

A variational multiscale approach to PDE-constrained optimization problems arising in Data-Driven Computational Mechanics

Ramon Codina^{1,2}, Roberto F. Ausas³, Pedro B. Bazon³, and Cristian G. Gebhardt^{*4}

¹*Universitat Politècnica de Catalunya, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain*

²*Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Campus Nord UPC, Gran Capità s/n, 08034 Barcelona, Spain*

³*Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, Av. Trab. São Carlsense 400, 13566-590 São Carlos-SP, Brazil*

⁴*University of Bergen, Geophysical Institute and Bergen Offshore Wind Centre (BOW), Allégaten 70, 5007 Bergen, Norway*

March 13, 2026

Abstract

We consider the primal and dual forms of the optimality conditions for PDE-constrained optimization problems arising in *Data-Driven Computational Mechanics* when specialized to the reaction-diffusion context. Starting with the continuous setting, we establish well-posedness of such concomitant formulations. Then, we propose stable and consistent finite element approximations for these underlying primal and dual problems relying on the *Variational MultiScale* framework. For quasi-uniform finite element partitions, we investigate approximations' general properties and establish well-posedness for two canonical choices of the sub-grid scales, i.e., the *Algebraic Sub-Grid Scale* and *Orthogonal Sub-Grid Scale*. Moreover, for continuous finite element functions, we are able to move back and forth between the discrete primal and dual formulations only by changing the design of the stabilization parameters. To conclude, we stress-test the proposed approximations through a series of progressively sophisticated cases, providing both a comparative and qualitative assessment of their numerical performance.

Keywords: data-driven computational mechanics, PDE-constrained optimization problems, diffusion-reaction problem, primal and dual formulations, stabilized finite elements, variational multiscale framework.

*In memoriam of Prof. Gustavo C. Buscaglia [1964-2025],
a dear friend and inspiring colleague.*

1 Introduction

Data-Driven Computational Mechanics (DDCM) constitutes the most recent paradigm shift in continuum mechanics, with certainly broad applicability. Specifically for the reaction-diffusion

^{*}Corresponding author. *E-mail-address:* cristian.gebhardt@uib.no

problem, DDCM seeks to find the concentration gradient and the associated flux vector field that best match an empirical extension of gradient-flux pairs. This is achieved by formulating an optimization problem where discrete and continuous fields are simultaneously sought to minimize their discrepancy, provided the continuous fields satisfy the physical conservation and geometrical compatibility laws as well as the Second Law of Thermodynamics.

A basic question that arises is the existence of optimal continuous fields when the discrete fields are held constant. This “optimization subproblem,” which is a continuous problem, can be analyzed and solved computationally using standard methods from the linear theory of *Partial Differential Equations* (PDEs). Addressing this specific subproblem is the primary objective of this work. To set the stage for our approach, we first review key related works, differentiate our methodology from existing literature and state the novelty of our present contribution.

DDCM is a highly active research area whose origins date back less than ten years. Its core methodologies have predominantly been established within solid and structural mechanics, as exemplified by its application to the static analysis of multi-bar structures [24]. Similar ideas led to the development of a data-driven magnetostatic solver [14]. This initial approach was later modified to improve robustness in the presence of constitutive data outliers [25] and subsequently extended to the dynamic analysis of structures featuring configuration-independent mass matrices [26]. In the interest of full rigor, we note that the question of existence of solutions has also been addressed in the context of data-driven elasticity by several studies [12, 13, 19]. Other research explores a hybrid strategy involving the offline reconstruction of a smooth constitutive manifold from empirical data. This concept was first implemented for time-independent analyses [18], then adapted to solve time-dependent problems [20]. The approach has been further generalized to tackle challenging *Mathematical Programs with Complementarity Constraints* (MPCCs), particularly relevant to contact mechanics simulations [15]. Recently in [16], the structural properties of the DDCM problem discretized by means of the *Finite Element Method* (FEM) is investigated and its global solvability is established. Based on the same principles, DDCM is applied to discrete electrical circuits in [17]. Remarkably, the present work addresses a significant gap, as there is only a single antecedent of DDCM in the context of the diffusion-reaction problem, developed previously by some of the authors of this contribution [4] where a *Galerkin/Least-Squares* (GLS) method for the stabilization of the resulting finite element discretization is proposed.

All previously introduced works that numerically solve the discrete-continuous optimization problem in DDCM rely on computing a solution to a “optimization subproblem”. This is because those works adopt an *Alternating Direction Method* (ADM). This work constitutes a crucial major improvement towards applying DDCM in the diffusion-reaction problem. Moreover, our analysis considers two or more spatial dimensions in contrast to contributions that focus on problems defined on closed intervals of the real line [18, 20, 15, 16]. We introduce novel finite element formulations and establish their well-posedness for the chosen discrete spaces. The formulations proposed are consistent, and therefore convergence is basically a consequence of stability, although we shall not pursue a convergence analysis of these formulations. The required stability properties are achieved by directly approximating the underlying saddle-point problem, an approach related to established methodologies such as those found in [28, 5, 2, 6].

Starting with the continuous setting, we present the primal and dual formulations for the PDE-constrained optimization problem and establish their well-posedness. In addition and by leveraging the necessary regularity of the problem’s functional setting, we successfully obtain the minimal irreducible forms of these two formulations. Then, we step into the discrete setting by adopting a finite element discretization. To ensure stability for both the discretized primal and dual problems when using continuous arbitrary approximations, we adopt the *Variational MultiScale* (VMS) framework to systematically derive the necessary stabilization terms [22, 23, 10]. In particular, we analyze some approximation’s general properties and establish the consistency and well-posedness of the resulting schemes when considering both the *Algebraic Sub-Grid Scale* (ASGS) and *Orthogonal Sub-Grid Scale* (OSGS) methods [7, 8]. We show that one can move from the primal to the dual formulation just by changing appropriately the design of the stabilization terms, as in [1] for Darcy’s problem or in [3] for the wave equation. We then stress-test the proposed discrete approximations.

The remainder of this article is structured as follows. Section 2 briefly establishes the context by reviewing the DDCM problem and derives through two main simplifications the PDE-constrained optimization problem targeted by the present contribution. Then, it culminates, in the continuous setting, with the formulation and analysis of the associated primal and dual forms. Section 3 formulates stable finite element approximations for the primal and dual forms and establishes their well-posedness. Section 4 details the numerical implementation, realized using the Firedrake finite element platform, and presents a series of progressively sophisticated examples to illustrate the proposed setting's properties. Section 5 provides concluding remarks, discusses limitations and suggests possible avenues for future work.

2 Theory

2.1 Data-Driven Computational Mechanics

Given a domain $\Omega \subset \mathbb{R}^{d_\Omega}$ such that $d_\Omega \in \{2, 3\}$ (a bounded and connected open subset with Lipschitz-continuous boundary $\partial\Omega$), consider a physical system that occupies the closure $\bar{\Omega}$ and that is subject to body sources, i.e., $q : \Omega \rightarrow \mathbb{R}$. All necessary conditions required to formulate the conservation and geometric laws of such a physical system up to constitutive dependencies can be described with the help of: *three primal fields*, e.g., a concentration or temperature $u : \Omega \rightarrow \mathbb{R}$, the concentration or temperature gradient $\mathbf{e} : \Omega \rightarrow \mathbb{R}^{d_\Omega}$ and the flux $\mathbf{s} : \Omega \rightarrow \mathbb{R}^{d_\Omega}$; and, *two dual fields* $\lambda : \Omega \rightarrow \mathbb{R}$ and $\boldsymbol{\mu} : \Omega \rightarrow \mathbb{R}^{d_\Omega}$, where the first one corresponds to the enforcement of the conservation law in itself and the second one to the enforcement of the geometric law relating the primal fields u and $\mathbf{e}$. In this work, we only consider those conditions to formulate the conservation and geometric laws up to constitutive dependencies that can be formulated as

$$\Phi((\lambda, \boldsymbol{\mu}), (u, \mathbf{e}, \mathbf{s}); q, \Omega) := (\lambda, \zeta u + \nabla \cdot \mathbf{s} - q) + (\boldsymbol{\mu}, \mathbf{e} - \nabla u) = 0 \quad \text{in } \Omega,$$

where $(\cdot, \cdot)$ represents a suitable inner product, the differential operator ∇ is such that the identity $(\nabla u, \mathbf{s}) = -(u, \nabla \cdot \mathbf{s})$ is identically satisfied for any admissible pair $(u, \mathbf{s})$ under appropriate boundary conditions [27, Thm. 6.2] and $\zeta \in L^\infty(\Omega)$ is a reaction coefficient as considered by [4], which gives origin to the coupling of all optimality conditions. To simplify the exposition, we shall consider ζ as a non-negative physical constant in what follows. In addition, achieving physical admissibility of the primal fields $(\mathbf{e}, \mathbf{s})$ requires considering the Second Law of Thermodynamics, i.e., $\mathbf{e} \cdot \mathbf{s} \leq 0$ almost everywhere in Ω .

To include the constitutive dependencies, let us consider the set $\mathcal{N}_m = \{n\}_{n=1}^{N_m}$, where N_m stands for the number of independent experimental measurements available and

$$\mathfrak{D}_m = \{(\boldsymbol{\epsilon}_n, \mathfrak{s}_n)\}_{n \in \mathcal{N}_m}$$

represents the set of all experimental pairs that lay, up to a certain tolerance, on the underlying constitutive manifold whose local or global representation is unknown. For $\mathcal{N}_a \subseteq \mathcal{N}_m$, the set that lists all “active” elements of $\mathfrak{D}_m$ (which are to be found subject to $\Phi((\lambda, \boldsymbol{\mu}), (u, \mathbf{e}, \mathbf{s}); q, \Omega)$ vanishes identically) such that $|\mathcal{N}_a| \leq |\mathcal{N}_m|$, let $P_a(\Omega) = \{\Omega_n\}_{n \in \mathcal{N}_a}$ be a partition of the domain Ω (which is also to be sought simultaneously) such that auxiliary fields $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ can be constructed, from the active constitutive pairs, over the whole domain in the form of a simple function, namely

$$(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})(\omega) = \sum_{n \in \mathcal{N}_a} \chi_{\Omega_n}(\omega) (\boldsymbol{\epsilon}_n, \mathfrak{s}_n), \quad \forall \omega \in \Omega,$$

where $\chi_{\Omega_n} : \Omega \rightarrow \mathbb{R}$ is the characteristic function of Ω_n defined as

$$\chi_{\Omega_n}(\omega) = \begin{cases} 1, & \omega \in \Omega_n, \\ 0, & \omega \notin \Omega_n. \end{cases}$$

This construction for $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ is a good approximation in the $L^2(\Omega)$ -sense.

By choosing proximity measures φ_e and φ_s that satisfy $\varphi_e(0) = \varphi_s(0) = 0$, a possible realization of the DDCM approach relies on solving a *Discrete-Continuous Nonlinear Optimization Problem* (DCNLP) given by

$$\inf_{(u, \mathbf{e}, \mathbf{s}, \mathcal{N}_a, P_a(\Omega), \nu)} \sup_{(\lambda, \boldsymbol{\mu})} \{ \varphi_e(\mathbf{e} - \tilde{\mathbf{e}}) + \varphi_s(\mathbf{s} - \tilde{\mathbf{s}}) + \Phi((\lambda, \boldsymbol{\mu}), (u, \mathbf{e}, \mathbf{s}); q, \Omega) \}$$

$$\text{s.t. } -\mathbf{e} \cdot \mathbf{s} \geq 0, \quad \nu \geq 0, \quad \nu \mathbf{e} \cdot \mathbf{s} = 0 \}.$$

Standard choices of the proximity measures are those induced by the $L^2(\Omega)$ -norm, e.g.,

$$\varphi_e(\mathbf{e} - \tilde{\mathbf{e}}) + \varphi_s(\mathbf{s} - \tilde{\mathbf{s}}) = \frac{1}{2} \|\mathbf{e} - \tilde{\mathbf{e}}\|^2 + \frac{\kappa}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2,$$

where $\kappa \in (0, \infty)$ is a constant providing unit consistency, required to be small enough when necessary (cf. Theorem 2 below).

Remark 1. For each data set, the unit consistency factor κ can be estimated, for instance, either as the geometric mean

$$\kappa_{\text{geom}} := r_{\max} \exp \left(\frac{1}{|\mathcal{N}_m^+|} \sum_{n \in \mathcal{N}_m^+} \ln \left(\frac{r_n}{r_{\max}} \right) \right)$$

or the arithmetic mean

$$\kappa_{\text{arith}} := \frac{1}{|\mathcal{N}_m^+|} \sum_{n \in \mathcal{N}_m^+} r_n,$$

where $r_n := \|\mathbf{e}_n\|_2 / \|\mathbf{s}_n\|_2$, $\mathcal{N}_m^+ := \{n \in \mathfrak{D}_m : \|\mathbf{s}\|_2 \neq 0\}$, $|\mathcal{N}_m^+|$ stands for its cardinality and $r_{\max}$ is the maximal element of $\{r_n\}_{n \in \mathcal{N}_m^+}$, where $[r_{\max}] = [s]^{-1}$ and $[\cdot]$ stands for physical units.

We can start with our argument, only after formally introducing the PDE-constrained optimization problem at given $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ and q , which is defined as follows:

Problem 1 (P_0). Let $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ and q be given. Find $(\mathbf{y}^*, \nu^*, \mathbf{z}^*) = ((u^*, \mathbf{e}^*, \mathbf{s}^*), \nu^*, (\lambda^*, \boldsymbol{\mu}^*)) \in Y \times N \times Z$ such that

$$(\mathbf{y}^*, \nu^*, \mathbf{z}^*) = \arg \left\{ \inf_{\mathbf{y} \in Y, \nu \in N} \sup_{\mathbf{z} \in Z} \mathcal{L}(\mathbf{y}, \mathbf{z}) \text{ s.t. } -\mathbf{e} \cdot \mathbf{s} \geq 0, \quad \nu \geq 0, \quad \nu \mathbf{e} \cdot \mathbf{s} = 0 \right\}, \quad (1)$$

where

$$\mathcal{L}(\mathbf{y}, \mathbf{z}) = \frac{1}{2} \|\mathbf{e} - \tilde{\mathbf{e}}\|^2 + \frac{\kappa}{2} \|\mathbf{s} - \tilde{\mathbf{s}}\|^2 + (\lambda, \nabla \cdot \mathbf{s} + \zeta u - q) + (\boldsymbol{\mu}, \mathbf{e} - \nabla u)$$

and subject to the specification of appropriate function spaces and boundary conditions.

Remark 2. Our description of P_0 will be completed by specifying: i) appropriate function spaces U, E, S ($Y = U \times E \times S$), L, M ($Z = L \times M$), N ; and, ii) appropriate boundary conditions for the fields u and/or $\mathbf{s}$; whichever these conditions are, in P_0 it is assumed that they are homogeneous, for simplicity.

The weak form of the optimality conditions for P_0 reads

$$\left\{ \begin{array}{ll} (\delta u, \zeta \lambda + \nabla \cdot \boldsymbol{\mu}) &= 0 & \forall \delta u \in U, \\ (\delta \mathbf{e}, \mathbf{e} + \boldsymbol{\mu} - \nu \mathbf{s}) &= (\delta \mathbf{e}, \tilde{\mathbf{e}}) & \forall \delta \mathbf{e} \in E, \\ (\delta \mathbf{s}, \kappa \mathbf{s} - \nabla \lambda - \nu \mathbf{e}) &= (\delta \mathbf{s}, \kappa \tilde{\mathbf{s}}) & \forall \delta \mathbf{s} \in S, \\ (\delta \lambda, \zeta u + \nabla \cdot \mathbf{s}) &= (\delta \lambda, q) & \forall \delta \lambda \in L, \\ (\delta \boldsymbol{\mu}, -\nabla u + \mathbf{e}) &= 0 & \forall \delta \boldsymbol{\mu} \in M, \\ -\mathbf{e} \cdot \mathbf{s} &\geq 0 & \text{a.e. in } \Omega, \\ \nu &\geq 0 & \text{a.e. in } \Omega, \\ \nu \mathbf{e} \cdot \mathbf{s} &= 0 & \text{a.e. in } \Omega, \end{array} \right. \quad (2)$$

provided that the chosen function spaces and considered boundary conditions are such that all boundary terms identically vanish, i.e., at least

$$\int_{\partial\Omega} \delta u(\boldsymbol{\mu} \cdot \hat{\mathbf{n}}) d\mathcal{H}^{d_\Omega-1}(\partial\Omega) = 0, \int_{\partial\Omega} (\delta \mathbf{s} \cdot \hat{\mathbf{n}}) \lambda d\mathcal{H}^{d_\Omega-1}(\partial\Omega) = 0 \quad \text{and} \quad \int_{\partial\Omega} (\delta \boldsymbol{\mu} \cdot \hat{\mathbf{n}}) u d\mathcal{H}^{d_\Omega-1}(\partial\Omega) = 0,$$

where $\mathcal{H}^{d_\Omega-1}(\partial\Omega)$ stands for the Hausdorff measure on the boundary of Ω .

Remark 3. *The consistency in units of P_0 is achieved by noting that the dimensional relationship between ζ and κ is given by*

$$[\zeta] = [\kappa]^{-1/2} [\ell_\Omega]^{-2},$$

where ℓ_Ω represents the characteristic length of the problem. In particular, since we assume ζ is a constant we may take $\ell_\Omega^4 \zeta^2 \kappa$ as a dimensionless physical constant, which we will consider $\ell_\Omega^4 \zeta^2 \kappa > 0$. In the case $\zeta = 0$ (or $\zeta \rightarrow 0$) our arguments can be easily adapted. Therefore, the cost function conjugated with constraints' enforcement terms will be dimension free.

In the remainder of this work, we specialize DDCM to the reaction-diffusion problem. Our analysis proceeds via two sequential levels of simplification to handle the resulting PDE-constrained optimization problem:

- We assume the thermodynamically-consistent auxiliary fields $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ are prescribed, which simplifies the original DCNLP problem to a MPCC.
- We further simplify the analysis by omitting the enforcement of the Second Law of Thermodynamics for the fields $(\mathbf{e}, \mathbf{s})$. This step linearizes the MPCC, yielding a simpler problem in saddle-point form. Nevertheless, the explicit inclusion of this Second Law that we have proposed in problem P_0 should serve to design more robust optimization schemes. We shall not explore this fact in this paper.

Although the linear saddle-point problem does not intrinsically guarantee thermodynamic consistency for $(\mathbf{e}, \mathbf{s})$, we verify *a posteriori* that any computed numerical solution fulfills the Second Law of Thermodynamics.

Finally, with the simplified setting adopted, we examine under which conditions this underlying linear problem guarantees a unique global minimizer $\mathbf{y}^* = (u^*, \mathbf{e}^*, \mathbf{s}^*)$ with unique multiplier $\mathbf{z}^* = (\lambda^*, \boldsymbol{\mu}^*)$, considering not only the continuous primal and dual formulations, but also their finite element discretizations. Crucially, we introduce and analyze stabilizations via the VMS approach, which offers an alternative to the GLS technique utilized in a previous work [4].

Remark 4. *In this work, we focus on the functional-analytic structure of the underlying linear problem and therefore do not enforce the Second Law of Thermodynamics, whose unilateral and strongly nonlinear nature would require a substantially different analytical framework as well as a different numerical strategy for solving the resulting optimality conditions. Understanding the linear setting – including its limitations – is essential before addressing this additional complexity. We acknowledge the importance of thermodynamic admissibility and are developing a consistent extension that incorporates the inequality constraint, though this lies primarily on the numerical rather than the functional-analytic side and thus lies outside the present scope. Moreover, to our knowledge, no existing DDCM formulation integrates this constraint together with the associated complementarity conditions. Addressing it here would weaken the clarity and purpose of the present linear analysis.*

2.2 Primal form

The primal formulation of the current problem is given then by the following equations:

$$\begin{cases} \zeta\lambda + \nabla \cdot \boldsymbol{\mu} &= 0 & \text{in } U^*, U = H_0^1(\Omega), \\ \boldsymbol{e} + \boldsymbol{\mu} &= \tilde{\boldsymbol{e}} & \text{in } E^* \simeq E = [L^2(\Omega)]^{d_\Omega}, \\ \kappa \boldsymbol{s} - \nabla \lambda &= \kappa \tilde{\boldsymbol{s}} & \text{in } S^* \simeq S = [L^2(\Omega)]^{d_\Omega}, \\ \zeta u + \nabla \cdot \boldsymbol{s} &= q & \text{in } L^*, L = H_0^1(\Omega), \\ -\nabla u + \boldsymbol{e} &= \mathbf{0} & \text{in } M^* \simeq M = [L^2(\Omega)]^{d_\Omega}, \end{cases}$$

where $\simeq$ stands for isomorphism between functional spaces. The standard notation is used for the Lebesgue and Sobolev spaces.

By taking $M = E$, $E = \nabla U$, and $\kappa S = \nabla L$, we arrive at

$$\begin{cases} \zeta\lambda - \Delta u &= -\nabla \cdot \tilde{\boldsymbol{e}} & \text{in } U^*, U = H_0^1(\Omega), \\ -\kappa\zeta u - \Delta \lambda &= \kappa \nabla \cdot \tilde{\boldsymbol{s}} - \kappa q & \text{in } L^*, L = H_0^1(\Omega). \end{cases} \quad (3)$$

Provided higher regularity in the function spaces involved and the domain, we have the reduced strong form

$$\kappa\zeta^2 u + \Delta \Delta u = \Delta \nabla \cdot \tilde{\boldsymbol{e}} - \zeta \kappa \nabla \cdot \tilde{\boldsymbol{s}} + \zeta \kappa q \quad \text{in } L^2(\Omega),$$

where

$$u \in H_0^4(\Omega), \quad \tilde{\boldsymbol{e}} \in [H^3(\Omega)]^{d_\Omega}, \quad \tilde{\boldsymbol{s}} \in [H^1(\Omega)]^{d_\Omega} \quad \text{and} \quad q \in L^2(\Omega),$$

which is a well-posed elliptic problem. In matrix-vector representation, problem (3) takes the form

$$-\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \Delta u \\ \Delta \lambda \end{pmatrix} + \begin{bmatrix} 0 & \zeta \\ -\zeta \kappa & 0 \end{bmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} \nabla \cdot \tilde{\boldsymbol{e}} \\ \kappa \nabla \cdot \tilde{\boldsymbol{s}} - \kappa q \end{pmatrix}.$$

The eigenvalues of the “reaction” matrix are $\pm i\sqrt{\kappa}\zeta$, showing that the problem cannot be uncoupled (for $\zeta \neq 0$) and behaves as a fourth order problem. For the primal formulation, the fields $\boldsymbol{s}$ and $\boldsymbol{\mu}$ are unnecessary. However, for the dual formulation, keeping them makes sense, as we shall see. Furthermore, $\boldsymbol{e}$ or $\boldsymbol{\mu}$ could be eliminated one in terms of the other, but this is not possible in the dual formulation.

Finally, the governing equations can be expressed in their weak form as

$$\begin{aligned} B_p(\boldsymbol{x}, \delta \boldsymbol{x}) &:= \zeta(\lambda, \delta u) - (\boldsymbol{\mu}, \nabla \delta u) \\ &\quad + (\boldsymbol{e}, \delta \boldsymbol{e}) + (\boldsymbol{\mu}, \delta \boldsymbol{e}) \\ &\quad + \kappa(\boldsymbol{s}, \delta \boldsymbol{s}) - (\nabla \lambda, \delta \boldsymbol{s}) \\ &\quad + \zeta(u, \delta \lambda) - (\boldsymbol{s}, \nabla \delta \lambda) \\ &\quad - (\nabla u, \delta \boldsymbol{\mu}) + (\boldsymbol{e}, \delta \boldsymbol{\mu}) \\ &= (\tilde{\boldsymbol{e}}, \delta \boldsymbol{e}) + \kappa(\tilde{\boldsymbol{s}}, \delta \boldsymbol{s}) + (q, \delta \lambda) =: L(\delta \boldsymbol{x}). \end{aligned}$$

Note that the space E acts as pivot space in this primal formulation.

Definition 1. Let the space Y_p be defined as

$$Y_p = \{\boldsymbol{y} = (u, \boldsymbol{e}, \boldsymbol{s}) \in H_0^1(\Omega) \times [L^2(\Omega)]^{d_\Omega} \times [L^2(\Omega)]^{d_\Omega}\}$$

and let us equip Y_p with the norm

$$\|\boldsymbol{y}\|_{Y_p} = \sqrt{\frac{1}{\ell_\Omega^2} \|u\|^2 + \|\nabla u\|^2 + \|\boldsymbol{e}\|^2 + \kappa \|\boldsymbol{s}\|^2}.$$

In addition, let the space Z_p be defined as

$$Z_p = \{\boldsymbol{z} = (\lambda, \boldsymbol{\mu}) \in H_0^1(\Omega) \times [L^2(\Omega)]^{d_\Omega}\}$$

and let us equip Z_p with the norm

$$\|z\|_{Z_p} = \sqrt{\frac{1}{\kappa \ell_\Omega^2} \|\lambda\|^2 + \frac{1}{\kappa} \|\nabla \lambda\|^2 + \|\mu\|^2}.$$

Finally, let us define $a_p : Y_p \times Y_p \rightarrow \mathbb{R}$ and $b_p : Y_p \times Z_p \rightarrow \mathbb{R}$ as

$$a_p(\mathbf{y}_1, \mathbf{y}_2) = (\mathbf{e}_1, \mathbf{e}_2) + \kappa(\mathbf{s}_1, \mathbf{s}_2) \quad (4)$$

$$b_p(\mathbf{y}, \mathbf{z}) = -(\nabla \lambda, \mathbf{s}) + \zeta(\lambda, u) + (\mu, \mathbf{e} - \nabla u). \quad (5)$$

Problem 2 (P_p). Find $\mathbf{x} = (\mathbf{y}, \mathbf{z}) \in X_p = Y_p \times Z_p$ such that

$$B_p(\mathbf{x}, \delta \mathbf{x}) = L(\delta \mathbf{x}), \quad \forall \delta \mathbf{x} = (\delta \mathbf{y}, \delta \mathbf{z}) \in X_p, \quad (6)$$

or, equivalently,

$$\begin{aligned} a_p(\mathbf{y}, \delta \mathbf{y}) + b_p(\delta \mathbf{y}, \mathbf{z}) &= (\tilde{\mathbf{e}}, \delta \mathbf{e}) + \kappa(\tilde{\mathbf{s}}, \delta \mathbf{s}) & \forall \delta \mathbf{y} \in Y_p, \\ b_p(\mathbf{y}, \delta \mathbf{z}) &= (q, \delta \lambda) & \forall \delta \mathbf{z} \in Z_p. \end{aligned}$$

Theorem 1. For a dimensionless physical constant $\kappa \zeta^2 \ell_\Omega^4 \in (0, \infty)$, problem P_p is well posed.

Proof. The bilinear and linear forms are continuous in X_p . From [4, Thm. 3.1], the necessary requirements on a_p and b_p to have a well-posed mixed saddle-point problem are met. Finally, the solution continuously depends on $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}}, q)$. $\square$

2.3 Dual form

The dual formulation of the current problem is given then by the following equations:

$$\left\{ \begin{array}{ll} \zeta \lambda + \nabla \cdot \mu &= 0 \quad \text{in } U^* \simeq U = L_0^2(\Omega), \\ \mathbf{e} + \mu &= \tilde{\mathbf{e}} \quad \text{in } E^* \simeq E = [L^2(\Omega)]^{d_\Omega}, \\ \kappa \mathbf{s} - \nabla \lambda &= \kappa \tilde{\mathbf{s}} \quad \text{in } S^*, S = H_0(\text{div}, \Omega), \\ \zeta u + \nabla \cdot \mathbf{s} &= q \quad \text{in } L^* \simeq L = L_0^2(\Omega), \\ -\nabla u + \mathbf{e} &= \mathbf{0} \quad \text{in } M^*, M = H_0(\text{div}, \Omega), \end{array} \right.$$

where $H_0(\text{div}, \Omega)$ is the space of functions in $[L^2(\Omega)]^{d_\Omega}$ with divergence in $L^2(\Omega)$ and vanishing normal trace on $\partial\Omega$.

Since we consider homogeneous boundary conditions for the normal components of $\mathbf{s}$ and μ , we have considered zero mean for u and λ , i.e.,

$$\sqrt{\kappa} U = L = L_0^2(\Omega) = \{v \in L^2(\Omega) : (v, 1) = 0\}.$$

That $(\lambda, 1) = 0$ follows directly from the first equation of the system to be solved, whereas that $(u, 1) = 0$ is a consequence of the fourth equation and the compatibility condition $(q, 1) = 0$ typically associated to the problem with zero prescribed fluxes. Obviously, the problem with zero fluxes does not correspond to the primal problem with zero prescribed primal variables on the boundary. Non-homogeneous boundary conditions for both problems could be analysed using standard lifting arguments. In fact, in the numerical examples non-homogeneous boundary conditions will be used.

By taking $M = E$, $U = \sqrt{\kappa} \ell_\Omega^2 \nabla \cdot S$, and $L = \sqrt{\kappa} \ell_\Omega^2 \nabla \cdot M$, we arrive at

$$\left\{ \begin{array}{ll} \zeta \mu - \nabla(\nabla \cdot \mathbf{s}) &= \zeta \tilde{\mathbf{e}} - \nabla q \quad \text{in } S^*, S = H_0(\text{div}, \Omega), \\ -\zeta \kappa \mathbf{s} - \nabla(\nabla \cdot \mu) &= -\zeta \kappa \tilde{\mathbf{s}} \quad \text{in } M^*, M = H_0(\text{div}, \Omega). \end{array} \right. \quad (7)$$

Provided higher regularity in the function spaces involved and the domain, we have the reduced strong form

$$\zeta^2 \kappa \mathbf{s} + \nabla(\nabla \cdot (\nabla(\nabla \cdot \mathbf{s}))) = -\zeta \nabla(\nabla \cdot \tilde{\mathbf{e}}) + \zeta^2 \kappa \nabla \cdot \tilde{\mathbf{s}} + \nabla(\nabla \cdot (\nabla q)) \quad \text{in } [L^2(\Omega)]^{d_\Omega},$$

where we may take

$$\mathbf{s} \in [H^4(\Omega)]^{d_\Omega}, \quad \tilde{\mathbf{e}} \in [H^2(\Omega)]^{d_\Omega}, \quad \tilde{\mathbf{s}} \in [H^1(\Omega)]^{d_\Omega} \quad \text{and} \quad q \in H^3(\Omega),$$

together with the boundary conditions

$$\nabla \cdot (\nabla(\nabla \cdot \mathbf{s})) = 0 \quad \text{and} \quad (\nabla(\nabla \cdot \mathbf{s})) \cdot \hat{\mathbf{n}} = 0$$

almost everywhere on all $\partial\Omega$, which is a well-posed elliptic problem. In matrix-vector representation, problem (7) takes the form

$$-\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{pmatrix} \nabla(\nabla \cdot \mathbf{s}) \\ \nabla(\nabla \cdot \boldsymbol{\mu}) \end{pmatrix} + \begin{bmatrix} 0 & \zeta I \\ -\zeta \kappa I & 0 \end{bmatrix} \begin{pmatrix} \mathbf{s} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \zeta \tilde{\mathbf{e}} - \nabla q \\ -\zeta \kappa \tilde{\mathbf{s}} \end{pmatrix}.$$

The eigenvalues of the “reaction” matrix are $\pm i \sqrt{\kappa} \zeta$, showing that the problem cannot be uncoupled (for $\zeta \neq 0$) and behaves as a fourth order problem. For the dual formulation, the fields u and λ are unnecessary. However, for the primal formulation, keeping them makes sense as seen above.

Finally, the governing equations can be expressed in their weak form as

$$\begin{aligned} B_d(\mathbf{x}, \delta \mathbf{x}) &:= \zeta(\lambda, \delta u) + (\nabla \cdot \boldsymbol{\mu}, \delta u) \\ &\quad + (\mathbf{e}, \delta \mathbf{e}) + (\boldsymbol{\mu}, \delta \mathbf{e}) \\ &\quad + \kappa(\mathbf{s}, \delta \mathbf{s}) + (\lambda, \nabla \cdot \delta \mathbf{s}) \\ &\quad + \zeta(u, \delta \lambda) + (\nabla \cdot \mathbf{s}, \delta \lambda) \\ &\quad + (u, \nabla \cdot \delta \boldsymbol{\mu}) + (\mathbf{e}, \delta \boldsymbol{\mu}) \\ &= (\tilde{\mathbf{e}}, \delta \mathbf{e}) + \kappa(\tilde{\mathbf{s}}, \delta \mathbf{s}) + (q, \delta \lambda) =: L(\delta \mathbf{x}). \end{aligned}$$

In principle, one could consider $E = H_0(\text{div}, \Omega)$, but as we shall see there is no way to control the divergence of $\mathbf{e}$. Thus, keeping $\mathbf{e}$ and $\boldsymbol{\mu}$ as unknowns simultaneously makes sense. Again, note that the space E acts as pivot space in this dual formulation.

Definition 2. Let the space Y_d be defined as

$$Y_d = \{\mathbf{y} = (u, \mathbf{e}, \mathbf{s}) \in L_0^2(\Omega) \times [L^2(\Omega)]^{d_\Omega} \times H_0(\text{div}, \Omega)\}$$

and let us equip Y_d with the norm

$$\|\mathbf{y}\|_{Y_d} = \sqrt{\frac{1}{\ell_\Omega^2} \|u\|^2 + \|\mathbf{e}\|^2 + \kappa \|\mathbf{s}\|^2 + \kappa \ell_\Omega^2 \|\nabla \cdot \mathbf{s}\|^2}.$$

In addition, let the space Z_d be defined as

$$Z_d = \{\mathbf{z} = (\lambda, \boldsymbol{\mu}) \in L_0^2(\Omega) \times H_0(\text{div}, \Omega)\}$$

and let us equip Z_d with the norm

$$\|\mathbf{z}\|_{Z_d} = \sqrt{\frac{1}{\kappa \ell_\Omega^2} \|\lambda\|^2 + \|\boldsymbol{\mu}\|^2 + \ell_\Omega^2 \|\nabla \cdot \boldsymbol{\mu}\|^2}.$$

Finally, let us define $a_d : Y_d \times Y_d \rightarrow \mathbb{R}$ and $b_d : Y_d \times Z_d \rightarrow \mathbb{R}$ as

$$a_d(\mathbf{y}_1, \mathbf{y}_2) = (\mathbf{e}_1, \mathbf{e}_2) + \kappa(\mathbf{s}_1, \mathbf{s}_2) \tag{8}$$

$$b_d(\mathbf{y}, \mathbf{z}) = (\lambda, \nabla \cdot \mathbf{s} + \zeta u) + (\boldsymbol{\mu}, \mathbf{e}) + (\nabla \cdot \boldsymbol{\mu}, u). \tag{9}$$

Problem 3 (P_d). Find $\mathbf{x} = (\mathbf{y}, \mathbf{z}) \in X_d = Y_d \times Z_d$ such that

$$B_d(\mathbf{x}, \delta \mathbf{x}) = L(\delta \mathbf{x}), \quad \forall \delta \mathbf{x} = (\delta \mathbf{x}, \delta \mathbf{y}) \in X_d, \quad (10)$$

or, equivalently,

$$\begin{aligned} a_d(\mathbf{y}, \delta \mathbf{y}) + b_d(\delta \mathbf{y}, \mathbf{z}) &= (\tilde{\mathbf{e}}, \delta \mathbf{e}) + \kappa(\tilde{\mathbf{s}}, \delta \mathbf{s}) & \forall \delta \mathbf{y} \in Y_d, \\ b_d(\mathbf{y}, \delta \mathbf{z}) &= (q, \delta \lambda) & \forall \delta \mathbf{z} \in Z_d. \end{aligned}$$

Theorem 2. For a dimensionless physical constant $\kappa \zeta^2 \ell_\Omega^4 \in (0, 2)$, problem P_d is well posed.

Proof. The bilinear forms in (8)-(9) and the linear form L are continuous. Moreover, from (8) we have that

$$a_d(\mathbf{y}, \mathbf{y}) = \|\mathbf{e}\|^2 + \kappa \|\mathbf{s}\|^2.$$

Let us proof the inf-sup condition for (9). Let us consider $(u_\mu, \mathbf{s}_\lambda, \mathbf{e}_\mu)$ such that

$$u_\mu = \ell_\Omega^2 \nabla \cdot \boldsymbol{\mu}, \quad \nabla \cdot \mathbf{s}_\lambda = \frac{1}{\kappa \ell_\Omega^2} \lambda \quad (\text{no } \mathbf{s}_\lambda \cdot \hat{\mathbf{n}} \text{ prescribed on } \partial\Omega) \quad \text{and} \quad \mathbf{e}_\mu = \boldsymbol{\mu}.$$

Since the divergence operator $\nabla \cdot : H_0(\text{div}, \Omega) \rightarrow L_0^2(\Omega)$ is surjective, i.e., its closed image is precisely the entire target space or equivalently $\text{Im}(\nabla \cdot) = \{\nabla \cdot \mathbf{s} : \mathbf{s} \in H_0(\text{div}, \Omega)\} = L_0^2(\Omega)$, the existence of $\mathbf{s}_\lambda$ is warranted. Then, by inserting $\mathbf{y}_z = (u_\mu, \mathbf{e}_\mu, \mathbf{s}_\lambda)$ in $b_d(\mathbf{y}, \mathbf{z})$ we obtain

$$b_d(\mathbf{y}_z, \mathbf{z}) = \frac{1}{\kappa \ell_\Omega^2} \|\lambda\|^2 + \|\boldsymbol{\mu}\|^2 + \zeta(\lambda, u_\mu) + \ell_\Omega^2 \|\nabla \cdot \boldsymbol{\mu}\|^2.$$

Now, we use Young's inequality as follows

$$\zeta(\lambda, u_\mu) = \ell_\Omega^2 (\zeta \lambda, \nabla \cdot \boldsymbol{\mu}) \geq -\frac{\ell_\Omega^2 \zeta^2}{2} \|\lambda\|^2 - \frac{\ell_\Omega^2}{2} \|\nabla \cdot \boldsymbol{\mu}\|^2$$

to arrive at

$$b_d(\mathbf{y}_z, \mathbf{z}) \geq \frac{2 - \kappa \zeta^2 \ell_\Omega^4}{2 \kappa \ell_\Omega^2} \|\lambda\|^2 + \|\boldsymbol{\mu}\|^2 + \frac{\ell_\Omega^2}{2} \|\nabla \cdot \boldsymbol{\mu}\|^2$$

and therefore,

$$\inf_{(\lambda, \boldsymbol{\mu}) \in Z_d} \sup_{(u, \mathbf{e}, \mathbf{s}) \in Y_d} \frac{(\lambda, \nabla \cdot \mathbf{s} + \zeta u) + (\boldsymbol{\mu}, \mathbf{e}) + (\nabla \cdot \boldsymbol{\mu}, u)}{\|(\lambda, \boldsymbol{\mu})\|_{Z_d} \|(u, \mathbf{e}, \mathbf{s})\|_{Y_d}} \geq k_{b_d}$$

with

$$k_{b_d} = \min \left\{ \frac{2 - \kappa \zeta^2 \ell_\Omega^4}{2}, \frac{1}{2} \right\} > 0, \quad \forall \kappa \in \left(0, \frac{2}{\zeta^2 \ell_\Omega^4} \right).$$

Let us consider now the coercivity of a_d in the kernel of b_d . Let

$$\text{Ker}(b_d(\cdot, \mathbf{z})) = \{\mathbf{y} = (u, \mathbf{e}, \mathbf{s}) \in Y_d : b_d(\mathbf{y}, \mathbf{z}) = 0 \quad \forall \mathbf{z} \in Z_d\}. \quad (11)$$

Let us pick $\mathbf{y}_\infty = (u_\infty, \mathbf{e}_\infty, \mathbf{s}_\infty)$, with $u_\infty \in C_0^\infty(\Omega)$. Then, $\mathbf{y}_\infty \in \text{Ker}(b_d(\cdot, \mathbf{z}))$ if $\nabla \cdot \mathbf{s}_\infty = -\zeta u_\infty$, $\mathbf{e}_\infty = \nabla u_\infty$, $u_\infty = 0$ on $\partial\Omega$. Using Poincaré's inequality, i.e.,

$$\|\nabla u_\infty\|^2 \geq \frac{1}{2} \|\nabla u_\infty\|^2 + \frac{1}{2C_{\text{PF}}^2 \ell_\Omega^2} \|u_\infty\|^2,$$

with C_{PF} a dimensionless constant, we arrive at

$$a_d(\mathbf{y}_\infty, \mathbf{y}_\infty) = \|\mathbf{e}_\infty\|^2 + \kappa \|\mathbf{s}_\infty\|^2 \geq \frac{1}{8C_{\text{PF}}^2 \ell_\Omega^2} \|u_\infty\|^2 + \frac{1}{2} \|\mathbf{e}_\infty\|^2 + \kappa \|\mathbf{s}_\infty\|^2 + \frac{1}{8\zeta^2 C_{\text{PF}}^2 \ell_\Omega^2} \|\nabla \cdot \mathbf{s}_\infty\|^2,$$

or in a more compact form

$$a_d(\mathbf{y}_\infty, \mathbf{y}_\infty) \geq k_{a_d} \|(\mathbf{u}_\infty, \mathbf{e}_\infty, \mathbf{s}_\infty)\|_{Y_d}^2, \quad (12)$$

with

$$k_{a_d} = \min \left\{ \frac{1}{8C_{PF}^2}, \frac{1}{2}, \frac{1}{8C_{PF}^2 \kappa \zeta^2 \ell_\Omega^4} \right\} > 0, \quad \forall \frac{1}{\kappa} \in \left(\frac{\zeta^2 \ell_\Omega^4}{2}, \infty \right),$$

which is consistent with the range of κ required to satisfy the previous inf-sup condition. Recall that we consider $\kappa \zeta^2 \ell_\Omega^4 > 0$ (cf. Remark 3).

The continuity of b_d and a standard compactness argument allows us to establish that (12) holds on the whole space in (11). We can conclude that

$$k_{a_d} = \min \left\{ \frac{1}{2}, \frac{1}{16C_{PF}^2} \right\} > 0.$$

Then, all the necessary requirements on a_d and b_d to have a well-posed mixed saddle-point problem are met.

Finally, the solution continuously depends on $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}}, q)$. $\square$

Remark 5. *The dual formulation imposes a more restrictive admissible range for the dimensionless physical constant $(\kappa \zeta^2 \ell_\Omega^4)$, reflecting the intrinsic trade-off of the primal-dual transition: reducing regularity requirements on the scalar fields in the primal problem enhances the regularity of the vector fields in the dual one, but at the cost of narrowing the permissible scaling parameter. Since the bounds arising from the corresponding inf-sup conditions are not sharp, we can only guarantee well-posedness within the stated ranges, although the actual interval of admissible values could be wider. The constant κ ensures unit consistency and provides a useful scaling mechanism for real data and numerical simulations. In practice, we recommend computing κ via either the geometric or the arithmetic mean and verifying that it falls within the safe range identified by the theory. Overall, both the primal and dual formulations perform robustly across different noise levels when parameters are chosen within these admissible ranges.*

3 Finite element approximation

For the spatial discretization of both the primal and dual problems, we employ in the following the finite element method. Let us consider then a finite element partition $\mathcal{P}_h = \{K\}$ of the domain Ω . Our objective is to consider the simplest possible setting, avoiding technicalities and focusing on the key difficulties of the problem. Thus, we shall only consider *quasi-uniform* finite element partitions of diameter h . Furthermore, the finite element spaces to be constructed from $\mathcal{P}_h$ will be considered to be made of $C^0(\Omega)$ functions, although the extension to spaces of discontinuous functions is not difficult. All finite element spaces in the following are based on $\mathcal{P}_h$, although their order can be different. We shall denote the finite element approximation to spaces U , E , S , L and M as U_h , E_h , S_h , L_h and M_h , respectively.

As usual, finite element functions and spaces will be indicated with the subscript h . This subscript will be used also for the broken $L^2(\Omega)$ -inner product and norm, respectively denoted by $(\cdot, \cdot)_h$ and $\|\cdot\|_h$.

We will make use of the inverse inequality

$$\|\nabla v_h\|_h \leq \frac{C_{\text{inv}}}{h} \|v_h\|_h, \quad (13)$$

where v_h is any finite element function and C_{inv} a dimensionless constant.

The problem to be approximated involves five variables. If the standard Galerkin method is used, the associated finite element spaces need to satisfy the global inf-sup condition, which

translates into different ('little') inf-sup conditions between the spaces in play. To achieve stable formulations that are independent of the particular finite element spaces chosen, thus avoiding these inf-sup conditions, we are required to include stabilization terms. To this end, we rely on the VMS framework to systematically accomplish this task. The key idea is to split the unknown as $\mathbf{x} = \mathbf{x}_h + \mathbf{x}'$, where $\mathbf{x}_h$ is the finite element solution to be computed and $\mathbf{x}' = (u', \mathbf{e}', \mathbf{s}', \lambda', \boldsymbol{\mu}')$ is the sub-grid scale, which cannot be captured and needs to be approximated in terms of $\mathbf{x}_h$. We shall not describe the rationale to model the sub-grid scales (see for example [10]). The model we propose is

$$\begin{cases} u' &= \tau_u P'(-\zeta \lambda_h - \nabla \cdot \boldsymbol{\mu}_h), \\ \mathbf{e}' &= \tau_e P'(\tilde{\mathbf{e}} - \mathbf{e}_h - \boldsymbol{\mu}_h), \\ \mathbf{s}' &= \tau_s P'(\kappa \tilde{\mathbf{s}} - \kappa \mathbf{s}_h + \nabla \lambda_h), \\ \lambda' &= \tau_\lambda P'(q - \zeta u_h - \nabla \cdot \mathbf{s}_h), \\ \boldsymbol{\mu}' &= \tau_\mu P'(\nabla u_h - \mathbf{e}_h), \end{cases} \quad (14)$$

where $\tau_u, \tau_e, \tau_s, \tau_\lambda$ and τ_μ are the stabilization parameters defined later and P' is a projection operator. Two canonical choices are: *i*) P' being the identity in $L^2(\Omega)$, leading to the ASGS method; and, *ii*) P' being the projection on the orthogonal complement (with respect to the $L^2(\Omega)$ -norm) of the finite element space, leading to the OSGS formulation.

Following the VMS approach, we can systematically derive a stabilized bilinear form of the problem associated to the projection P' , which will be of the form:

$$B'_{\text{stab}}(\mathbf{x}_h, \delta \mathbf{x}_h) := B(\mathbf{x}_h, \delta \mathbf{x}_h) + S'_h(\mathbf{x}_h, \delta \mathbf{x}_h),$$

where $B(\mathbf{x}_h, \delta \mathbf{x}_h)$ is the original bilinear form associated to the continuous problem and $S'_h(\mathbf{x}_h, \delta \mathbf{x}_h)$ is the stabilization that depends, in principle, on the problem, i.e., either primal or dual form, and the particular operator P' selected. We shall write it as $S'_h = S_h$ when $P' = I$ and as $S'_h = S_h^\perp$ when $P' = P^\perp = I - P_h$, where P_h is the projection onto the appropriate finite element space. Appropriate subscripts will be used for the primal and the dual problems. To simplify the notation, we will not specify the space onto which P' is applied, being it clear by the context.

Likewise, the linear form of the right-hand-side will also be modified to $L_{\text{stab}}(\delta \mathbf{x}_h)$. This modification has the same expression for the primal and the dual problems, and is given by

$$\begin{aligned} L'_{\text{stab}}(\delta \mathbf{x}_h) &= L(\delta \mathbf{x}_h) - (\tau_\lambda P'(q), \zeta \delta u_h + \nabla \cdot \delta \mathbf{s}_h) - (\tau_e P'(\tilde{\mathbf{e}}), \delta \mathbf{e}_h + \delta \boldsymbol{\mu}_h) \\ &\quad - (\tau_s P'(\kappa \tilde{\mathbf{s}}), \kappa \delta \mathbf{s}_h - \nabla \delta \lambda_h). \end{aligned}$$

Remark 6. When applied to continuous finite element functions it is easily checked that

$$B_p(\mathbf{x}_h, \delta \mathbf{x}_h) = B_d(\mathbf{x}_h, \delta \mathbf{x}_h).$$

Therefore, the difference between the primal and the dual formulation is only in the expression of $S(\mathbf{x}_h, \delta \mathbf{x}_h)$ and, since the element residuals are identical, in the stabilization parameters. In other words, we will be able to move from the primal to the dual formulations only changing the design of the stabilization parameters.

3.1 Discretized primal problem

The stabilized bilinear form for the discrete primal formulation can be written as

$$B'_{\text{stab,p}}(\mathbf{x}_h, \delta \mathbf{x}_h) := B_p(\mathbf{x}_h, \delta \mathbf{x}_h) + S'_{p,h}(\mathbf{x}_h, \delta \mathbf{x}_h),$$

where $S'_{p,h}(\mathbf{x}_h, \delta \mathbf{x}_h)$ is, in principle, a mesh-dependent term providing the desired stability.

In addition, let us consider:

$$\tau_u = k_u h^2, \quad \tau_e = k_e, \quad \tau_s = \frac{k_s}{\kappa}, \quad \tau_\lambda = -\frac{k_\lambda h^2}{\zeta^2 \ell_\Omega^4}, \quad \tau_\mu = -k_\mu,$$

where k_u , k_e , k_s , k_λ and k_μ are algorithmic (dimensionless) constants. Then, the stabilization term is as follows:

$$\begin{aligned}
S'_{p,h}(\mathbf{x}_h, \delta \mathbf{x}_h) &:= -k_u h^2 [\zeta^2(P' \lambda_h, \delta \lambda_h) + \zeta((P' \nabla \cdot \boldsymbol{\mu}_h, \delta \lambda_h)_h + (P' \lambda_h, \nabla \cdot \delta \boldsymbol{\mu}_h)_h) + (P' \nabla \cdot \boldsymbol{\mu}_h, \nabla \cdot \delta \boldsymbol{\mu}_h)_h] \\
&\quad - k_e [(P' \mathbf{e}_h, \delta \mathbf{e}_h) + (P' \boldsymbol{\mu}_h, \delta \mathbf{e}_h) + (P' \mathbf{e}_h, \delta \boldsymbol{\mu}_h) + (P' \boldsymbol{\mu}_h, \delta \boldsymbol{\mu}_h)] \\
&\quad - \frac{k_s}{\kappa} [(P' \nabla \lambda_h, \nabla \delta \lambda_h) - \kappa(P' \mathbf{s}_h, \nabla \delta \lambda_h) - \kappa(P' \nabla \lambda_h, \delta \mathbf{s}_h) + \kappa^2(P' \mathbf{s}_h, \delta \mathbf{s}_h)] \\
&\quad + \frac{k_\lambda h^2}{\zeta^2 \ell_\Omega^4} [\zeta^2(P' u_h, \delta u_h) + \zeta((P' \nabla \cdot \mathbf{s}_h, \delta u_h)_h + (P' u_h, \nabla \cdot \delta \mathbf{s}_h)_h) + (P' \nabla \cdot \mathbf{s}_h, \nabla \cdot \delta \mathbf{s}_h)_h] \\
&\quad + k_\mu [(P' \nabla u_h, \nabla \delta u_h) - (P' \mathbf{e}_h, \nabla \delta u_h) - (P' \nabla u_h, \delta \mathbf{e}_h) + (P' \mathbf{e}_h, \delta \mathbf{e}_h)].
\end{aligned}$$

Note that for the continuous finite element interpolations we are considering, we could in fact replace $(\cdot, \cdot)_h$ by $(\cdot, \cdot)$.

Particularity, for $P' = P^\perp$ we have that

$$\begin{aligned}
S_{p,h}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h) &= k_\mu (P^\perp(\nabla u_h), P^\perp(\nabla \delta u_h)) \\
&\quad + \frac{k_\lambda h^2}{\zeta^2 \ell_\Omega^4} (P^\perp(\nabla \cdot \mathbf{s}_h), P^\perp(\nabla \cdot \delta \mathbf{s}_h))_h \\
&\quad - \frac{k_s}{\kappa} (P^\perp(\nabla \lambda_h), P^\perp(\nabla \delta \lambda_h)) \\
&\quad - k_u h^2 (P^\perp(\nabla \cdot \boldsymbol{\mu}_h), P^\perp(\nabla \cdot \delta \boldsymbol{\mu}_h))_h.
\end{aligned}$$

Problem 4 ($P_{p,h}$). Find $\mathbf{x}_h \in X_h$ such that

$$B'_{\text{stab},p}(\mathbf{x}_h, \delta \mathbf{x}_h) = L'_{\text{stab}}(\delta \mathbf{x}_h), \quad \forall \delta \mathbf{x}_h \in X_{p,h}. \quad (15)$$

Definition 3. Let $X_{p,h}$ be defined as

$$X_{p,h} = U_h \times E_h \times S_h \times L_h \times M_h$$

and equipped with the norm defined by

$$\begin{aligned}
\|\mathbf{x}_h\|_{X_{p,h}}^2 &= \frac{1}{\ell_\Omega^2} \|u_h\|^2 + \|\nabla u_h\|^2 + \|\mathbf{e}_h\|^2 + \kappa \|\mathbf{s}_h\|^2 + \kappa h^2 \|\nabla \cdot \mathbf{s}_h\|_h^2 \\
&\quad + \frac{1}{\kappa \ell_\Omega^2} \|\lambda_h\|^2 + \frac{1}{\kappa} \|\nabla \lambda_h\|^2 + \|\boldsymbol{\mu}_h\|^2 + h^2 \|\nabla \cdot \boldsymbol{\mu}_h\|_h^2.
\end{aligned}$$

Theorem 3. The discretized primal problem $P_{p,h}$ is well posed for P' being either the identity (ASGS) or the $L^2(\Omega)$ -projection on the orthogonal complement of the finite element space chosen (OSGS). In addition, it consistently approximates the continuous primal problem P_p .

Proof. The continuity of $B'_{\text{stab},p}(\mathbf{x}_h, \delta \mathbf{x}_h)$ and $L'_{\text{stab}}(\delta \mathbf{x}_h)$ in the norm of $X_{p,h}$ is straightforward. For the inf-sup stability of $B_{\text{stab},p}(\mathbf{x}_h, \delta \mathbf{x}_h)$, let P' be I and set $\delta \mathbf{x}_h^0 = (u_h, \mathbf{e}_h, \mathbf{s}_h, -\lambda_h, -\boldsymbol{\mu}_h)$. In addition, let us consider the following inequalities:

$$\begin{aligned}
-2k_\mu(\mathbf{e}_h, \nabla u_h) &\geq -\frac{1}{2} \|\mathbf{e}_h\|^2 - 4k_\mu^2 \|\nabla u_h\|^2, \quad (\text{Young}), \\
\frac{2}{\zeta} (\nabla \cdot \mathbf{s}_h, u_h)_h &\geq -\frac{k_1}{\zeta^2} \|\nabla \cdot \mathbf{s}_h\|_h^2 - \frac{1}{k_1} \|u_h\|^2, \quad \forall k_1 > 0, \quad (\text{Young}), \\
2\zeta (\nabla \cdot \boldsymbol{\mu}_h, \lambda_h)_h &\geq -k_2 \|\nabla \cdot \boldsymbol{\mu}_h\|_h^2 - \frac{\zeta^2}{k_2} \|\lambda_h\|^2, \quad \forall k_2 > 0, \quad (\text{Young}), \\
\|\nabla u_h\|^2 &\geq k_3 \|\nabla u_h\|^2 + \frac{(1-k_3)}{C_{\text{PF}}^2 \ell_\Omega^2} \|u_h\|^2, \quad \forall k_3 \in (0, 1), \quad (\text{Poincaré}), \\
\|\nabla \lambda_h\|^2 &\geq k_4 \|\nabla \lambda_h\|^2 + \frac{(1-k_4)}{C_{\text{PF}}^2 \ell_\Omega^2} \|\lambda_h\|^2, \quad \forall k_4 \in (0, 1), \quad (\text{Poincaré}).
\end{aligned}$$

After some algebraic manipulations, we arrive at

$$B_{\text{stab,p}}(\mathbf{x}_h, \delta \mathbf{x}_h^0) \geq \tau_1 \|u_h\|^2 + \tau_2 \|\mathbf{e}_h\|^2 + \tau_3 \|\mathbf{s}_h\|^2 + \tau_4 \|\lambda_h\|^2 + \tau_5 \|\boldsymbol{\mu}_h\|^2 + \tau_6 \|\nabla u_h\|^2 + \tau_7 \|\nabla \cdot \mathbf{s}_h\|_h^2 + \tau_8 \|\nabla \lambda_h\|^2 + \tau_9 \|\nabla \cdot \boldsymbol{\mu}_h\|_h^2,$$

where

$$\begin{aligned} \tau_1 &= \left(1 - \frac{1}{k_1}\right) \frac{k_\lambda h^2}{\ell_\Omega^4} + \frac{1 - k_3}{C_{\text{PF}}^2 \ell_\Omega^2} k_\mu (1 - 4k_\mu), \\ \tau_2 &= k_\mu - k_e + \frac{3}{4}, \\ \tau_3 &= \kappa(1 - k_s), \\ \tau_4 &= \left(1 - \frac{1}{k_2}\right) k_u h^2 \zeta^2 + \frac{1 - k_4}{C_{\text{PF}}^2 \ell_\Omega^2} \frac{k_s}{\kappa}, \\ \tau_5 &= k_e, \\ \tau_6 &= k_3 k_\mu (1 - 4k_\mu), \\ \tau_7 &= (1 - k_1) \frac{k_\lambda h^2}{\ell_\Omega^4 \zeta^2}, \\ \tau_8 &= k_4 \frac{k_s}{\kappa}, \\ \tau_9 &= (1 - k_2) k_u h^2. \end{aligned}$$

Expression $B_{\text{stab,p}}(\mathbf{x}_h, \delta \mathbf{x}_h^0)$ is positive provided that

$$\begin{aligned} k_u &\geq 0, k_e \in (0, k_e^*), k_s \in (0, 1), k_\lambda \geq 0, k_\mu \in \left(0, \frac{1}{4}\right), \\ k_1 &\in (k_1^*, 1], \quad k_2 \in (k_2^*, 1], \quad k_3 \in (0, 1), \quad k_4 \in (0, 1), \end{aligned}$$

where

$$k_e^* = k_\mu + \frac{3}{4}, \quad k_1^* = \frac{1}{1 + \frac{k_\mu(1-4k_\mu)}{k_\lambda} \frac{(1-k_3)}{C_{\text{PF}}^2} \frac{\ell_\Omega^2}{h^2}}, \quad k_2^* = \frac{1}{1 + \frac{k_s}{k_u} \frac{(1-k_4)}{C_{\text{PF}}^2} \frac{1}{\kappa h^2 \zeta^2 \ell_\Omega^2}}.$$

For $P' = P^\perp$, i.e., the projector onto the $L^2(\Omega)$ -orthogonal complement of the chosen particular finite element space, we have

$$\begin{aligned} B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^0) &\geq \|\mathbf{e}_h\|^2 + \kappa \|\mathbf{s}_h\|^2 + \\ &k_\mu \|P^\perp(\nabla u_h)\|^2 + \frac{k_\lambda h^2}{\zeta^2 \ell_\Omega^4} \|P^\perp(\nabla \cdot \mathbf{s}_h)\|_h^2 + \frac{k_s}{\kappa} \|P^\perp(\nabla \lambda_h)\|^2 + k_u h^2 \|P^\perp(\nabla \cdot \boldsymbol{\mu}_h)\|_h^2. \end{aligned}$$

Full control on ∇u_h and $\nabla \lambda_h$ can be obtained using the argument briefly summarized next. The task we need to accomplish is to get control over $\|P(\nabla u_h)\|$ and $\|P(\nabla \lambda_h)\|$. For such an end, let us take $\delta \mathbf{x}_h^1 = (0, \mathbf{0}, -k_s \kappa^{-1} P(\nabla \lambda_h), \mathbf{0}, -k_\mu P(\nabla u_h))$, which yields

$$\begin{aligned} B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^1) &= +k_\mu \|P(\nabla u_h)\|^2 + \frac{k_s}{\kappa} \|P(\nabla \lambda_h)\|^2 \\ &\quad - k_\mu (\mathbf{e}_h, P(\nabla u_h)) - k_s (\mathbf{s}_h, P(\nabla \lambda_h)) \\ &\quad + k_u k_\mu h^2 (P^\perp(\nabla \cdot \boldsymbol{\mu}_h), \nabla \cdot (P(\nabla u_h)))_h \\ &\quad - \frac{k_s k_\lambda h^2}{\kappa \zeta^2 \ell_\Omega^4} (P^\perp(\nabla \cdot \mathbf{s}_h), \nabla \cdot (P(\nabla \lambda_h)))_h. \end{aligned}$$

By employing Young's inequality for the third and fourth terms and by combining Young's and inverse inequalities for the fifth and sixth terms, namely

$$+k_u k_\mu h^2 (P^\perp(\nabla \cdot \boldsymbol{\mu}_h), \nabla \cdot (P(\nabla u_h)))_h \geq -\frac{k_7 C_{\text{inv}}^2 h^2 k_u^2 k_\mu}{2} \|P^\perp(\nabla \cdot \boldsymbol{\mu}_h)\|_h^2 - \frac{k_\mu}{2k_7} \|P(\nabla u_h)\|^2$$

and

$$-\frac{k_s k_\lambda h^2}{\kappa \zeta^2 \ell_\Omega^4} (P^\perp(\nabla \cdot \mathbf{s}_h), \nabla \cdot (P(\nabla \lambda_h)))_h \geq -\frac{k_8 C_{\text{inv}}^2 h^2 k_\lambda^2 k_s}{2\zeta^8 \ell_\Omega^8 \kappa} \|P^\perp(\nabla \cdot \mathbf{s}_h)\|_h^2 - \frac{k_s}{2k_8 \kappa} \|P(\nabla \lambda_h)\|^2,$$

we arrive at

$$\begin{aligned} B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^1) &\geq + \left(1 - \frac{1}{2k_5} - \frac{1}{2k_7}\right) k_\mu \|P(\nabla u_h)\|^2 \\ &\quad + \left(1 - \frac{1}{2k_6} - \frac{1}{2k_8}\right) \frac{k_s}{\kappa} \|P(\nabla \lambda_h)\|^2 \\ &\quad - \frac{k_5 k_\mu}{2} \|\mathbf{e}_h\|^2 - \frac{k_6 k_s \kappa}{2} \|\mathbf{s}_h\|^2 \\ &\quad - \frac{k_7 k_u^2 k_\mu C_{\text{inv}}^2 h^2}{2} \|P^\perp(\nabla \cdot \boldsymbol{\mu}_h)\|_h^2 \\ &\quad - \frac{k_8 k_s k_\lambda^2 C_{\text{inv}}^2 h^2}{2\zeta^4 \ell_\Omega^8 \kappa} \|P^\perp(\nabla \cdot \mathbf{s}_h)\|_h^2 \\ &\gtrsim + \|P(\nabla u_h)\|^2 + \frac{1}{\kappa} \|P(\nabla \lambda_h)\|^2 \\ &\quad - \|\mathbf{e}_h\|^2 - \kappa \|\mathbf{s}_h\|^2 \\ &\quad - \kappa h^2 \|P^\perp(\nabla \cdot \mathbf{s}_h)\|_h^2 - h^2 \|P^\perp(\nabla \cdot \boldsymbol{\mu}_h)\|_h^2, \end{aligned}$$

where the control over $\|P(\nabla u_h)\|$ and $\|P(\nabla \lambda_h)\|$ is warranted for sufficiently large k_i (for all i in $\{5, 6, 7, 8\}$). The additional control over $\|\mathbf{e}_h\|$ and $\|\lambda_h\|$ is achieved by employing Poincaré's inequality. Finally, setting $\delta \mathbf{x}_h^\beta = \delta \mathbf{x}_h^0 + \beta_1 \delta \mathbf{x}_h^1$ and taking β_1 sufficiently small yields

$$B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^\beta) \gtrsim \|\mathbf{x}_h\|_{X_{p,h}}^2,$$

which holds true due to conjugation of the estimates for $B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^0)$ and $B_{\text{stab,p}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^1)$. Here and below, $\gtrsim$ stands for $\geq$ up to constants, and likewise for $\lesssim$. Then, the desired inf-sup condition follows after checking that $\|\mathbf{x}_h\|_{X_{p,h}} \gtrsim \|\delta \mathbf{x}_h^\beta\|_{X_{p,h}}$. $\square$

Proposition 1. *Let $k_u = k_\lambda = 0$, $P' = I$ and $M_h = E_h$. Then, the following residuum orthogonalities hold:*

$$\begin{aligned} (\kappa \mathbf{s} - \nabla \lambda_h - \kappa \tilde{\mathbf{s}}, \delta \mathbf{s}_h) &= 0 \quad (\text{orthogonality w.r.t. } S_h), \\ (\mathbf{e}_h + \boldsymbol{\mu}_h - \tilde{\mathbf{e}}, \delta \mathbf{e}_h) &= 0 \quad (\text{orthogonality w.r.t. } E_h), \\ (-\nabla u_h + \mathbf{e}_h, \delta \boldsymbol{\mu}_h) &= 0 \quad (\text{orthogonality w.r.t. } M_h). \end{aligned}$$

Proof. First, by setting $k_u = k_\lambda = 0$, it follows

$$\begin{aligned} (\kappa \mathbf{s} + \kappa \tau_s P'(\kappa \tilde{\mathbf{s}} - \kappa \mathbf{s}_h - \nabla \lambda_h) + \nabla \lambda_h - \kappa \tilde{\mathbf{s}}, \delta \mathbf{s}_h) &= 0, \\ (\mathbf{e}_h + \tau_e P'(\tilde{\mathbf{e}} - \mathbf{e}_h - \boldsymbol{\mu}_h) + \boldsymbol{\mu}_h + \tau_\mu P'(\nabla u_h - \mathbf{e}_h) - \tilde{\mathbf{e}}, \delta \mathbf{e}_h) &= 0 \end{aligned}$$

and

$$(-\nabla u_h + \mathbf{e}_h + \tau_e P'(\tilde{\mathbf{e}} - \mathbf{e}_h - \boldsymbol{\mu}_h), \delta \boldsymbol{\mu}_h) = 0.$$

Second, by setting $P' = I$, we have

$$(1 - \tau_s)(\kappa \mathbf{s} - \nabla \lambda_h - \kappa \tilde{\mathbf{s}}, \delta \mathbf{s}_h) = 0,$$

which is proposition's first orthogonality result,

$$((\tau_e - 1)(\tilde{e} - e_h - \mu_h) + \tau_\mu(\nabla u_h - e_h), \delta e_h) = 0$$

and

$$(-(\nabla u_h - e_h) + \tau_e(\tilde{e} - e_h - \mu_h), \delta \mu_h) = 0.$$

By choosing

$$\delta \mu_h = \tau_\mu \delta e_h$$

and after some algebraic manipulations on the two last equations, we arrive at

$$(e_h + \mu_h - \tilde{e}, \delta e_h) = 0,$$

which is proposition's second orthogonality result. Similarly, by choosing

$$\delta e_h = \frac{1}{\tau_\mu} \delta \mu_h$$

and after some algebraic manipulations, we arrive at

$$(-\nabla u_h + e_h, \delta \mu_h) = 0,$$

which is proposition's third orthogonality result. $\square$

Proposition 2 (minimally stabilized discrete primal formulation). *Let h be small enough with respect to ℓ_Ω . For $k_u = k_\lambda = 0$ and $M_h = E_h$, the following assertions hold true: i) the stabilization term is mesh-independent, i.e., $S_{p,h}(\mathbf{x}_h, \delta \mathbf{x}_h)$ can be written as $S_p(\mathbf{x}_h, \delta \mathbf{x}_h)$, with*

$$\begin{aligned} S_p(\mathbf{x}_h, \delta \mathbf{x}_h) := & -k_e((e_h, \delta e_h) + (\mu_h, \delta e_h) + (e_h, \delta \mu_h) + (\mu_h, \delta \mu_h)) \\ & - \frac{k_s}{\kappa}((\nabla \lambda_h, \delta \nabla \lambda_h) - \kappa(s_h, \nabla \delta \lambda_h) - \kappa(\nabla \lambda_h, s_h) + \kappa^2(s_h, \delta s_h)) \\ & + k_\mu((\nabla u_h, \nabla \delta u_h) - (e_h, \nabla \delta u_h) - (\nabla u_h, \delta e_h) + (e_h, \delta e_h)); \end{aligned}$$

and, ii) the ASGS and the OSGS formulations are equivalent.

Proof. The first assertion follows intermediately. The second one is a direct consequence of the fact that all subscales are (by design) either zero or orthogonal to the corresponding finite element space (due to the previous proposition). $\square$

3.2 Discretized dual problem

The stabilized bilinear form for the discrete dual formulation is

$$B'_{\text{stab,d}}(\mathbf{x}_h, \delta \mathbf{x}_h) := B_d(\mathbf{x}_h, \delta \mathbf{x}_h) + S'_d(\mathbf{x}_h, \delta \mathbf{x}_h),$$

where $S'_{d,h}(\mathbf{x}_h, \delta \mathbf{x}_h)$ is, in principle, a mesh-dependent term providing the desired stability. In addition, let us consider:

$$\tau_u = k_u \ell_\Omega^2, \quad \tau_e = k_e, \quad \tau_s = \frac{k_s h^2}{\kappa \ell_\Omega^2}, \quad \tau_\lambda = -\frac{k_\lambda}{\zeta^2 \ell_\Omega^2}, \quad \tau_\mu = -\frac{k_\mu h^2}{\ell_\Omega^2}.$$

Then, the stabilization term is as follows:

$$\begin{aligned} S'_{d,h}(\mathbf{x}_h, \delta \mathbf{x}_h) := & -k_u \ell_\Omega^2 [\zeta^2 (P' \lambda_h, \delta \lambda_h) + \zeta ((P' \nabla \cdot \mu_h, \delta \lambda_h) + (P' \lambda_h, \nabla \cdot \delta \mu_h)) + (P' \nabla \cdot \mu_h, \nabla \cdot \delta \mu_h)] \\ & - k_e [(P' e_h, \delta e_h) + (P' \mu_h, \delta e_h) + (P' e_h, \delta \mu_h) + (P' \mu_h, \delta \mu_h)] \\ & - \frac{k_s h^2}{\kappa \ell_\Omega^2} [(P' \nabla \lambda_h, \delta \nabla \lambda_h)_h - \kappa (P' s_h, \nabla \delta \lambda_h)_h - \kappa (P' \nabla \lambda_h, \delta s_h)_h + \kappa^2 (P' s_h, \delta s_h)] \\ & + \frac{k_\lambda}{\zeta^2 \ell_\Omega^2} [\zeta^2 (P' u_h, \delta u_h) + \zeta ((P' \nabla \cdot s_h, \delta u_h) + (P' u_h, \nabla \cdot \delta s_h)) + (P' \nabla \cdot s_h, \nabla \cdot \delta s_h)] \\ & + \frac{k_\mu h^2}{\ell_\Omega^2} [(P' \nabla u_h, \nabla \delta u_h)_h - (P' e_h, \nabla \delta u_h)_h - (P' \nabla u_h, \delta e_h)_h + (P' e_h, \delta e_h)]. \end{aligned}$$

Particularity, for $P' = P^\perp$ we have that

$$\begin{aligned} S_{d,h}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h) = & + \frac{k_\mu h^2}{\ell_\Omega^2} (P^\perp(\nabla u_h), P^\perp(\nabla \delta u_h))_h \\ & + \frac{k_\lambda}{\zeta^2 \ell_\Omega^2} (P^\perp(\nabla \cdot \mathbf{s}_h), P^\perp(\nabla \cdot \delta \mathbf{s}_h)) \\ & - \frac{k_s h^2}{\kappa \ell_\Omega^2} (P^\perp(\nabla \lambda_h), P^\perp(\nabla \delta \lambda_h))_h \\ & - k_u \ell_\Omega^2 (P^\perp(\nabla \cdot \boldsymbol{\mu}_h), P^\perp(\nabla \cdot \delta \boldsymbol{\mu}_h)). \end{aligned}$$

Problem 5 ($P_{d,h}$). Find $\mathbf{x}_h \in X_{d,h}$ such that

$$B'_{\text{stab},d}(\mathbf{x}_h, \delta \mathbf{x}_h) = L'_{\text{stab}}(\delta \mathbf{x}_h), \quad \forall \delta \mathbf{x}_h \in X_{d,h}. \quad (16)$$

Definition 4. Let $X_{d,h}$ be defined as

$$X_{d,h} = U_h \times E_h \times S_h \times L_h \times M_h$$

and equipped with the norm defined by

$$\begin{aligned} \|\mathbf{x}_h\|_{X_{d,h}}^2 = & \frac{1}{\ell_\Omega^2} \|u_h\|^2 + \frac{h^2}{\ell_\Omega^2} \|\nabla u_h\|_h^2 + \|\mathbf{e}_h\|^2 + \kappa \|\mathbf{s}_h\|^2 + \kappa \ell_\Omega^2 \|\nabla \cdot \mathbf{s}_h\|^2 \\ & + \frac{1}{\kappa \ell_\Omega^2} \|\lambda_h\|^2 + \frac{h^2}{\kappa \ell_\Omega^2} \|\nabla \lambda_h\|_h^2 + \|\boldsymbol{\mu}_h\|^2 + \ell_\Omega^2 \|\nabla \cdot \boldsymbol{\mu}_h\|^2. \end{aligned}$$

Theorem 4. The discretized dual problem ($P_{d,h}$) is well posed for P' being either the identity (ASGS) or the $L^2(\Omega)$ -projection on the orthogonal complement of the finite element space chosen (OSGS). In addition, it consistently approximates the continuous problem P_d .

Proof. The continuity of $B'_{\text{stab},d}(\mathbf{x}_h, \delta \mathbf{x}_h)$ and $L'_{\text{stab}}(\delta \mathbf{x}_h)$ in the norm of $X_{d,h}$ is straightforward. For the inf-sup stability of $B_{\text{stab},d}(\mathbf{x}_h, \delta \mathbf{x}_h)$, let P' be equal to I and set $\delta \mathbf{x}_h^0 = (u_h, \mathbf{e}_h, \mathbf{s}_h, -\lambda_h, -\boldsymbol{\mu}_h)$. In addition, let us consider the following inequalities:

$$\begin{aligned} -2k_\mu(\mathbf{e}_h, \nabla u_h)_h & \geq -\frac{1}{2} \|\mathbf{e}_h\|^2 - 4k_\mu^2 \|\nabla u_h\|_h^2, \quad (\text{Young}), \\ \frac{2}{\zeta}(\nabla \cdot \mathbf{s}_h, u_h) & \geq -\frac{k_1}{\zeta^2} \|\nabla \cdot \mathbf{s}_h\|^2 - \frac{1}{k_1} \|u_h\|^2, \quad \forall k_1 > 0, \quad (\text{Young}), \\ 2\zeta(\nabla \cdot \boldsymbol{\mu}_h, \lambda_h) & \geq -k_2 \|\nabla \cdot \boldsymbol{\mu}_h\|^2 - \frac{\zeta^2}{k_2} \|\lambda_h\|^2, \quad \forall k_2 > 0, \quad (\text{Young}), \\ \|\nabla u_h\|_h^2 & \geq k_3 \|\nabla u_h\|_h^2 + \frac{(1-k_3)}{C_{\text{PF}}^2 \ell_\Omega^2} \|u_h\|^2, \quad \forall k_3 \in (0, 1), \quad (\text{Poincaré}), \\ \|\nabla \lambda_h\|_h^2 & \geq k_4 \|\nabla \lambda_h\|_h^2 + \frac{(1-k_4)}{C_{\text{PF}}^2 \ell_\Omega^2} \|\lambda_h\|^2, \quad \forall k_4 \in (0, 1), \quad (\text{Poincaré}). \end{aligned}$$

After some algebraic manipulations, we arrive at

$$\begin{aligned} B_{\text{stab},d}(\mathbf{x}_h, \delta \mathbf{x}_h^0) \geq & \tau_1 \|u_h\|^2 + \tau_2 \|\mathbf{e}_h\|^2 + \tau_3 \|\mathbf{s}_h\|^2 + \tau_4 \|\lambda_h\|^2 + \tau_5 \|\boldsymbol{\mu}_h\|^2 \\ & + \tau_6 \|\nabla u_h\|_h^2 + \tau_7 \|\nabla \cdot \mathbf{s}_h\|^2 + \tau_8 \|\nabla \lambda_h\|_h^2 + \tau_9 \|\nabla \cdot \boldsymbol{\mu}_h\|^2, \end{aligned}$$

where

$$\begin{aligned}
\tau_1 &= \left(1 - \frac{1}{k_1}\right) \frac{k_\lambda}{\ell_\Omega^2} + \frac{(1-k_3)}{C_{\text{PF}}^2 \ell_\Omega^2} k_\mu (1-4k_\mu) \frac{h^2}{\ell_\Omega^2}, \\
\tau_2 &= \left(k_\mu - \frac{1}{4}\right) \frac{h^2}{\ell_\Omega^2} - k_e + 1, \\
\tau_3 &= \kappa \left(1 - k_s \frac{h^2}{\ell_\Omega^2}\right), \\
\tau_4 &= \left(1 - \frac{1}{k_2}\right) k_u \ell_\Omega^2 \zeta^2 + \frac{(1-k_4)}{C_{\text{PF}}^2 \ell_\Omega^2} \frac{k_s h^2}{\kappa \ell_\Omega^2}, \\
\tau_5 &= k_e, \\
\tau_6 &= k_3 k_\mu (1-4k_\mu) \frac{h^2}{\ell_\Omega^2}, \\
\tau_7 &= (1-k_1) \frac{k_\lambda}{\ell_\Omega^2 \zeta^2}, \\
\tau_8 &= k_4 \frac{k_s h^2}{\kappa \ell_\Omega^2}, \\
\tau_9 &= (1-k_2) k_u \ell_\Omega^2.
\end{aligned}$$

Expresssion $B_{\text{stab,d}}(\mathbf{x}_h, \delta \mathbf{x}_h^0)$ is positive provided that

$$\begin{aligned}
k_u > 0, k_e \in (0, k_e^*), k_s \in [0, k_s^*), k_\lambda > 0, k_\mu \in \left[0, \frac{1}{4}\right], \\
k_1 \in (k_1^*, 1), \quad k_2 \in (k_2^*, 1), \quad k_3 \in (0, 1), \quad k_4 \in (0, 1),
\end{aligned}$$

where

$$\begin{aligned}
k_e^* &= 1 + \left(k_\mu - \frac{1}{4}\right) \frac{h^2}{\ell_\Omega^2}, \quad k_s^* = \frac{\ell_\Omega^2}{h^2}, \\
k_1^* &= \frac{1}{1 + \frac{k_\mu(1-4k_\mu)}{k_\lambda} \frac{(1-k_3)}{C_{\text{PF}}^2} \frac{h^2}{\ell_\Omega^2}}, \quad k_2^* = \frac{1}{1 + \frac{k_s}{k_u} \frac{(1-k_4)}{C_{\text{PF}}^2} \frac{h^2}{\kappa \zeta^2 \ell_\Omega^6}}.
\end{aligned}$$

For $P' = P^\perp$, i.e., the projector onto the $L^2(\Omega)$ -orthogonal complement of the chosen particular finite element space, we have

$$\begin{aligned}
B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^0) &\geq \|\mathbf{e}_h\|^2 + \kappa \|\mathbf{s}_h\|^2 \\
&+ \frac{k_\mu h^2}{\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 + \frac{k_\lambda}{\zeta^2 \ell_\Omega^2} \|P^\perp(\nabla \cdot \mathbf{s}_h)\|^2 + \frac{k_s h^2}{\kappa \ell_\Omega^2} \|P^\perp(\nabla \lambda_h)\|_h^2 + k_u \ell_\Omega^2 \|P^\perp(\nabla \cdot \boldsymbol{\mu}_h)\|^2.
\end{aligned}$$

Stability on the whole norm of $\nabla \cdot \mathbf{s}_h$ and $\nabla \cdot \boldsymbol{\mu}_h$ is obtained using an argument similar to that of Theorem 3. The first task we need to accomplish is to get control over $\|P(\nabla \cdot \mathbf{s}_h)\|$ and $\|P(\nabla \cdot \boldsymbol{\mu}_h)\|$. For such an end, let us take $\delta \mathbf{x}_h^1 = (k_u \ell_\Omega^2 P(\nabla \cdot \boldsymbol{\mu}_h), \mathbf{0}, \mathbf{0}, k_\lambda \ell_\Omega^{-2} \zeta^{-2} P(\nabla \cdot \mathbf{s}_h), \mathbf{0})$, which yields

$$\begin{aligned}
B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^1) &= + \frac{k_\lambda}{\ell_\Omega^2 \zeta^2} \|P(\nabla \cdot \mathbf{s}_h)\|^2 + k_u \ell_\Omega^2 \|P(\nabla \cdot \boldsymbol{\mu}_h)\|^2 \\
&+ \frac{k_\lambda}{\ell_\Omega^2 \zeta^2} (u_h, P(\nabla \cdot \mathbf{s}_h)) + k_u \zeta \ell_\Omega^2 (\lambda_h, P(\nabla \cdot \boldsymbol{\mu}_h)) \\
&+ k_u k_\mu h^2 (P^\perp(\nabla u_h), \nabla(P(\nabla \cdot \boldsymbol{\mu}_h)))_h \\
&- \frac{k_s k_\lambda h^2}{\kappa \zeta^2 \ell_\Omega^4} (P^\perp(\nabla \lambda_h), \nabla(P(\nabla \cdot \mathbf{s}_h)))_h.
\end{aligned}$$

By employing Young's inequality for the third and fourth terms and by combining Young's and inverse inequalities for the fifth and sixth terms, namely,

$$+k_u k_\mu h^2 (P^\perp(\nabla u_h), \nabla(P(\nabla \cdot \mu_h)))_h \geq -\frac{k_7 C_{\text{inv}}^2 h^2 k_u k_\mu^2}{2\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 - \frac{k_u \ell_\Omega^2}{2k_7} \|P(\nabla \cdot \mu_h)\|_h^2$$

and

$$-\frac{k_s k_\lambda h^2}{\kappa \zeta^2 \ell_\Omega^4} (P^\perp(\nabla \lambda_h), \nabla(P(\nabla \cdot s_h)))_h \geq -\frac{k_8 C_{\text{inv}}^2 h^2 k_s^2 k_\lambda}{2\zeta^2 \ell_\Omega^6 \kappa^2} \|P^\perp(\nabla \lambda_h)\|_h^2 - \frac{k_\lambda}{2k_8 \ell_\Omega^2 \zeta^2} \|P(\nabla \cdot s_h)\|_h^2,$$

we arrive at

$$\begin{aligned} B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^1) &\geq + \left(1 - \frac{1}{2k_5} - \frac{1}{2k_8}\right) \frac{k_\lambda}{\ell_\Omega^2 \zeta^2} \|P(\nabla \cdot s_h)\|^2 \\ &\quad + \left(1 - \frac{1}{2k_6} - \frac{1}{2k_7}\right) k_u \ell_\Omega^2 \|P(\nabla \cdot \mu_h)\|^2 \\ &\quad - \frac{k_5 k_\lambda}{2\ell_\Omega^2} \|u_h\|^2 - \frac{k_7 C_{\text{inv}}^2 h^2 k_u^2 k_\mu}{2\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 \\ &\quad - \frac{k_6 k_u \ell_\Omega^2 \zeta^2}{2} \|\lambda_h\|^2 - \frac{k_8 C_{\text{inv}}^2 h^2 k_s^2 k_\lambda}{2\kappa^2 \ell_\Omega^6 \zeta^2} \|P^\perp(\nabla \lambda_h)\|_h^2 \\ &\gtrsim + \kappa \ell_\Omega^2 \|P(\nabla \cdot s_h)\|^2 + \ell_\Omega^2 \|P(\nabla \cdot \mu_h)\|^2 \\ &\quad - \frac{1}{2\ell_\Omega^2} \|u_h\|^2 - \frac{h^2}{\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 \\ &\quad - \frac{1}{\kappa \ell_\Omega^2} \|\lambda_h\|^2 - \frac{h^2}{\kappa \ell_\Omega^2} \|P^\perp(\nabla \lambda_h)\|_h^2, \end{aligned}$$

where the control over $\|P(\nabla \cdot s_h)\|$ and $\|P(\nabla \cdot \mu_h)\|$ is warranted for sufficiently large k_i (for all i in $\{5, 6, 7, 8\}$).

The second task we need to accomplish, is to get control over $\|P_{L_h}(u_h)\|$ and $\|P_{U_h}(\lambda_h)\|$, where P_{L_h} is the $L^2(\Omega)$ -projection onto L_h and P_{U_h} is the $L^2(\Omega)$ -projection onto U_h , which now need to be distinguished. The natural choice would be to take $\sqrt{\kappa}U_h = L_h$ (recall that both spaces are made of continuous functions), but for generality we consider the case in which these spaces are different. For such an end, let us take $\delta \mathbf{x}_h^2 = (\zeta^{-1} \ell_\Omega^{-2} \kappa^{-1} P_{U_h}(\lambda_h), \mathbf{0}, \mathbf{0}, \zeta^{-1} \ell_\Omega^{-2} P_{L_h}(u_h), \mathbf{0})$, which yields

$$\begin{aligned} B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^2) &= + \frac{1}{\ell_\Omega^2} \|P_{L_h}(u_h)\|^2 + \frac{1}{\kappa \ell_\Omega^2} \|P_{U_h}(\lambda_h)\|^2 \\ &\quad + \frac{k_\mu h^2}{\kappa \ell_\Omega^4 \zeta} (P^\perp(\nabla u_h), \nabla(P_{U_h}(\lambda_h)))_h - \frac{k_s h^2}{\kappa \ell_\Omega^4 \zeta} (\nabla(P_{L_h}(u_h)), P^\perp(\nabla \lambda_h))_h \\ &\quad + \frac{1}{\ell_\Omega^2 \zeta} (P_{L_h}(u_h), \nabla \cdot s_h) + \frac{1}{\kappa \ell_\Omega^2 \zeta} (P_{U_h}(\lambda_h), \nabla \cdot \mu_h). \end{aligned}$$

By combining Young's and inverse inequalities for the third and fourth terms and employing

Young's inequality to the fifth and sixth terms, we arrive at

$$\begin{aligned}
B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^2) &\geq + \left(1 - \frac{1}{2k_9} - \frac{k_s}{2k_{11}\kappa\ell_\Omega^4\zeta^2}\right) \frac{1}{\ell_\Omega^2} \|P_{L_h}(u_h)\|^2 - \frac{k_{11}k_\mu h^2}{2\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 \\
&\quad + \left(1 - \frac{1}{2k_{10}} - \frac{k_\mu}{2k_{12}\kappa\ell_\Omega^4\zeta^2}\right) \frac{1}{\kappa\ell_\Omega^2} \|P_{U_h}(\lambda_h)\|^2 - \frac{k_{12}k_s h^2}{2\kappa\ell_\Omega^2} \|P^\perp(\nabla \lambda_h)\|_h^2 \\
&\quad - \frac{k_9}{2\ell_\Omega^2\zeta^2} \|\nabla \cdot \mathbf{s}_h\|^2 - \frac{k_{10}}{2\kappa\ell_\Omega^2\zeta^2} \|\nabla \cdot \boldsymbol{\mu}_h\|^2 \\
&\gtrsim + \frac{1}{\ell_\Omega^2} \|P_{L_h}(u_h)\|^2 - \frac{h^2}{\ell_\Omega^2} \|P^\perp(\nabla u_h)\|_h^2 \\
&\quad + \frac{1}{\kappa\ell_\Omega^2} \|P_{U_h}(\lambda_h)\|^2 - \frac{h^2}{\kappa\ell_\Omega^2} \|P^\perp(\nabla \lambda_h)\|_h^2 \\
&\quad - \kappa\ell_\Omega^2 \|\nabla \cdot \mathbf{s}_h\|^2 - \ell_\Omega^2 \|\nabla \cdot \boldsymbol{\mu}_h\|^2,
\end{aligned}$$

where the control over $\|P_{L_h}(u_h)\|$ and $\|P_{U_h}(\lambda_h)\|$ is warranted for sufficiently large k_i (for all i in $\{9, 10, 11, 12\}$) and sufficiently small k_s and k_μ .

Finally, setting $\delta \mathbf{x}_h^\beta = \delta \mathbf{x}_h^0 + \beta_1 \delta \mathbf{x}_h^1 + \beta_2 \delta \mathbf{x}_h^2$ and taking β_1 and β_2 positive and sufficiently small yields

$$B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^\beta) \gtrsim \|\mathbf{x}_h\|_{X_{\text{d},h}^*}^2,$$

which holds true due to the conjugation of the estimates for $B_{\text{stab,d}}^\perp(\mathbf{x}_h, \delta \mathbf{x}_h^i)$ (for all i in $\{0, 1, 2\}$), and the norm $\|\cdot\|_{X_{\text{d},h}^*}$ is the same as $\|\cdot\|_{X_{\text{d},h}}$ replacing the $L^2(\Omega)$ -norm of u_h and λ_h by $P_{L_h}(u_h)$ and $P_{U_h}(\lambda_h)$, respectively. Then, the desired inf-sup condition follows after checking that $\|\mathbf{x}_h\|_{X_{\text{d},h}^*} \gtrsim \|\delta \mathbf{x}_h^\beta\|_{X_{\text{d},h}^*}$. $\square$

Remark 7. When $\sqrt{\kappa}U_h \neq L_h$ the norm $\|\cdot\|_{X_{\text{d},h}^*}$ is slightly weaker than $\|\cdot\|_{X_{\text{d},h}}$. However, this does not alter the well-posedness result stated in Theorem 4.

Proposition 3. Let $k_s = k_\mu = 0$, $P' = I$ and $\sqrt{\kappa}U_h = L_h$. Then, the following residuum orthogonalities hold:

$$\begin{aligned}
(\zeta \lambda_h + \nabla \cdot \boldsymbol{\mu}_h, \delta u_h) &= 0 && (\text{orthogonality w.r.t. } U_h), \\
(\mathbf{e}_h + \boldsymbol{\mu}_h - \tilde{\mathbf{e}}, \delta \mathbf{e}_h) &= 0 && (\text{orthogonality w.r.t. } E_h), \\
(\zeta u_h + \nabla \cdot \mathbf{s}_h - q, \delta \lambda_h) &= 0 && (\text{orthogonality w.r.t. } L_h).
\end{aligned}$$

Proof. First, by setting $k_s = k_\mu = 0$, it follows that

$$(\zeta \lambda_h + \zeta \tau_\lambda P'(q - \zeta u_h - \nabla \cdot \mathbf{s}_h) + \nabla \cdot \boldsymbol{\mu}_h, \delta u_h) = 0$$

and

$$(\zeta u_h + \zeta \tau_u P'(-\zeta \lambda_h - \nabla \cdot \boldsymbol{\mu}_h) + \nabla \cdot \mathbf{s}_h - q, \delta \lambda_h) = 0.$$

Second, by setting $P' = I$, we have

$$(\zeta \lambda_h + \zeta \tau_\lambda (q - \zeta u_h - \nabla \cdot \mathbf{s}_h) + \nabla \cdot \boldsymbol{\mu}_h, \delta u_h) = 0$$

and

$$(\zeta u_h + \zeta \tau_u (-\zeta \lambda_h - \nabla \cdot \boldsymbol{\mu}_h) + \nabla \cdot \mathbf{s}_h - q, \delta \lambda_h) = 0.$$

By choosing

$$\delta \lambda_h = \zeta \tau_\lambda \delta u_h$$

and after some algebraic manipulations on the two last equations, we arrive at

$$(\zeta \lambda_h + \nabla \cdot \boldsymbol{\mu}_h, \delta u_h) = 0,$$

which is proposition's first orthogonality result. Similarly, by choosing

$$\delta u_h = \zeta \tau_u \delta \lambda_h$$

and after some algebraic manipulations, we arrive at

$$(\zeta u_h + \nabla \cdot \mathbf{s}_h - q, \delta \lambda_h) = 0,$$

which is proposition's third orthogonality result. Finally, considering $P_{d,h}$ and noting that $\boldsymbol{\mu}' = 0$ (by design), we arrive at

$$(\mathbf{e}_h - \boldsymbol{\mu}_h - \tilde{\mathbf{e}}, \delta \mathbf{e}_h) = 0,$$

which is proposition's second orthogonality result. $\square$

Proposition 4 (minimally stabilized discrete dual formulation). *Let h be small enough with respect to ℓ_Ω . For $k_s = k_\mu = 0$ and $\sqrt{\kappa}U_h = L_h$, the following assertions hold true: i) the stabilization term is mesh-independent, i.e., $S_{d,h}(\mathbf{x}_h, \delta \mathbf{x}_h)$ can be written as $S_d(\mathbf{x}_h, \delta \mathbf{x}_h)$ with*

$$\begin{aligned} S_d(\mathbf{x}_h, \delta \mathbf{x}_h) := & -k_u \ell_\Omega^2 (\zeta^2(\lambda_h, \delta \lambda_h) + \zeta((\nabla \cdot \boldsymbol{\mu}_h, \delta \lambda_h) + (\lambda_h, \nabla \cdot \delta \boldsymbol{\mu}_h)) + (\nabla \cdot \boldsymbol{\mu}_h, \nabla \cdot \delta \boldsymbol{\mu}_h)) \\ & - k_e ((\mathbf{e}_h, \delta \mathbf{e}_h) + (\boldsymbol{\mu}_h, \delta \mathbf{e}_h) + (\mathbf{e}_h, \delta \boldsymbol{\mu}_h) + (\boldsymbol{\mu}_h, \delta \boldsymbol{\mu}_h)) \\ & + \frac{k_\lambda}{\zeta^2 \ell_\Omega^2} (\zeta^2(u_h, \delta u_h) + \zeta((\nabla \cdot \mathbf{s}_h, \delta u_h) + (u_h, \nabla \cdot \delta \mathbf{s}_h)) + (\nabla \cdot \mathbf{s}_h, \nabla \cdot \delta \mathbf{s}_h)); \end{aligned}$$

and, ii) the ASGS and the OSGS formulations are equivalent.

Proof. The first assertion follows intermediately. The second is a direct consequence of the fact that all subscales are (by design) either zero or orthogonal to the corresponding finite element space (due to the previous proposition). $\square$

3.3 Expected order of convergence

Theorems 3 and 4 essentially state that both the primal and the dual discrete formulations proposed are inf-sup stable, the former in the norm $\|\cdot\|_{X_{p,h}}$ introduced in Definition 3 and the latter in the norm $\|\cdot\|_{X_{d,h}}$ in Definition 4. From this and the consistency of the formulations, a standard analysis yields optimal convergence in these norms. For the sake of conciseness, we will skip this analysis, which would be worth carrying out in a more general setting than the one considered in this work. Nevertheless, the outcome is that the error of the discrete solution in these norms behaves as the interpolation error, also in these norms.

Let p_u, p_e, p_s, p_λ and p_μ be the interpolation orders of $u, \mathbf{e}, \mathbf{s}, \lambda$ and $\boldsymbol{\mu}$, respectively. Recall that we have considered all finite element spaces as made of continuous functions. From the expression of $\|\cdot\|_{X_{p,h}}$ in Definition 3 and assuming enough regularity of the continuous solutions, the error function of the primal formulation will behave as

$$\begin{aligned} E_p(h) \sim & \frac{1}{\ell_\Omega} h^{p_u+1} + h^{p_u} + h^{p_e+1} + \sqrt{\kappa} h^{p_s+1} + \sqrt{\kappa} h h^{p_s} \\ & + \frac{1}{\sqrt{\kappa} \ell_\Omega} h^{p_\lambda+1} + \frac{1}{\sqrt{\kappa}} h^{p_\lambda} + h^{p_\mu+1} + h h^{p_\mu}, \end{aligned}$$

which for equal order interpolation p for all the unknowns yields $E_p(h) \sim h^p$. In particular, this implies that

$$\begin{aligned} \|\nabla u - \nabla u_h\| &= \mathcal{O}(h^p), \quad \|\nabla \lambda - \nabla \lambda_h\| = \mathcal{O}(h^p), \\ \|\nabla \cdot \mathbf{s} - \nabla \cdot \mathbf{s}_h\| &= \mathcal{O}(h^{p-1}), \quad \|\nabla \cdot \boldsymbol{\mu} - \nabla \cdot \boldsymbol{\mu}_h\| = \mathcal{O}(h^{p-1}). \end{aligned}$$

On the other hand, from the expression of $\|\cdot\|_{X_{d,h}}$ in Definition 4 and assuming again enough regularity of the continuous solutions, the error function of the dual formulation will behave as

$$\begin{aligned} E_d(h) &\sim \frac{1}{\ell_\Omega} h^{p_u+1} + \frac{1}{\ell_\Omega} h h^{p_u} + h^{p_e+1} + \sqrt{\kappa} h^{p_s+1} + \sqrt{\kappa} \ell_\Omega h^{p_s} \\ &\quad + \frac{1}{\sqrt{\kappa} \ell_\Omega} h^{p_\lambda+1} + \frac{1}{\sqrt{\kappa} \ell_\Omega} h h^{p_\lambda} + h^{p_\mu+1} + \ell_\Omega h^{p_\mu}, \end{aligned}$$

which for equal interpolation order p for all the unknowns yields also $E_d(h) \sim h^p$. Now this implies that

$$\begin{aligned} \|\nabla u - \nabla u_h\| &= \mathcal{O}(h^{p-1}), \quad \|\nabla \lambda - \nabla \lambda_h\| = \mathcal{O}(h^{p-1}), \\ \|\nabla \cdot \mathbf{s} - \nabla \cdot \mathbf{s}_h\| &= \mathcal{O}(h^p), \quad \|\nabla \cdot \boldsymbol{\mu} - \nabla \cdot \boldsymbol{\mu}_h\| = \mathcal{O}(h^p). \end{aligned}$$

This clearly highlights the transfer of one order of convergence from ∇u_h and $\nabla \lambda_h$ to $\nabla \cdot \mathbf{s}_h$ and $\nabla \cdot \boldsymbol{\mu}_h$ when moving from the primal to the dual formulation. For the $L^2(\Omega)$ -norm of the unknowns themselves we expect one order more of convergence than for their derivatives (either gradients or divergences). This should be proved using classical duality arguments under appropriate conditions on the computational domain.

4 Numerical results

This section presents a numerical assessment of both the primal and dual formulations introduced above. Two numerical examples are considered. The first aims to evaluate the convergence properties of the methods by computing relevant error norms against a smooth manufactured solution. These results will be compared with those from a *Natural* formulation, a *Fully stabilized* equal-order formulation based on GLS and an *Unstabilized* equal-order formulation, all recently discussed in [4].

The second example addresses a more challenging problem featuring an internal reaction boundary layer, in order to assess the approximation properties of the proposed methods in the presence of large solution gradients. This case clearly illustrates key differences between the primal and dual formulations and further allows us to examine the effect of different choices for the data fields $(\tilde{e}, \tilde{\mathbf{s}})$ making this final experiment particularly relevant for DDCM applications.

We consider a sequence of conforming, shape-regular finite element partitions $\mathcal{P}_h$ of the computational domain, with characteristic mesh size h . A primary motivation for the VMS approach is the ability to use equal-order approximations. Accordingly, we employ the following choice of discrete finite element spaces:

$$(u_h, \mathbf{e}_h, \mathbf{s}_h, \lambda_h, \boldsymbol{\mu}_h) \in \text{CG}_k \times [\text{CG}_k]^{d_\Omega} \times [\text{CG}_k]^{d_\Omega} \times \text{CG}_k \times [\text{CG}_k]^{d_\Omega},$$

where d_Ω denotes the spatial dimension of the computational domain Ω , $k > 0$ and

$$\text{CG}_k := \{v \in C^0(\Omega) : v|_K \in \mathbb{P}_k(K) \ \forall K \in \mathcal{P}_h\}$$

where $\mathbb{P}_k(K)$ denotes the space of polynomials of degree k defined on each element $K \in \mathcal{P}_h$. In the ASGS method, being P' the identity operator, no additional unknowns are required, since the subscales are simply evaluated locally at the element level. In contrast, for the OSGS method, with P' denoting the $L^2(\Omega)$ -projector onto the orthogonal complement of the finite element space, one additional unknown per field must be considered. That is, if we denote by $w \in W$ any of the unknown fields, such that $w = w_h + w'$ with $w_h \in W_h \subsetneq W$, the subscale component w' is in this case given by

$$w' = \tau_w P' (r(w)) = \tau_w [r_w - P_{W_h} (r_w)] \quad (17)$$

where r_w is the algebraic residual equation of the variable w and P_{W_h} stands for the $L^2(\Omega)$ -projection onto W_h . Since this leads to an excessive computational burden, in practice the explicit

computation of the subscales is avoided by applying some sort of fixed-point iteration scheme. In our implementation, these variables are included as unknowns and solved for monolithically. This approach yields cleaner code, simplifies the implementation and eliminates the numerical error associated with an additional iterative procedure. We anticipate, however, that no significant differences in results were observed between the ASGS and OSGS methods for either the primal or dual formulations. Consequently, all results presented hereafter are obtained by using the ASGS method. Finally, the implementation of both formulations, including the two variants, ASGS and OSGS, was carried out using the finite element platform Firedrake [21] and is available in <https://github.com/pedrobbazon/DDCM.git>.

Remark 8. *In this proof-of-concept study, we consider two-dimensional numerical examples of relatively small size and rely on direct solvers for the resulting linear systems, which allows us to avoid additional challenges associated with numerical linear algebra, particularly those stemming from the saddle-point structure of the problem. For larger systems, however, direct methods quickly become impractical; scalable performance will therefore require efficient iterative solvers equipped with suitable preconditioning strategies.*

4.1 Example 1: Convergence to a manufactured solution

Let us first describe the construction of a manufactured solution with respect to which the relevant error norms will be computed. Its construction deserves some attention, although it essentially follows the idea presented in [4]. By examining the optimality conditions in their strong form, it is straightforward to devise a manufactured solution that resembles a relevant scenario for DDCM.

The manufactured solution is illustrated in Figure 1, which shows all the primal fields $(u, \mathbf{s}, \mathbf{e})$, the dual fields $(\lambda, \boldsymbol{\mu})$, and the data fields $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$. To construct this solution, we begin by defining the primal scalar field

$$u(x, y) = \cos(\pi x) \cos(\pi y), \quad (18)$$

whose gradient is given by

$$\mathbf{e}(x, y) = -\pi [\sin(\pi x) \cos(\pi y), \cos(\pi x) \sin(\pi y)]^\top. \quad (19)$$

We then define $\tilde{\mathbf{e}}$ as a perturbation of $\mathbf{e}$ as follows

$$\tilde{\mathbf{e}}(x, y) = \mathbf{e}(x, y) + \frac{1}{20} [\sin(4\pi x), \sin(4\pi y)]^\top \quad (20)$$

Then, for the primal flux $\mathbf{s}$ we consider

$$\mathbf{s} = -\frac{\mathbf{e}}{\|\mathbf{e}\|} \left(\|\mathbf{e}\| - \frac{1}{40} \|\mathbf{e}\|^3 \right). \quad (21)$$

such as to mimic a nonlinear constitutive relation between $\mathbf{s}$ and $\mathbf{e}$, as plotted in Figure 1 (f), which shows the corresponding relation between the magnitude of $\mathbf{s}$ and the magnitude of $\mathbf{e}$ obtained by sampling the solution over a computational mesh. For the dual scalar field λ we take the function

$$\lambda(x, y) = -\frac{1}{40\pi} [\cos(2\pi x) + \cos(2\pi y)], \quad (22)$$

and finally from the equation for $\mathbf{e}$ we simply obtain $\boldsymbol{\mu}$ as

$$\boldsymbol{\mu} = \tilde{\mathbf{e}} - \mathbf{e} \quad (23)$$

We need to introduce a (dual) source term on the right-hand-side of the equation for u , which must satisfy

$$\zeta \lambda + \nabla \cdot \boldsymbol{\mu} = f, \quad (24)$$

but note that this is done solely for the purpose of assessing the convergence properties of the method using such a manufactured solution. Finally, $\tilde{s}$ is obtained from

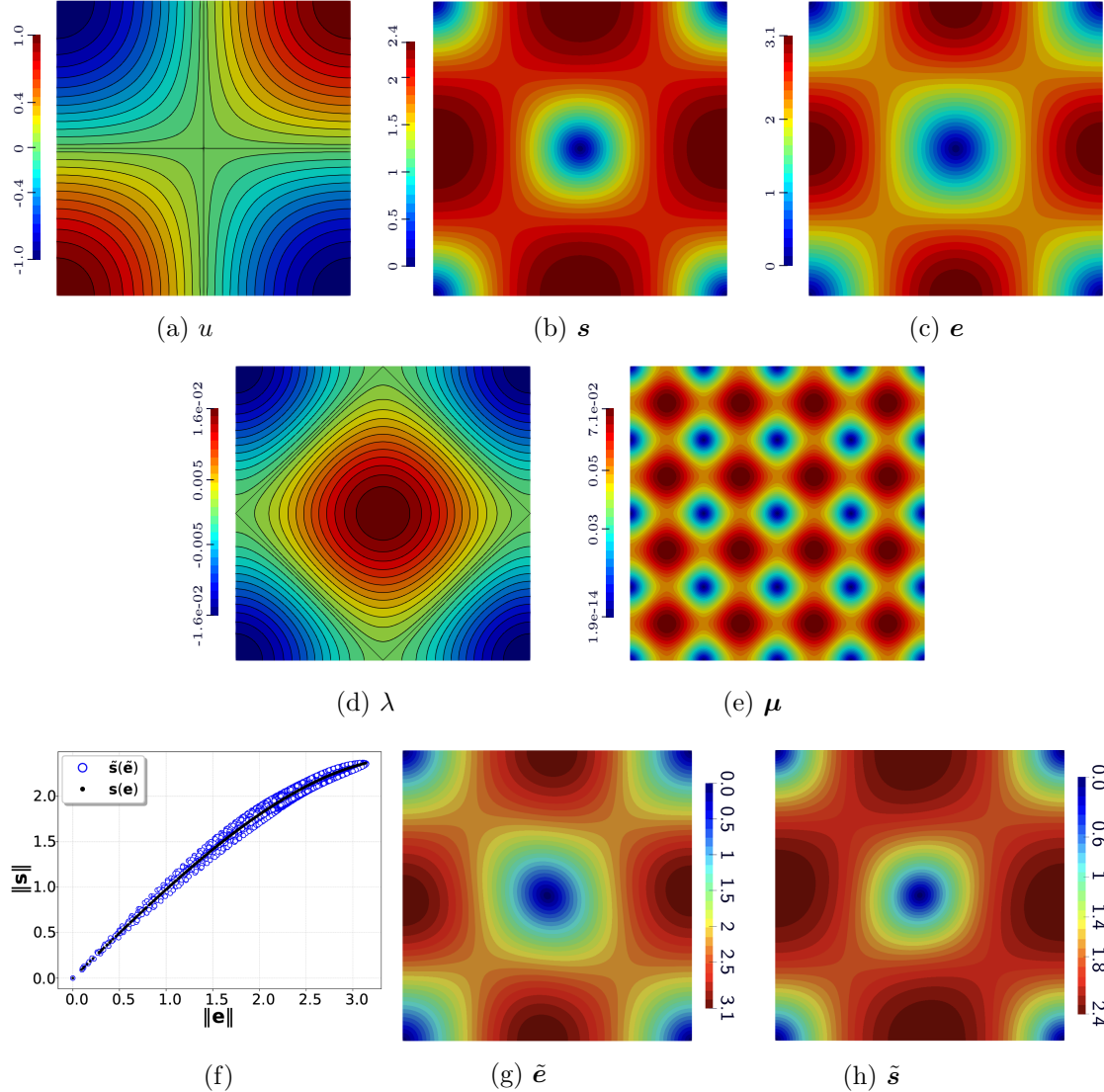
$$\tilde{s} = s - \kappa^{-1} \nabla \lambda \quad (25)$$

and the source term q is obtained by substituting the proposed fields into the balance equation

$$q = \zeta u + \nabla \cdot s. \quad (26)$$

Figure 1 (f) also displays, using blue dots, the sampled values of the magnitude of $\tilde{s}$ as a function of the magnitude of $\tilde{e}$, taken over the computational mesh. The computational domain Ω is chosen as the unit square $[0, 1]^2$, with the scale parameter and reaction coefficient both set to unity, i.e., $\kappa = 1$ and $\zeta = 1$.

Figure 1: Manufactured solution for the example 1 showing the primal and dual fields as well as the data fields $\tilde{s}$ and $\tilde{e}$.



Boundary conditions

Regarding boundary conditions on $\partial\Omega$, we consider two scenarios to ensure a fair comparison between the primal and dual formulations: In the first case, we assume the entire boundary is of Dirichlet type and strongly impose the boundary conditions on the scalar unknowns u and λ , i.e.,

$$\begin{cases} u = g_u & \text{on } \partial\Omega, \\ \lambda = g_\lambda & \text{on } \partial\Omega, \end{cases} \quad (27)$$

while in the latter case, we assume the entire boundary is of Neumann type and strongly impose the boundary conditions on the normal component of the vector unknowns $\mathbf{s}$ and $\boldsymbol{\mu}$, i.e.,

$$\begin{cases} \mathbf{s} \cdot \hat{\mathbf{n}} = h_s & \text{on } \partial\Omega, \\ \boldsymbol{\mu} \cdot \hat{\mathbf{n}} = h_\mu & \text{on } \partial\Omega, \end{cases} \quad (28)$$

where the values on the right-hand-side are evaluated from the manufactured fields at the locations of the corresponding degrees of freedom.

Choice of stabilization parameters

From Sections 3.1 and 3.2, we observe considerable flexibility in selecting the stabilization parameters for both the primal and dual formulations. To better guide the practical selection of stabilization parameters in the numerical experiments, we conducted a parametric study examining the sensitivity of the matrix condition number, $\mathcal{K}$, to variations in these parameters for both the primal and dual formulations with the identity projector operator (ASGS). From these numerical experiments, we observed that for the Primal-ASGS formulation, $\mathcal{K}$ remains nearly constant over a wide range of admissible values of k_u and k_λ , while it increases smoothly with k_s and k_μ . In addition, $\mathcal{K}$ was found to be essentially insensitive to k_e . For the Dual-ASGS formulation, on the other hand, $\mathcal{K}$ was found to be nearly insensitive to k_s and to decrease with increasing k_μ , while it increases with k_u , k_e , or k_λ . The same qualitative behavior was observed in reaction-dominated regimes (e.g., $\zeta = 1000$), where the magnitude of the condition number increases but its dependence on the stabilization parameters remains essentially unchanged. These results are omitted here solely for the sake of brevity.

In this example, for the primal formulation, the stabilization parameters are chosen as

$$k_u = k_\lambda = 0, \quad k_e = k_s = k_\mu = \frac{1}{8}.$$

This choice is convenient, as it reduces the number of parameters to be selected to just three. It also exploits the residual orthogonality property established earlier, ensuring that the results are independent of the choice of the projector operator P' . On the other hand, for the dual formulation we take

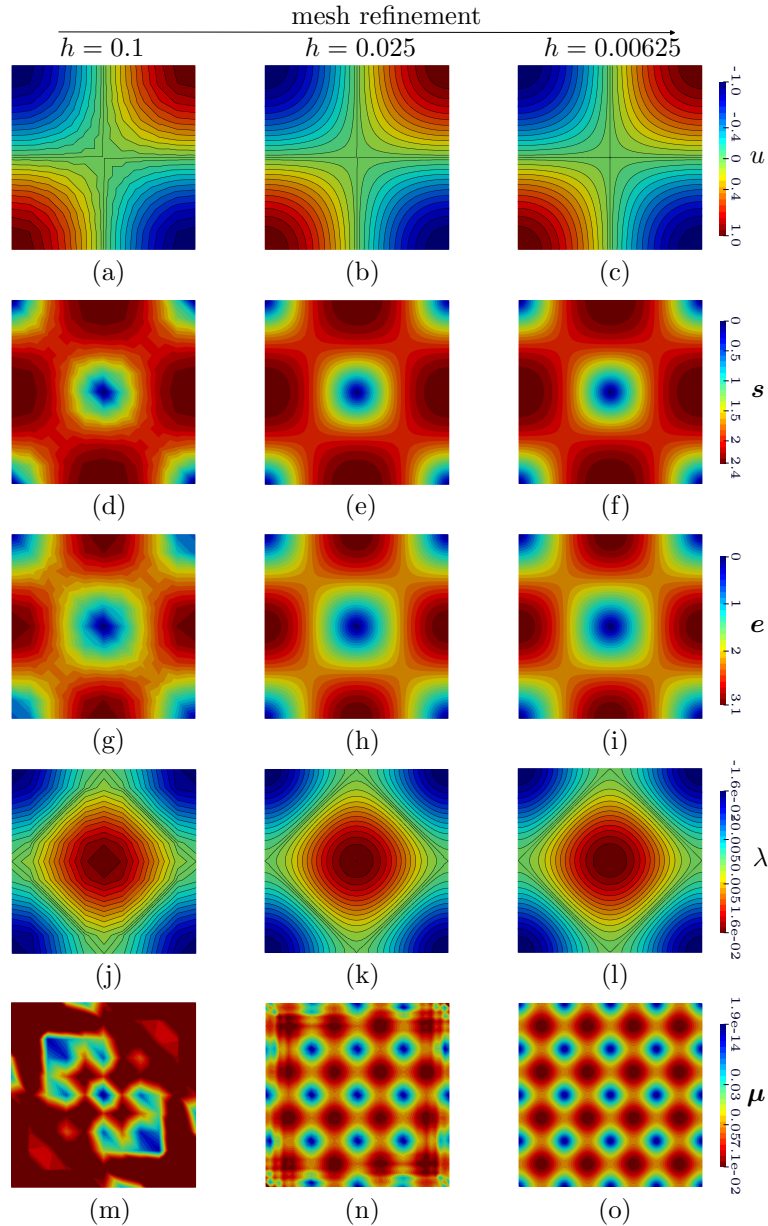
$$k_u = k_\lambda = k_e = k_\mu = \frac{1}{4}, \quad k_s = 1.$$

Although we could also have selected $k_\mu = k_s = 0$, the choice $k_\mu = \frac{1}{4}$ makes the selection of k_e mesh-independent. This is convenient for assessing the convergence rates of the method across the full range of mesh refinements tested. As a final remark, the convergence studies presented below employ linear and quadratic polynomial approximations (i.e., $k \in \{1, 2\}$). Therefore, in the following discussion, we consider the algebraic implementation of the methods, which we refer to as *Primal-ASGS* and *Dual-ASGS*.

We begin the numerical analysis with a qualitative comparison of the unknown fields across successive mesh refinements. Three refinement levels are examined: a very coarse mesh ($h = 0.1$), an intermediate mesh ($h = 0.025$), and a very fine mesh ($h = 0.00625$). The solutions for the *Primal-ASGS* and *Dual-ASGS* formulations are shown in Figures 2 and 3, respectively. These results should be compared against the reference solution in Figure 1. A eye-catching observation from these plots is that both formulations perform well on the finest mesh. The only minor exception is the field λ , which exhibits slight numerical artifacts near the boundaries, see Figure 3 (l). These artifacts, however, diminish with further refinement and do not impact the

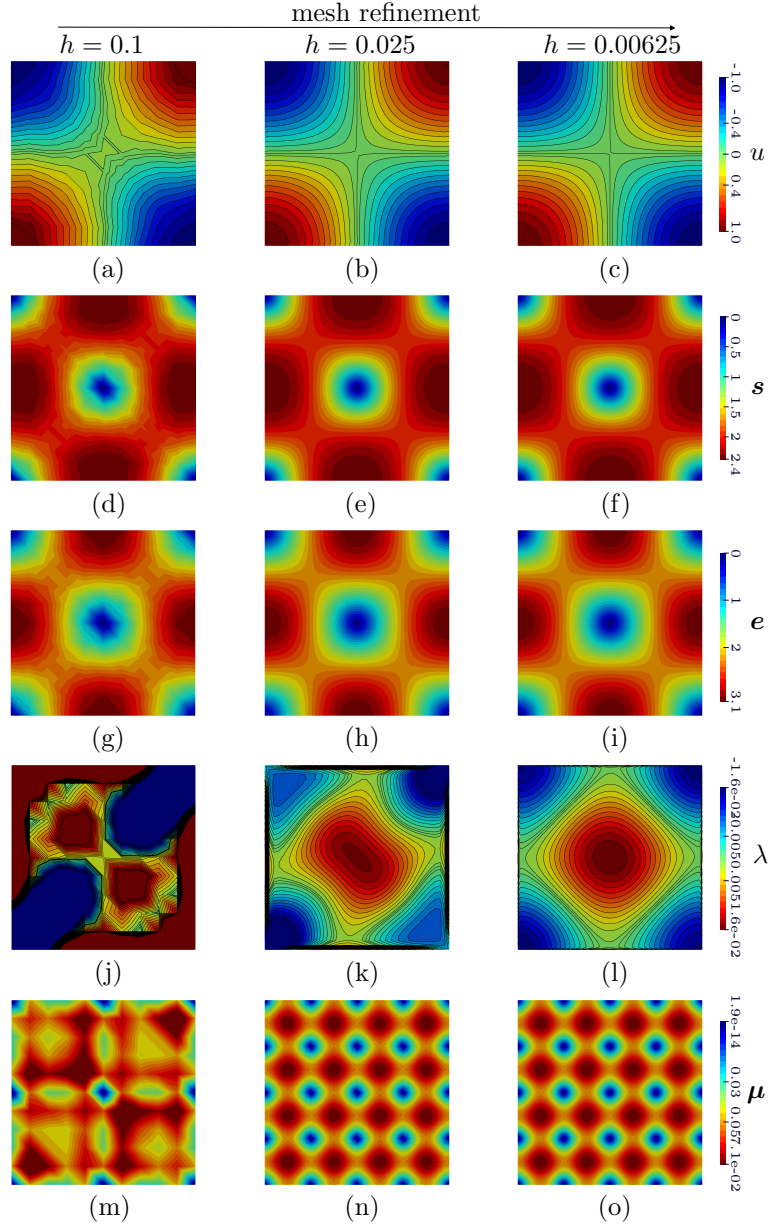
convergence of the method, as will be demonstrated. For the coarsest and intermediate mesh levels, the differences between the two formulations become more evident. While the primal formulation does a reasonable job of capturing the scalar fields, it struggles to accurately resolve particularly the μ field, Figure 2 (m–n). As expected, the situation is reversed for the dual formulation, which performs significantly better in capturing μ but exhibits poorer accuracy for the scalar fields, see Figure 3 (a, j–k), where this behavior is most evident). This qualitative assessment suggests that both formulations capture the field s well, a finding later corroborated by the $L^2(\Omega)$ -error norm in the numerical convergence analysis. However, the situation differs when the error is measured in the $H(\text{div}, \Omega)$ -norm, a distinction not apparent from this preliminary evaluation.

Figure 2: Contours of fields u , s , e , λ and μ for the *Primal-ASGS* formulation for different levels of mesh refinement.



We now numerically assess the convergence properties of the discrete formulations. To this

Figure 3: Contours of fields u , s , e , λ and μ for the *Dual-ASGS* for different levels of mesh refinement.



end, Figures 4 to 6 present the error norms for all fields using linear elements ($k = 1$), while Figures 7 to 9 show the corresponding errors obtained with quadratic elements ($k = 2$). Recall that we are also comparing our results against the different formulations recently evaluated in [4]: namely, a (stable) Galerkin-based *Natural* formulation in which $\nabla U_h \subset M_h$, $\nabla L_h \subset \kappa S_h$, and $E_h = M_h$; a *Fully stabilized* GLS formulation permitting equal-order interpolation; and finally, an *Unstabilized* equal-order plain Galerkin formulation.

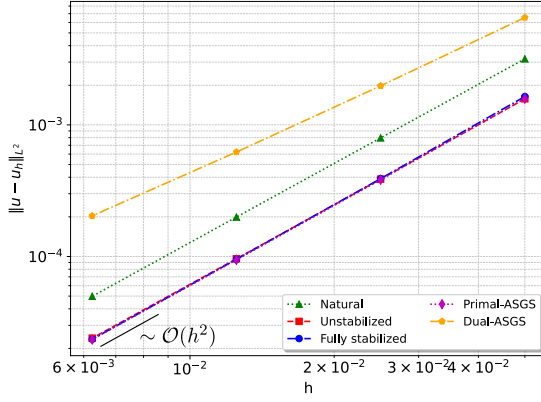
For linear elements, see Figures 4 to 6, all formulations achieve optimal convergence rates: second-order for the scalar fields in the $L^2(\Omega)$ -norm and first-order in the $H^1(\Omega)$ -norm. However, the errors for the *Dual-ASGS* formulation are significantly larger, highlighting the superior accuracy of the primal formulation to approximate the scalar fields. Note also that, in this case, only

the $L^2(\Omega)$ -norm is assessed for the *Dual-ASGS* formulation. For the vector fields, see Figure 5, all formulations under consideration, except for the *Natural* formulation, exhibit superlinear convergence in the $L^2(\Omega)$ -norm, nearly achieving $\mathcal{O}(h^2)$ despite a theoretical estimate of $\mathcal{O}(h)$. Results for the $H(\text{div}, \Omega)$ -norm are shown in Figure 6 for all formulations except the *Natural* one. Recall that only the dual formulation guarantees convergence for the vector fields $\mathbf{s}$ and $\boldsymbol{\mu}$ in the $H(\text{div}, \Omega)$ -norm. Nevertheless, for linear elements, all cases exhibit $\mathcal{O}(h)$ convergence, with very similar errors for $\mathbf{s}$. For $\boldsymbol{\mu}$ (yellow lines), the dual formulation yields significantly better results and is the only one achieving linear convergence.

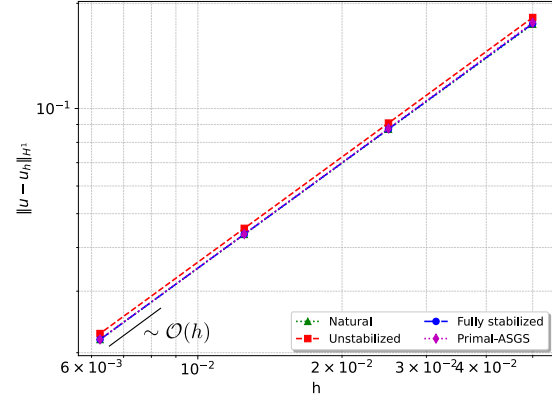
Moving to quadratic elements ($k = 2$), the primal formulation achieved $\mathcal{O}(h^2)$ convergence in the $H^1(\Omega)$ -norm for the scalar fields u and λ , which is the theoretical optimal rate, see Figure 7. Its results are essentially identical to those of the *Natural* and *Fully stabilized* GLS formulations. The *Unstabilized* formulation exhibited suboptimal convergence rates, underscoring the importance of stabilization, that was not evident for linear elements. For these fields in the $L^2(\Omega)$ -norm, the dual formulation shows the largest errors among all formulations since it only attains $\mathcal{O}(h^2)$ in contrast to the primal and GLS formulations that achieved $\mathcal{O}(h^3)$.

For the vector fields $\mathbf{s}$, $\boldsymbol{\mu}$ and $\mathbf{e}$, all formulations achieved second-order convergence in the $L^2(\Omega)$ -norm, see Figure 8. Notably, the *Fully stabilized* GLS formulation exhibited near-cubic ($\mathcal{O}(h^3)$) superconvergence for $\mathbf{s}$. This behavior may, however, be specific to this particular example. As anticipated, the primary advantage of the dual formulation is evident in the $H(\text{div}, \Omega)$ -norm errors, see Figure 9, where it is the only method to achieve second-order convergence for $\boldsymbol{\mu}$. Although theoretical estimates for $\mathbf{e}$ in this norm are unavailable, we note that the dual formulation yields significantly lower errors for $\mathbf{e}$ compared to all other methods. These numerical results confirm the well posedness of the discrete formulations and validate our implementation of both the primal and dual methods introduced earlier.

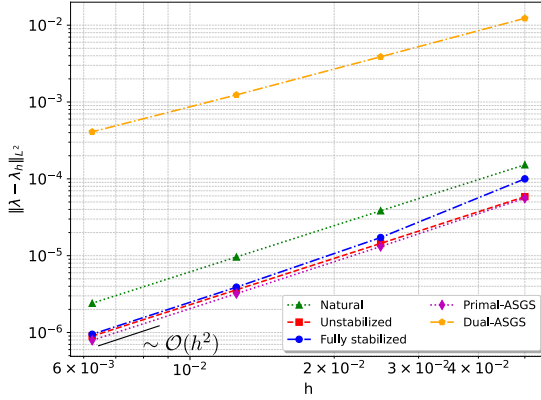
Figure 4: Convergence of scalar fields u and λ for $k = 1$. Figures (a–b) show the results for u and (c–d) for λ , considering all formulations: *Natural*, equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS*. Errors are evaluated in the $L^2(\Omega)$ - and $H^1(\Omega)$ -norms, as indicated. For the *Dual-ASGS* formulation, only the $L^2(\Omega)$ -norm is assessed.



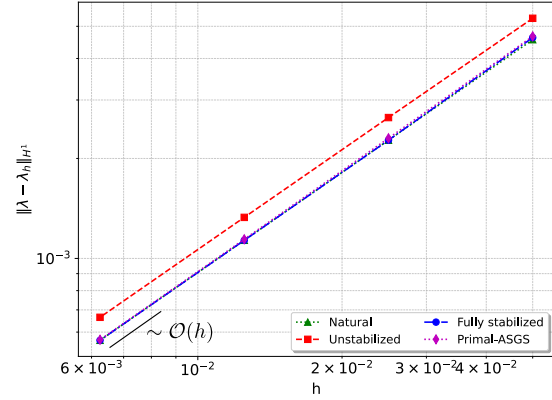
(a)



(b)

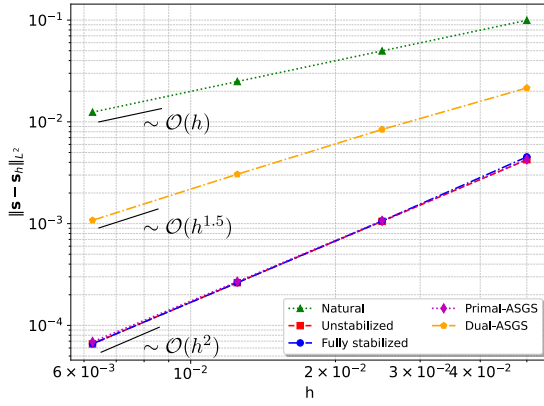


(c)

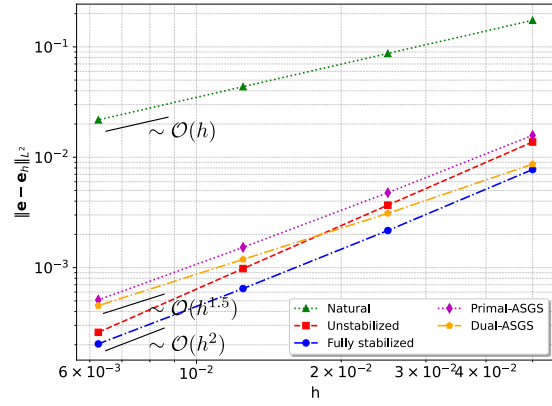


(d)

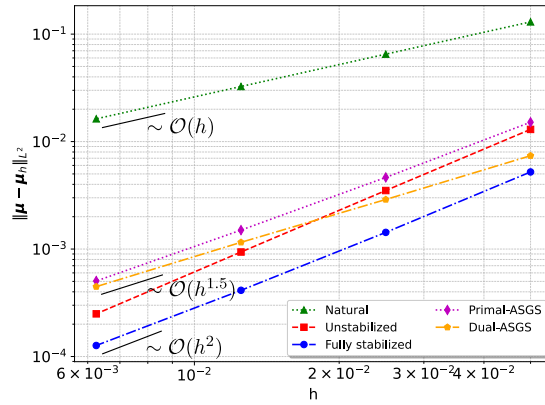
Figure 5: Convergence of vector fields $\mathbf{s}$, $\mathbf{e}$ and $\boldsymbol{\mu}$ with $k = 1$, evaluated in the $L^2(\Omega)$ -norm considering all formulations: *Natural*, equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS*.



(a)



(b)



(c)

Figure 6: Convergence of vector fields $\mathbf{s}$, $\mathbf{e}$ and $\boldsymbol{\mu}$ with $k = 1$, evaluated in the $H(\text{div}, \Omega)$ -norm for equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS* formulations.

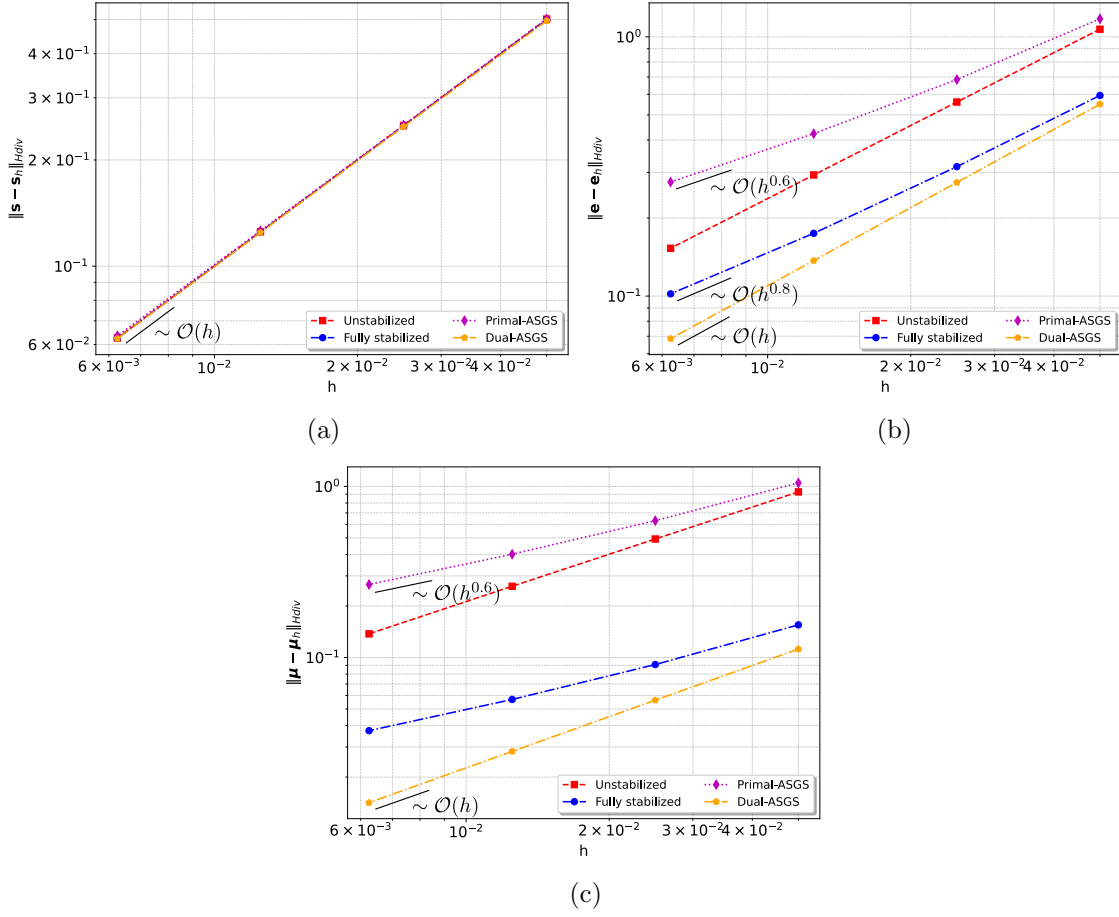


Figure 7: Convergence of scalar fields u and λ for $k = 2$. Figures (a–b) show the results for u , and (c–d) for λ , considering all formulations: *Natural*, equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS*. Errors are evaluated in the $L^2(\Omega)$ - and $H^1(\Omega)$ -norms, as indicated. For the *Dual-ASGS* formulation, only the $L^2(\Omega)$ -norm is assessed.

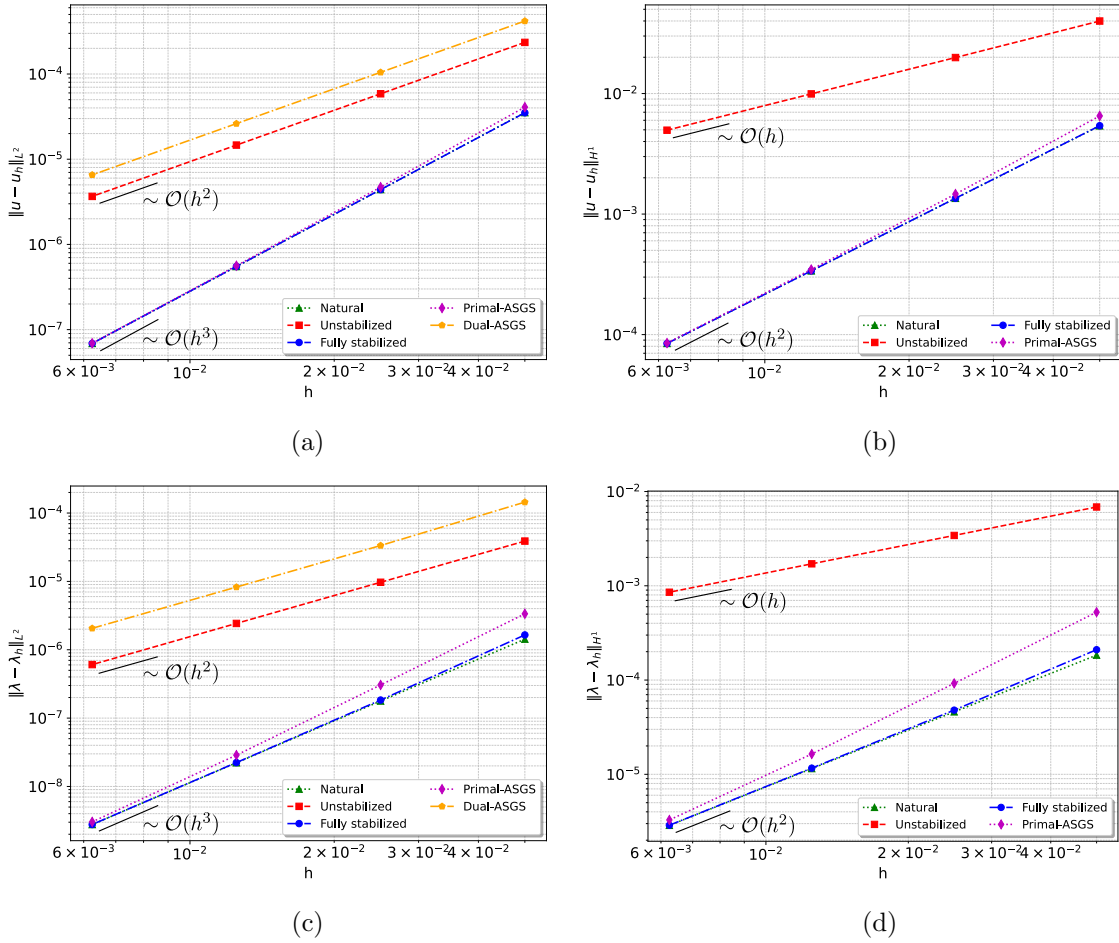


Figure 8: Convergence of vector fields $\mathbf{s}$, $\mathbf{e}$ and $\boldsymbol{\mu}$ with $k = 2$, evaluated in the $L^2(\Omega)$ -norm considering all formulations: *Natural*, equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS*.

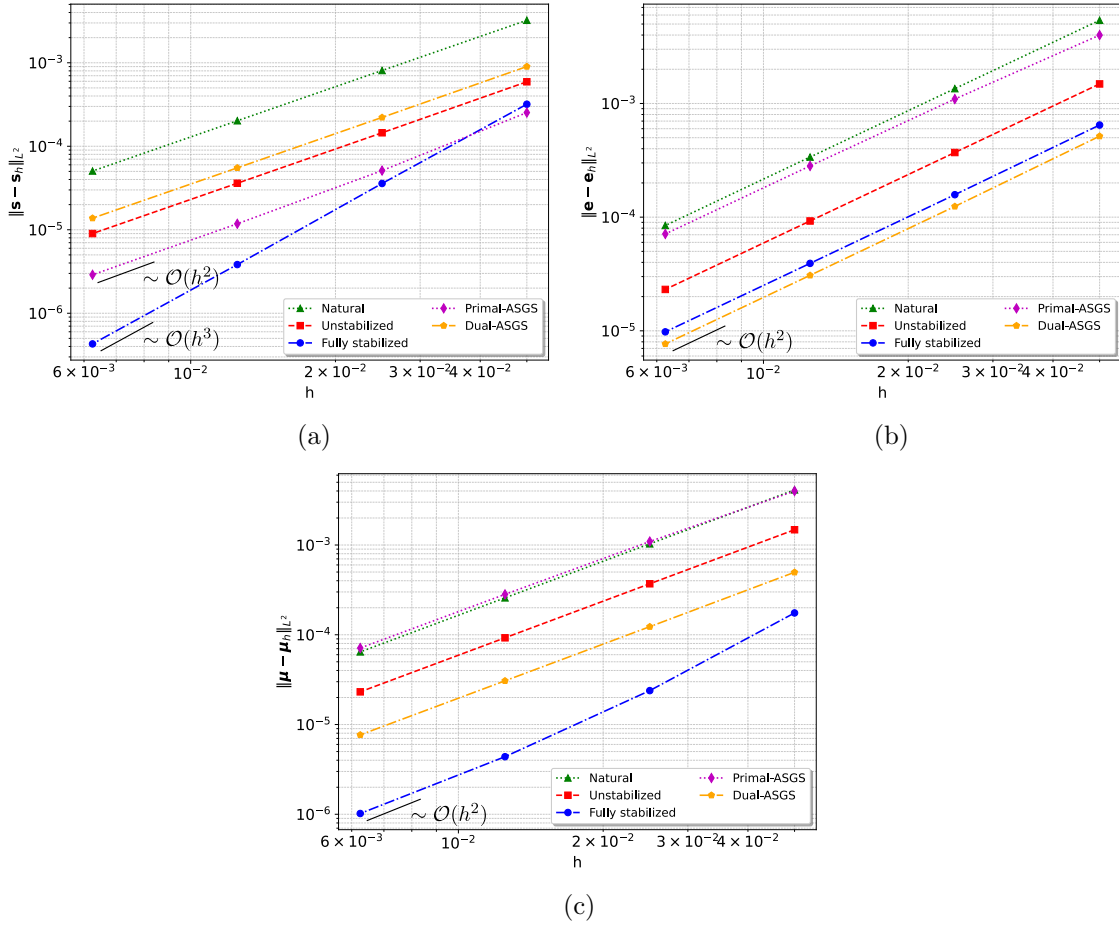
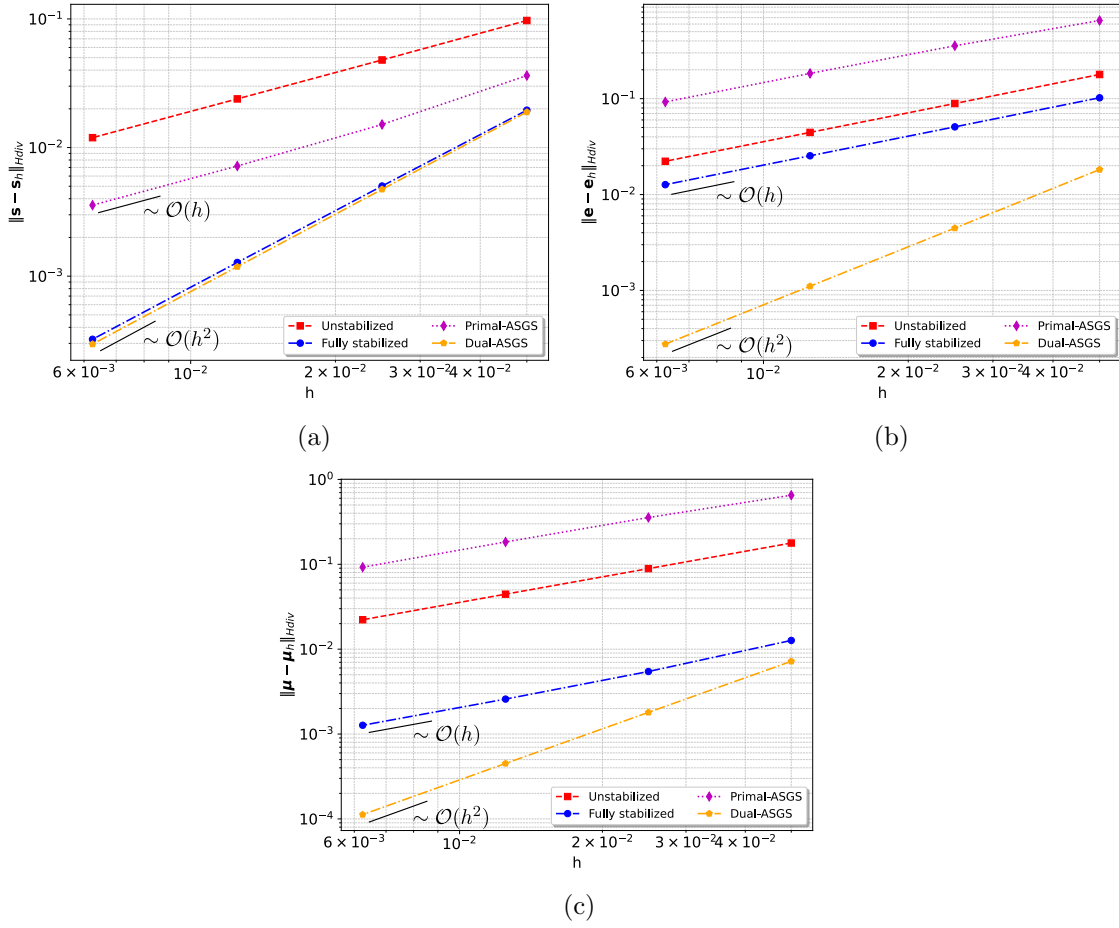


Figure 9: Convergence of vector fields $\mathbf{s}$, $\mathbf{e}$ and $\boldsymbol{\mu}$ with $k = 2$, evaluated in the $H(\text{div}, \Omega)$ -norm for equal-order *Unstabilized*, equal-order *Fully stabilized*, *Primal-ASGS* and *Dual-ASGS* formulations.



4.2 Example 2: Performance assessment in the presence of large gradients

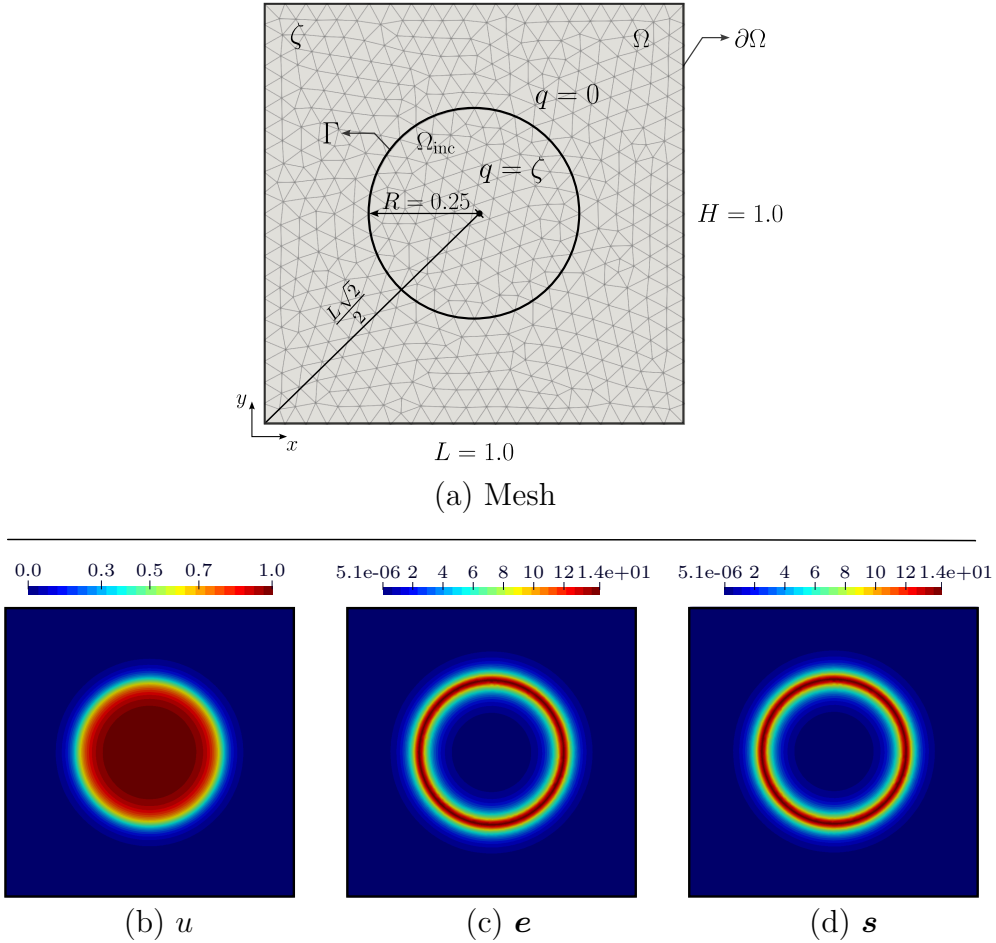
This final example aims to further assess the performance of the primal and dual formulations on a more challenging problem characterized by large internal gradients, typical of reaction-diffusion systems. We introduce a standard problem that serves a dual purpose: it will generate the synthetic data fields $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}$ required for the data-driven formulation while also providing a reference solution for comparative evaluation.

Let $\Omega = (0, 1) \times (0, 1)$ be the computational domain. We consider a circular inclusion immersed in Ω , centered at $\mathbf{x}_c = (\frac{1}{2}, \frac{1}{2})$ and defined by

$$\Omega_{\text{inc}} = \{ \mathbf{x} \in \Omega : \|\mathbf{x} - \mathbf{x}_c\| < R \}, \quad R = 0.25,$$

whose boundary is denoted by Γ . Figure 10 (a) presents a schematic of the computational domain along with the finite element mesh. As shown, the interface Γ is aligned with the element edges.

Figure 10: Computational domain and the associated finite element mesh. Figures (b–d) illustrate the corresponding reference (exact) physical fields: the potential u , the flux $\mathbf{s}$, and the gradient $\mathbf{e}$, respectively.



The standard problem therefore consists of solving the following diffusion–reaction equation: Find $u : \Omega \rightarrow \mathbb{R}$ such that

$$\begin{cases} -\nabla \cdot (\gamma \nabla u) + \zeta u &= q & \text{in } \Omega, \\ \nabla u \cdot \hat{\mathbf{n}} &= g & \text{on } \partial\Omega, \end{cases} \quad (29)$$

where the reaction coefficient $\zeta > 0$ is uniform and the source term q is piecewise constant, i.e.,

$$q(\mathbf{x}) = \begin{cases} \zeta, & \mathbf{x} \in \Omega_{\text{inc}}, \\ 0, & \mathbf{x} \in \Omega \setminus \Omega_{\text{inc}}. \end{cases}$$

Assuming uniform conductivity γ , symmetry implies the solution depends only on the radial coordinate, i.e., $u = u(r)$ with $r = \sqrt{(x - \frac{1}{2})^2 + (y - \frac{1}{2})^2}$. Under this assumption, the problem reduces to an ordinary differential equation in r , whose solution is

$$u(r) = \begin{cases} 1 - \alpha R K_1(\alpha R) I_0(\alpha r), & 0 \leq r \leq R, \\ \alpha R I_1(\alpha R) K_0(\alpha r), & r > R, \end{cases} \quad \alpha = \sqrt{\frac{\zeta}{\gamma}}. \quad (30)$$

Here, I_0 and I_1 are the modified Bessel functions of the first kind (orders 0 and 1, respectively), and K_0 and K_1 are the corresponding modified Bessel functions of the second kind. Figure 10 (b) shows the exact potential field u , while Figures 10 (c–d) illustrate the corresponding gradient and flux fields, $\mathbf{e}$ and $\mathbf{s}$, respectively. Several observations are noteworthy: first, note that this solution satisfies the following physical conditions across the interface Γ :

$$\text{at } r = R : \quad \begin{cases} u(R^+) = u(R^-), \\ \gamma u'(R^+) = \gamma u'(R^-), \end{cases}$$

ensuring continuity of u and of the normal diffusive flux. Second, the solution decays to zero at infinity, and its boundedness at the origin implies $u'(0) = 0$. Third, the sharpness of the transition region is controlled by adjusting the ratio α of the reaction to the diffusion coefficient. In our numerical experiments, we set $\gamma = 1$ and $\zeta = 1000$, which yields a transition region with a characteristic length of approximately 0.03. This behavior is clearly illustrated in Figure 10 (b). Also, Figures 10 (c–d) show the very sharp gradients in the vector fields $\mathbf{e}$ and $\mathbf{s}$.

Before presenting the numerical results, we specify the following computational setup:

- (i) A finite element mesh with characteristic size $h = 0.025$.
- (ii) Stabilization parameters set as

$$k_u = k_\lambda = 1, \quad k_s = k_e = k_\mu = \frac{1}{8}$$

for the *Primal-ASGS* formulation, and

$$k_u = k_\lambda = k_e = \frac{1}{4}, \quad k_s = 1, \quad k_\mu = \frac{1}{8}$$

for the *Dual-ASGS* formulation. It is worth mentioning that, although the same set of stabilization parameters used in Example 1 could also be employed in this second experiment, we instead explore different parameter choices, always within the admissible ranges.

- (iii) Strong imposition of boundary conditions: for the primal formulation, $u = 0$ and $\lambda = 0$; for the dual formulation, $\mathbf{s} \cdot \hat{\mathbf{n}} = 0$ and $\boldsymbol{\mu} \cdot \hat{\mathbf{n}} = 0$. These homogenous values are safe since the solution has already sufficiently decayed at the boundary.
- (iv) The data fields $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}$ are defined as elementwise constants. They are generated by sampling a finite element solution of problem (29), which is computed using a standard primal Galerkin formulation with continuous piecewise linear elements on the same mesh used for the data-driven problem. Unlike the previous example, where data fields were treated as continuously available, here the data are provided only at discrete locations, that is, one data-pair per element. This brings the test closer to realistic data-driven mechanics applications.

Turning to the numerical results, Figure 11 illustrates the physical fields u and $\mathbf{s}$ obtained using both the *Primal*- and *Dual*-ASGS formulations. Figures 11 (a–b) demonstrate that the primal formulation provides a superior approximation for the scalar potential field, whereas the dual formulation shows considerable distortion compared to the reference solution plotted in Figure 10. On the other hand, the dual formulation approximates the flux field more effectively. This is evident in Figures 11 (c–d), where the largest distortions are observed in the primal result. Unlike the previous test, where improvements were confined to the $H(\text{div}, \Omega)$ -norm of the error, here the advantage of the dual formulation is already visible by simple inspection. This trend is further evident in Figures 12 and 13, which show profiles of u and $\mathbf{s}$ along the domain diagonal and along circumferential curves at different radii indicated in Figure 13 (a), namely, $r = 0.1, 0.2$ and 0.4 . Figure 12 (a) and the top part of Figure 13 reveal that the primal formulation (magenta lines) clearly outperforms the dual formulation (green lines) in capturing u . Conversely, Figure 12 (b) and the bottom part of Figure 13 reveal the opposite situation for $\mathbf{s}$. A closer view of these differences is provided in the insets of Figure 12. See also the root mean square error reported for the circumferential profiles for each formulation in Figure 13. For further clarity, it is instructive to examine the pointwise error distribution across the computational domain. Figure 14 presents these error fields: parts (a–b) show the error in $u - u_h$ while parts (c–d) show the error in the flux $\mathbf{s} - \mathbf{s}_h$. Results for the *Primal*-ASGS formulation appear in the left column; those for the *Dual*-ASGS formulation appear on the right.

Numerical results in the presence of noisy data

This experiment is well suited for examining the methods' behavior under conditions more representative of data-driven applications. To this end, we introduce an additional ingredient: incorporating noise into the data fields $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}$, which were originally sampled from a finite element solution of problem (29). Specifically, we perturb these fields by adding uniform noise:

$$\tilde{\mathbf{e}}^\delta = \tilde{\mathbf{e}} + \delta \boldsymbol{\xi}_e, \quad \tilde{\mathbf{s}}^\delta = \tilde{\mathbf{s}} + \delta \boldsymbol{\xi}_s, \quad (31)$$

where $\delta > 0$ represents the noise level and $\boldsymbol{\xi}_e$ and $\boldsymbol{\xi}_s$ are random vector fields with components independently drawn from a uniform distribution:

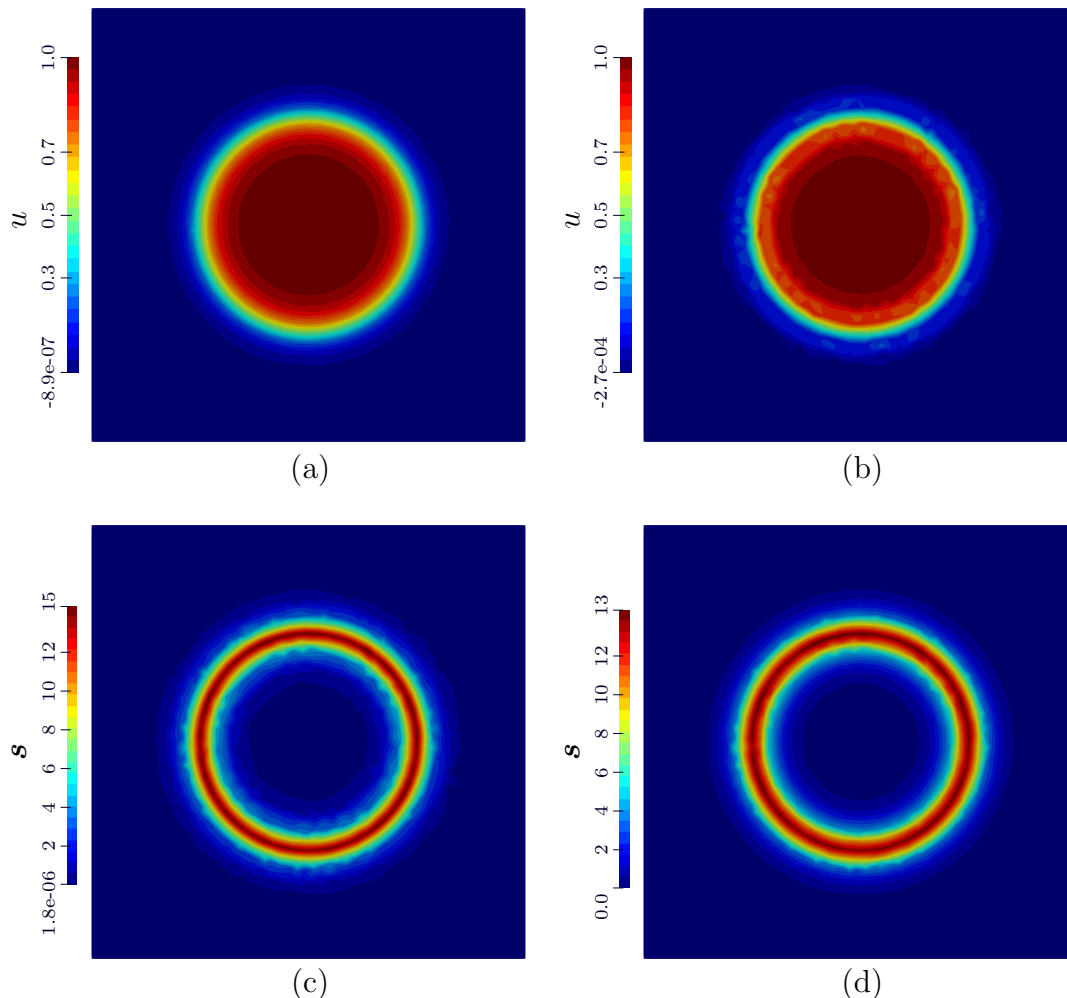
$$\boldsymbol{\xi}_e, \boldsymbol{\xi}_s \sim \mathcal{U}(-1, 1). \quad (32)$$

This corresponds to adding a uniform perturbation in the interval $[-\delta, \delta]$ to each component of the gradient and flux fields. Figure 15 show the data field $\tilde{\mathbf{s}}^\delta$ for different noise levels, namely, (a) $\delta = 0.25$, (b) $\delta = 0.5$ and (c) $\delta = 1$. Figure 16 (a–c) further illustrates both, the noise-free data pairs $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$, which form a straight line, since the flux and the gradient are related by a linear constitutive law, and the corresponding perturbed data pairs sampled from the finite element mesh. The effect of increasing noise on the numerical approximation of u is shown in Figure 16 (d–i) for both the *Primal*- and *Dual*-ASGS formulations. Corresponding results for $\mathbf{s}$ appear in Figure 17. These results reinforce the trends observed earlier: the primal formulation remains more robust for approximating the scalar field, while the dual formulation continues to better capture the vector field, even at high noise levels as seen in Figure 16 (f) and Figure 17 (f).

The presence of noise introduces an additional consideration: the loss of thermodynamic consistency between $\mathbf{e}$ and $\mathbf{s}$, an effect present in both formulations. This is visualized in Figure 18 via contours of the scalar field $\mathbf{s} \cdot \mathbf{e}$. The onset of violation is indicated by the zero contour, shown as a black line. The data near the origin exhibit a larger relative noise, leading to a higher concentration of elements violating the thermodynamic consistency in regions where the solution gradient is low, that is, far from the interface Γ . Nothing in the formulation itself precludes these violations, but their occurrence underscores a central issue for reliable data-driven computations in practical applications. The development of numerical methods capable of filtering potential experimental errors from the data fields and ensuring thermodynamic consistency is essential but left for future work.

Finally, we verified through an additional parametric study that moderate variations of the stabilization parameters within their admissible ranges do not significantly affect the results, even

Figure 11: Figures (a–b) illustrate the field u for the *Primal*- and *Dual-ASGS* formulations. Likewise, Figures (c–d) show the field s for the *Primal*- and *Dual-ASGS* formulations.



in the presence of noisy data. This indicates that the proposed formulation is robust with respect to the choice of stabilization parameters in the context of perturbed data. Therefore, we retain the previously used set of parameters.

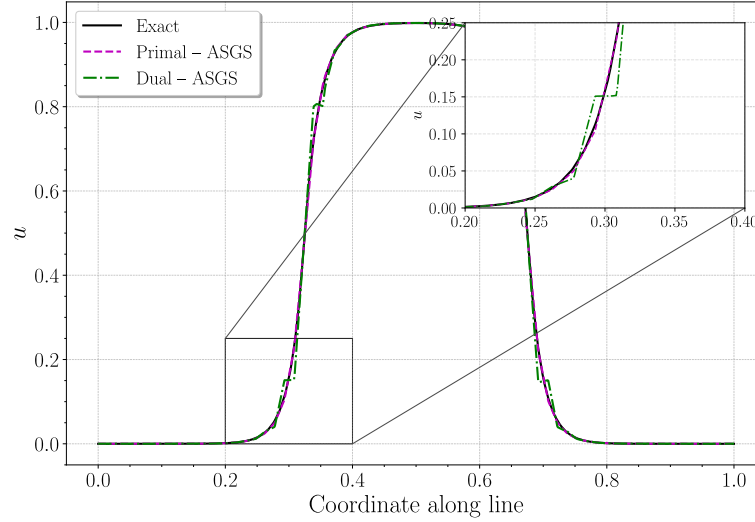
5 Conclusion

In this work, we have proposed a finite element approximation of a PDE-constrained optimization problem arising in reaction-diffusion problems within the framework of Data-Driven Computational Mechanics (DDCM) based on the Variational Multiscale (VMS) concept.

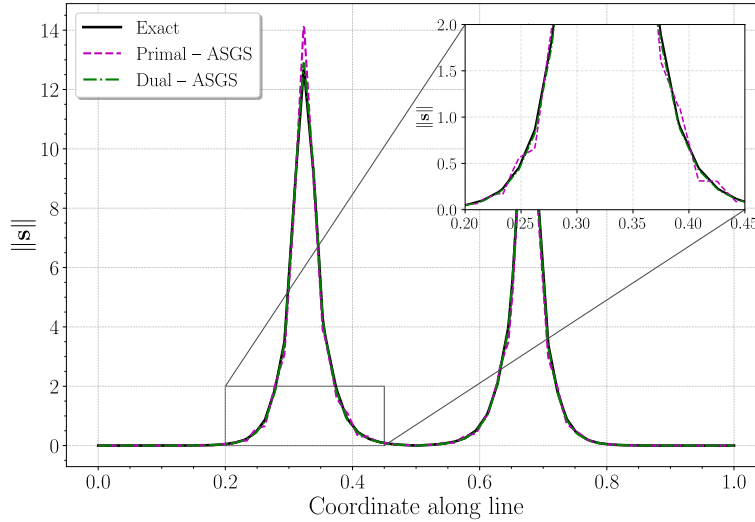
We have established that both the primal and dual forms of the optimality conditions are well-posed in the continuous setting. This has been achieved by proving that the necessary requirements for mixed saddle-point problems are satisfied.

Regarding the finite element approximation, to overcome the limitations of standard Galerkin methods—which would require complex, problem-specific inf-sup conditions for the finite element spaces—we adopted the VMS framework. This approach allowed us to derive stable and consistent finite element approximations by systematically modeling the sub-grid scales. Specifically, we established the well-posedness of both the Algebraic Sub-Grid Scale (ASGS) and the Orthogonal

Figure 12: Figure (a) displays the u profile along the coordinate diagonal for the exact reference solution, together with the corresponding profiles obtained from the *Primal*- and *Dual*-ASGS formulations. Likewise, Figure (b) displays the s profile along the coordinate diagonal for the exact reference solution and for the *Primal*- and *Dual*-ASGS formulations.



(a)



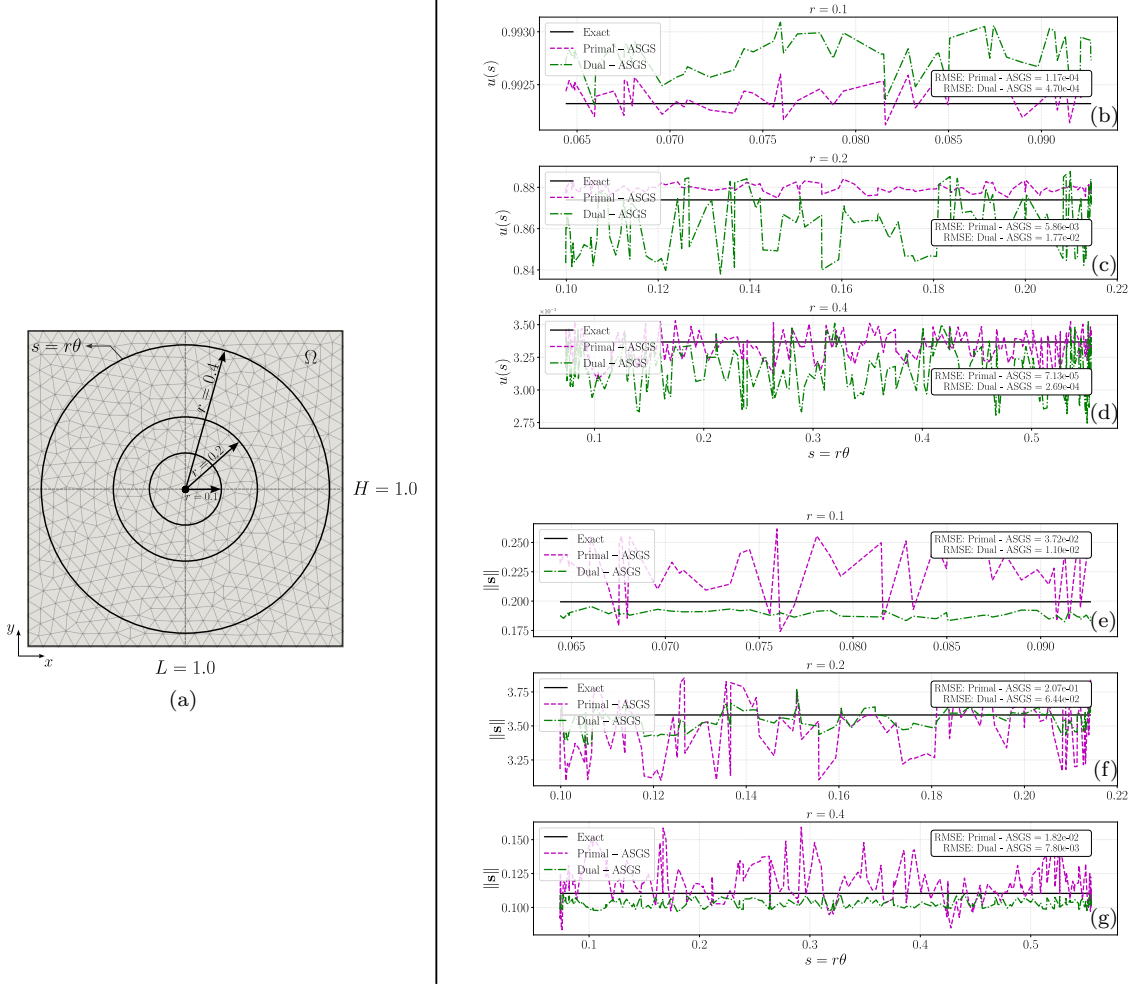
(b)

Sub-Grid Scale (OSGS) methods.

For the VMS approximation, we have demonstrated that one can transition between the primal and the dual form of the problem simply by adjusting the design of the stabilization parameters, providing a unified view of the problem's discrete structure.

The numerical results presented have provided a robust qualitative and comparative assessment, confirming the stability and consistency of our stabilized finite element schemes. These

Figure 13: Figures (b–d) present the u profiles along three circumferences of radius $r = 0.1$, $r = 0.2$ and $r = 0.4$, illustrated schematically in Figure (a). Each plot compares the exact reference solution with the corresponding profiles obtained from the *Primal*- and *Dual*-ASGS formulations. Likewise, Figures (e–g) display the $\|\mathbf{s}\|$ profiles along the same radii for both the exact reference solution and the *Primal*- and *Dual*-ASGS formulations.



results underscore the effectiveness of our approach in handling the inherent challenges of data-driven models.

While this work represents a major improvement in applying DDCM to diffusion-reaction problems, we acknowledge certain simplifications. In this study, we linearized the problem by omitting the explicit enforcement of the Second Law of Thermodynamics, though we verified its fulfillment a posteriori. We believe that the inclusion of this law directly into future optimization schemes will lead to even more robust formulations.

At the numerical level, we have focused on continuous finite element functions and quasi-uniform partitions; the extension of this VMS approach to discontinuous spaces and arbitrary mesh refinements remains a promising avenue for future research. Discontinuous (but conforming) interpolations could be considered by introducing the concept of sub-grid scales on the element boundaries [11], while non-uniform finite refinements could be addressed by using the techniques in [9]. Likewise, hp -convergence in the norm of the continuous problem and improved convergence estimates in $L^2(\Omega)$ represent important topics for subsequent research.

Figure 14: Figures (a–b) present the error field $u - u_h$ for the *Primal*- and *Dual-ASGS* formulations. In the same way, Figures (c–d) present the error field $\mathbf{s} - \mathbf{s}_h$ for the *Primal*- and *Dual-ASGS* formulations.

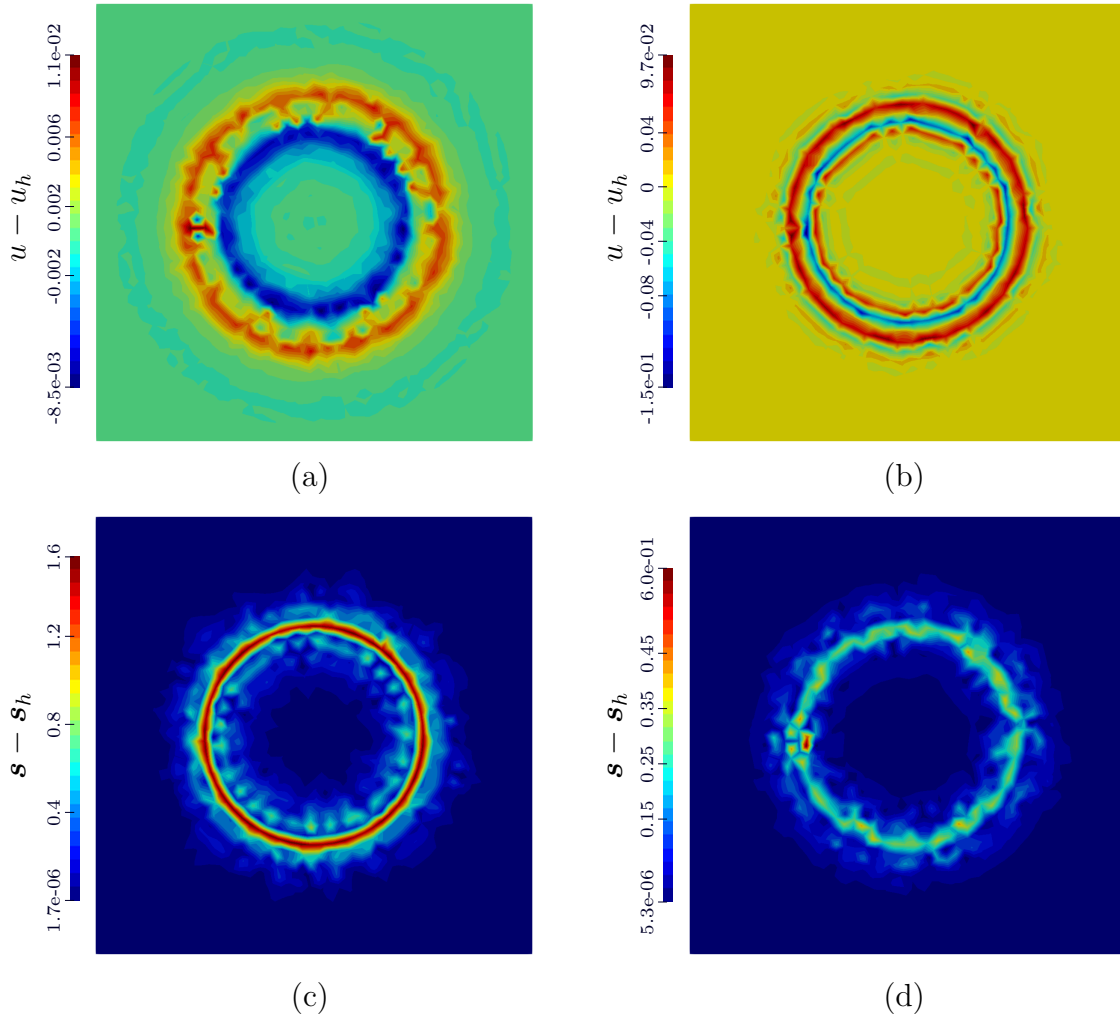


Figure 15: Data field $\tilde{s}$ sampled from a finite element solution of problem (29) to which uniform noise has been added. Different noise levels are considered: (a) $\delta = 0.25$, (b) $\delta = 0.5$ and (c) $\delta = 1$.

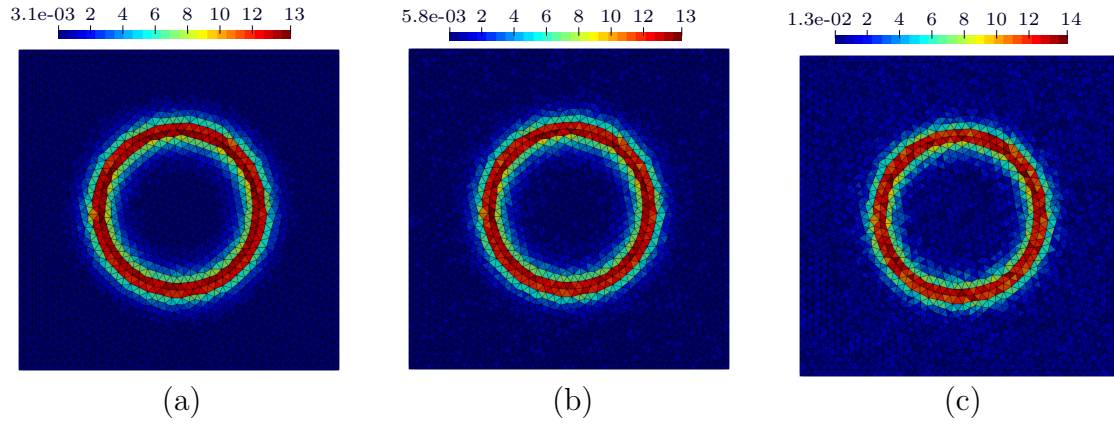


Figure 16: Figures (a–c) present the noise-free data pairs $(\tilde{\mathbf{e}}, \tilde{\mathbf{s}})$ together with the corresponding noisy pairs for three different values of δ , namely $\delta = 0.25$, $\delta = 0.5$, and $\delta = 1.0$, respectively. Figures (d–f) show the corresponding approximations of the scalar field u obtained with the *Primal-ASGS* formulation. Similarly, Figures (g–i) display the approximations of the same field computed using the *Dual-ASGS* formulation.

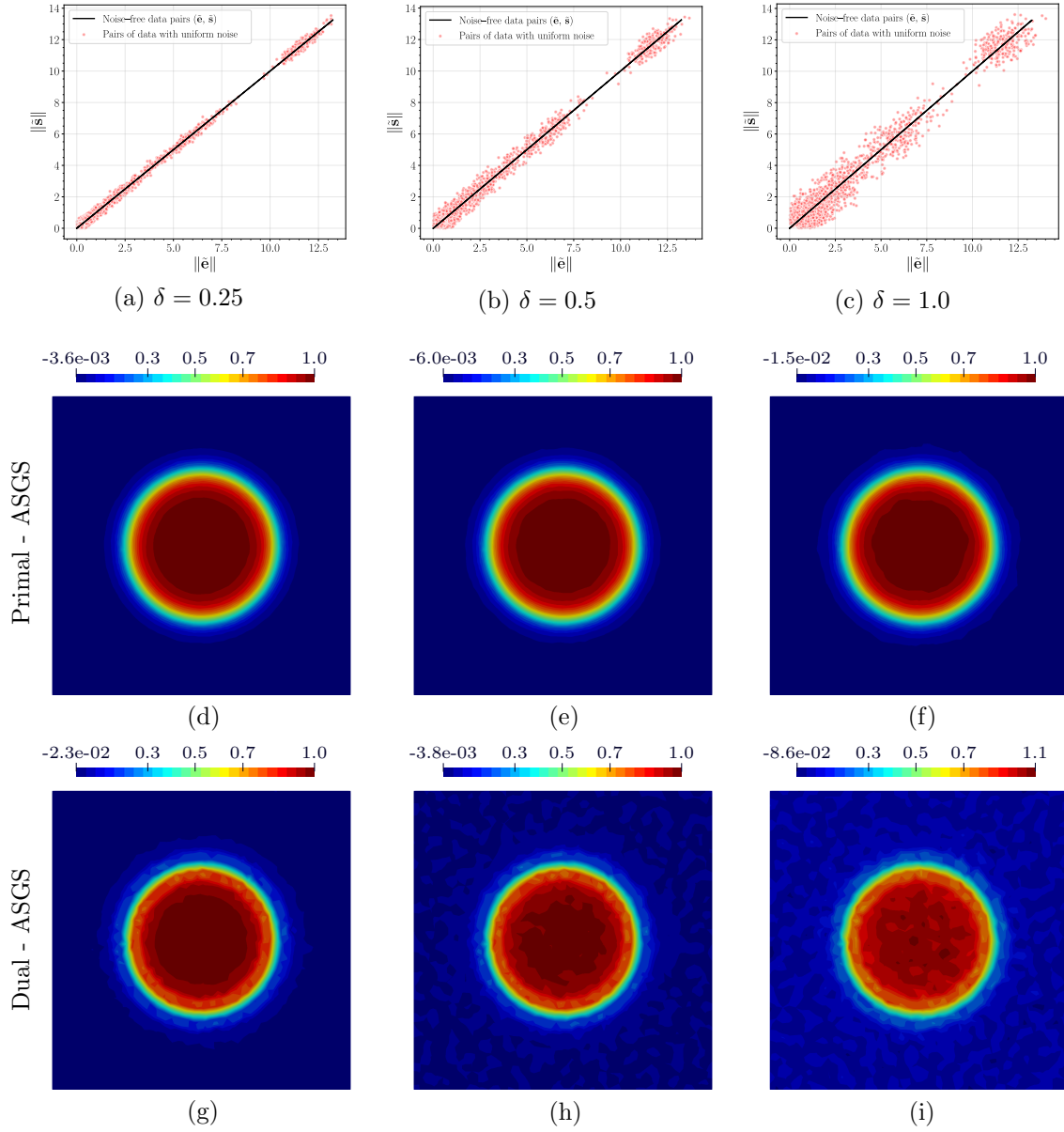


Figure 17: Approximations of the vector field $\mathbf{s}$ obtained with the *Primal-ASGS* formulation (parts (a)–(c)) and with the *Dual-ASGS* formulation (parts (d)–(f)): $\delta = 0.25$ (left), $\delta = 0.5$ (middle), $\delta = 1$ (right)

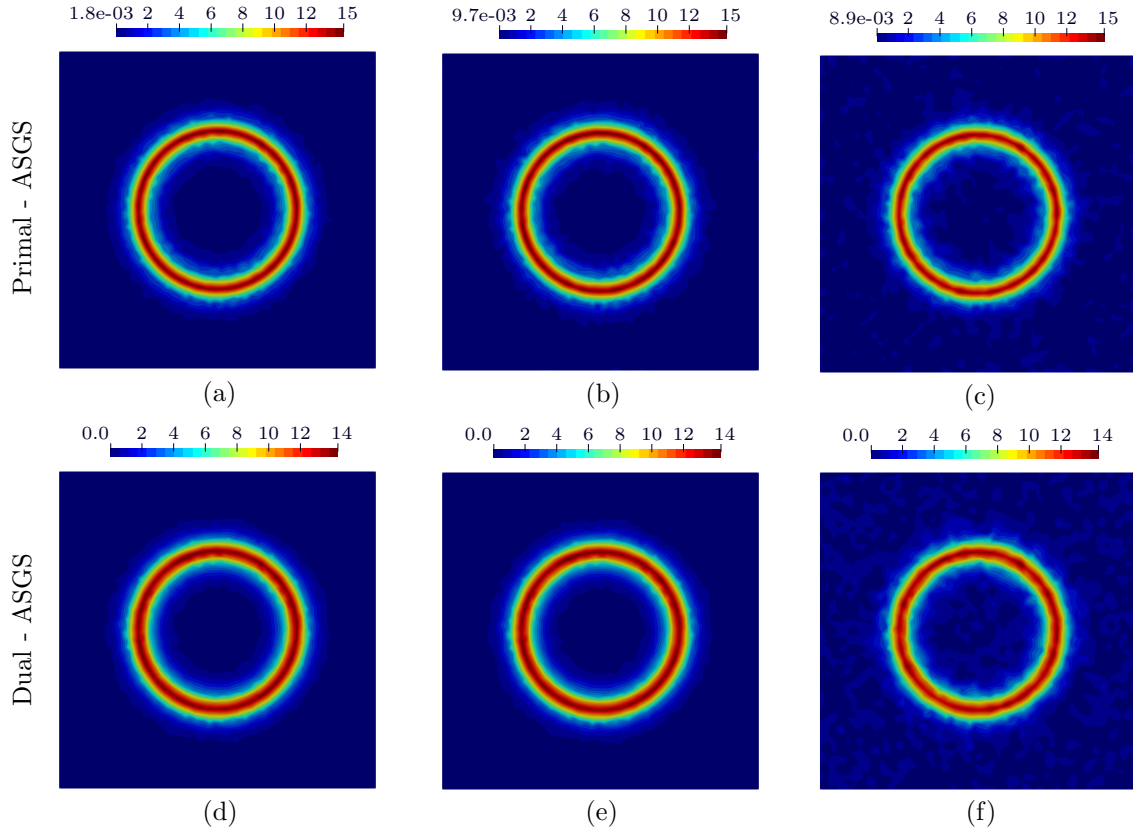
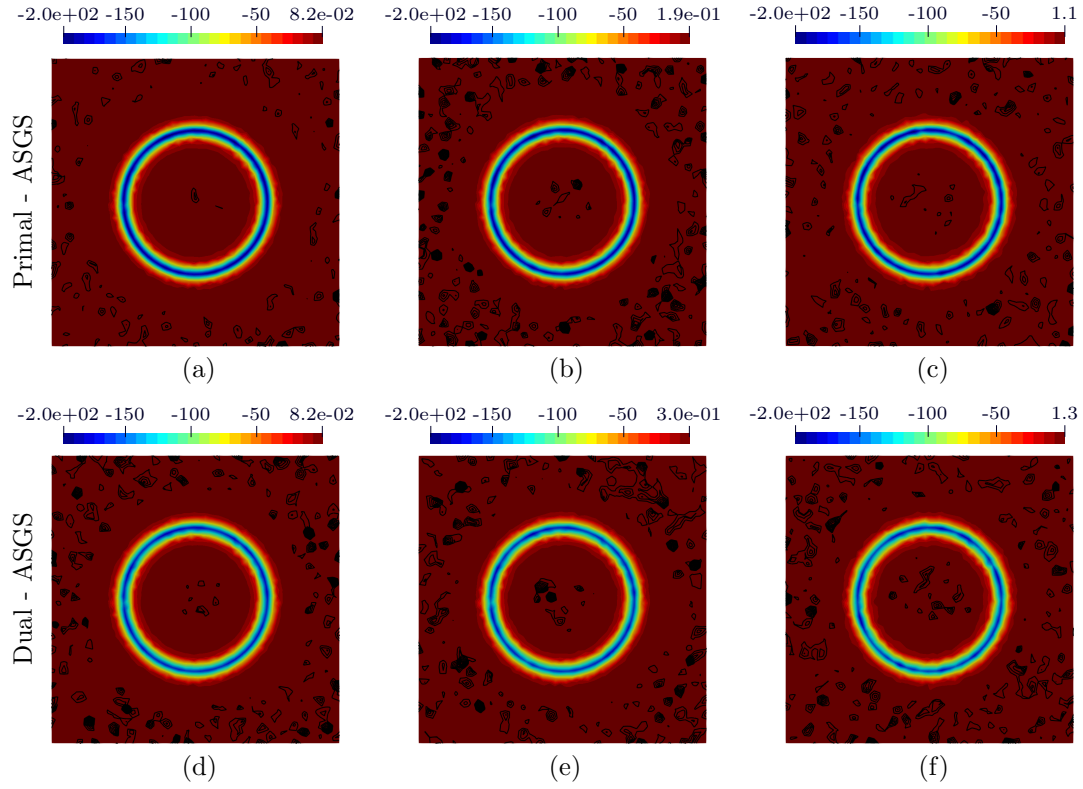


Figure 18: Contours of $\mathbf{s} \cdot \mathbf{e}$ to illustrate the violation of thermodynamic consistency across the domain for different levels of noise in the data fields $\tilde{\mathbf{e}}$ and $\tilde{\mathbf{s}}$: $\delta = 0.25$ (left), $\delta = 0.5$ (middle), $\delta = 1$ (right).



CRedit authorship contribution statement

Ramon Codina: Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization;

Roberto F. Ausas: Writing – original draft, Software, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization;

Pedro B. Bazon: Writing – original draft, Software, Validation, Visualization, Investigation;

Cristian G. Gebhardt: Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization;

Acknowledgments

RC gratefully acknowledges the support received from the ICREA Acadèmia Program, from the Catalan Government.

RFA thankfully acknowledges financial support from the São Paulo Research Foundation (FAPESP) under grants 2013/07375-0 and 2025/00975-9, and from the National Council for Scientific and Technological Development - CNPq under grant 308704/2022-3.

PBB gratefully acknowledges financial support from the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES), Finance Code 88887.975792/2024-00, and the São Paulo Research Foundation (FAPESP) under Grant No. 2025/27460-9.

CGG gratefully acknowledges the financial support from the European Research Council through the ERC Consolidator Grant “DATA-DRIVEN OFFSHORE” (Project ID 101083157).

References

- [1] S. Badia and R. Codina. Unified stabilized finite element formulations for the Stokes and the Darcy problems. *SIAM Journal on Numerical Analysis*, 47:1971–2000, 2009.
- [2] S. Badia and R. Codina. Stabilized continuous and discontinuous Galerkin techniques for Darcy flow. *Computer Methods in Applied Mechanics and Engineering*, 199:1654–1667, 2010.
- [3] S. Badia, R. Codina, and H. Espinoza. Stability, convergence, and accuracy of stabilized finite element methods for the wave equation in mixed form. *SIAM Journal on Numerical Analysis*, 52:1729–1752, 2014.
- [4] P. B. Bazon, C. G. Gebhardt, G. C. Buscaglia, and R. F. Ausas. Numerical approximation of a PDE-constrained Optimization problem that appears in Data-Driven Computational Mechanics. *International Journal for Numerical Methods in Engineering*, 127:e70283, 2026.
- [5] P. Bochev and C. Dohrmann. A computational study of stabilized, low-order C^0 finite element approximations of Darcy equations. *Computational Mechanics*, 38:323–333, 2006.
- [6] E. Burman, P. Hansbo, and M. Larson. The augmented Lagrangian method as a framework for stabilised methods in computational mechanics. *Archives of Computational Methods in Engineering*, 30:2579–2604, 2023.
- [7] R. Codina. Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 190:1579–1599, 2000.
- [8] R. Codina. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer Methods in Applied Mechanics and Engineering*, 191:4295–4321, 2002.
- [9] R. Codina. Analysis of a stabilized finite element approximation of the Oseen equations using orthogonal subscales. *Applied Numerical Mathematics*, 58:264–283, 2008.

- [10] R. Codina, S. Badia, J. Baiges, and J. Principe. *Variational Multiscale Methods in Computational Fluid Dynamics*, in Encyclopedia of Computational Mechanics (eds E. Stein, R. Borst and T.J.R. Hughes), pages 1–28. John Wiley & Sons Ltd., 2017.
- [11] R. Codina, J. Principe, and J. Baiges. Subscales on the element boundaries in the variational two-scale finite element method. *Computer Methods in Applied Mechanics and Engineering*, 198:838–852, 2009.
- [12] S. Conti, S. Müller, and M. Ortiz. Data-driven problems in elasticity. *Archive for Rational Mechanics and Analysis*, 229:79–123, 2018.
- [13] S. Conti, S. Müller, and M. Ortiz. Data-driven finite elasticity. *Archive for Rational Mechanics and Analysis*, 237:1–33, 2020.
- [14] A. Galetzka, D. Loukrezis, and H. De Gerssem. Data-driven solvers for strongly nonlinear material response. *International Journal for Numerical Methods in Engineering*, 122:1538–1562, 2021.
- [15] C. G. Gebhardt, S. Lange, and M. C. Steinbach. Formulating and heuristic solving of contact problems in hybrid data-driven computational mechanics. *Communications in Nonlinear Science and Numerical Simulation*, 134:108031, 2024.
- [16] C. G. Gebhardt, S. Lange, and M. C. Steinbach. On the mathematical structure and numerical solution of discrete-continuous optimization problems in DDCM. *Zeitschrift für Angewandte Mathematik und Physik - ZAMP*, 76:177, 2025.
- [17] C. G. Gebhardt, B. A. Roccia, J. M. Bossio, and G. R. Bossio. A framework for data-driven simulation of electrical circuits based on discrete-continuous optimization. *IEEE Access*, 13:168743–168760, 2025.
- [18] C. G. Gebhardt, D. Schillinger, M. C. Steinbach, and R. Rolfes. A framework for data-driven structural analysis in general elasticity based on nonlinear optimization: The static case. *Computer Methods in Applied Mechanics and Engineering*, 365, 2020.
- [19] C. G. Gebhardt and M. C. Steinbach. On the mathematical structure and solvability of certain Hilbert space optimization problems in data-driven elasticity. *Applied Mathematics Letters*, 172:109739, 2026.
- [20] C. G. Gebhardt, M. C. Steinbach, D. Schillinger, and R. Rolfes. A framework for data-driven structural analysis in general elasticity based on nonlinear optimization: The dynamic case. *International Journal for Numerical Methods in Engineering*, 121:5447–5468, 2020.
- [21] D. A. Ham, P. H. J. Kelly, L. Mitchell, C. J. Cotter, R. C. Kirby, K. Sagiya, N. Bouziani, S. Vorderwuelbecke, T. J. Gregory, J. Betteridge, D. R. Shapero, R. B. Nixon-Hill, C. J. Ward, P. E. Farrell, P. D. Brubeck, I. Marsden, T. H. Gibson, M. Homolya, T. Sun, A. T. T. McRae, F. Loporini, A. Gregory, M. Lange, S. W. Funke, F. Rathgeber, G.-T. Bercea, and G. R. Markall. *Firedrake User Manual*. Imperial College London, University of Oxford, Baylor University and University of Washington, first edition, 2023.
- [22] T.J.R. Hughes. Multiscale phenomena: Green’s function, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations. *Computer Methods in Applied Mechanics and Engineering*, 127:387–401, 1995.
- [23] T.J.R. Hughes, G.R. Feijóo, L. Mazzei, and J.B. Quincy. The variational multiscale method—A paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166:3–24, 1998.
- [24] T. Kirchdoerfer and M. Ortiz. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 304:81–101, 2016.

- [25] T. Kirchdoerfer and M. Ortiz. Data driven computing with noisy material data sets. *Computer Methods in Applied Mechanics and Engineering*, 326:622–641, 2017.
- [26] T. Kirchdoerfer and M. Ortiz. Data-driven computational dynamics. *International Journal for Numerical Methods in Engineering*, 113:1697–1710, 2018.
- [27] M. Kurula and H. Zwart. The duality between the gradient and divergence operators on bounded Lipschitz domains. Technical report, Department of Applied Mathematics, University of Twente, 2012.
- [28] A. Masud and T. Hughes. A stabilized mixed finite element method for Darcy flow. *Computer Methods in Applied Mechanics and Engineering*, 191:4341–4370, 2002.