Oden Institute REPORT 21-10

July 2021

Discontinuous Galerkin methods through the lens of variational multiscale analysis

by

Stein K.F. Stoter, B. Cockburn, T. J. R. Hughes, and Dominik Schillinger

Oden Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Discontinuous Galerkin methods through the lens of variational multiscale analysis

Stein K.F. Stoter[a*], B. Cockburn[b], T. J. R. Hughes[c], Dominik Schillinger[a]

[a] *Institute of Mechanics and Computational Mechanics, Leibniz Universät Hannover, Hanover, Germany*
[b] *School of Mathematics, University of Minnesota, United States of America*
[c] *Oden Institute, The University of Texas at Austin, United States of America*

Abstract

In this article, we present a theoretical framework for integrating discontinuous Galerkin methods in the variational multiscale paradigm. Our starting point is a projector-based multiscale decomposition of a generic variational formulation that uses broken Sobolev spaces and Lagrange multipliers to accommodate the non-conforming nature at the boundaries of discontinuous Galerkin elements. We show that existing discontinuous Galerkin formulations, including their penalty terms, follow immediately from a specific choice of multiscale projector. We proceed by defining the "fine-scale closure function", which captures the closure relation between the remaining fine-scale term in the discontinuous Galerkin formulation and the coarse-scale solution via a single integral expression for each basis function in the coarse-scale test space. We show that the projectors that correspond to discontinuous Galerkin methods lead to fine-scale closure functions with more compact support and smaller amplitudes compared to the fine-scale closure function of the classical (conforming) finite element method. This observation provides a new perspective on the natural stability of discontinuous Galerkin methods for hyperbolic problems, and may open the door to rigorously designed variational multiscale based fine-scale models that are suitable for DG methods.

Keywords: Variational multiscale method, discontinuous Galerkin methods, local discontinuous Galerkin method, advection-diffusion equation, fine-scale closure function, fine-scale Green's function

---

*Corresponding author;

Institute of Mechanics and Computational Mechanics, Leibniz Universität Hannover, Appelstraße 9a, 30167 Hanover, Germany.

Email addresses: `Stein.Stoter@ibnm.uni-hannover.de` (Stein K.F. Stoter[a]), `Cockburn@math.umn.edu` (B. Cockburn[b]), `Hughes@oden.utexas.edu` (T. J. R. Hughes[c]), `Schillinger@ibnm.uni-hannover.de` (Dominik Schillinger[a])

Contents

1. Introduction

Discontinuous Galerkin (DG) methods were introduced by Reed and Hill in 1973 [1]. The first theoretical analysis of their method was performed shortly after by Lesaint and Raviart in 1974 [2]. Since then, DG methods have evolved considerably and in the last decades they have challenged the position that finite volume methods have assumed in computational fluid dynamics [3, 4, 5, 6, 7]. In contrast to finite volume methods, discontinuous Galerkin methods enable higher-order accuracy on unstructured meshes [8, 9, 10], and allow efficient parallelization through hybridization [11, 12, 13]. Compared to continuous finite element methods, they show superior stability and robustness [14, 4, 15, 16, 17]. While continuous finite element methods exhibit similar conservation properties [17, 18], discontinuous Galerkin methods enable the simple and explicit derivation and modification of the flux expressions that govern element-wise conservation [16, 19, 20]. Moreover, it has been shown that DG concepts, for instance in the form of Nitsche's method, can improve the performance of continuous finite element methods. An important example is the treatment of boundary layers, where the use of Nitsche's method shows significant accuracy benefits for the imposition of no-slip boundary conditions when the boundary layer is underresolved [21, 22, 23]. Another example is the coupling of patches with DG formulations to enable advanced discretization techniques on unstructured meshes (e.g., isogeometric analysis [24, 25, 26, 27], structure preserving bases/mimetic methods [28, 29, 30, 31], and reduced basis methods [32, 33]).

In any numerical scheme, a turbulence model is required to accurately approximate solutions to the Navier-Stokes equations in the sense of large eddy simulation. Turbulence models represent some form of fine-scale closure, accounting for the effect of the scales that cannot be resolved by the discrete solution. Many turbulence models that had originally been designed for finite volume methods have also been adopted in DG research codes [7, 34, 35, 36, 37, 38, 39, 40]. In contrast to many of the turbulence models adopted from earlier finite volume methods, the variational multiscale (VMS) method offers a finite element specific, mathematically rigorous approach to large eddy simulation [41, 42, 43, 44, 45, 46]. The availability of approximation spaces as discrete subspaces of Sobolev spaces permits the filtering operation in large eddy simulation to be defined as a projection operation. Thereby, the fine-scale solution and the approximation error become one and the same. This unifies the design of the closure model and the design of the finite element formulation. The fine-scale models can be derived from the fine-scale systems of equations, either via analytical approximation [43, 47, 48, 49] or computational approximation [50, 51, 52], and are free of phenomenological parameters.

Along the same lines, the VMS paradigm has shed light on the origins of stabilized methods [53, 54, 55]. Classical residual-based stabilization terms can be interpreted as homogenized fine-scale effects through averaging of the fine-scale Green's functions [56, 57], and new stabilized methods have been designed as approximate fine-scale closures (the approaches presented in [58, 59, 50] are of particular interest due to their DG related focus). This illustrates that even for linear systems that arise in fluid mechanics (e.g., the advection-diffusion equation [60, 54], the Darcy equations [61, 62], the Stokes equations [63, 54], and the Oseen equations [64, 65, 66]) it is the unresolved fine-scale nature of the true solution

that is of key importance for the stability of finite element schemes.

Despite the prominent roles that discontinuous Galerkin methods and variational multiscale analysis have assumed in the design of efficient approximation methods for fluid mechanics problems, a comprehensive unifying framework is still missing. This can be attributed to a number of apparent conflicts between the DG and VMS concepts:

1. The foremost aspect that disconnects the two concepts is that DG methods are "nonconforming" with respect to traditional weak formulations. This means that the approximation spaces are not subspaces of the Sobolev spaces relevant to those weak formulations. The classical VMS method, however, relies on a decomposition of the infinite dimensional spaces into a finite dimensional coarse-scale (approximation) space and a fine-scale complement.

2. The numerical fluxes in discontinuous Galerkin formulations involve operators that are unbounded in the corresponding Sobolev spaces. Examples are penalties on jumps of functions in L^2 and penalties on jumps of normal gradients of functions in H^1. While these operations are permitted in the finite dimensional DG approximation spaces, the corresponding trace operators are unbounded in the complete Sobolev spaces. It is unclear how these operators would occur in the coarse-scale formulation when they are not permitted in the full-scale weak formulation.

3. Historically, the variational multiscale method is strongly linked to *residual based* stabilized methods. Any well-developed discontinuous Galerkin method, however, is stable by design [14], which may seem to eliminate the need for VMS analysis. Still, the relation between DG methods and residual-based stabilization has been explored from various angles, see e.g., [67, 14, 4]. We thus argue that the variational multiscale method could offer a valuable perspective on the stabilization mechanism in DG formulations, and, looking ahead, be further developed into an indispensable framework for the rigorous design of appropriate turbulence models.

4. Variational multiscale methods rely on a projector that defines the coarse-scale approximation. Often, the H_0^1-projector is used, as it exhibits favorable properties in terms of solution behavior and scale interaction [57]. However, since this projection operator only involves gradients, it is not a well-posed mapping when DG approximation spaces are involved: any collection of piecewise constants is a zero-energy mode.

5. Classical variational multiscale methods result in residual-based closure models that hinge on sufficient continuity of the finite element basis [68]. To make the theory applicable for conventional $\mathcal{C}^0$-continuous finite element basis functions, the fine scales must incorporate inter-element Dirac layers [53, 69] or mimic their effect [70]. To avoid the associated computational and implementation efforts, the fine scales are typically assumed to vanish on element boundaries [54, 71, 55, 43]. This assumption cannot hold in the case of DG approximations, where solutions jump by design [72].

6. Discontinuous Galerkin methods involve penalty parameters, which, at first sight, seems in contradiction with variational multiscale methods and their intrinsically defined parameters. In [73, 69, 74, 75], such penalty terms are interpreted from the perspective of a fine-scale model. The fine-scale model should follow from an underlying scale decomposition. That scale decomposition lies at the core of the variational multiscale method through the projection operator. This thus implies a more involved relation between different discontinuous Galerkin methods and projectors.

In this article we pursue two objectives: firstly, we aim to produce a rigorous derivation of discontinuous Galerkin methods following the variational multiscale principles (tackling points 1, 2, 4 and 6). Secondly, we develop the "fine-scale closure function", strongly related to the "fine-scale Green's function" from [57], which encodes the exact influence of the remaining fine-scale terms on the finite element formulation for a given DG formulation (relating to points 3, 5 and 6). As an illustrative example, we make use of the reduced order form of the advection–diffusion equation as a model problem and we derive the local discontinuous Galerkin (LDG) formulation [76] from our VMS perspective.

Our strategy for achieving these two objective is the following. In Section 2, we pose a domain decomposed form of the partial differential equation. This starting point makes DG approximation spaces conforming (point 1). The patches (elements) are coupled with double valued Lagrange multiplier fields. These Lagrange multiplier fields provide access to trace quantities from either side of an element interface that would otherwise require unbounded operators (point 2). Next, in Section 3, we perform a direct sum decomposition of the Cartesian-product space into coarse scales and fine scales. A direct sum decomposition implies a projection operator and vice versa. In Section 4, we propose a carefully constructed decomposition projector that eliminates the Lagrange multiplier field from the coarse-scale approximation space, and eliminates the majority of the fine-scale occurrences from the coarse-scale equations (point 4). By leveraging the Lagrange multiplier field in the decomposition projector, we are able to introduce penalty terms in the coarse-scale equation that would be unbounded in the full-scale equation (points 2 and 5). The result is a discontinuous Galerkin formulation derived completely from the perspective of the variational multiscale method. Based on this derivation we know the exact remaining fine-scale terms in the DG method. We are interested in the magnitude of the fine-scale effects, and the extent to which the scale interaction localizes. Both these properties can be inferred from the fine-scale closure function, introduced in Section 5. This function is projector dependent and thus different for each DG formulation. Understanding the relation between the parameters in DG formulations and the properties of the scale interaction can guide the design of fine-scale models that are suitable for DG finite element frameworks (points 3 and 5). Within the scope of this article, we compute the fine-scale closure function in Section 6 for various cases to quantify the sensitivity of the fine-scale influence in the LDG method for different parameter choices, and we compare fine-scale closure functions with those corresponding to a conforming mixed method.

2. A well-posed mixed formulation with broken spaces

The variational multiscale paradigm bridges the gap between a mathematical model of a physical problem set in Sobolev spaces, and the finite element approximation thereof. To apply this strategy for discontinuous Galerkin methods, we must first determine the appropriate weak statement as our point of departure. Consider the following reduced order form of the advection–diffusion problem:

$$\kappa^{-1}\boldsymbol{a}\phi - \nabla\phi = \kappa^{-1}\boldsymbol{\sigma} \qquad \text{in} \quad \Omega\,, \tag{1a}$$

$$\nabla \cdot \boldsymbol{\sigma} = f \qquad \text{in} \quad \Omega\,, \tag{1b}$$

$$\phi = \phi_D \qquad \text{on} \quad \partial\Omega_D\,, \tag{1c}$$

$$-\boldsymbol{\sigma} \cdot \boldsymbol{n} = g_N = 0 \qquad \text{on} \quad \partial\Omega_N \subset \partial\Omega^-\,, \tag{1d}$$

$$\boldsymbol{a} \cdot \boldsymbol{n}\,\phi - \boldsymbol{\sigma} \cdot \boldsymbol{n} = g_R \qquad \text{on} \quad \partial\Omega_R \subset \partial\Omega^+\,, \tag{1e}$$

where the auxiliary variable $\boldsymbol{\sigma}$ is introduced as the total flux. We assume a strictly positive diffusion coefficient ($\kappa > 0$) and a solenoidal advective field ($\nabla \cdot \boldsymbol{a} = 0$). The superscripts $+$ and $-$ indicate outflow ($\boldsymbol{a} \cdot \boldsymbol{n} > 0$) and inflow ($\boldsymbol{a} \cdot \boldsymbol{n} \leq 0$) parts of the boundary, respectively. The Dirichlet condition of Eq. (1c) is a natural condition in the mixed formulation that follows. Equation (1e) is an essential condition and is chosen to be homogeneous for ease of presentation. Equation (1d) is a mixed (Robin) condition and may only be specified on the outflow part of the boundary. Regarding the data, we require $f \in L^2(\Omega)$, $\phi_D \in H^{1/2}(\partial\Omega_D)$ and $g_R \in H^{-1/2}(\partial\Omega_R)$.

An equivalent mixed formulation reads:

$$\text{Find } \phi, \boldsymbol{\sigma} \in L^2(\Omega) \times \boldsymbol{H}_0(\text{div}, \Omega) \text{ s.t. } \forall\, w, \boldsymbol{\tau} \in L^2(\Omega) \times \boldsymbol{H}_0(\text{div}, \Omega):$$

$$\begin{aligned}
\left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{\sigma}\right)_\Omega &- \left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{a}\phi\right)_\Omega - \left(\nabla \cdot \boldsymbol{\tau}, \phi\right)_\Omega + \left\langle\boldsymbol{\tau} \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}\boldsymbol{\sigma} \cdot \boldsymbol{n}\right\rangle_{\partial\Omega_R} \\
&= -\left\langle\boldsymbol{\tau} \cdot \boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} - \left\langle\boldsymbol{\tau} \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R},
\end{aligned} \tag{2a}$$

$$\left(w, \nabla \cdot \boldsymbol{\sigma}\right)_\Omega = \left(w, f\right)_\Omega, \tag{2b}$$

where $(\cdot, \cdot)_\Omega = (\cdot, \cdot)_{L^2(\Omega)}$ denotes the L^2-inner product on Ω and $\left\langle\cdot, \cdot\right\rangle_{\partial\Omega}$ is the $H^{-1/2}(\partial\Omega) \times H^{1/2}(\partial\Omega) \to \mathbb{R}$ duality pairing. The space $\boldsymbol{H}_0(\text{div}, \Omega)$ is defined as $\{\boldsymbol{\tau} \in \boldsymbol{H}(\text{div}, \Omega) : \boldsymbol{\tau} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega_N\}$. This weak formulation is well-posed in the sense that it permits a unique solution $(\phi, \boldsymbol{\sigma})$ for all possible right-hand-side data and that this solution depends continuously on the data [77, 78]. Given these critical properties we can use the variational multiscale method to derive finite element formulations.

In regards to the saddle-point structure of the formulation, finite element formulations based on Eq. (2) would either require compatible discretization [79, 80] or addition of stabilization terms [63, 61]. In both cases, conforming finite element approximation spaces are usually chosen and they thus fall naturally within the variational multiscale school of thought. This is different for DG methods based on Eq. (2), where the approximation space for $\boldsymbol{\sigma}$ is not $\boldsymbol{H}(\text{div})$ conforming. This concerns the first challenge raised in the introduction. To overcome this challenge, Eq. (2) needs to be modified.

In order to do so, we decompose the domain Ω into a collection of open subdomains $\mathcal{T} = \{K_i\}_{i=1}^n$. These will represent the elements of the mesh later on. We define the union of open subdomains $\tilde{\Omega}$ and their interface skeleton Γ:

$$\tilde{\Omega} := \bigcup_{i=1}^n K_i \,, \tag{3a}$$

$$\Gamma := \bigcup_{i=1}^n \partial K_i \,. \tag{3b}$$

An integral over $\tilde{\Omega}$ is defined as the sum of the integrals over the subdomains.

We denote the interior part of the interface skeleton as Γ_0, the interior interfaces plus the Dirichlet boundary as Γ_D and the interior interfaces plus the inflow part of the Neumann boundary as Γ_N:

$$\Gamma_0 := \Gamma \setminus \partial\Omega \,, \tag{4a}$$

$$\Gamma_D := \Gamma_0 \cup \partial\Omega_D \,, \tag{4b}$$

$$\Gamma_N := \Gamma_0 \cup \partial\Omega_N \,, \tag{4c}$$

If we now define a broken variant of the space $\boldsymbol{H}(\mathrm{div}, \Omega)$ as:

$$\boldsymbol{H}(\mathrm{div}, \tilde{\Omega}) := \{\boldsymbol{\sigma} \in [L^2(\Omega)]^d : \boldsymbol{\sigma}\big|_K \in \boldsymbol{H}(\mathrm{div}, K) \;\forall\, K \in \mathcal{T}\} \,, \tag{5}$$

then we can interpret $\boldsymbol{H}_0(\mathrm{div}, \Omega)$ as a constrained subspace of $\boldsymbol{H}(\mathrm{div}, \tilde{\Omega})$. The constraints are continuity conditions imposed on the interior interfaces and a homogeneous boundary condition on $\partial\Omega_N$:

$$\boldsymbol{H}(\mathrm{div}, \Omega) = \{\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, \tilde{\Omega}) : [\![\boldsymbol{\sigma}]\!] = 0 \text{ on } \Gamma_N\} \,, \tag{6}$$

where, as is customary in the literature on DG methods, the jump is defined as:

$$[\![\boldsymbol{\sigma}]\!] := \begin{cases} \boldsymbol{\sigma}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{\sigma}^- \cdot \boldsymbol{n}^- & \text{on } \Gamma_0 \,, \\ \boldsymbol{\sigma}^- \cdot \boldsymbol{n}^- & \text{on } \partial\Omega \,. \end{cases} \tag{7}$$

In this definition, the superscripts $+$ and $-$ denote arbitrary sides of an interface. Similarly, we define the jump of a scalar as:

$$[\![\phi]\!] := \begin{cases} \phi^+ \boldsymbol{n}^+ + \phi^- \boldsymbol{n}^- & \text{on } \Gamma_0 \,, \\ \phi^- \boldsymbol{n}^- & \text{on } \partial\Omega \,. \end{cases} \tag{8}$$

and average of either a scalar or a vector as:

$$\{\!\{\phi\}\!\} := \begin{cases} \frac{1}{2}(\phi^+ + \phi^-) & \text{on } \Gamma_0 \,, \\ \phi^- & \text{on } \partial\Omega \,. \end{cases} \tag{9}$$

The constraints that are enforced strongly in $\boldsymbol{H}(\operatorname{div}, \Omega)$ according to Eq. (6) may alternatively be imposed weakly with Lagrange multipliers [81, 82]. This results in the following weak form:

Find $\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+ \in U$ s.t. $\forall w, \boldsymbol{\tau}, q^-, q^+ \in U$:

$$
\begin{aligned}
\left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{\sigma}\right)_{\tilde{\Omega}} - \left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{a}\phi\right)_{\tilde{\Omega}} - \left(\nabla \cdot \boldsymbol{\tau}, \phi\right)_{\tilde{\Omega}} + \left\langle \boldsymbol{\tau} \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}\boldsymbol{\sigma} \cdot \boldsymbol{n}\right\rangle_{\partial\Omega_R} & \\
+ \left\langle [\![\boldsymbol{\tau}]\!], \{\!\{\lambda\}\!\}\right\rangle_{\Gamma_N} = -\left\langle \boldsymbol{\tau} \cdot \boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} - \left\langle \boldsymbol{\tau} \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R}, &
\end{aligned}
\tag{10a}
$$

$$
\left(w, \nabla \cdot \boldsymbol{\sigma}\right)_{\tilde{\Omega}} = \left(w, f\right)_{\tilde{\Omega}},
\tag{10b}
$$

$$
\left\langle [\![\boldsymbol{\sigma}]\!], \{\!\{q\}\!\}\right\rangle_{\Gamma_N} = 0,
\tag{10c}
$$

$$
\left([\![\lambda]\!], [\![q]\!]\right)_{H^{1/2}(\Gamma_D)} = \left(\phi_D, q^-\right)_{H^{1/2}(\partial\Omega_D)}.
\tag{10d}
$$

The suitable function spaces may be identified as:

$$
U = \mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0 = L^2(\Omega) \times \boldsymbol{H}(\operatorname{div}, \tilde{\Omega}) \times H^{1/2}(\Gamma \setminus \partial\Omega_R) \times H^{1/2}(\Gamma_0).
\tag{11a}
$$

The weak formulation of Eq. (10) is well-posed and it is equivalent to the original statement of Eq. (1). Different from the earlier weak formulation, discontinuous Galerkin approximation spaces are now conforming subspaces.

Remark 2.1. *Equation* (10d) *immediately implies that* $[\![\lambda]\!] = 0$ *on* Γ_0, *and thus that* $\lambda^- = \lambda^+ =: \lambda$. *The reason for introducing a double valued Lagrange multiplier will become clear in Section 4.*

Remark 2.2. *From Eq.* (1) *it follows that the solution* ϕ *lies in* $H^1(\Omega)$ *and the equivalence of Eq.* (10) *and Eq.* (1) *then dictates that the Lagrange multiplier will satisfy* $\lambda = \phi$. *The latter statement hinges on the former as it requires a well-defined and single-valued trace of* ϕ *on* Γ_N.

3. Scale decomposition

The objective of a variational multiscale formulation is to find the discrete solution that is the "optimal" approximation of the true solution based on some definition of optimality. This is defined to be the coarse-scale solution. The choice of optimality condition defines a mapping from every pair $(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+)$ to a coarse-scale solution pair $(\phi^h, \boldsymbol{\sigma}^h, \lambda^{-h}, \lambda^{+h})$ in the finite dimensional coarse-scale space $U^h = (\mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0)^h$:

$$
\begin{aligned}
\mathscr{P}: \quad U \quad &\to \quad U^h := \operatorname{ran}(\mathscr{P}) \\
(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+) &\mapsto \quad \mathscr{P}(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+).
\end{aligned}
\tag{12}
$$

Any true solution that lies in the coarse-scale solution space should map to itself, making $\mathscr{P}$ an idempotent mapping and hence a projector:

$$
\mathscr{P}(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+) = (\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+) \quad \forall \phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+ \in U^h
\tag{13}
$$

8

The coarse-scale and fine-scale components of ϕ, $\boldsymbol{\sigma}$, λ^- and λ^+ are then defined as:

$$(\phi^h, \boldsymbol{\sigma}^h, \lambda^{-h}, \lambda^{+h}) := \mathscr{P}(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+), \tag{14a}$$

$$(\phi', \boldsymbol{\sigma}', \lambda^{-\prime}, \lambda^{+\prime}) := (\mathscr{I} - \mathscr{P})(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+), \tag{14b}$$

with $\mathscr{I}$ the identity operator $\mathscr{I}(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+) = (\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+)$.

If the projector $\mathscr{P}$ is linear, then it defines a direct sum decomposition of the original mixed space:

$$U = \mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0 = (\mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0)^h \oplus (\mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0)' = U^h \oplus U' \tag{15}$$

It follows from Eq. (14b) that the range of $\mathscr{I} - \mathscr{P}$, or the kernel of $\mathscr{P}$, is the fine-scale space:

$$(\phi', \boldsymbol{\sigma}', \lambda^{-\prime}, \lambda^{+\prime}) \in \ker \mathscr{P} =: (\mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0)' = U'. \tag{16}$$

We note that the coarse and fine-scale spaces do not necessarily decompose as Cartesian products of coarse-scale respectively fine-scale spaces. This means that the different subequations in Eq. (10) can no longer be guaranteed to decouple as such. Instead, the general scale decomposition of the weak formulation reads:

Find $\phi^h, \boldsymbol{\sigma}^h, \lambda^{-h}, \lambda^{+h} \in U^h$ and $\phi', \boldsymbol{\sigma}', \lambda^{-\prime}, \lambda^{+\prime} \in U'$

$\quad$ s.t. $\forall\, w^h, \boldsymbol{\tau}^h, q^{-h}, q^{+h} \in U^h$ and $\forall\, w', \boldsymbol{\tau}', q^{-\prime}, q^{+\prime} \in U'$:

$$\left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h + \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} - \left(\kappa^{-1}\boldsymbol{a}\cdot\boldsymbol{\tau}^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle\boldsymbol{\tau}^h\cdot\boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}(\boldsymbol{\sigma}^h\cdot\boldsymbol{n} + \boldsymbol{\sigma}'\cdot\boldsymbol{n})\right\rangle_{\partial\Omega_R}$$
$$- \left(\nabla\cdot\boldsymbol{\tau}^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle[\![\boldsymbol{\tau}^h]\!], \{\!\{\lambda'\}\!\} + \{\!\{\lambda^h\}\!\}\right\rangle_{\Gamma_N} - \left(w^h, \nabla\cdot\boldsymbol{\sigma}^h + \nabla\cdot\boldsymbol{\sigma}'\right)_{\tilde{\Omega}}$$
$$+ \left\langle[\![\boldsymbol{\sigma}^h + \boldsymbol{\sigma}']\!], \{\!\{q^h\}\!\}\right\rangle_{\Gamma_N} + \left([\![\lambda^h]\!] + [\![\lambda']\!], [\![q^h]\!]\right)_{H^{1/2}(\Gamma_D)} = -\left\langle\boldsymbol{\tau}^h\cdot\boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} \tag{17a}$$
$$- \left\langle\boldsymbol{\tau}^h\cdot\boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R} - \left(w^h, f\right)_{\tilde{\Omega}} + \left(\phi_D, q^{-h}\right)_{H^{1/2}(\partial\Omega_D)},$$

$$\left(\kappa^{-1}\boldsymbol{\tau}', \boldsymbol{\sigma}^h + \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} - \left(\kappa^{-1}\boldsymbol{a}\cdot\boldsymbol{\tau}', \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle\boldsymbol{\tau}'\cdot\boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}(\boldsymbol{\sigma}^h\cdot\boldsymbol{n} + \boldsymbol{\sigma}'\cdot\boldsymbol{n})\right\rangle_{\partial\Omega_R}$$
$$- \left(\nabla\cdot\boldsymbol{\tau}', \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle[\![\boldsymbol{\tau}']\!], \{\!\{\lambda'\}\!\} + \{\!\{\lambda^h\}\!\}\right\rangle_{\Gamma_N} - \left(w', \nabla\cdot\boldsymbol{\sigma}^h + \nabla\cdot\boldsymbol{\sigma}'\right)_{\tilde{\Omega}}$$
$$\left\langle[\![\boldsymbol{\sigma}^h + \boldsymbol{\sigma}']\!], \{\!\{q'\}\!\}\right\rangle_{\Gamma_N} + \left([\![\lambda^h]\!] + [\![\lambda']\!], [\![q']\!]\right)_{H^{1/2}(\Gamma_D)} = -\left\langle\boldsymbol{\tau}'\cdot\boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} \tag{17b}$$
$$- \left\langle\boldsymbol{\tau}'\cdot\boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R} - \left(w', f\right)_{\tilde{\Omega}} + \left(\phi_D, q^{-\prime}\right)_{H^{1/2}(\partial\Omega_D)},$$

The eventual finite element formulations correspond to the weak formulations with the coarse-scale test functions (i.e., Eq. (17a)). These equations still involve fine scales, for which some form of closure needs to be substituted. Different closure models then encompass turbulence models [43], residual-based stabilized methods [53], residual-free bubble stabilized methods [56], and, as we intend to show in this article, also discontinuous Galerkin methods.

4. Discontinuous Galerkin methods as projectors

Discontinuous Galerkin formulations of Eq. (1) compute approximations $\boldsymbol{\sigma}^h$ and ϕ^h in the discrete spaces:

$$\mathcal{W}^h = \{w \in L^2(\Omega) : w|_K \in \mathbb{P}^p(K) \ \forall K \in \mathcal{T}\} \subset \mathcal{W}, \tag{18a}$$

$$\boldsymbol{\Sigma}^h = \{\boldsymbol{\sigma} \in [L^2(\Omega)]^d : \boldsymbol{\sigma}|_K \in [\mathbb{P}^q(K)]^d \ \forall K \in \mathcal{T}\} \subset \boldsymbol{\Sigma}, \tag{18b}$$

The discontinuous Galerkin methods that we focus on in this article do not compute an approximation to the Lagrange multiplier. The size of our finite element system of equation thus corresponds to the dimensionality of $\mathcal{W}^h \times \Sigma^h$. Since the finite element formulation should follow from the coarse-scale equation the coarse-scale space must satisfy $\dim U^h = \dim(\mathcal{W}^h \times \Sigma^h)$, which is a condition on the range of the chosen projector. This requirement is satisfied by projectors of the following form, the motivation for which will follow shortly:

$$\mathscr{P}(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+) = (\phi^h, \boldsymbol{\sigma}^h, \phi^{h-}, \phi^{h+}). \tag{19}$$

Only $(\phi^h, \boldsymbol{\sigma}^h)$ still needs to be defined. As addressed in point four of the list of challenges from the introduction, optimality in the standard H^1-seminorm no longer induces an applicable projector. Consider instead the following operation:

$$\phi^h, \boldsymbol{\sigma}^h = \arg\inf_{\boldsymbol{\sigma}^h \in \boldsymbol{\Sigma}^h} \arg\sup_{\phi^h \in \mathcal{W}^h} \frac{1}{2}\big(\kappa^{-1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \boldsymbol{\sigma} - \boldsymbol{\sigma}^h\big)_{\tilde{\Omega}} - \big(\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h), \phi - \phi^h\big)_{\tilde{\Omega}}$$

$$+ \frac{1}{2}\big\langle \boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{\sigma}^h \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}(\boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{\sigma}^h \cdot \boldsymbol{n})\big\rangle_{\partial\Omega_R} + \big\langle [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!], \{\!\{\lambda - \phi^h\}\!\}\big\rangle_{\Gamma_N}$$

$$+ \big\langle [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!], \boldsymbol{\beta} \cdot [\![\lambda - \phi^h]\!]\big\rangle_{\Gamma_0} + \frac{1}{2}\big\langle [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!], \kappa^{-1}C[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!]\big\rangle_{\Gamma_0} + \frac{1}{2}\big\langle \kappa\eta[\![\lambda - \phi^h]\!], [\![\lambda - \phi^h]\!]\big\rangle_{\Gamma_D}. \tag{20}$$

Well-posedness of this saddle-point problem relies on the appropriate choice of operators η, $\boldsymbol{\beta}$ and C [19, 76, 83, 4, 84]. Of course, this choice affects the definition of the coarse-scale solution for any given true solution. The qualitative impact of these parameters on the coarse-scale solution has, for instance, been studied in [84].

The first three terms in Eq. (20) relate to the potential for which the saddle point defines a mixed formulation of a Poisson problem. Their minimization produces optimal results if one were to use compatible discretization, e.g., with Raviart-Thomas or Brezzi-Douglas-Marini elements for the auxiliary solution field [79, 78]. The pressure-stabilizing Petrov-Galerkin (PSPG) and least-squares on incompressibility constraint (LSIC) stabilization methods required to stabilize equal order interpolation pairs would follow from the addition of volumetric weightings [63]. Now that we wish to make use of DG coarse-scale spaces, we have added weighted interface terms.

Note that the Lagrange multiplier $\lambda^\pm$ takes the role of $\phi^\pm$ in the boundary terms in Eq. (20). This is a subtle but important detail and relates to the second point raised in the introduction. The trace operator is unbounded when its domain is all of $L^2(\tilde{\Omega})$. It can not be used in the definition of the projector since the direct sum decomposition in Eq. (15)

requires that the domain of the projector is all of U. While we solve this issue by using $\lambda^{\pm}$ in the definition of $(\phi^h, \boldsymbol{\sigma}^h)$, we now appear to loose the idempotency property of the projector. An idempotent operator is retrieved by our definition of the coarse-scale Lagrange multipliers as $\lambda^{\pm,h} = \phi^{h,\pm}$ in Eq. (19). This choice, in turn, is possible due to the use of a *double valued* Lagrange multiplier in Eq. (10).

Solution pairs $(\phi^h, \boldsymbol{\sigma}^h)$ are equilibrium points of the functional in Eq. (20). They may be determined by simultaneously equating to zero the first variations with respect to $\boldsymbol{\sigma}^h$ and ϕ^h. We thereby obtain the following two optimality conditions which have to be satisfied for all $\boldsymbol{\nu}^h \in \boldsymbol{\Sigma}^h$ and all $v^h \in \mathcal{W}^h$:

$$
\begin{aligned}
&- \left(\kappa^{-1}\boldsymbol{\nu}^h, \boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left(\nabla \cdot \boldsymbol{\nu}^h, \phi - \phi^h\right)_{\tilde{\Omega}} - \left\langle \boldsymbol{\nu}^h \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}(\boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{\sigma}^h \cdot \boldsymbol{n})\right\rangle_{\partial\Omega_R} \\
&\quad - \left\langle [\![\boldsymbol{\nu}^h]\!], \{\!\!\{\lambda - \phi^h\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\nu}^h]\!], \boldsymbol{\beta} \cdot [\![\lambda - \phi^h]\!]\right\rangle_{\Gamma_0} - \left\langle [\![\boldsymbol{\nu}^h]\!], \kappa^{-1}C[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!]\right\rangle_{\Gamma_0} = 0 \,,
\end{aligned}
\tag{21a}
$$

$$
\begin{aligned}
&\left(v^h, \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^h)\right)_{\tilde{\Omega}} - \left\langle [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!], \{\!\!\{v^h\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!], \boldsymbol{\beta} \cdot [\![v^h]\!]\right\rangle_{\Gamma_0} \\
&\quad - \left\langle \kappa\eta[\![v^h]\!], [\![\lambda - \phi^h]\!]\right\rangle_{\Gamma_D} = 0 \,,
\end{aligned}
\tag{21b}
$$

or, by replacing all occurrences of $\boldsymbol{\sigma} - \boldsymbol{\sigma}^h$, $\phi - \phi^h$, $[\![\lambda - \phi^h]\!]$ and $\{\!\!\{\lambda - \phi^h\}\!\!\}$ by $\boldsymbol{\sigma}'$, ϕ', $[\![\lambda']\!]$ and $\{\!\!\{\lambda'\}\!\!\}$ respectively:

$$
\begin{aligned}
&- \left(\kappa^{-1}\boldsymbol{\nu}^h, \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} + \left(\nabla \cdot \boldsymbol{\nu}^h, \phi'\right)_{\tilde{\Omega}} - \left\langle [\![\boldsymbol{\nu}^h]\!], \{\!\!\{\lambda'\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle \boldsymbol{\nu}^h \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}\boldsymbol{\sigma}' \cdot \boldsymbol{n}\right\rangle_{\partial\Omega_R} \\
&\quad - \left\langle [\![\boldsymbol{\nu}^h]\!], \boldsymbol{\beta} \cdot [\![\lambda']\!]\right\rangle_{\Gamma_0} - \left\langle [\![\boldsymbol{\nu}^h]\!], \kappa^{-1}C[\![\boldsymbol{\sigma}']\!]\right\rangle_{\Gamma_0} = 0 \,,
\end{aligned}
\tag{22a}
$$

$$
\left(v^h, \nabla \cdot \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} - \left\langle [\![\boldsymbol{\sigma}']\!], \{\!\!\{v^h\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\sigma}']\!], \boldsymbol{\beta} \cdot [\![v^h]\!]\right\rangle_{\Gamma_0} - \left\langle \kappa\eta[\![v^h]\!], [\![\lambda']\!]\right\rangle_{\Gamma_D} = 0 \,.
\tag{22b}
$$

which represents the fine-scale constraint that defines all the functions in $\ker(\mathscr{P})$.

The properties of the chosen projector, i.e., $\lambda^{-h} = \phi^{h-}$, $\lambda^{+h} = \phi^{h+}$, $q^{+h} = w^{h+}$ and $q^{-h} = w^{h-}$ from Eq. (19) and the fine-scale constraints of Eq. (22), may directly be used as a partial closure in the coarse-scale problem. By substituting all these identities into Eq. (17a), we obtain:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ s.t. $\forall w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$
\begin{aligned}
&\left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} - \left(\kappa^{-1}\boldsymbol{a} \cdot \boldsymbol{\tau}^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}\boldsymbol{\sigma}^h \cdot \boldsymbol{n}\right\rangle_{\partial\Omega_R} \\
&\quad - \left(\nabla \cdot \boldsymbol{\tau}^h, \phi^h\right)_{\tilde{\Omega}} + \left\langle [\![\boldsymbol{\tau}^h]\!], \{\!\!\{\phi^h\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\tau}^h]\!], \boldsymbol{\beta} \cdot [\![\lambda']\!]\right\rangle_{\Gamma_0} - \left\langle [\![\boldsymbol{\tau}^h]\!], \kappa^{-1}C[\![\boldsymbol{\sigma}']\!]\right\rangle_{\Gamma_0} \\
&\qquad = -\left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} - \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, (\boldsymbol{a} \cdot \boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R},
\end{aligned}
\tag{23a}
$$

$$
\begin{aligned}
&- \left(w^h, \nabla \cdot \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left\langle [\![\boldsymbol{\sigma}^h]\!], \{\!\!\{w^h\}\!\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\sigma}']\!], \boldsymbol{\beta} \cdot [\![w^h]\!]\right\rangle_{\Gamma_0} - \left\langle \kappa\eta[\![w^h]\!], [\![\lambda']\!]\right\rangle_{\Gamma_D} \\
&\quad + \left([\![\phi^h]\!] + [\![\lambda']\!], [\![w^h]\!]\right)_{H^{1/2}(\Gamma_D)} = -\left(w^h, f\right)_{\tilde{\Omega}} + \left(\phi_D, w^h\right)_{H^{1/2}(\partial\Omega_D)} \,.
\end{aligned}
\tag{23b}
$$

Many of the remaining fine scales can still be related to coarse scales by making use of known solution behavior. From the original problem formulation, Eq. (1), we know that the true solutions ϕ and $\boldsymbol{\sigma}$ will lie in $H^1(\Omega)$ and $\boldsymbol{H}(\mathrm{div}, \Omega)$ respectively, and from Remark 2.2 we know $\lambda = \phi$. As a result, the true solution satisfies $[\![\lambda']\!] = -[\![\phi^h]\!]$ and $[\![\boldsymbol{\sigma}']\!] = -[\![\boldsymbol{\sigma}^h]\!]$ on Γ_0. Aditionally, from the imposed boundary conditions we know that the

true solution satisfies $\lambda' = \phi_D - \phi^h$ on the Dirichlet boundary and $\boldsymbol{\sigma}' \cdot \boldsymbol{n} = -\boldsymbol{\sigma}^h \cdot \boldsymbol{n}$ on the (homogeneous) Neumann boundary. By making use of this known solution behavior, we arrive at the following weak formulation for the coarse-scale solution:

$$\text{Find } \phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h \text{ s.t. } \forall\, w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h :$$

$$\left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} - \left(\kappa^{-1}\boldsymbol{a}\cdot\boldsymbol{\tau}^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle\boldsymbol{\tau}^h \cdot \boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}\boldsymbol{\sigma}^h \cdot \boldsymbol{n}\right\rangle_{\partial\Omega_R}$$

$$- \left(\nabla\cdot\boldsymbol{\tau}^h, \phi^h\right)_{\tilde{\Omega}} + \left\langle[\![\boldsymbol{\tau}^h]\!], \{\!\{\phi^h\}\!\}\right\rangle_{\Gamma_N} + \left\langle[\![\boldsymbol{\tau}^h]\!], \boldsymbol{\beta}\cdot[\![\phi^h]\!]\right\rangle_{\Gamma_0} + \left\langle[\![\boldsymbol{\tau}^h]\!], \kappa^{-1}C[\![\boldsymbol{\sigma}^h]\!]\right\rangle_{\Gamma_0} \quad (24a)$$

$$= -\left\langle\boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D} - \left\langle\boldsymbol{\tau}^h \cdot \boldsymbol{n}, (\boldsymbol{a}\cdot\boldsymbol{n})^{-1}g_R\right\rangle_{\partial\Omega_R},$$

$$- \left(w^h, \nabla\cdot\boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left\langle[\![\boldsymbol{\sigma}^h]\!], \{\!\{w^h\}\!\}\right\rangle_{\Gamma_N} + \left\langle[\![\boldsymbol{\sigma}^h]\!], \boldsymbol{\beta}\cdot[\![w^h]\!]\right\rangle_{\Gamma_0} + \left\langle\kappa\eta\,[\![w^h]\!], [\![\phi^h]\!]\right\rangle_{\Gamma_D}$$

$$= -\left(w^h, f\right)_{\tilde{\Omega}} + \left\langle\kappa\eta\,w^h, \phi_D\right\rangle_{\partial\Omega_D}. \qquad\qquad (24b)$$

Apart from the one remaining fine-scale term, this is a discontinuous Galerkin method.

Remark 4.1. *Despite the mixed origin of the formulation, only the fine scales of the primal solution show up in Eq. (24). This fine-scale solution occurs in a single term: a weighted volumetric average. In contrast, the original decomposed form of Eq. (17) includes six fine-scale terms. From the variational multiscale perspective, the additional terms in a DG method are fine-scale closures. There are two important messages: firstly, this derivation shows that it is not appropriate to substitute fine-scale models for each of the fine-scale terms in Eq. (17) and additionally add the terms from the DG formulation. This relates to the sixth point raised in the introduction. Secondly, our approach illustrates that the decomposition projector can be designed with the objective of simplifying the remaining scale interaction (atop the usual objective of yielding suitable coarse-scale solutions). This relates to the third point in the introduction, and we explore this concept more in Section 6.3.*

Remark 4.2. *When $\boldsymbol{a} = \boldsymbol{0}$, we retrieve the general form of the local discontinuous Galerkin (LDG) formulation for a Poisson problem [76]. Of course, this is an immediate result of Eq. (20) as the chosen projector. If $(\phi, \boldsymbol{\sigma}, \lambda^-, \lambda^+)$ solves a Poisson problem, then the Euler-Lagrange equations of the functional in Eq. (20) are the LDG formulation. We have put this knowledge in a VMS derivation that applies also when the diffusive term in only part of the partial differential equation. This is similar to the use of the H_0^1-projector for conforming finite element methods in a primal formulation: optimality in the H_0^1-norm corresponds to the solution that satisfies the associated Poisson problem. Finite element methods tend to be very effective in treating elliptic equations, and this form of VMS formulation aims to maintain that effectiveness when non-elliptic terms are added.*

Remark 4.3. *Even though our focus is on LDG-type formulations, other DG methods would fit in our framework by changing either the weak statement or the projector. For example, by removing all penalty terms from the projection operator one would obtain the (unstable) first Bassi-Rebay method [85]. By subsequently adding the term $\sum_{e\in\Gamma}\int_\Omega \frac{1}{2}\eta\,r_e([\![\lambda - \phi^h]\!])^2$ (where $r_e : [L^2(e)]^d \to \boldsymbol{\Sigma}^h$ is a so-called lifting operator [19]), one retrieves the second Bassi-Rebay method [86]. The form of Eq. (24) also encompasses HDG formulations. By*

moving the advective term from Eq. (1a) to Eq. (1b) and by adding a penalty of $\frac{1}{2}|\boldsymbol{a}\cdot\boldsymbol{n}|$ to η in the projector, one would obtain upwind formulations such as the one described in [87]. Lastly, if one uses a primal formulation in place of Eq. (1), then our framework may yield the symmetric and non-symmetric interior penalty formulations and the Baumann-Oden method [69].

5. The fine-scale closure function

As a result of the previous formalism, discontinuous Galerkin formulations fall within the variational multiscale paradigm in the following general sense:

Find $u^h \in U^h$ and $u' \in U'$ s.t. $\forall w^h \in U^h$ and $\forall w' \in U'$:

$$B_{DG}(u^h, w^h) + \mathcal{S}(u', w^h) = L_{DG}(w^h),\tag{25a}$$

$$B(u', w') = L(w') - B(u^h, w') =: \mathrm{Res}_{u^h}(w'),\tag{25b}$$

where, for the mixed advection–diffusion model problem, the coarse scales are defined as $u^h := (\phi^h, \boldsymbol{\sigma}^h, \phi^{h-}, \phi^{h+}) \in U^h$, and the fine scales as $u' := (\phi', \boldsymbol{\sigma}', \lambda^{-\prime}, \lambda^{+\prime}) \in U'$. The bilinear form $B(\cdot, \cdot)$ and linear form $L(\cdot)$ correspond to those of the original variational problem (i.e., Eq. (10)). The bilinear form $B_{DG}(\cdot, \cdot)$ and linear form $L_{DG}(\cdot)$ are different and correspond to discontinuous Galerkin formulations. Compared to a standard discontinuous Galerkin formulation, the scale-interaction term $\mathcal{S}(u', w^h)$ appears as a new term. For conforming finite element formulations, the scale-interaction term $\mathcal{S}(u', w^h)$ would simply be $B(u', w^h)$. However, per the derivation of Sections 2 to 4 many fine-scale terms were eliminated from the coarse-scale equations, thereby producing the discontinuous Galerkin (bi)linear forms $B_{DG}(u^h, w^h)$ and $L_{DG}(w^h)$ in Eq. (25a). Hence, for discontinuous Galerkin methods the scale-interaction term is necessarily different from $B(u', w^h)$.

It is *only* through this new scale-interaction term that the fine scales influence the coarse scales. For linear problems, different approximations to this term lead to different stabilization methods. For the non-linear Navier-Stokes equations, this term represents the unresolved Reynolds stresses and different approximation approaches (e.g., the Boussinesq's approximation) lead to different turbulence models. However, in the variational multiscale method we can formally and explicitly define a "fine-scale closure function" which encodes the exact scale interaction. This fine-scale closure function would ideally play a central role in the design of the fine-scale model. Within the scope of this article, we investigate the fine-scale closure function to clarify the stable nature of DG methods from a VMS perspective, thereby addressing the third point raised in the introduction.

Even though the fine-scale solution appears in the scale-interaction term, we need not know it explicitly to close this term. Rather, it is the *action* that it has on the coarse-scale test functions that is of importance. This action on the coarse scales can be incorporated while at the same time eliminating the fine-scale solution by a Schur complement type approach. We write $U^h = \mathrm{span}\{w_j^h\}_{j=1}^N$, and consider the coarse-scale basis function w_i^h.

Then, we formulate an adjoint problem:

$$\text{Find } g_i' \in U' \text{ s.t. } \forall\, \xi' \in U' :$$
$$B\big(\xi', g_i'\big) = -\mathcal{S}\big(\xi', w_i^h\big). \tag{26}$$

This equation produces one fine-scale function g_i' for every coarse-scale basis function w_i^h. If we now choose $\xi' = u'$ in Eq. (26) and $w' = g_i'$ in Eq. (25b), then we obtain the following identity that relates the fine-scale solution u' in $\mathcal{S}\big(u', w_i^h\big)$ to the coarse-scale solution u^h:

$$\mathcal{S}\big(u', w_i^h\big) = -B\big(u', g_i'\big) = -\text{Res}_{u^h}(g_i') = B\big(u^h, g_i'\big) - L(g_i'). \tag{27}$$

The right-hand expression of Eq. (27) no longer depends on the fine-scale solution. It is the exact closure of the fine scales that is relevant to the DG formulation. We thus call the fine-scale function g_i' that solves Eq. (26) the fine-scale closure function corresponding to basis w_i^h.

Due to the linearity of Eq. (26), the fine-scale closure function for $w^h = \sum_i c_i w_i^h \in U^h$ becomes $g_{w^h}' = \sum_i c_i g_i'$. Then, the closed discontinuous Galerkin formulation that incorporates the exact scale interaction and produces the coarse-scale solution reads:

$$\text{Find } u^h \in U^h \text{ s.t. } \forall\, w^h \in U^h :$$
$$B_{DG}\big(u^h, w^h\big) + B\big(u^h, g_{w^h}'\big) = L_{DG}\big(w_i^h\big) + L(g_{w^h}'). \tag{28}$$

Remark 5.1. *The fine-scale closure function defined by Eq. (26) strongly relates to the "fine-scale Green's function" introduced in [57]. The fine-scale Green's function is a two-point correlation function that follows from the solution of an adjoint problem like Eq. (26), but where the linear form is induced by a Dirac-delta distribution at a variable location in the domain Ω. The resulting fine-scale closure in the coarse-scale equations involves an integration on $\Omega \times \Omega$. Differently, the fine-scale closure function is a correlation between a location in the domain Ω and a coarse-scale test function. The complete action of the fine scales on the coarse scales is given by the single integration shown in Eq. (27). Figures 1 and 2 illustrate the two different approaches for eliminating the fine scales from the coarse-scale equation.*

6. Scale interaction visualized

Since the fine-scale closure function is a member of the fine-scale function space, it depends on the projection operator. From Section 4, we know that different discontinuous Galerkin methods correspond to different projectors. As a result, each basis function in a discontinuous Galerkin method has a natural correspondence to a fine-scale closure function. In this section, we graph the fine-scale closure functions for different local discontinuous Galerkin methods and for conforming mixed methods. A large spread in the fine-scale closure function means a global impact of the fine scales on the coarse-scales, whereas a small spread means that the fine scales have a localized influence. Large values in the fine-scale closure function mean that the scale-interaction term has a significant influence,
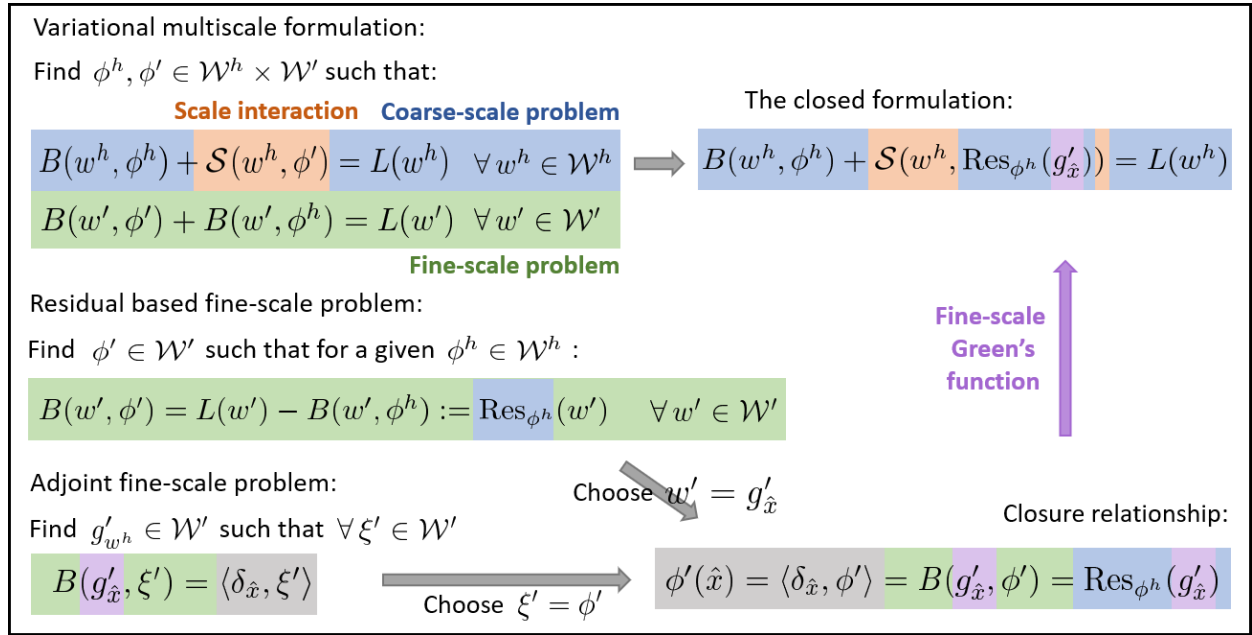
Figure 1: Eliminating the fine scales from the coarse-scale equation with the fine-scale Green's function.
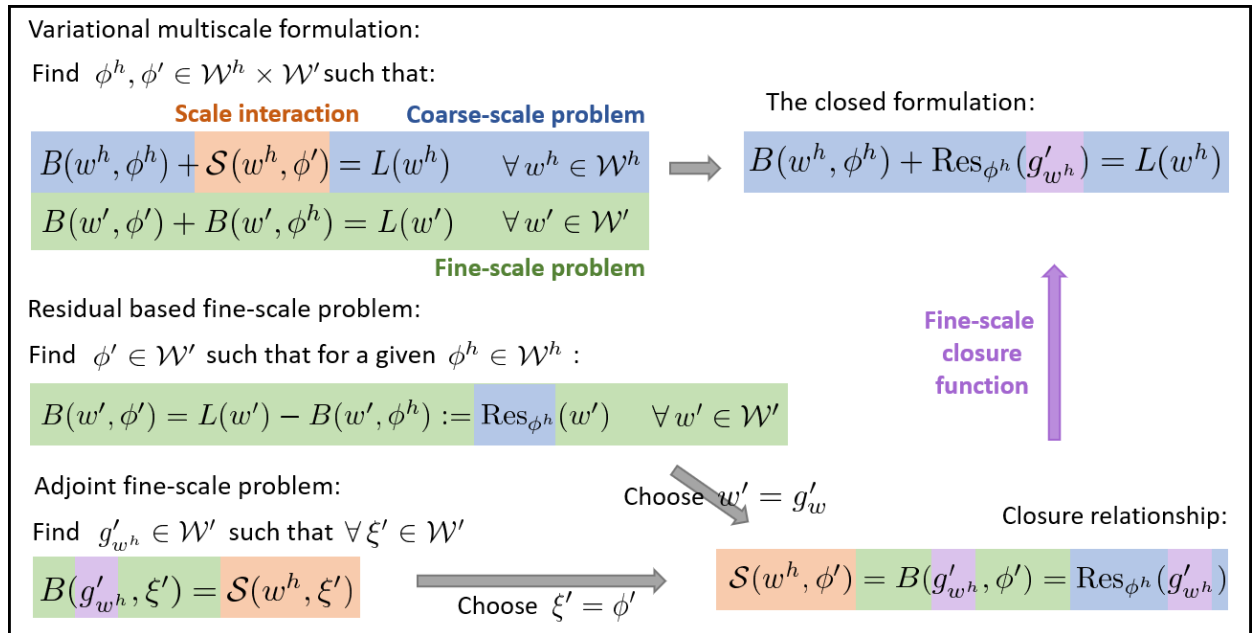

Figure 2: Eliminating the fine scales from the coarse-scale equation with the fine-scale closure function.

whereas small values imply that the finite element formulation is less sensitive to the fine-scale effects. From this perspective, a discontinuous Galerkin formulation that corresponds to a localized and moderately valued fine-scale closure function is favorable over a method with a large spread and large peak values.

We compute the fine-scale closure function by solving the adjoint problem of Eq. (26) on a heavily refined mesh. Recall that the fine-scale space U' is the subspace of $\mathcal{W} \times \mathbf{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0$ subjected to the constraints from Eq. (22). To simplify the computational implementation, we release the constraints from U' and enforce them with new Lagrange multiplier terms. This results in:

Find $g', \boldsymbol{h}', l^{-\prime}, l^{+\prime} \in \mathcal{W} \times \mathbf{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0$ and $\zeta^h, \boldsymbol{\zeta}^h \in \mathcal{W}^h \times \mathbf{\Sigma}^h$

$\quad$ s.t. $\forall \xi', \boldsymbol{\xi}', r^{-\prime}, r^{+\prime} \in \mathcal{W} \times \mathbf{\Sigma} \times \mathcal{Q} \times \mathcal{Q}_0$ and $\forall v^h, \boldsymbol{\nu}^h \in \mathcal{W}^h \times \mathbf{\Sigma}^h :$

$$B\big((\xi', \boldsymbol{\xi}', r^{-\prime}, r^{+\prime}), (g', \boldsymbol{h}', l^{-\prime}, l^{+\prime})\big) + C\big((\xi', \boldsymbol{\xi}', r^{-\prime}, r^{+\prime}), (\zeta^h, \boldsymbol{\zeta}^h)\big) = \big(\kappa^{-1}\boldsymbol{a} \cdot \boldsymbol{\tau}^h, \xi'\big)_{\tilde{\Omega}}, \quad (29a)$$

$$C\big((g', \boldsymbol{h}', l^{-\prime}, l^{+\prime}), (v^h, \boldsymbol{\nu}^h)\big) = 0, \quad (29b)$$

where the bilinear form $C\big((g', \boldsymbol{h}', l^{-\prime}, l^{+\prime}), (v^h, \boldsymbol{\nu}^h)\big)$ corresponds directly to that from Eq. (22). The new variables $(\zeta^h, \boldsymbol{\zeta}^h)$ are the Lagrange multipliers associated to those constraints. The bilinear form $B(\cdot, \cdot)$ is still that of Eq. (10).

6.1. One-dimensional verification

We begin with a verification of the fine-scale elimination procedure of Sections 4 and 5. For this purpose, we consider a one-dimensional model problem on a domain $\Omega = [0, 1]$, with a positive unit advective speed, a diffusivity $\kappa = 0.01$, a constant source function $f = 1$ and homogeneous Dirichlet conditions at both the inflow and the outflow boundary. The solution $(\phi, \boldsymbol{\sigma}, \lambda^+, \lambda^-)$ to the mixed formulation of Eq. (10) is:

$$\phi = \frac{f}{\|\boldsymbol{a}\|}\Big(x - \frac{e^{\frac{\|\boldsymbol{a}\|}{\kappa}x} - 1}{e^{\frac{\|\boldsymbol{a}\|}{\kappa}} - 1}\Big) \qquad \text{in } \tilde{\Omega} \qquad (30a)$$

$$\boldsymbol{\sigma} = f\Big(x + \frac{1}{e^{\frac{\|\boldsymbol{a}\|}{\kappa}} - 1} - \frac{\kappa}{\|\boldsymbol{a}\|}\Big) \qquad \text{in } \tilde{\Omega} \qquad (30b)$$

$$\lambda^- = \phi \qquad \qquad \text{on } \Gamma \qquad (30c)$$

$$\lambda^+ = \phi \qquad \qquad \text{on } \Gamma_0 \qquad (30d)$$

Next, we construct coarse-scale spaces $\mathcal{W}^h$ and $\mathbf{\Sigma}^h$ from first-order discontinuous basis functions on five uniformly spaced elements. To compute the fine-scale closure function, we approximate the spaces $\mathcal{W}$ and $\mathbf{\Sigma}$ with piece-wise constant discontinuous and first-order continuous basis functions respectively on a fine-scale mesh with 1024 elements in each coarse-scale element. The fine-scale closure function corresponding to the sixth basis function of $\mathbf{\Sigma}^h$ is shown in Fig. 3. The parameters in the scale-decomposition projector of Eq. (20) are chosen as $\eta = 0.1$, $\beta = -0.5$ and $C = 0.01$. The figure shows an extreme localization of the scale interaction: the influence of the fine scales is almost contained in a single element. We explore the footprint of fine-scale closure functions more in Sections 6.2 and 6.3.

After also computing the fine-scale closure functions for the remaining 9 basis functions that span $\mathbf{\Sigma}^h$ we can solve the system of equations of Eq. (28). Figure 4 shows the resulting coarse-scale solution dotted in red. The figure also shows the true solution of Eq. (30) in
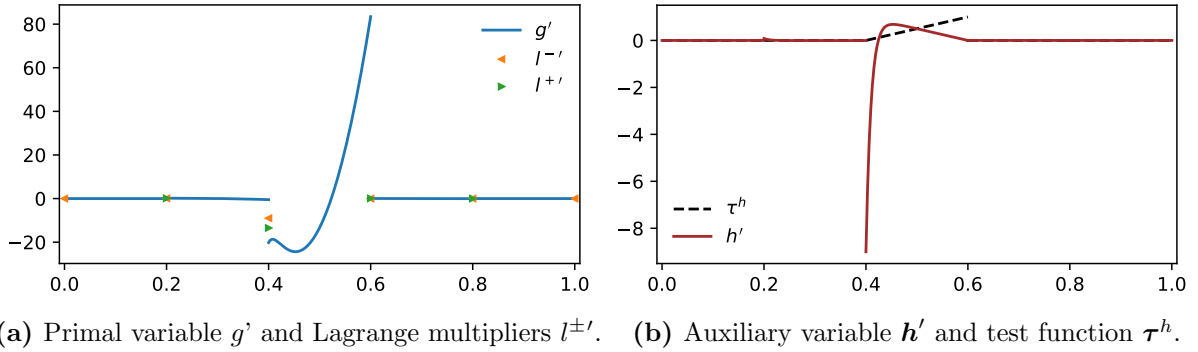
(a) Primal variable g' and Lagrange multipliers $l^{\pm'}$. **(b)** Auxiliary variable $\boldsymbol{h}'$ and test function $\boldsymbol{\tau}^h$.

Figure 3: Fine-scale closure function for $\kappa = 0.01$ and the LDG-projector with $\eta = 0.1$, $\beta = -0.5$ and $C = 0.01$.

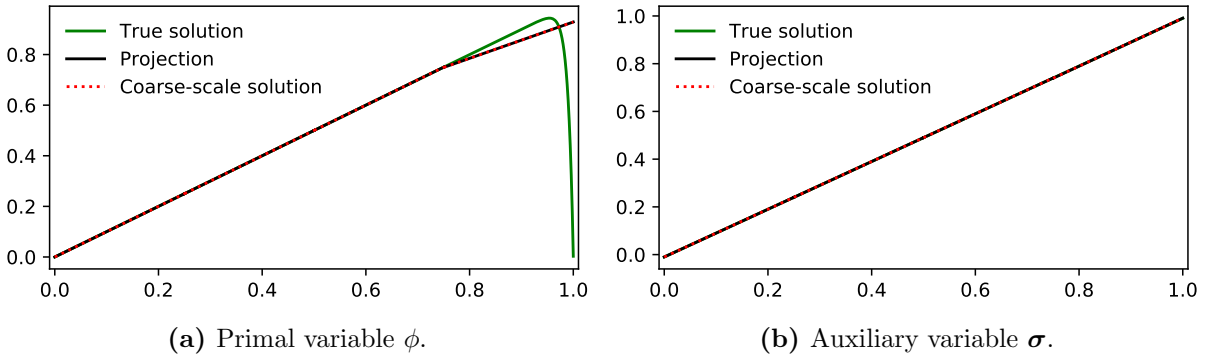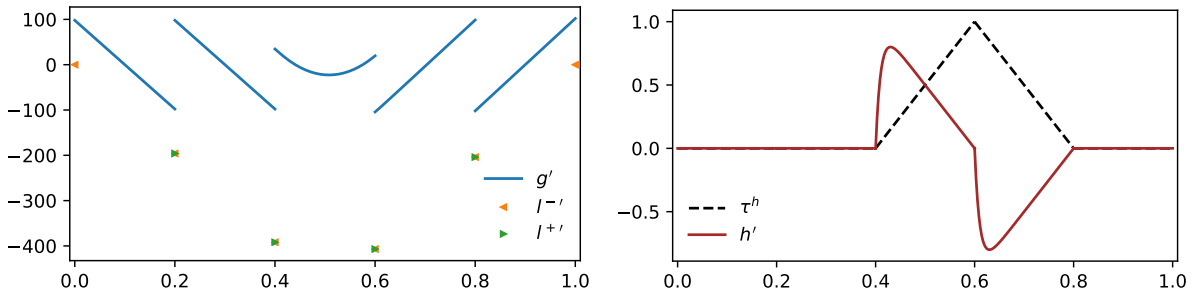**(a)** Primal variable ϕ. **(b)** Auxiliary variable $\boldsymbol{\sigma}$.

Figure 4: Full-scale and coarse-scale solution for the LDG-projection.

green, and the projection of that true solution onto the coarse-scale space in black. The figure confirms that the "optimal discontinuous Galerkin formulation" of Eq. (28) indeed yields the sought-after coarse-scale solution. This corroborates the legitimacy of our procedure for adopting discontinuous Galerkin formulations in the variational multiscale framework per Section 4, as well as the procedure for eliminating the remaining fine scales by means of the fine-scale closure function per Section 5.

6.2. Comparison with a mixed method

To put the impact of a discontinuous Galerkin formulation in perspective, we compare fine-scale closure functions to those of a conforming method. We repeat the experiment corresponding to Fig. 3, but we replace the constraint equations by constraints corresponding to a mixed formulation: the penalty parameters in the projection operator of Eq. (20) are set to zero and the coarse-scale space is comprised of piece-wise constants and first-order continuous basis functions for $\mathcal{W}^h$ and $\boldsymbol{\Sigma}^h$ respectively. Figure 5 shows the resulting fine-scale closure function for the coarse-scale test function shown in Fig. 5b. When compared to

(a) Primal variable g' and Lagrange multipliers $l^{\pm\prime}$. **(b)** Auxiliary variable $\boldsymbol{h}'$ and test function $\boldsymbol{\tau}^h$.

Figure 5: Fine-scale closure function for the (conforming) mixed method.

the closure function of the discontinuous Galerkin method from Fig. 3, the closure function of the conforming mixed formulation shows a much wider support. This is particularly clear in the primal field, where both g' and $l^{\pm}$ are completely global. In the auxiliary field, the fine-scale closure function is exactly contained within the elements corresponding to the coarse-scale test function.

For a two-dimensional comparison, we consider the model problem described in [57]. The physical domain is a unit square and the advective field is $\boldsymbol{a} = [1/\sqrt{5}, 2/\sqrt{5}]^{\mathrm{T}}$. We use a 16×16 mesh of triangles for the construction of all following coarse-scale spaces, and we approximate the spaces $\mathcal{W}$ and $\boldsymbol{\Sigma}$ by subdividing each coarse-scale element into 1024 piecewise constant discontinuous Galerkin and first-order Brezzi-Douglas-Marini (BDM) elements respectively. Both the coarse-scale mesh and the fine-scale mesh are shown in Fig. 6a. For the source term in the adjoint formulation we consider the coarse-scale $\boldsymbol{\tau}^h$ Brezzi-Douglas-Marini basis function shown in Fig. 6b. In the unconstrained case (i.e., $C(\cdot, \cdot) = 0$) and for a diffusion coefficient $\kappa = 0.01$ this source term in the adjoint problem would produce the solution ϕ shown in Fig. 6a.

To compute the fine-scale closure function of the two-dimensional conforming mixed formulation, we construct $\mathcal{W}^h$ and $\boldsymbol{\Sigma}^h$ on the coarse-scale mesh with piece-wise constant DG elements and first-order BDM elements respectively. The fine-scale closure function for the mixed method follows from adding the constraints of Eq. (22) with the choices $\eta = C = 0$ and $\boldsymbol{\beta} = \boldsymbol{0}$. The solution fields g' and $\boldsymbol{h}'$ for a diffusion coefficient $\kappa = 0.01$ are shown in Fig. 7. This figure again illustrates that the fine-scale closure function corresponding to a plain mixed method has a widespread support. Within each individual element the fine-scale closure function changes sign. The function g' is almost linear within upwind and downwind elements, and the gradient in upwind elements points in the opposite direction compared to the gradient in downwind elements. The same solution behavior can be observed in the earlier one-dimensional fine-scale closure function of Fig. 5. Just like for that one-dimensional fine-scale closure function, the auxiliary variable $\boldsymbol{h}'$ is much more localized than ϕ' and appears smooth. Different from the one-dimensional case, $\boldsymbol{h}'$ is no longer fully contained within the elements corresponding to $\boldsymbol{\tau}^h$.

18

(a) Solution ϕ in the unconstrained case, $\kappa = 0.01$.

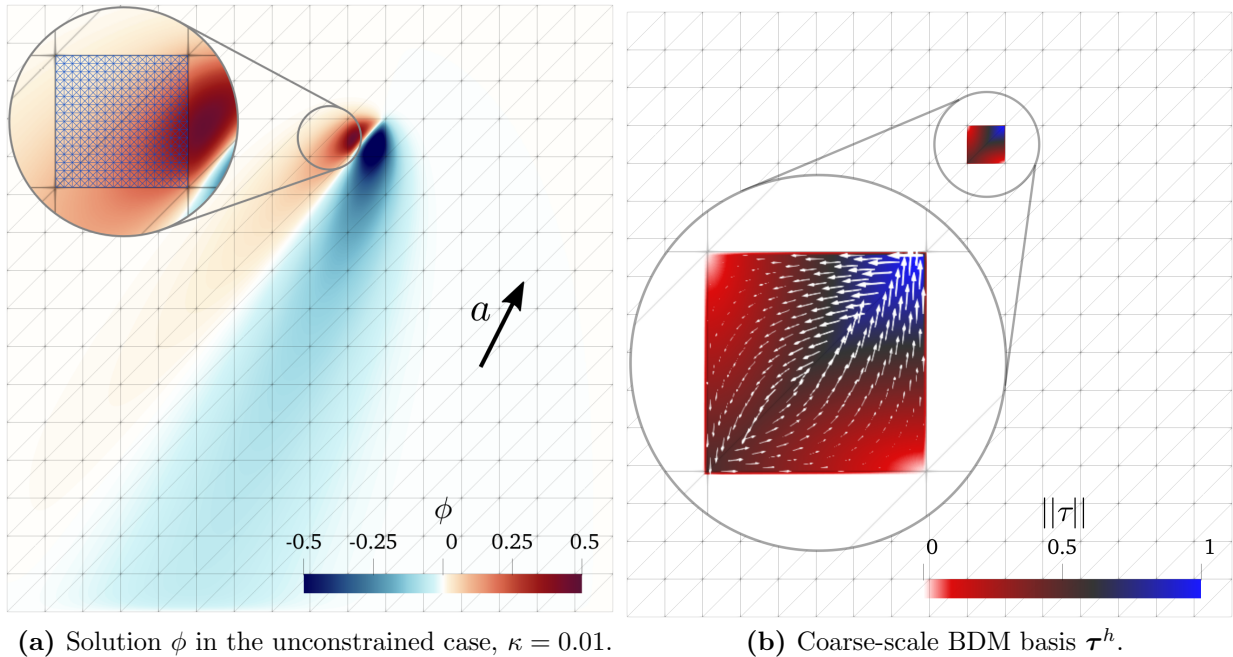**(b)** Coarse-scale BDM basis $\boldsymbol{\tau}^h$.

Figure 6: Formulation of the adjoint problem and the example solution in the unconstrained case. Illustration of the coarse-scale mesh (grey) and the fine-scale mesh (blue).

Next, we consider an LDG formulation in the same setting. We use the same mesh for the construction of the coarse-scale spaces $\boldsymbol{\Sigma}^h$ and $\mathcal{W}^h$, but now use first-order DG elements for both spaces. To avoid bias, we make use of the same coarse-scale function $\boldsymbol{\tau}^h$ from Fig. 6b in the adjoint problem. The penalty parameters are chosen as $\eta = 20$, $C = 0$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$, where $\boldsymbol{n}^+$ is the normal vector on Γ for which $\boldsymbol{a} \cdot \boldsymbol{n}^+ \geq 0$. The corresponding fine-scale closure function for a diffusivity of $\kappa = 0.01$ is shown in Fig. 8.

The change to a DG method results in a drastic decrease in the size of the footprint of the fine-scale closure function. Recall that the fine-scale closure functions shown in these graphs capture the fine-scale effects that would need to be incorporated in the finite element formulation via Eq. (28). The difference between Figs. 7 and 8 can then be interpreted from two perspectives: producing a fine-scale model that mimics the effects of Fig. 8 appears more viable than attempting to include the effects of Fig. 7, which suggests that a DG framework is particularly suitable when the physical problem exhibits severe inherent multiscale effects (e.g. for large eddy simulation of turbulent flow). On the other hand, if no fine-scale model is included, then the figures illustrate the fine-scale closure that is neglected. Omitting the closure of Fig. 7 is a more severe modeling approximation than omitting the closure of Fig. 8. This provides a new perspective on the natural stability of DG methods when it comes to hyperbolic terms in partial differential equations: the harmful impact of the missing closure terms is significantly reduced.
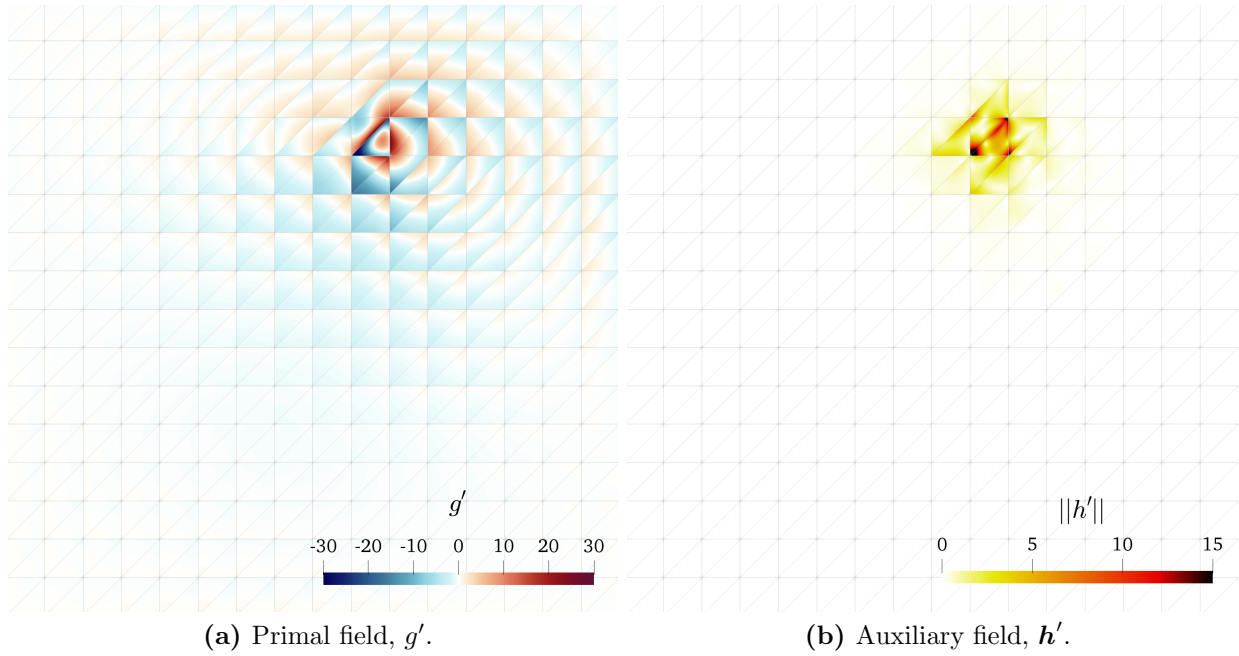
19

(a) Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 7: Fine-scale closure functions for the (conforming) mixed method and $\kappa = 0.01$.

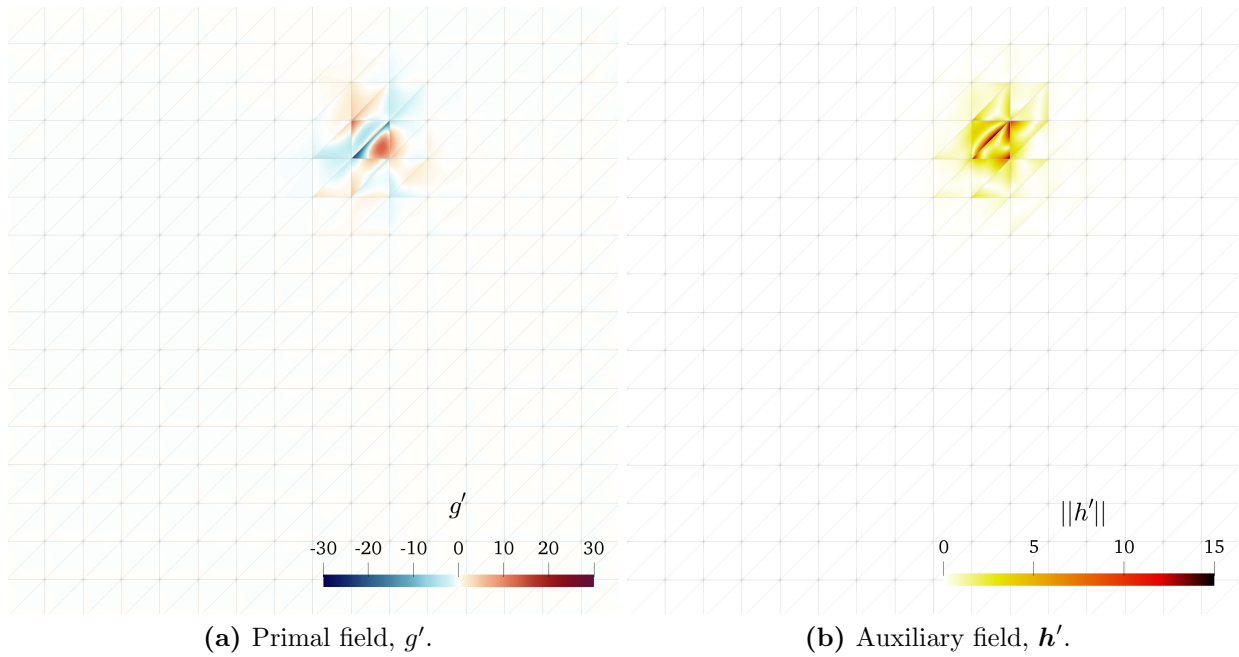**(a)** Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 8: Fine-scale closure functions for the LDG method and $\kappa = 0.01$, with $\eta = 20$, $C = 0$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.
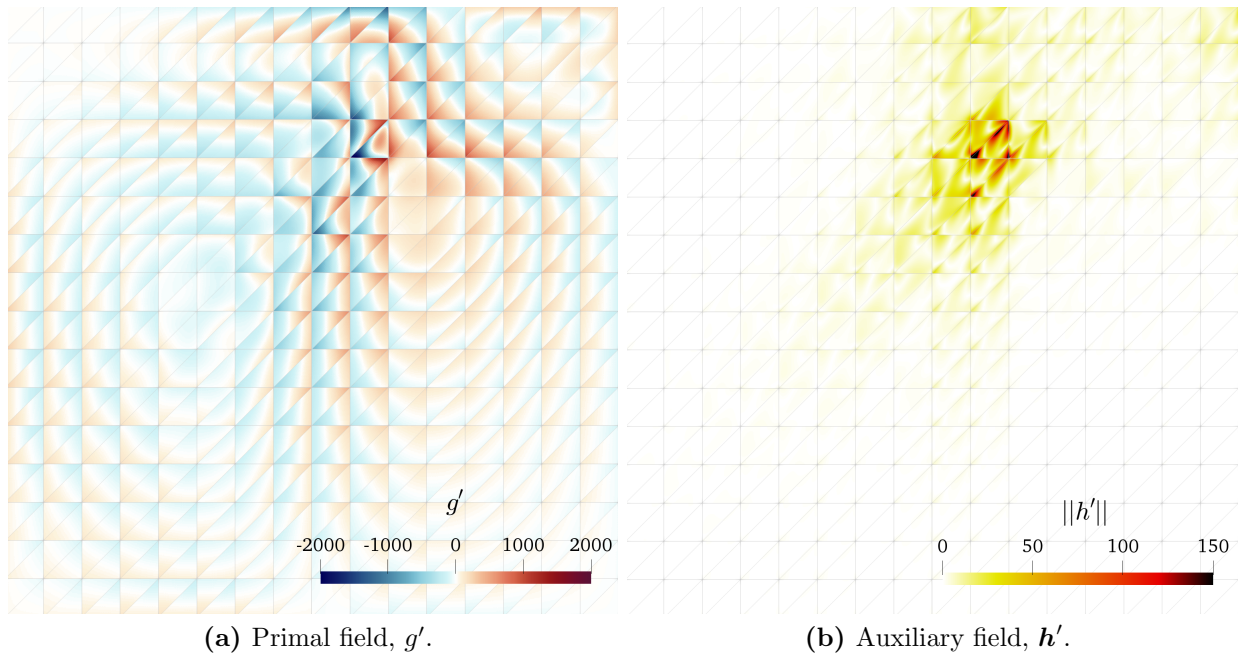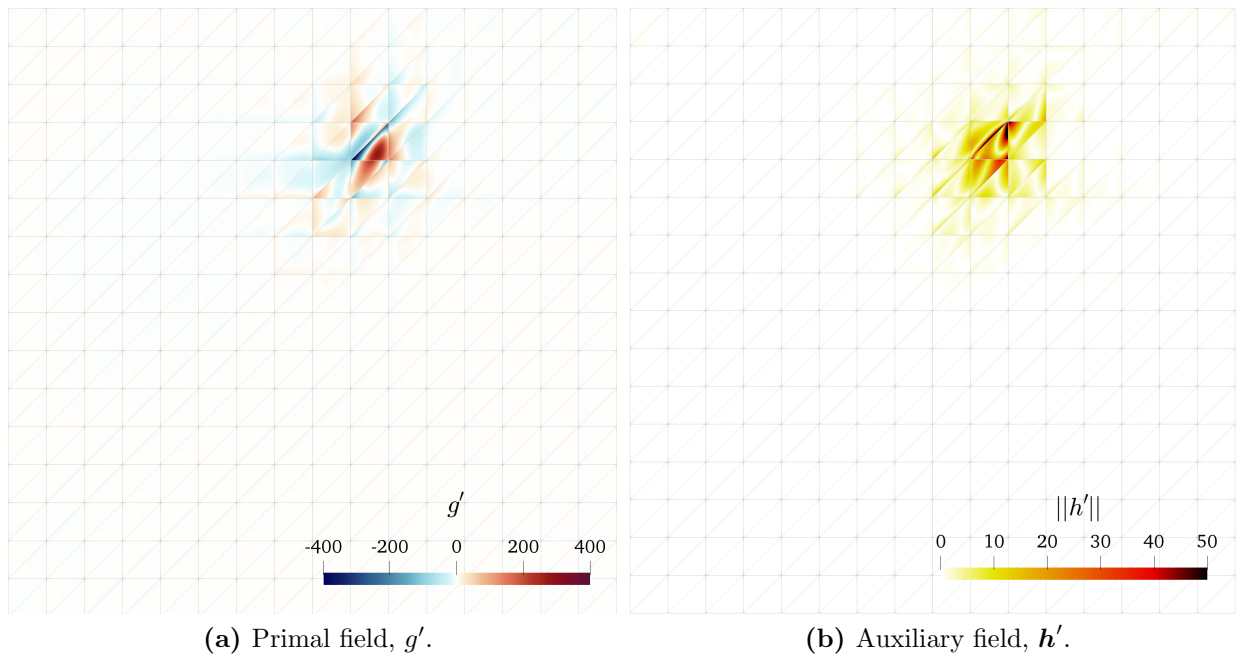
20

(a) Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 9: Fine-scale closure functions for the (conforming) mixed method and $\kappa = 0.002$.

(a) Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 10: Fine-scale closure functions for the LDG method and $\kappa = 0.002$, with $\eta = 20$, $C = 0$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

To further corroborate this point, we reduce the diffusivity by a factor of five to $\kappa = 0.002$. The fine-scale closure functions corresponding to these finite element formulations are shown in Figs. 9 and 10. Even more so than before do we observe a widespread and high-valued fine-scale closure function for the mixed method in comparison to the compact moderately valued function for the DG method. Also note the change in scale of the color bar in this regard. While the footprint of the auxiliary field $\boldsymbol{h}'$ for the mixed method was relatively small before, with the decrease in diffusivity it has grown significantly. In contrast, the auxiliary field $\boldsymbol{h}'$ for the fine-scale closure function of the discontinuous Galerkin formulation shows roughly the same spread as the primal variable, just like it did for the larger diffusion coefficient. In general, the fine-scale closure function corresponding to the DG method appears smooth and well-behaved, implying that a fine-scale model can potentially be effective.

6.3. Localization of the fine-scale closure function in LDG methods

Next, we explore the effect of the parameters in the LDG method on the fine-scale closure function. The free parameters are C, η and $\boldsymbol{\beta}$. Different choices lead to different LDG methods with different convergence characteristics [4, 88, 89]. Three example methods are collected in Table 1, optimal convergence rates of which are proven in [90, 89]. These three methods illustrate three extrema in terms of parameter dependencies: as the mesh size h tends to zero, the parameter η shrinks to 0 for the LDG-1 method and grows to ∞ for the LDG-3 method, whereas the parameter C grows to ∞ and shrinks to zero. From our variational multiscale viewpoint, these different choices lead to different decomposition projections (Eq. (20)) and thereby to different closure relations (Eqs. (26) and (28)). From that perspective, the motivation for choosing one particular parameter set over the other could be to simplify the remaining closure relationship, or more specifically, to yield a compact fine-scale closure function.

Table 1: Example parameter choices of local discontinuous Galerkin methods [4, 90]. The domain width W is introduced for dimensional consistency.

Name	η	$\boldsymbol{\beta}$	C
LDG-1	$\mathcal{O}(h/W^2)$	$S^+\boldsymbol{n}^+ + S^-\boldsymbol{n}^-$	$\mathcal{O}(W^2/h)$
LDG-2	$\mathcal{O}(1/W)$	$S^+\boldsymbol{n}^+ + S^-\boldsymbol{n}^-$	$\mathcal{O}(W)$
LDG-3	$\mathcal{O}(1/h)$	$S^+\boldsymbol{n}^+ + S^-\boldsymbol{n}^-$	$\mathcal{O}(h)$

To exemplify such a VMS-based design driving approach, we compute the support of the fine-scale closure function for a wide range of parameters in the LDG formulation. We do this for the one-dimensional model problem, and subsequently verify that the fine-scale closure function of a two dimensional problem with the same element Péclet number exhibits similar characteristic behavior. To reduce the size of the search space, we choose $\boldsymbol{\beta}$ as either $\frac{1}{2}\boldsymbol{n}^+$ or $\frac{1}{2}\boldsymbol{n}^-$. These are two common choices of the 'switch' parameter $S^\pm$ from Table 1 as they

reduce the bandwidth of the resulting stiffness matrix [91]. For each one-dimensional fine-scale closure function, we measure the support as the fraction of the degrees of freedom in the fine-scale space for which the absolute value of g' or h' is larger than 1 or 0.1 respectively. A smaller fraction means a more localized fine-scale closure function. Figure 11 shows the resulting heatmap when the coarse-scale space is constructed with a mesh of five elements, which corresponds to the earlier Fig. 3. In the two plots on the left, the parameter $\boldsymbol{\beta}$ is chosen as half the inflow normal $\frac{1}{2}\boldsymbol{n}^-$, and in the plots on the right as half the outflow normal $\frac{1}{2}\boldsymbol{n}^+$. The top row shows results for a diffusion coefficient $\kappa = 0.01$ and the bottom row for $\kappa = 0.001$. Figure 12 shows the same four graphs for a coarse-scale mesh with 16 elements, which relates to the earlier two-dimensional experiments of Figs. 8 and 10. The different figures thus span a range of coarse-scale element Péclet numbers.
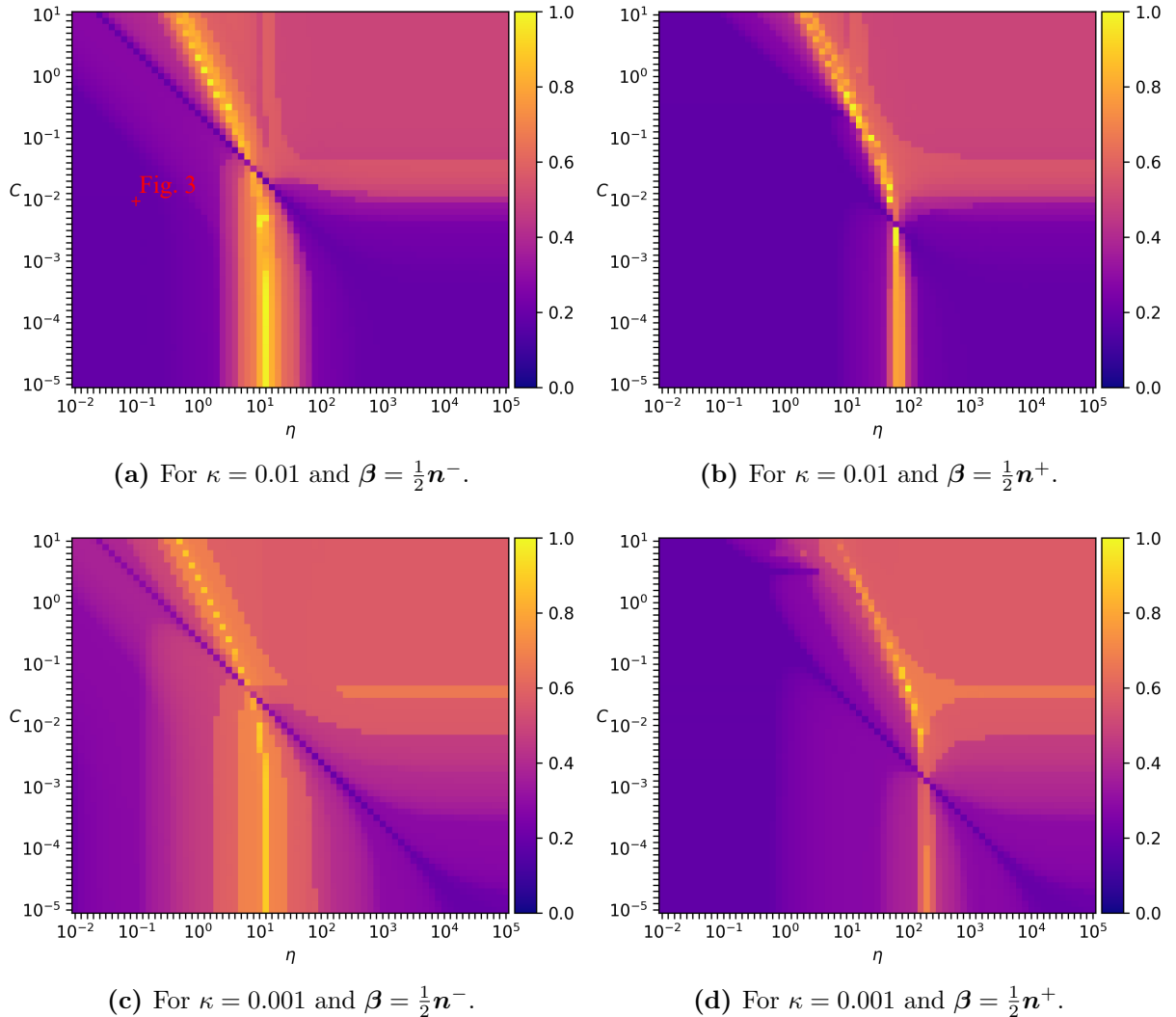

(a) For $\kappa = 0.01$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$.   **(b)** For $\kappa = 0.01$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

(c) For $\kappa = 0.001$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$.   **(d)** For $\kappa = 0.001$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

Figure 11: Support of the fine-scale closure function as a fraction of the complete domain for different parameters in the LDG-projector and different diffusion coefficients. The fine-scale closure function is computed for the central basis function in a one-dimensional domain discretized with five coarse-scale elements.
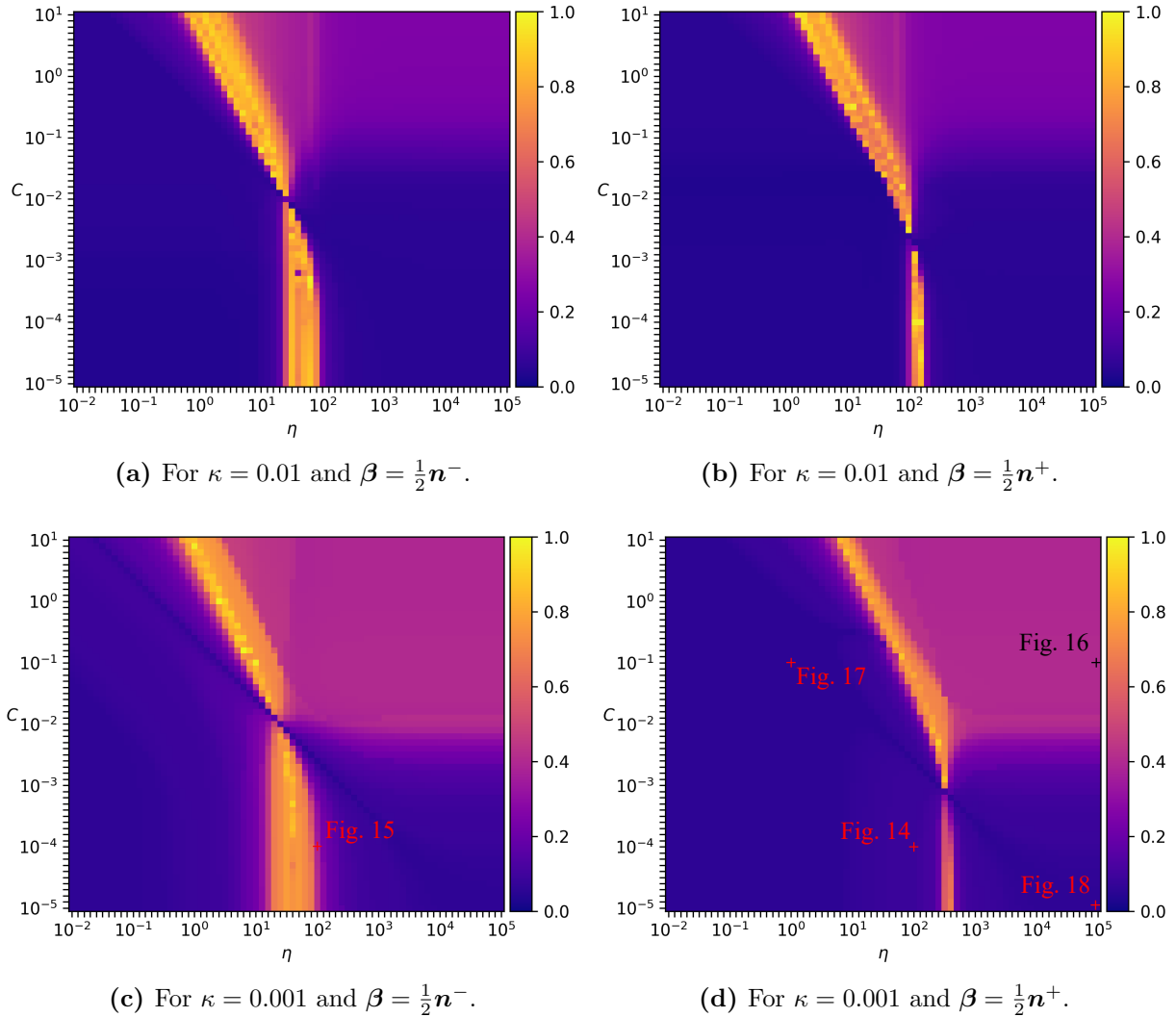
23

(a) For $\kappa = 0.01$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$.

(b) For $\kappa = 0.01$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

(c) For $\kappa = 0.001$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$.

(d) For $\kappa = 0.001$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

Figure 12: Support of the fine-scale closure function as a fraction of the complete domain for different parameters in the LDG-projector and different diffusion coefficients. The fine-scale closure function is computed for the central basis function in a one-dimensional domain discretized with 16 coarse-scale elements.

All eight graphs show the same three distinct regions: there are bands of combinations of η and C for which the fine-scale closure function is almost global (indicated in orange and yellow), in the upper right corner of the graphs the fine-scale closure functions have a footprint of roughly half the domain (indicated in purple and pink), and the left and bottom-right sections of the graphs show localized fine-scale closure functions (indicated in dark blue). Based on these results, favorable parameter choices would be a small η and a large C or vice versa. When comparing the left and right columns in Figs. 11 and 12, we observe that the choice $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$ (with $\boldsymbol{n}^+$ the outflow normal) is preferable over $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$. The right figures consistently show thinner orange-yellow bands and wider dark blue regions, indicating more localized scale interaction overall. This choice of $\boldsymbol{\beta}$ is particularly important

24

(a) Solution ϕ in the unconstrained case, $\kappa = 0.001$.

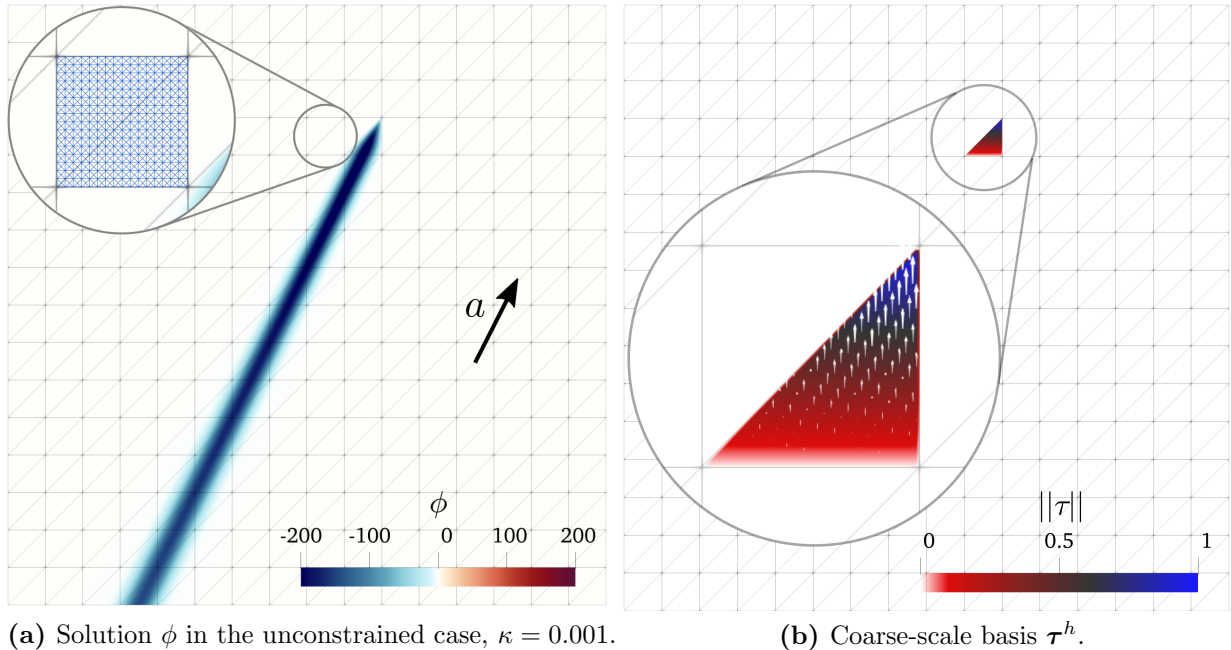**(b)** Coarse-scale basis τ^h.

Figure 13: Formulation of the adjoint problem and the example solution in the unconstrained case. Illustration of the coarse-scale mesh (grey) and the fine-scale mesh (blue).

for moderate choices of η and C, as it most significantly reduces the footprint of the fine-scale closure function for $C < 0.01$ and around the range $1 < \eta < 100$. Parameter ranges with more extreme values of η and C produce localized fine-scale closure functions irrespective of the choice of $\boldsymbol{\beta}$.

Next, we compute two-dimensional fine-scale closure functions corresponding to distinctive points in Figs. 12c and 12d. To maintain approximately the same coarse-scale element Péclet number, we use the same model problem as in Section 6.2 with a reduced diffusion coefficient of $\kappa = 0.001$. The discontinuous Galerkin coarse-scale basis function depicted in Fig. 13b is used for the computation of all following fine-scale closure functions. In the unconstrained case, this adjoint problem would yield the primal solution shown in Fig. 13a. Figure 13a also shows the highly refined fine-scale mesh on which the fine-scale closure functions are computed.

Figures 14 and 15 show the fine-scale closure functions for $\eta = 100$, $C = 10^{-4}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$ respectively. These parameter points are drawn into Figs. 12c and 12d. As may be observed from Fig. 12c, the choice $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^-$ falls within the yellow band and should result in a highly non-local fine-scale closure function, whereas the choice $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$ should produce a function with a more local support. The simulation results of Figs. 14 and 15 indeed demonstrate the anticipated behavior. This confirms that choosing $\boldsymbol{\beta}$ as half the outflow normal is preferable to half the inflow normal, also for a two-dimensional case. The result implies that the general characteristics of the fine-scale closure functions are determined more by the coarse-scale element Péclet number than by the spatial dimension of the problem.
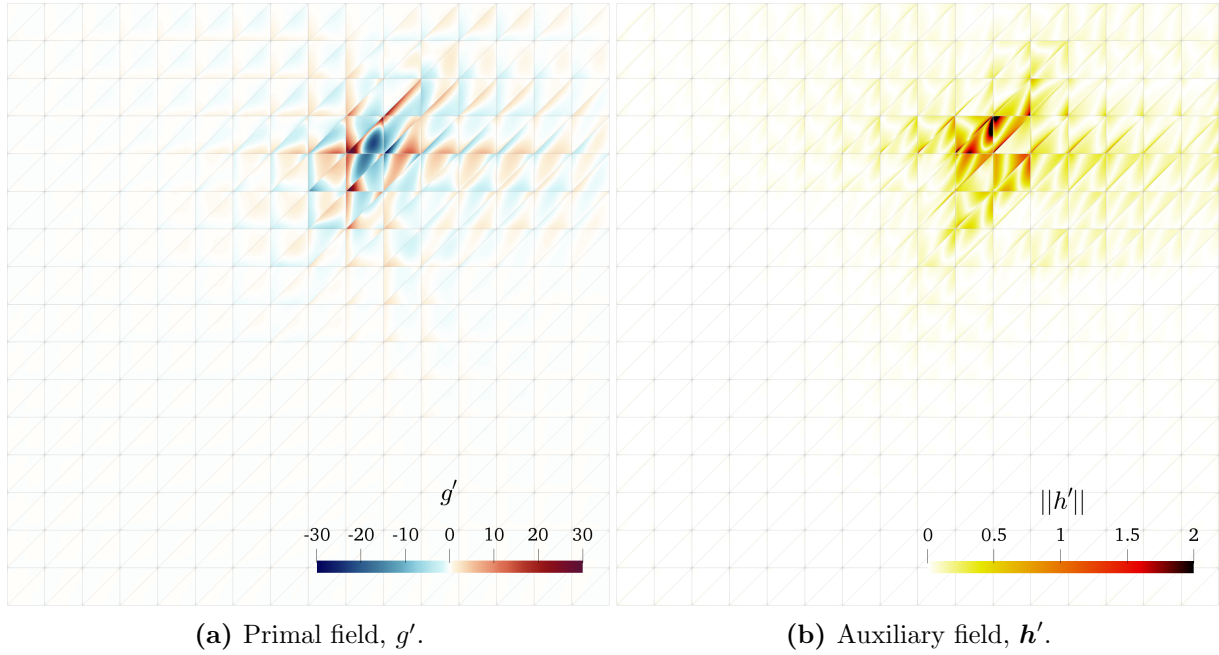
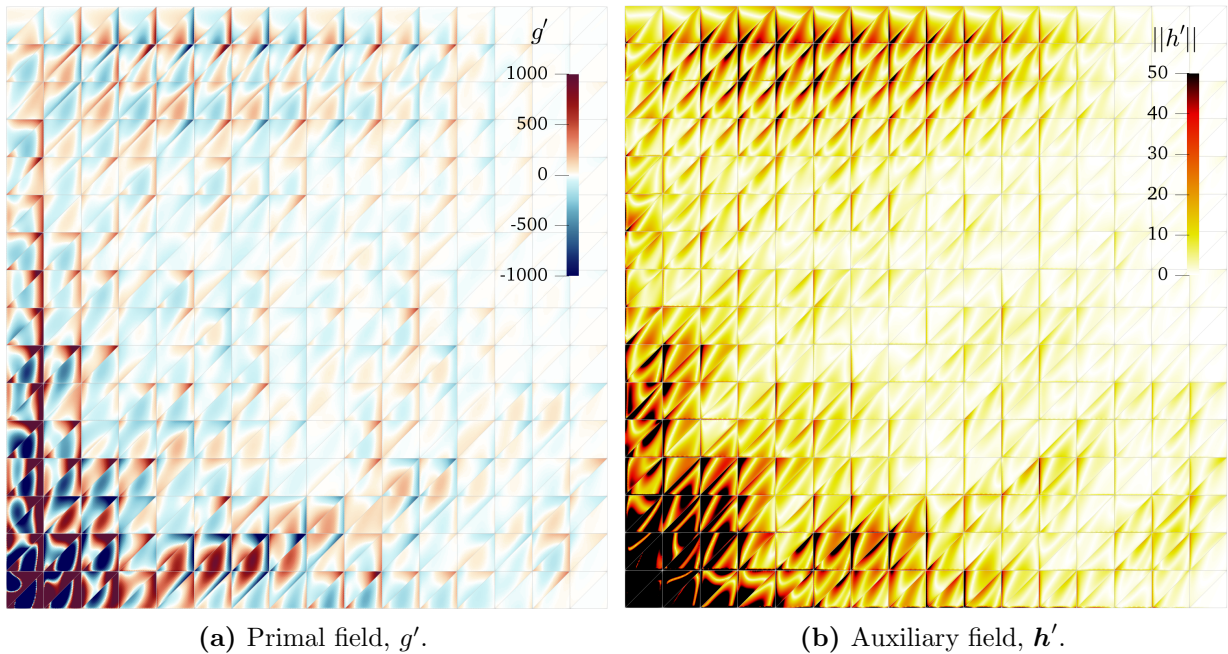**(a)** Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 14: Fine-scale closure function for $\kappa = 0.001$ with parameters $\eta = 100$, $C = 10^{-4}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^{+}$. Corresponding to the middle point in Fig. 12d.

(a) Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 15: Fine-scale closure function for $\kappa = 0.001$ with parameters $\eta = 100$, $C = 10^{-4}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^{-}$. Corresponding to the middle point in Fig. 12c.

26

(a) Primal field, g'.   (b) Auxiliary field, h'.
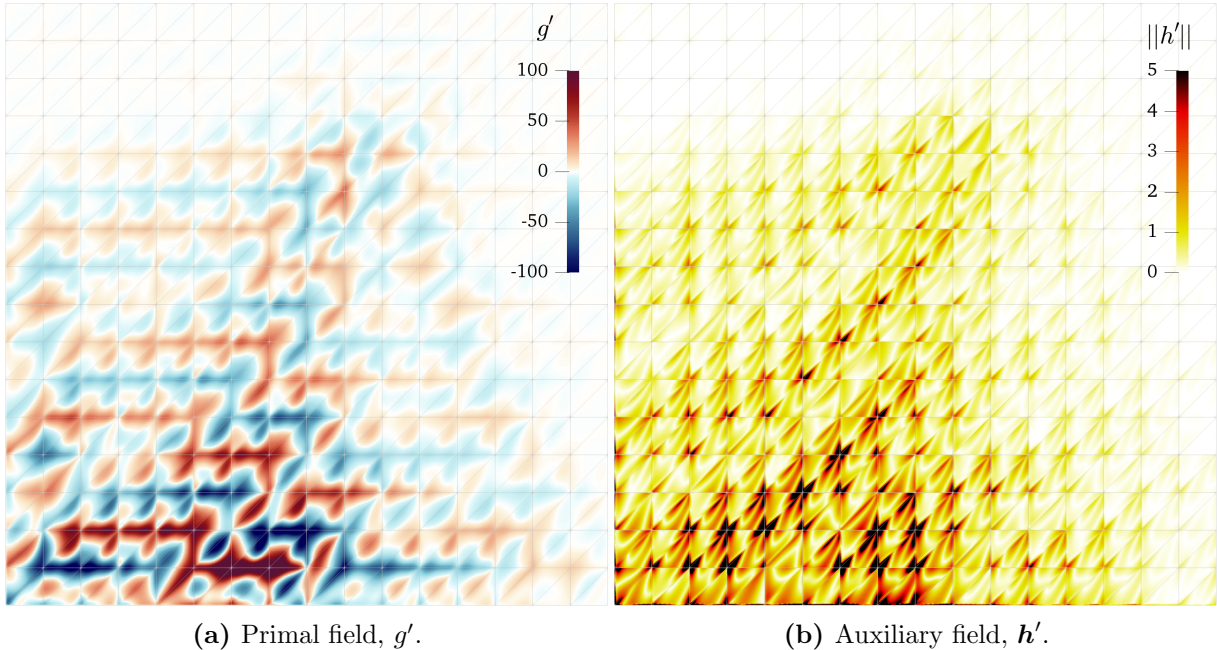
Figure 16: Fine-scale closure function for $\kappa = 0.001$ with parameters $\eta = 10^5$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. Corresponding to the upper-right point in Fig. 12d.

Secondly, we choose a parameter point in the upper right corner of Fig. 12d, namely $\eta = 10^5$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. The resulting fine-scale closure function depicted in Fig. 16 again shows a wide-spread support. Different from the earlier function shown in Fig. 15, the support of the new function is limited to the upwind region of the coarse-scale basis function. This explains the values of 0.5 on the top-right corners of the graphs in Figs. 11 and 12: the complete upwind half of the one-dimensional domain affects the central basis function. The large penalty values on both the primal and auxiliary fields enforces an incompatible continuity and thereby results in a non-local fine-scale closure function.

Next, we consider the upper-left and lower-right corners of Fig. 12d. Figure 17 shows the result for $\eta = 1$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$ and Fig. 18 shows the result for $\eta = 10^5$, $C = 10^{-5}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. As anticipated, both parameter choices produce a localized fine-scale closure function, although quite different in nature. The function in Fig. 17 appears more erratic compared to the smooth function in Fig. 18. Also note the change in colorbar-range for the primal field in that regard.

Finally, we investigate the coarse-scale solution corresponding to the parameter choices of Figs. 17 and 18. Consider the benchmark problem described in Fig. 19, adopted from [57]. The solution exhibits steep internal and boundary layers, making this a challenging function to approximate well on a coarse mesh. Figure 20 shows the coarse-scale solution (i.e., the solution to Eq. (28) including the exact scale interaction) for the parameter choices $\eta = 1$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$, and Fig. 21 shows the coarse-scale solution for $\eta = 10^5$, $C = 10^{-5}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.

27

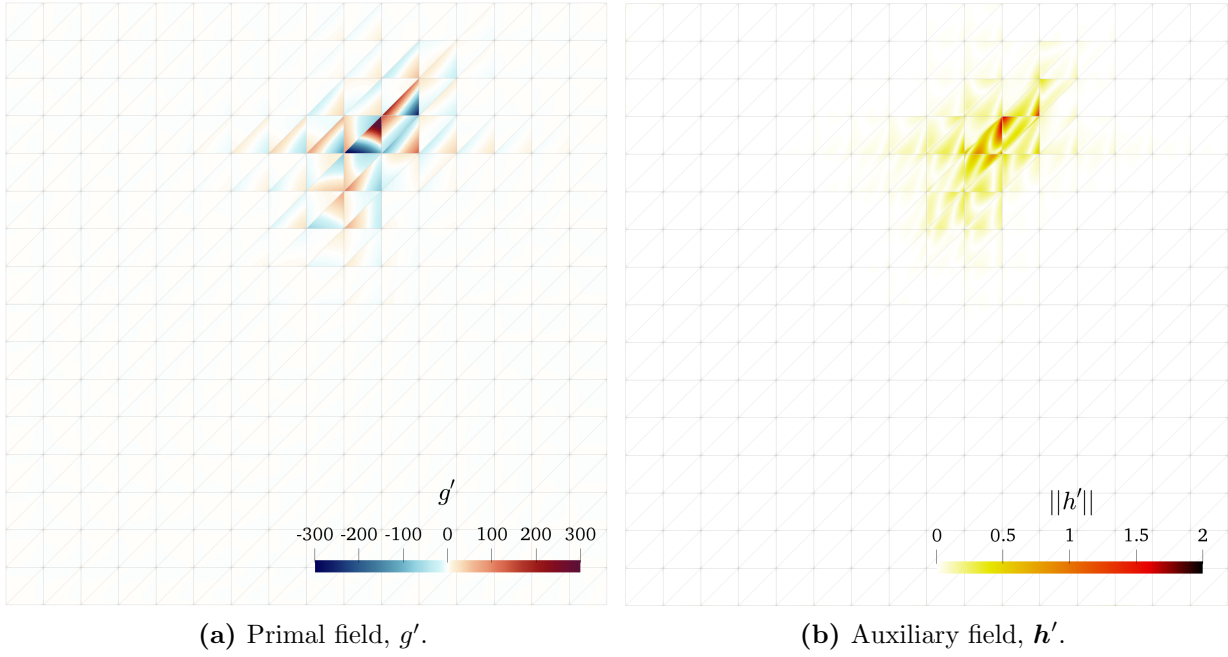**(a)** Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 17: Fine-scale closure function for $\kappa = 0.001$ with parameters $\eta = 1$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. Corresponding to the upper-left point in Fig. 12d.
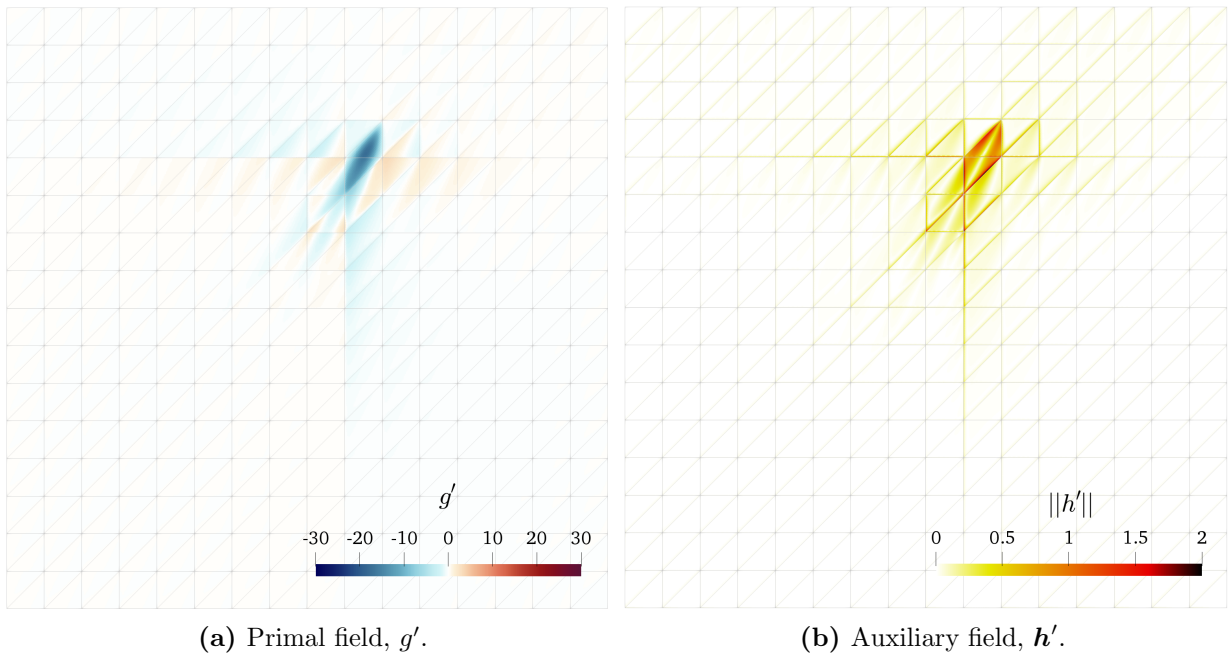

(a) Primal field, g'.

(b) Auxiliary field, $\boldsymbol{h}'$.

Figure 18: Fine-scale closure function for $\kappa = 0.001$ with parameters $\eta = 10^5$, $C = 10^{-5}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. Corresponding to the lower-right point in Fig. 12d.

28

The two different parameter choices produce radically different coarse-scale solutions. In Fig. 20, η is chosen small. This puts little weight on the continuity of the coarse-scale solution, permitting large jumps between elements. As a result, the coarse-scale solution is free to mirror the high gradients of the true solution, while maintaining smoothness where the true solution is smooth. An immediate consequence of these large (changes in) gradients is that the solution exhibits significant overshoots and undershoots. These severe overshoots and undershoots and the surrounding oscillations make the parameter choice $\eta = 1$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$ an unlikely candidate for use in practical computations, despite the favorable localization of the scale interaction per Fig. 17.

Differently, the case with large η produces a near continuous solution field. The steep layers in the true solution are smoothed out, and the undershoots and overshoots are minor. In the "eyeball norm", the coarse-scale solution of Fig. 21 even outperforms those corresponding to an H_0^1-projector and an L^2-projector on a similarly coarse grid (the focus of study in [57]). The same parameter choice ($\eta = 10^5$, $C = 10^{-5}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$) also produces the most localized fine-scale closure function, see Fig. 18, and this closure function is more compact than the fine-scale Green's functions for either the H_0^1-projector or the L^2-projector (also illustrated in [57]). Both these arguments speak in favor of using the DG-type projector of Eq. (20) with a large value for η and a small value for C in the scale decomposition in a variational multiscale framework. It might also be worthwhile to explore modifications to this projector as a starting point for developing new discontinuous Galerkin methods. For example, one could add a small L^2-weighting to the primal field to balance out the excess smoothing (yielding PSPG-like additional terms in the LDG formulation), or one could incorporate the additional terms proposed in [92] to introduce shock-capturing terms.
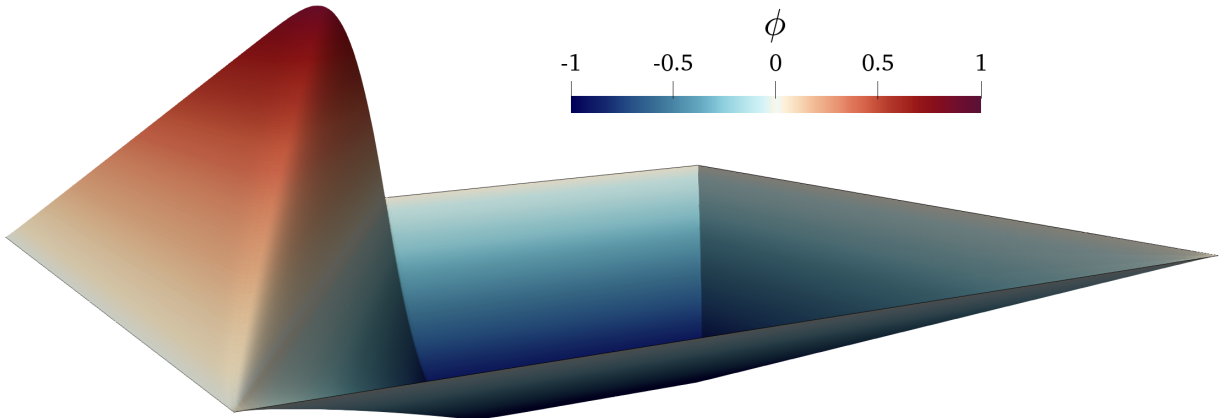

Figure 19: Solution to the advection-diffusion model problem on a unit-square domain, with $\kappa = 0.001$, $\boldsymbol{a} = [1/\sqrt{5}, 2/\sqrt{5}]^{\mathrm{T}}$ and $f = 1$ if $2x < y$ or $f = -1$ if $2x > y$.
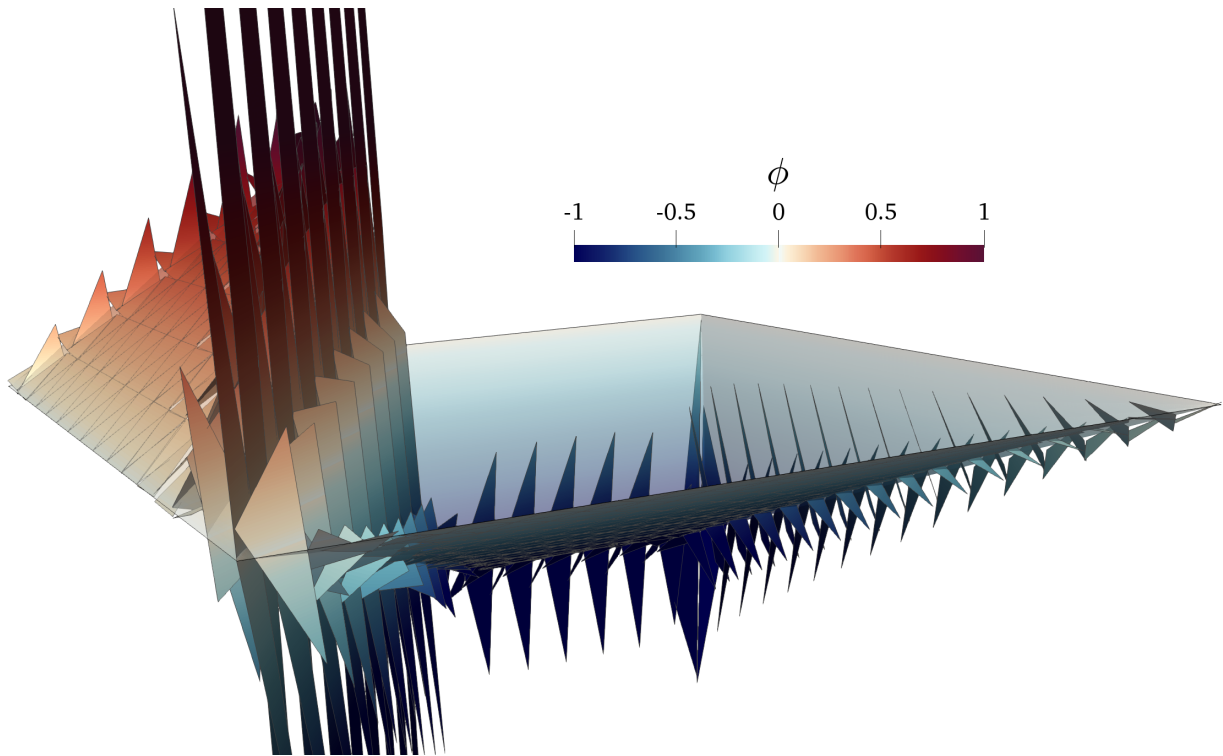
Figure 20: Coarse-scale solution for the problem of Fig. 19, with parameters $\eta = 1$, $C = 0.1$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. Corresponding to the fine-scale closure function of Fig. 17.
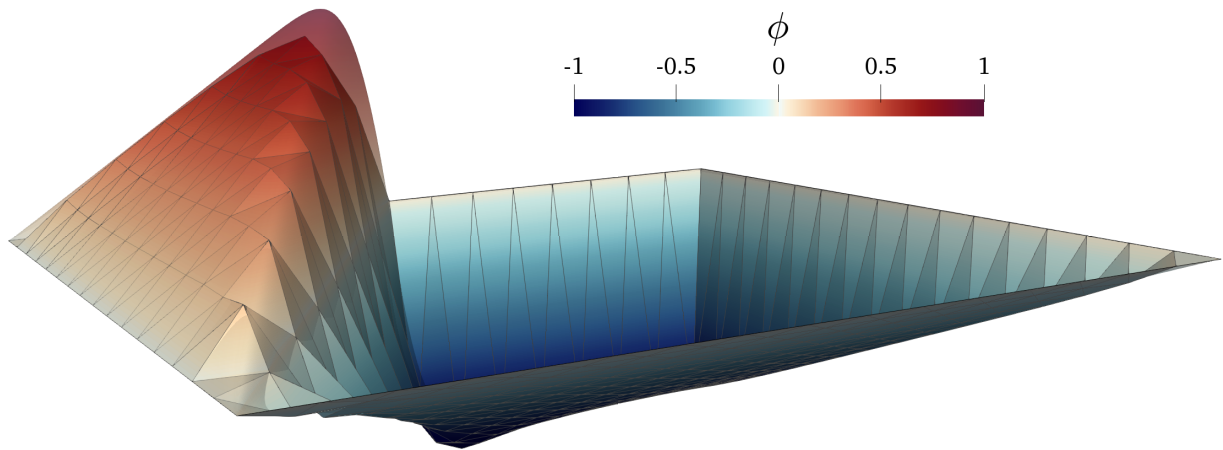

Figure 21: Coarse-scale solution for the problem of Fig. 19, with parameters $\eta = 10^5$, $C = 10^{-5}$ and $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$. Corresponding to the fine-scale closure function of Fig. 18.

7. Conclusion and outlook

Discontinuous Galerkin methods continue to gain popularity in the field of fluid mechanics. These methods *(i)* are non-conforming with respect to traditional weak formulations, *(ii)* involve operators that are unbounded in the continuous limit, *(iii)* are stable by careful design of numerical fluxes, their discontinuous basis functions *(iv)* render the typical H_0^1-multiscale projector invalid and *(v)* conflict with the continuity requirements of classical residual-based fine-scale models, and *(vi)* they involve penalty terms. These are six characteristics that obfuscate the connection between discontinuous Galerkin methods and the variational multiscale paradigm. In this article, we showed that given the right perspective, discontinuous Galerkin methods fit well within the variational multiscale view on finite element formulations. To this end, we presented a new approach that tackles all six points, thus enabling the derivation of discontinuous Galerkin formulations from a variational multiscale perspective.

In particular, we tackle point *(i)* by interpreting the standard Sobolev spaces ($H^1(\Omega)$, $\boldsymbol{H}(\text{div}, \Omega)$, etc.) as constrained spaces, where the constraints are the continuity conditions across element boundaries. We release these constraints from the space, and reapply them in the weak formulation with a Lagrange multiplier. Discontinuous Galerkin approximation spaces are subspaces of these broken Sobolev spaces, and the variational multiscale decomposition is now applicable. The central role of the Lagrange multiplier field in our derivation permits us to avoid unbounded trace operations (point *(ii)*). Points *(iv)* and *(v)* are addressed by augmenting the H_0^1-projector with weighted boundary terms. Discontinuous Galerkin formulations then emerge directly (i.e., without any approximation or fine-scale modeling) from special choices of projection operators. Understanding this inherent connection between the finite element formulation and the scale decomposition is of key importance for handling the remaining fine-scale terms.

When we apply this derivation in the context of an advection-diffusion problem to obtain a local discontinuous Galerkin formulation, then (only) a single (volumetric) fine-scale term remains. We use the newly proposed "fine-scale closure function" to compare the influence of the remaining fine-scale term on the discontinuous Galerkin method with that on a conforming mixed finite element method. The results show a localized and moderately valued fine-scale closure for the discontinuous Galerkin method and a widespread and large fine-scale closure for the conforming formulation. The stable nature of discontinuous Galerkin methods stated in point *(iii)* above can thus be substantiated from the perspective of the variational multiscale method.

We used our formalism to explore the dependency of the required fine-scale closure on the parameters in the local discontinuous Galerkin formulation. Our results indicate that these parameters severely affect the nature of the interaction with the unresolved scales (relating to points *(i)* and *(vi)*). Appropriately chosen parameters may yield highly localized closure relationships (Fig. 18) and unsuitably chosen parameters may yield closure relationships that are virtually impossible to model effectively (Figs. 15 and 16). With the adaptation of discontinuous Galerkin methods in the variational multiscale framework, we have shown that such considerations may serve as design driving factors. Additionally, it

provides the foundation for VMS-based fine-scale modeling appropriate for discontinuous Galerkin methods, for example via approximation of the fine-scale closure function.

References

[1] W. H. Reed and T. R. Hill, "Triangular Mesh Methods for the Neutron Transport Equation," *Proceedings of the American Nuclear Society*, vol. 836, pp. 1–23, 1973.

[2] P. Lesaint and P. Raviart, "On a Finite Element Method for Solving the Neutron Transport Equation," in *Mathematical Aspects of Finite Elements in Partial Differential Equations*, 1974.

[3] B. Cockburn, G. E. Karniadakis, and C.-W. Shu (eds.), *Discontinuous Galerkin methods: theory, computation and applications*, vol. 11 of *Lecture Notes in Computational Science and Engineering*. Berlin, Heidelberg: Springer Science & Business Media, 2000.

[4] B. Cockburn, "Discontinuous Galerkin methods for computational fluid dynamics," in *Encyclopedia of Computational Mechanics Second Edition* (E. Stein, R. Borst, and T. Hughes, eds.), vol. 5, pp. 141–203, Chichester, U.K.: John Wiley & Sons, Ltd., 2018.

[5] A. de Montlaur, S. Fernández-Méndez, and A. Huerta, "High-Order discontinuous Galerkin methods for incompressible flows," *European Conference on Computational Fluid Dynamics*, pp. 1–20, 2010.

[6] Z. J. Wang, K. Fidkowski, R. Abgrall, F. Bassi, D. Caraeni, A. Cary, H. Deconinck, R. Hartmann, K. Hillewaert, H. T. Huynh, N. Kroll, G. May, P.-O. Persson, B. van Leer, and M. Visbal, "High-Order CFD Methods: Current Status and Perspective," *International Journal for Numerical Methods in Fluids*, vol. 72, no. 8, pp. 811–845, 2013.

[7] "HiOCFD4 4th International Workshop on High-Order CFD Methods." https://how4.cenaero.be/, 2016.

[8] D. Mavriplis, C. Nastase, K. Shahbazi, L. Wang, and N. Burgess, "Progress in high-order Discontinuous Galerkin methods for aerospace applications," *47th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, vol. 601, 2009.

[9] J. Peraire and P.-O. Persson, "High-Order Discontinuous Galerkin Methods for CFD," in *Adaptive high-order methods in computational fluid dynamics* (Z. J. Wang, ed.), vol. 2, ch. 5, pp. 119–152, World Scientific, 2011.

[10] A. Huerta, A. Angeloski, X. Roca, and J. Peraire, "Efficiency of high-order elements for continuous and discontinuous Galerkin methods," *International Journal for Numerical Methods in Engineering*, vol. 96, no. 9, pp. 529–560, 2013.

[11] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, "Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems," *SIAM Journal on Numerical Analysis*, vol. 47, no. 2, pp. 1319–1365, 2009.

[12] D. Moro, N. C. Nguyen, and J. Peraire, "Navier-stokes solution using Hybridizable discontinuous Galerkin methods," in *20th AIAA Computational Fluid Dynamics Conference 2011*, p. 3407, 2011.

[13] R. M. Kirby, S. J. Sherwin, and B. Cockburn, "To CG or to HDG: A comparative study," *Journal of Scientific Computing*, vol. 51, no. 1, pp. 183–212, 2012.

[14] F. Brezzi, B. Cockburn, L. D. Marini, and E. Süli, "Stabilization mechanisms in discontinuous Galerkin finite element methods," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 25-28, pp. 3293–3310, 2006.

[15] T. J. R. Hughes, G. Engel, L. Mazzei, and M. G. Larson, "A comparison of discontinuous and continuous Galerkin methods based on error estimates, conservation, robustness and efficiency," in *Discontinuous Galerkin methods: theory, computation and applications* (B. Cockburn, G. E. Karniadakis, and C.-W. Shu, eds.), vol. 11 of *Lecture Notes in Computational Science and Engineering*, pp. 135–146, Springer, 2000.

[16] J. S. Hesthaven and T. Warburton, *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications.* Springer Science & Business Media., 2007.

[17] B. Cockburn, J. Gopalakrishnan, and H. Wang, "Locally conservative fluxes for the continuous Galerkin method," *SIAM Journal on Numerical Analysis*, vol. 45, pp. 1742–1776, 2007.

[18] T. J. R. Hughes, G. Engel, L. Mazzei, and M. G. Larson, "The continuous Galerkin method is locally conservative," *Journal of Computational Physics*, vol. 163, no. 2, pp. 467–488, 2000.

[19] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, "Unified analysis of discontinuous Galerkin methods for elliptic problems," *SIAM Journal on Numerical Analysis*, vol. 39, no. 5, pp. 1749–1779, 2002.

[20] B. Cockburn, "Devising discontinuous Galerkin methods for non-linear hyperbolic conservation laws," *Journal of Computational and Applied Mathematics*, vol. 128, no. 1-2, pp. 187–204, 2001.

[21] Y. Bazilevs and T. J. R. Hughes, "Weak imposition of Dirichlet boundary conditions in fluid mechanics," *Computers and Fluids*, vol. 36, no. 1, pp. 12–26, 2007.

[22] Y. Bazilevs, C. Michler, V. M. Calo, and T. J. R. Hughes, "Weak Dirichlet boundary conditions for wall-bounded turbulent flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 49-52, pp. 4853–4862, 2007.

[23] F. Xu, D. Schillinger, D. Kamensky, V. Varduhn, C. Wang, and M.-C. Hsu, "The tetrahedral finite cell method for fluids: Immersogeometric analysis of turbulent flow around complex geometries," *Computers & Fluids*, vol. 141, pp. 135–154, 2016.

[24] U. Langer, A. Mantzaflaris, S. E. Moore, and I. Toulopoulos, "Multipatch discontinuous Galerkin isogeometric analysis," in *Isogeometric Analysis and Applications* (B. Jüttler and B. Simeon, eds.), vol. 107 of *Lecture Notes in Computational Science and Engineering*, pp. 1–32, Springer, 2015.

[25] M. Ruess, D. Schillinger, A. I. Oezcan, and E. Rank, "Weak coupling for isogeometric analysis of non-matching and trimmed multi-patch geometries," *Computer Methods in Applied Mechanics and Engineering*, vol. 269, pp. 46–71, 2014.

[26] D. Schillinger, I. Harari, M.-C. Hsu, D. Kamensky, S. K. Stoter, Y. Yu, and Y. Zhao, "The non-symmetric Nitsche method for the parameter-free imposition of weak boundary and coupling conditions in immersed finite elements," *Computer Methods in Applied Mechanics and Engineering*, vol. 309, pp. 625–652, 2016.

[27] Y. Guo, J. Heller, T. J. Hughes, M. Ruess, and D. Schillinger, "Variationally consistent isogeometric analysis of trimmed thin shells at finite deformations, based on the step exchange format," *Computer Methods in Applied Mechanics and Engineering*, vol. 336, pp. 39–79, 2018.

[28] B. Cockburn, F. Li, and C. W. Shu, "Locally divergence-free discontinuous Galerkin methods for the Maxwell equations," *Journal of Computational Physics*, vol. 194, no. 2, pp. 588–610, 2004.

[29] B. Cockburn and F.-J. Sayas, "Divergence–conforming HDG methods for Stokes flow," *Mathematics of Computation*, vol. 83, pp. 1571–1598, 2014.

[30] J. A. Evans and T. J. R. Hughes, "Isogeometric divergence-conforming B-splines for the unsteady Navier–Stokes equations," *Journal of Computational Physics*, vol. 241, pp. 141–167, 2013.

[31] C. Lehrenfeld and J. Schöberl, "High order exactly divergence-free hybrid discontinuous Galerkin methods for unsteady incompressible flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 307, pp. 339–361, 2016.

[32] J. Baiges, R. Codina, and S. Idelsohn, "A domain decomposition strategy for reduced order models. Application to the incompressible Navier-Stokes equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 267, pp. 23–42, 2013.

[33] S. Kaulmann, M. Ohlberger, and B. Haasdonk, "A new local reduced basis discontinuous Galerkin approach for heterogeneous multiscale problems," *Comptes Rendus Mathematique*, vol. 349, no. 23-24, pp. 1233–1238, 2011.

[34] C. C. de Wiart, K. Hillewaert, L. Bricteux, and G. Winckelmans, "Implicit LES of free and wall-bounded turbulent flows based on the discontinuous Galerkin/symmetric interior penalty method," *International Journal for Numerical Methods in Fluids*, vol. 78, no. 6, pp. 335–354, 2015.

[35] B. Landmann, M. Kessler, S. Wagner, and E. Krämer, "A parallel, high-order discontinuous Galerkin

code for laminar and turbulent flows," *Computers & Fluids*, vol. 37, no. 4, pp. 427–438, 2008.

[36] N. C. Nguyen, P. O. Persson, and J. Peraire, "RANS solutions using high order discontinuous Galerkin methods," in *45th AIAA Aerospace Sciences Meeting and Exhibit*, no. 914, 2007.

[37] A. Frère, C. C. de Wiart, K. Hillewaert, P. Chatelain, and G. Winckelmans, "Application of wall-models to discontinuous Galerkin LES," *Physics of Fluids*, vol. 29, no. 8, p. 85111, 2017.

[38] F. Bassi, L. Botti, A. Colombo, A. Crivellini, A. Ghidoni, and F. Massa, "On the development of an implicit high-order discontinuous Galerkin method for DNS and implicit LES of turbulent flows," *European Journal of Mechanics - B/Fluids*, vol. 55, no. 2, pp. 367–379, 2016.

[39] K. Sengupta, F. Mashayek, and G. B. Jacobs, "Large-eddy simulation using a discontinuous Galerkin spectral element method," *45th AIAA Aerospace Sciences Meeting*, no. 402, 2007.

[40] A. Uranga, P.-O. Persson, M. Drela, and J. Peraire, "Implicit large eddy simulation of transition to turbulence at low Reynolds numbers using a discontinuous Galerkin method," *International Journal for Numerical Methods in Engineering*, vol. 87, no. 1-5, pp. 232–261, 2011.

[41] T. J. R. Hughes, A. A. Oberai, and L. Mazzei, "Large eddy simulation of turbulent channel flows by the variational multiscale method," *Physics of Fluids*, vol. 13, no. 6, pp. 1784–1799, 2001.

[42] T. J. R. Hughes, V. M. Calo, and G. Scovazzi, "Variational and Multiscale Methods in Turbulence," *Mechanics of the 21st Century*, pp. 153–163, 2005.

[43] Y. Bazilevs, V. M. Calo, J. A. Cottrell, T. J. R. Hughes, A. Reali, and G. Scovazzi, "Variational multi-scale residual-based turbulence modeling for large eddy simulation of incompressible flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 1-4, pp. 173–201, 2007.

[44] R. Codina, "Stabilized finite element approximation of transient incompressible flows using orthogonal subscales," *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 39-40, pp. 4295–4321, 2002.

[45] A. A. Oberai, J. Liu, D. Sondak, and T. J. R. Hughes, "A residual based eddy viscosity model for the large eddy simulation of turbulent flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 282, pp. 54–70, 2014.

[46] N. Ahmed, T. C. Rebollo, V. John, and S. Rubino, "A review of variational multiscale methods for the simulation of turbulent incompressible flows," *Archives of Computational Methods in Engineering*, vol. 24, no. 1, pp. 115–164, 2017.

[47] T. E. Tezduyar and Y. Osawa, "Finite element stabilization parameters computed from element matrices and vectors," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 3-4, pp. 411–430, 2000.

[48] G. Hauke and A. García-Olivares, "Variational subgrid scale formulations for the advection-diffusion-reaction equation," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 51-52, pp. 6847–6865, 2001.

[49] R. Codina, J. Principe, O. Guasch, and S. Badia, "Time dependent subscales in the stabilized finite element approximation of incompressible flow problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 21-24, pp. 2413–2430, 2007.

[50] C. Coley and J. A. Evans, "Variational multiscale modeling with discontinuous subscales: analysis and application to scalar transport," *Meccanica*, vol. 53, no. 6, pp. 1241–1269, 2018.

[51] A. Masud and R. Calderer, "A variational multiscale method for incompressible turbulent flows: Bubble functions and fine scale fields," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 33-36, pp. 2577–2593, 2011.

[52] V. Gravemeier, W. A. Wall, and E. Ramm, "Large eddy simulation of turbulent incompressible flows by a three-level finite element method," *International Journal for Numerical Methods in Fluids*, vol. 48, no. 10, pp. 1067–1099, 2005.

[53] T. J. R. Hughes, G. Scovazzi, and L. P. Franca, "Multiscale and Stabilized Methods," in *Encyclopedia of Computational Mechanics* (E. Stein, R. De Borst, and T. J. R. Hughes, eds.), vol. 3, ch. 4, pp. 1–64, John Wiley & Sons, Ltd, second ed., 2018.

[54] T. J. R. Hughes, "Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods," *Computer Methods in Applied*

Mechanics and Engineering, vol. 127, no. 1-4, pp. 387–401, 1995.

[55] T. J. R. Hughes, G. R. Feijóo, L. Mazzei, and J. B. Quincy, "The variational multiscale method - A paradigm for computational mechanics," *Computer Methods in Applied Mechanics and Engineering*, vol. 166, no. 1-2, pp. 3–24, 1998.

[56] F. Brezzi, L. P. Franca, T. J. R. Hughes, and A. Russo, "$b = \int G.$," *Computer Methods in Applied Mechanics and Engineering*, vol. 145, no. 145, pp. 329–339, 1997.

[57] T. J. R. Hughes and G. Sangalli, "Variational multiscale analysis: The fine-scale Green's function, projection, optimization, localization, and stabilized methods," *SIAM Journal on Numerical Analysis*, vol. 45, no. 2, pp. 539–557, 2007.

[58] P. Bochev, T. J. R. Hughes, and G. Scovazzi, "A multiscale discontinuous Galerkin method," in *International Conference on Large-Scale Scientific Computing*, pp. 84–93, Springer, 2005.

[59] T. J. R. Hughes, G. Scovazzi, P. B. Bochev, and A. Buffa, "A multiscale discontinuous Galerkin method with the computational structure of a continuous Galerkin method," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 19, pp. 2761–2787, 2006.

[60] A. N. Brooks and T. J. R. Hughes, "Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 32, no. 1-3, pp. 199–259, 1982.

[61] A. Masud and T. J. R. Hughes, "A stabilized mixed finite element method for Darcy flow," *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 39-40, pp. 4341–4370, 2002.

[62] K. B. Nakshatrala, D. Z. Turner, K. D. Hjelmstad, and A. Masud, "A stabilized mixed finite element method for Darcy flow based on a multiscale decomposition of the solution," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 33-36, pp. 4036–4049, 2006.

[63] T. J. R. Hughes, L. P. Franca, and M. Balestra, "A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation of the stokes problem accommodating equal-order interpolations," *Computer Methods in Applied Mechanics and Engineering*, vol. 59, no. 1, pp. 85–99, 1986.

[64] M. Braack and E. Burman, "Local projection stabilization for the Oseen problem and its interpretation as a variational multiscale method," *SIAM Journal on Numerical Analysis*, vol. 43, no. 6, pp. 2544–2566, 2006.

[65] M. Braack, E. Burman, V. John, and G. Lube, "Stabilized finite element methods for the generalized Oseen problem," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 4-6, pp. 853–866, 2007.

[66] R. Codina, "Analysis of a stabilized finite element approximation of the Oseen equations using orthogonal subscales," *Applied Numerical Mathematics*, vol. 58, no. 3, pp. 264–283, 2008.

[67] C. Johnson, U. Nävert, and J. Pitkäranta, "Finite element methods for linear hyperbolic problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 45, no. 1-3, 1984.

[68] I. Akkerman, Y. Bazilevs, V. M. Calo, T. J. R. Hughes, and S. Hulshoff, "The role of continuity in residual-based variational multiscale modeling of turbulence," *Computational Mechanics*, vol. 41, no. 3, pp. 371–378, 2008.

[69] S. K. F. Stoter, S. R. Turteltaub, S. J. Hulshoff, and D. Schillinger, "A discontinuous Galerkin residual-based variational multiscale method for modeling subgrid-scale behavior of the viscous Burgers equation," *International Journal for Numerical Methods in Fluids*, vol. 88, no. 5, pp. 217–238, 2018.

[70] K. E. Jansen, S. S. Collis, C. Whiting, and F. Shakib, "A better consistency for low-order stabilized finite element methods," *Computer Methods in Applied Mechanics and Engineering*, vol. 174, no. 1-2, pp. 153–170, 1999.

[71] T. J. R. Hughes and J. R. Stewart, "A space-time formulation for multiscale phenomena," *Journal of Computational and Applied Mathematics*, vol. 74, no. 1-2, pp. 217–229, 1996.

[72] S. K. F. Stoter, M. F. P. ten Eikelder, F. de Prenter, I. Akkerman, E. H. van Brummelen, C. V. Verhoosel, and D. Schillinger, "Nitsche's method as a variational multiscale formulation and a resulting boundary layer fine-scale model," *Computer Methods in Applied Mechanics and Engineering*, vol. 382, 2021.

[73] S. K. F. Stoter, S. R. Turteltaub, S. J. Hulshoff, and D. Schillinger, "Residual-based variational multiscale modeling in a discontinuous Galerkin framework," *Multiscale Modeling and Simulation*, vol. 16, no. 3, pp. 1333–1364, 2018.

[74] R. Codina, J. Principe, and J. Baiges, "Subscales on the element boundaries in the variational two-scale finite element method," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 5-8, pp. 838–852, 2009.

[75] R. Codina and J. Baiges, "Finite element approximation of transmission conditions in fluids and solids introducing boundary subgrid scales," *International Journal for Numerical Methods in Engineering*, vol. 87, no. 1-5, pp. 386–411, 2011.

[76] B. Cockburn and C. W. Shu, "The local discontinuous Galerkin method for time-dependent convection-diffusion systems," *SIAM Journal on Numerical Analysis*, vol. 35, no. 6, pp. 2440–2463, 1998.

[77] F.-J. Sayas, T. S. Brown, and M. E. Hassell, *Variational Techniques for Elliptic Partial Differential Equations: Theoretical Tools and Advanced Applications*. CRC Press, 2019.

[78] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, vol. 44. Springer, 2013.

[79] D. N. Arnold, R. S. Falk, and R. Winther, "Finite element exterior calculus: From Hodge theory to numerical stability," *Bulletin of the American Mathematical Society*, vol. 47, no. 2, pp. 281–354, 2010.

[80] D. N. Arnold, P. B. Bochev, R. B. Lehoucq, R. A. Nicolaides, and M. Shashkov, *Compatible Spatial Discretizations*, vol. 142. Springer Science & Business Media, 2006.

[81] I. Babuška, "The finite element method with Lagrangian multipliers," *Numerische Mathematik*, vol. 20, no. 3, pp. 179–192, 1973.

[82] I. Babuška and G. N. Gatica, "On the mixed finite element method with Lagrange multipliers," *Numerical Methods for Partial Differential Equations*, vol. 19, no. 2, pp. 192–210, 2003.

[83] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, "An a priori error analysis of the local discontinuous Galerkin method for elliptic problems," *SIAM Journal on Numerical Analysis*, vol. 38, pp. 1676–1706, 2000.

[84] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau, "Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids," *SIAM Journal on Numerical Analysis*, vol. 39, no. 1, pp. 264–285, 2002.

[85] F. Bassi and S. Rebay, "A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations," *Journal of Computational Physics*, vol. 131, no. 2, pp. 267–279, 1997.

[86] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini, "A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows," in *Proceedings of 2nd European Conference on Turbomachinery, Fluid Dynamics and Thermodynamicst* (R. Decuypere and G. Dibelius, eds.), (Antwerpen, Belgium), pp. 99–108, 1997.

[87] B. Cockburn and B. Dong, "An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems," *Journal of Scientific Computing*, vol. 32, no. 2, pp. 233–262, 2007.

[88] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, "A projection-based error analysis of HDG methods," *Mathematics of Computation*, vol. 79, no. 271, pp. 1351–1367, 2010.

[89] G. Fu, W. Qiu, and W. Zhang, "An analysis of HDG methods for convection-dominated diffusion problems," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 49, no. 1, 2015.

[90] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, "An a priori error analysis of the local discontinuous Galerkin method for elliptic problems," *SIAM Journal on Numerical Analysis*, vol. 38, no. 5, pp. 1676–1706, 2001.

[91] J. Peraire and P.-O. Persson, "The compact discontinuous Galerkin (CDG) method for elliptic problems," *SIAM Journal on Scientific Computing*, vol. 30, no. 4, pp. 1806–1824, 2007.

[92] M. F. P. ten Eikelder, Y. Bazilevs, and I. Akkerman, "A theoretical framework for discontinuity capturing: Joining variational multiscale analysis and variation entropy theory," *Computer Methods in Applied Mechanics and Engineering*, p. 112664, 2019.