\mathbf{r}_T^{k+1}} \underline{u}_T, \nabla_{\mathbf{r}_T^{k+1}} \underline{v}_T)_T + \sum_{T \in \mathcal{T}_h} (\mathbf{S}_T \underline{u}_T, \mathbf{S}_T \underline{v}_T)_T = \int_{\Omega} f \tilde{\Pi}_{\mathcal{D}_h} \underline{v}_h. \quad (62)$$

For all methods except the ncMFD with  $k = 1$ , accounting for **(14)** when  $k = 0$  and  $l = -1$ , we have  $L_{f,T}(\underline{v}_T) = (f, \underline{v}_T)_T$  and thus the choice  $\tilde{\Pi}_{\mathcal{D}_h} = \Pi_{\mathcal{D}_h}$  defined in **(16)** ensures that the right-hand sides of **(60)** and **(62)** coincide. For the ncMFD with  $k = 1$ , a slightly different discretisation of the right-hand side has to be considered in order to ensure optimal  $L^2$ -error estimates under elliptic regularity; see [39, Section 2.7] for further details.

To prove that **(60)** can be written as **(62)**, it remains to show that for any stabilisation  $\text{Stab}_T$  as above, there exists  $\mathbf{S}_T$  satisfying **(S1)–(S3)** and such that

$$\forall \underline{u}_T, \underline{v}_T \in \underline{U}_T^{k,l}, (\mathbf{S}_T \underline{u}_T, \mathbf{S}_T \underline{v}_T)_T = \text{Stab}_T(\underline{u}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{u}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T). \quad (63)$$

Let us fix an initial  $\mathbf{S}_T^0$  satisfying the design properties (for example, the stabilisation defined by **(36)**). The property **(S1)** on  $\mathbf{S}_T^0$  and Lemma 20 in Appendix A.2 below show that  $\mathbf{S}_T^0 : \underline{U}_T^{k,l} \rightarrow \text{Im}(\mathbf{S}_T^0)$  and  $(\text{Id} - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1}) : \underline{U}_T^{k,l} \rightarrow \text{Im}(\text{Id} - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1})$  have the same kernel. As shown by **(61)** and Lemma 20,  $\text{Stab}_T$  is an inner product on  $\text{Im}(\text{Id} - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1})$ . Applying thus [28, Lemma A.3] produces an inner product  $\langle \cdot, \cdot \rangle$  on  $\text{Im}(\mathbf{S}_T^0)$  such that

$$\forall \underline{u}_T, \underline{v}_T \in \underline{U}_T^{k,l}, \langle \mathbf{S}_T^0 \underline{u}_T, \mathbf{S}_T^0 \underline{v}_T \rangle = \text{Stab}_T(\underline{u}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{u}_T, \underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T). \quad (64)$$

Using then [29, Lemma 5.2] with the inner products  $\langle \cdot, \cdot \rangle$  and  $(\cdot, \cdot)_T$  on  $\text{Im}(\mathbf{S}_T^0)$ , we find an isomorphism  $\mathcal{L}_T$  of  $\text{Im}(\mathbf{S}_T^0)$  such that

$$\forall \underline{u}_T, \underline{v}_T \in \underline{U}_T^{k,l}, \langle \mathbf{S}_T^0 \underline{u}_T, \mathbf{S}_T^0 \underline{v}_T \rangle = (\mathcal{L}_T \mathbf{S}_T^0 \underline{u}_T, \mathcal{L}_T \mathbf{S}_T^0 \underline{v}_T)_T. \quad (65)$$

Combining **(64)** and **(65)** shows that **(63)** holds with

$$\mathbf{S}_T = \mathcal{L}_T \mathbf{S}_T^0.$$

The proof that  $\mathbf{S}_T$  defined above satisfies the design properties is easy. The  $L^2$ -stability and boundedness **(S1)** is a direct consequence of **(63)** with  $\underline{v}_T = \underline{u}_T$  and of **(61)**. The image of  $\mathbf{S}_T^0$  is, by assumption,  $L^2(T)^d$ -orthogonal to  $\mathbb{P}^k(T)^d$  and contained in some  $\mathbb{P}^{k_S}(\mathcal{P}_T)$ . Since  $\mathcal{L}_T$  is an isomorphism of  $\text{Im}(\mathbf{S}_T^0)$ , we have  $\text{Im}(\mathbf{S}_T) = \text{Im}(\mathbf{S}_T^0)$  and Properties **(S2)** and **(S3)** on  $\mathbf{S}_T$  therefore follow.

*Remark 17* (Formulations based on  $\mathbf{G}_T^k$ ). With minor modifications, it is easy to construct a stabilisation  $\mathbf{S}_T$  that satisfies  $\tilde{\mathbf{(S2)}}$  instead of **(S2)**. This can be done at a reduced cost, as discussed in Section 4.1. The reasoning above can also be easily adapted to the first formulation of ncVEM, provided that the terms  $\nabla_{\mathbf{r}_T^{k+1}}$  in **(60)** are replaced with  $\mathbf{G}_T^k$ .

## A Proofs of the results on DSGDs

This section contains the proofs of Theorem 8 and Proposition 9 preceeded by some preliminary results: the study of the properties of stabilising contributions satisfying **(S1)–(S3)** and uniform equivalences of discrete  $W^{1,p}$ -seminorms. We also include lemmas used in Section 4 to show that the HHO method, ncMFD method and ncVEM are GDMs.

## A.1 Properties of the stabilising contribution

**Proposition 18** (Properties of  $\mathbf{S}_T$ ). *Let  $\{\mathbf{S}_T : T \in \mathcal{T}_h\}$  be a family of stabilising contributions satisfying assumptions **(S1)**–**(S3)**. Then, the following properties hold:*

(i)  *$L^p$ -stability and boundedness. For all  $T \in \mathcal{T}_h$  and all  $\underline{v}_T \in \underline{U}_T^{k,l}$ ,*

$$\|\mathbf{S}_T \underline{v}_T\|_{L^p(T)^d} \simeq |\underline{v}_T|_{p,\partial T}, \quad (66)$$

*with hidden constant as in (25) and additionally depending on  $p$  and  $k_S$ .*

(ii) *Consistency. For all  $T \in \mathcal{T}_h$  and all  $v \in W^{k+2,p}(T)$ ,*

$$\|\mathbf{S}_T \underline{I}_T^{k,l} v\|_{L^p(T)^d} \lesssim h_T^{k+1} |v|_{W^{k+2,p}(T)}, \quad (67)$$

*where  $a \lesssim b$  means  $a \leq Cb$  with real number  $C > 0$  independent of both  $h$  and  $T$ , but possibly depending on  $d, \varrho, k, l, p$  and  $k_S$ .*

*As a consequence of (67), if  $v \in \mathbb{P}^{k+1}(T)$ , then  $\mathbf{S}_T \underline{I}_T^{k,l} v = 0$ .*

In the proof, we will need the following direct and reverse Lebesgue embeddings, proved in [16, Lemma 5.1]: Let  $X$  denote a measurable subset of  $\mathbb{R}^d$  with inradius  $r_X$  and diameter  $h_X$ , and let two reals  $r, s \in [1, +\infty]$  and an integer  $\ell \in \mathbb{N}$  be fixed. Then, for all  $q \in \mathbb{P}^\ell(X)$ , it holds that

$$\|q\|_{L^r(X)} \simeq |X|^{\frac{1}{r} - \frac{1}{s}} \|q\|_{L^s(X)}, \quad (68)$$

where  $a \simeq b$  means  $Ca \leq b \leq C^{-1}a$  with real number  $C > 0$  only depending on  $d$ , a lower bound of the ratio  $\frac{r_X}{h_X}$ ,  $r, s$ , and  $\ell$ .

We will also need the following  $L^p$ -trace inequality (see, e.g., [16, Eq. (A.10)]): For all  $T \in \mathcal{T}_h$  and all  $v \in W^{1,p}(T)$ ,

$$h_T^{\frac{1}{p}} \|v\|_{L^p(\partial T)} \lesssim \|v\|_{L^p(T)} + h_T \|\nabla v\|_{L^p(T)^d}, \quad (69)$$

where  $a \lesssim b$  means  $a \leq Cb$  with real number  $C > 0$  independent of  $h$  and of  $T$ , but possibly depending on  $d, p$ , and  $\varrho$ . When  $v \in \mathbb{P}^\ell(T)$  for some integer  $\ell \geq 0$ , combining (69) with the following inverse inequality (see, e.g., [16, Remark A.2]):

$$\|\nabla v\|_{L^p(T)^d} \lesssim h_T^{-1} \|v\|_{L^p(T)}, \quad (70)$$

yields

$$h_T^{\frac{1}{p}} \|v\|_{L^p(\partial T)} \lesssim \|v\|_{L^p(T)}, \quad (71)$$

where the hidden multiplicative constant in (70) and (71) can additionally depend on  $\ell$ .

We are now ready to prove Proposition 18.

*Proof of Proposition 18.* (i)  *$L^p$ -stability and boundedness.* When  $p = 2$ , (66) coincides with (25). Let us now consider the case  $p \neq 2$  and recall the following inequalities valid for all integers  $n \geq 1$  and all reals  $q \in [1, +\infty)$ ,  $a_i \geq 0$  ( $1 \leq i \leq n$ ):

$$\sum_{i=1}^n a_i^q \leq \left( \sum_{i=1}^n a_i \right)^q \leq n^{q-1} \sum_{i=1}^n a_i^q, \quad n^{\frac{1-q}{q}} \sum_{i=1}^n a_i^{\frac{1}{q}} \leq \left( \sum_{i=1}^n a_i \right)^{\frac{1}{q}} \leq \sum_{i=1}^n a_i^{\frac{1}{q}}. \quad (72)$$

We also notice that, owing to (27), it holds for all  $T \in \mathcal{T}_h$  that

$$|P| \simeq |T| \quad \forall P \in \mathcal{P}_T, \quad \text{card}(\mathcal{P}_T) \simeq 1. \quad (73)$$

To prove  $|P| \simeq |T|$ , it suffices to observe that  $|P| \simeq h_P^d \simeq h_T^d \simeq |T|$ , where we have used, respectively, the first and second conditions in (27) and the mesh regularity to conclude. The bound on  $\text{card}(\mathcal{P}_T)$  follows by writing  $|T| = \sum_{P \in \mathcal{P}_T} |P| \simeq \sum_{P \in \mathcal{P}_T} |T| = \text{card}(\mathcal{P}_T)|T|$ .

Let now  $T \in \mathcal{T}_h$  and  $\underline{v}_T \in \underline{U}_T^{k,l}$  be fixed. Since  $\mathbf{S}_T \underline{v}_T \in \mathbb{P}^{ks}(P)$  for all  $P \in \mathcal{P}_T$ , we have that

$$\begin{aligned} \|\mathbf{S}_T \underline{v}_T\|_{L^p(T)^d}^p &= \sum_{P \in \mathcal{P}_T} \|\mathbf{S}_T \underline{v}_T\|_{L^p(P)^d}^p \\ &\simeq \sum_{P \in \mathcal{P}_T} |P|^{p\left(\frac{1}{p}-\frac{1}{2}\right)} \|\mathbf{S}_T \underline{v}_T\|_{L^2(P)^d}^p && \text{Eqs. (27) and (68)} \\ &\simeq |T|^{p\left(\frac{1}{p}-\frac{1}{2}\right)} \sum_{P \in \mathcal{P}_T} \|\mathbf{S}_T \underline{v}_T\|_{L^2(P)^d}^p && \text{Eq. (73)}. \end{aligned}$$

In the second line, condition (27) is invoked to use  $\varrho$  as a lower bound for  $\frac{r_P}{h_P}$  when applying (68). Using the first pair of inequalities in (72) with  $q = \frac{p}{2}$  if  $p \geq 2$ , the second pair of inequalities in (72) with  $q = \frac{2}{p}$  if  $p < 2$  and, in both cases,  $a_i = \|\mathbf{S}_T \underline{v}_T\|_{L^2(P)^d}^2$  and  $n = \text{card}(\mathcal{P}_T) \simeq 1$ , we infer

$$\|\mathbf{S}_T \underline{v}_T\|_{L^p(T)^d}^p \simeq |T|^{p\left(\frac{1}{p}-\frac{1}{2}\right)} \left( \sum_{P \in \mathcal{P}_T} \|\mathbf{S}_T \underline{v}_T\|_{L^2(P)^d}^2 \right)^{\frac{p}{2}} = |T|^{p\left(\frac{1}{p}-\frac{1}{2}\right)} \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}^p.$$

Taking the  $p$ th root of the above relation, we arrive at

$$\|\mathbf{S}_T \underline{v}_T\|_{L^p(T)^d} \simeq |T|^{\frac{1}{p}-\frac{1}{2}} \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}. \quad (74)$$

Proceeding similarly using (68) repeatedly on the faces of  $T$  and using  $|F|h_F \simeq |T|$ , we can prove that

$$|\underline{v}_T|_{2,\partial T} \simeq |T|^{\frac{1}{2}-\frac{1}{p}} |\underline{v}_T|_{p,\partial T}. \quad (75)$$

Combining (74) and (75) with (25), (66) follows.

(ii) *Consistency.* Let a mesh element  $T \in \mathcal{T}_h$  be fixed and set, for the sake of brevity,  $\hat{\underline{v}}_T := \underline{I}_T^{k,l} v$ . Using the uniform equivalence (66) proved in the first point, and recalling the definition (23) of the  $|\cdot|_{p,\partial T}$ -seminorm, it is inferred that

$$\|\mathbf{S}_T \hat{\underline{v}}_T\|_{L^p(T)^d}^p \lesssim |\hat{\underline{v}}_T|_{p,\partial T}^p = \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|(\delta_{TF}^k - \delta_T^l) \hat{\underline{v}}_T\|_{L^p(F)}^p.$$

For all  $F \in \mathcal{F}_T$ , using the triangle inequality followed by the discrete trace inequality (71), we infer that

$$\|(\delta_{TF}^k - \delta_T^l) \hat{\underline{v}}_T\|_{L^p(F)}^p \lesssim \|\delta_{TF}^k \hat{\underline{v}}_T\|_{L^p(F)}^p + \|\delta_T^l \hat{\underline{v}}_T\|_{L^p(F)}^p \lesssim \|\delta_{TF}^k \hat{\underline{v}}_T\|_{L^p(F)}^p + h_T^{-1} \|\delta_T^l \hat{\underline{v}}_T\|_{L^p(T)}^p.$$

Using the above inequality together with the uniform bound on the number of faces of  $T$  (see [19, Lemma 1.41]) for the second term, and expanding the difference operators according to their definitions (21), we arrive at

$$\|\mathbf{S}_T \hat{\underline{v}}_T\|_{L^p(T)^d}^p \lesssim h_T^{-p} \|\pi_T^{0,l}(\mathbf{r}_T^{k+1} \hat{\underline{v}}_T - v)\|_{L^p(T)}^p + \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|\pi_F^{0,k}(\mathbf{r}_T^{k+1} \hat{\underline{v}}_T - v)\|_{L^p(F)}^p.$$



We notice that the projectors in the above bound can be removed invoking the  $L^p(T)$ -boundedness of  $\pi_T^{0,l}$  for the first term and the  $L^p(F)$ -boundedness of  $\pi_F^{0,k}$  for the second (see [16, Lemma 3.2]). Combining this observation with the optimal approximation properties of  $\mathbf{r}_T^{k+1} \circ \underline{\mathbf{I}}_T^{k,l}$  discussed in Remark 4, we then conclude that

$$\|\mathbf{S}_T \widehat{\underline{v}}_T\|_{L^p(T)^d}^p \lesssim h_T^{-p} \|\mathbf{r}_T^{k+1} \widehat{\underline{v}}_T - v\|_{L^p(T)}^p + \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|\mathbf{r}_T^{k+1} \widehat{\underline{v}}_T - v\|_{L^p(F)}^p \lesssim h_T^{p(k+1)} |v|_{W^{k+2,p}(T)}^p. \quad \square$$

## A.2 Uniform equivalence of discrete $W^{1,p}$ -seminorms

The second preliminary result is the uniform equivalence of various  $W^{1,p}$ -seminorms on the global space of discrete unknowns  $\underline{U}_h^{k,l}$ .

**Proposition 19** (Uniform equivalence of discrete  $W^{1,p}$ -seminorms). *Define the discrete seminorm  $\|\cdot\|_{1,p,h}$  such that, for all  $\underline{v}_h \in \underline{U}_h^{k,l}$ ,*

$$\begin{aligned} \|\underline{v}_h\|_{1,p,h}^p &:= \sum_{T \in \mathcal{T}_h} \|\underline{v}_T\|_{1,p,T}^p, \\ \|\underline{v}_T\|_{1,p,T}^p &:= \|\nabla v_T\|_{L^p(T)^d}^p + \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|v_F - v_T\|_{L^p(F)}^p \quad \forall T \in \mathcal{T}_h. \end{aligned} \quad (76)$$

Then,  $\|\cdot\|_{1,p,h}$  is a norm on the subspace  $\underline{U}_{h,0}^{k,l}$  and, denoting by  $\{\mathbf{S}_T : T \in \mathcal{T}_h\}$  a family of stabilising contributions that satisfy (S1)–(S3) and defining  $\nabla_{\mathcal{D}_h}$  by (24), it holds for all  $\underline{v}_h \in \underline{U}_h^{k,l}$  and all  $T \in \mathcal{T}_h$ ,

$$\|\underline{v}_T\|_{1,p,T} \simeq \|\underline{v}_T\|_{1,p,T} \simeq \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(T)^d}, \quad (77)$$

where  $a \simeq b$  means  $Ca \leq b \leq C^{-1}a$  with real number  $C > 0$  independent of  $T$  and  $h$ , but possibly depending on  $d, p, \varrho, k, l$ , and  $k_S$ . As a consequence, for all  $\underline{v}_h \in \underline{U}_h^{k,l}$ ,

$$\|\underline{v}_h\|_{1,p,h} \simeq \|\underline{v}_h\|_{1,p,h} \simeq \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}. \quad (78)$$

*Proof.* The fact that  $\|\cdot\|_{1,p,h}$  is a norm on  $\underline{U}_{h,0}^{k,l}$  can be proved in a similar manner as for the case  $p = 2$  and  $l = k$  considered in [20, Proposition 5]. The local seminorm equivalence  $\|\underline{v}_h\|_{1,p,T} \simeq \|\underline{v}_h\|_{1,p,T}$  valid for all  $T \in \mathcal{T}_h$ , on the other hand, is proved in [16, Lemma 5.2] (see also references therein) for the case  $l = k$ , and the same reasoning extends to  $l = k - 1$  and  $l = k + 1$ .

Let now a mesh element  $T \in \mathcal{T}_h$  be fixed. We have that

$$\begin{aligned} \|\nabla_{\mathcal{D}} \underline{v}_h\|_{L^p(T)^d} &= \|\mathcal{G}_T \underline{v}_T\|_{L^p(T)^d} \simeq |T|^{\frac{1}{p} - \frac{1}{2}} \|\mathcal{G}_T \underline{v}_T\|_{L^2(T)^d} \\ &= |T|^{\frac{1}{p} - \frac{1}{2}} \left( \|\mathbf{G}_T^k \underline{v}_T\|_{L^2(T)^d}^2 + \|\mathbf{S}_T \underline{v}_T\|_{L^2(T)^d}^2 \right)^{\frac{1}{2}} \\ &\simeq |T|^{\frac{1}{p} - \frac{1}{2}} \left( \|\mathbf{G}_T^k \underline{v}_T\|_{L^2(T)^d} + |\underline{v}_T|_{2,\partial T} \right) \\ &\simeq \left( \|\mathbf{G}_T^k \underline{v}_T\|_{L^p(T)^d} + |\underline{v}_T|_{p,\partial T} \right) \simeq \|\underline{v}_T\|_{1,p,T}, \end{aligned} \quad (79)$$

where we have used a reasoning similar to the one leading to (74) in the first line (recall that  $\mathcal{G}_T \underline{v}_T$  is piecewise polynomial on  $T$  owing to (S3)), the orthogonality property (S2)

in the second line, the stability and boundedness property **(S1)** in the third line, and (75) together with the discrete Lebesgue embeddings (68) in the last line. This concludes the proof of (77). The global version (78) follows by raising (77) to the power  $p$  and summing over  $T \in \mathcal{T}_h$ .  $\square$

The following lemma, which justifies the importance of the seminorm  $|\cdot|_{p,\partial T}$ , was used in Section 4.5 to prove that HHO, ncMFD and ncVEM are GDM.

**Lemma 20.** *For any  $T \in \mathcal{T}_h$  and any  $\underline{v}_T \in \underline{U}_T^{k,l}$ , it holds that*

$$\|\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T\|_{1,p,T} \simeq |\underline{v}_T|_{p,\partial T}, \quad (80)$$

where  $a \simeq b$  means  $C^{-1}a \leq b \leq Ca$  with real number  $C > 0$  depending only on  $d, \varrho, p, k$ , and  $l$ . As a consequence,  $\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T = 0$  if and only if  $|\underline{v}_T|_{p,\partial T} = 0$ .

*Proof.* Here,  $a \lesssim b$  means that  $a \leq Cb$  for some real number  $C > 0$  as in the statement. Using direct and inverse Lebesgue inequalities as in the proof of Proposition 18, we deduce from (43) that  $\|\nabla \delta_T^l \underline{v}_T\|_{L^p(T)^d} \lesssim |\underline{v}_T|_{p,\partial T}$ . Hence, using the relation (22) together with the definitions (76) of  $\|\cdot\|_{1,p,T}$  and (23) of  $|\underline{v}_T|_{p,\partial T}$ , we obtain

$$\|\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T\|_{1,p,T}^p = \|\nabla \delta_T^l \underline{v}_T\|_{L^p(T)^d}^p + \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|(\delta_{TF}^k - \delta_T^l) \underline{v}_T\|_{L^p(F)}^p \simeq |\underline{v}_T|_{p,\partial T}^p,$$

which is (80). If  $\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T = 0$ , the relation above shows that  $|\underline{v}_T|_{p,\partial T} = 0$ . Conversely, if  $|\underline{v}_T|_{p,\partial T} = 0$ , then  $\|\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T\|_{1,p,T} = 0$  which implies that  $\underline{v}_T$  is constant equal to  $c$  and that  $v_F = v_T = c$  for all  $F \in \mathcal{F}_T$ . Then  $\mathbf{r}_T^{k+1} \underline{v}_T = \mathbf{r}_T^{k+1} \underline{I}_T^{k,l} c = \pi_T^{1,k+1} c = c$  (see Remark 4 and additionally observe, for  $k = 0$  and  $l = -1$ , that  $v_T = c$  owing to the choice of the weights  $\omega_{TF}$ ). Thus  $\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T = \underline{I}_T^{k,l} c - \underline{I}_T^{k,l} c = 0$ .  $\square$

The last lemma of this section was used in Section 4.5 to analyse the ncVEM stabilisation in the context of the GDM.

**Lemma 21.** *Let  $\mathcal{T}_h$  be a mesh and  $V_k^{h,nc}$  be the non-conforming VEM space defined in Section 4.4. For all  $T \in \mathcal{T}_h$  and all  $\mathbf{v}_h \in V_k^{h,nc}(T)$ , it holds that*

$$\|\nabla \mathbf{v}_h\|_{L^2(T)^d} \simeq \|\underline{I}_T^{k,k-1} \mathbf{v}_h\|_{1,2,T}, \quad (81)$$

where  $\|\cdot\|_{1,2,T}$  is defined in (76) and  $a \simeq b$  means  $C^{-1}a \leq b \leq Ca$  with real number  $C > 0$  independent on  $h$ , but possibly depending on  $d, k, l$  and  $\varrho$ . As a consequence,

$$\text{for all } \mathbf{v}_h \in V_k^{h,nc}(T) \text{ such that } \pi_T^{1,k} \mathbf{v}_h = 0, \|\nabla \mathbf{v}_h\|_{L^2(T)^d} \simeq |\underline{I}_T^{k,k-1} \mathbf{v}_h|_{1,2,T}. \quad (82)$$

*Proof.* Here,  $a \lesssim b$  means  $a \leq Cb$  with  $C$  as in the statement. Since  $\nabla \mathbf{v}_h \cdot \mathbf{n}_{TF} \in \mathbb{P}^k(F)$  for all  $F \in \mathcal{F}_T$  and  $\Delta \mathbf{v}_h \in \mathbb{P}^{k-1}(T)$ , integrating by parts and setting  $\underline{v}_T = \underline{I}_T^{k,k-1} \mathbf{v}_h$ ,

$$\begin{aligned} \|\nabla \mathbf{v}_h\|_{L^2(T)^d}^2 &= -(\mathbf{v}_h, \Delta \mathbf{v}_h)_T + \sum_{F \in \mathcal{F}_T} (\mathbf{v}_h, \nabla \mathbf{v}_h \cdot \mathbf{n}_{TF})_F \\ &= (-v_T, \Delta \mathbf{v}_h)_T + \sum_{F \in \mathcal{F}_T} (v_F, \nabla \mathbf{v}_h \cdot \mathbf{n}_{TF})_F \\ &= (\nabla v_T, \nabla \mathbf{v}_h)_T + \sum_{F \in \mathcal{F}_T} (v_F - v_T, \nabla \mathbf{v}_h \cdot \mathbf{n}_{TF})_F. \end{aligned}$$

The Cauchy–Schwarz inequality and the trace inequality (71) applied on  $\nabla \mathbf{v}_h \cdot \mathbf{n}_{TF}$  then yields  $\|\nabla \mathbf{v}_h\|_{L^2(T)^d} \lesssim \|\underline{v}_T\|_{1,2,T}$ , which is half of (81).

To prove the second half for  $k \geq 1$ , recall first that  $v_T = \pi_T^{0,k-1} \mathbf{v}_h$ . A triangle inequality and (11a) with  $\ell = k - 1$ ,  $p = 2$ ,  $\alpha = 0$  and  $r = s = 1$  thus show that

$$\|\nabla v_T\|_{L^2(T)^d} \lesssim \|\nabla \mathbf{v}_h\|_{L^2(T)^d}. \quad (83)$$

Then, write

$$\|v_F - v_T\|_{L^2(F)} = \|\pi_F^{0,k}(\mathbf{v}_h - \pi_T^{0,k-1} \mathbf{v}_h)\|_{L^2(F)} \leq \|\mathbf{v}_h - \pi_T^{0,k-1} \mathbf{v}_h\|_{L^2(F)} \leq h_T^{\frac{1}{2}} \|\nabla \mathbf{v}_h\|_{L^2(T)^d},$$

where (11b) was used with  $\ell = k - 1$ ,  $p = 2$ ,  $\alpha = 0$ ,  $r = 0$  and  $s = 1$ . Combining this estimate with (83) concludes the proof of (81). The above argument can be extended along similar lines to the case  $k = 0$ , the only variation being linked to the fact that, in this case,  $v_T$  is not an  $L^2$ -orthogonal projection of  $\mathbf{v}_h$  (see (14)).

To prove (82), notice that if  $\pi_T^{1,k+1} \mathbf{v}_h = 0$  then  $\mathbf{r}_T^{k+1} \underline{v}_T = \mathbf{r}_T^{k+1} \underline{I}_T^{k,l} \mathbf{v}_h = \pi_T^{1,k+1} \mathbf{v}_h = 0$ . Hence, (81) yields

$$\|\nabla \mathbf{v}_h\|_{L^2(T)^d} \simeq \|\underline{v}_T - \underline{I}_T^{k,l} \mathbf{r}_T^{k+1} \underline{v}_T\|_{1,2,T},$$

and the proof is complete by invoking (80).  $\square$

### A.3 Main results

We are now ready to prove the main results stated in Section 3.5.

#### A.3.1 Properties of Discontinuous Skeletal Gradient Discretisations

*Proof of Theorem 8.* We use a polytopal toolbox in the spirit of [30] and [27, Section 7.2]. Let

$$X_{\mathcal{T}_h,0} := \left\{ \underline{w}_h = ((w_T)_{T \in \mathcal{T}_h}, (w_F)_{F \in \mathcal{F}_h}) : w_T \in \mathbb{R}, w_F \in \mathbb{R}, w_F = 0 \quad \forall F \in \mathcal{F}_h^b \right\} = \underline{U}_{h,0}^{0,0}$$

and  $\Pi_{\mathcal{T}_h} : X_{\mathcal{T}_h,0} \rightarrow L^p(\Omega)$ ,  $\nabla_{\mathcal{T}_h} : X_{\mathcal{T}_h,0} \rightarrow L^p(\Omega)^d$  be defined by, for all  $\underline{w}_h \in X_{\mathcal{T}_h,0}$  and all  $T \in \mathcal{T}_h$ ,

$$(\Pi_{\mathcal{T}_h} \underline{w}_h)|_T = w_T \quad \text{and} \quad (\nabla_{\mathcal{T}_h} \underline{w}_h)|_T = \frac{1}{|T|} \sum_{F \in \mathcal{F}_T} |F| w_F \mathbf{n}_{TF}.$$

By [27, Corollary 7.13], the coercivity, limit-conformity, consistency, and compactness of  $(\mathcal{D}_h)_{h \in \mathcal{H}}$  follow if we find a mapping  $\Phi : \underline{U}_{h,0}^{k,l} \rightarrow X_{\mathcal{T}_h,0}$  (“control” of  $\mathcal{D}_h$ ) such that, recalling the definition (76) of  $\|\cdot\|_{1,p,h}$  and setting

$$\|\Phi\|_{\mathcal{D}_h, \mathcal{T}_h} := \max_{\underline{v}_h \in \underline{U}_{h,0}^{k,l} \setminus \{0_h\}} \frac{\|\Phi(\underline{v}_h)\|_{1,p,h}}{\|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}}, \quad (84a)$$

$$\omega^\Pi(\mathcal{D}_h, \mathcal{T}_h, \Phi) := \max_{\underline{v}_h \in \underline{U}_{h,0}^{k,l} \setminus \{0_h\}} \frac{\|\Pi_{\mathcal{D}_h} \underline{v}_h - \Pi_{\mathcal{T}_h} \Phi(\underline{v}_h)\|_{L^p(\Omega)}}{\|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}}, \quad (84b)$$

$$\omega^\nabla(\mathcal{D}_h, \mathcal{T}_h, \Phi) := \max_{\underline{v}_h \in \underline{U}_{h,0}^{k,l} \setminus \{0_h\}} \frac{\left( \sum_{T \in \mathcal{T}_h} |T|^{1-p} \left| \int_T (\nabla_{\mathcal{D}_h} \underline{v}_h - \nabla_{\mathcal{T}_h} \Phi(\underline{v}_h)) \right|^p \right)^{\frac{1}{p}}}{\|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}}, \quad (84c)$$

we have

$$\|\Phi\|_{\mathcal{D}_h, \mathcal{T}_h} \lesssim 1 \text{ and, as } h \rightarrow 0, \omega^\Pi(\mathcal{D}_h, \mathcal{T}_h, \Phi) \rightarrow 0 \text{ and } \omega^\nabla(\mathcal{D}_h, \mathcal{T}_h, \Phi) \rightarrow 0, \quad (85)$$

where  $a \lesssim b$  means  $a \leq Cb$  with real number  $C > 0$  independent of  $h$ , but possibly depending on  $d, \rho, k, l$ , and  $k_S$ .

Let  $\Phi$  be defined the following way. For all  $\underline{v}_h = ((v_T)_{T \in \mathcal{T}_h}, (v_F)_{F \in \mathcal{F}_h}) \in \underline{U}_h^{k,l}$ , we let  $\Phi(\underline{v}_h) := \underline{v}_h^0 = ((v_T^0)_{T \in \mathcal{T}_h}, (v_F^0)_{F \in \mathcal{F}_h}) \in X_{\mathcal{T}_h, 0}$  be such that  $v_T^0 = \pi_T^{0,0} v_T$  for all  $T \in \mathcal{T}_h$  and  $v_F^0 = \pi_F^{0,0} v_F$  for all  $F \in \mathcal{F}_h$ . Properties (85) follow if we establish that, for all  $\underline{v}_h \in \underline{U}_{h,0}^{k,l}$ ,

$$\|\underline{v}_h^0\|_{1,p,h} \lesssim \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}, \quad (86a)$$

$$\|\Pi_{\mathcal{D}_h} \underline{v}_h - \Pi_{\mathcal{T}_h} \underline{v}_h^0\|_{L^p(T)} \lesssim h \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}, \quad (86b)$$

and, for all  $T \in \mathcal{T}_h$  and all  $v_T \in \underline{U}_T^{k,l}$ ,

$$(\mathcal{G}_T \underline{v}_T, \boldsymbol{\eta})_T = \sum_{F \in \mathcal{F}_T} (v_F^0, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F \quad \forall \boldsymbol{\eta} \in \mathbb{P}^0(T)^d. \quad (86c)$$

Indeed, (86a) gives a bound on  $\|\Phi\|_{\mathcal{D}_h, \mathcal{T}_h}$ , (86b) gives an  $\mathcal{O}(h)$  estimate on  $\omega^\Pi(\mathcal{D}_h, \mathcal{T}_h, \Phi)$ , and (86c) shows that  $\omega^\nabla(\mathcal{D}_h, \mathcal{T}_h, \Phi) = 0$ .

(i) *Proof of (86a)*. By the definition (76) of the  $\|\cdot\|_{1,p,h}$ -seminorm along with that of  $\underline{v}_h^0$ , we have that

$$\|\underline{v}_h^0\|_{1,p,h}^p = \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|v_F^0 - v_T^0\|_{L^p(F)}^p. \quad (87)$$

Let now a mesh element  $T \in \mathcal{T}_h$  be fixed and observe that, for all  $F \in \mathcal{F}_T$ ,

$$\begin{aligned} \|v_F^0 - v_T^0\|_{L^p(F)} &= \|\pi_F^{0,0}(v_F - v_T)\|_{L^p(F)} \leq \|v_F - v_T\|_{L^p(F)} \\ &\leq \|v_F - v_T\|_{L^p(F)} + \|v_T - v_T^0\|_{L^p(F)} \\ &\lesssim \|v_F - v_T\|_{L^p(F)} + h_T^{-\frac{1}{p}} \|v_T - v_T^0\|_{L^p(T)} \\ &\lesssim \|v_F - v_T\|_{L^p(F)} + h_T^{1-\frac{1}{p}} \|\nabla v_T\|_{L^p(T)^d}, \end{aligned} \quad (88)$$

where we have used the  $L^p$ -boundedness of the  $L^2$ -orthogonal projector (see [16, Lemma 3.2]) in the first line, the triangle inequality in the second line, the discrete  $L^p$ -trace inequality (71) in the third line, and a local Poincaré–Wirtinger inequality which can be inferred from (11a) with  $\alpha = \ell = r = 0$  and  $s = 1$  to conclude.

Taking the  $p$ th power of (88), multiplying by  $h_F^{1-p} \simeq h_T^{1-p}$  and summing over  $F \in \mathcal{F}_T$  and  $T \in \mathcal{T}_h$  leads to  $\|\underline{v}_h^0\|_{1,p,h} \lesssim \|\underline{v}_h\|_{1,p,h}$ . Estimate (86a) follows by using (78).

(ii) *Proof of (86b)*. Using a local Poincaré–Wirtinger inequality as above we infer, for all  $T \in \mathcal{T}_h$ ,  $\|v_T - v_T^0\|_{L^p(T)} \lesssim h_T \|\nabla v_T\|_{L^p(T)^d}$ . Taking the  $p$ th power of this inequality, summing over  $T \in \mathcal{T}_h$ , and using  $h_T \leq h$  and the uniform norm equivalence (78) to bound the right-hand side, (86b) follows.

(iii) *Proof of (86c)*. Let  $\boldsymbol{\eta} \in \mathbb{P}^0(T)^d \subset \mathbb{P}^k(T)^d$ . Since  $\nabla \cdot \boldsymbol{\eta} = 0$ , using the orthogonality property (S2) followed by the definition (17) of  $\mathcal{G}_T^k$  we infer that  $(\mathcal{G}_T \underline{v}_T, \boldsymbol{\eta})_T = \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F$ . Equation (86c) then follows by noticing that,  $\boldsymbol{\eta} \cdot \mathbf{n}_{TF}$  being constant on  $F$ ,  $(v_F, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F = (\pi_F^{0,0} v_F, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F = (v_F^0, \boldsymbol{\eta} \cdot \mathbf{n}_{TF})_F$ .

The GD-consistency follows from Proposition 9 (proved below), and from [27, Lemma 2.13] which shows that the consistency holds provided that  $S_{\mathcal{D}_h}(\phi) \rightarrow 0$  for all  $\phi$  in a dense subset of  $W_0^{1,p}(\Omega)$ .  $\square$

*Remark 22* (Condition (86a)). In [27], a slightly different norm is considered in the argument of the maximum in (84a). The original expression is obtained replacing  $\|\underline{v}_h^0\|_{1,p,h}^p$  by

$$\sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} d_{TF}^{1-p} \|v_F^0 - v_T^0\|_{L^p(F)}^p,$$

where the only difference with respect to (87) is that the role of the local length scale is played by  $d_{TF}$ , the orthogonal distance between a point  $\mathbf{x}_T$  inside  $T$  and the face  $F$ , instead of  $h_F$ . If, for all  $T \in \mathcal{T}_h$ , we choose the point  $\mathbf{x}_T$  such that condition (33) is verified, it can easily be proved that  $d_{TF} \simeq h_F$ , and the two norms are uniformly equivalent.

### A.3.2 Estimates on $S_{\mathcal{D}}$ and $W_{\mathcal{D}}$

*Proof of Proposition 9.* (i) *Estimates on the addends in  $S_{\mathcal{D}_h}$ .* Take  $\phi \in W_0^{1,p}(\Omega) \cap W^{l+1,p}(\mathcal{T}_h)$  and let  $\underline{v}_h = \underline{I}_h^{k,l} \phi \in \underline{U}_h^{k,l}$ . For all  $T \in \mathcal{T}_h$ , if  $l \geq 0$  then  $(\Pi_{\mathcal{D}_h} \underline{v}_h)|_T = v_T = \pi_T^{0,l} \phi$  on  $T$  so the approximation estimate (11a) applied with  $\alpha = 0$ ,  $\ell = l$ ,  $s = \ell + 1$  and  $r = 0$  yields

$$\|\Pi_{\mathcal{D}_h} \underline{v}_h - \phi\|_{L^p(T)} \lesssim h_T^{l+1} |\phi|_{W^{l+1,p}(T)}. \quad (89)$$

On the other hand, if  $l = -1$ , the specific choice (14) of  $v_T$  yields

$$\|\Pi_{\mathcal{D}_h} \underline{v}_h - \phi\|_{L^p(T)} \lesssim h_T |\phi|_{W^{1,p}(T)}. \quad (90)$$

Combining (89) and (90), taking the  $p$ th power, summing over  $T \in \mathcal{T}_h$ , and taking the  $p$ th root of the resulting inequality gives (28a).

For a fixed mesh element  $T \in \mathcal{T}_h$ , use the definition (24) of  $\nabla_{\mathcal{D}_h}$ , the commutativity property (18) of  $\mathbf{G}_T^k$ , the approximation property (11a) of  $\pi_T^{0,k}$  with  $s = k + 1$  and  $r = 0$ , and the consistency (67) of  $\mathbf{S}_T$  to obtain

$$\begin{aligned} \|\nabla_{\mathcal{D}_h} \underline{v}_h - \nabla \phi\|_{L^p(T)^d} &\leq \|\mathbf{G}_T^k \underline{v}_T - \nabla \phi\|_{L^p(T)^d} + \|\mathbf{S}_T \underline{v}_T\|_{L^p(T)^d} \\ &\lesssim h_T^{k+1} |\nabla \phi|_{W^{k+1,p}(T)^d} + h_T^{k+1} |\phi|_{W^{k+2,p}(T)} \lesssim h_T^{k+1} |\phi|_{W^{k+2,p}(T)}. \end{aligned}$$

The estimate (28b) follows taking the  $p$ th power, summing over  $T \in \mathcal{T}_h$ , and taking the  $p$ th root of the resulting inequality.

If  $l \in \{k, k+1\}$  or  $l = -1$  (in which case  $k = 0$ ), the estimate (29) on  $S_{\mathcal{D}_h}(\phi)$  is an immediate consequence of (28). Consider now  $l = k - 1$  and  $k \geq 1$ . An easy modification of the proof above shows that, for all  $\phi \in W_0^{1,p}(\Omega) \cap W^{k+1,p}(\mathcal{T}_h)$ ,  $\|\nabla_{\mathcal{D}_h} \underline{v}_h - \nabla \phi\|_{L^p(\Omega)^d} \lesssim h^k \|\phi\|_{W^{k+1,p}(\mathcal{T}_h)}$ . Then (29) follows from this modified version of (28b) and from (28a).

(ii) *Estimate on  $W_{\mathcal{D}_h}(\psi)$ .* For all  $\underline{v}_h \in X_{\mathcal{D}_h,0}$ ,

$$\begin{aligned} \int_{\Omega} \nabla_{\mathcal{D}_h} \underline{v}_h(\mathbf{x}) \cdot \psi(\mathbf{x}) \, d\mathbf{x} &= \sum_{T \in \mathcal{T}_h} (\nabla_{\mathcal{D}_h} \underline{v}_h, \psi - \pi_T^{0,k} \psi)_T + \sum_{T \in \mathcal{T}_h} (\nabla_{\mathcal{D}_h} \underline{v}_h, \pi_T^{0,k} \psi)_T \\ &=: \sum_{T \in \mathcal{T}_h} \mathcal{A}_T + \sum_{T \in \mathcal{T}_h} \mathcal{B}_T. \end{aligned} \quad (91)$$

The approximation property (11a) of  $\pi_T^{0,k}$  with  $s = k + 1$ ,  $r = 0$  and  $p'$  instead of  $p$  yields

$$\begin{aligned}
\sum_{T \in \mathcal{T}_h} |\mathcal{A}_T| &\lesssim \sum_{T \in \mathcal{T}_h} h_T^{k+1} \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(T)^d} \|\boldsymbol{\psi}\|_{W^{k+1,p'}(T)^d} \\
&\leq h^{k+1} \left( \sum_{T \in \mathcal{T}_h} \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(T)^d}^p \right)^{1/p} \left( \sum_{T \in \mathcal{T}_h} \|\boldsymbol{\psi}\|_{W^{k+1,p'}(T)^d}^{p'} \right)^{1/p'} \\
&= h^{k+1} \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d} \|\boldsymbol{\psi}\|_{W^{k+1,p'}(\mathcal{T}_h)^d}.
\end{aligned} \tag{92}$$

By definitions (24) and (17) of  $\nabla_{\mathcal{D}_h}$  and  $\mathbf{G}_T^k$ , and by the orthogonality property (S2) of  $\mathbf{S}_T$ ,

$$\begin{aligned}
\mathcal{B}_T &= (\mathbf{G}_T^k \underline{v}_T, \boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi})_T = - (v_T, \nabla \cdot \boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi})_T + \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F \\
&= - (v_T, \nabla \cdot (\boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} - \boldsymbol{\psi}))_T + \sum_{F \in \mathcal{F}_T} (v_F, (\boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \mathbf{n}_{TF})_F \\
&\quad - (v_T, \nabla \cdot \boldsymbol{\psi})_T + \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F \\
&= (\nabla v_T, (\boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} - \boldsymbol{\psi}))_T + \sum_{F \in \mathcal{F}_T} (v_F - v_T, (\boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \mathbf{n}_{TF})_F \\
&\quad - (\Pi_{\mathcal{D}_h} \underline{v}_h, \nabla \cdot \boldsymbol{\psi})_T + \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F \\
&=: \mathcal{B}_{T,1} - (\Pi_{\mathcal{D}_h} \underline{v}_h, \nabla \cdot \boldsymbol{\psi})_T + \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F,
\end{aligned} \tag{93}$$

where we used an integration-by-parts and the definition (16) of  $\Pi_{\mathcal{D}_h}$  in the penultimate line. For any interface  $F$  with  $T_1, T_2$  as neighbouring mesh elements, since  $\boldsymbol{\psi} \in \mathbf{W}^{p'}(\text{div}; \Omega)$  we have  $\boldsymbol{\psi} \cdot \mathbf{n}_{T_1 F} + \boldsymbol{\psi} \cdot \mathbf{n}_{T_2 F} = 0$  on  $F$ . Moreover,  $v_F = 0$  whenever  $F$  is a boundary face. Hence

$$\sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F = \sum_{F \in \mathcal{F}_h^i} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{T_1 F} + \boldsymbol{\psi} \cdot \mathbf{n}_{T_2 F})_F + \sum_{F \in \mathcal{F}_h^b} (v_F, \boldsymbol{\psi} \cdot \mathbf{n}_{TF})_F = 0.$$

Summing (93) over  $T \in \mathcal{T}_h$  and using the previous relation leads to

$$\sum_{T \in \mathcal{T}_h} \mathcal{B}_T = \sum_{T \in \mathcal{T}_h} \mathcal{B}_{T,1} - \int_{\Omega} \Pi_{\mathcal{D}_h} \underline{v}_h(\mathbf{x}) \nabla \cdot \boldsymbol{\psi}(\mathbf{x}) \, d\mathbf{x}. \tag{94}$$

Recalling the definition (10) of  $\boldsymbol{\pi}_T^{0,k}$ , it is readily inferred that the first term in  $\mathcal{B}_{T,1}$  is zero since  $\nabla v_T \in \nabla \mathbb{P}^l(T) \subset \mathbb{P}^k(T)^d$ . Moreover, using again the approximation properties (11) of  $\pi_T^{0,k}$  with  $r = 0$ ,  $s = k + 1$  and  $p'$  instead of  $p$ , we can write

$$\begin{aligned}
|\mathcal{B}_{T,1}| &\lesssim \sum_{F \in \mathcal{F}_T} \|v_F - v_T\|_{L^p(F)} \|\boldsymbol{\pi}_T^{0,k} \boldsymbol{\psi} - \boldsymbol{\psi}\|_{L^{p'}(F)^d} \\
&\lesssim \sum_{F \in \mathcal{F}_T} h_T^{k+1 - \frac{1}{p'}} \|v_F - v_T\|_{L^p(F)} |\boldsymbol{\psi}|_{W^{k+1,p'}(T)^d} \\
&\lesssim h^{k+1} |\boldsymbol{\psi}|_{W^{k+1,p'}(T)^d} \left( \sum_{F \in \mathcal{F}_T} h_F^{\frac{1}{p} - 1} \|v_F - v_T\|_{L^p(F)} \right),
\end{aligned}$$

where we used  $h_F \leq h_T$  in the last line. Sum over  $T \in \mathcal{T}_h$  and invoke Hölder's inequality, the property  $\text{Card}(\mathcal{F}_T) \lesssim 1$ , and the norm equivalence (78) to deduce

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} |\mathcal{B}_{T,1}| &\lesssim h^{k+1} |\boldsymbol{\psi}|_{W^{k+1,p'}(\mathcal{T}_h)^d} \left( \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} h_F^{1-p} \|v_F - v_T\|_{L^p(F)}^p \right)^{1/p} \\ &\lesssim h^{k+1} \|\boldsymbol{\psi}\|_{W^{k+1,p'}(\mathcal{T}_h)^d} \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}. \end{aligned} \quad (95)$$

Finally, using (92), (94) and (95) in (91), we get

$$\left| \int_{\Omega} \nabla_{\mathcal{D}_h} \underline{v}_h(\mathbf{x}) \cdot \boldsymbol{\psi}(\mathbf{x}) \, d\mathbf{x} + \int_{\Omega} \Pi_{\mathcal{D}_h} \underline{v}_h(\mathbf{x}) \nabla \cdot \boldsymbol{\psi}(\mathbf{x}) \, d\mathbf{x} \right| \lesssim h^{k+1} \|\boldsymbol{\psi}\|_{W^{k+1,p'}(\mathcal{T}_h)^d} \|\nabla_{\mathcal{D}_h} \underline{v}_h\|_{L^p(\Omega)^d}.$$

The estimate (30) follows immediately.  $\square$

*Remark 23* (Choice of the gradient reconstruction). An inspection of the above proof shows that using (41) in place of (24) can lead to significant losses in the order of convergence for  $W_{\mathcal{D}_h}(\boldsymbol{\psi})$  (while the convergence expressed by (9) still holds true). As a matter of fact, with this choice one would have to replace throughout the proof  $\pi_T^{0,k} \boldsymbol{\psi}$  by the  $L^2$ -orthogonal projection of  $\boldsymbol{\psi}$  on  $\nabla \mathbb{P}^{k+1}(T)$ . The latter quantity has optimal approximation properties only if either  $k = 0$  (since  $\nabla \mathbb{P}^1(T) = \mathbb{P}^0(T)^d$ ) or there exists  $w \in L^{p'}(\Omega)$  such that  $\boldsymbol{\psi}|_T = \nabla w|_T$  for all  $T \in \mathcal{T}_h$ . Recalling Theorem 2 with  $p = 2$  and  $\boldsymbol{\sigma}(\mathbf{x}, u, \nabla u) = \nabla u$ , we see that for the Poisson equation,  $W_{\mathcal{D}_h}$  is applied to  $\boldsymbol{\psi} = \nabla u$ . In this case, the gradient reconstruction (41) leads to optimal convergence rates. However, there is a real loss of estimate for more general problems for which error estimates are written in terms of  $W_{\mathcal{D}_h}(\boldsymbol{\Lambda} \nabla u)$  (for anisotropic linear diffusion, see [27, Theorem 3.2]) or  $W_{\mathcal{D}_h}(|\nabla u|^{p-2} \nabla u)$  (for the  $p$ -Laplace equation, see Theorem 2).

**Acknowledgements** The first author was supported by the ANR grant HHOMM (ANR-15-CE40-0005). The second author was supported by the ARC Discovery Projects funding scheme (project number DP170100605). The third author was funded by the Laboratory Directed Research and Development program, under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy by Los Alamos National Laboratory, operated by Los Alamos National Security LLC under contract DE-AC52-06NA25396.

## References

- [1] Stable Discontinuous Galerkin FEM without penalty parameter, 2015.
- [2] B. Ayuso de Dios, K. Lipnikov, and G. Manzini. The nonconforming virtual element method. *ESAIM: Math. Model. Numer. Anal.*, 50(3):879–904, 2016.
- [3] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131(2):267–279, 1997.
- [4] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci.*, 199(23):199–214, 2013.
- [5] L. Beirão da Veiga, K. Lipnikov, and G. Manzini. *The mimetic finite difference method for elliptic problems*, volume 11 of *MS&A. Modeling, Simulation and Applications*. Springer, Cham, 2014.
- [6] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.



- [7] D. Boffi and D. A. Di Pietro. Unified formulation and analysis of mixed and primal discontinuous skeletal methods on polytopal meshes. *ESAIM: Math. Model Numer. Anal.*, 2017. Accepted for publication. Preprint [arXiv: 1609.04601](https://arxiv.org/abs/1609.04601).
- [8] J. Bonelle, D. A. Di Pietro, and A. Ern. Low-order reconstruction operators on polyhedral meshes: application to compatible discrete operator schemes. *Computer Aided Geometric Design*, 35–36:27–41, 2015.
- [9] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896, 2005.
- [10] F. Brezzi, K. Lipnikov, and V. Simoncini. A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Mathematical Models and Methods in Applied Sciences*, 15(10), 2005.
- [11] A. Cangiani, V. Gyrya, and G. Manzini. The nonconforming virtual element method for the Stokes equations. *SIAM Journal on Numerical Analysis*, 54(6):3411–3435, 2016.
- [12] A. Cangiani, G. Manzini, and O. J. Sutton. Conforming and nonconforming virtual element methods for elliptic problems. *IMA J. Numer. Anal.*, 2016. Published online, doi: 10.1093/imanum/drw036.
- [13] B. Cockburn, D. A. Di Pietro, and A. Ern. Bridging the Hybrid High-Order and Hybridizable Discontinuous Galerkin methods. *ESAIM: Math. Model. Numer. Anal.*, 50(3):635–650, 2016.
- [14] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.
- [15] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.
- [16] D. A. Di Pietro and J. Droniou. A Hybrid High-Order method for Leray–Lions elliptic equations on general meshes. *Math. Comp.*, 86(307):2159–2191, 2017.
- [17] D. A. Di Pietro and J. Droniou.  $W^{s,p}$ -approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a Hybrid High-Order discretisation of Leray–Lions problems. *Math. Models Methods Appl. Sci.*, 27(5):879–908, 2017.
- [18] D. A. Di Pietro, J. Droniou, and A. Ern. A discontinuous-skeletal method for advection-diffusion-reaction on general meshes. *SIAM J. Numer. Anal.*, 53(5):2135–2157, 2015.
- [19] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2012.
- [20] D. A. Di Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Meth. Appl. Mech. Engrg.*, 283:1–21, 2015.
- [21] D. A. Di Pietro and A. Ern. Hybrid high-order methods for variable-diffusion problems on general meshes. *C. R. Acad. Sci. Paris, Ser. I*, 353:31–34, 2015.
- [22] D. A. Di Pietro, A. Ern, and S. Lemaire. An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Meth. Appl. Math.*, 14(4):461–472, 2014.
- [23] D. A. Di Pietro and S. Lemaire. An extension of the Crouzeix–Raviart space to general meshes with application to quasi-incompressible linear elasticity and Stokes flow. *Math. Comp.*, 84(291):1–31, 2015.
- [24] D. A. Di Pietro and R. Tittarelli. *Numerical methods for PDEs. Lectures from the fall 2016 thematic quarter at Institut Henri Poincaré*, chapter An introduction to Hybrid High-Order methods. SEMA SIMAI series. Springer, 2017. Accepted for publication. Preprint [arXiv: 1703.05136](https://arxiv.org/abs/1703.05136) [math.NA].
- [25] J. I. Diaz and F. de Thelin. On a nonlinear parabolic problem arising in some models related to turbulent flows. *SIAM J. Math. Anal.*, 25(4):1085–1111, 1994.
- [26] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105(1):35–71, 2006.
- [27] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The gradient discretisation method: A framework for the discretisation and numerical analysis of linear and nonlinear elliptic and parabolic problems*. Maths & Applications. Springer, 2017. To appear. Preprint [hal-01382358](https://hal.archives-ouvertes.fr/hal-01382358), version 3.
- [28] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Math. Models Methods Appl. Sci.*, 20(2):1–31, 2010.

- [29] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations. *Math. Models Methods Appl. Sci.*, 23(13):2395–2432, 2013.
- [30] J. Droniou, R. Eymard, and R. Herbin. Gradient schemes: Generic tools for the numerical analysis of diffusion equations. *ESAIM Mathematical Modelling and Numerical Analysis*, 50(3):749–781, 2016.
- [31] J. Droniou and B. P. Lamichhane. Gradient schemes for linear and non-linear elasticity equations. *Numer. Math.*, 129(2):251–277, 2015.
- [32] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, NY, 2004.
- [33] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces. *IMA J. Numer. Anal.*, 30(4):1009–1043, 2010.
- [34] A. Friedman. The Stefan problem in several space variables. *Trans. Amer. Math. Soc.*, 133:51–87, 1968.
- [35] R. Glowinski. *Numerical methods for nonlinear variational problems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.
- [36] R. Glowinski and J. Rappaz. Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology. *ESAIM: Math. Model Numer. Anal. (M2AN)*, 37(1):175–186, 2003.
- [37] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In R. Eymard and J.-M. Hérard, editors, *Finite Volumes for Complex Applications V*, pages 659–692. John Wiley & Sons, 2008.
- [38] C. Lehrenfeld. *Hybrid Discontinuous Galerkin methods for solving incompressible flow problems*. PhD thesis, Rheinisch-Westfälischen Technischen Hochschule Aachen, 2010.
- [39] K. Lipnikov and G. Manzini. A high-order mimetic method on unstructured polyhedral meshes for the diffusion equation. *J. Comput. Phys.*, 272:360–385, 2014.
- [40] J.-C. Nédélec. Mixed finite elements in  $\mathbf{R}^3$ . *Numer. Math.*, 35(3):315–341, 1980.
- [41] P.-A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. pages 292–315. *Lecture Notes in Math.*, Vol. 606, 1977.