

# Nitsche's method as a variational multiscale formulation and a resulting boundary layer fine-scale model

Stein K.F. Stoter<sup>a,\*</sup>, Marco F.P. ten Eikelder<sup>b</sup>, Frits de Prenter<sup>c</sup>, Ido Akkerman<sup>b</sup>,  
E. Harald van Brummelen<sup>c</sup>, Clemens V. Verhoosel<sup>c</sup>, Dominik Schillinger<sup>a</sup>

<sup>a</sup> Institute of Mechanics and Computational Mechanics, Leibniz Universität Hannover, Hanover, Germany

<sup>b</sup> Department of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands

<sup>c</sup> Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

Received 14 September 2019; received in revised form 1 February 2021; accepted 18 April 2021

Available online 5 May 2021

## Abstract

We show that in the variational multiscale framework, the weak enforcement of essential boundary conditions via Nitsche's method corresponds directly to a particular choice of projection operator. The consistency, symmetry and penalty terms of Nitsche's method all originate from the fine-scale closure dictated by the corresponding scale decomposition. As a result of this formalism, we are able to determine the exact fine-scale contributions in Nitsche-type formulations. In the context of the advection–diffusion equation, we develop a residual-based model that incorporates the non-vanishing fine scales at the Dirichlet boundaries. This results in an additional boundary term with a new model parameter. We then propose a parameter estimation strategy for all parameters involved that is also consistent for higher-order basis functions. We illustrate with numerical experiments that our new augmented model mitigates the overly diffusive behavior that the classical residual-based fine-scale model exhibits in boundary layers at boundaries with weakly enforced essential conditions.

© 2021 Elsevier B.V. All rights reserved.

**Keywords:** Variational multiscale method; Nitsche's method; Weak boundary conditions; Boundary layer accuracy; Fine-scale Green's function; Higher-order basis functions

## 1. Introduction

The variational multiscale (VMS) method was established in the 1990s by Hughes and coworkers as a universal framework for developing and classifying stabilized methods [1–5]. It was hypothesized that the unresolved fine-scale nature of the solution of the partial differential equation is of key importance for the stability of finite-element schemes. The VMS methodology offers a means to ascertain the effect of the fine-scale solution onto the resolved finite element solution. As exact expressions for the fine scales are often not available, these fine-scale effects must be modeled. A particularly prevalent class of fine-scale models is that of the residual-based

\* Correspondence to: Institute of Mechanics and Computational Mechanics, Leibniz Universität Hannover, Appelstraße 9a, 30167 Hanover, Germany.

E-mail addresses: [Stein.Stoter@ibnm.uni-hannover.de](mailto:Stein.Stoter@ibnm.uni-hannover.de) (S.K.F. Stoter), [M.F.P.tenEikelder@tudelft.nl](mailto:M.F.P.tenEikelder@tudelft.nl) (M.F.P. ten Eikelder), [F.d.Prenter@tue.nl](mailto:F.d.Prenter@tue.nl) (F. de Prenter), [I.Akkerman@tudelft.nl](mailto:I.Akkerman@tudelft.nl) (I. Akkerman), [E.H.v.Brummelen@tue.nl](mailto:E.H.v.Brummelen@tue.nl) (E.H. van Brummelen), [C.V.Verhoosel@tue.nl](mailto:C.V.Verhoosel@tue.nl) (C.V. Verhoosel), [Schillinger@ibnm.uni-hannover.de](mailto:Schillinger@ibnm.uni-hannover.de) (D. Schillinger).

models [6,7]. By assuming vanishing fine scales on element boundaries, the associated fine-scale problem can be solved on each element locally. The fine-scale solution is then approximated as the residual of the coarse-scale finite element solution, multiplied by the averaged Green's function. It has been shown that the resulting stabilized formulation is closely related (and often equivalent) to classical stabilized methods [8,9]. Typical examples include streamline-upwind Petrov–Galerkin (SUPG) [8–11], Galerkin least-squares (GLS) [9,11,12], and pressure-stabilized Petrov–Galerkin (PSPG) methods [13]. More recently, residual-based modeling of the fine scales has found its use as an effective turbulence model for finite element implementations of the Navier–Stokes equations [7,14–18]. Of particular interest to our current work are the fine-scale models that include element boundaries contributions [19,20]. Fine-scale modeling techniques have also been developed for capturing sharp internal and boundary layers [5,21–25]. Since all these models are mathematically inspired by the fine-scale equations, their parameters are clearly defined and they yield consistent formulations. Both these points are in contrast with typical eddy viscosity models, which are phenomenologically inspired and variationally inconsistent.

Another approach that has been shown to yield favorable results for fluid-mechanics applications is the use of weakly imposed Dirichlet boundary conditions [26–30]. Typically, Nitsche's method is the method of choice for weakly enforcing essential boundary conditions. While Nitsche's method was initially proposed in relation to energy minimization functionals [31], both its symmetric and nonsymmetric variants have since been studied extensively in fluid-mechanics applications [27,28,30,32]. One of the main drivers for the significant recent interest in weakly enforced boundary conditions is their importance in immersed finite element methods. Notable references in the context of fluid mechanics include [33–35], and specifically in the context of isogeometric analysis [36–41]. For immersed finite element methods, the approximation space is no longer tailored to fit the domain boundary. The essential boundary conditions can thus not easily be satisfied strongly. Hence, there is a need for weakly enforcing the Dirichlet boundary condition in the weak formulation.

At first glance, the variational multiscale method and Nitsche's method appear to be at conflict: the basis for a variational multiscale decomposition is a well-posed continuous weak formulation, but Nitsche's method involves spaces that are unconstrained at Dirichlet boundaries and requires penalty terms that become unbounded in the continuous limit. Additionally, the fine-scale solution does, by design, not vanish on the Dirichlet boundary, which violates one of the key assumptions on which traditional residual-based fine-scale models are built. In previous work, we focused on discontinuous Galerkin methods, where the discontinuities between elements give rise to similar issues [42,43]. The goal of this article is to completely eliminate these issues for weak boundary imposition in Nitsche-type formulations.

The remainder of this article is structured as follows. In Section 2, we derive a variational multiscale finite element formulation of the advection–diffusion equation and we show that Nitsche's method arises from a particular choice of fine-scale closure. In Section 3, we develop the fine-scale model that takes into account the non-vanishing fine scales at the Dirichlet boundary and provide estimates for the involved model parameters. The complete formulation is summarized in Section 3.5. Next, in Section 4, we show that the resulting bilinear form is coercive. In Section 5, we verify the theory for a one-dimensional model problem, and in Section 6 we computationally investigate the performance for a two-dimensional model problem that involves multiple boundary layers. In Section 7, we present concluding remarks.

## 2. A variational multiscale derivation of the finite element formulation

The classical model problem for variational multiscale analysis is the steady advection–diffusion equation. Let  $\Omega$  denote the spatial domain with boundary  $\partial\Omega$ . The governing equations in strong form read:

$$\mathbf{a} \cdot \nabla \phi - \nabla \cdot \kappa \nabla \phi = f \quad \text{in } \Omega, \quad (1a)$$

$$\phi = \phi_D \quad \text{on } \partial\Omega_D, \quad (1b)$$

$$\kappa \partial_n \phi = g_N \quad \text{on } \partial\Omega_N^+, \quad (1c)$$

$$\kappa \partial_n \phi - \mathbf{a} \cdot \mathbf{n} \phi = g_N \quad \text{on } \partial\Omega_N^-, \quad (1d)$$

where we consider source functions  $f \in H^{-1}(\Omega)$ , Dirichlet data  $\phi_D \in H^{\frac{1}{2}}(\partial\Omega_D)$  and Neumann (or Robin) data  $g_N \in H^{-\frac{1}{2}}(\partial\Omega_N)$ . The advective velocity  $\mathbf{a} = \mathbf{a}(\mathbf{x})$  is a given solenoidal vector field ( $\nabla \cdot \mathbf{a} = 0$ ) and the diffusivity  $\kappa$  is a strictly positive constant. The Dirichlet and Neumann (or Robin) boundaries  $\partial\Omega_D$  and  $\partial\Omega_N$  are complementary

subsets of the boundary  $\partial\Omega$ , i.e.  $\partial\Omega = \overline{\partial\Omega_D} \cup \overline{\partial\Omega_N}$ , and the superscripts  $+$  and  $-$  indicate outflow ( $\mathbf{a} \cdot \mathbf{n} \geq 0$ ) and inflow parts ( $\mathbf{a} \cdot \mathbf{n} < 0$ ) respectively. On the boundary, the normal gradient is denoted  $\partial_n \phi = \mathbf{n} \cdot \nabla \phi$ . By convention,  $\mathbf{n}$  denotes the outward facing unit normal vector.

Approximation of the solution  $\phi$  by means of a finite element method is challenging when the advection dominates. Residual-based stabilization techniques can be highly effective, but they become insufficient when steep boundary layers are under-resolved. In [27], this problem is addressed by using Nitsche's method to weakly impose the essential boundary conditions. The following finite element formulation is proposed:

$$\begin{aligned} \text{Find } \phi^h \in \mathcal{W}^h \text{ s.t. } \forall w^h \in \mathcal{W}^h : \\ - (\mathbf{a} \cdot \nabla w^h, \phi^h)_{\Omega} + (\nabla w^h, \kappa \nabla \phi^h)_{\Omega} - (\tau \mathbf{a} \cdot \nabla w^h, \mathcal{R}_{\phi^h})_{\tilde{\Omega}} + \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi^h \rangle_{\partial\Omega^+} \\ - \langle \kappa \partial_n \phi^h, w^h \rangle_{\partial\Omega_D} - \langle \kappa \partial_n w^h, \phi^h \rangle_{\partial\Omega_D} + \langle \kappa \beta w^h, \phi^h \rangle_{\partial\Omega_D} = (w^h, f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi_D \rangle_{\partial\Omega_D^-} \\ + \langle g_N, w^h \rangle_{\partial\Omega_N} - \langle \kappa \partial_n w^h, \phi_D \rangle_{\partial\Omega_D} + \langle \kappa \beta w^h, \phi_D \rangle_{\partial\Omega_D}, \end{aligned} \quad (2)$$

where  $(\cdot, \cdot)_{\Omega}$  denotes the  $L^2$ -inner product on domain  $\Omega$ , and  $\langle \cdot, \cdot \rangle_{\partial\Omega}$  denotes the duality pairing with functions defined on the indicated boundary. The finite element approximation space  $\mathcal{W}^h$  is constructed on a mesh  $\mathcal{T} = \{K_i\}_{i=1}^{N_{el}}$ , and  $\tilde{\Omega}$  is the union of open element domains.

The third term in Eq. (2) represents the SUPG stabilization term, where  $\tau$  is the stabilization parameter and  $\mathcal{R}_{\phi^h} = \mathbf{a} \cdot \nabla \phi^h - \nabla \cdot \kappa \nabla \phi^h - f$  the residual. From the variational multiscale perspective, this term represents the interaction with the unresolved fine scales. All terms on the Dirichlet boundary correspond to Nitsche's method. Straightforward addition of these Nitsche terms to the SUPG-stabilized formulation goes against the philosophy of the variational multiscale methodology: ideally, all terms in the finite element formulation follow from a scale decomposition of the weak statement of the PDE. Our objective is to derive Eq. (2) completely from a variational multiscale treatment of Eq. (1). This sets the stage for an improved fine-scale model that includes boundary effects.

### 2.1. Variational multiscale weak formulation

The first step of deriving Eq. (2) from a VMS perspective is to obtain a suitable weak formulation of Eq. (1). Normally, one would take the standard weak formulation of the PDE which makes use of  $H_0^1(\Omega)$  Sobolev spaces for the test and trial functions. However, such a weak formulation is not suitable for our objective of deriving Eq. (2) since the finite element space  $\mathcal{W}^h$  in Eq. (2) is unconstrained on the Dirichlet boundary and hence not a subspace of  $H_0^1(\Omega)$ . This makes the Nitsche method non-conformal. Instead, we use  $H^1(\Omega)$  as our test and trial space. The essential condition must now be enforced with a Lagrange multiplier:

$$\begin{aligned} \text{Find } (\phi, \lambda) \in \mathcal{W} \times \mathcal{Q} \text{ s.t. } \forall (w, q) \in \mathcal{W} \times \mathcal{Q} : \\ - (\mathbf{a} \cdot \nabla w, \phi)_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w, \phi \rangle_{\partial\Omega^+} + (\nabla w, \kappa \nabla \phi)_{\Omega} + \langle \lambda, w \rangle_{\partial\Omega_D} \\ = (w, f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w, \phi_D \rangle_{\partial\Omega_D^-} + \langle g_N, w \rangle_{\partial\Omega_N} \end{aligned} \quad (3a)$$

$$\langle q, \phi \rangle_{\partial\Omega_D} = \langle q, \phi_D \rangle_{\partial\Omega_D}, \quad (3b)$$

where  $\mathcal{W} = H^1(\Omega)$  and  $\mathcal{Q} = H^{-\frac{1}{2}}(\partial\Omega_D)$ . Equivalence of the strong form and this hybrid formulation dictates that  $\lambda = -\kappa \partial_n \phi$ .

Stabilized finite element formulations based on the hybrid form in Eq. (3) have existed for decades [44]. However, we do not intent to discretize the Lagrange multiplier: our objective of deriving Nitsche's method means that we have to eliminate this auxiliary field later on. We start with a decomposition of the trial and test function spaces into coarse and fine scales:

$$\mathcal{W} \times \mathcal{Q} = (\mathcal{W} \times \mathcal{Q})^h \oplus (\mathcal{W} \times \mathcal{Q})', \quad (4)$$

We note that, in general, the coarse-scale and fine-scale spaces themselves do not decouple as Cartesian products [45]. In fact, we will make use of a decomposition for which  $(\mathcal{W} \times \mathcal{Q})^h \neq \mathcal{W}^h \times \mathcal{Q}^h$ . Rather, since we aim to derive a finite element formulation set in  $\mathcal{W}^h$ , we must have  $\dim((\mathcal{W} \times \mathcal{Q})^h) = \dim(\mathcal{W}^h)$ .

The components of the solutions and test functions decouple as

$$(\phi, \lambda) = (\phi^h, \lambda^h) + (\phi', \lambda'), \quad (5a)$$

$$(w, q) = (w^h, q^h) + (w', q'), \quad (5b)$$

with coarse scales  $(\phi^h, \lambda^h) \in (\mathcal{W} \times \mathcal{Q})^h$  and fine scales  $(\phi', \lambda') \in (\mathcal{W} \times \mathcal{Q})'$ . The direct sum decomposition in Eq. (4) is associated with a projection operator:

$$(w^h, q^h) = \mathcal{P}(w, q) \in (\mathcal{W} \times \mathcal{Q})^h, \quad (6a)$$

$$(w', q') = (\mathcal{I} - \mathcal{P})(w, q) \in (\mathcal{W} \times \mathcal{Q})', \quad (6b)$$

where  $\mathcal{P} : \mathcal{W} \times \mathcal{Q} \rightarrow (\mathcal{W} \times \mathcal{Q})^h$  is the projector and  $\mathcal{I} : \mathcal{W} \times \mathcal{Q} \rightarrow \mathcal{W} \times \mathcal{Q}$  is the identity operator. Formally, this projector is incorporated in the weak formulation through the definition of the fine-scale space  $(\mathcal{W} \times \mathcal{Q})'$ . It follows from Eq. (6b) that  $(\mathcal{W} \times \mathcal{Q})' = \text{Im}(\mathcal{I} - \mathcal{P}) = \text{Ker} \mathcal{P}$ . Then, the direct sum decomposition of Eq. (4) ensures the unique decomposition that satisfies Eqs. (5) and (6). Using this two-scale split we arrive at the following alternative – equivalent – weak statement:

Find  $(\phi^h, \lambda^h) \in (\mathcal{W} \times \mathcal{Q})^h$  and  $(\phi', \lambda') \in (\mathcal{W} \times \mathcal{Q})'$  s.t.  $\forall (w^h, q^h) \in (\mathcal{W} \times \mathcal{Q})^h$  and  $\forall (w', q') \in (\mathcal{W} \times \mathcal{Q})'$ :

$$-(\mathbf{a} \cdot \nabla w^h, \phi^h + \phi')_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi^h + \phi' \rangle_{\partial \Omega^+} + (\nabla w^h, \kappa \nabla \phi^h + \kappa \nabla \phi')_{\Omega} + \langle \lambda^h + \lambda', w^h \rangle_{\partial \Omega_D} = (w^h, f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi_D \rangle_{\partial \Omega_D^-} + \langle g_N, w^h \rangle_{\partial \Omega_N}, \quad (7a)$$

$$-(\mathbf{a} \cdot \nabla w', \phi^h + \phi')_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w', \phi^h + \phi' \rangle_{\partial \Omega^+} + (\nabla w', \kappa \nabla \phi^h + \kappa \nabla \phi')_{\Omega} + \langle \lambda^h + \lambda', w' \rangle_{\partial \Omega_D} = (w', f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w', \phi_D \rangle_{\partial \Omega_D^-} + \langle g_N, w' \rangle_{\partial \Omega_N}, \quad (7b)$$

$$\langle q^h, \phi^h + \phi' \rangle_{\partial \Omega_D} = \langle q^h, \phi_D \rangle_{\partial \Omega_D}, \quad (7c)$$

$$\langle q', \phi^h + \phi' \rangle_{\partial \Omega_D} = \langle q', \phi_D \rangle_{\partial \Omega_D}. \quad (7d)$$

Eqs. (7a) and (7c) are the ‘coarse-scale problem’, and can be interpreted as a relation for  $(\phi^h, \lambda^h)$  for a given  $(\phi', \lambda')$ . Similarly, the ‘fine-scale problem’ of Eqs. (7b) and (7d) is a relation for the fine-scale component  $(\phi', \lambda') \in (\mathcal{W} \times \mathcal{Q})'$  for given coarse scales  $(\phi^h, \lambda^h)$ . This space, however, is infinite-dimensional and is thus not amenable to discrete implementation. Hence, in Eq. (7a) a closure model will be substituted in place of the fine-scale solution.

## 2.2. Nitsche’s method as a partial fine-scale closure

Next, we show that a particular scale decomposition (i.e. projector) leads to Nitsche’s classical formulation. Our goal is to illustrate that Nitsche’s method and the accompanying penalty terms are not in conflict with the VMS theory, but rather can be interpreted naturally in the VMS framework as a particular choice of fine-scale closure. Consider the following projection operator, which we will refer to as the Nitsche projector:

$$\begin{aligned} \mathcal{P}_N : \mathcal{W} \times \mathcal{Q} &\rightarrow (\mathcal{W} \times \mathcal{Q})^h \\ (\phi, \lambda) &\mapsto (\phi^h, -\kappa \partial_n \phi^h), \end{aligned} \quad (8)$$

with

$$\begin{aligned} \phi^h = \arg \inf_{\phi^h \in \mathcal{W}^h} & \frac{1}{2} (\kappa (\nabla \phi - \nabla \phi^h), \nabla \phi - \nabla \phi^h)_{\Omega} + \langle \lambda + \kappa \partial_n \phi^h, \phi - \phi^h \rangle_{\partial \Omega_D} \\ & + \frac{1}{2} \langle \kappa \beta (\phi - \phi^h), \phi - \phi^h \rangle_{\partial \Omega_D} + \frac{1}{2} \langle \mathbf{a} \cdot \mathbf{n} (\phi - \phi^h), \phi - \phi^h \rangle_{\partial \Omega^+}. \end{aligned} \quad (9)$$

With these definitions,  $\mathcal{P}_N$  is a linear and bounded operator on all of  $\mathcal{W} \times \mathcal{Q}$ . Moreover, it is idempotent, making it a projector. For Eq. (9) to yield a unique coarse-scale solution for arbitrary data (that is, an arbitrary pair  $(\phi, \lambda)$ ), there is a lower limit to  $\beta$  [31,46].

The range of  $\mathcal{P}_N$  can be identified as:

$$\text{Ran}(\mathcal{P}_N) =: (\mathcal{W} \times \mathcal{Q})^h = \{(w, -\kappa \partial_n w) : w \in \mathcal{W}^h\}, \quad (10)$$

such that indeed  $\dim(\mathcal{W} \times \mathcal{Q})^h = \dim(\mathcal{W}^h)$ . This is a result of the strict relation between the coarse-scale solutions  $\phi^h$  and  $\lambda^h$ . From Eq. (8) we have:

$$\lambda^h = -\kappa \partial_n \phi^h, \quad (11a)$$

$$q^h = -\kappa \partial_n w^h. \quad (11b)$$

By taking the Gâteaux derivative of the potential in Eq. (9) and subsequently replacing  $\phi - \phi^h$  by  $\phi'$  and  $\lambda + \kappa \partial_n \phi^h$  by  $\lambda'$ , we obtain the following associated optimality condition for the fine scales:

$$\begin{aligned} & -(\nabla v^h, \kappa \nabla \phi')_{\Omega} - \langle \lambda', v^h \rangle_{\partial \Omega_D} + \langle \kappa \partial_n v^h, \phi' \rangle_{\partial \Omega_D} \\ & - \langle \kappa \beta v^h, \phi' \rangle_{\partial \Omega_D} - \langle \mathbf{a} \cdot \mathbf{n} v^h, \phi' \rangle_{\partial \Omega^+} = 0 \quad \forall v^h \in \mathcal{W}^h. \end{aligned} \quad (12)$$

Typically, in the variational multiscale framework, one would attempt to invert the fine-scale problem of Eq. (7b) while satisfying the requirement posed by Eq. (12). However, the particular structures of the coarse-scale problem and the Nitsche projector allow for a more direct closure of (part of) the fine scales in the coarse-scale equation [43]. After making the choice  $v^h = w^h$  in Eq. (12) we can directly substitute the identities of Eqs. (11) and (12) into Eqs. (7a) and (7c) to arrive at:

$$\begin{aligned} & -(\mathbf{a} \cdot \nabla w^h, \phi^h + \phi')_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi^h \rangle_{\partial \Omega^+} + (\nabla w^h, \kappa \nabla \phi^h)_{\Omega} - \langle \kappa \partial_n \phi^h, w^h \rangle_{\partial \Omega_D} - \langle \kappa \partial_n w^h, \phi^h \rangle_{\partial \Omega_D} \\ & - \langle \kappa \beta w^h, \phi' \rangle_{\partial \Omega_D} = (w^h, f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi_D \rangle_{\partial \Omega_D^-} + \langle g_N, w^h \rangle_{\partial \Omega_N} - \langle \kappa \partial_n w^h, \phi_D \rangle_{\partial \Omega_D}, \end{aligned} \quad (13)$$

which holds for all test functions  $w^h \in \mathcal{W}^h$ .

Lastly, we use the identity  $\phi' = \phi - \phi^h = \phi_D - \phi^h$  on the Dirichlet boundary. Then, we are left with Nitsche's formulation of the advection–diffusion problem:

Find  $\phi^h \in \mathcal{W}^h$  s.t.  $\forall w^h \in \mathcal{W}^h$ :

$$\begin{aligned} & -(\mathbf{a} \cdot \nabla w^h, \phi^h + \phi')_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi^h \rangle_{\partial \Omega^+} + (\nabla w^h, \kappa \nabla \phi^h)_{\Omega} - \langle \kappa \partial_n \phi^h, w^h \rangle_{\partial \Omega_D} \\ & - \langle \kappa \partial_n w^h, \phi^h \rangle_{\partial \Omega_D} + \langle \kappa \beta w^h, \phi^h \rangle_{\partial \Omega_D} = (w^h, f)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi_D \rangle_{\partial \Omega_D^-} + \langle g_N, w^h \rangle_{\partial \Omega_N} \\ & - \langle \kappa \partial_n w^h, \phi_D \rangle_{\partial \Omega_D} + \langle \kappa \beta w^h, \phi_D \rangle_{\partial \Omega_D}, \end{aligned} \quad (14)$$

where only a single (advective) fine-scale term remains. We emphasize that both the symmetric term of Nitsche's method as well as its penalty term find their origin in the scale decomposition per the projection operator of Eqs. (8) and (9).

**Remark 1.** If we write the remaining fine-scale term as  $(\mathcal{L}_{\text{stab}} w^h, \phi')_{\Omega}$ , then we observe that the appropriate fine-scale model (read: stabilization technique) in conjunction with Nitsche's method involves the SUPG operator  $\mathcal{L}_{\text{stab}} = -\mathbf{a} \cdot \nabla$ . Typically, in variational multiscale stabilized methods, the adjoint differential operator  $\mathcal{L}_{\text{stab}}^* = -\mathbf{a} \cdot \nabla - \nabla \cdot \kappa \nabla$  is proposed. However, since the diffusive part of the fine-scale terms has vanished due to the small scale orthogonality given in Eq. (12), the diffusive operator no longer occurs in the fine-scale term.

### 2.3. Analysis of the Nitsche projector

The projector defines the scale decomposition of Eq. (4) via Eq. (6). This means that it impacts the appropriate modeling choices in the development of the fine-scale model. Before continuing the fine-scale modeling, we thus first dedicate a short study on the Nitsche projector. The ensuing analysis will heavily rely on the work by Hughes and Sangalli in [47], who performed the same analysis using the  $H_0^1$  and  $L^2$  projectors.

A projector may be defined by a finite number of functional constraints; as many as the dimension of  $\mathcal{W}^h$ :

$$\begin{aligned} \mathcal{P}_N(\phi, \lambda) &= (\phi^h, -\kappa \partial_n \phi^h) \text{ such that:} \\ \langle \mu_i, (\phi - \phi^h, \lambda + \kappa \partial_n \phi^h) \rangle &= 0 \quad i = 1, \dots, \dim(\mathcal{W}^h). \end{aligned} \quad (15)$$

The  $\mu_i$  functionals lie in the dual space of  $\mathcal{W} \times \mathcal{Q}$ . For the Nitsche projector they may be inferred from the weak statement of Eq. (12). By performing integration by parts on its first term while interpreting the resulting integral in the sense of distributions, the functional constraints follow from substituting the various candidates for  $v^h$ . For a one-dimensional domain with a set of nodes  $\Gamma = \{x_1, x_2, \dots, x_n\}$  and element domains  $\mathcal{T} = \{[x_1, x_2], \dots, [x_{n-1}, x_n]\}$

and a typical nodal finite element construction of  $\mathcal{W}^h$  with polynomial order  $P$ , we find:

$\mathcal{P}_N(\phi, \lambda) = (\phi^h, -\kappa \partial_n \phi^h)$  such that:

$$\phi(x_i) - \phi^h(x_i) = 0 \quad \text{for } x_i \in \Gamma \setminus \partial\Omega_D, \quad (16a)$$

$$\int_K (\phi - \phi^h) x^p = 0 \quad \text{for } K \in \mathcal{T} \text{ and } 0 \leq p \leq P-2, \quad (16b)$$

$$(\lambda + \kappa \partial_n \phi^h)|_{x_i} + \kappa \beta (\phi - \phi^h)|_{x_i} = 0 \quad \text{for } x_i \in \partial\Omega_D^-, \quad (16c)$$

$$(\lambda + \kappa \partial_n \phi^h)|_{x_i} + (\kappa \beta + \mathbf{a} \cdot \mathbf{n})(\phi - \phi^h)|_{x_i} = 0 \quad \text{for } x_i \in \partial\Omega_D^+. \quad (16d)$$

The first requirement dictates nodal exactness of the finite element formulation, and together with the second requirement these coincide with the constraints from the  $H_0^1$  projector [47]. The last two originate from the extra degrees of freedom on the Dirichlet boundary.

**Remark 2.** The central role of the projector can be useful for the interpretation of the obtained finite element solution. For instance, for the current example of Nitsche's method, a better approximation of the exact diffusive flux (the solution field  $\lambda$ ) on the Dirichlet boundary could be obtained by rewriting Eqs. (16c) and (16d):

$$\lambda = -\kappa \partial_n \phi = -\kappa \partial_n \phi^h + \kappa \beta (\phi^h - \phi_D) \quad \text{on } \partial\Omega_D^-, \quad (17a)$$

$$\lambda = -\kappa \partial_n \phi = -\kappa \partial_n \phi^h + (\kappa \beta + \mathbf{a} \cdot \mathbf{n})(\phi^h - \phi_D) \quad \text{on } \partial\Omega_D^+. \quad (17b)$$

The same expressions were also proposed in [27] by Bazilevs et al., although motivated based on discrete conservation laws.

**Remark 3.** In [48], Harari and Albocher perform a spectral analysis of Nitsche's formulation. Their work shows that its spectrum consists of (i) traditional modes that are independent of the Nitsche parameter and which vanish on Dirichlet boundaries, and (ii) modes that depend on the Nitsche parameter and are locally supported in a layer along the Dirichlet boundaries. In Eq. (16), we observe a similar split in the (functional) constraints imposed by the Nitsche projector.

### 3. A residual-based multiscale model with non-vanishing fine-scale boundary values

For our finite element scheme to yield solutions close to  $\phi^h = \mathcal{P}_N \phi$ , the weighted integral of  $\phi'$  in formulation (14) needs to be taken into account. The closure relation for these remaining fine scales follows from the fine-scale problem of Eq. (7b). In general, the fine-scale problem cannot be solved exactly and modeling approximations are necessary. Following [8,47], we will only make approximations that are consistent with the one-dimensional case. We take the following steps:

- Section 3.1: we write the fine-scale solution in terms of the coarse-scale residual. This relation also involves the fine-scale Lagrange multiplier.
- Section 3.2: we introduce the fine-scale Green's function. We eliminate the fine-scale Lagrange multiplier by making use of the fine-scale Green's function corresponding to the  $H_0^1$ -projector. The resulting explicit relation between  $\phi'$  and  $\phi^h$  is still exact, and it now includes a fine-scale boundary term.
- Section 3.3: we consider the one-dimensional case to simplify the expressions. The fine-scale Green's function localizes and the integrals can be rewritten to yield parameters  $\tau$  and  $\gamma$ . The latter is the new modeling parameter that is a result from the non-vanishing fine scales on the Dirichlet boundary.
- Section 3.4: we generalize the exact expression from the one-dimensional case to obtain a practical method for multi-dimensional simulations. Based on the one-dimensional case, we derive explicit expressions for the effective model parameters for lower and higher order approximation spaces.

We provide a summary of the finite element formulation and all model parameters in Section 3.5.

#### 3.1. Inversion of the fine-scale problem

We start by developing an explicit relation that determines  $\phi'$  from a given coarse-scale function  $\phi^h$  and a given Lagrange multiplier  $\lambda$ . For any  $\lambda$ , we can define a fine-scale space  $\mathcal{W}'$  as:

$$\mathcal{W}' = \{w \in \mathcal{W} : \mathcal{P}_N(w, \lambda) = (0, 0)\}. \quad (18)$$



We then consider the following form of Eq. (7b), the fine-scale problem:

$$\begin{aligned} &\text{Find } \phi' \in \mathcal{W}' \text{ s.t. } \forall w' \in \mathcal{W}' : \\ &\langle \mathcal{L}^* w', \phi' \rangle = \langle \mathcal{R}_{\phi^h}, w' \rangle - \langle \lambda, w' \rangle_{\partial \Omega_D}, \end{aligned} \quad (19)$$

where we introduce the adjoint operator  $\mathcal{L}^*$  and the residual  $\mathcal{R}_{\phi^h}$ , whose definitions follow from Eq. (7b) as:

$$\langle \mathcal{L}^* w', \cdot \rangle = -(\mathbf{a} \cdot \nabla w', \cdot)_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w', \cdot \rangle_{\partial \Omega^+} + (\nabla w', \kappa \nabla \cdot)_{\Omega}, \quad (20a)$$

$$\begin{aligned} \langle \mathcal{R}_{\phi^h}, w' \rangle &= (w', f)_{\Omega} + \langle w', g_N \rangle_{\partial \Omega_N} - \langle \mathbf{a} \cdot \mathbf{n} w', \phi_D \rangle_{\partial \Omega_D^-} \\ &\quad + (\mathbf{a} \cdot \nabla w', \phi^h)_{\Omega} - \langle \mathbf{a} \cdot \mathbf{n} w', \phi^h \rangle_{\partial \Omega^+} - (\nabla w', \kappa \nabla \phi^h)_{\Omega}. \end{aligned} \quad (20b)$$

We hypothesize that one can find a function  $w'(y) \in \mathcal{W}'$  such that  $\mathcal{L}^* w'(y)$  acts on any function in  $v \in \mathcal{W}$  in the following way:

$$\langle \mathcal{L}^* w', v \rangle = \langle \delta_x, v \rangle + \sum_{i=1}^{\dim(\mathcal{W}^h)} c_i \langle \mu_i, v \rangle, \quad (21)$$

where this particular function is called the fine-scale Green's function  $g'(x, y)$ . The  $\delta_x$  denotes the Dirac delta distribution at  $x$ , and the  $\mu_i$  functionals refer to the functional constraints from Eq. (15). When we substitute this fine-scale Green's function as  $w'(y)$  in (19), we obtain:

$$\begin{aligned} \langle \mathcal{L}^* g'(x, y), \phi' \rangle &= \langle \delta_x, \phi' \rangle + \sum_{i=1}^{\dim(\mathcal{W}^h)} \langle c_i \mu_i, \phi' \rangle \\ &= \phi'(x) = \langle \mathcal{R}_{\phi^h}, g'(x, y) \rangle - \langle \lambda, g'(x, y) \rangle_{\partial \Omega_D}, \end{aligned} \quad (22)$$

where the summation vanishes due to Eq. (15).

### 3.2. Adoption of the $H_0^1$ fine-scale Green's function

The fine-scale Green's function in Eqs. (21) and (22) corresponds to the Nitsche projector. However, most literature on the variational multiscale method focuses on a scale decomposition by means of the  $H_0^1$  projector. To maintain the connection with existing fine-scale models, we reintroduce the  $H_0^1$  fine-scale Green's function as follows:

$$g'_N(x, y) = g'_{H_0^1}(x, y) + \tilde{g}'(x, y). \quad (23)$$

The newly added subscripts indicate the projector with which the fine-scale Green's function is associated.

The similarity between the  $H_0^1$  projector and the Nitsche projector can be expressed in terms of their imposed functional constraints, i.e., those from Eq. (15). The set of functions  $\mu_i$  corresponding to the projector  $\mathcal{P}_{H_0^1}$  is a subset of those of  $\mathcal{P}_N$ . If we order the set  $\mu_i$  such that coinciding occurrences come first, then we may write for the  $H_0^1$  fine-scale Green's function:

$$\langle \mathcal{L}^* g'_{H_0^1}(x, y), v \rangle = \langle \delta_x, v \rangle + \sum_{i=1}^{\dim(\mathcal{W}^h \cap H_0^1)} d_i(x) \langle \mu_i, v \rangle - \langle \mathbb{H} g'_{H_0^1}, v \rangle_{\partial \Omega_D} \quad \forall v \in \mathcal{W}, \quad (24)$$

where the last term is due to the non-zero boundary values of  $v$  on  $\partial \Omega_D$ . For our advection–diffusion model problem the operator  $\mathbb{H}$  is:

$$\mathbb{H} = -\kappa \mathbf{n} \cdot \nabla_y. \quad (25)$$

After substitution of Eq. (24) in Eq. (21), we obtain a result for  $\tilde{g}'(x, y)$  that holds for all  $v \in \mathcal{W}$ :

$$\langle \mathcal{L}^* \tilde{g}'(x, y), v \rangle = \sum_{i=1}^{\dim(\mathcal{W}^h \cap H_0^1)} (c_i(x) - d_i(x)) \langle \mu_i, v \rangle + \sum_{i=\dim(\mathcal{W}^h \cap H_0^1)+1}^{\dim(\mathcal{W}^h)} c_i(x) \langle \mu_i, v \rangle + \langle \mathbb{H} g'_{H_0^1}, v \rangle_{\partial \Omega_D}. \quad (26)$$

Substitution of Eq. (23) into Eq. (22) then gives:

$$\begin{aligned}
 \phi'(x) &= \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} + \tilde{g}' \rangle - \langle \lambda, g'_{H_0^1} + \tilde{g}' \rangle_{\partial\Omega_D}, \\
 &= \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \mathcal{R}_{\phi^h}, \tilde{g}' \rangle - \langle \lambda, \tilde{g}' \rangle_{\partial\Omega_D}, \\
 &= \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \mathcal{L}^* \tilde{g}', \phi' \rangle, \\
 &= \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \mathbb{H} g'_{H_0^1}, \phi' \rangle_{\partial\Omega_D},
 \end{aligned} \tag{27}$$

where we make use of  $g'_{H_0^1} = 0$  on  $\partial\Omega_D$  in the second equality, of Eq. (19) in the third equality, and of Eq. (26) followed by Eq. (15) in the fourth equality.

**Remark 4.** The inversion posed by Eq. (27) is no longer unique; it is satisfied for solutions  $\phi^h$  from  $\mathcal{P}_N(\phi, -\kappa \partial_n \phi)$ , but also for solutions  $\phi^h = \mathcal{P}_{H_0^1} \phi$ . However, the partial fine-scale closure discussed in Section 2.2 is not satisfied by  $\phi^h = \mathcal{P}_{H_0^1} \phi$ . The formulation obtained after substitution of Eq. (27) into Nitsche's coarse-scale formulation of Eq. (14) will hence be uniquely satisfied by the coarse-scale solution defined by the Nitsche projector.

The final result of Eq. (27) may be rewritten in a somewhat more familiar form by substituting the definition of the weak residual from Eq. (20b) and by performing reverse integration by parts. We obtain the following expression, to be interpreted in the sense of distributions:

$$\phi'(x) = \int_{\Omega} \mathcal{R}_{\phi^h} g'_{H_0^1}(x, y) dy + \int_{\partial\Omega_D} \mathbb{H} g'_{H_0^1}(x, y) (\phi_D - \phi^h) dy \tag{28}$$

The original residual-based model assumes that the fine-scale solution vanishes on element boundaries [7], and hence that the second term in Eq. (28) cancels. We aim not to make this assumption, and this will be the key term that we retain to obtain a more suitable fine-scale model.

The only term in the coarse-scale formulation of Eq. (14) in which  $\phi'$  appears is the advective term. We thus finally obtain the fine-scale contribution to the coarse-scale equation as:

$$\begin{aligned}
 - \int_{\Omega} \mathbf{a} \cdot \nabla w^h \phi' dx &= - \int_{\Omega} \int_{\Omega} \mathbf{a} \cdot \nabla w^h(x) g'_{H_0^1}(x, y) \mathcal{R}_{\phi^h}(y) dy dx \\
 &\quad - \int_{\Omega} \int_{\partial\Omega_D} \mathbf{a} \cdot \nabla w^h(x) \mathbb{H} g'_{H_0^1}(x, y) (\phi_D(y) - \phi^h(y)) dy dx.
 \end{aligned} \tag{29}$$

The double integration and the limited availability of Green's functions make the closure relation of Eq. (29), albeit exact, unsuitable for computational use. Simplifications via approximations need to be introduced. The strategy that we will employ repeatedly in this article is to reformulate such that exactness is maintained in the case of constant physical parameters on a one-dimensional domain, while ease of implementation is established in the general case.

### 3.3. The classical one-dimensional case

In the one-dimensional case, the nodal exactness induced by the  $\mathcal{P}_{H_0^1}$  projector results in an element-local fine-scale Green's function [47]. The double integrals in Eq. (29) can thus be split in contributions of individual elements. The newly added term only affects elements that lie adjacent to the Dirichlet boundary, where the precise value of the fine-scale solution is known as  $\phi' = \phi_D - \phi^h$ . Let us consider the contribution of one element that shares a node with the Dirichlet boundary:

$$\begin{aligned}
 - \int_K \mathbf{a} \cdot \nabla w^h \phi' dx &= - \int_K \int_K \mathbf{a} \cdot \nabla w^h(x) g'_{H_0^1}(x, y) \mathcal{R}_{\phi^h}(y) dy dx \\
 &\quad - \int_K \int_{\partial K \cap \partial\Omega_D} \mathbf{a} \cdot \nabla w^h(x) \mathbb{H} g'_{H_0^1}(x, y) (\phi_D(y) - \phi^h(y)) dy dx.
 \end{aligned} \tag{30}$$

To simplify, we use the polynomial representation of the test function  $w^h|_K$  and the residual  $\mathcal{R}_{\phi^h}|_K$ . When  $\mathcal{W}^h$  is constructed with  $P$ -order nodal elements and the source function  $f$  is at most polynomial order  $P - 1$ , then:

$$-\mathbf{a} \cdot \nabla w^h(x) = -a \sum_{i=1}^P \hat{w}_i i x^{i-1}; \quad \mathcal{R}_{\phi^h}(y) = \sum_{j=1}^P \hat{R}_j y^{j-1} \quad \text{for } x \in K, y \in K. \tag{31}$$



We also have the following properties of the fine-scale Green's function  $g'_{H_0^1}(x, y)$ :

$$\int_K \int_K x^q g'_{H_0^1}(x, y) y^r dy dx = 0 \quad \text{when } q < P-1 \text{ or } r < P-1, \quad (32a)$$

$$\int_K \int_{\partial K} x^q \mathbb{H} g'_{H_0^1}(x, y) dy dx = 0 \quad \text{when } q < P. \quad (32b)$$

Property (32a) is shown in [47], and we prove (32b) in A.1. Substituting Eq. (31) into Eq. (30) while using Eq. (32) yields:

$$\begin{aligned} - \int_K \mathbf{a} \cdot \nabla w^h \phi' dx &= - \int_K \int_K a (\hat{w}_P P x^{P-1}) g'_{H_0^1}(x, y) (\hat{R}_P y^{P-1}) dy dx \\ &\quad - \int_K \int_{\partial K \cap \partial \Omega_D} a (\hat{w}_P P x^{P-1}) \mathbb{H} g'_{H_0^1}(x, y) (\phi_D(y) - \phi^h(y)) dy dx. \end{aligned} \quad (33)$$

We assume  $a$  to be constant in  $K$ , and we extract  $a$  and all other constants from the double integration. We then integrate the two right hand side terms over  $K$  and  $F := \partial K \cap \partial \Omega_D$  respectively, and we divide them by  $|K|$  and  $|F|$ :

$$\begin{aligned} - \int_K \mathbf{a} \cdot \nabla w^h \phi' dx &= - \int_K a \hat{w}_P P h^{P-1} \left[ \frac{1}{|K|} \int_K \int_K \frac{x^{P-1}}{h^{P-1}} g'_{H_0^1}(x, y) \frac{y^{P-1}}{h^{P-1}} dy dx \right] \hat{R}_P h^{P-1} d\hat{x} \\ &\quad - \int_F a \hat{w}_P P h^{P-1} \left[ \frac{1}{|F|} \int_K \int_F \frac{x^{P-1}}{h^{P-1}} \mathbb{H} g'_{H_0^1}(x, y) dy dx \right] (\phi_D(\hat{x}) - \phi^h(\hat{x})) d\hat{x}, \end{aligned} \quad (34)$$

where we can identify the following model parameters:

$$\tau = \frac{1}{|K|} \int_K \int_K \frac{x^{P-1}}{h^{P-1}} g'_{H_0^1}(x, y) \frac{y^{P-1}}{h^{P-1}} dy dx, \quad (35a)$$

$$\gamma = \frac{1}{|F|} \int_K \int_F \frac{x^{P-1}}{h^{P-1}} \mathbb{H} g'_{H_0^1}(x, y) dy dx. \quad (35b)$$

The multiplication and division by  $h^{P-1}$  in Eq. (34) ( $h$  being the element size) ensures that the parameters in Eq. (35) remain dimensionally consistent with varying polynomial order.

### 3.4. Fine-scale closure generalization

Up until now, all derivations have been exact. To make use of the integral expressions in Eq. (34) on multi-dimensional domains, we approximate them by the following inner products:

$$- \int_K a \hat{w}_P P h^{P-1} \tau \hat{R}_P h^{P-1} d\hat{x} \approx -(\mathbf{a} \cdot \nabla w^h, \tau_{\text{eff}} \mathcal{R}_{\phi^h})_K, \quad (36a)$$

$$- \int_F a \hat{w}_P P h^{P-1} \gamma (\phi_D - \phi^h) d\hat{x} \approx -(\mathbf{a} \cdot \nabla w^h, \gamma_{\text{eff}} (\phi_D - \phi^h))_{F^+}. \quad (36b)$$

As Eq. (36b) indicates, we only make use of the newly proposed term at the outflow Dirichlet boundary  $F^+ := \partial K \cap \partial \Omega_D^+$ . This is where the boundary layers occur, and where the weak enforcement of the Dirichlet conditions results in impactful fine-scale boundary values.

All the approximations involved in the final finite element formulation may be traced back to these two equations. Essentially, they shift the modeling effort onto the effective stabilization parameters  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$ . We design  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$  such that these are approximations of  $\tau$  and  $\gamma$  that take into account the change of (bi)linear forms, while being suitable for multi-dimensional computations for arbitrary order polynomial basis functions.

#### 3.4.1. Estimation of operator impact

In the one-dimensional case, the bilinear forms of the left-hand and right-hand sides of Eq. (36) may be written as:

$$B_{\text{vol}}(w^h, \phi^h) = (a \left( \frac{h^{P-1}}{(P-1)!} \right) \frac{\partial^P}{\partial x^P} w^h, \tau \left( \frac{h^{P-1}}{(P-1)!} \right) \frac{\partial^{P-1}}{\partial x^{P-1}} \mathcal{L} \phi^h)_K, \quad (37a)$$

$$\tilde{B}_{\text{vol}}(w^h, \phi^h) = \left( a \frac{\partial}{\partial x} w^h, \tau_{\text{eff}} \mathcal{L} \phi^h \right)_K, \quad (37b)$$

$$B_{\text{bdy}}(w^h, \phi^h) = \left\langle a \left( \frac{h^{P-1}}{(P-1)!} \right) \frac{\partial^P}{\partial x^P} w^h, \gamma \phi^h \right\rangle_{\partial K \cap \partial \Omega_D^+}, \quad (37c)$$

$$\tilde{B}_{\text{bdy}}(w^h, \phi^h) = \left\langle a \frac{\partial}{\partial x} w^h, \gamma_{\text{eff}} \phi^h \right\rangle_{\partial K \cap \partial \Omega_D^+}. \quad (37d)$$

The impact of these different bilinear forms may be quantified by considering their norms:

$$\|B\| := \sup_{\frac{\partial}{\partial x} w^h \neq 0, \frac{\partial}{\partial x} \phi^h \neq 0} \frac{|B(w^h, \phi^h)|}{\left\| \frac{\partial}{\partial x} w^h \right\|_{L^2(K)} \left\| \frac{\partial}{\partial x} \phi^h \right\|_{L^2(K)}}, \quad (38)$$

where we have chosen to define the norm of the bilinear forms with respect to the  $H^1$ -seminorm of its arguments, as this seminorm is one of the terms in the optimality condition induced by the Nitsche projector according to Eq. (9).

We choose  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$  such that the impact of these bilinear forms equal:  $\|B_{\text{vol}}\| = \|\tilde{B}_{\text{vol}}\|$  and  $\|B_{\text{bdy}}\| = \|\tilde{B}_{\text{bdy}}\|$ . If we assume constant parameters in  $K$ , and the advective dominant case such that  $\mathcal{L} \phi^h$  may be approximated by  $a \frac{\partial}{\partial x} \phi^h$ , then we obtain:

$$\tau_{\text{eff}} \approx \tau \left( \frac{h^{P-1}}{(P-1)!} \right)^2 \frac{\left\| \left( \frac{\partial^P}{\partial x^P} \cdot, \frac{\partial^P}{\partial x^P} \cdot \right)_K \right\|}{\left\| \left( \frac{\partial}{\partial x} \cdot, \frac{\partial}{\partial x} \cdot \right)_K \right\|} = \tau \left( \frac{h^{P-1}}{(P-1)!} \right)^2 \sup_{\frac{\partial}{\partial x} w^h \neq 0} \frac{\left\| \frac{\partial^P}{\partial x^P} w^h \right\|_{L^2(K)}^2}{\left\| \frac{\partial}{\partial x} w^h \right\|_{L^2(K)}^2}, \quad (39a)$$

$$\gamma_{\text{eff}} \approx \gamma \left( \frac{h^{P-1}}{(P-1)!} \right) \frac{\left\| \left( \frac{\partial^P}{\partial x^P} \cdot, \cdot \right)_F \right\|}{\left\| \left( \frac{\partial}{\partial x} \cdot, \cdot \right)_F \right\|} = \gamma \left( \frac{h^{P-1}}{(P-1)!} \right) \frac{\sup_{\frac{\partial}{\partial x} w^h \neq 0} \left( \left| \frac{\partial^P}{\partial x^P} w^h \right|_F / \left\| \frac{\partial}{\partial x} w^h \right\|_{L^2(K)} \right)}{\sup_{\frac{\partial}{\partial x} w^h \neq 0} \left( \left| \frac{\partial}{\partial x} w^h \right|_F / \left\| \frac{\partial}{\partial x} w^h \right\|_{L^2(K)} \right)}. \quad (39b)$$

The inverse inequalities in these expressions are computable by hand [49]. For linear, quadratic and cubic coarse-scale basis functions, the relations between the parameters  $\tau_{\text{eff}}$  and  $\tau$ , and  $\gamma_{\text{eff}}$  and  $\gamma$  become:

$$\tau_{\text{eff}}, \gamma_{\text{eff}} \approx \begin{cases} \tau, & \gamma & \text{for } P = 1, \\ 12\tau, & \sqrt{3}\gamma & \text{for } P = 2, \\ 180\tau, & 2\sqrt{5}\gamma & \text{for } P = 3. \end{cases} \quad (40)$$

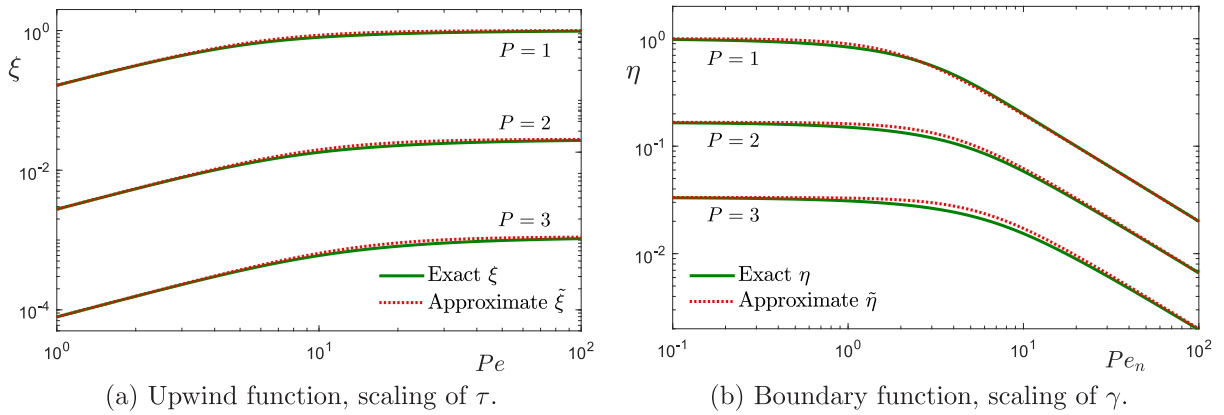
### 3.4.2. $\tau$ -parameter approximation for $P \in \{1, 2, 3\}$

In the literature, we find that the  $\tau$ -parameter that is used with higher-order basis functions is often the same as that for  $P = 1$  (i.e., obtained from the element local Green's function), sometimes with a  $P$ -dependent scaling (see, e.g., [7,27,50]). In this section, we propose an approximation of  $\tau$  for linear, quadratic and cubic elements based on the actual fine-scale Green's functions. These are devised such that they limit to the exact expressions in the advection ( $\tau_a$ ) or diffusion ( $\tau_d$ ) dominated cases. Using the definition of  $\tau$  from Eq. (35a), together with the fine-scale Green's functions from the Appendix, the following exact expressions for  $\tau$  may be computed for  $P = 1, 2$  and 3 respectively [1,10,47]:

$$\tau_1 = \frac{h}{2|a|} \left( \frac{2 + Pe - (2 - Pe) \exp(Pe)}{-Pe + Pe \exp(Pe)} \right) =: \frac{h}{2|a|} \xi_1(Pe), \quad (41a)$$

$$\tau_2 = \frac{h}{72|a|} \left( \frac{12 + 6Pe + Pe^2 - (12 - 6Pe + Pe^2) \exp(Pe)}{-2Pe - Pe^2 + (2Pe - Pe^2) \exp(Pe)} \right) =: \frac{h}{2|a|} \xi_2(Pe), \quad (41b)$$

$$\begin{aligned} \tau_3 &= \frac{h}{1800|a|} \left( \frac{120 + 60Pe + 12Pe^2 + Pe^3 - (120 - 60Pe + 12Pe^2 - Pe^3) \exp(Pe)}{-12Pe - 6Pe^2 - Pe^3 + (12Pe - 6Pe^2 + Pe^3) \exp(Pe)} \right) \\ &=: \frac{h}{2|a|} \xi_3(Pe), \end{aligned} \quad (41c)$$



**Fig. 1.** Exact  $\xi$  and  $\eta$  functions and their approximations for the one-dimensional case. Showing the correct scaling behavior.

where  $Pe = \frac{|a|h}{\kappa}$  is the element Péclet number and  $\xi$  is the *upwind function*. From these equations we obtain the following advective and diffusive limits:

$$\tau_{1,a} := \lim_{Pe \rightarrow \infty} \tau_1 = \frac{h}{2|a|}, \quad \tau_{1,d} := \lim_{Pe \rightarrow 0^+} \tau_1 = \frac{h^2}{12\kappa}, \quad (42a)$$

$$\tau_{2,a} = \frac{h}{72|a|}, \quad \tau_{2,d} = \frac{h^2}{720\kappa}, \quad (42b)$$

$$\tau_{3,a} = \frac{h}{1800|a|}, \quad \tau_{3,d} = \frac{h^2}{25200\kappa}. \quad (42c)$$

The following approximation strategy for  $\tau$  satisfies the above limits and is used frequently in stabilized methods [7,13,51–53]:

$$\tau \approx \frac{1}{\sqrt{\tau_a^{-2} + \tau_d^{-2}}}. \quad (43)$$

To determine the effectiveness of the scaling for the various polynomial orders, we substitute Eq. (42) into (43). In all cases, we can rewrite the expression to obtain the effective approximate upwind function. For example, for linear elements:

$$\tau_1 \approx \frac{1}{\sqrt{\frac{4|a|^2}{h^2} + \frac{144\kappa^2}{h^4}}} = \frac{h}{2|a|} \frac{1}{\sqrt{1 + 36Pe^{-2}}} =: \frac{h}{2|a|} \tilde{\xi}_1(Pe). \quad (44)$$

Fig. 1(a) illustrates how the approximate upwind functions  $\tilde{\xi}(Pe)$  relates to the exact upwind functions of Eq. (41). The figure shows that the approximation of  $\tau$  according to Eq. (43) has the correct asymptotic limits, and converges to these limits at the correct rates. We observe that this holds for each polynomial order.

### 3.4.3. $\gamma$ -parameter approximation for $P \in \{1, 2, 3\}$

In a similar sense, we wish to construct an approximate  $\gamma$ , based on generic (spatial dimension independent) parameters, that share the asymptotic scaling behavior of the exact one-dimensional expression. The exact expressions of  $\gamma$  may be computed from Eq. (35b) as:

$$\gamma_1 = \frac{h}{2} \left( \frac{2 + 2Pe_n - 2\exp(Pe_n)}{Pe_n - Pe_n \exp(Pe_n)} \right) =: \frac{h}{2} \eta_1(Pe_n), \quad (45a)$$

$$\gamma_2 = \frac{h}{12} \left( \frac{12 + 8Pe_n + 2Pe_n^2 - (12 - 4Pe_n)\exp(Pe_n)}{2Pe_n + Pe_n^2 - (2Pe_n - Pe_n^2)\exp(Pe_n)} \right) =: \frac{h}{2} \eta_2(Pe_n), \quad (45b)$$

$$\gamma_3 = \frac{h}{60} \left( \frac{120 + 72Pe_n + 18Pe_n^2 + 2Pe_n^3 - (120 - 48Pe_n + 6Pe_n^2)\exp(Pe_n)}{12Pe_n + 6Pe_n^2 + Pe_n^3 - (12Pe_n - 6Pe_n^2 + Pe_n^3)\exp(Pe_n)} \right) =: \frac{h}{2} \eta_3(Pe_n), \quad (45c)$$

for  $P = 1, 2$  and  $3$  respectively.  $Pe_n = \frac{\mathbf{a} \cdot \mathbf{n} h}{\kappa}$  is a boundary type element Péclet number. Since we only make use of  $\gamma$  on the outflow boundary ( $\mathbf{a} \cdot \mathbf{n} \geq 0$ ), we exclusively consider  $Pe_n \geq 0$  in the following. The advective and diffusive limits of these expressions are:

$$\gamma_{1,a} := \lim_{Pe_n \rightarrow \infty} \gamma_1 = 0, \quad \gamma_{1,d} := \lim_{Pe_n \rightarrow 0^+} \gamma_1 = \frac{h}{2}, \quad (46a)$$

$$\gamma_{2,a} = 0, \quad \gamma_{2,d} = \frac{h}{12}, \quad (46b)$$

$$\gamma_{3,a} = 0, \quad \gamma_{3,d} = \frac{h}{60}. \quad (46c)$$

Just like for the parameter  $\tau$  we require an approximation strategy that satisfies these limits. The strategy of Eq. (43) is not viable as the limits in Eq. (46) do not depend on  $\mathbf{a}$  nor on  $\kappa$ . Instead, we deduce the following scaling of  $\gamma$  based on its definition from Eq. (35b):

$$\gamma \propto |\mathbb{H}| \hat{\tau}, \quad (47)$$

where  $|\mathbb{H}|$  represents the scaling induced by  $\mathbb{H}$  as defined in Eq. (25), and  $\hat{\tau}$  is a modified  $\tau$ -like quantity that takes into account the difference in domain of integration between  $\gamma$  and  $\tau$  in Eq. (35).  $\mathbb{H}$  scales linearly with  $\kappa$  and inversely with length, such that we may choose  $|\mathbb{H}| \propto \sqrt{\kappa \tau_d^{-1}}$ . By adopting a similar approximation for  $\hat{\tau}$  as Eq. (43), we obtain:

$$\gamma \approx c_s \sqrt{\kappa \tau_d^{-1}} \sqrt{\frac{1}{c_1 \tau_a^{-2} + c_2 \tau_d^{-2}}} = c_s \sqrt{\frac{\kappa}{c_1 \tau_d \tau_a^{-2} + c_2 \tau_d^{-1}}}. \quad (48)$$

$c_s$  takes into account the shape effect of the element as the measure of the relevant boundary versus the measure of the element interior:

$$c_s := \frac{h|F|}{|K|}, \quad (49)$$

where the multiplication with  $h$  ensures a mesh size independent scaling, and  $h$  should be the representative element size that is used in the approximation of  $\tau$  through  $\tau_a$  and  $\tau_d$ . In this article we use the longest element edge. In the one-dimensional case  $|F| = 1$  and  $|K| = h$ , such that  $c_s = 1$ .

The coefficients  $c_1$  and  $c_2$  are introduced to capture the difference in scaling between  $\tau$  and  $\hat{\tau}$ . They can be determined from Eq. (48) by ensuring the correct limiting behavior of  $\gamma$  in the one-dimensional case. For example, for linear elements, substituting the limits of Eq. (42a) results in:

$$\gamma_1 \approx \frac{h}{2} \sqrt{\frac{1}{\frac{1}{12} c_1 Pe_n^2 + 3 c_2}} =: \frac{h}{2} \tilde{\eta}_1(Pe_n). \quad (50)$$

By ensuring that Eq. (50) has the same asymptotic limits as Eq. (45) and also has the same convergence rate towards zero, we obtain  $c_1 = 3$  and  $c_2 = 1/3$ . The same strategy results in  $c_1 = 1.25$  and  $c_2 = 0.2$  for  $P = 2$  and  $c_1 = 7/9$  and  $c_2 = 1/7$  for  $P = 3$ . Fig. 1(b) shows the approximate and exact boundary functions and confirms that the approximation of  $\gamma$  displays the correct asymptotic scaling behavior for all polynomial orders.

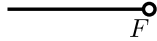
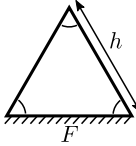
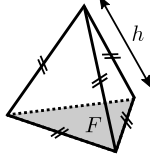
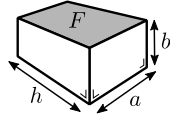
### 3.5. Complete finite element formulation

With all the modeling terms included, the finite element formulation becomes:

Find  $\phi^h \in \mathcal{W}^h$  s.t.  $\forall w^h \in \mathcal{W}^h$  :

$$\begin{aligned} B(w^h, \phi^h) &= B_A(w^h, \phi^h) + B_D(w^h, \phi^h) + B_{\text{VMS}, \tilde{\Omega}}(w^h, \phi^h) + B_{\text{VMS}, \partial \Omega_D^+}(w^h, \phi^h) \\ &= (w^h, f)_{\Omega} + \langle g_N, w^h \rangle_{\partial \Omega_N} - \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi_D \rangle_{\partial \Omega_D^-} - \langle \kappa \partial_n w^h, \phi_D \rangle_{\partial \Omega_D} + \langle \kappa \beta w^h, \phi_D \rangle_{\partial \Omega_D} \\ &\quad + (\tau_{\text{eff}} \mathbf{a} \cdot \nabla w^h, f)_{\Omega} + \langle \gamma_{\text{eff}} \mathbf{a} \cdot \nabla w^h, \phi_D \rangle_{\partial \Omega_D^+}, \end{aligned} \quad (51)$$

**Table 1**Overview of  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$  expressions for different elements and polynomial degrees.

	$P = 1$	$P = 2$	$P = 3$	
$\tau_{\text{eff}}:$	$\sqrt{\frac{1}{\tau_{1,a}^{-2} + \tau_{1,d}^{-2}}}$	$\sqrt{\frac{1}{9\tau_{1,a}^{-2} + 25\tau_{1,d}^{-2}}}$	$\sqrt{\frac{1}{25\tau_{1,a}^{-2} + \frac{1225}{9}\tau_{1,d}^{-2}}}$	
$\gamma_{\text{eff}}:$	$c_s \sqrt{\frac{\kappa}{3\tau_{1,d}\tau_{1,a}^{-2} + \frac{1}{3}\tau_{1,d}^{-1}}}$	$c_s \sqrt{\frac{\kappa}{9\tau_{1,d}\tau_{1,a}^{-2} + 4\tau_{1,d}^{-1}}}$	$c_s \sqrt{\frac{\kappa}{15\tau_{1,d}\tau_{1,a}^{-2} + 15\tau_{1,d}^{-1}}}$	
Element:				
$c_s$ (Eq. (49)):	1	$\frac{4}{\sqrt{3}}$	$2\sqrt{\frac{2}{3}}$	$h/b$

where the advection and diffusion parts of the bilinear form are:

$$B_A(w^h, \phi^h) = -(\mathbf{a} \cdot \nabla w^h, \phi^h)_\Omega + \langle \mathbf{a} \cdot \mathbf{n} w^h, \phi^h \rangle_{\partial\Omega^+}, \quad (52a)$$

$$B_D(w^h, \phi^h) = (\kappa \nabla w^h, \nabla \phi^h)_\Omega - \langle \kappa \partial_n \phi^h, w^h \rangle_{\partial\Omega_D} - \langle \kappa \partial_n w^h, \phi^h \rangle_{\partial\Omega_D} + \langle \kappa \beta w^h, \phi^h \rangle_{\partial\Omega_D}, \quad (52b)$$

and where the two variational multiscale components are:

$$B_{\text{VMS}, \tilde{\Omega}}(w^h, \phi^h) = (\tau_{\text{eff}} \mathbf{a} \cdot \nabla w^h, \mathbf{a} \cdot \nabla \phi^h - \nabla \cdot \kappa \nabla \phi^h)_{\tilde{\Omega}}, \quad (53a)$$

$$B_{\text{VMS}, \partial\Omega_D^+}(w^h, \phi^h) = \langle \gamma_{\text{eff}} \mathbf{a} \cdot \nabla w^h, \phi^h \rangle_{\partial\Omega_D^+}, \quad (53b)$$

with  $\tilde{\Omega}$  the sum of open element domains.

Expressions for the parameters  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$  are collected in Table 1. These expressions take into account all the considerations discussed in Sections 3.4.1–3.4.3. As the table shows, we have formulated all model parameters such that they depend exclusively on  $\tau_{1,a}$  and  $\tau_{1,d}$ , i.e., those relating to linear elements. The approach for computing these limiting values remains flexible. For example, one could incorporate the Jacobian of the element mapping [7,28,40], use local length-scales and Péclet numbers (taking into account element distortions and aspect ratios) [53], or use the analytical expressions of Eq. (42a) and the element diameter. The requirements on  $\beta$  are discussed in the following section.

#### 4. Analysis of existence and uniqueness

As touched upon in Remark 4 in Section 3.2, the uniqueness of the fine-scale inversion is not a trivial issue. Additionally, the required simplification and modeling steps discussed in Section 3.4 may raise further questions regarding the existence of the approximate coarse-scale solution. In a finite-dimensional functional setting, existence and uniqueness follow directly from the coercivity of the bilinear form, which we analyze in this section.

In the analysis, we assume that  $\mathbf{a}$  and  $\kappa$  are constant in  $\Omega$  and that the grid is (quasi) uniform, such that also  $\beta$ ,  $\tau_a$  and  $\tau_d$  can be chosen as global constants. The derivations can trivially be modified for non-uniform grids or non-constant  $\mathbf{a}$  and  $\kappa$  and element-wise parameters  $\beta$ ,  $\tau_a$  and  $\tau_d$ . We further assume that the parameter  $\tau_{\text{eff}}$  satisfies:

$$\tau_{\text{eff}} \leq C_d \tau_{1,d} \leq \inf_{w^h \in \mathcal{W}^h} \frac{1}{2\kappa} \frac{\|\nabla w^h\|_\Omega^2}{\|\Delta w^h\|_\Omega^2} \propto \frac{h^2}{\kappa} \quad (54)$$

The first inequality is satisfied due to the harmonic mean structure of  $\tau_{\text{eff}}$ , where the  $C_d$ 's follow directly from Table 1 as 1,  $\frac{1}{5}$  and  $\frac{3}{35}$  for linear, quadratic and cubic basis functions respectively. On a one-dimensional mesh, this means that  $C_d \tau_{1,d} = \frac{h^2}{12\kappa}$ ,  $\frac{h^2}{60\kappa}$  and  $\frac{h^2}{140\kappa}$ . In this one-dimensional case, the inverse estimate in Eq. (54) may be explicitly computed, resulting in maxima of  $\infty$ ,  $\frac{h^2}{24\kappa}$  and  $\frac{h^2}{120\kappa}$  for polynomial orders of 1, 2 and 3. The condition of Eq. (54) is thus satisfied.

Finally, we require that  $\beta$  satisfies:

$$\beta \geq 4(T_1 + c_s^2 T_2) \propto \frac{1}{h}, \quad (55)$$

with:

$$T_1 = \sup_{w^h \in \mathcal{W}^h} \frac{\|\partial_n w^h\|_{\partial\Omega_D}^2}{\|\nabla w^h\|_{\Omega}^2} \propto \frac{1}{h}, \quad (56a)$$

$$T_2 = \sup_{w^h \in \mathcal{W}^h} \frac{\|\mathbf{a} \cdot \nabla w^h\|_{\partial\Omega_D^+}^2}{\|\mathbf{a} \cdot \nabla w^h\|_{\Omega}^2} \propto \frac{1}{h}. \quad (56b)$$

We carry out the coercivity proof for the different components of the bilinear forms separately, and then look at the formulation as a whole. We start with determining a relation between the model parameters  $\gamma_{\text{eff}}$  and  $\tau_{\text{eff}}$ .

**Lemma 4.1.** *For  $P = 1, 2$  or  $3$ , the expressions from Table 1 are such that  $\gamma_{\text{eff}}$  is bounded by  $\tau_{\text{eff}}$  according to:*

$$\gamma_{\text{eff}}^2 \leq 3c_s^2 \kappa \tau_{\text{eff}}. \quad (57)$$

**Proof.** We write the  $\gamma_{\text{eff}}$  and  $\tau_{\text{eff}}$  expressions from Table 1 in the following general form:

$$\tau_{\text{eff}} = \sqrt{\frac{1}{(C_a \tau_{1,a})^{-2} + (C_d \tau_{1,d})^{-2}}} \quad (58)$$

$$\gamma_{\text{eff}} = c_s \sqrt{\frac{\kappa}{\tau_{1,d}}} \sqrt{\frac{1}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}}} \quad (59)$$

Dividing  $\gamma_{\text{eff}}$  by  $\tau_{\text{eff}}$  and squaring gives:

$$\left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = c_s^2 \frac{\kappa}{\tau_{1,d}} \frac{(C_a \tau_{1,a})^{-2} + (C_d \tau_{1,d})^{-2}}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}} \quad (60)$$

After multiplying both sides by  $C_d \tau_{1,d}$  and using  $C_d \tau_{1,d} \geq \tau$ , we obtain:

$$C_d \tau_{1,d} \left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = c_s^2 \kappa \frac{C_d C_a^{-2} \tau_{1,a}^{-2} + C_d^{-1} \tau_{1,d}^{-2}}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}} \geq \tau_{\text{eff}} \left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = \frac{\gamma_{\text{eff}}^2}{\tau_{\text{eff}}} \quad (61)$$

The fraction may be bound from above as:

$$\frac{\gamma_{\text{eff}}^2}{\tau_{\text{eff}}} \leq c_s^2 \kappa \frac{\max(\frac{C_d}{C_a^2}, \frac{1}{C_d})(\tau_{1,a}^{-2} + \tau_{1,d}^{-2})}{\min(\frac{1}{C_1^2}, \frac{1}{C_2^2})(\tau_{1,a}^{-2} + \tau_{1,d}^{-2})} = c_s^2 \kappa \frac{\max(\frac{C_d}{C_a^2}, \frac{1}{C_d})}{\min(\frac{1}{C_1^2}, \frac{1}{C_2^2})} = \begin{cases} 3 c_s^2 \kappa & \text{for } P = 1, \\ \frac{5}{4} c_s^2 \kappa & \text{for } P = 2, \\ \frac{7}{9} c_s^2 \kappa & \text{for } P = 3. \end{cases} \quad (62)$$

Eq. (57) follows from the maximum of the three cases.  $\square$

**Lemma 4.2.** *The bilinear form in Eq. (52a) satisfies the following coercivity result:*

$$B_A(w^h, w^h) \geq \frac{1}{2} \|\sqrt{|\mathbf{a} \cdot \mathbf{n}|} w^h\|_{\partial\Omega}^2 \quad \forall w^h \in \mathcal{W}^h, \quad (63)$$

where the norms are  $L^2$  norms on the indicated domains.

**Proof.** Direct substitution of  $\phi^h = w^h$  into Eq. (52a) results in:

$$B_A(w^h, w^h) = -(\mathbf{a} \cdot \nabla w^h, w^h)_{\Omega} + \langle \mathbf{a} \cdot \mathbf{n} w^h, w^h \rangle_{\partial\Omega^+}. \quad (64)$$

Making use of the property  $\nabla \cdot \mathbf{a} = 0$ , the first term may be rewritten as follows:

$$-\int_{\Omega} \nabla \cdot (\frac{1}{2} \mathbf{a} (w^h)^2) = -\frac{1}{2} \int_{\partial\Omega} \mathbf{a} \cdot \mathbf{n} (w^h)^2 = \frac{1}{2} \int_{\partial\Omega^-} |\mathbf{a} \cdot \mathbf{n}| (w^h)^2 - \frac{1}{2} \int_{\partial\Omega^+} |\mathbf{a} \cdot \mathbf{n}| (w^h)^2. \quad (65)$$

Substitution into Eq. (64) completes the proof.  $\square$

**Lemma 4.3.** Under the condition posed by Eq. (54), the bilinear form in Eq. (52b) satisfies the following lower bound:

$$B_D(w^h, w^h) \geq \frac{1}{2}\kappa \|\nabla w^h\|_\Omega^2 + (\beta - 2T_1)\kappa \|w^h\|_{\partial\Omega_D}^2 \quad \forall w^h \in \mathcal{W}^h, \quad (66)$$

where the norms are  $L^2$  norms on the indicated domains, and  $T_1$  is given in Eq. (56a).

**Proof.** Direct substitution of  $\phi^h = w^h$  into Eq. (52b) results in:

$$B_D(w^h, w^h) = \kappa \|\nabla w^h\|_\Omega^2 - 2 \langle \kappa \partial_n w^h, w^h \rangle_{\partial\Omega_D} + \beta \kappa \|w^h\|_{\partial\Omega_D}^2. \quad (67)$$

By using Young's inequality to bound the nonsymmetric term, we obtain:

$$\begin{aligned} -2 \langle \kappa \partial_n w^h, w^h \rangle_{\partial\Omega_D} &\geq -\epsilon \kappa \|\partial_n w^h\|_{\partial\Omega_D}^2 - \frac{\kappa}{\epsilon} \|w^h\|_{\partial\Omega_D}^2, \\ &\geq -\epsilon T_1 \kappa \|\nabla w^h\|_\Omega^2 - \frac{\kappa}{\epsilon} \|w^h\|_{\partial\Omega_D}^2, \end{aligned} \quad (68)$$

where  $T_1$  is defined in Eq. (56a). Choosing the parameter from Young's inequality as  $\epsilon = 1/(2T_1)$  completes the proof.  $\square$

**Lemma 4.4.** Under the condition posed by Eq. (54), the volumetric variational multiscale term in Eq. (53a) satisfies the following lower bound:

$$B_{VMS, \tilde{\Omega}}(w^h, w^h) \geq \frac{1}{2}\tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 - \frac{1}{4}\kappa \|\nabla w^h\|_\Omega^2 \quad \forall w^h \in \mathcal{W}^h, \quad (69)$$

where the norms are  $L^2$  norms on the indicated domains.

**Proof.** Direct substitution of  $\phi^h = w^h$  results in:

$$B_{VMS, \tilde{\Omega}}(w^h, w^h) = \tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 + (\sqrt{\tau_{\text{eff}}} \mathbf{a} \cdot \nabla w^h, \sqrt{\tau_{\text{eff}}} \kappa \Delta w^h)_{\tilde{\Omega}}. \quad (70)$$

With Young's inequality we bound the second term from below:

$$B_{VMS, \tilde{\Omega}}(w^h, w^h) \geq \tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 - \frac{1}{2}\tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 - \frac{1}{2}\tau_{\text{eff}} \kappa^2 \|\Delta w^h\|_\Omega^2. \quad (71)$$

Using the assumed bound of  $\tau_{\text{eff}}$  from Eq. (54) completes the proof.  $\square$

**Lemma 4.5.** The boundary variational multiscale term in Eq. (53b) satisfies the following lower bound:

$$B_{VMS, \partial\Omega_D^+}(w^h, \phi^h) \geq -\frac{1}{4}\tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 - 3c_s^2 T_2 \kappa \|w^h\|_{\partial\Omega_D^+}^2 \quad \forall w^h \in \mathcal{W}^h, \quad (72)$$

where the norms are  $L^2$  norms on the indicated domains.

**Proof.** After substitution of  $\phi^h = w^h$ , we obtain:

$$B_{VMS, \partial\Omega_D^+}(w^h, w^h) = \langle \mathbf{a} \cdot \nabla w^h, \gamma_{\text{eff}} w^h \rangle_{\partial\Omega_D^+} \quad (73)$$

$$\begin{aligned} &\geq -\frac{1}{2}\varepsilon \|\mathbf{a} \cdot \nabla w^h\|_{\partial\Omega_D^+}^2 - \frac{1}{2}\frac{\gamma_{\text{eff}}^2}{\varepsilon} \|w^h\|_{\partial\Omega_D^+}^2 \\ &\geq -\frac{1}{2}\varepsilon T_2 \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 - \frac{1}{2}\frac{\gamma_{\text{eff}}^2}{\varepsilon} \|w^h\|_{\partial\Omega_D^+}^2. \end{aligned} \quad (74)$$

The first inequality follows from Young's inequality with parameter  $\varepsilon$ , and the second inequality as well as the parameter  $T_2$  originate from the inverse estimate of Eq. (56b). Choosing the parameter  $\varepsilon = \tau_{\text{eff}}/(2T_2)$  and using the result of Lemma 4.1 to relate  $\tau_{\text{eff}}$  and  $\gamma_{\text{eff}}$  completes the poof.  $\square$

**Theorem 4.1.** The combined bilinear form of Eq. (51) satisfies the following coercivity result:

$$B(w^h, w^h) \geq \frac{1}{4}\tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_\Omega^2 + \frac{1}{2}\|\sqrt{|\mathbf{a} \cdot \mathbf{n}|} w^h\|_{\partial\Omega}^2 + \frac{1}{4}\kappa \|\nabla w^h\|_\Omega^2 + \frac{1}{4}\beta \kappa \|w^h\|_{\partial\Omega_D}^2 \quad \forall w^h \in \mathcal{W}^h, \quad (75)$$

where the norms are  $L^2$  norms on the indicated domains.



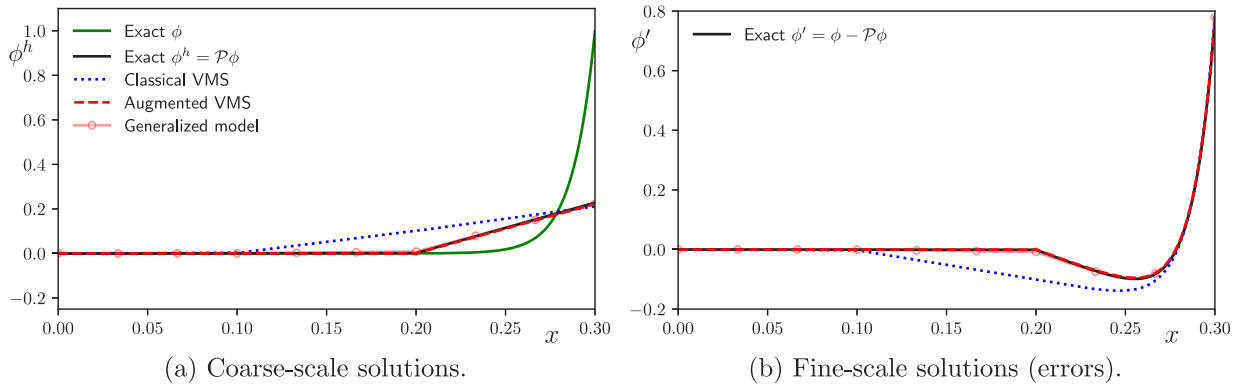


Fig. 2. One-dimensional results using three linear elements.

**Proof.** Direct substitution of  $\phi^h = w^h$  in the bilinear form, and using the results of Lemmas 4.2 to 4.5, results in:

$$\begin{aligned} B(w^h, w^h) &= B_A(w^h, w^h) + B_D(w^h, w^h) + B_{\text{VMS}, \bar{\Omega}}(w^h, w^h) + B_{\text{VMS}, \partial\Omega_D^+}(w^h, w^h) \\ &\geq \frac{1}{2} \|\sqrt{|\mathbf{a} \cdot \mathbf{n}|} w^h\|_{\partial\Omega}^2 + \frac{1}{2} \kappa \|\nabla w^h\|_{\Omega}^2 + (\beta - 2T_1) \kappa \|w^h\|_{\partial\Omega_D}^2 + \frac{1}{2} \tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_{\Omega}^2 \\ &\quad - \frac{1}{4} \kappa \|\nabla w^h\|_{\Omega}^2 - \frac{1}{4} \tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_{\Omega}^2 - 3 c_s^2 T_2 \kappa \|w^h\|_{\partial\Omega_D^+}^2 \end{aligned} \quad (76)$$

$$\begin{aligned} &= \frac{1}{4} \tau_{\text{eff}} \|\mathbf{a} \cdot \nabla w^h\|_{\Omega}^2 + \frac{1}{2} \|\sqrt{|\mathbf{a} \cdot \mathbf{n}|} w^h\|_{\partial\Omega}^2 + \frac{1}{4} \kappa \|\nabla w^h\|_{\Omega}^2 \\ &\quad + (\beta - 2T_1 - 3 c_s^2 T_2) \kappa \|w^h\|_{\partial\Omega_D}^2. \end{aligned} \quad (77)$$

Using the assumption on  $\beta$  from Eq. (55) completes the proof.  $\square$

**Remark 5.** Note that both Nitsche's method (Lemma 4.3) and the VMS method (Lemma 4.4) rely on the term  $\kappa \|\nabla w\|_{\Omega}^2$  for their stability. As a result, the combined use of VMS and weakly enforced boundary conditions requires a larger penalty factor  $\beta$  compared to the typical choice of  $\beta = 2T_1$  for the standard Nitsche's method [36].

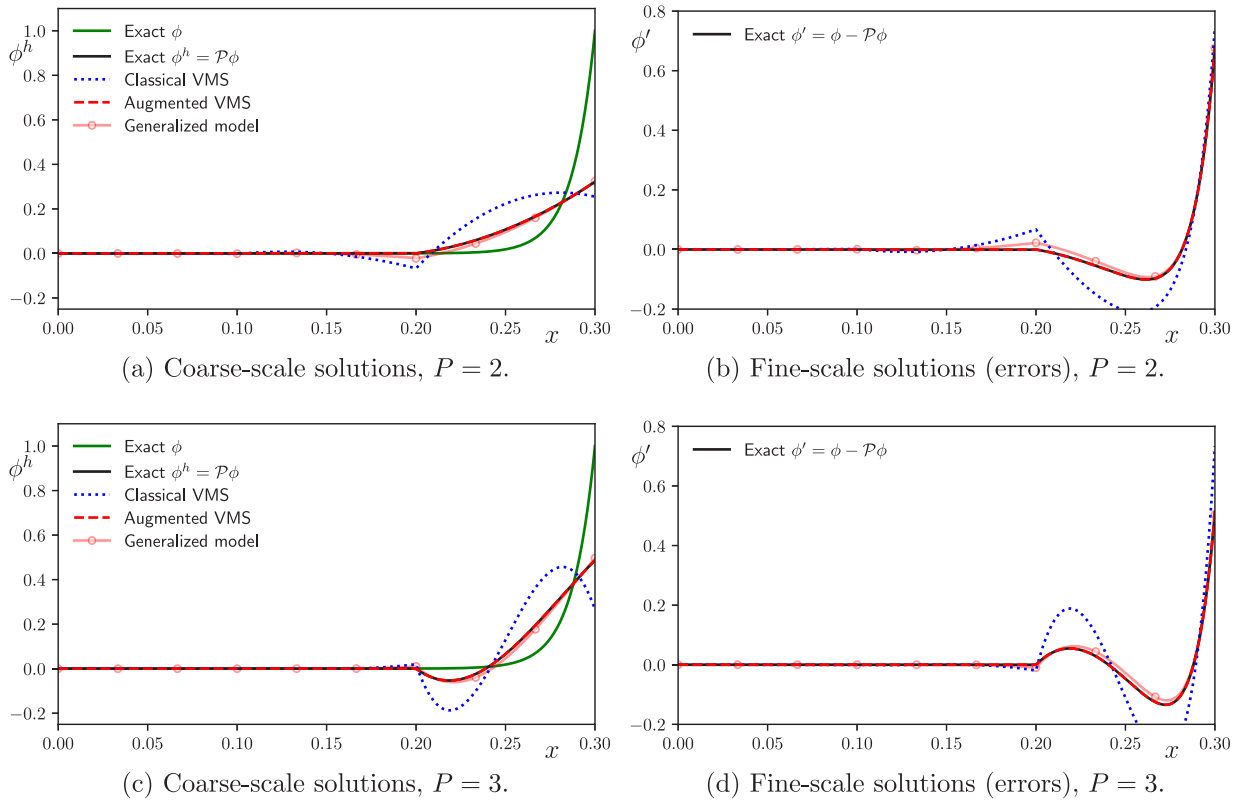
## 5. Numerical verification for a one-dimensional model problem

We present a number of numerical experiments to verify the derivation from Sections 3.1–3.3, and to investigate the accuracy improvement that may be achieved by using the new residual-based fine-scale model of Sections 3.4 and 3.5. For all subsequent simulations, we choose  $\beta$  such that it satisfies the requirements determined in Section 4.

### 5.1. Linear basis functions

Fig. 2 shows the result for a simulation with  $\mathbf{a} = 0.8$ ,  $\kappa = 0.02$  and  $\beta = 2/h$  on the domain  $\Omega = [0, 0.3]$ , discretized with three linear elements. The global Péclet number is 12 and the element local Péclet number is 4. The solid green line shows the exact solution  $\phi$ . With the current discretization, the boundary layer falls completely within a single element. We obtain the exact coarse-scale solution, indicated with the black line, by projecting the exact solution onto the finite element mesh with the Nitsche projector from Section 2.2. The blue dotted line is obtained by only using the classical VMS term, equivalent to  $\gamma = 0$ , whereas the red dashed line incorporates the exact augmented VMS model from Eq. (34) with the exact parameter definitions from Eq. (35). Finally, the line with the circular markers shows the result when the generalized model of Sections 3.4 and 3.5.

It is a celebrated fact that for this model problem the VMS method with strongly enforced boundary conditions results in nodally exact solutions [3]. The results of Fig. 2(a) show that this property is lost when the boundary conditions are enforced weakly, and that the under-resolved boundary layer affects the approximation on a large part of the domain. Fig. 2(b) shows that the magnitude of the fine-scale solution on the outflow boundary of the domain is considerable. Hence, the assumption of vanishing fine scales, which is critical in the derivation of the classical VMS model, is severely violated. In contrast, the augmented VMS model is exactly the Nitsche projection



**Fig. 3.** One-dimensional results using three quadratic or cubic elements.

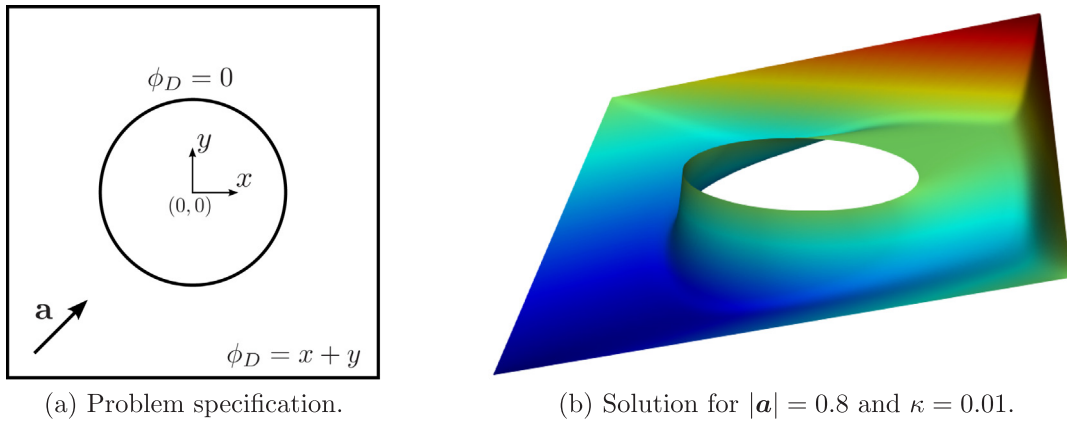
of the exact solution. As a result, the nodal exactness of the computational solution is retrieved and the adverse effect of the boundary layer is constrained to a single element. Due to the relative simplicity of this problem, the solution obtained with the generalized model is only affected by the estimation of the model parameters. As it is nearly identical to the exact coarse-scale solution we conclude that, at least for this simple case, the estimation strategy is effective.

**Remark 6.** It is well known that for the current case the classical VMS term simplifies to a (consistent) diffusion term. Interestingly, in a similar sense the augmented term in the VMS formulation simplifies to a reduced diffusion in the symmetric part of Nitsche's formulation. In this context, the solution obtained by using the classical VMS model may be interpreted as excessively diffusive in the boundary layer, which is (consistently) counteracted by the augmented VMS term.

## 5.2. Higher-order basis functions

We use the same problem formulation but discretize the domain with three higher-order elements. With quadratic basis functions and  $\beta = 3/h$  we obtain the solutions from Figs. 3(a) and 3(b), and with cubic basis function and  $\beta = 6/h$  we obtain the solutions from Figs. 3(c) and 3(d). We can largely draw the same conclusions as for the linear basis functions: the fine-scale solution deviates significantly from zero at the outflow boundary. As a result, the solution quality of the classical VMS model is spoiled. By using the augmented VMS model we obtain the Nitsche projection of the exact solution. This gives a nodally exact solution, where the adverse effect of the boundary layer is contained within the boundary element. These points are all compliant with the theory of Section 2.2.

Additionally, we observe that the solution for the approximate augmented VMS model is very close to the exact coarse-scale solution. For the linear basis functions we concluded that the estimations of the model parameters are



**Fig. 4.** Two-dimensional model problem for linear basis functions.

effective. We can now also conclude that the approximation of the differential operator described Section 3.4.1 is effective, at least for this one-dimensional case.

## 6. Numerical experiments for a two-dimensional model problem

Next, we present numerical experiments for a two-dimensional domain. All the model approximations become important, and their effectiveness can be assessed.

### 6.1. Linear basis functions, high and low advective dominance

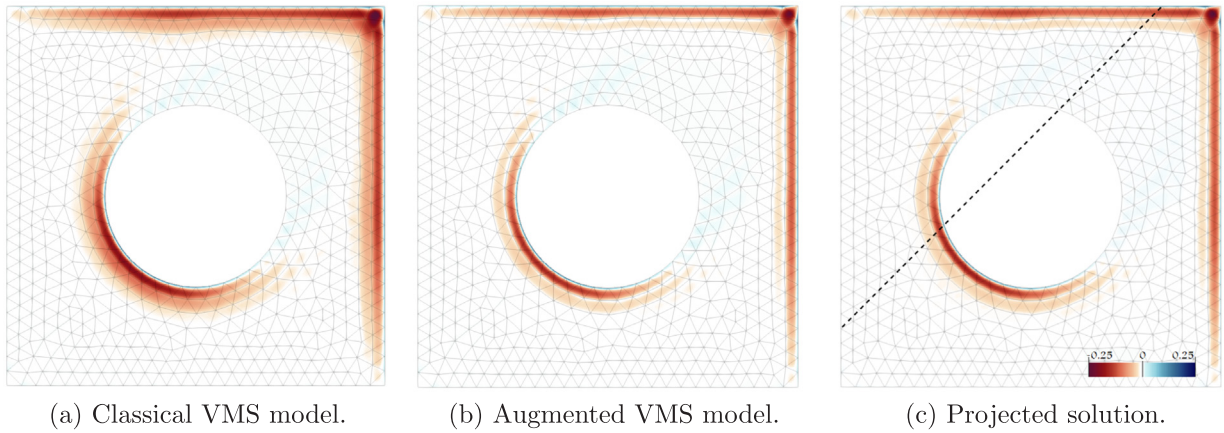
We consider a model problem of a unit square with a circular hole of radius 0.24 in the center. Dirichlet conditions are enforced on all boundaries;  $\phi_D = 0$  around the circular cut-out, and  $\phi_D = x + y$  around the square. The advective field acts across the diagonal with a magnitude of 0.8, the diffusivity is  $\kappa = 0.01$  or  $\kappa = 0.003$  (global Péclet numbers of 80 and 800/3 respectively), and we use  $\beta = 10/h$ . Fig. 4(a) schematically illustrates the model problem, and Fig. 4(b) shows the solution for  $\kappa = 0.01$  obtained on a highly refined mesh with a standard Galerkin finite element method. The solution features multiple boundary layers with various heights and thicknesses and at various orientations with respect to the advective vector.

The performance of the models can most clearly be assessed by investigating the resulting fine-scale solutions. These are shown in Figs. 5 and 6 for  $\kappa = 0.01$  and  $\kappa = 0.003$  respectively. In Figs. 5(a) and 6(a) classical VMS stabilization is used (i.e.  $\gamma = 0$ ), and Figs. 5(b) and 6(b) show the results for the augmented model on a boundary fitted mesh. Additionally, we project the overrefined solution onto the coarse-scale function space using the Nitsche projector, and show the resulting fine-scale solution in Figs. 5(c) and 6(c). This represents the ‘exact’ fine-scale solution.

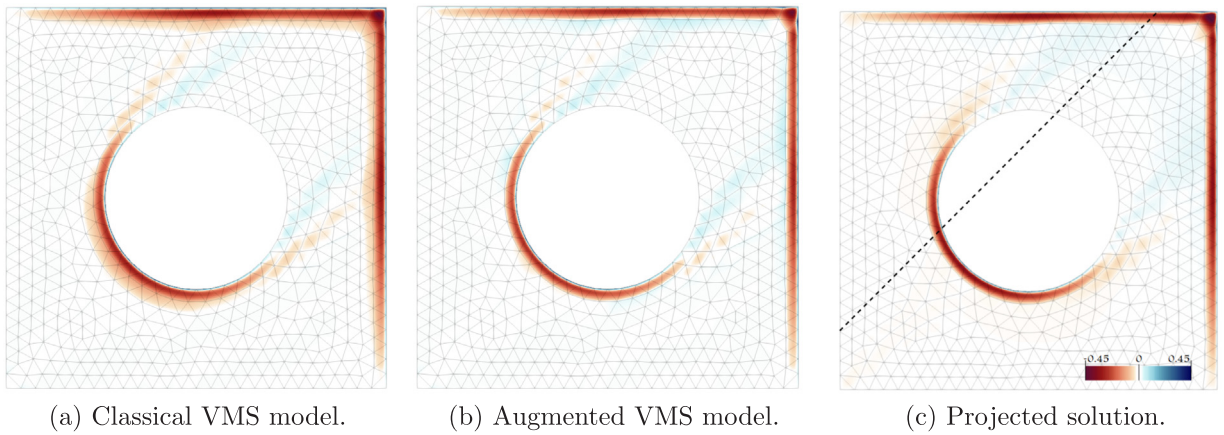
The observations made for the one-dimensional case almost directly transfer to this two-dimensional problem. The Nitsche projector aims to constrain the impact of the high gradients to the boundary layer elements, without spoiling the results further into the domain. This is illustrated by large fine-scale solutions in only a single row of elements adjacent to the outflow boundary. When the classical VMS model is used, we observe a significant thickening of the range of nonzero fine scales; interpretable as excessive diffusion in the coarse-scale solution. When we add the additional modeling term this thickening is decreased, which leads to nearly the same solution quality as that obtained with the Nitsche projection. We observe these effects irrespective of the Péclet number.

To further illustrate the significance of the change, we show all three solutions on a cut-plane in Fig. 7. Note, in particular, the similarity of Fig. 7(b) and the corresponding figure for the one-dimensional case (Fig. 2(b)). The fine-scale solution corresponding to a completely non-stabilized computation is also plotted to put the overall improvement of the solution quality into context.

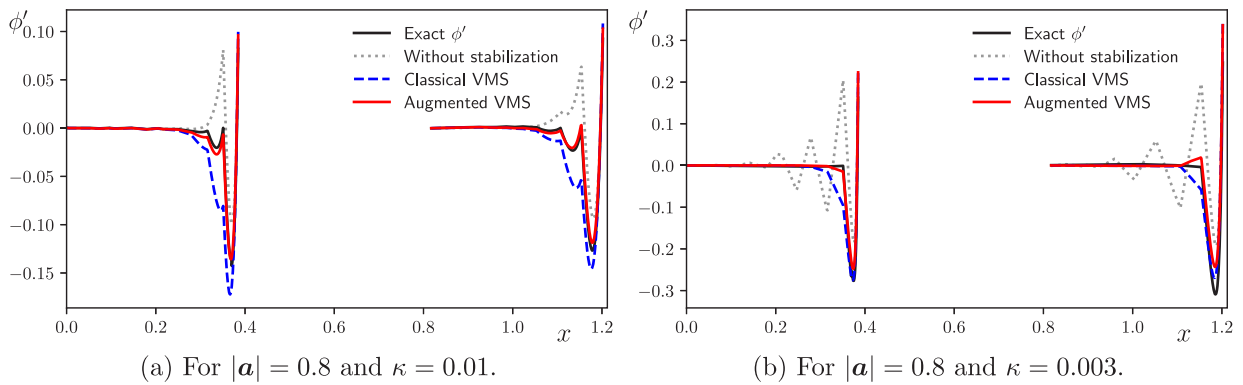
Convergence in the  $L^2$  or  $H^1$  (semi)norms is not indicative of solution quality for the current case;  $L^2$  projections of shocks lead to highly oscillatory solutions such that the non-stabilized solution often achieves the lowest  $L^2$  error,



**Fig. 5.** Fine-scale solutions  $\phi - \phi^h$  (errors) for linear basis functions and  $|a| = 0.8$  and  $\kappa = 0.01$ . Element Péclet number of approximately 3.

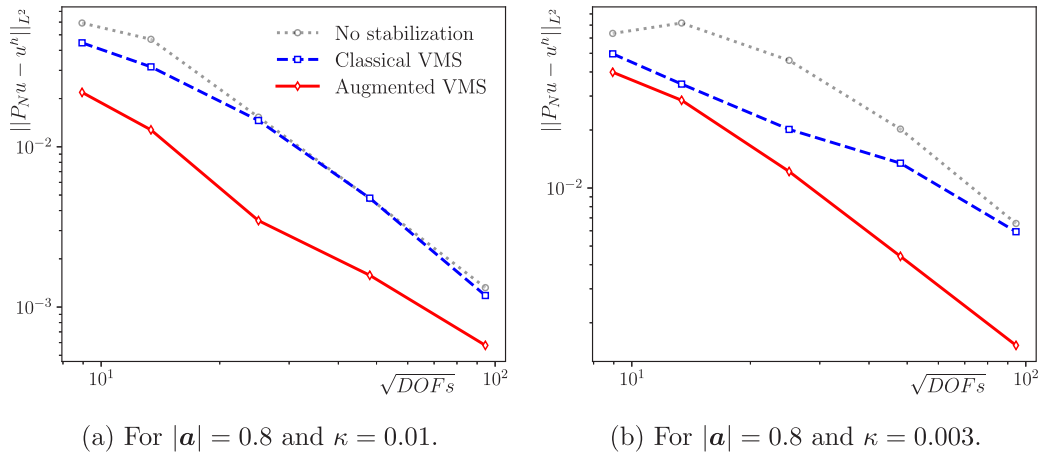


**Fig. 6.** Fine-scale solutions  $\phi - \phi^h$  (errors) for linear basis functions and  $|a| = 0.8$  and  $\kappa = 0.003$ . Element Péclet number of approximately 10.



**Fig. 7.** Fine-scale solutions on the cut-planes from Figs. 5(c) and 6(c).

and neither the classical nor the augmented VMS model aims to achieve optimality in the  $H^1$  seminorm as the boundary conditions are not enforced strongly. Rather, the use of weakly enforced boundary conditions implies the



**Fig. 8.** Error with respect to the exact coarse-scale solution, using linear basis functions.

optimality condition of Eq. (12), satisfied by solutions that minimize Eq. (9). The error of interest is thus the one with respect to the ‘optimal’ solution; the exact coarse-scale solution obtained with the Nitsche projector. This error indicates the performance of the fine-scale model. This error is plotted for different mesh densities in Figs. 8(a) and 8(b). Figs. 5 to 7 correspond to the third data point in these convergence graphs. Both graphs show a considerable reduction in error when the augmented VMS model is used, which persists throughout mesh refinement. We observe that the ( $L^2$ ) error reduction from classical to augmented VMS is often of the same magnitude, if not larger, than from non-stabilized to classical VMS.

**Remark 7.** Analogous to the one-dimensional case, the augmented VMS term decreases the diffusivity in the symmetry part of the Nitsche formulation. Different from the one-dimensional case, this becomes vector-valued and the formulation becomes a streamline directed modified diffusion on the boundary. One could interpret this as a boundary equivalent of the streamline diffusion that the classical VMS terms revert to for the same case.

## 6.2. Higher-order basis functions

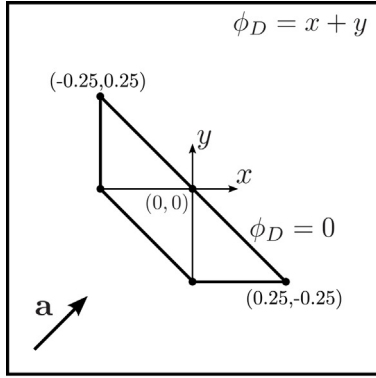
Next, we change the geometry to a square with a polygonal exclusion, as depicted in Fig. 9. An exact geometry representation can be achieved, which, for these higher-order basis functions, is important for accurately computing the boundary integrals for the Nitsche projection  $\mathcal{P}_N \phi$ . We focus on the advection dominated case of  $|a| = 0.8$  and  $\kappa = 0.003$ , and we use  $\beta = 4P^2/h$ .

Figs. 10(a) to 10(c) show the fine-scale solutions for quadratic basis functions obtained with the classical VMS model, the augmented VMS model and the Nitsche projector respectively. We observe that the classical VMS model with the parameter estimation developed in Sections 3.4.1 and 3.4.2 already performs remarkably well. The boundary layers are almost exclusively contained in a single row of elements.

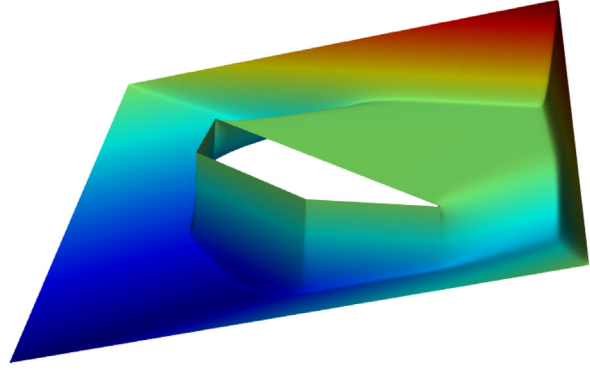
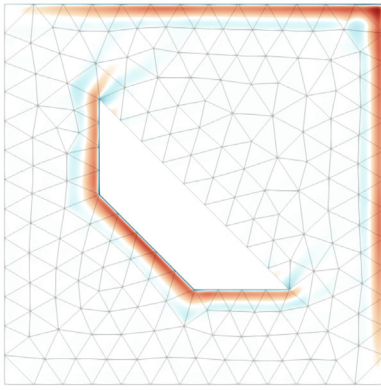
When we add the augmented term in the VMS model, the obtained error field qualitatively more closely resembles the exact fine-scale solution shown in Fig. 10(c). We do, however, also observe some small oscillations. This is consistent with the decreased diffusion interpretation proposed in Remarks 6 and 7. A more detailed analysis of the resulting error confirms that the solution obtained with the augmented VMS model more closely resembles the exact coarse-scale solution defined by the Nitsche projector. This is shown in Fig. 11(a), where the resulting fine-scale solution is plotted along a cut-plane, as well as in Fig. 11(b), which plots the  $L^2$ -error with respect to the exact coarse-scale solution for various mesh densities.

Finally, if we use cubic basis functions, we obtain the results from Figs. 12 and 13. Similar conclusions may be drawn as for the case of quadratic basis functions: the classical model with the parameters from Table 1 leads to a coarse-scale solution where the error is contained in the first row of elements. Adding the augmented term results in a solution that exhibits small oscillations, but nonetheless bears closer resemblance to the exact coarse-scale solution, as measured qualitatively in Fig. 13(a) and quantitatively in Fig. 13(b).

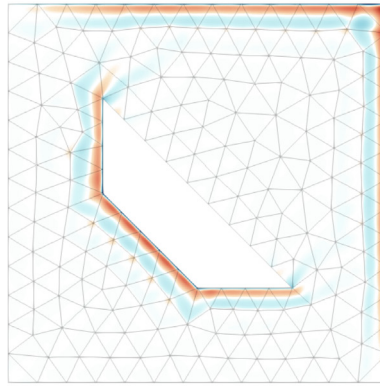




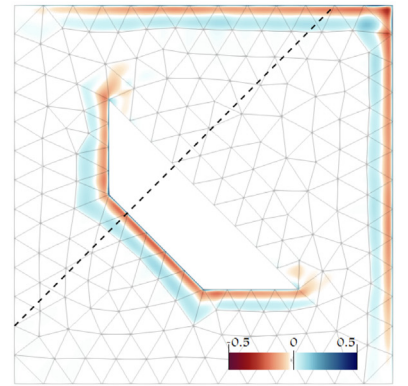
(a) Problem specification.

(b) Solution for  $|\mathbf{a}| = 0.8$  and  $\kappa = 0.003$ .**Fig. 9.** Two-dimensional model problem for higher-order basis functions.

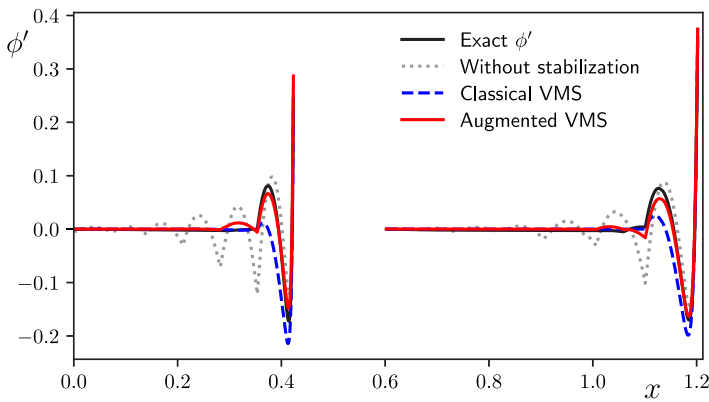
(a) Classical VMS model.



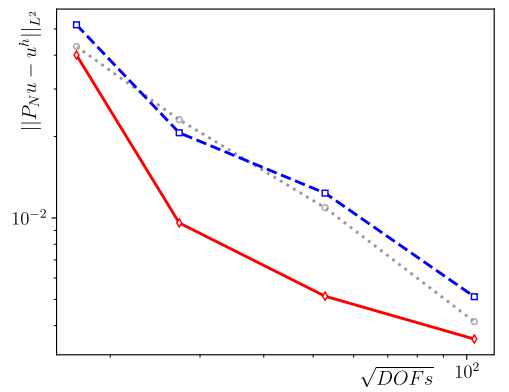
(b) Augmented VMS model.



(c) Projected solution.

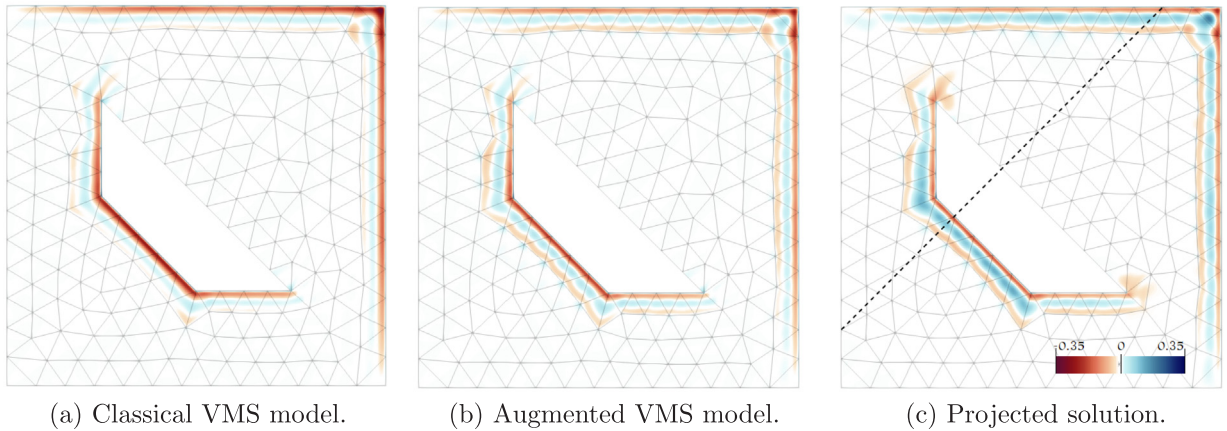
**Fig. 10.** Fine-scale solutions  $\phi - \phi^h$  (errors) for quadratic basis functions and  $|\mathbf{a}| = 0.8$  and  $\kappa = 0.003$ . Element Péclet number of approximately 20.

(a) Solutions on the cut-plane illustrated in fig. 10c.

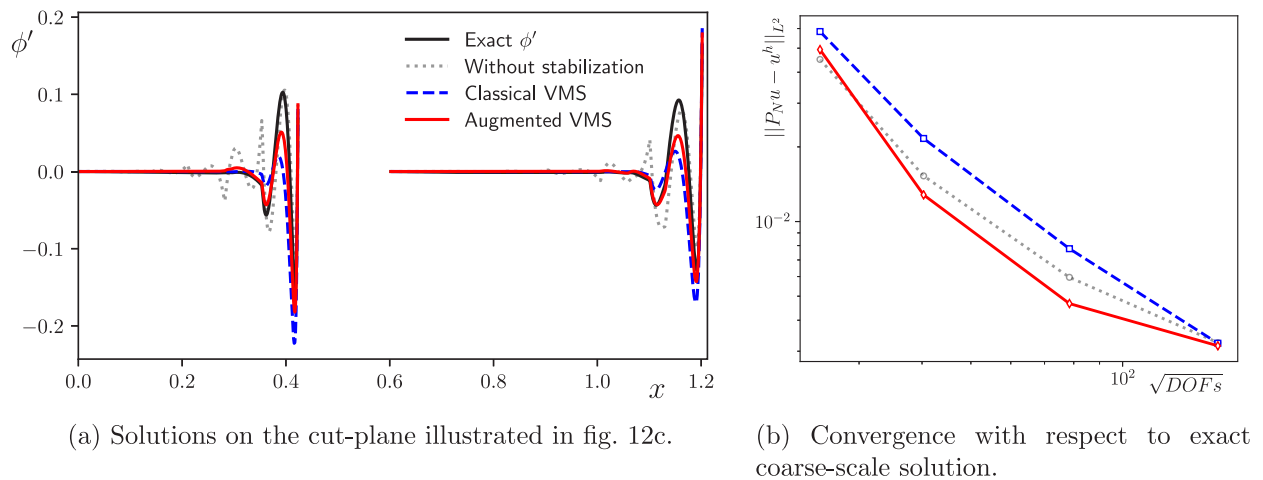


(b) Convergence with respect to exact coarse-scale solution.

**Fig. 11.** Detailed error behavior for quadratic basis functions.



**Fig. 12.** Fine-scale solutions  $\phi - \phi^h$  (errors) for cubic basis functions and  $|a| = 0.8$  and  $\kappa = 0.003$ . Element Péclet number of approximately 20.



**Fig. 13.** Detailed error behavior for cubic basis functions.

It should also be noted that the Dirichlet boundary conditions are more closely satisfied with these cubic basis functions, as shown in Fig. 13(a). This is, at least in part, due to the larger penalty parameter  $\beta \propto P^2$ . The near-strong enforcement of the Dirichlet condition leaves a small fine-scale boundary value. The new term in the augmented model has almost no effect, and the classical VMS model suffices. Indeed, the difference between the augmented and classical models is not as pronounced as it was in earlier simulations. These results convey that the augmented model provides fine-scale corrections in the pre-asymptotic regime, and vanishes (asymptotically) when such fine-scale corrections cease to be relevant.

## 7. Conclusion

In this article, we place Nitsche's method within the variational multiscale framework. Individually, the variational multiscale method and Nitsche's method have shown great value in the context of fluid mechanics. With their merger, we are in a position to develop a fine-scale model that is appropriate for use in combination with Nitsche's method.

When the Dirichlet boundary conditions are enforced weakly, the standard  $H_0^1$  projector is no longer applicable for the scale decomposition around which the variational multiscale method revolves. Instead, we propose a new projector, which we call the Nitsche projector. We show that adoption of this projector in the multiscale formulation



naturally leads to Nitsche's formulation. That is, both the penalty term and the symmetry term in Nitsche's formulation automatically fall into place as part of the variational multiscale scale decomposition.

The model for the remaining fine-scale terms is based on the inversion of the fine-scale problem, which, in turn, is formally posed in the kernel space of the projector. We show that the functional constraints that define the  $H_0^1$  projector are a subset of those corresponding to the Nitsche projector, such that we can largely base the inversion of the fine-scale problem on existing theory. An important difference in the context of weakly enforced boundary conditions is that the assumption of vanishing fine scales on element boundaries is no longer applicable for elements adjacent to the Dirichlet boundary. As a result, the fine-scale model that we obtain is the classical VMS model plus an additional boundary term. This 'augmented' term takes into account the non-vanishing fine scales on the Dirichlet boundary. It may be interpreted as a consistent streamline diffusion in the symmetry term of Nitsche's formulation.

Additionally, we develop approximations for the modeling terms  $\tau$  and  $\gamma$  based on the fine-scale Green's functions. These expressions and approximation strategies are also suitable for discretization with higher-order elements.

With this new model, and these new parameter definitions, we retrieve nodally exact solutions on one-dimensional meshes for all polynomial orders. This is an important property of the classical VMS model, which is lost when the boundary conditions are enforced weakly. On two-dimensional domains, we observe that the augmented model more closely resembles the actual coarse-scale solution defined by the Nitsche projector, as measured in an  $L^2$  sense. This holds for all polynomial orders. For quadratic and cubic basis functions, the model without the augmented term already performs very well with the newly developed  $\tau$  approximations. The error due to the boundary layer is contained in a single row of elements. For linear basis functions, however, the classical VMS model leads to a too thick boundary layer. This is almost completely mitigated when the augmented model is added to the formulation.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

D. Schillinger gratefully acknowledges support from the National Science Foundation via the NSF CAREER Award No. 1651577 and from the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) via the Emmy Noether Award SCH 1249/2-1. M.F.P. ten Eikelder and I. Akkerman are grateful for the support of Delft University of Technology.

### Appendix. Fine-scale Green's functions and $\gamma$ 's vanishing moments

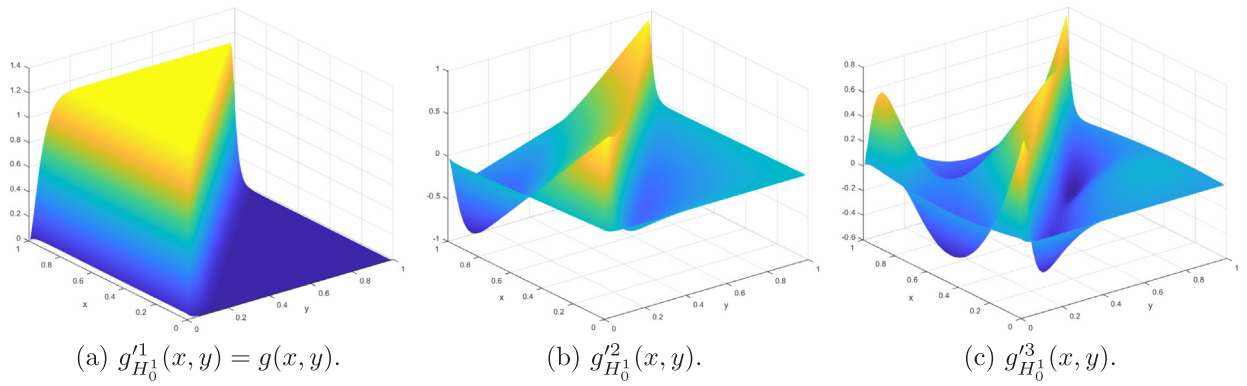
In this appendix we draw conclusions on the  $\gamma$  expression for the fine-scale Green's functions corresponding to  $P = 2$  and  $P = 3$  polynomial basis functions. Recall the definition of  $\gamma$ :

$$\gamma = \frac{1}{|F|} \int_K \int_F \frac{x^{P-1}}{h^{P-1}} \mathbb{H} g'_{H_0^1}{}^P(x, y) dy dx, \quad (\text{A.1})$$

where we now add the superscript  $P$  to the fine-scale Green's function to denote the polynomial order of the coarse-scale basis functions on which  $\mathcal{P}_{H_0^1}$  projects.

The fine-scale Green's function associated to the  $H_0^1$ -projector has been studied extensively in [47]. The authors prove the element local nature of  $g'_{H_0^1}(x, y)$  in the one-dimensional case. They also show that in a single element, the fine-scale Green's function can then be obtained from the element local classical Green's function:

$$g'_{H_0^1}{}^P(x, y) = g(x, y) - \left[ \int_0^h g(x, y) dy \quad \cdots \quad \int_0^h y^{P-2} g(x, y) dy \right] \\ \left[ \begin{array}{ccc} \int_0^h \int_0^h g(x, y) dx dy & \cdots & \int_0^h \int_0^h y^{P-2} g(x, y) dx dy \\ \vdots & \ddots & \vdots \\ \int_0^h \int_0^h x^{P-2} g(x, y) dx dy & \cdots & \int_0^h \int_0^h x^{P-2} y^{P-2} g(x, y) dx dy \end{array} \right]^{-1} \left[ \begin{array}{c} \int_0^h g(x, y) dx \\ \vdots \\ \int_0^h x^{P-2} g(x, y) dx \end{array} \right]. \quad (\text{A.2})$$



**Fig. A.14.** Fine-scale Green's functions on one element for different polynomial coarse-scale basis functions. Using  $\kappa = 0.02$  and  $a = 0.8$  on an element of size  $h = 1$ .

Refer to [42] or [9] for the expression for  $g(x, y)$ . The resulting functions for  $P = 1, 2$  and  $3$  are plotted in Fig. A.14. For this particular case, the fine-scale Green's function for  $P = 1$  is exactly the element local classical Green's function  $g(x, y)$ .

The derivation in Section 3.3 requires vanishing 'moments' of the fine-scale Green's function, as stated in Eq. (32b). We prove that this holds for  $g_{H_0}^{'P}(x, y)$  in the following theorem.

**Theorem A.1.** Define a  $\gamma$ -like parameter that depends on coarse-scale polynomial order  $P$  and a  $Q$ th moment:

$$\gamma^{Q,P} := \frac{1}{|F|} \int_0^h \int_F \frac{x^{Q-1}}{h^{Q-1}} \mathbb{H} g_{H_0}^{'P}(x, y) dy dx, \quad (\text{A.3})$$

then from the definition of  $g_{H_0}^{'P}(x, y)$  in Eq. (A.2) it follows that:

$$\gamma^{Q,P} = \begin{cases} 0 & \text{if } Q < P, \\ \gamma & \text{if } Q = P. \end{cases} \quad (\text{A.4})$$

**Proof.** The equality  $\gamma^{Q,P} = \gamma$  for  $Q = P$  follows directly from the definition of  $\gamma$ . For  $Q < P$  we substitute the definition of the fine-scale Green's function from Eq. (A.2). After carrying out the integration in Eq. (A.2), the first vector becomes independent of the  $y$ -variable, and the last vector independent of the  $x$ -variable. The center matrix is filled with constants. This means that the differential operator and integration from Eq. (A.3) act on different vectors and they can thus be separated. After re-ordering of derivatives and integrals we obtain:

$$\gamma^{Q,P} = \gamma^{Q,1} - \begin{bmatrix} \int_0^h \int_0^h x^{Q-1} g(x, y) dx dy & \cdots & \int_0^h \int_0^h x^{Q-1} y^{P-2} g(x, y) dx dy \\ \int_0^h \int_0^h g(x, y) dx dy & \cdots & \int_0^h \int_0^h y^{P-2} g(x, y) dx dy \\ \vdots & \ddots & \vdots \\ \int_0^h \int_0^h x^{P-2} g(x, y) dx dy & \cdots & \int_0^h \int_0^h x^{P-2} y^{P-2} g(x, y) dx dy \end{bmatrix}^{-1} \begin{bmatrix} \gamma^{1,1} \\ \vdots \\ \gamma^{P-2,1} \end{bmatrix}. \quad (\text{A.5})$$

For ease of notation we denote the involved vectors and matrix  $\boldsymbol{\zeta}^T$ ,  $\mathbf{C}^{-1}$  and  $\boldsymbol{\xi}$ . By recognizing that  $\boldsymbol{\zeta}^T$  is the  $Q$ th row of  $\mathbf{C}$  we can write  $\boldsymbol{\zeta}^T = \mathbf{e}_Q^T \mathbf{C}$ , where  $\mathbf{e}_Q$  is a vector of zeros with a 1 at the  $Q$ th row. Substitution into the matrix–vector multiplication yields:

$$\gamma_F^{Q,P} = \gamma_F^{Q,1} - \boldsymbol{\zeta}^T \mathbf{C}^{-1} \boldsymbol{\xi} = \gamma_F^{Q,1} - \mathbf{e}_Q^T \mathbf{C} \mathbf{C}^{-1} \boldsymbol{\xi} = \gamma_F^{Q,1} - \mathbf{e}_Q^T \boldsymbol{\xi} = \gamma_F^{Q,1} - \gamma_F^{Q,1} = 0. \quad (\text{A.6})$$

Note that this only holds for  $Q < P$ , since  $\mathbf{C}$  has  $P - 1$  rows.  $\square$

## References

- [1] T.J.R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, *Comput. Methods Appl. Mech. Engrg.* 127 (1–4) (1995) 387–401.
- [2] T.J.R. Hughes, J.R. Stewart, A space-time formulation for multiscale phenomena, *J. Comput. Appl. Math.* 74 (1–2) (1996) 217–229.
- [3] T.J.R. Hughes, G.R. Feijóo, L. Mazzei, J.-B. Quincy, The variational multiscale method – a paradigm for computational mechanics, *Comput. Methods Appl. Mech. Engrg.* 166 (1–2) (1998) 3–24.
- [4] L.P. Franca, G. Hauke, A. Masud, Revisiting stabilized finite element methods for the advective–diffusive equation, *Comput. Methods Appl. Mech. Engrg.* 195 (13–16) (2006) 1560–1572.
- [5] C. Coley, J.A. Evans, Variational multiscale modeling with discontinuous subscales: Analysis and application to scalar transport, *Meccanica* 53 (6) (2018) 1241–1269.
- [6] R. Codina, Comparison of some finite element methods for solving the diffusion-convection-reaction equation, *Comput. Methods Appl. Mech. Engrg.* 156 (1–4) (1998) 185–210.
- [7] Y. Bazilevs, V.M. Calo, J.A. Cottrell, T.J.R. Hughes, A. Reali, G. Scovazzi, Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows, *Comput. Methods Appl. Mech. Engrg.* 197 (1–4) (2007) 173–201.
- [8] F. Brezzi, L.P. Franca, T.J.R. Hughes, A. Russo,  $b = \int G$ , *Comput. Methods Appl. Mech. Engrg.* 145 (3–4) (1997) 329–339.
- [9] T.J.R. Hughes, G. Scovazzi, L.P. Franca, Multiscale and stabilized methods, in: E. Stein, R. De Borst, T.J.R. Hughes (Eds.), *Encyclopedia of Computational Mechanics*, John Wiley & Sons, Ltd, 2004, chapter 4.
- [10] A.N. Brooks, T.J.R. Hughes, Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 32 (1–3) (1982) 199–259.
- [11] M.F.P. ten Eikelder, I. Akkerman, Correct energy evolution of stabilized formulations: The relation between vms, supg and gls via dynamic orthogonal small-scales and isogeometric analysis. I: The convective–diffusive context, *Comput. Methods Appl. Mech. Engrg.* 331 (2018) 259–280.
- [12] T.J.R. Hughes, L.P. Franca, G.M. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective–diffusive equations, *Comput. Methods Appl. Mech. Engrg.* 73 (2) (1989) 173–189.
- [13] T.E. Tezduyar, Stabilized finite element formulations for incompressible flow computations, *Adv. Appl. Mech.* 28 (C) (1991) 1–44.
- [14] R. Codina, J. Principe, O. Guasch, S. Badia, Time dependent subscales in the stabilized finite element approximation of incompressible flow problems, *Comput. Methods Appl. Mech. Engrg.* 196 (21) (2007) 2413–2430.
- [15] A. Masud, R.A. Khurram, A multiscale finite element method for the incompressible Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 195 (13–16) (2006) 1750–1777.
- [16] K. Chang, T.J.R. Hughes, V.M. Calo, Isogeometric variational multiscale large-eddy simulation of fully-developed turbulent flow over a wavy wall, *Comput. & Fluids* 68 (2012) 94–104.
- [17] V. Gravemeier, W.A. Wall, Residual-based variational multiscale methods for laminar, transitional and turbulent variable-density flow at low mach number, *Internat. J. Numer. Methods Fluids* 65 (10) (2011) 1260–1278.
- [18] M.F.P. ten Eikelder, I. Akkerman, Correct energy evolution of stabilized formulations: The relation between VMS, SUPG and GLS via dynamic orthogonal small-scales and isogeometric analysis. II: The incompressible Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 340 (2018) 1135–1154.
- [19] R. Codina, J. Principe, J. Baiges, Subscales on the element boundaries in the variational two-scale finite element method, *Comput. Methods Appl. Mech. Engrg.* 198 (5–8) (2009) 838–852.
- [20] R. Codina, J. Baiges, Finite element approximation of transmission conditions in fluids and solids introducing boundary subgrid scales, *Internat. J. Numer. Methods Engrg.* 87 (1–5) (2011) 386–411.
- [21] C. Farhat, I. Harari, L.P. Franca, The discontinuous enrichment method, *Comput. Methods Appl. Mech. Engrg.* 190 (48) (2001) 6455–6479.
- [22] T.J. Truster, A. Masud, Primal interface formulation for coupling multiple PDEs: A consistent derivation via the variational multiscale method, *Comput. Methods Appl. Mech. Engrg.* 268 (2014) 194–224.
- [23] N.C.B. Arruda, R.C. Almeida, E.G.D. do Carmo, Discontinuous subgrid formulations for transport problems, *Comput. Methods Appl. Mech. Engrg.* 199 (49–52) (2010) 3227–3236.
- [24] M.F.P. ten Eikelder, Y. Bazilevs, I. Akkerman, A theoretical framework for discontinuity capturing: Joining variational multiscale analysis and variation entropy theory, *Comput. Methods Appl. Mech. Engrg.* 359 (2020) 112664.
- [25] A. Masud, M. Anguiano, I. Harari, Modeling of steep layers in singularly perturbed diffusion–reaction equation via flexible fine-scale basis, *Comput. Methods Appl. Mech. Engrg.* 372 (2020) 113343.
- [26] R. Stenberg, On some techniques for approximating boundary conditions in the finite element method, *J. Comput. Appl. Math.* 63 (1–3) (1995) 139–148.
- [27] Y. Bazilevs, T.J.R. Hughes, Weak imposition of Dirichlet boundary conditions in fluid mechanics, *Comput. & Fluids* 36 (1) (2007) 12–26.
- [28] Y. Bazilevs, C. Michler, V.M. Calo, T.J.R. Hughes, Weak Dirichlet boundary conditions for wall-bounded turbulent flows, *Comput. Methods Appl. Mech. Engrg.* 196 (49–52) (2007) 4853–4862.
- [29] M. Juntunen, R. Stenberg, Nitsche's method for general boundary conditions, *Math. Comput.* 78 (267) (2009) 1353–1374.
- [30] Y. Bazilevs, C. Michler, V.M. Calo, T.J.R. Hughes, Isogeometric variational multiscale modeling of wall-bounded turbulent flows with weakly enforced boundary conditions on unstretched meshes, *Comput. Methods Appl. Mech. Engrg.* 199 (13–16) (2010) 780–790.
- [31] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, in: *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, vol. 36, Springer, 1971, pp. 9–15.

- [32] E. Burman, A penalty-free nonsymmetric Nitsche-type method for the weak imposition of boundary conditions, *SIAM J. Numer. Anal.* 50 (4) (2012) 1959–1981.
- [33] A. Hansbo, P. Hansbo, An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems, *Comput. Methods Appl. Mech. Engrg.* 191 (47–48) (2002) 5537–5552.
- [34] J. Dolbow, I. Harari, An efficient finite element method for embedded interface problems, *Internat. J. Numer. Methods Engrg.* 78 (2) (2009) 229–252.
- [35] E. Burman, P. Hansbo, Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method, *Appl. Numer. Math.* 62 (4) (2012) 328–341.
- [36] A. Embar, J. Dolbow, I. Harari, Imposing Dirichlet boundary conditions with Nitsche’s method and spline based finite elements, *Internat. J. Numer. Methods Engrg.* 83 (7) (2010) 877–898.
- [37] Y. Bazilevs, M.-C. Hsu, M.A. Scott, Isogeometric fluid–structure interaction analysis with emphasis on non-matching discretizations, and with application to wind turbines, *Comput. Methods Appl. Mech. Engrg.* 249–252 (2012) 28–41.
- [38] F. Xu, D. Schillinger, D. Kamensky, V. Varduhn, C. Wang, M.-C. Hsu, The tetrahedral finite cell method for fluids: Immersogeometric analysis of turbulent flow around complex geometries, *Comput. & Fluids* 141 (2016) 135–154.
- [39] M.-C. Hsu, C. Wang, F. Xu, A.J. Herrema, A. Krishnamurthy, Direct immersogeometric fluid flow analysis using B-rep CAD models, *Comput. Aided Geom. Design* 43 (2016) 143–158.
- [40] D. Kamensky, M.-C. Hsu, D. Schillinger, J.A. Evans, A. Aggarwal, Y. Bazilevs, M.S. Sacks, T.J.R. Hughes, An immersogeometric variational framework for fluid–structure interaction: Application to bioprosthetic heart valves, *Comput. Methods Appl. Mech. Engrg.* 284 (2015) 1005–1053.
- [41] T. Hoang, C.V. Verhoosel, C.-Z. Qin, F. Auricchio, A. Reali, E.H. van Brummelen, Skeleton-stabilized immersogeometric analysis for incompressible viscous flow problems, *Comput. Methods Appl. Mech. Engrg.* 344 (2019) 421–450.
- [42] S.K.F. Stoter, S.R. Turteltaub, S.J. Hulshoff, D. Schillinger, Residual-based variational multiscale modeling in a discontinuous Galerkin framework, *Multiscale Model. Simul.* 16 (3) (2018) 1333–1364.
- [43] S.K.F. Stoter, S.R. Turteltaub, S.J. Hulshoff, D. Schillinger, A discontinuous Galerkin residual-based variational multiscale method for modeling subgrid-scale behavior of the viscous Burgers equation, *Internat. J. Numer. Methods Fluids* 88 (5) (2018) 217–238.
- [44] H.J.C. Barbosa, T.J.R. Hughes, The finite element method with Lagrange multipliers on the boundary: Circumventing the Babuška-Brezzi condition, *Comput. Methods Appl. Mech. Engrg.* 85 (1) (1991) 109–128.
- [45] S.K.F. Stoter, *The Variational Multiscale Method for Discontinuous Galerkin Type Finite Element Formulations* (Ph.D. thesis), University of Minnesota, 2019,.
- [46] T. Warburton, J.S. Hesthaven, On the constants in hp-finite element trace inverse inequalities, *Comput. Methods Appl. Mech. Engrg.* 192 (25) (2003) 2765–2773.
- [47] T.J.R. Hughes, G. Sangalli, Variational multiscale analysis: The fine-scale Green’s function, projection, optimization, localization, and stabilized methods, *Soc. Ind. Appl. Math.* 45 (2) (2007) 539–557.
- [48] I. Harari, U. Albocher, Spectral investigations of Nitsche’s method, *Finite Elem. Anal. Des.* 145 (2018) 20–31.
- [49] I. Harari, T.J.R. Hughes, What are C and h?: Inequalities for the analysis and design of finite element methods, *Comput. Methods Appl. Mech. Engrg.* 97 (2) (1992) 157–192.
- [50] I. Akkerman, Y. Bazilevs, V.M. Calo, T.J.R. Hughes, S. Hulshoff, The role of continuity in residual-based variational multiscale modeling of turbulence, *Comput. Mech.* 41 (3) (2008) 371–378.
- [51] T.J.R. Hughes, M. Mallet, A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems, *Comput. Methods Appl. Mech. Engrg.* 58 (3) (1986) 305–328.
- [52] F. Shakib, T.J.R. Hughes, Z. Johan, A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 89 (1–3) (1991) 141–219.
- [53] T.E. Tezduyar, Y. Osawa, Finite element stabilization parameters computed from element matrices and vectors, *Comput. Methods Appl. Mech. Engrg.* 190 (3–4) (1999) 411–430.