A FULLY DISCRETE FRAMEWORK FOR THE ADAPTIVE SOLUTION OF INVERSE PROBLEMS
TECHNICAL REPORT TR-11-26
COMPUTER SCIENCE DEPARTMENT, VIRGINIA TECH

MIHAI ALEXE* AND ADRIAN SANDU*

Abstract. We investigate and contrast the differences between the *discretize–then–differentiate* and *differentiate–then–discretize* approaches to the numerical solution of parameter estimation problems. The former approach is attractive in practice due to the use of automatic differentiation for the generation of the dual and optimality equations in the first-order KKT system. The latter strategy is more versatile, in that it allows one to formulate efficient mesh-independent algorithms over suitably chosen function spaces. However, it is significantly more difficult to implement, since automatic code generation is no longer an option. The starting point is a classical elliptic inverse problem. An *a priori* error analysis for the discrete optimality equation shows consistency and stability are not inherited automatically from the primal discretization. Similar to the concept of dual consistency, We introduce the concept of *optimality consistency*. However, the convergence properties can be restored through suitable consistent modifications of the target functional. Numerical tests confirm the theoretical convergence order for the optimal solution. We then derive *a posteriori* error estimates for the infinite dimensional optimal solution error, through a suitably chosen error functional. This estimates are constructed using second order derivative information for the target functional. For computational efficiency, the Hessian is replaced by a low order BFGS approximation. The efficiency of the error estimator is confirmed by a numerical experiment with multigrid optimization.

1. Introduction.

1.1. Background and objectives. Inverse problems use *a priori* measurements to determine approximate values of one or more important parameters in a given model. They are most frequently formulated as numerical optimization problems [27], where the optimal parameter values, i.e., the inverse solution, correspond to the minimum point of a carefully selected cost functional. This target functional incorporates both the model and measurement information. The optimization problems are constrained by the model equations. Often, inequality-type constraints apply to both the inversion variables and the model state, to ensure feasibility of the inverse solution. Regularization terms may also be added to the cost functional to guarantee that the resulting constrained problem is well posed.

It is important to make the distinction between two types of inverse problems. Solvers for model calibration problems [8] seek to minimize the target functional $\mathcal{J}$. This is the case in aerospace applications, where one aims to find wing shapes that maximize the lift coefficient, or minimize drag [18]. Note that here the objective $\mathcal{J}$ is considered to depend on inversion parameters only indirectly. The optimal control of functional-type objectives, with PDE constraints using the finite element method, is a well developed field of research (see, e.g., the excellent review [6], and references therein). In parameter estimation problems, the objective is to explicitly retrieve the optimal values of the parameters. The control of errors in inverse problem solution itself through *a posteriori* error estimation for the primal and dual solutions, is not well developed. Note that the objective functional depends on the optimal solution only indirectly, through the model constraint equation. Hence, the theory for model

*Computational Science Laboratory, Computer Science Department, Virginia Polytechnic Institute and State University, 2202 Kraft Drive, Blacksburg, VA, 24061, USA. Email: {sandu,malexe}@vt.edu.

calibration problems cannot be applied directly for parameter identification [7].

This paper considers parameter identification problems. We start from the first order necessary conditions for a local minimum of the continuous and discrete optimization problems. The dual consistency framework given in [16] is then extended to the discrete optimality condition. This equation, part of the discrete KKT set, is obtained by linearization of the primal model along the inverse variables. Consistency and stability are both essential for convergence of the discrete optimal solution to its continuous counterpart. Similar to dual consistency, consistency of the discrete optimality equation does not hold for all discretizations. However, the analysis shows that consistent modifications [16] to the target functional, may be used to restore stability and consistency of the linearized discretization. The consistency analysis results are confirmed by the numerical experiments.

The second major contribution of this paper concerns error analysis and estimation for adaptive mesh refinement in parameter identification problems. Energy norm estimates for the control error have been derived for a rather general class of elliptic inverse problems [10, 11]. However, such estimates are of very limited use in practice, since they rely on problem dependent stability constants, and coercivity estimates for the saddle–point problems stemming from the KKT equations [23]. Another approach is proposed by Becker and Vexler in [7]. The error estimates are constructed using dual–weighted residuals [23], and apply to a modified problem defined in terms of an error functional. This error functional may be chosen to be a weighted average of the optimal solution. The authors of [7] assume that the parameter space is finite dimensional. We follow [7], and generalize their approach to the case of an infinite dimensional control space. For high computational efficiency, the exact Hessian of the modified problem is replaced by a quasi-Newton approximation. Numerical results demonstrate the efficiency of the error estimation strategy, when used in a multigrid inversion algorithm.

We consider only time independent inverse problems. However, the results presented here can be extended to the time-dependent problems, by leveraging dual consistency results of Runge–Kutta DG discretizations [1, 26].

1.2. Organization. The outline for this paper is as follows. Section 2 gives the general continuous and discrete formulations of parameter identification problems, together with the associated first order necessary conditions. The model problem and cost functional used in our numerical experiments are given in section 3, while section 4 discusses their discrete counterparts. Equations are given for the reduced gradients of the target functional with respect to the control variables, to be later used in the optimization procedure. Section 3 also contains a consistency proof for the discrete optimality condition, and the associated reduced gradient. The framework for *a posteriori* error estimation for infinite dimensional controls, based on a suitably chosen error functional, is discussed in detail in section 7. The numerical results for the multigrid optimization, the error adaptation strategy, and the consistency tests for the discrete optimality equation, are all given in section 9. Finally, section 10 discusses the conclusions, and summarizes directions of future research.

2. Continuous and discrete formulations of parameter identification problems. Given the model state $\mathbf{u} \in \mathcal{U}$, the inversion parameters $\mathbf{q} \in \mathcal{Q}$, and the real-valued target functional $\mathcal{J}$, the general continuous formulation of an inverse problem reads as follows:

$$\text{Find } \mathbf{q}_* = \arg \min_{\mathbf{q}\in\mathcal{Q},\mathbf{u}\in\mathcal{U}} \mathcal{J}\left[\mathbf{u},\mathbf{q}\right] \,, \qquad \text{subject to} \quad \mathcal{A}\left[\mathbf{u},\,\mathbf{q}\right]\,(\boldsymbol{\chi}) = 0\,, \;\; \forall\, \boldsymbol{\chi} \in \mathcal{U}\,. \tag{2.1}$$

Here $\mathcal{Q}$ and $\mathcal{U}$ are appropriate function spaces, and $\mathcal{A}[\cdot,\,\cdot](\cdot)$ is a semi-linear form (linear in the test functionals χ). We use square brackets for the nonlinear arguments, and round parentheses for the linear arguments.

2.1. The continuous optimality system. The numerical solution of (2.1) is the objective of the *differentiate–then–discretize* approach [2]. Here, one leverages the first order necessary conditions for a local optimum [22] to obtain a linearized system of optimality equations that are satisfied by all local solutions to (2.1). Given the optimal solution pair $\{\mathbf{u}_*, \mathbf{q}_*\}$ for (2.1), constrained optimization theory [22] guarantees, under suitable *a priori* assumptions on $\mathcal{J}$ and $\mathcal{A}$ [21], the existence of Lagrange multipliers $\boldsymbol{\lambda}_*$ such that the following Karush–Kuhn–Tucker (KKT) first order necessary conditions hold for the triplet $\boldsymbol{\xi}_* := \{\mathbf{u}_*,\,\boldsymbol{\lambda}_*,\,\mathbf{q}_*\} \in \mathcal{X} := \mathcal{U}\times\mathcal{U}\times\mathcal{Q}$:

$$\mathcal{A}\left[\mathbf{u}_*,\mathbf{q}_*\right](\boldsymbol{\psi}_{\boldsymbol{\lambda}}) = 0\,, \quad \forall\, \boldsymbol{\psi}_{\boldsymbol{\lambda}} \in \mathcal{U}\,, \tag{2.2a}$$

$$\mathcal{A}_{\mathbf{u}}\left[\mathbf{u}_*,\mathbf{q}_*\right](\boldsymbol{\psi}_{\mathbf{u}},\boldsymbol{\lambda}_*) = \mathcal{J}_{\mathbf{u}}\left[\mathbf{u}_*,\mathbf{q}_*\right](\boldsymbol{\psi}_{\mathbf{u}})\,, \quad \forall\, \boldsymbol{\psi}_{\mathbf{u}} \in \mathcal{U}\,, \tag{2.2b}$$

$$\mathcal{A}_{\mathbf{q}}\left[\mathbf{u}_*,\mathbf{q}_*\right](\boldsymbol{\psi}_{\mathbf{q}},\boldsymbol{\lambda}_*) = \mathcal{J}_{\mathbf{q}}\left[\mathbf{u}_*,\mathbf{q}_*\right](\boldsymbol{\psi}_{\mathbf{q}})\,, \quad \forall\, \boldsymbol{\psi}_{\mathbf{q}} \in \mathcal{Q}\,. \tag{2.2c}$$

The subscripts of the semi-linear form $\mathcal{A}$ denote partial derivatives. The small subscripts of the test functions denote the components of $\boldsymbol{\psi} = \{\boldsymbol{\psi}_{\mathbf{u}}, \boldsymbol{\psi}_{\boldsymbol{\lambda}}, \boldsymbol{\psi}_{\mathbf{q}}\} \in \mathcal{X}$.

These analytical optimality equations are obtained using the duality theory of differential operators, and Fréchet differentiation over suitable function spaces. Once derived, the analytical KKT system (2.2) is solved numerically as a set of coupled PDEs, using a suitable space-time discretization. The discretization step comes after the analytical differentiation of the primal model, hence the name of the *differentiate - first* approach.

Assumption. For any admissible parameter function $\mathbf{q} \in \mathcal{Q}^{\text{adm}}$ the forward system has a unique solution. The cost functional $\mathcal{J}$ and the primal weak–form PDE are compatible [1]. Hence, the adjoint system is well posed and has a unique solution as well. We denote the primal and dual solutions by

$$\mathbf{u} = U[\mathbf{q}]\,, \quad \boldsymbol{\lambda} = \Lambda(\mathbf{q})\,, \tag{2.3}$$

respectively. The reduced cost functional depends only on $\mathbf{q}$ as follows:

$$j[\mathbf{q}] = \mathcal{J}\left[U[\mathbf{q}],\,\mathbf{q}\right]\,. \tag{2.4}$$

Consider the Lagrangian functional $\mathcal{L} : \mathcal{X} \to \mathbb{R}$ [22] associated with the constrained problem (2.1):

$$\mathcal{L}[\boldsymbol{\xi}] := \mathcal{J}[\mathbf{u},\,\mathbf{q}] - \mathcal{A}\left[\mathbf{u},\mathbf{q}\right](\boldsymbol{\lambda})\,. \tag{2.5}$$

The KKT system (2.2) can be written compactly as:

$$\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\psi}) = 0\,,\; \forall\, \boldsymbol{\psi} \in \mathcal{X}\,. \tag{2.6}$$

Here $\mathcal{L}_{\boldsymbol{\xi}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ denotes the first variation of the Lagrangian, and is equal to

$$\begin{aligned}\mathcal{L}_{\boldsymbol{\xi}}\left[\mathbf{u},\, \boldsymbol{\lambda},\, \mathbf{q}\right]\left(\psi_{\mathrm{u}}, \psi_{\boldsymbol{\lambda}}, \psi_{\mathrm{q}}\right) &= \mathcal{J}_{\mathbf{u}}[\mathbf{u},\, \mathbf{q}](\psi_{\mathrm{u}}) + \mathcal{J}_{\mathbf{q}}[\mathbf{u},\, \mathbf{q}](\psi_{\mathrm{q}}) \\ &\quad - \mathcal{A}\left[\mathbf{u}, \mathbf{q}\right]\,(\psi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\psi_{\mathrm{u}}, \boldsymbol{\lambda}) - \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\psi_{\mathrm{q}}, \boldsymbol{\lambda}) \\ &= -\mathcal{A}\left[\mathbf{u}, \mathbf{q}\right](\psi_{\boldsymbol{\lambda}}) + \left\{\mathcal{J}_{\mathbf{u}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\psi_{\mathrm{u}}) \\ &\quad + \left\{\mathcal{J}_{\mathbf{q}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\psi_{\mathrm{q}})\,.\end{aligned}$$

Consider a second test function $\phi = \{\phi_{\mathrm{u}}, \phi_{\boldsymbol{\lambda}}, \phi_{\mathrm{q}}\} \in \mathcal{X}$. The second variation of the Lagrangian reads:

$$\begin{aligned}&\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}\left[\mathbf{u},\, \boldsymbol{\lambda},\, \mathbf{q}\right]\left(\phi_{\mathrm{u}}, \phi_{\boldsymbol{\lambda}}, \phi_{\mathrm{q}}\,;\, \psi_{\mathrm{u}}, \psi_{\boldsymbol{\lambda}}, \psi_{\mathrm{q}}\right) \\ &\quad= \mathcal{J}_{\mathbf{u},\mathbf{u}}[\mathbf{u},\, \mathbf{q}](\phi_{\mathrm{u}}, \psi_{\mathrm{u}}) + \mathcal{J}_{\mathbf{u},\mathbf{q}}[\mathbf{u},\, \mathbf{q}](\phi_{\mathrm{q}}, \psi_{\mathrm{u}}) \\ &\quad+ \mathcal{J}_{\mathbf{q},\mathbf{u}}[\mathbf{u},\, \mathbf{q}](\phi_{\mathrm{u}}, \psi_{\mathrm{q}}) + \mathcal{J}_{\mathbf{q},\mathbf{q}}[\mathbf{u},\, \mathbf{q}](\phi_{\mathrm{q}}, \psi_{\mathrm{q}}) \\ &\quad- \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{u}}, \psi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{q}}, \psi_{\boldsymbol{\lambda}}) \\ &\quad- \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\psi_{\mathrm{u}}, \phi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{u},\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{u}}, \psi_{\mathrm{u}}, \boldsymbol{\lambda}) - \mathcal{A}_{\mathbf{u},\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{q}}, \psi_{\mathrm{u}}, \boldsymbol{\lambda}) \\ &\quad- \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\psi_{\mathrm{q}}, \phi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{q},\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{u}}, \psi_{\mathrm{q}}, \boldsymbol{\lambda}) - \mathcal{A}_{\mathbf{q},\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{q}}, \psi_{\mathrm{q}}, \boldsymbol{\lambda})\,.\end{aligned}$$

This equation can be rearranged as follows:

$$\begin{aligned}\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}\left[\boldsymbol{\xi}\right](\phi, \psi) &= \left\{\mathcal{J}_{\mathbf{u},\mathbf{u}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{u},\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\phi_{\mathrm{u}}, \psi_{\mathrm{u}}) \\ &\quad+ \left\{\mathcal{J}_{\mathbf{u},\mathbf{q}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{u},\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\phi_{\mathrm{q}}, \psi_{\mathrm{u}}) \\ &\quad+ \left\{\mathcal{J}_{\mathbf{q},\mathbf{u}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{q},\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\phi_{\mathrm{u}}, \psi_{\mathrm{q}}) \\ &\quad+ \left\{\mathcal{J}_{\mathbf{q},\mathbf{q}}[\mathbf{u},\, \mathbf{q}] - \mathcal{A}_{\mathbf{q},\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\boldsymbol{\lambda})\right\}\,(\phi_{\mathrm{q}}, \psi_{\mathrm{q}}) \\ &\quad- \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{u}}, \psi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{u}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\boldsymbol{\lambda}}, \psi_{\mathrm{u}}) \\ &\quad- \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\mathrm{q}}, \psi_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{q}}\left[\mathbf{u}, \mathbf{q}\right](\phi_{\boldsymbol{\lambda}}, \psi_{\mathrm{q}})\,.\end{aligned} \tag{2.7}$$

The primal, dual, and optimality equations are often solved simultaneously (*all–at–once*), to yield a new search direction for the nonlinear solution algorithm. For example, the Newton update with the solution increment $\boldsymbol{\delta\xi}_k$ at iteration k reads:

$$\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_k](\boldsymbol{\delta\xi}_k\,, \psi) = -\mathcal{L}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_k](\psi)\,, \quad \forall\, \psi \in \mathcal{X}\,, \tag{2.8}$$

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi}_k + \boldsymbol{\delta\xi}_k\,. \tag{2.9}$$

In the *continuous approach* (the *"differentiate–then–discretize"* strategy) to solving the inverse problem the KKT equations (2.6) are first derived analytically in an infinite dimensional setting, and then are discretized with a numerical method of choice, resulting in a system of nonlinear equations. The *"optimize–then–discretize"* approach goes one step further, in that the full iterative solution algorithm for the optimality system is derived in an infinite dimensional setting (2.8); the iterations (2.8) are then discretized with the method of choice.

The advantage of the continuous formulation lies in its flexibility. The complete solution algorithm, i.e., both the model equations, and the nonlinear minimization algorithm for $\mathcal{J}$, can be formulated in function spaces [2]. This allows arbitrary choices of finite-element type discretizations once the problem has been fully specified in a functional space setting. Space-time meshes can be changed between nonlinear iterations, and convergence can be quantified in a mesh-independent fashion. There are also few restrictions on the mesh types and trial function spaces [2, 4]. The main drawback of this method is its additional complexity in both derivation and implementation, since it is not amenable to automatic code generation.

2.2. The discrete optimality equations. The *discrete approach* (the *"discretize–then–differentiate"* strategy) to solving the inverse problem starts from the discrete counterpart of (2.1), obtained using the discontinuous Galerkin finite element method [17]:

$$\text{Find } \mathbf{q}^h_* = \arg \min_{\mathbf{q}^h \in \mathcal{Q}^h, \mathbf{u}^h \in \mathcal{U}^h} \mathcal{J}^h[\mathbf{u}^h, \mathbf{q}^h] \,, \qquad \text{subject to} \quad \mathcal{A}^h[\mathbf{u}^h, \mathbf{q}^h]\,(\boldsymbol{\chi}^h) = 0 \,, \; \forall\, \boldsymbol{\chi}^h \in \mathcal{U}^h \,. \tag{2.10}$$

The discrete operators, variables , and function spaces are denoted by the superscript h. The discrete function spaces are $\mathcal{U}^h \subset \mathcal{U}$, and $\mathcal{Q}^h \subset \mathcal{Q}$. The weak form $\mathcal{A}^h$ is linear in $\boldsymbol{\chi}^h$, but may be nonlinear in both $\mathbf{q}^h$ and $\mathbf{u}^h$. The discrete weak formulation of the primal model is *a priori* assumed to be a consistent and stable discretization of the original weak-form PDE in (2.1).

Assumption. The primal discretization is convergent, and the solution $\mathbf{u}^h$ is unique for any given set of admissible inversion variables $\mathbf{q}^h$.

Again, we assume there exists at least one locally unique solution to (2.10). Such a solution $\boldsymbol{\xi}^h_* := \{\mathbf{u}^h_*, \boldsymbol{\lambda}^h_*, \mathbf{q}^h_*\} \in \mathcal{X}^h = \mathcal{U}^h \times \mathcal{U}^h \times \mathcal{Q}^h$ is required to satisfy the discrete KKT necessary conditions:

$$\mathcal{A}^h[\mathbf{u}^h_*, \mathbf{q}^h_*](\boldsymbol{\psi}^h_{\boldsymbol{\lambda}}) = 0 \,, \quad \forall\, \boldsymbol{\psi}^h_{\boldsymbol{\lambda}} \in \mathcal{U}^h \,, \tag{2.11a}$$

$$\mathcal{A}^h_{\mathbf{u}^h}[\mathbf{u}^h_*, \mathbf{q}^h_*](\boldsymbol{\psi}^h_{\mathbf{u}}, \boldsymbol{\lambda}^h_*) = \mathcal{J}^h_{\mathbf{u}^h}[\mathbf{u}^h_*, \mathbf{q}^h_*](\boldsymbol{\psi}^h_{\mathbf{u}}) \,, \quad \forall\, \boldsymbol{\psi}^h_{\mathbf{u}} \in \mathcal{U}^h \,, \tag{2.11b}$$

$$\mathcal{A}^h_{\mathbf{q}^h}[\mathbf{u}^h_*, \mathbf{q}^h_*](\boldsymbol{\psi}^h_{\mathbf{q}}, \boldsymbol{\lambda}^h) = \mathcal{J}^h_{\mathbf{q}^h}[\mathbf{u}^h_*, \mathbf{q}^h_*](\boldsymbol{\psi}^h_{\mathbf{q}}) \,, \quad \forall\, \boldsymbol{\psi}^h_{\mathbf{q}} \in \mathcal{Q}^h \,, \tag{2.11c}$$

or, in a more compact notation

$$\mathcal{L}^h_{\boldsymbol{\xi}^h}[\boldsymbol{\xi}^h_*](\boldsymbol{\psi}^h) = 0 \,, \quad \forall\, \boldsymbol{\psi}^h \in \mathcal{X}^h \,,$$

with the discrete Lagrangian functional $\mathcal{L}^h : \mathcal{X}^h \to \mathbb{R}$ [22]

$$\mathcal{L}^h[\boldsymbol{\xi}^h] := \mathcal{J}^h[\mathbf{u}^h, \mathbf{q}^h] - \mathcal{A}^h[\mathbf{u}^h, \mathbf{q}^h](\boldsymbol{\lambda}^h) \,. \tag{2.12}$$

The subscripts denote partial derivatives with respect to the discrete variables.

While the discretize–first approach lacks the flexibility of its differentiate–first counterpart, it has the important advantage that the KKT system can be generated with relatively low human effort using automatic differentiation [12]. The grid transfer operators used for mesh refinement and coarsening operations in discontinuous Galerkin are also amenable to automatic differentiation [1]. This can significantly reduce the software development time required for a full implementation. Moreover, many practical problems require the reuse of legacy software; and for many applications the solution algorithm needs to interface with existing numerical optimization or ODE solvers that require the inputs to be discrete mesh variables. In such cases discretize–first is the only feasible approach to solving the inverse problem. Finally, we mention that the discrete adjoint approach is a natural fit for multigrid optimization [7, 20], which compensates for its inability to adapt the mesh between consecutive nonlinear iterations.

As previously indicated in the literature (see, e.g., [14, 16]), the linearization and discretization steps do not generally commute. In the limit of the discretization, one hopes for convergence of the discrete optimal solution, but this happens only if the discretize–first approach yields a set of discrete KKT equations that are stable and consistent discretizations of their continuous counterparts.

2.3. Consistency of the discretizations. We now define consistency for the primal, dual, and optimality equation discretizations (2.11a)–(2.11c). The primal and dual consistency definitions follow the ones given in [16]. In addition, we propose to consider the consistency of the optimality equation discretization (definition 2.3); this type of consistency will prove to be a crucial requirement for the convergence of the discrete optimal solution $\mathbf{q}^h$ to its analytical counterpart $\mathbf{q}$.

DEFINITION 2.1 (Primal consistency). *The primal discretization (2.11a) is consistent if the exact solutions* $\mathbf{u}$ *and* $\mathbf{q}$ *of the weak form primal equation (2.2a) satisfy:*

$$\mathcal{A}^h[\mathbf{u},\mathbf{q}](\psi_{\boldsymbol{\lambda}}) = 0\,, \quad \forall\, \psi_{\boldsymbol{\lambda}} \in \mathcal{U}\,.$$

DEFINITION 2.2 (Dual consistency). *The primal discretization (2.11a) is dual consistent, if any triplet* $\boldsymbol{\xi} = \{\mathbf{u}, \boldsymbol{\lambda}, \mathbf{q}\} \in \mathcal{X}$ *that verifies the weak-form primal and dual equations (2.2a)–(2.2b), also satisfies:*

$$\mathcal{A}^h_{\mathbf{u}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{u}}, \boldsymbol{\lambda}) = \mathcal{J}^h_{\mathbf{u}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{u}})\,, \quad \forall\, \psi_{\mathbf{u}} \in \mathcal{U}\,.$$

DEFINITION 2.3 (Optimality equation consistency). *The discretization (2.11c) is said to be consistent, if any triplet* $\boldsymbol{\xi} = \{\mathbf{u}, \boldsymbol{\lambda}, \mathbf{q}\} \in \mathcal{X}$ *that verifies (2.2c), also satisfies the equation:*

$$\mathcal{A}^h_{\mathbf{q}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{q}}, \boldsymbol{\lambda}) = \mathcal{J}^h_{\mathbf{q}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{q}})\,, \quad \forall\, \psi_{\mathbf{q}} \in \mathcal{Q}\,.$$

The definitions above can be extended to time–dependent problems, where the time dimension is discretized using a Runge–Kutta quadrature [1].

3. The model problem. Let $\Omega \subset \mathbb{R}^d$ be a closed convex polyhedral domain with $d \in \{2, 3\}$. Γ is the boundary of Ω. Consider the following elliptic boundary-value problem, henceforth referred to as the *primal model*:

$$\begin{aligned} -\nabla \cdot (\mathbf{q}\, \nabla \mathbf{u}) &= \mathbf{f}\,,\ \mathbf{x} \in \Omega\,, \\ \mathbf{u} &= \mathbf{g}\,,\ \mathbf{x} \in \Gamma\,. \end{aligned} \tag{3.1}$$

Smoothness assumptions. Let the volume forcing $\mathbf{f} \in \mathcal{L}^2(\Omega)$, and the Dirichlet boundary data $\mathbf{g} \in \mathcal{H}^{3/2}(\Gamma)$. Also, assume the inversion variables $\mathbf{q} \in \mathcal{H}^2(\Omega)$. The formulation of the primal problem requires that $\mathbf{q} \geq 0$, *a.e.* on Ω. Also, assume that any additional conditions on the domain boundary Γ [13] are satisfied, such that the smoothness of the exact primal solution is $\mathbf{u} \in \mathcal{U} \subset \mathcal{H}^2(\Omega)$.

The smoothness of the solution space $\mathcal{U}$ guarantees existence and boundedness of the following trace operators for any function $\mathbf{v} \in \mathcal{U}$ [24]:

$$\gamma^0 : \mathcal{H}^2(\Omega) \to \mathcal{H}^{3/2}(\Gamma)\,, \quad \gamma^0(\mathbf{v}) = \mathbf{v}|_{\Gamma}\,, \tag{3.2a}$$

$$\gamma^1 : \mathcal{H}^2(\Omega) \to \mathcal{H}^{1/2}(\Gamma)\,, \quad \gamma^1(\mathbf{v}) = \nabla \mathbf{v} \cdot \vec{\mathbf{n}}|_{\Gamma}\,. \tag{3.2b}$$

Inner products of functions in $\mathcal{U}$ on Ω and Γ are defined as follows:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\Omega} := \int_{\Omega} \mathbf{u}\, \mathbf{v}\, d\mathbf{x}\,, \qquad \langle \mathbf{u}, \mathbf{v} \rangle_{\Gamma} := \int_{\Gamma} \gamma^0(\mathbf{u})\, \gamma^0(\mathbf{v})\, d\mathbf{s}\,, \qquad \forall\, \mathbf{u}, \mathbf{v} \in \mathcal{U}\,.$$

They induce the corresponding norms on $\mathcal{U}$ and on the space of traces of functions in $\mathcal{U}$:

$$\|\mathbf{u}\|_{\mathcal{L}^2(\Omega)} := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_\Omega}\,, \quad \|\mathbf{u}\|_{\mathcal{L}^2(\Gamma)} := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_\Gamma}\,, \quad \forall\, \mathbf{u} \in \mathcal{U}\,.$$

Inner products on the space $\mathcal{Q}$, and on the product space $\mathcal{U} \times \mathcal{Q}$, are defined in a similar fashion.

The target functional reads:

$$\mathcal{J}[\mathbf{u}, \mathbf{q}] = \frac{1}{2} \|\mathcal{H}\mathbf{u} - \mathbf{o}\|^2_{\mathcal{L}^2(\Omega)} + \frac{1}{2} \|\nabla(\mathbf{q} - \mathbf{q}_B)\|^2_{\mathcal{L}^2(\Omega)} + \frac{\beta}{2} \|\mathbf{q} - \mathbf{q}_B\|^2_{\mathcal{L}^2(\Omega)}\,. \tag{3.3}$$

The background control $\mathbf{q}_B \in \mathcal{Q}$ is positive *a.e.* on Ω. $\mathcal{H} : \mathcal{U} \to \mathcal{O} \subset \mathcal{L}^2(\Omega)$ is a linear and continuous observation operator. The exact choice of regularization parameter $\beta > 0$ depends on the discretization, as well as on the properties of primal model and cost functional, and is the topic of current research: see, e.g., [15, 19, 29], and references therein. We will assume a constant regularization parameter.

With this choice of model and cost functional, the inverse problem (2.1) has the following form:

$$\text{Find } \mathbf{q}_* = \arg\min_{\mathbf{q} \in \mathcal{Q}} \mathcal{J}[\mathbf{u}, \mathbf{q}]\,, \text{ subject to (3.1)}\,. \tag{3.4}$$

Note that there are no explicit bound constraints on the inversion variables $\mathbf{q}$. Rather, the positivity of the diffusion coefficient is enforced indirectly, by solving (3.4) in a sufficiently small neighborhood of the reference profile. Another approach to guarantee positivity is to explicitly enforce the bounds for the discrete parameter values as constraints in the optimization; this will be the topic of future work.

3.1. The continuous KKT system for the model problem. In what follows, adjoint operators will be denoted by a $*$ superscript. Assume no boundary terms arise in the definition of the adjoint operator $\mathcal{H}^*$ (compatibility condition 1 in [1]):

$$\langle \mathcal{H}\mathbf{u}, \mathbf{o} \rangle_\Omega = \langle \mathbf{u}, \mathcal{H}^* \mathbf{o} \rangle_\Omega\,, \; \forall\, \mathbf{u} \in \mathcal{U}, \mathbf{o} \in \mathcal{O}\,.$$

With this, the adjoint system (2.2b) associated with the model problem (3.1) reads:

$$\begin{aligned} -\nabla \cdot (\mathbf{q}\, \nabla \boldsymbol{\lambda}) &= \mathcal{H}^*(\mathcal{H}\mathbf{u} - \mathbf{o})\,, \; \mathbf{x} \in \Omega \\ \boldsymbol{\lambda} &= 0\,, \; \mathbf{x} \in \Gamma\,. \end{aligned} \tag{3.5}$$

Remark. With the smoothness assumptions on $\mathbf{q}$, $\mathbf{o}$, and $\mathbf{u}$, the exact dual solution is $\boldsymbol{\lambda} \in \mathcal{H}^2(\Omega)$.

The optimality equation (2.2c) associated with the model problem (3.1) reads:

$$\langle \nabla\mathbf{q}, \nabla\mathbf{z} \rangle_\Omega - \langle \nabla\mathbf{q}_B, \nabla\mathbf{z} \rangle_\Omega + \beta\, \langle \mathbf{q} - \mathbf{q}_B, \mathbf{z} \rangle_\Omega + \langle \nabla \cdot (\mathbf{z}\, \nabla\mathbf{u}), \boldsymbol{\lambda} \rangle_\Omega = 0\,, \;\; \forall \mathbf{z} \in \mathcal{Q}\,.$$

The strong form of the optimality equation is then obtained through an integration by parts:

$$\begin{aligned} -\Delta\, \mathbf{q} + \beta\, (\mathbf{q} - \mathbf{q}_B) &= -\Delta \mathbf{q}_B + \nabla\mathbf{u} \cdot \nabla\boldsymbol{\lambda}\,, \; \mathbf{x} \in \Omega\,, \\ \nabla\mathbf{q} \cdot \vec{\mathbf{n}} &= \nabla\mathbf{q}_B \cdot \vec{\mathbf{n}}\,, \; \mathbf{x} \in \Gamma\,. \end{aligned} \tag{3.6}$$

Remark. Using the strong maximum principle [24], this Helmholtz boundary value problem can be shown to be well posed. Since the volume forcing $\left(-\Delta \mathbf{q}_B + \nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda}\right) \in \mathcal{L}^2(\Omega)$, and the Neumann boundary data $\nabla \mathbf{q}_B \cdot \vec{\mathbf{n}} \in \mathcal{H}^{1/2}(\Gamma)$, the exact solution $\mathbf{q}_* \in \mathcal{H}^2(\Omega)$. By the Sobolev imbedding theorem [24], $\mathbf{q}_*$ is bounded and continuous on Ω for $d \in \{2, 3\}$. Furthermore, in a small neighborhood of a sufficiently large $\mathbf{q}_B$, the solution $\mathbf{q}_* > 0$ *a.e.* on Ω.

3.2. The reduced gradient. This section illustrates in detail the derivation of the gradient of the reduced cost functional (2.4). Consider an arbitrary function $\boldsymbol{\delta}\mathbf{q} \in \mathcal{Q}$, and let $\boldsymbol{\delta}\mathbf{u} := \dfrac{\partial U}{\partial \mathbf{q}}[\mathbf{q}](\boldsymbol{\delta}\mathbf{q})$, with $\boldsymbol{\delta}\mathbf{u} \in \mathcal{U}$. The reduced gradient $\nabla_{\mathbf{q}} \mathcal{J}$ is defined through identification from the following equality:

$$
\left\langle \nabla_{\mathbf{q}} \mathcal{J}, \boldsymbol{\delta}\mathbf{q} \right\rangle_{\Omega} := \frac{\partial \mathcal{J}}{\partial \mathbf{q}}[\mathbf{u}, \mathbf{q}](\boldsymbol{\delta}\mathbf{q}) + \frac{\partial \mathcal{J}}{\partial \mathbf{u}}[\mathbf{u}, \mathbf{q}](\boldsymbol{\delta}\mathbf{u}) \,, \ \forall\, \boldsymbol{\delta}\mathbf{q} \in \mathcal{Q} \,. \tag{3.7}
$$

We make use of the tangent linear equation, which is obtained from the primal model by differentiation in the direction $(\boldsymbol{\delta}\mathbf{u}, \boldsymbol{\delta}\mathbf{q}) \in \mathcal{U} \times \mathcal{Q}$:

$$
\begin{aligned}
-\nabla \cdot (\boldsymbol{\delta}\mathbf{q}\, \nabla \mathbf{u}) - \nabla \cdot (\mathbf{q}\, \nabla\, \boldsymbol{\delta}\mathbf{u}) &= 0 \,, \ \mathbf{x} \in \Omega \\
\boldsymbol{\delta}\mathbf{u} &= 0 \,, \ \mathbf{x} \in \Gamma \,.
\end{aligned} \tag{3.8}
$$

Using integration by parts, the adjoint equation (3.5), and the tangent linear model (3.8), the reduced gradient formula becomes:

$$
\begin{aligned}
\left\langle \nabla_{\mathbf{q}} \mathcal{J}, \boldsymbol{\delta}\mathbf{q} \right\rangle_{\Omega} &= \int_{\Omega} \mathcal{H}^*(\mathcal{H}\mathbf{u} - \mathbf{o})\, \boldsymbol{\delta}\mathbf{u}\, \mathrm{d}\mathbf{x} + \int_{\Omega} \nabla \mathbf{q}\, \nabla \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} \\
&\quad - \int_{\Omega} \nabla \mathbf{q}_B\, \nabla \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} + \boldsymbol{\beta} \int_{\Omega} (\mathbf{q} - \mathbf{q}_B)\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{s} \\
&= - \int_{\Omega} \nabla \boldsymbol{\lambda} \cdot \nabla \mathbf{u}\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} - \int_{\Omega} \Delta \mathbf{q}\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} + \int_{\Omega} \Delta \mathbf{q}_B\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} \\
&\quad + \boldsymbol{\beta} \int_{\Omega} (\mathbf{q} - \mathbf{q}_B)\ \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{x} + \int_{\Gamma} \nabla \mathbf{q} \cdot \vec{\mathbf{n}}\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{s} - \int_{\Gamma} \nabla \mathbf{q}_B \cdot \vec{\mathbf{n}}\, \boldsymbol{\delta}\mathbf{q}\, \mathrm{d}\mathbf{s} \\
&:= \left\langle \nabla_{\mathbf{q}} \mathcal{J}\big|_{\Omega}, \boldsymbol{\delta}\mathbf{q} \right\rangle_{\Omega} + \left\langle \nabla_{\mathbf{q}} \mathcal{J}\big|_{\Gamma}, \boldsymbol{\delta}\mathbf{q} \right\rangle_{\Gamma} \,.
\end{aligned} \tag{3.9}
$$

We can immediately see that the gradient is identically zero at the optimal solution $\mathbf{q}_*$, which satisfies (3.6).

4. *A priori* error analysis for the discrete optimality system. This section investigates the consistency properties of the optimality system for the discrete counterpart of problem (3.4). The discontinuous Galerkin method [17] is used for the spatial discretization.

4.1. Notation and preliminaries for the discrete problems. Assume the discrete domains cover exactly the analytical ones, i.e., $\Omega^h \equiv \Omega$, and $\Gamma^h \equiv \Gamma$. Let $\mathcal{U}_p^h = \mathrm{span}\{\chi_j(\mathbf{x})\}_{j=1\ldots P} \subset \mathcal{U}$ denote (for any integer $p \geq 1$) the discrete solution space consisting on discontinuous piecewise polynomial functions of degree less than or equal to $p \geq 1$. Similarly, $\mathcal{Q}_r^h = \mathrm{span}\{\chi_i^{\mathbf{q}}(\mathbf{x})\}_{i=1\ldots R} \subset \mathcal{Q}$ is the space of piecewise polynomials of maximal degree $r \geq 1$.

Following the notation in [16], the domain Ω is divided into shape-regular meshes $\mathcal{T}^h = \bigcup_{j=1}^{J} \kappa_j$, and $\mathcal{T}_{\mathrm{q}}^h = \bigcup_{i=1}^{I} \kappa_i^q$, comprised of polyhedral elements. The primal and dual discrete variables $\mathbf{u}^h$ and $\boldsymbol{\lambda}^h$ are defined on the mesh $\mathcal{T}^h$, while the triangulation $\mathcal{T}_{\mathrm{q}}^h$ holds the discrete inversion variables $\mathbf{q}^h$. Let $\Gamma_{\mathcal{I}}$ and $\Gamma_{\mathcal{I}}^{\mathrm{q}}$ be the union of all distinct interior edges of $\mathcal{T}^h$ and $\mathcal{T}_{\mathrm{q}}^h$, respectively.

From the definition of $\mathcal{U}_p^h$, the restriction of any mesh function $\mathbf{v}^h$ defined on $\mathcal{T}^h$ to an arbitrary element κ can be written as follows:

$$\mathbf{v}^h(\mathbf{x})\Big|_{\kappa} = \sum_{j=1}^{P} v_{\kappa}^j\, \chi_j(\mathbf{x}) , \quad \forall\, \mathbf{u}^h \in \mathcal{U}_p^h , \, \forall\, \kappa \in \mathcal{T}^h . \tag{4.1}$$

The coefficients $\{v_{\kappa}^j\}$, $j = 1 \dots P$, fully determine the numerical solution inside κ. Similarly, for arbitrary mesh functions $\mathbf{z}^h \in \mathcal{Q}_r^h$ defined on the triangulation $\mathcal{T}_{\mathrm{q}}^h$, one has:

$$\mathbf{z}^h(\mathbf{x})\Big|_{\kappa^q} = \sum_{i=1}^{R} z_{\kappa^q}^i\, \chi_i^{\mathbf{q}}(\mathbf{x}) , \quad \forall\, \mathbf{z}^h \in \mathcal{Q}_r^h , \, \forall\, \kappa^q \in \mathcal{T}_{\mathrm{q}}^h . \tag{4.2}$$

Let h_κ denote the diameter of $\kappa \in \mathcal{T}^h$:

$$h_\kappa := \max_{\mathbf{x},\mathbf{y} \in \kappa} \|\mathbf{x} - \mathbf{y}\|_2 .$$

The area (or volume) of the element κ is $|\kappa|$, and its boundary is denoted by $\partial\kappa$. Furthermore, $h := \max_{\kappa \in \mathcal{T}^h} h_\kappa$, and $h_{\min} := \min_{\kappa \in \mathcal{T}^h} h_\kappa$, with the assumption that the ratio $h/h_{\min}$ is a constant independent of h. Given two neighboring elements κ_- and κ_+ (with a common edge $e = \kappa_- \cap \kappa_+$), we let $\mathbf{u}_\pm^h := \mathbf{u}^h\Big|_{\partial\kappa_\pm}$ denote the trace of $\mathbf{u}^h$ on e, taken from the interior of $\kappa_\pm$, respectively.

Let $e = \kappa_+ \cup \kappa_-$ denote an edge (or face) between two neighboring elements κ_+ and κ_-. We denote by $|e|$ the length (or area) of e. Furthermore, the jump in the solution over e is given by:

$$[\![\mathbf{u}^h]\!] := \mathbf{u}_+^h \vec{\mathbf{n}}_+ + \mathbf{u}_-^h \vec{\mathbf{n}}_- ,$$

whereas the solution average at $\mathbf{x}^h \in \kappa_- \cap \kappa_+$ is

$$\{\mathbf{u}^h\} := \frac{\mathbf{u}_-^h + \mathbf{u}_+^h}{2} .$$

On a boundary edge $e \subset \Gamma$ the definitions become $\{\mathbf{u}^h\} := \mathbf{u}_+^h$ and $[\![\mathbf{u}]\!] := \mathbf{u}_+^h\, \vec{\mathbf{n}}_+$.

Define the following two discrete inner product functionals over the space $\mathcal{U}_p^h$:

$$\left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_\kappa := \int_\kappa \mathbf{u}^h\, \mathbf{v}^h\, \mathrm{d}\mathbf{x} , \, \forall\, \kappa \in \mathcal{T}^h ,$$
$$\left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_e := \int_e \mathbf{u}^h\, \mathbf{v}^h\, \mathrm{d}\mathbf{s} , \, \forall\, e \in \Gamma \cup \Gamma_{\mathcal{I}} .$$

They are valid for any mesh functions $\mathbf{u}^h$, $\mathbf{v}^h \in \mathcal{U}_p^h$. Inner products on Ω and inner/outer boundaries are defined by extending the definitions above:

$$\left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_\Omega = \sum_{\kappa \in \mathcal{T}^h} \left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_\kappa ,$$

$$\left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_{\Gamma} = \sum_{e\in\Gamma} \left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_e ,$$

$$\left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_{\Gamma_{\mathcal{I}}} = \sum_{e\in\Gamma_{\mathcal{I}}} \left\langle \mathbf{u}^h, \mathbf{v}^h \right\rangle_e , \forall\, \mathbf{u}^h, \mathbf{v}^h \in \mathcal{U}_p^h .$$

The inner products induce corresponding $\mathcal{L}^2$ norms on mesh elements ($\|\cdot\|_\kappa$) and on mesh edges ($\|\cdot\|_e$) .

Assume that the mesh $\mathcal{T}^h$ is regular enough such that, for some $\alpha_1, \alpha_2 > 0$,

$$\alpha_1 h_\kappa^{-1} \le \frac{|e|}{|\kappa|} \le \alpha_2 h_\kappa^{-1} , \quad \forall \kappa \in \mathcal{T}^h .$$

Then the following trace inequalities hold for polynomial functions on κ [9, 17, 25]

$$\|\mathbf{u}^h\|_{\partial\kappa} \le C_1(p)\, h_\kappa^{-1/2}\, \|\mathbf{u}^h\|_\kappa , \tag{4.3a}$$

$$\|\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}\|_{\partial\kappa} \le C_2(p)\, h_\kappa^{-1/2} \|\nabla \mathbf{u}^h\|_\kappa , \quad \forall\, \mathbf{u}^h \in \mathcal{U}_p^h . \tag{4.3b}$$

Here the constants $C_1(p)$ and $C_2(p)$ depend only on the polynomial basis order p.

Let $\mathcal{H}^s(\mathcal{T}^h)$ denote the broken Sobolev space, for any real number s:

$$\mathcal{H}^s(\mathcal{T}^h) := \left\{ \mathbf{v} \in \mathcal{L}^2(\Omega) \mid \forall\, \kappa \in \mathcal{T}^h, \mathbf{v}|_\kappa \in \mathcal{H}^s(\kappa) \right\} .$$

This broken space is endowed with the following norm:

$$\|\mathbf{v}\|_{\mathcal{H}^s(\mathcal{T}^h)} := \left(\sum_{\kappa\in\mathcal{T}^h} \|\mathbf{v}\|^2_{\mathcal{H}^s(\kappa)} \right)^{1/2} . \tag{4.4}$$

We have that $\mathcal{U}_p^h \subset \mathcal{H}^s(\mathcal{T}^h)$ for any $s \ge 0$.

Remark. Similar assumptions, definitions, and notations hold for mesh functions in $\mathcal{Q}_r^h$ defined on the triangulation $\mathcal{T}_\mathrm{q}^h$.

4.1.1. Nested meshes and combined inner products. To be able to use different meshes for the parameters and primal/dual variables, we make some simplifying assumptions on the structure of the triangulations $\mathcal{T}^h$ and $\mathcal{T}_\mathrm{q}^h$ [2]. Specifically, we assume that $\mathcal{T}^h$ can be obtained from $\mathcal{T}_\mathrm{q}^h$ by hierarchical mesh refinement, and that $p \ge r$. This implies that the basis functions for $\mathcal{Q}_r^h$ can be written as linear combinations of the bases for $\mathcal{U}_p^h$:

$$\chi_\ell^q = \sum_{j=1}^{P} \mathbf{M}_{\ell j}\, \chi_j , \quad \ell = 1, \ldots, R .$$

With (4.1) and (4.2), we define the combined inner product of a function $\mathbf{u}^h \in \mathcal{U}_p^h$ with a function $\mathbf{z}^h \in \mathcal{Q}_r^h$ on an arbitrary element $\kappa \in \mathcal{T}^h$ as follows:

$$\left\langle \mathbf{u}^h, \mathbf{z}^h \right\rangle_\kappa := \int_\kappa \mathbf{u}^h\, \mathbf{z}^h\, \mathrm{d}\mathbf{s} = \sum_{i=1}^{P} \sum_{\ell=1}^{R} \sum_{j=1}^{P} u_\kappa^i\, z_{\kappa^q}^\ell\, \mathbf{M}_{\ell j} \left\langle \chi_i, \chi_j \right\rangle_\kappa .$$

Remark. A similar definition holds on any element $\kappa_q \in \mathcal{T}_{\mathrm{q}}^h$, using the reverse mapping $\mathbf{M}^q \in \mathbb{R}^{P \times R}$.

These definitions can be extended to cover the entire discrete domain Ω, boundary Γ, or inner faces $\Gamma_{\mathcal{I}}$, and $\Gamma_{\mathcal{I}}^{\mathrm{q}}$, respectively. To do so, it suffices to note that since $\mathcal{T}_{\mathrm{q}}^h$ can be obtained by coarsening $\mathcal{T}^h$, the set of inner edges for the parameter mesh $\mathcal{T}_{\mathrm{q}}^h$ is included in that of the primal/dual

$$\Gamma_{\mathcal{I}}^{\mathrm{q}} \subset \Gamma_{\mathcal{I}} \,. \tag{4.5}$$

Equation (4.5) will also be useful in defining and analyzing the discrete weak formulations discussed in later sections.

4.2. The discrete KKT system. Consider the primal problem (3.1). Its symmetric interior penalty discontinuous Galerkin (SIPG) discretization reads [16]:

(4.6a) Find $\mathbf{u}^h \in \mathcal{U}_p^h$ such that:

$$\mathcal{N}^h[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) = \int_\Omega \mathbf{f}^h\, \mathbf{w}^h\, \mathrm{d}\mathbf{x} + \mathcal{B}^h[\mathbf{g}^h, \mathbf{q}^h](\mathbf{w}^h) \,,\ \forall\, \mathbf{w}^h \in \mathcal{U}_p^h \,,$$

with

$$\begin{aligned} \mathcal{N}^h[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) := & \int_\Omega \mathbf{q}^h\, (\nabla \mathbf{u}^h \cdot \nabla \mathbf{w}^h)\, \mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \phi\, [\![\mathbf{u}^h]\!]\, [\![\mathbf{w}^h]\!]\, \mathrm{d}\mathbf{s} \\ & - \int_{\Gamma_{\mathcal{I}} \cup \Gamma} [\![\mathbf{u}^h]\!]\, \left\{ \mathbf{q}^h\, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}) \right\}\, \mathrm{d}\mathbf{s} \\ & - \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \left\{ \mathbf{q}^h\, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \right\}\, [\![\mathbf{w}^h]\!]\, \mathrm{d}\mathbf{s} \,, \end{aligned} \tag{4.6b}$$

$$\mathcal{B}^h[\mathbf{g}^h, \mathbf{q}^h](\mathbf{w}^h) := - \int_\Gamma \mathbf{q}^h\, \mathbf{g}^h\, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}})\, \mathrm{d}\mathbf{s} + \int_\Gamma \phi\, \mathbf{g}^h\, \mathbf{w}^h\, \mathrm{d}\mathbf{s} \,. \tag{4.6c}$$

Recall that $\Gamma_{\mathcal{I}}$ denotes the union of all interior edges. The penalty parameter is

$$\phi = \widehat{\phi}\, h_e^{-1} \,, \quad h_e := |e|^{1/(d-1)} \,, \tag{4.7}$$

for any edge e, and dimension $d \in \{2, 3\}$. From (4.3)

$$|e| \le \alpha_2 h_\kappa^{-1}\, |\kappa| \le \alpha_2 h_\kappa^{d-1} \le \alpha_2 h^{d-1} \quad \Rightarrow \quad h_e \le \alpha_2^{1/(d-1)}\, h \,, \quad \forall \kappa \in \mathcal{T}^h \,, \tag{4.8}$$

where the constant depends on the regularity of the mesh but not on the particular element κ.

Using the notation above, the energy norm (or *natural* DG norm) [17, 25] is defined as:

$$\|\mathbf{v}\|_{\mathrm{DG}} := \left(\sum_{\kappa \in \mathcal{T}^h} \int_\kappa \mathbf{q}\, \nabla \mathbf{v} \cdot \nabla \mathbf{v}\, \mathrm{d}\mathbf{x} + \sum_{e \in \Gamma_{\mathcal{I}} \cup \Gamma} \widehat{\phi}\, h_e^{-1} \int_e [\![\mathbf{v}]\!] \cdot [\![\mathbf{v}]\!]\, \mathrm{d}\mathbf{s} \right)^{1/2} ,\ \forall\, \mathbf{v} \in \mathcal{U} \,. \tag{4.9}$$

With this, we have the following result.

THEOREM 4.1. *Assume the exact solution to (3.1) belongs to $\mathcal{H}^s(\Omega)$, $s \ge 2$. For sufficiently a large penalty parameter $\widehat{\phi} > 0$, there exists a constant C independent of h such that*

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathrm{DG}} \le C\, h^{\min(p+1,s)-1} \|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}^h)} \,, \tag{4.10}$$

and

$$\|\mathbf{u}-\mathbf{u}^h\|_{\mathcal{L}^2(\Omega)} \le C\, h^{\min(p+1,s)} \|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,, \tag{4.11}$$

Proof. See [25, Chapter 2]. □

It is useful to note that (4.9) and (4.8) imply:

$$\begin{aligned}\|\mathbf{v}\|_{\mathrm{DG}} &\ge \left(\sum_{e\in\Gamma_{\mathcal{I}}\cup\Gamma} \widehat{\phi}\, h_e^{-1} \int_e [\![\mathbf{v}]\!]\cdot[\![\mathbf{v}]\!]\,\mathrm{d}\mathbf{s}\right)^{1/2}\\ &\ge \widehat{\phi}^{1/2}\,\alpha_2^{-1/(2d-2)}\, h^{-1/2} \left(\sum_{e\in\Gamma_{\mathcal{I}}\cup\Gamma} \int_e [\![\mathbf{v}]\!]\cdot[\![\mathbf{v}]\!]\,\mathrm{d}\mathbf{s}\right)^{1/2},\end{aligned}$$

and therefore

$$h^{-1/2}\left(\sum_{e\in\Gamma_{\mathcal{I}}\cup\Gamma} \int_e [\![\mathbf{u}-\mathbf{u}^h]\!]\cdot[\![\mathbf{u}-\mathbf{u}^h]\!]\,\mathrm{d}\mathbf{s}\right)^{1/2} \le C\,\left\|\mathbf{u}-\mathbf{u}^h\right\|_{\mathrm{DG}}. \tag{4.12a}$$

From (4.12a) we have that

$$h^{-1/2}\left\|\mathbf{u}-\mathbf{u}^h\right\|_{\mathcal{L}^2(\Gamma)} \le C(p)\,\left\|\mathbf{u}-\mathbf{u}^h\right\|_{\mathrm{DG}}, \tag{4.12b}$$

$$h^{-1/2}\left\|\,[\![\mathbf{u}-\mathbf{u}^h]\!]\,\right\|_{\mathcal{L}^2(\Gamma_{\mathcal{I}})} \le C(p)\,\left\|\mathbf{u}-\mathbf{u}^h\right\|_{\mathrm{DG}}. \tag{4.12c}$$

The discrete counterpart of $\mathcal{J}$ (3.3) reads:

$$\begin{aligned}\mathcal{J}^h[\mathbf{u}^h,\mathbf{q}^h] = &\frac{1}{2}\int_\Omega (\mathcal{H}^h\mathbf{u}^h-\mathbf{o}^h)^T(\mathcal{H}^h\mathbf{u}^h-\mathbf{o}^h)\,\mathrm{d}\mathbf{x}\\ &+\frac{\boldsymbol{\beta}}{2}\int_\Omega (\mathbf{q}^h-\mathbf{q}^h_B)^T(\mathbf{q}^h-\mathbf{q}^h_B)\,\mathrm{d}\mathbf{x}\\ &+\frac{1}{2}\int_\Omega (\nabla\mathbf{q}^h-\nabla\mathbf{q}^h_B)^T\,(\nabla\mathbf{q}^h-\nabla\mathbf{q}^h_B)\,\mathrm{d}\mathbf{x}\\ &+\mathcal{R}^h[\mathbf{u}^h,\mathbf{q}^h]\,.\end{aligned} \tag{4.13}$$

The additional term $\mathcal{R}^h$ [16] is a consistent modification of the cost functional which allows to establish the consistency of the discrete dual and optimality equations. We have the following definition [16]

Definition 4.2 (Consistent modification of the objective functional). *The term $\mathcal{R}^h[\mathbf{u}^h,\mathbf{q}^h]$ indicates a modification to the discrete functional $\mathcal{J}^h$. $\mathcal{R}^h$ must be Fréchet differentiable in both of its arguments. This modification is said to be consistent if $\mathcal{R}^h$ cancels when evaluated at the exact solutions $\mathbf{u}$, $\mathbf{q}$ of (2.2a)–(2.2c):*

$$\mathcal{R}^h[\mathbf{u},\,\mathbf{q}] = 0\,.$$

Remark. Any nontrivial consistent modification in (4.13) does impact both the value and the gradient of the discrete cost functional $\mathcal{J}^h$. The modification term $\mathcal{R}^h$ can only be expected to vanish in the limit of the discretization. Both $\mathcal{R}^h$ and its Fréchet derivative are nonzero for fixed values of $h > 0$.

As shown in [16] consistent function modifications are required for certain inverse problems to guarantee dual consistency (of the primal discretization). This concept is extended here to the optimality equation. Specifically, for the model problem (3.4), consistent modifications to $\mathcal{J}^h$ are used to introduce the stabilization terms necessary for convergence of the optimality equation discretization (2.11c).

Given (4.13) and (4.6), we can formulate the discrete inverse problem as:

$$\text{Find } \mathbf{q}_*^h = \arg\min_{\mathbf{q}^h \in \mathcal{Q}_r^h} \mathcal{J}^h[\mathbf{u}^h, \mathbf{q}^h] \text{ , subject to } (4.6) \text{ .} \tag{4.14}$$

The discrete Lagrangian functional for (4.14)–(4.6) reads:

$$\begin{aligned} \mathcal{L}^h[\mathbf{u}^h, \boldsymbol{\lambda}^h, \mathbf{q}^h] := {} & \mathcal{J}^h[\mathbf{u}^h, \mathbf{q}^h] - \mathcal{N}^h[\mathbf{u}^h, \mathbf{q}^h](\boldsymbol{\lambda}^h) \\ & + \left\langle \mathbf{f}^h, \boldsymbol{\lambda}^h \right\rangle_\Omega + \mathcal{B}^h[\mathbf{g}^h, \mathbf{q}^h](\boldsymbol{\lambda}^h) \text{ .} \end{aligned} \tag{4.15}$$

Then, the discrete adjoint problem corresponding to (4.13)–(4.6) is:

$$\begin{aligned} & \text{Find } \boldsymbol{\lambda} \in \mathcal{U}_p^h \text{ such that:} \\ & \frac{\partial \mathcal{N}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h, \boldsymbol{\lambda}^h) = \frac{\partial \mathcal{J}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) \text{ ,} \quad \forall\, \mathbf{w}^h \in \mathcal{U}^h \text{ ,} \end{aligned} \tag{4.16}$$

where, due to the special structure of $\mathcal{N}^h$

$$\begin{aligned} \frac{\partial \mathcal{N}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h, \boldsymbol{\lambda}^h) & := \mathcal{N}^h[\mathbf{w}^h, \mathbf{q}^h](\boldsymbol{\lambda}^h) \\ & := \int_\Omega \mathbf{q}^h\, \nabla \mathbf{w}^h \cdot \nabla \boldsymbol{\lambda}^h \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathcal{I} \cup \Gamma} \phi\, [\![\mathbf{w}^h]\!] \cdot [\![\boldsymbol{\lambda}^h]\!] \,\mathrm{d}\mathbf{s} \\ & \quad - \int_{\Gamma_\mathcal{I} \cup \Gamma} \Big([\![\mathbf{w}^h]\!] \cdot \{ \mathbf{q}^h\, \nabla \boldsymbol{\lambda}^h \} + \{ \mathbf{q}^h\, \nabla \mathbf{w}^h \} \cdot [\![\boldsymbol{\lambda}^h]\!] \Big) \,\mathrm{d}\mathbf{s} \end{aligned} \tag{4.17}$$

$$:= \mathcal{N}^h[\boldsymbol{\lambda}^h, \mathbf{q}^h](\mathbf{w}^h) \text{ ,} \tag{4.18}$$

and

$$\frac{\partial \mathcal{J}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) := \int_\Omega (\mathcal{H}^h\, \mathbf{w}^h)^T (\mathcal{H}^h \mathbf{u}^h - \mathbf{o}^h) \,\mathrm{d}\mathbf{x} + \frac{\partial \mathcal{R}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) \text{ .} \tag{4.19}$$

The SIPG discretization (4.6) applied to (3.5) gives the following equation, which defines the continuous adjoint variables $\overline{\boldsymbol{\lambda}}$

$$\begin{aligned} & \text{Find } \overline{\boldsymbol{\lambda}}^h \in \mathcal{U}_p^h \text{ such that:} \\ & \mathcal{N}^h[\overline{\boldsymbol{\lambda}}^h, \mathbf{q}^h](\mathbf{w}^h) = \int_\Omega (\mathcal{H}^h\, \mathbf{w}^h)^T (\mathcal{H}\mathbf{u} - \mathbf{o})^h \,\mathrm{d}\mathbf{x} \text{ ,} \ \forall\, \mathbf{w}^h \in \mathcal{U}_p^h \text{ .} \end{aligned} \tag{4.20}$$

It is shown in [16] that SIPG is *dual consistent* with no modification of the cost function ($\mathcal{R}^h = 0$ in (4.13)), i.e., its discrete adjoint (4.16) represents a consistent discretization of the continuous adjoint PDE (3.5). Given that the strong form of the adjoint equation (3.5) is also an elliptic problem, and the dual discretization is of SIPG form, the $\mathcal{L}^2$-error bound in Theorem 4.1 transfers immediately to the dual problem. Specifically, we have the following corollary:

COROLLARY 4.3. *Consider the case where the modification of the cost function does not depend on* $\mathbf{u}^h$, *i.e.,* $\mathcal{R}^h[\mathbf{u}^h, \mathbf{q}^h] = \mathcal{R}^h[\mathbf{q}^h]$.

For a sufficiently large penalty parameter $\widehat{\phi} > 0$, *the dual discretization (4.16) is a consistent and stable discretization of the adjoint PDE. Moreover, the following a priori error bounds hold:*

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^h\|_{\mathrm{DG}} \leq C\, h^{\min(p+1,s)-1} \|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,, \tag{4.21}$$

and

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^h\|_{\mathcal{L}^2(\Omega)} \leq C\, h^{\min(p+1,s)} \|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,. \tag{4.22}$$

Here C is a constant independent of the mesh size h.

Proof.

Theorem 4.1 applied to (4.20) gives

$$\|\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}^h\|_{\mathrm{DG}} \leq C\, h^{\min(p+1,s)-1} \|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,,$$
$$\|\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}^h\|_{\mathcal{L}^2(\Omega)} \leq C\, h^{\min(p+1,s)} \|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,,$$

with C a constant independent of the mesh size h.

$$\frac{\partial \mathcal{R}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h) = 0\,.$$

Subtracting (4.20) from (4.16) leads to the equation

$$\mathcal{N}^h[\boldsymbol{\lambda}^h - \overline{\boldsymbol{\lambda}}^h, \mathbf{q}^h](\mathbf{w}^h) = \text{obs. discretization residual} + \frac{\partial \mathcal{R}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h)\,.$$

which allows to bound the difference between the discrete and the continuos adjoint solutions as follows.

□

Remark. The discrete adjoint solution may be superconvergent in practice. However, we shall use the conservative error bounds (4.21)–(4.22) throughout our derivations.

It remains to be determined whether the discrete optimality condition (2.11c) is a stable and consistent discretization of the strong form optimality condition (2.2c). The optimality equation in the discrete KKT system of the problem (4.14) reads:

$$\frac{\partial \mathcal{N}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h, \boldsymbol{\lambda}^h) - \frac{\partial \mathcal{B}^h}{\partial \mathbf{q}^h}[\mathbf{g}^h, \mathbf{q}^h](\mathbf{z}^h, \boldsymbol{\lambda}^h) = \frac{\partial \mathcal{J}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h)\,,\ \forall\, \mathbf{z}^h \in \mathcal{Q}^h_r\,. \tag{4.23}$$

Noting that the admissible test functionals $\mathbf{z}^h \in \mathcal{Q}^h_r$ represent directions of differentiation, one has:

$$\frac{\partial \mathcal{N}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h, \boldsymbol{\lambda}^h) := \int_\Omega \mathbf{z}^h\, \nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h\, \mathrm{d}\mathbf{x}$$

$$- \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \left([\![\mathbf{u}^h]\!] \, \{\mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\} + [\![\boldsymbol{\lambda}^h]\!] \, \{\mathbf{z}^h \, \nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}\} \right) \, \mathrm{d}\mathbf{s} \; ,$$

$$\frac{\partial \mathcal{B}^h}{\partial \mathbf{q}^h}[\mathbf{g}^h, \mathbf{q}^h](\mathbf{z}^h, \boldsymbol{\lambda}^h) := - \int_{\Gamma} \mathbf{z}^h \, \mathbf{g}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \, \mathrm{d}\mathbf{s} \; ,$$

and from (4.13)

$$\frac{\partial \mathcal{J}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h) = \beta \int_{\Omega} \left(\mathbf{q}^h - \mathbf{q}_B^h \right) \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \int_{\Omega} \left(\nabla \mathbf{q}^h - \nabla \mathbf{q}_B^h \right) \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x}$$
$$+ \frac{\partial \mathcal{R}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h) \; .$$

The discrete optimality condition (4.23) reads:

$$\int_{\Omega} \nabla \left(\mathbf{q}^h - \mathbf{q}_B^h \right) \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \boldsymbol{\beta} \int_{\Omega} (\mathbf{q}^h - \mathbf{q}_B^h) \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \frac{\partial \mathcal{R}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h)$$
$$= \int_{\Omega} \left(\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h \right) \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} - \int_{\Gamma_{\mathcal{I}}} \left([\![\mathbf{u}^h]\!] \, \left\{ \mathbf{z}^h \, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}) \right\} + [\![\boldsymbol{\lambda}^h]\!] \, \left\{ \mathbf{z}^h \, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \right\} \right) \, \mathrm{d}\mathbf{s}$$
$$- \int_{\Gamma} \left(\mathbf{u}^h - \mathbf{g}^h \right) \mathbf{z}^h \, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}) \, \mathrm{d}\mathbf{s} - \int_{\Gamma} \boldsymbol{\lambda}^h \, \mathbf{z}^h \, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \, \mathrm{d}\mathbf{s} \; , \qquad \forall \, \mathbf{z}^h \in \mathcal{Q}_r^h \; .$$

We recast this as an equation for

$$\widehat{\mathbf{q}}^h := \mathbf{q}^h - \mathbf{q}_B^h \; ,$$

to get

$$\int_{\Omega} \nabla \widehat{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \boldsymbol{\beta} \int_{\Omega} \widehat{\mathbf{q}}^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \frac{\partial \mathcal{R}^h}{\partial \widehat{\mathbf{q}}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h) \tag{4.24}$$
$$= \int_{\Omega} \left(\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h \right) \, \mathbf{z}^h \, \mathrm{d}\mathbf{x}$$
$$- \int_{\Gamma_{\mathcal{I}}} \left([\![\mathbf{u}^h]\!] \, \left\{ \mathbf{z}^h \, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}) \right\} + [\![\boldsymbol{\lambda}^h]\!] \, \left\{ \mathbf{z}^h \, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \right\} \right) \, \mathrm{d}\mathbf{s}$$
$$- \int_{\Gamma} \left(\mathbf{u}^h - \mathbf{g}^h \right) \mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \, \mathrm{d}\mathbf{s} - \int_{\Gamma} \boldsymbol{\lambda}^h \, \nabla \mathbf{u}^h \cdot \vec{\mathbf{n}} \, \mathbf{z}^h \, \mathrm{d}\mathbf{s} \; , \quad \forall \, \mathbf{z}^h \in \mathcal{Q}_r^h \; .$$

4.3. ***A priori*** **analysis of the discrete optimality equation.** The SIPG discretization of the continuous optimality equation (3.6) reads [25]:

(4.25) Find $\widetilde{\mathbf{q}}^h \in \mathcal{Q}_r^h$ such that, for all $\mathbf{z}^h \in \mathcal{Q}_r^h$:

$$\int_{\Omega} \nabla \widetilde{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \boldsymbol{\beta} \int_{\Omega} \widetilde{\mathbf{q}}^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x}$$
$$- \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \left([\![\widetilde{\mathbf{q}}^h]\!] \, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\} + [\![\mathbf{z}^h]\!] \, \{\nabla \widetilde{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\} \right) \, \mathrm{d}\mathbf{s} + \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widetilde{\mathbf{q}}^h]\!] \, [\![\mathbf{z}^h]\!] \, \mathrm{d}\mathbf{s}$$
$$= \int_{\Omega} (\nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda})^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} \; .$$

Here the penalty parameter ϕ^{q} is defined for any edge $e_{\mathrm{q}} \in \Gamma_{\mathcal{I}}^{\mathrm{q}}$, and dimension $d \in \{2, 3\}$:

$$\phi^{\mathrm{q}} := \widehat{\phi}^{\mathrm{q}} \, \left| e_{\mathrm{q}} \right|^{1/(d-1)} \; ,$$

and $\widehat{\phi}^{\mathrm{q}} > 0$ sufficiently large for the discretization to be convergent (see, e.g., [25]).

Comparing the SIPG discretization (4.25) of the strong form optimality problem (3.6) against the discrete equation (4.24), we note that the stabilization terms, as well as the fluxes on interior faces, are not present in the linearized discretization (4.24). Hence, despite being consistent according to Definition 2.3, the discrete optimality equation (4.24) is not stable, and cannot be expected to yield useful results.

To add the necessary terms back into (4.24), a consistent modification of the functional (4.13) $\mathcal{J}^h$ is needed. More precisely, let:

$$\mathcal{R}^h[\mathbf{u}^h, \widehat{\mathbf{q}}^h] := \frac{1}{2}\int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widehat{\mathbf{q}}^h]\!] \, [\![\widehat{\mathbf{q}}^h]\!] \, \mathrm{d}\mathbf{s} - \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} [\![\widehat{\mathbf{q}}^h]\!] \, \{\nabla \widehat{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\} \, \mathrm{d}\mathbf{s} \,. \tag{4.26}$$

For $\mathbf{q}$, $\mathbf{q}_B \in \mathcal{H}^2(\Omega)$, the consistency of the functional $\mathcal{J}^h$ is preserved. Note that the jump in $\mathbf{q}^h$ over any edge $e \in \Gamma_{\mathcal{I}} \backslash \Gamma_{\mathcal{I}}^{\mathrm{q}}$ is zero. This follows from the continuity of the discrete optimal solution inside any element $\kappa^q \in \mathcal{T}_h^q$.

Remark. The modification (4.26) does not depend on $\mathbf{u}$. Consequently

$$\frac{\partial \mathcal{R}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \widehat{\mathbf{q}}^h] = 0 \,,$$

and this modification does not contribute to (4.19), therefore it does not change the dual equation (4.16). The dual consistency property is maintained, and the conclusions of the Corrolary 4.3 remain valid.

Remark. The residual $\mathcal{R}^h$ in (4.26) is directly computable from the discrete approximations $\mathbf{u}^h$ and $\mathbf{q}^h$, and it does not depend on the dual variable $\boldsymbol{\lambda}^h$. Moreover, its Fréchet derivative with respect to $\mathbf{q}^h$ along $\mathbf{z}^h \in \mathcal{Q}_p^h$ introduces suitable interior edge and penalty terms in the discrete optimality equation (4.24), while leaving the discrete adjoint equation (4.16) unmodified:

$$\begin{aligned} \frac{\partial \mathcal{R}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h) = & \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widehat{\mathbf{q}}^h]\!] \, [\![\mathbf{z}^h]\!] \, \mathrm{d}\mathbf{s} \\ & - \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \Big([\![\mathbf{z}^h]\!] \, \{\nabla \widehat{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\} + [\![\widehat{\mathbf{q}}^h]\!] \, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\} \Big) \, \mathrm{d}\mathbf{s} \,. \end{aligned} \tag{4.27}$$

With (4.26) the discrete optimality equation (4.24) reads:

$$\begin{aligned} & \int_{\Omega} \nabla \widehat{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \beta \int_{\Omega} \widehat{\mathbf{q}}^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} \\ & \quad - \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \Big([\![\widehat{\mathbf{q}}^h]\!] \, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\} + [\![\mathbf{z}^h]\!] \, \{\nabla \widehat{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\} \Big) \, \mathrm{d}\mathbf{s} + \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widehat{\mathbf{q}}^h]\!] \, [\![\mathbf{z}^h]\!] \, \mathrm{d}\mathbf{s} \\ & = \int_{\Omega} (\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h) \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} - \int_{\Gamma} \boldsymbol{\lambda}^h \, \mathbf{z}^h \, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \, \mathrm{d}\mathbf{s} \\ & \quad - \int_{\Gamma_{\mathcal{I}}} [\![\mathbf{u}^h]\!] \, \{\mathbf{z}^h \, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\} \, \mathrm{d}\mathbf{s} - \int_{\Gamma_{\mathcal{I}}} [\![\boldsymbol{\lambda}^h]\!] \, \{\mathbf{z}^h \, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}})\} \, \mathrm{d}\mathbf{s} \\ & \quad - \int_{\Gamma} (\mathbf{u}^h - \mathbf{g}^h) \mathbf{z}^h \, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}) \, \mathrm{d}\mathbf{s} \,, \qquad \forall \, \mathbf{z}^h \in \mathcal{Q}_r^h \,. \end{aligned} \tag{4.28}$$

Note that (4.28) is linear in $\mathbf{z}^h$, and is valid under any scaling of the test variables. Without loss of generality we can consider only test functions with $\left\| \mathbf{z}^h \right\|_{\mathcal{L}^\infty(\Gamma_{\mathcal{I}} \cup \Gamma)} = 1$.

We regard (4.28) as a numerical scheme applied to solve the continuous optimality equation (3.6), and seek to derive error bounds for its solution. For this, consider the following integral terms

$$\mathbb{I}_\Gamma^1 := \int_\Gamma (\mathbf{u}^h - \mathbf{g}^h)\, \mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\, \mathrm{d}\mathbf{s}\,,$$
$$\mathbb{I}_\Gamma^2 := \int_\Gamma \boldsymbol{\lambda}^h\, \mathbf{z}^h\, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}})\, \mathrm{d}\mathbf{s}\,,$$

and

$$\mathbb{I}_{\Gamma_\mathcal{I}}^1 := \int_{\Gamma_\mathcal{I}} [\![\mathbf{u}^h]\!]\, \{\mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\}\, \mathrm{d}\mathbf{s}\,,$$
$$\mathbb{I}_{\Gamma_\mathcal{I}}^2 := \int_{\Gamma_\mathcal{I}} [\![\boldsymbol{\lambda}^h]\!]\, \{\mathbf{z}^h\, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}})\}\, \mathrm{d}\mathbf{s}\,,$$

defined for arbitrary $\mathbf{z}^h \in \mathcal{Q}_r^h$. We have the following *a priori* estimates for the magnitude of these integral terms.

LEMMA 4.4. *Assume* $\mathbf{u}, \boldsymbol{\lambda} \in \mathcal{H}^s(\Omega)$, *and* $\mathbf{g} \in \mathcal{H}^{\widehat{s}}(\Gamma)$, *with* $\widehat{s} = s - 1/2$. *The following upper bounds hold:*

$$\left|\mathbb{I}_\Gamma^1\right| \le C(p)\, \left(h^{1/2} \cdot \left\|\mathbf{u} - \mathbf{u}^h\right\|_{DG} + h^{\min(p+1,\widehat{s})}\, |\mathbf{g}|_{\mathcal{H}^{\widehat{s}}(\Gamma)}\right) \cdot \left\|\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\right\|_{\mathcal{L}^2(\Gamma)} \left\|\mathbf{z}^h\right\|_{\mathcal{L}^\infty(\Gamma_\mathcal{I}\cup\Gamma)}\,,$$
$$\left|\mathbb{I}_\Gamma^2\right| \le C(p)\, h^{1/2} \cdot \left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^h\right\|_{DG} \left\|\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}\right\|_{\mathcal{L}^2(\Gamma)} \|\mathbf{z}^h\|_{\mathcal{L}^\infty(\Gamma_\mathcal{I}\cup\Gamma)}\,,$$
$$\left|\mathbb{I}_{\Gamma_\mathcal{I}}^1\right| \le C(p)\, \left\|\mathbf{u} - \mathbf{u}^h\right\|_{\mathrm{DG}} \left\|\nabla \boldsymbol{\lambda}^h\right\|_{\mathcal{L}^2(\mathcal{T}^h)} \|\mathbf{z}^h\|_{L^\infty(\Gamma_\mathcal{I}\cup\Gamma)}\,, \tag{4.29a}$$
$$\left|\mathbb{I}_{\Gamma_\mathcal{I}}^2\right| \le C(p)\, \left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^h\right\|_{\mathrm{DG}} \left\|\nabla \mathbf{u}^h\right\|_{\mathcal{L}^2(\mathcal{T}^h)} \|\mathbf{z}^h\|_{L^\infty(\Gamma_\mathcal{I}\cup\Gamma)}\,. \tag{4.29b}$$

The constants $\mathcal{C}(p,r)$, $\widehat{\mathcal{C}}(p,r)$, $\mathcal{C}_1(p,r)$, *and* $\mathcal{C}_2(p,r)$ *depend only on the orders* p *and* r *of the polynomial bases.*

Proof.

We can write:

$$\mathbb{I}_\Gamma^1 = \int_\Gamma \Big[(\mathbf{u}^h - \mathbf{u}) + (\mathbf{u} - \mathbf{g}) + (\mathbf{g} - \mathbf{g}^h)\Big]\, \mathbf{z}^h\, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\, \mathrm{d}\mathbf{s}$$
$$= \underbrace{\int_\Gamma (\mathbf{u}^h - \mathbf{u})\, \mathbf{z}^h\, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\, \mathrm{d}\mathbf{s}}_{\mathbb{I}_\Gamma^{1,1}} + \underbrace{\int_\Gamma (\mathbf{g} - \mathbf{g}^h)\, \mathbf{z}^h\, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\, \mathrm{d}\mathbf{s}}_{\mathbb{I}_\Gamma^{1,2}}\,.$$

It follows from (4.12b) that:

$$\begin{aligned}
\left|\mathbb{I}_\Gamma^{1,1}\right| &= \left|\int_\Gamma (\mathbf{u} - \mathbf{u}^h)\, \mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\, \mathrm{d}\mathbf{s}\right| \\
&\le \left\|\mathbf{u} - \mathbf{u}^h\right\|_{\mathcal{L}^2(\Gamma)} \left\|\mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\right\|_{\mathcal{L}^2(\Gamma)} \\
&\le \left\|\mathbf{u} - \mathbf{u}^h\right\|_{\mathcal{L}^2(\Gamma)} \left\|\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\right\|_{\mathcal{L}^2(\Gamma)} \left\|\mathbf{z}^h\right\|_{\mathcal{L}^\infty(\Gamma)} \\
\text{by (4.12}b) &\le C(p)\, h^{1/2} \cdot \left\|\mathbf{u} - \mathbf{u}^h\right\|_{DG} \left\|\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}\right\|_{\mathcal{L}^2(\Gamma)} \left\|\mathbf{z}^h\right\|_{\mathcal{L}^\infty(\Gamma)}\,.
\end{aligned}$$

The integral $\mathbb{I}_{\Gamma}^{1,2}$ can be bounded using classical results from polynomial approximation theory for finite element methods. Given that the boundary data $\mathbf{g} \in \mathcal{H}^{s}(\Gamma)$, the following upper bound holds for the interpolation error on Γ [9]:

$$\left\| \mathbf{g} - \mathbf{g}^h \right\|_{\mathcal{L}^2(\Gamma)} \leq C_{\mathbf{g}}(p)\, h^{\min(p+1,\widehat{s})} \, |\mathbf{g}|_{\mathcal{H}^{\widehat{s}}(\Gamma)} \,, \tag{4.30}$$

with a mesh-independent constant $C_{\mathbf{g}}$. The functional $|\cdot|$ is the broken Sobolev semi–norm on Γ. The Cauchy–Schwartz inequality gives:

$$\begin{aligned} \left| \mathbb{I}_{\Gamma}^{1,2} \right| &\leq \left\| \mathbf{g} - \mathbf{g}^h \right\|_{\mathcal{L}^2(\Gamma)} \left\| \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\|_{\mathcal{L}^2(\Gamma)} \left\| \mathbf{z}^h \right\|_{\mathcal{L}^\infty(\Gamma)} \\ &\leq C_{\mathbf{g}}(p)\, h^{\min(p+1,\widehat{s})} \, |\mathbf{g}|_{\mathcal{H}^{\widehat{s}}(\Gamma)} \left\| \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\|_{\mathcal{L}^2(\Gamma)} \left\| \mathbf{z}^h \right\|_{\mathcal{L}^\infty(\Gamma)} \,. \end{aligned}$$

This proves inequality (4.29a). Note that (4.29a) follows immediately from an identical argument by reversing the roles of $\mathbf{u}$ and $\boldsymbol{\lambda}$, since $\boldsymbol{\lambda}|_{\Gamma} = 0$, by equation (3.5).

Consider now the integral $\mathbb{I}_{\Gamma_{\mathcal{I}}}^{1}$. Since the jump operator $[\![\cdot]\!]$ is linear, and $[\![\mathbf{u}]\!] = 0$ (for any $\mathbf{u} \in \mathcal{U}$), we can write:

$$\begin{aligned} \left| \mathbb{I}_{\Gamma_{\mathcal{I}}}^{1} \right| &= \left| \int_{\Gamma_{\mathcal{I}}} [\![\mathbf{u}^h - \mathbf{u}]\!] \left\{ \mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \mathbf{ds} \right| \\ &\leq \int_{\Gamma_{\mathcal{I}}} \left| [\![\mathbf{u}^h - \mathbf{u}]\!] \left\{ \mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right| \mathbf{ds} \\ &\leq \int_{\Gamma_{\mathcal{I}}} \left| h^{-1/2} [\![\mathbf{u}^h - \mathbf{u}]\!] \, h^{1/2} \left\{ \mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right| \mathbf{ds} \\ &\leq \left\| h^{-1/2} [\![\mathbf{u}^h - \mathbf{u}]\!] \right\|_{\mathcal{L}^2(\Gamma_{\mathcal{I}})} h^{1/2} \left\| \left\{ \mathbf{z}^h \, \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right\|_{L^2(\Gamma_{\mathcal{I}})} \\ &\leq \left\| h^{-1/2} [\![\mathbf{u}^h - \mathbf{u}]\!] \right\|_{\mathcal{L}^2(\Gamma_{\mathcal{I}})} h^{1/2} \left\| \left\{ \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right\|_{L^2(\Gamma_{\mathcal{I}})} \left\| \left\{ \mathbf{z}^h \right\} \right\|_{L^\infty(\Gamma_{\mathcal{I}})} \\ &\leq \left\| h^{-1/2} [\![\mathbf{u}^h - \mathbf{u}]\!] \right\|_{\mathcal{L}^2(\Gamma_{\mathcal{I}})} \left(h \sum_{e \in \Gamma_{\mathcal{I}}} \left\| \left\{ \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right\|_{\mathcal{L}^2(\mathbf{e})}^2 \right)^{1/2} \left\| \left\{ \mathbf{z}^h \right\} \right\|_{L^\infty(\Gamma_{\mathcal{I}})} \end{aligned}$$

Since

$$\left(\frac{a+b}{2} \right)^2 \leq \frac{a^2}{2} + \frac{b^2}{2}$$

and therefore

$$\begin{aligned} \left\| \left\{ \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\} \right\|_{\mathcal{L}^2(\mathbf{e})}^2 &= \int_{\mathbf{e}} \left(\frac{(\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_+ + (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_-}{2} \right)^2 \mathbf{ds} \\ &\leq \frac{1}{2} \int_{\mathbf{e}} \left((\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_+ \right)^2 \mathbf{ds} + \frac{1}{2} \int_{\mathbf{e}} \left((\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_- \right)^2 \mathbf{ds} \\ &= \frac{1}{2} \left\| (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_+ \right\|_{\mathcal{L}^2(\mathbf{e})}^2 + \frac{1}{2} \left\| (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})_- \right\|_{\mathcal{L}^2(\mathbf{e})}^2 \end{aligned}$$

we have

$$h \sum_{e \in \Gamma_{\mathcal{I}}} \left\| \{ \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \} \right\|^2_{\mathcal{L}^2(\mathbf{e})} \leq \frac{h}{2} \sum_{\kappa \in \mathcal{T}^h} \left\| \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \right\|^2_{\mathcal{L}^2(\partial\kappa)}$$
$$(\text{trace}) \leq C(p)\, h \sum_{\kappa \in \mathcal{T}^h} h^{-1} \left\| \nabla \boldsymbol{\lambda}^h \right\|^2_{\mathcal{L}^2(\kappa)}$$
$$\leq C(p) \left\| \nabla \boldsymbol{\lambda}^h \right\|^2_{\mathcal{L}^2(\mathcal{T}^h)}$$
$$\leq C(p) \left\| \boldsymbol{\lambda}^h \right\|^2_{\mathcal{H}^s(\mathcal{T}^h)} .$$

Consequently

$$\left| \mathbb{I}^1_{\Gamma_{\mathcal{I}}} \right| \leq C(p) \left\| h^{-1/2} [\![\mathbf{u}^h - \mathbf{u}]\!] \right\|_{\mathcal{L}^2(\Gamma_{\mathcal{I}})} \left\| \nabla \boldsymbol{\lambda}^h \right\|_{\mathcal{L}^2(\mathcal{T}^h)} \left\| \{ \mathbf{z}^h \} \right\|_{L^\infty(\Gamma_{\mathcal{I}})}$$
$$\text{by (4.12}c) \ \leq C(p) \left\| \mathbf{u} - \mathbf{u}^h \right\|_{\text{DG}} \left\| \nabla \boldsymbol{\lambda}^h \right\|_{\mathcal{L}^2(\mathcal{T}^h)} \left\| \{ \mathbf{z}^h \} \right\|_{L^\infty(\Gamma_{\mathcal{I}})} .$$

Thus (4.29a) holds. Since the SIPG discretization is dual consistent, (4.29b) follows by reversing the roles of $\mathbf{u}$ and $\boldsymbol{\lambda}$,.

□

An immediate consequence of the lemma is the following corollary:

COROLLARY 4.5. *For all $p \geq 1$ and $s > 3/2$, the discrete optimality equation associated to the problem (4.14) is asymptotically consistent.*

Remark. With our specific smoothness assumptions ($s = 2$, $\widehat{s} = 3/2$), the integrals (4.29a)–(4.29b) are $\mathcal{O}(h^{1/2})$.

***A priori* convergence order.** Let $\mathbf{e}^h_{\mathbf{q}} = \widehat{\mathbf{q}}^h - \widetilde{\mathbf{q}}^h$, $\mathbf{e}_{\mathbf{u}} = \mathbf{u} - \mathbf{u}^h$, and $\mathbf{e}_{\boldsymbol{\lambda}} = \boldsymbol{\lambda} - \boldsymbol{\lambda}^h$. Subtracting (4.25) from (4.28) gives the following equation for $\mathbf{e}^h_{\mathbf{q}}$:

SIPG discretization of the continuous KKT system. The SIPG discretization of the forward problem:

$$\mathcal{N}^h[\widetilde{\mathbf{u}}^h, \widetilde{\mathbf{q}}^h](\mathbf{w}^h) = \int_\Omega \widetilde{\mathbf{q}}^h \, (\nabla \widetilde{\mathbf{u}}^h \cdot \nabla \mathbf{w}^h) \, \mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \phi \, [\![\widetilde{\mathbf{u}}^h]\!] \, [\![\mathbf{w}^h]\!] \, \mathrm{d}\mathbf{s}$$
$$Ref.(4.6): \quad - \int_{\Gamma_{\mathcal{I}} \cup \Gamma} [\![\widetilde{\mathbf{u}}^h]\!] \left\{ \widetilde{\mathbf{q}}^h \, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}) \right\} \mathrm{d}\mathbf{s}$$
$$- \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \left\{ \widetilde{\mathbf{q}}^h \, (\nabla \widetilde{\mathbf{u}}^h \cdot \vec{\mathbf{n}}) \right\} [\![\mathbf{w}^h]\!] \, \mathrm{d}\mathbf{s}$$
$$\mathcal{N}^h[\widetilde{\mathbf{u}}^h, \widetilde{\mathbf{q}}^h](\mathbf{w}^h) = \int_\Omega \mathbf{f}^h \, \mathbf{w}^h \, \mathrm{d}\mathbf{x} - \int_\Gamma \widetilde{\mathbf{q}}^h \, \mathbf{g}^h \, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}) \, \mathrm{d}\mathbf{s} + \int_\Gamma \phi \, \mathbf{g}^h \, \mathbf{w}^h \, \mathrm{d}\mathbf{s} .$$

The SIPG discretization of the continuous adjoint equation is:

$$Ref.(4.20) \ \ \mathcal{N}^h[\widetilde{\boldsymbol{\lambda}}^h, \widetilde{\mathbf{q}}^h](\mathbf{w}^h) = \int_\Omega (\mathcal{H}^h \, \mathbf{w}^h)^T (\mathcal{H} \widetilde{\mathbf{u}} - \mathbf{o})^h \, \mathrm{d}\mathbf{x} \, , \ \forall \, \mathbf{w}^h \in \mathcal{U}^h_p .$$

The SIPG discretization of the continuous optimality equation:

$$Ref.(4.25) \ \ \int_\Omega \nabla \widetilde{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \beta \int_\Omega \widetilde{\mathbf{q}}^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x}$$

$$
\begin{aligned}
&- \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \Big([\![\widetilde{\mathbf{q}}^h]\!] \left\{ \nabla \mathbf{z}^h \cdot \vec{\mathbf{n}} \right\} + [\![\mathbf{z}^h]\!] \left\{ \nabla \widetilde{\mathbf{q}}^h \cdot \vec{\mathbf{n}} \right\} \Big) \,\mathrm{d}\mathbf{s} + \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widetilde{\mathbf{q}}^h]\!] \, [\![\mathbf{z}^h]\!] \,\mathrm{d}\mathbf{s} \\
&= \int_{\Omega} (\nabla \widetilde{\mathbf{u}} \cdot \nabla \widetilde{\boldsymbol{\lambda}})^h \, \mathbf{z}^h \,\mathrm{d}\mathbf{x} \,.
\end{aligned}
$$

Discrete KKT system. The discrete forward problem reads:

$$
\begin{aligned}
\mathcal{N}^h[\widehat{\mathbf{u}}^h, \widehat{\mathbf{q}}^h](\mathbf{w}^h) = & \int_{\Omega} \widehat{\mathbf{q}}^h \, (\nabla \widehat{\mathbf{u}}^h \cdot \nabla \mathbf{w}^h) \,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \phi \, [\![\widehat{\mathbf{u}}^h]\!] \, [\![\mathbf{w}^h]\!] \,\mathrm{d}\mathbf{s} \\
Ref.(4.6): \quad & - \int_{\Gamma_{\mathcal{I}} \cup \Gamma} [\![\widehat{\mathbf{u}}^h]\!] \, \left\{ \widehat{\mathbf{q}}^h \, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}) \right\} \,\mathrm{d}\mathbf{s} \\
& - \int_{\Gamma_{\mathcal{I}} \cup \Gamma} \left\{ \widehat{\mathbf{q}}^h \, (\nabla \widehat{\mathbf{u}}^h \cdot \vec{\mathbf{n}}) \right\} \, [\![\mathbf{w}^h]\!] \,\mathrm{d}\mathbf{s} \\
\mathcal{N}^h[\widehat{\mathbf{u}}^h, \widehat{\mathbf{q}}^h](\mathbf{w}^h) = & \int_{\Omega} \mathbf{f}^h \, \mathbf{w}^h \,\mathrm{d}\mathbf{x} - \int_{\Gamma} \widehat{\mathbf{q}}^h \, \mathbf{g}^h \, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}) \,\mathrm{d}\mathbf{s} + \int_{\Gamma} \phi \, \mathbf{g}^h \, \mathbf{w}^h \,\mathrm{d}\mathbf{s} \,.
\end{aligned}
$$

The discrete adjoint problem is:

$$
Ref.(4.16) \quad \mathcal{N}^h[\widehat{\boldsymbol{\lambda}}^h, \widehat{\mathbf{q}}^h](\mathbf{w}^h) = \int_{\Omega} (\mathcal{H}^h \, \mathbf{w}^h)^T (\mathcal{H}^h \widehat{\mathbf{u}}^h - \mathbf{o}^h) \,\mathrm{d}\mathbf{x}
$$

The discrete optimality condition reads:

$$
\begin{aligned}
Ref.(4.28) \quad & \int_{\Omega} \nabla \widehat{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \,\mathrm{d}\mathbf{x} + \beta \int_{\Omega} \widehat{\mathbf{q}}^h \, \mathbf{z}^h \,\mathrm{d}\mathbf{x} \\
& - \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \Big([\![\widehat{\mathbf{q}}^h]\!] \left\{ \nabla \mathbf{z}^h \cdot \vec{\mathbf{n}} \right\} + [\![\mathbf{z}^h]\!] \left\{ \nabla \widehat{\mathbf{q}}^h \cdot \vec{\mathbf{n}} \right\} \Big) \,\mathrm{d}\mathbf{s} + \int_{\Gamma_{\mathcal{I}}^{\mathrm{q}}} \phi^{\mathrm{q}} \, [\![\widehat{\mathbf{q}}^h]\!] \, [\![\mathbf{z}^h]\!] \,\mathrm{d}\mathbf{s} \\
& = \int_{\Omega} (\nabla \widehat{\mathbf{u}}^h \cdot \nabla \widehat{\boldsymbol{\lambda}}^h) \, \mathbf{z}^h \,\mathrm{d}\mathbf{x} + R_{\mathbb{I}}
\end{aligned}
$$

SIPG of continuous forward eqn:	Discrete forward problem:
$\mathcal{N}^h[\overline{\mathbf{u}}^h, \overline{\mathbf{q}}^h](\mathbf{w}^h) = \int_\Omega \mathbf{f}^h\, \mathbf{w}^h \,\mathrm{d}\mathbf{x}$ $- \int_\Gamma \overline{\mathbf{q}}^h\, \mathbf{g}^h\, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}})\,\mathrm{d}\mathbf{s}$ $+ \int_\Gamma \phi\, \mathbf{g}^h\, \mathbf{w}^h \,\mathrm{d}\mathbf{s}\,.$	$\mathcal{N}^h[\widehat{\mathbf{u}}^h, \widehat{\mathbf{q}}^h](\mathbf{w}^h) = \int_\Omega \mathbf{f}^h\, \mathbf{w}^h \,\mathrm{d}\mathbf{x}$ $- \int_\Gamma \widehat{\mathbf{q}}^h\, \mathbf{g}^h\, (\nabla \mathbf{w}^h \cdot \vec{\mathbf{n}})\,\mathrm{d}\mathbf{s}$ $+ \int_\Gamma \phi\, \mathbf{g}^h\, \mathbf{w}^h \,\mathrm{d}\mathbf{s}\,.$
SIPG of continuous adjoint eqn: $\mathcal{N}^h[\overline{\boldsymbol{\lambda}}^h, \overline{\mathbf{q}}^h](\mathbf{w}^h) =$ $\int_\Omega (\mathcal{H}^h\, \mathbf{w}^h)^T (\mathcal{H}\overline{\mathbf{u}} - \mathbf{o})^h \,\mathrm{d}\mathbf{x}$	Discrete adjoint problem: $\mathcal{N}^h[\widehat{\boldsymbol{\lambda}}^h, \widehat{\mathbf{q}}^h](\mathbf{w}^h) =$ $\int_\Omega (\mathcal{H}^h\, \mathbf{w}^h)^T (\mathcal{H}^h\widehat{\mathbf{u}}^h - \mathbf{o}^h) \,\mathrm{d}\mathbf{x}$
SIPG of continuous optimality eqn: $\int_\Omega \nabla\overline{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \,\mathrm{d}\mathbf{x} + \beta \int_\Omega \overline{\mathbf{q}}^h\, \mathbf{z}^h \,\mathrm{d}\mathbf{x}$ $- \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\overline{\mathbf{q}}^h]\!]\, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\}\,\mathrm{d}\mathbf{s}$ $- \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\mathbf{z}^h]\!]\, \{\nabla \overline{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\}\,\mathrm{d}\mathbf{s}$ $+ \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \phi^{\mathrm{q}}\, [\![\overline{\mathbf{q}}^h]\!]\, [\![\mathbf{z}^h]\!]\,\mathrm{d}\mathbf{s}$ $= \int_\Omega (\nabla \overline{\mathbf{u}} \cdot \nabla \overline{\boldsymbol{\lambda}})^h\, \mathbf{z}^h \,\mathrm{d}\mathbf{x}$	Discrete optimality condition: $\int_\Omega \nabla\widehat{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \,\mathrm{d}\mathbf{x} + \beta \int_\Omega \widehat{\mathbf{q}}^h\, \mathbf{z}^h \,\mathrm{d}\mathbf{x}$ $- \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\widehat{\mathbf{q}}^h]\!]\, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\}\,\mathrm{d}\mathbf{s}$ $- \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\mathbf{z}^h]\!]\, \{\nabla \widehat{\mathbf{q}}^h \cdot \vec{\mathbf{n}}\}\,\mathrm{d}\mathbf{s}$ $+ \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \phi^{\mathrm{q}}\, [\![\widehat{\mathbf{q}}^h]\!]\, [\![\mathbf{z}^h]\!]\,\mathrm{d}\mathbf{s}$ $= \int_\Omega (\nabla \widehat{\mathbf{u}}^h \cdot \nabla \widehat{\boldsymbol{\lambda}}^h)\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} + R_{\mathbb{I}}$

5. *A priori* convergence order. Let $\mathbf{e}^h_{\mathbf{q}} = \widehat{\mathbf{q}}^h - \widetilde{\mathbf{q}}^h$, $\mathbf{e}_{\mathbf{u}} = \mathbf{u} - \mathbf{u}^h$, and $\mathbf{e}_{\boldsymbol{\lambda}} = \boldsymbol{\lambda} - \boldsymbol{\lambda}^h$. Subtracting (4.25) from (4.28) gives the following equation for $\mathbf{e}^h_{\mathbf{q}}$:

Find $\mathbf{e}^h_{\mathbf{q}} \in \mathcal{Q}^h_p$ such that, for all $\mathbf{z}^h \in \mathcal{Q}^h_r$:

$$\int_\Omega \nabla \mathbf{e}^h_{\mathbf{q}} \cdot \nabla \mathbf{z}^h \,\mathrm{d}\mathbf{x} + \beta \int_\Omega \mathbf{e}^h_{\mathbf{q}}\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} - \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \Big([\![\mathbf{e}^h_{\mathbf{q}}]\!]\, \{\nabla \mathbf{z}^h \cdot \vec{\mathbf{n}}\} + \{\nabla \mathbf{e}^h_{\mathbf{q}} \cdot \vec{\mathbf{n}}\}\, [\![\mathbf{z}^h]\!] \Big)\,\mathrm{d}\mathbf{s} + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \phi^{\mathrm{q}}\, [\![\mathbf{e}$$

$$= \int_\Omega (\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h - \nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda})\, \mathbf{z}^h \,\mathrm{d}\mathbf{x}$$
$$- \int_\Gamma \boldsymbol{\lambda}^h\, \mathbf{z}^h\, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}})\,\mathrm{d}\mathbf{s} - \int_\Gamma (\mathbf{u}^h - \mathbf{g}^h)\mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\,\mathrm{d}\mathbf{s}$$
$$- \int_{\Gamma_{\mathcal{I}}} [\![\mathbf{u}^h]\!]\, \{\mathbf{z}^h\, (\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}})\}\,\mathrm{d}\mathbf{s} - \int_{\Gamma_{\mathcal{I}}} [\![\boldsymbol{\lambda}^h]\!]\, \{\mathbf{z}^h\, (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}})\}\,\mathrm{d}\mathbf{s}$$
$$= \int_\Omega (\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h)\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} - \int_\Omega (\nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda}^h)\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} + \int_\Omega (\nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda}^h)\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} - \int_\Omega (\nabla \mathbf{u} \cdot \nabla \boldsymbol{\lambda})\, \mathbf{z}^h \,\mathrm{d}$$
$$- \mathbb{I}^1_\Gamma - \mathbb{I}^2_\Gamma - \mathbb{I}^1_{\Gamma_{\mathcal{I}}} - \mathbb{I}^2_{\Gamma_{\mathcal{I}}}\,.$$
$$= -\int_\Omega (\nabla \mathbf{e}_{\mathbf{u}} \cdot \nabla \boldsymbol{\lambda}^h)\, \mathbf{z}^h \,\mathrm{d}\mathbf{x} - \int_\Omega (\nabla \mathbf{u} \cdot \nabla \mathbf{e}_{\boldsymbol{\lambda}})\, \mathbf{z}^h \,\mathrm{d}\mathbf{x}$$
$$- \mathbb{I}^1_\Gamma - \mathbb{I}^2_\Gamma - \mathbb{I}^1_{\Gamma_{\mathcal{I}}} - \mathbb{I}^2_{\Gamma_{\mathcal{I}}}\,.$$

$$= -\int_{\Omega}(\nabla \mathbf{e_u}\cdot\nabla\boldsymbol{\lambda}^h)\,\mathbf{z}^h\,\mathrm{d}\mathbf{x} - \int_{\Omega}(\nabla \mathbf{u}^h\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}})\,\mathbf{z}^h\,\mathrm{d}\mathbf{x} - \int_{\Omega}(\nabla \mathbf{e_u}\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}})\,\mathbf{z}^h\,\mathrm{d}\mathbf{x}$$
$$-\mathbb{I}^1_{\Gamma} - \mathbb{I}^2_{\Gamma} - \mathbb{I}^1_{\Gamma_{\mathcal{I}}} - \mathbb{I}^2_{\Gamma_{\mathcal{I}}}$$

Consider the following elliptic error equation defined over the broken Sobolev space $\mathcal{H}^s(\mathcal{T}^h)$:

$$\text{(5.1)}\qquad \begin{aligned} -\Delta\,\overline{\mathbf{e}}_{\mathbf{q}} + \beta\,\overline{\mathbf{e}}_{\mathbf{q}} &= -\nabla\mathbf{e_u}\cdot\nabla\boldsymbol{\lambda}^h - \nabla\mathbf{u}^h\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}} - \nabla\mathbf{e_u}\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}}\,,\ \mathbf{x}\in\Omega\,,\\ \nabla\overline{\mathbf{e}}_{\mathbf{q}}\cdot\vec{\mathbf{n}} &= 0\,,\ \mathbf{x}\in\Gamma\,. \end{aligned}$$

Under our solution regularity assumptions, the Lax–Milgram theorem [25, Theorem 2.8] guarantees the existence of a constant C independent on h, such that:

$$\begin{aligned} \|\overline{\mathbf{e}}_{\mathbf{q}}\|_{\mathcal{H}^s(\mathcal{T}^h)} &\le C\left\|\nabla\mathbf{e_u}\cdot\nabla\boldsymbol{\lambda}^h + \nabla\mathbf{u}^h\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}} + \nabla\mathbf{e_u}\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}}\right\|_{\mathcal{L}^2(\mathcal{T}^h)}\\ &\le C\sum_{\kappa\in\mathcal{T}^h}\left|\int_{\mathcal{T}_k}\left(\nabla\mathbf{e_u}\cdot\nabla\boldsymbol{\lambda}^h + \nabla\mathbf{u}^h\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}} + \nabla\mathbf{e_u}\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}}\right)\mathrm{d}\mathbf{x}\right|\\ &\le C\sum_{\kappa\in\mathcal{T}^h}\Big(\|\nabla\mathbf{e_u}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\boldsymbol{\lambda}^h\|_{\mathcal{L}^2(\kappa)} + \|\nabla\mathbf{e}_{\boldsymbol{\lambda}}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\mathbf{u}^h\|_{\mathcal{L}^2(\kappa)}\\ &\qquad + \|\nabla\mathbf{e_u}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\mathbf{e}_{\boldsymbol{\lambda}}\|_{\mathcal{L}^2(\kappa)}\Big)\,. \end{aligned}\tag{5.2}$$

Now, from the discrete Cauchy inequality, and Theorem 4.1, one has:

$$\begin{aligned} \sum_{\kappa\in\mathcal{T}^h}\|\nabla\mathbf{e_u}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\boldsymbol{\lambda}^h\|_{\mathcal{L}^2(\kappa)} &\le \left(\sum_{\kappa\in\mathcal{T}^h}\|\nabla\mathbf{e_u}\|^2_{\mathcal{L}^2(\kappa)}\right)^{1/2}\left(\sum_{\kappa\in\mathcal{T}^h}\|\nabla\boldsymbol{\lambda}^h\|^2_{\mathcal{L}^2(\kappa)}\right)^{1/2}\\ &= \|\nabla\mathbf{e_u}\|_{\mathcal{L}^2(\mathcal{T}^h)}\cdot\|\nabla\boldsymbol{\lambda}^h\|_{\mathcal{L}^2(\mathcal{T}^h)}\\ &\le C_{\mathbf{u}}(p)\,h^{\min(p+1,s)-1}\,\|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,\|\nabla\boldsymbol{\lambda}^h\|_{\mathcal{L}^2(\mathcal{T}^h)}\,. \end{aligned}$$

Similarly, using the convergence result for the dual problem (Corollary 4.3), we obtain:

$$\sum_{\kappa\in\mathcal{T}^h}\|\nabla\mathbf{e}_{\boldsymbol{\lambda}}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\mathbf{u}^h\|_{\mathcal{L}^2(\kappa)} \le C_{\boldsymbol{\lambda}}(p)\,h^{\min(p+1,s)-1}\,\|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,\|\nabla\mathbf{u}^h\|_{\mathcal{L}^2(\mathcal{T}^h)}\,,$$

and

$$\sum_{\kappa\in\mathcal{T}^h}\|\nabla\mathbf{e_u}\|_{\mathcal{L}^2(\kappa)}\,\|\nabla\mathbf{e}_{\boldsymbol{\lambda}}\|_{\mathcal{L}^2(\kappa)} \le C_{\mathbf{u}}(p)\,C_{\boldsymbol{\lambda}}(p)\,h^{2\min(p+1,s)-2}\,\|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,\|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}^h)}\,.$$

Substituting these upper bounds in equation (5.2), it follows that:

$$\|\overline{\mathbf{e}}_{\mathbf{q}}\|_{\mathcal{H}^s(\mathcal{T}^h)} \le C\,h^{\min(p+1,s)-1}\left(\|\mathbf{u}^h\|_{\mathcal{H}^s(\mathcal{T}^h)} + \|\boldsymbol{\lambda}^h\|_{\mathcal{H}^s(\mathcal{T}^h)}\right)\,.\tag{5.3}$$

The unperturbed SIPG discretization of (5.1) is:

(5.4) Find $\overline{\mathbf{e}}^h_{\mathbf{q}}\in\mathcal{Q}^h_p$ such that, for all $\mathbf{z}^h\in\mathcal{Q}^h_r$:

$$\begin{aligned} &\int_{\Omega}\nabla\overline{\mathbf{e}}^h_{\mathbf{q}}\cdot\nabla\mathbf{z}^h\,\mathrm{d}\mathbf{x} + \beta\int_{\Omega}\overline{\mathbf{e}}^h_{\mathbf{q}}\,\mathbf{z}^h\,\mathrm{d}\mathbf{x}\\ &\quad - \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}}\left([\![\overline{\mathbf{e}}^h_{\mathbf{q}}]\!]\,\{\nabla\mathbf{z}^h\cdot\vec{\mathbf{n}}\} + \{\nabla\overline{\mathbf{e}}^h_{\mathbf{q}}\cdot\vec{\mathbf{n}}\}\,[\![\mathbf{z}^h]\!]\right)\mathrm{d}\mathbf{s} + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}}\phi^{\mathrm{q}}\,[\![\overline{\mathbf{e}}^h_{\mathbf{q}}]\!]\cdot[\![\mathbf{z}^h]\!]\,\mathrm{d}\mathbf{s}\\ &= -\int_{\Omega}\nabla\mathbf{e_u}\cdot\nabla\boldsymbol{\lambda}^h\,\mathbf{z}^h\,\mathrm{d}\mathbf{x} - \int_{\Omega}\nabla\mathbf{u}^h\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}}\,\mathbf{z}^h\,\mathrm{d}\mathbf{x} - \int_{\Omega}\nabla\mathbf{e_u}\cdot\nabla\mathbf{e}_{\boldsymbol{\lambda}}\,\mathbf{z}^h\,\mathrm{d}\mathbf{x}\,. \end{aligned}$$

The discrete solution $\overline{\mathbf{e}}^h_{\mathbf{q}}$ of (5.4) verifies the *a priori* error estimate given in Theorem 4.1, hence:

$$\begin{aligned}\|\overline{\mathbf{e}}_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\|_{\mathrm{DG}} &\leq C\,h^{\min(p+1,s)-1}\|\overline{\mathbf{e}}_{\mathbf{q}}\|_{\mathcal{H}^s(\mathcal{T}^h)}\\ &\leq C\,h^{2\min(p+1,s)-2}\left(\|\mathbf{u}^h\|_{\mathcal{H}^s(\mathcal{T}^h)} + \|\boldsymbol{\lambda}^h\|_{\mathcal{H}^s(\mathcal{T}^h)}\right)\,.\end{aligned}$$

One obtains:

$$\begin{aligned}\|\overline{\mathbf{e}}^h_{\mathbf{q}}\|_{\mathrm{DG}} &\leq \|\overline{\mathbf{e}}_{\mathbf{q}}\|_{\mathrm{DG}} + \|\overline{\mathbf{e}}_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\|_{\mathrm{DG}}\\ &\leq C\,h^{\min(p+1,s)-1}\left(\|\mathbf{u}^h\|_{\mathcal{H}^s(\mathcal{T}^h)} + \|\boldsymbol{\lambda}^h\|_{\mathcal{H}^s(\mathcal{T}^h)}\right)\,.\end{aligned}$$

The equation for the discrete optimal solution error, denoted by $\mathbf{e}^h_{\mathbf{q}}$, obtained by subtracting (4.25) from (4.28), is a perturbed SIPG discretization. The perturbation is given by the sum of the boundary and inner face integrals

$$R_I := \mathbb{I}^1_{\Gamma} + \mathbb{I}^2_{\Gamma} + \mathbb{I}^1_{\Gamma_{\mathcal{I}}} + \mathbb{I}^2_{\Gamma_{\mathcal{I}}}\,.$$

Subtracting the SIPG discretization of the error equation from the equation for $\mathbf{e}^h_{\mathbf{q}}$, leads to:

$$\begin{aligned}(5.5)\quad \mathcal{N}_{\mathbf{e}_{\mathbf{q}}}\left(\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}, \mathbf{z}^h\right) &= \int_{\Omega}\nabla\left(\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\right)\cdot\nabla\mathbf{z}^h\,\mathrm{d}\mathbf{x} + \boldsymbol{\beta}\int_{\Omega}\left(\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\right)\mathbf{z}^h\,\mathrm{d}\mathbf{x}\\ &\quad - \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}}\left(\llbracket\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\rrbracket\cdot\{\nabla\mathbf{z}^h\} + \{\nabla\left(\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\right)\}\cdot\llbracket\mathbf{z}^h\rrbracket\right)\mathrm{d}\mathbf{s}\\ &\quad + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}}\phi^{\mathrm{q}}\,\llbracket\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\rrbracket\cdot\llbracket\mathbf{z}^h\rrbracket\,\mathrm{d}\mathbf{s}\\ &= -R_I\,.\end{aligned}$$

The bilinear form $\mathcal{N}_{\mathbf{e}_{\mathbf{q}}}(\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}, \mathbf{z}^h) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is of SIPG form. Thus, it is continuous and coercive [25]. The right hand side of equation (5.5) is linear in the test functionals $\mathbf{z}^h$. By the Lax–Milgram theorem applied on $\mathbb{R}^N$, and (4.29a)–(4.29b), we obtain the following a priori bound:

$$\begin{aligned}(5.6)\quad \|\mathbf{e}^h_{\mathbf{q}} - \overline{\mathbf{e}}^h_{\mathbf{q}}\|_{\mathcal{H}^s(\mathcal{T}^h_{\mathrm{q}})} &\leq C\,|\mathbb{I}^1_{\Gamma} + \mathbb{I}^2_{\Gamma} + \mathbb{I}^1_{\Gamma_{\mathcal{I}}} + \mathbb{I}^2_{\Gamma_{\mathcal{I}}}|\\ &\leq C(r,p)\,h^{\min(p+1,s)-3/2}\left(\|\mathbf{u}\|_{\mathcal{L}^2(\mathcal{T}^h)} + \|\boldsymbol{\lambda}\|_{\mathcal{L}^2(\mathcal{T}^h)}\right)\,.\end{aligned}$$

6. Solution of the optimization problem. We seek to solve the optimization problem by a gradient based method such as quasi-Newton on nonlinear conjugate gradients. Such methods are well suited for large scale optimization problems. In a discrete framework we need to compute the reduced gradient $\nabla_{\mathbf{q}^h}\mathcal{J}^h$ of the discrete cost functional (4.26). The application of this approach requires only the forward and the adjoint models, but not the explicit optimality equation.

6.1. The discrete reduced gradient. The gradient $\nabla_{\mathbf{q}^h}\mathcal{J}^h$ of (4.26) is defined by the following identity:

$$(6.1)\quad \left\langle\nabla_{\mathbf{q}^h}\mathcal{J}^h, \boldsymbol{\delta}\mathbf{q}^h\right\rangle_{\Omega} := \mathcal{J}^h_{\mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\boldsymbol{\delta}\mathbf{q}^h) + \mathcal{J}^h_{\mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\boldsymbol{\delta}\mathbf{u}^h)\,,\ \forall\,\boldsymbol{\delta}\mathbf{q}^h \in \mathcal{Q}^h_r\,,$$

where

$$\boldsymbol{\delta}\mathbf{u}^h := \frac{\partial \mathbf{u}^h}{\partial \mathbf{q}^h}[\mathbf{q}^h](\boldsymbol{\delta}\mathbf{q}^h) \;\in\; \mathcal{Q}_r^h \,.$$

The discrete adjoint solution $\boldsymbol{\lambda}^h$ is a valid test function for the primal model (4.6). The Fréchet derivative of (4.6), tested with $\boldsymbol{\lambda}^h$, in the direction $(\boldsymbol{\delta}\mathbf{u}^h, \boldsymbol{\delta}\mathbf{q}^h) \in \mathcal{U}_p^h \times \mathcal{Q}_r^h$, gives the discrete tangent linear model:

$$\begin{aligned}(6.2)\qquad & \frac{\partial \mathcal{N}^h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \boldsymbol{\lambda}^h](\boldsymbol{\delta}\mathbf{q}^h) + \frac{\partial \mathcal{N}^h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \boldsymbol{\lambda}^h](\boldsymbol{\delta}\mathbf{u}^h) \\ & - \frac{\partial \mathcal{B}^h}{\partial \mathbf{q}^h}[\mathbf{g}^h, \boldsymbol{\lambda}^h](\boldsymbol{\delta}\mathbf{q}^h) - \frac{\partial \mathcal{B}^h}{\partial \mathbf{u}^h}[\mathbf{g}^h, \boldsymbol{\lambda}^h](\boldsymbol{\delta}\mathbf{u}^h) = 0 \,.\end{aligned}$$

From the adjoint equation (4.16) we get:

$$(6.3)\qquad \mathcal{N}^{h,*}(\boldsymbol{\delta}\mathbf{u}^h, \boldsymbol{\lambda}^h) = \mathcal{J}^h_{\mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\boldsymbol{\delta}\mathbf{u}^h) \,.$$

Then, using (6.3) and (6.1), and the integration by parts formula on $\mathcal{T}^h_{\mathrm{q}}$,

$$\begin{aligned}\int_\Omega \nabla \mathbf{q}^h \cdot \nabla \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} = & - \int_\Omega \Delta \mathbf{q}^h \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} \\ & + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}} \cup \Gamma} \{\nabla \mathbf{q}^h\} \cdot [\![\boldsymbol{\delta}\mathbf{q}^h]\!] \,\mathrm{d}\mathbf{s} + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\nabla \mathbf{q}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h\} \,\mathrm{d}\mathbf{s} \,,\end{aligned}$$

we find that:

$$\begin{aligned}(6.4)\qquad \left\langle \nabla_{\mathbf{q}^h} \mathcal{J}^h, \boldsymbol{\delta}\mathbf{q}^h \right\rangle_\Omega := & \int_\Omega \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_\Omega \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} + \int_\Gamma \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_\Gamma \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} \\ & + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} \,.\end{aligned}$$

The cell, boundary and interior face contributions to the gradient, are computed, respectively, through the following equations:

$$\begin{aligned}\int_\Omega \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_\Omega \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} := & - \int_\Omega \Delta \mathbf{q}^h \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} + \int_\Omega \Delta \mathbf{q}^h_B \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} \\ & + \beta \int_\Omega (\mathbf{q}^h - \mathbf{q}^h_B) \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} - \int_\Omega \nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{x} \,, \\ \int_\Gamma \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_\Gamma \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} := & \int_\Gamma \Big(\mathbf{u}^h \,(\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}}) + (\nabla \mathbf{u}^h \cdot \vec{\mathbf{n}}) \,\boldsymbol{\lambda}^h \Big) \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} \\ & + \int_\Gamma (\nabla \mathbf{q}^h \cdot \vec{\mathbf{n}} \;-\; \nabla \mathbf{q}^h_B \cdot \vec{\mathbf{n}}) \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} - \int_\Gamma \mathbf{g}^h \,\nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \,\boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} \,, \\ \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \nabla_{\mathbf{q}^h} \mathcal{J}^h \Big|_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \boldsymbol{\delta}\mathbf{q}^h \,\mathrm{d}\mathbf{s} := & \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} [\![\nabla \mathbf{q}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h\} \,\mathrm{d}\mathbf{s} + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \phi \,[\![\boldsymbol{\delta}\mathbf{q}^h]\!] \cdot [\![\mathbf{q}^h]\!] \,\mathrm{d}\mathbf{s} \\ & + \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \{\nabla \mathbf{q}^h\} \cdot [\![\boldsymbol{\delta}\mathbf{q}^h]\!] \,\mathrm{d}\mathbf{s} - \int_{\Gamma^{\mathrm{q}}_{\mathcal{I}}} \{\nabla \mathbf{q}^h_B\} \cdot [\![\boldsymbol{\delta}\mathbf{q}^h]\!] \,\mathrm{d}\mathbf{s} \\ & + \int_{\Gamma_{\mathcal{I}}} \Big([\![\mathbf{u}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h \,\nabla \boldsymbol{\lambda}^h\} + \{\boldsymbol{\delta}\mathbf{q}^h \,\nabla \mathbf{u}^h\} \cdot [\![\boldsymbol{\lambda}^h]\!] \Big) \,\mathrm{d}\mathbf{s} \,.\end{aligned}$$

LEMMA 6.1. *If we have convergence of the discrete primal and dual solutions, then*

$$\lim_{h \to 0} \nabla_{\mathbf{q}^h} \mathcal{J}^h = \nabla_{\mathbf{q}} \mathcal{J} \,,$$

i.e., the reduced discrete gradient is asymptotically consistent.

Proof. Use equations (3.9), (6.4), and Lemma 4.4. ☐

Remark. In addition to cell and boundary components, which are present in the continuous gradient equation (3.9), the discrete gradient (6.4) is also influenced by traces on the interior element faces. Even though the contribution defined on interior faces $\nabla_{\mathbf{q}^h}\mathcal{J}^h\Big|_{\Gamma^{\mathrm{q}}_{\mathcal{I}}}$ vanishes in the limit of the discretization, it is nonzero for fixed $h > 0$.

6.1.1. Computation of the discrete gradient. We now discuss the computation of the reduced gradient from an implementation perspective. The more straightforward, albeit significantly less efficient approach is the component–wise calculation of gradient entries. Let N be the total number of degrees of freedom (DoFs) for the mesh $\mathcal{T}^h_{\mathrm{q}}$, and the size of the discrete solution $\mathbf{q}^h$. It follows that $N = I \times R$, where I is the total number of elements defined on $\mathcal{T}^h_{\mathrm{q}}$, and the constant R denotes the number of degrees of freedom per element.

Assume the following ordering in the components of the discrete reduced gradient: the $\big((i-1)\times R + j\big)$-th entry corresponds to the j-th DoF defined in the i-th mesh element κ^q_i. Note that all the discontinuous Galerkin DoFs are defined inside the mesh elements, given that the numerical solution is discontinuous at the inter-element boundaries.

Let

$$\boldsymbol{\delta}\mathbf{q}^h(\mathbf{x}) := \begin{cases} \chi^q_j(\mathbf{x}) & , \quad \mathbf{x} \in \kappa^q_i \\ 0 & , \quad \mathbf{x} \notin \kappa^q_i \end{cases} , \tag{6.5}$$

in (6.4), and any indices $i = 1 \dots I$, and $j = 1 \dots R$. Then

$$\left\langle \nabla_{\mathbf{q}^h}\mathcal{J}^h, \boldsymbol{\delta}\mathbf{q}^h \right\rangle_\Omega = \left(\nabla_{\mathbf{q}^h}\mathcal{J}^h\right)_{(i-1)\times R+j} .$$

With this observation, $I \times R$ evaluations of the functionals (6.4) are needed for a complete gradient computation. For practical values of I, we need a significantly better approach.

The *sparse* structure of the test functional (6.5) can be exploited to dramatically increase the efficiency of the gradient computation. To see this, use equation (6.5) in (6.4), to get:

$$(\nabla_{\mathbf{q}^h}\mathcal{J}^h)_{(i-1)\times R+j} := \mathcal{G}_\Omega + \mathcal{G}_\Gamma + \mathcal{G}_{\Gamma_{\mathcal{I}}} , \tag{6.6}$$

where the volume and boundary integrals reduce to

$$\mathcal{G}_\Omega := \int_{\kappa^q_i} \nabla_{\mathbf{q}^h}\mathcal{J}^h\Big|_\Omega \, \chi^q_j \, \mathrm{d}\mathbf{x} ,$$

and

$$\mathcal{G}_\Gamma := \sum_{e \in \Gamma \cap \partial\kappa^q_i} \int_e \nabla_{\mathbf{q}^h}\mathcal{J}^h\Big|_\Gamma \, \chi^q_j \, \mathrm{d}\mathbf{s} .$$

The inner face integrals warrant further examination. Consider for brevity only the integral below (the rest of the inner face terms in (6.4) are treated similarly):

$$\int_{\Gamma^q_{\mathcal{I}}} [\![\nabla \mathbf{q}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h\} \, \mathbf{ds} = \sum_{e \in \Gamma^q_{\mathcal{I}}} \int_e [\![\nabla \mathbf{q}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h\} \, \mathbf{ds}$$
$$= \sum_{e \in \Gamma^q_{\mathcal{I}}} \frac{1}{2} \int_e \left(\nabla \mathbf{q}^h_+ \cdot \vec{\mathbf{n}} - \nabla \mathbf{q}^h_- \cdot \vec{\mathbf{n}} \right) \left(\boldsymbol{\delta}\mathbf{q}^h_+ + \boldsymbol{\delta}\mathbf{q}^h_- \right) \mathbf{ds} \, ,$$

where we let the edge $e = \kappa_+ \cap \kappa_-$. Let $\kappa^q_i \equiv \kappa_+$. Then the integral above reduces to

$$\int_{\Gamma^q_{\mathcal{I}}} [\![\nabla \mathbf{q}^h]\!] \cdot \{\boldsymbol{\delta}\mathbf{q}^h\} \, \mathbf{ds} = \sum_{e \in \partial\kappa^q_i \cap \Gamma^q_{\mathcal{I}}} \int_e \left(\nabla \mathbf{q}^h_+ \cdot \vec{\mathbf{n}} - \nabla \mathbf{q}^h_- \cdot \vec{\mathbf{n}} \right) \boldsymbol{\delta}\mathbf{q}^h_+ \, \mathbf{ds} \, .$$

This locality implies that all three integrals in (6.6) can be assembled in parallel over all DoF indices $1 \dots N$. Hence, the cost of a reduced gradient computation becomes comparable to that of evaluating a single gradient component on the triangulation $\mathcal{T}^h_{\mathrm{q}}$ using equation (6.4).

7. Targeted *a posteriori* error estimation. Aposteriori estimates are important in practical computations, in order to steer the grid adaptation such as to obtain the target accuracy with small computational effort. In the inverse problem setting one needs to adjust the grids for the forward problem, the adjoint problem, and the optimality equation, in order to control the error in the inverse problem solution. Note that controlling the local errors in each of the problems (primal, dual, and optimal) is likely to be an inefficient strategy, since only the impact of these local errors on the accuracy of the inverse solution is of interest.

Our approach to posterior error estimation follows Becker and Vexler [7]. We generalize their work to cover the case where the parameter space $\mathcal{Q}$ is infinite dimensional. From this we derive a computational procedure applicable to our formulation using discrete adjoints. Our derivation of the error estimates uses an approach based on perturbations of the KKT system, and better explains the impact and of different numerical errors and the magnitude of different error estimation terms.

7.1. Aposteriori error estimation based on an error functional. Consider an error functional

$$\mathcal{E}[\mathbf{q}] : \mathcal{Q} \to \mathbb{R} \, . \tag{7.1}$$

We assume an infinite dimensional parameter space $\mathcal{Q}$. The gradient of $\mathcal{E}$ is defined by identification from

$$\left\langle \nabla_{\mathbf{q}} \mathcal{E}, \boldsymbol{\delta}\mathbf{q} \right\rangle = \mathcal{E}_{\mathbf{q}}[\mathbf{q}](\boldsymbol{\delta}\mathbf{q}) \, , \; \forall \, \boldsymbol{\delta}\mathbf{q} \in \mathcal{Q} \, .$$

As in [7], we seek a error representation of the type:

$$\mathcal{E}[\mathbf{q}^h] - \mathcal{E}[\mathbf{q}] \approx E^h + \text{h.o.t.} \, ,$$

where E^h is computable in practice, and the higher order terms (h.o.t) are ignored. This representation captures one aspect of the difference between $\mathbf{q}$ and $\mathbf{q}^h$; the adaptivity will *target* a reduction in this aspect of the error.

Consider now a perturbation of the optimality system (2.6) in the form of a small added residual $\boldsymbol{\rho}$. The perturbed optimality conditions have the solution $\boldsymbol{\xi}_*^{\boldsymbol{\rho}}$,

$$\left(\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^{\boldsymbol{\rho}}\right]+\boldsymbol{\rho}\right)(\boldsymbol{\psi})=0\,,\quad \forall\,\boldsymbol{\psi}\in\mathcal{X}\,. \tag{7.2}$$

For the unperturbed case $\boldsymbol{\rho}=0$ we obtain the solution $\boldsymbol{\xi}_*^0=\boldsymbol{\xi}_*$ of (2.6).

Let $\mathcal{M}^{\boldsymbol{\rho}}$ denote the Lagrangian associated with the functional $\mathbf{E}$ and with the perturbed optimality conditions [7]:

$$\mathcal{M}^{\boldsymbol{\rho}}[\boldsymbol{\xi},\boldsymbol{\sigma}]=\mathcal{E}[\boldsymbol{\xi}]-\left(\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}\right]+\boldsymbol{\rho}\right)(\boldsymbol{\sigma})\,. \tag{7.3}$$

The Lagrange multipliers associated with the constraints posed by the perturbed optimality conditions (7.2) are $\boldsymbol{\sigma}$. Let $(\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*)$ be a stationary point of $\mathcal{M}^0[\boldsymbol{\xi},\boldsymbol{\sigma}]=\mathcal{E}[\boldsymbol{\xi}]-\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}\right](\boldsymbol{\sigma})$ for the unperturbed case $\boldsymbol{\rho}=0$. We have that

$$\mathcal{M}^0_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*](\boldsymbol{\psi})=\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\boldsymbol{\psi})-\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\psi},\boldsymbol{\sigma}_*)=0\,,\quad\forall\boldsymbol{\psi}\in\mathcal{X}\,, \tag{7.4a}$$

$$\mathcal{M}^0_{\boldsymbol{\sigma}}[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*](\boldsymbol{\phi})=-\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\phi})=0\,,\quad\forall\boldsymbol{\phi}\in\mathcal{X}\,. \tag{7.4b}$$

Equation (7.4b) states that if $(\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*)$ is a stationary point of $\mathcal{M}^0$, then $\boldsymbol{\xi}_*$ is a stationary point of $\mathcal{L}$, since the first order optimality conditions for $\mathcal{L}$ are imposed as constraints in $\mathcal{M}^0$.

Consider now the residual $\boldsymbol{\rho}^h$ for which the optimal solution of the continuous problem equals the optimal solution of the discrete problem, $\boldsymbol{\xi}_*^{\boldsymbol{\rho}}=\boldsymbol{\xi}_*^h$:

$$\begin{aligned}\boldsymbol{\rho}^h&=\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*\right]-\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right]=-\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right]\\ &\Rightarrow\quad\left(\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right]+\boldsymbol{\rho}^h\right)(\boldsymbol{\psi})=0\,,\quad\forall\boldsymbol{\psi}\in\mathcal{X}\,.\end{aligned} \tag{7.5}$$

The error in the cost functional (7.1) reads:

$$\begin{aligned}\mathcal{E}[\boldsymbol{\xi}_*]-\mathcal{E}[\boldsymbol{\xi}_*^h]&=\mathcal{M}^0[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*]-\mathcal{M}^{\boldsymbol{\rho}}[\boldsymbol{\xi}_*^h,\boldsymbol{\sigma}_*]\\ &=\mathcal{E}[\boldsymbol{\xi}_*]-\mathcal{E}[\boldsymbol{\xi}_*^h]-\left(\mathcal{L}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*]-\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right]-\boldsymbol{\rho}^h\right)(\boldsymbol{\sigma}_*)\\ &=\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\boldsymbol{\xi}_*-\boldsymbol{\xi}_*^h)-\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\boldsymbol{\xi}_*-\boldsymbol{\xi}_*^h,\boldsymbol{\sigma}_*)+\langle\boldsymbol{\rho},\boldsymbol{\sigma}_*\rangle+h.o.t.\\ &=\left\langle\boldsymbol{\rho}^h,\boldsymbol{\sigma}_*\right\rangle+h.o.t.\,,\end{aligned}$$

where the first two terms disappear due to the stationarity of $\mathcal{M}^0$ condition (7.4a).

Combining the equations we are lead to the following error estimate:

$$\begin{aligned}\mathcal{E}[\boldsymbol{\xi}_*^h]-\mathcal{E}[\boldsymbol{\xi}_*]&=\mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right](\boldsymbol{\sigma}_*)+h.o.t.\\ &=\mathcal{A}[\mathbf{u}^h,\mathbf{q}^h](\Delta\boldsymbol{\sigma}_*^{\mathbf{u}})\\ &\quad+\mathcal{J}_{\mathbf{u}}[\mathbf{u}^h,\mathbf{q}^h](\Delta\boldsymbol{\sigma}_*^{\boldsymbol{\lambda}})-\mathcal{A}_{\mathbf{u}}[\mathbf{u}^h,\mathbf{q}^h](\Delta\boldsymbol{\sigma}_*^{\boldsymbol{\lambda}},\boldsymbol{\lambda}^h)\\ &\quad+\mathcal{J}_{\mathbf{q}}[\mathbf{u}^h,\mathbf{q}^h](\Delta\boldsymbol{\sigma}_*^{\mathbf{q}})-\mathcal{A}_{\mathbf{q}}[\mathbf{u}^h,\mathbf{q}^h](\Delta\boldsymbol{\sigma}_*^{\mathbf{q}},\boldsymbol{\lambda}^h)+h.o.t.\end{aligned}$$

A more accurate estimator is obtained as follows. Define

$$\begin{aligned}\mathbf{e}&=\boldsymbol{\xi}_*^h-\boldsymbol{\xi}_*\\ N(s)&=\mathcal{E}[\boldsymbol{\xi}_*+s\,\mathbf{e}]-\left(\mathcal{L}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*+s\,\mathbf{e}]+s\boldsymbol{\rho}^h\right)(\boldsymbol{\sigma}_*)\\ \mathcal{M}^{\boldsymbol{\rho}}[\boldsymbol{\xi}_*^h,\boldsymbol{\sigma}_*]-\mathcal{M}^0[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*]&=N(1)-N(0)=\int_0^1 N'(s)\,ds\end{aligned}$$

Therefore we have that

$$
\begin{aligned}
\mathcal{E}[\boldsymbol{\xi}_*^h] - \mathcal{E}[\boldsymbol{\xi}_*] &= \mathcal{M}^{\rho}[\boldsymbol{\xi}_*^h, \boldsymbol{\sigma}_*] - \mathcal{M}^0[\boldsymbol{\xi}_*, \boldsymbol{\sigma}_*] \\
&= \int_0^1 \left(\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_* + s\mathbf{e}](\mathbf{e}) - \mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_* + s\mathbf{e}](\mathbf{e}, \boldsymbol{\sigma}_*) - \left\langle \boldsymbol{\rho}^h, \boldsymbol{\sigma}_* \right\rangle \right) ds \\
&= -\left\langle \boldsymbol{\rho}^h, \boldsymbol{\sigma}_* \right\rangle + \frac{1}{2}\left(\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\mathbf{e}) - \mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\mathbf{e}, \boldsymbol{\sigma}_*) \right) \\
&\quad + \frac{1}{2}\left(E_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h](\mathbf{e}) - \mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h](\mathbf{e}, \boldsymbol{\sigma}_*) \right) + R(\mathbf{e}, \mathbf{e}, \mathbf{e}) \\
&= -\left\langle \boldsymbol{\rho}^h, \boldsymbol{\sigma}_* \right\rangle + \frac{1}{2}\left(\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h](\mathbf{e}) - \mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h](\mathbf{e}, \boldsymbol{\sigma}_*) \right) + R(\mathbf{e}, \mathbf{e}, \mathbf{e})
\end{aligned}
$$

where the residual of the trapezoidal integration method is of order three with respect to $\mathbf{e}$. Using the fact that $\boldsymbol{\xi}_*^h = \boldsymbol{\xi}_* + \mathbf{e}$, expanding $\mathcal{E}_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h]$ and $\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*^h]$ in Taylor series about $\boldsymbol{\xi}_*$, and using the stationarity condition (7.4a), we obtain:

$$
\mathcal{E}[\boldsymbol{\xi}_*^h] - \mathcal{E}[\boldsymbol{\xi}_*] = -\left\langle \boldsymbol{\rho}^h, \boldsymbol{\sigma}_* \right\rangle + \frac{1}{2}\mathcal{E}_{\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\mathbf{e}, \mathbf{e}) - \frac{1}{2}\mathcal{L}_{\boldsymbol{\xi},\boldsymbol{\xi},\boldsymbol{\xi}}[\boldsymbol{\xi}_*](\mathbf{e}, \mathbf{e}, \boldsymbol{\sigma}_*) + h.o.t.
$$

where the higher order terms are at least of order three. This analysis reveals the structure of the second order error terms. It also shows that the additional terms considered by Becker are the second order terms. These second order terms are the residual of the optimality equation (7.4a) for the optimal discrete solution. This residual is weighed against the error $\mathbf{e}$.

Remark. We note the following:

1. All the residuals in the above equation are for the continuous operators.
2. By performing the integration by parts in reverse order, we arrive at the strong form (high derivatives of the polynomials inside each element) dot product with the $\boldsymbol{\sigma}$.
3. Each residual can be written as a sum of residuals within each element. Similarly, each element-wise residual can be computed in strong form, by differentiating the polynomials, to arrive at polynomial residuals.

7.2. The first order posterior estimator. Using the general residual formula (7.5) the first order posterior error estimator reads

$$
\begin{aligned}
E^h &= -\left\langle \boldsymbol{\rho}^h, \boldsymbol{\sigma}_* \right\rangle \\
&= \left\langle \mathcal{L}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi}_*^h\right], \boldsymbol{\sigma}_* \right\rangle \\
&= \left\langle \mathcal{L}_{\boldsymbol{\lambda}}\left[\boldsymbol{\xi}_*^h\right], \boldsymbol{\sigma}_*^{\boldsymbol{\lambda}} \right\rangle + \left\langle \mathcal{L}_{\mathbf{u}}\left[\boldsymbol{\xi}_*^h\right], \boldsymbol{\sigma}_*^{\mathbf{u}} \right\rangle + \left\langle \mathcal{L}_{\mathbf{q}}\left[\boldsymbol{\xi}_*^h\right], \boldsymbol{\sigma}_*^{\mathbf{q}} \right\rangle \\
&= \boldsymbol{\rho}_{\mathbf{u}}[\boldsymbol{\xi}_*^h]\left(\boldsymbol{\sigma}_*^{\boldsymbol{\lambda}}\right) + \boldsymbol{\rho}_{\boldsymbol{\lambda}}[\boldsymbol{\xi}_*^h]\left(\boldsymbol{\sigma}_*^{\mathbf{u}}\right) + \boldsymbol{\rho}_{\mathbf{q}}[\boldsymbol{\xi}_*^h]\left(\boldsymbol{\sigma}_*^{\mathbf{q}}\right).
\end{aligned}
$$

The residuals $\boldsymbol{\rho}$ are given by the continuous Lagrangian derivatives evaluated at the discrete solutions. Specifically, $\boldsymbol{\rho}_{\mathbf{u}}$, $\boldsymbol{\rho}_{\boldsymbol{\lambda}}$, $\boldsymbol{\rho}_{\mathbf{q}}$, represent the residuals in the primal equation, dual equation, and optimality equation, respectively:

$$
\begin{aligned}
\boldsymbol{\rho}_{\mathbf{u}}[\boldsymbol{\xi}_*^h]\,(\cdot) &:= -\mathcal{A}[\mathbf{u}^h, \mathbf{q}^h](\cdot)\,, \\
\boldsymbol{\rho}_{\boldsymbol{\lambda}}[\boldsymbol{\xi}_*^h]\,(\cdot) &:= \mathcal{J}_{\mathbf{u}}[\mathbf{u}^h, \mathbf{q}^h](\cdot) - \mathcal{A}_{\mathbf{u}}[\mathbf{u}^h, \mathbf{q}^h](\cdot, \boldsymbol{\lambda}^h)\,, \\
\boldsymbol{\rho}_{\mathbf{q}}[\boldsymbol{\xi}_*^h]\,(\cdot) &:= \mathcal{J}_{\mathbf{q}}[\mathbf{u}^h, \mathbf{q}^h](\cdot) - \mathcal{A}_{\mathbf{q}}[\mathbf{u}^h, \mathbf{q}^h](\cdot, \boldsymbol{\lambda}^h)\,.
\end{aligned}
$$

For example, consider the weak form of our elliptic model problem (3.1)

$$\mathcal{A}[\mathbf{q},\mathbf{u}](\phi) = a[\mathbf{q}](\mathbf{u},\phi) - \ell(\phi)\,, \quad \forall \phi \in \mathcal{U}\,,$$

where a is bilinear, and ℓ is linear. The corresponding residuals are

$$\begin{aligned}
\boldsymbol{\rho}_{\mathbf{u}}[\boldsymbol{\xi}^h_*](\cdot) &= -a\left[\mathbf{q}^h_*\right](\mathbf{u}^h_*,\cdot) + \ell(\cdot)\,,\\
\boldsymbol{\rho}_{\boldsymbol{\lambda}}[\boldsymbol{\xi}^h_*](\cdot) &= \mathcal{J}_{\mathbf{u}}\left[\mathbf{u}^h_*,\mathbf{q}^h_*\right](\cdot) - a_{\mathbf{u}}\left[\mathbf{q}^h_*\right](\cdot,\boldsymbol{\lambda}^h_*)\,,\\
\boldsymbol{\rho}_{\mathbf{q}}[\boldsymbol{\xi}^h_*](\cdot) &= \mathcal{J}_{\mathbf{q}}\left[\mathbf{u}^h_*,\mathbf{q}^h_*\right](\cdot) - a_{\mathbf{q}}\left[\mathbf{q}^h_*\right](\cdot,\mathbf{u}^h_*,\boldsymbol{\lambda}^h_*)\,.
\end{aligned}$$

Explain here why $\Delta\boldsymbol{\sigma}$ and not just $\boldsymbol{\sigma}$ The multipliers $\boldsymbol{\sigma}$ that weigh the residuals are estimated through local higher–order interpolation [23]. Consider, for example, the error weight

$$\Delta\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_* := \boldsymbol{\sigma}^{\boldsymbol{\lambda}}_* - \mathcal{I}_h\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*\,.$$

Here $\mathcal{I}_h : \mathcal{U} \to \mathcal{U}^h$ denotes the projection operator onto the discrete solution space. We can approximate the weight $\Delta\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*$ through a patch-wise higher–order interpolation of $\boldsymbol{\sigma}^h_{\boldsymbol{\lambda}}$ onto a coarser mesh. Suppose $\boldsymbol{\sigma}^h_{\boldsymbol{\lambda}}$ is defined on a regular mesh $\mathcal{T}^q_h$ (see section 4 for details on the notation). The higher-order interpolant $\boldsymbol{\sigma}^H_{\boldsymbol{\lambda}}$ is defined on a coarser mesh $\mathcal{T}^q_H$ that is obtained directly from $\mathcal{T}^q_h$ through pure hierarchical coarsening (e.g., $H := 2h$). Then, for all mesh elements $\kappa \in \mathcal{T}^h_{\mathrm{q}}$, we set:

$$\Delta\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*\big|_\kappa \approx \left((\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*)^H - (\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*)^h\right)\Big|_\kappa\,.$$

Similar computations are done for the other weights $\Delta\boldsymbol{\sigma}^{\mathbf{u}}_*$, $\Delta\boldsymbol{\sigma}^{\mathbf{q}}_*$, etc.

7.3. Calculation of the multipliers $\boldsymbol{\sigma}_*$. We are now left with the task of calculating $\boldsymbol{\sigma}_*$ from the stationarity condition (7.4a). These additional multipliers have components for the forward, dual, and optimality problems, $\boldsymbol{\sigma}_* = \{\boldsymbol{\sigma}^{\mathbf{u}}_*, \boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*, \boldsymbol{\sigma}^{\mathbf{q}}_*\}$. The Lagrangian $\mathcal{M}^0$ (7.3) has the form

$$\begin{aligned}
\mathcal{M}^0[\boldsymbol{\xi},\boldsymbol{\sigma}] &= \mathcal{E}[\mathbf{q}] - \mathcal{L}_{\mathbf{q}}[\boldsymbol{\xi}](\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{u}}[\boldsymbol{\xi}](\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{L}_{\boldsymbol{\lambda}}[\boldsymbol{\xi}](\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*)\\
&= \mathcal{E}[\mathbf{q}] - \mathcal{L}_{\mathbf{q}}[\boldsymbol{\xi}](\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{u}}[\boldsymbol{\xi}](\boldsymbol{\sigma}^{\mathbf{u}}_*)
\end{aligned}$$

where the last equality holds true if $\mathbf{u} = U[\mathbf{q}]$. The first variation of this Lagrangian $\mathcal{M}^0$ (7.4a) about the optimal point $\boldsymbol{\xi}_* = \{\mathbf{u}_*,\ \boldsymbol{\lambda}_*,\ \mathbf{q}_*\}$ can be expanded as follows

$$\begin{aligned}
\mathcal{M}^0_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*](\delta\boldsymbol{\xi}) &= \mathcal{E}_{\mathbf{q}}[\mathbf{q}_*](\delta\mathbf{q})\\
&\quad -\mathcal{L}_{\mathbf{q},\mathbf{q}}[\boldsymbol{\xi}_*](\delta\mathbf{q}\,,\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{q},\mathbf{u}}[\boldsymbol{\xi}_*](\delta\mathbf{u}\,,\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{q},\boldsymbol{\lambda}}[\boldsymbol{\xi}_*](\delta\boldsymbol{\lambda}\,,\boldsymbol{\sigma}^{\mathbf{q}}_*)\\
&\quad -\mathcal{L}_{\mathbf{u},\mathbf{q}}[\boldsymbol{\xi}_*](\delta\mathbf{q}\,,\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{L}_{\mathbf{u},\mathbf{u}}[\boldsymbol{\xi}_*](\delta\mathbf{u}\,,\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{L}_{\mathbf{u},\boldsymbol{\lambda}}[\boldsymbol{\xi}_*](\delta\boldsymbol{\lambda}\,,\boldsymbol{\sigma}^{\mathbf{u}}_*)\\
&= \left\{\mathcal{E}_{\mathbf{q}}[\mathbf{q}_*] - \mathcal{L}_{\mathbf{q},\mathbf{q}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{u},\mathbf{q}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{u}}_*)\right\}(\delta\mathbf{q})\\
&\quad -\left\{\mathcal{L}_{\mathbf{q},\mathbf{u}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{q}}_*) + \mathcal{L}_{\mathbf{u},\mathbf{u}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{u}}_*)\right\}(\delta\mathbf{u})\\
&\quad -\left\{\mathcal{L}_{\mathbf{q},\boldsymbol{\lambda}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{q}}_*) + \mathcal{L}_{\mathbf{u},\boldsymbol{\lambda}}[\boldsymbol{\xi}_*](\boldsymbol{\sigma}^{\mathbf{u}}_*)\right\}(\delta\boldsymbol{\lambda})
\end{aligned} \tag{7.6}$$

Using (2.7), and a compact notation for the nonlinear arguments, we have for the elliptic equation

$$\begin{aligned}\mathcal{M}^0_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*](\psi) = &\left\{\mathcal{E}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right] - \mathcal{J}_{\mathbf{u},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{J}_{\mathbf{q},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\right.\\ &\left.+\mathcal{A}_{\mathbf{q},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*) + \mathcal{A}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*) + \mathcal{A}_{\mathbf{u},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*)\right\}(\psi_{\mathbf{q}})\\ &-\left\{\mathcal{J}_{\mathbf{u},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) + \mathcal{J}_{\mathbf{q},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\right.\\ &\left.+\mathcal{A}_{\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*) + \mathcal{A}_{\mathbf{u},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) + \mathcal{A}_{\mathbf{q},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\right\}(\psi_{\mathbf{u}})\\ &+\left\{\mathcal{A}_{\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) + \mathcal{A}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\right\}(\psi_{\boldsymbol{\lambda}})\end{aligned}$$

where the components of the test function $\psi = (\psi_{\mathbf{u}}, \psi_{\boldsymbol{\lambda}}, \psi_{\mathbf{q}})$ appear explicitly within different terms. The variation (7.6) of the Lagrangian $\mathcal{M}^0$ is zero at its stationary points. The equation for $\psi_{\mathbf{q}}$ reads

$$\begin{aligned}0 = &\,\mathcal{E}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right] - \mathcal{J}_{\mathbf{u},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{J}_{\mathbf{q},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*) + \mathcal{A}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*)\\ &+\mathcal{A}_{\mathbf{q},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*) + \mathcal{A}_{\mathbf{u},\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*)\,.\end{aligned} \tag{7.7}$$

The equation for $\psi_{\boldsymbol{\lambda}}$ is the tangent linear model evaluated at the optimal solution

$$0 = \mathcal{A}_{\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) + \mathcal{A}_{\mathbf{q}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*) \quad \Leftrightarrow \quad \boldsymbol{\sigma}^{\mathbf{u}}_* = U'[\mathbf{q}_*]\,\boldsymbol{\sigma}^{\mathbf{q}}_*\,. \tag{7.8}$$

The equation for $\psi_{\mathbf{u}}$ is the second order adjoint model evaluated at the optimal solution:

$$\begin{aligned}0 = &\,\mathcal{J}_{\mathbf{u},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) + \mathcal{J}_{\mathbf{q},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\\ &-\mathcal{A}_{\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*) - \mathcal{A}_{\mathbf{u},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{u}}_*) - \mathcal{A}_{\mathbf{q},\mathbf{u}}\left[\boldsymbol{\xi}_*\right](\boldsymbol{\sigma}^{\mathbf{q}}_*)\,.\end{aligned} \tag{7.9}$$

To derive the computational procedure we revisit the reduced cost function, which is equal to the reduced form of the Lagrangian (2.5) for any $\mathbf{q}$ and for any $\boldsymbol{\lambda}$:

$$\begin{aligned}j[\mathbf{q};\boldsymbol{\lambda}] &= \mathcal{L}\left[\xi_2\right] = \mathcal{J}\left[U[\mathbf{q}],\mathbf{q}\right] - \mathcal{A}\left[U[\mathbf{q}],\mathbf{q}\right](\boldsymbol{\lambda})\,,\\ \xi_2[\mathbf{q},\boldsymbol{\lambda}] &= (\mathbf{q},U[\mathbf{q}],\boldsymbol{\lambda})\,, \quad \forall \mathbf{q}\in\mathcal{Q},\ \boldsymbol{\lambda}\in\mathcal{U}\,.\end{aligned}$$

The notation ξ_2 reminds that this $\boldsymbol{\xi}$ depends on only two parameters, $\mathbf{q}$ and $\boldsymbol{\lambda}$. The reduced gradient is

$$\begin{aligned}j_{\mathbf{q}}[\mathbf{q};\boldsymbol{\lambda}](\phi) &= \mathcal{L}_{\mathbf{q}}[\xi_2](\phi) + \mathcal{L}_{\mathbf{u}}[\xi_2]\left(U'[\mathbf{q}]\phi\right) + \mathcal{L}_{\boldsymbol{\lambda}}[\xi_2]\left(\Lambda'[\mathbf{q}]\phi\right)\\ &= \mathcal{L}_{\mathbf{q}}[\xi_2](\phi) + \mathcal{L}_{\mathbf{u}}[\xi_2]\left(U'[\mathbf{q}]\phi\right)\,, \quad \phi\in\mathcal{Q}\,.\end{aligned} \tag{7.10}$$

The term $\mathcal{L}_{\boldsymbol{\lambda}}$ is the forward equation and vanishes identically since $\mathbf{u} = U[\mathbf{q}]$ in ξ_2. Note that

$$j_{\mathbf{q},\boldsymbol{\lambda}}[\mathbf{q};\boldsymbol{\lambda}](\psi_{\boldsymbol{\lambda}},\phi) = \mathcal{L}_{\mathbf{q},\boldsymbol{\lambda}}[\xi_2](\psi_{\boldsymbol{\lambda}},\phi) + \mathcal{L}_{\mathbf{u},\boldsymbol{\lambda}}[\xi_2]\left(\psi_{\boldsymbol{\lambda}},U'[\mathbf{q}]\phi\right)\,. \tag{7.11}$$

The reduced Hessian is

$$\begin{aligned}j_{\mathbf{q},\mathbf{q}}[\mathbf{q};\boldsymbol{\lambda}](\psi,\phi) = &\,\mathcal{L}_{\mathbf{q},\mathbf{q}}[\xi_2](\psi,\phi) + \mathcal{L}_{\mathbf{q},\mathbf{u}}[\xi_2](U'[\mathbf{q}]\psi,\phi)\\ &+\mathcal{L}_{\mathbf{u},\mathbf{q}}[\xi_2]\left(\psi,U'[\mathbf{q}]\phi\right) + \mathcal{L}_{\mathbf{u},\mathbf{u}}[\xi_2]\left(U'[\mathbf{q}]\psi,U'[\mathbf{q}]\phi\right)\\ &+\mathcal{L}_{\mathbf{u}}[\xi_2]\left(U''[\mathbf{q}](\psi,\phi)\right)\,, \quad \psi,\phi\in\mathcal{Q}\,.\end{aligned}$$

With $\boldsymbol{\lambda} = \Lambda(\mathbf{q})$ the adjoint equation $\mathcal{L}_{\mathbf{u}} = 0$ makes the last term vanish identically. Let $\boldsymbol{\xi}_1[\mathbf{q}] = (\mathbf{q}, U[\mathbf{q}], \Lambda(\mathbf{q}))$, which depends on just one parameter, $\mathbf{q}$. The reduced Hessian reads

$$j_{\mathbf{q},\mathbf{q}}[\mathbf{q}](\psi,\phi) = \mathcal{L}_{\mathbf{q},\mathbf{q}}[\boldsymbol{\xi}_1](\psi,\phi) + \mathcal{L}_{\mathbf{q},\mathbf{u}}[\boldsymbol{\xi}_1](U'[\mathbf{q}]\psi,\phi) \\ +\mathcal{L}_{\mathbf{u},\mathbf{q}}[\boldsymbol{\xi}_1]\left(\psi,U'[\mathbf{q}]\phi\right) + \mathcal{L}_{\mathbf{u},\mathbf{u}}[\boldsymbol{\xi}_1]\left(U'[\mathbf{q}]\psi,U'[\mathbf{q}]\phi\right)\,, \quad \psi,\phi\in\mathcal{Q}\,.$$

Consider the variation of the Lagrangian $\mathcal{M}^0_{\boldsymbol{\xi}}$ (7.6) along the direction $\psi \in \mathcal{X}$ with the following components: $\psi_{\mathbf{q}} \in \mathcal{Q}$ arbitrary, $\psi_{\mathbf{u}} = U'[\mathbf{u}_*]\psi_{\mathbf{q}}$, and $\psi_{\boldsymbol{\lambda}} = 0$. Using (7.8) we replace $\boldsymbol{\sigma}^{\mathbf{u}}_* = U'[\mathbf{q}_*]\,\boldsymbol{\sigma}^{\mathbf{q}}_*$. Note that $\boldsymbol{\xi}_1[\mathbf{q}_*] = \boldsymbol{\xi}_*$. With these substitutions, the variation of the Lagrangian (7.6) reads

$$\begin{aligned}\mathcal{M}^0_{\boldsymbol{\xi}}[\boldsymbol{\xi}_*,\boldsymbol{\sigma}_*]\,(\psi) &= \mathcal{E}_{\mathbf{q}}[\mathbf{q}_*]\,(\psi_{\mathbf{q}}) - \mathcal{L}_{\mathbf{q},\mathbf{q}}\,[\boldsymbol{\xi}_*]\,(\psi_{\mathbf{q}},\boldsymbol{\sigma}^{\mathbf{q}}_*) - \mathcal{L}_{\mathbf{u},\mathbf{q}}\,[\boldsymbol{\xi}_*]\,(\psi_{\mathbf{q}},U'[\mathbf{q}_*]\,\boldsymbol{\sigma}^{\mathbf{q}}_*) \\ &\quad -\mathcal{L}_{\mathbf{q},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,\left(U'[\mathbf{u}_*]\psi_{\mathbf{q}},\boldsymbol{\sigma}^{\mathbf{q}}_*\right) - \mathcal{L}_{\mathbf{u},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,\left(U'[\mathbf{u}_*]\psi_{\mathbf{q}},U'[\mathbf{q}_*]\,\boldsymbol{\sigma}^{\mathbf{q}}_*\right) \\ &= \mathcal{E}_{\mathbf{q}}[\mathbf{q}_*]\,(\psi_{\mathbf{q}}) - j_{\mathbf{q},\mathbf{q}}[\mathbf{q}_*]\,(\psi_{\mathbf{q}},\boldsymbol{\sigma}^{\mathbf{q}}_*)\;.\end{aligned}$$

The variation of the Lagrangian $\mathcal{M}^0$ (7.6) along the direction ψ must be zero for any $\psi_{\mathbf{q}}$, which leads to the following equation for the multiplier $\boldsymbol{\sigma}^{\mathbf{q}}_*$

$$j_{\mathbf{q},\mathbf{q}}[\mathbf{q}_*]\,(\phi,\boldsymbol{\sigma}^{\mathbf{q}}_*) = \mathcal{E}_{\mathbf{q}}[\mathbf{q}_*]\,(\phi)\,, \quad \forall\phi\in\mathcal{Q}\,. \tag{7.12}$$

Equation (7.12) can be solved approximately by using a quasi-Newton approximation of the reduced Hessian. This approximation is based on the sequence of reduced gradients obtained during the optimization. The computational procedure is given in Algorithm 7.1.

Algorithm 7.1 Error estimation using the second order adjoint solution

1: Solve equation (7.12) for $\boldsymbol{\sigma}^{\mathbf{q}}_*$ using a quasi-Newton approximation of $j_{\mathbf{q},\mathbf{q}}$.
2: Given $\boldsymbol{\sigma}^{\mathbf{q}}_*$, solve the tangent linear model (7.8) to obtain $\boldsymbol{\sigma}^{\mathbf{u}}_*$.
3: Given $\boldsymbol{\sigma}^{\mathbf{q}}_*$ and $\boldsymbol{\sigma}^{\mathbf{u}}_*$, solve the second order adjoint model (7.9) to obtain $\boldsymbol{\sigma}^{\boldsymbol{\lambda}}_*$.

8. Numerical results for a distributed control test problem. Consider the case where the inversion variable is the volume forcing $\mathbf{f}$ of the primal problem. Assume the discrete variables $\mathbf{f}^h \in \mathcal{Q}^h_r \equiv \mathcal{U}^h_p$. We define the cost functional $\mathcal{J}$ as follows:

$$\mathcal{J} := \frac{1}{2}\,\|\mathcal{H}\mathbf{u} - \mathbf{o}\|^2_{\mathcal{L}^2(\Omega)}\,\mathrm{d}\mathbf{x} + \frac{\beta}{2}\,\|\mathbf{f} - \mathbf{f}_B\|^2_{\mathcal{L}^2(\Omega)}\,\mathrm{d}\mathbf{x}\,. \tag{8.1}$$

The primal and adjoint equations remain unchanged. However, the optimality condition (3.6) becomes

$$\beta\,\mathbf{f} = -\boldsymbol{\lambda}\,,\ \mathbf{x}\in\Omega\,, \tag{8.2}$$

whereas (4.23) now reads:

$$\beta\,\left\langle \mathbf{f}^h,\,\delta\mathbf{f}^h \right\rangle_\Omega + \left\langle \boldsymbol{\lambda}^h,\,\delta\mathbf{f}^h \right\rangle_\Omega = 0\,,\ \forall\,\delta\mathbf{f}^h \in \mathcal{U}^h_p\,. \tag{8.3}$$

For this problem the discrete optimality condition is obviously a consistent discretization of the continuous formulation. We analyze the convergence of the problem for an adaptive discretization. Let $\mathcal{H}$ be the identity operator, $\mathbf{o} := \mathcal{H}\mathbf{u}_*$,

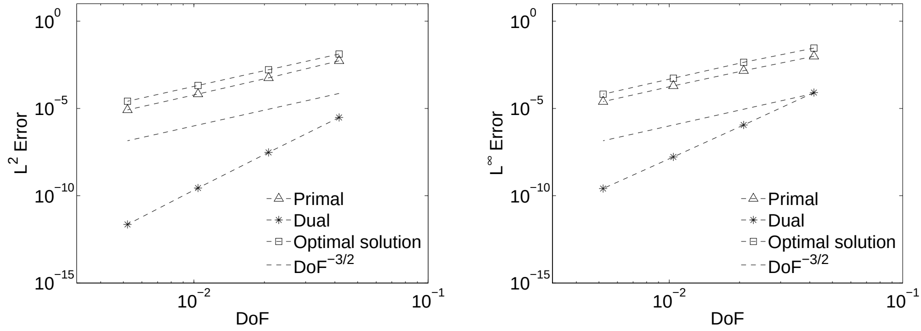

FIG. 8.1. $\mathcal{L}^2$ *(left), and* $\mathcal{L}^\infty$ *(right) convergence for the distributed control problem (3.1)–(8.1), with* $r = p = 2$.

$\boldsymbol{\beta} = 100$, and

$$\mathbf{q}(\mathbf{x}) := (1 + x^2 + y^2)/10$$
$$\mathbf{u}_*(\mathbf{x}) := 10 \exp\left(-\frac{(10x-5)^2}{4}\right) \exp\left(-\frac{(10y-5)^2}{4}\right) ,$$
$$\mathbf{f}_*(\mathbf{x}) = \mathbf{f}_B := -\nabla \cdot (\mathbf{q}\, \nabla \mathbf{u}_*)(\mathbf{x}) \ , \ \forall\ \mathbf{x} = (x, y) \in \Omega = [0,\, 1]^2 \ .$$

The analytical cost functional is then zero at the optimal solution: $\mathcal{J}(\mathbf{u}_*,\, \mathbf{q}_*) = 0$.

Figure 8 shows the convergence of the discrete primal, dual, and optimal solutions to their exact values, upon mesh refinement. All variables are discretized on the same mesh, using a quadratic Lagrange basis ($p = 2$), with square elements, and hierarchical mesh refinement. The theoretical estimate $\mathcal{O}(h^3)$ is verified numerically. The dual solution is super-convergent: $\|\boldsymbol{\lambda}^h - \boldsymbol{\lambda}\|_{\mathcal{L}^\infty(\mathcal{T}^h)} \le C\, h^6$.

9. Numerical results for the coefficient identification problem. Let $\mathcal{H}$ be the identity operator, $\mathbf{o} := \mathcal{H}\mathbf{u}_*$, $\boldsymbol{\beta} = 100$, and $\Omega := [0,\, 1] \times [0,\, 1]$. The first numerical test makes use of

$$\text{Test A:} \quad \mathbf{q}_B(\mathbf{x}) := 1 + x^2 + y^2 \ , \tag{9.1}$$
$$\mathbf{u}_*(\mathbf{x}) := 10 \exp\left(-\frac{(10x-5)^2}{4}\right) \exp\left(-\frac{(10y-5)^2}{4}\right) ,$$

and the initial guess

$$\mathbf{q}_0(\mathbf{x}) := 1 + 7x + 7y \ , \ \mathbf{x} \in \Omega \ . \tag{9.2}$$

The second set of experiments uses the following functions:

$$\text{Test B:} \quad \mathbf{q}_B(\mathbf{x}) := 5 + \sin\left(\frac{\pi x}{2}\right) \cos(2\pi y) \ , \tag{9.3}$$
$$\mathbf{u}_*(\mathbf{x}) := 10 \exp\left(-\frac{(10x-5)^2}{4}\right) \exp\left(-\frac{(10y-5)^2}{4}\right) ,$$

and a constant initial guess

$$\mathbf{q}_0 := 1 \ . \tag{9.4}$$

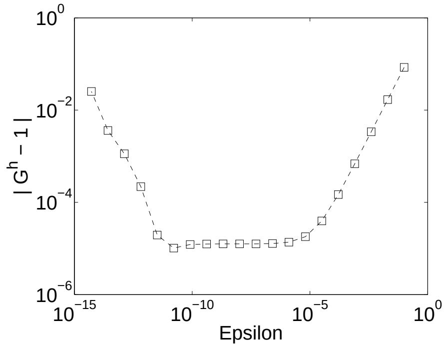

FIG. 9.1. *Asymptotic consistency test for the discrete gradient $\nabla_{\mathbf{q}^h}\mathcal{J}^h$ using the first order Taylor approximation (9.5), and the functions (9.1)–(9.2).*

For both experiments we choose

$$\mathbf{f}(\mathbf{x}) := -\nabla \cdot (\mathbf{q}_B \, \nabla \mathbf{u}_*)(\mathbf{x}) \; , \; \forall \, \mathbf{x} \in \Omega \; .$$

Then, $\mathbf{q}_* = \mathbf{q}_B$, and $\mathcal{J}(\mathbf{u}_*, \, \mathbf{q}_*) = 0$. The discrete solver uses the `deal.II` library [3] for finite element computations. The spaces $\left(\mathcal{U}^h\right)^p$ and $\mathcal{Q}_r^h$ are spanned by the second order Lagrange basis functions ($p = r = 2$). Moreover, $\mathcal{T}^h \equiv \mathcal{T}_{\mathrm{q}}^h$. Note that for the first set (9.1)–(9.2), the exact optimal solution $\mathbf{q}_*$ can be represented exactly by the polynomial basis spanning $\mathcal{Q}_r^h$.

9.1. Consistency of the discrete gradient. An important step in adjoint code development is validation of the discrete gradient (6.4). The classic approach to adjoint code validation is through the numerical verification of a truncated Taylor expansion [28]. For smooth $\mathcal{J}_h$, and small perturbations $\varepsilon \, \boldsymbol{\delta}\mathbf{q}^h$ around a reference state $\mathbf{q}^h$, the Taylor theorem gives:

$$\mathcal{J}^h(\mathbf{q}^h + \varepsilon \, \boldsymbol{\delta}\mathbf{q}^h) = \mathcal{J}^h(\mathbf{q}^h) + \varepsilon \, \left\langle \nabla_{\mathbf{q}^h}\mathcal{J}^h, \, \boldsymbol{\delta}\mathbf{q}^h \right\rangle_\Omega + \mathcal{O}(\varepsilon^2) \; .$$

We verify numerically the following limit for small values of ε:

$$\lim_{\varepsilon \to 0} G^h(\varepsilon) := \lim_{\varepsilon \to 0} \frac{\mathcal{J}^h(\mathbf{q}^h + \varepsilon \, \boldsymbol{\delta}\mathbf{q}^h) - \mathcal{J}^h(\mathbf{q}^h)}{\varepsilon \, \left\langle \nabla_{\mathbf{q}^h}\mathcal{J}^h, \, \boldsymbol{\delta}\mathbf{q}^h \right\rangle_\Omega} = 1 \; . \tag{9.5}$$

As shown in figure 9.1, the discrete reduced gradient is found to be numerically consistent. For $\varepsilon < 10^{-11}$, truncation errors start to degrade the quality of the finite difference approximation (9.5).

9.2. The numerical behavior of the discrete optimality condition. We now consider the solution of equation (4.24). Figure 9.2(a) shows the decrease with $h \sim \mathrm{DoF}^{-1/2}$ of the integrals $\mathbb{I}_\Gamma^1$, $\mathbb{I}_\Gamma^2$, $\mathbb{I}_{\Gamma_\mathcal{I}}^1$, and $\mathbb{I}_{\Gamma_\mathcal{I}}^2$ in (4.29a)–(4.29b). For $p = r = 2$, and the $\mathcal{C}^\infty(\Omega)$ exact solutions (9.1)–(9.2), the integrals vanish asymptotically at least as fast as $h^{3/2}$, verifying the analytical bounds. Similar results are reported in figure 9.2(b) for the problem (9.3)–(9.4).

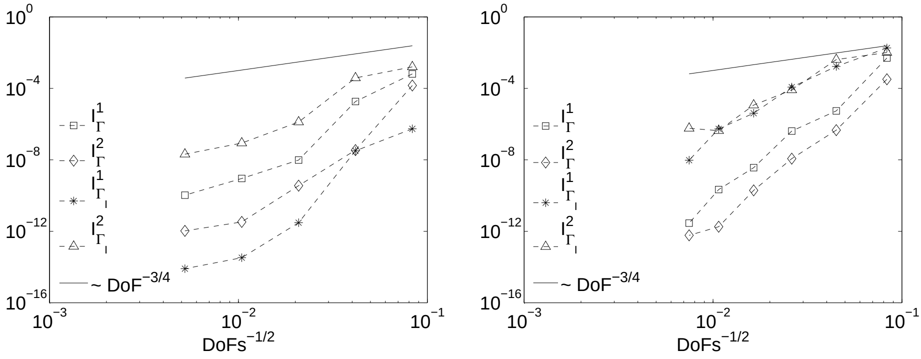

FIG. 9.2. *Asymptotic behavior of the integrals (4.29a)–(4.29b) for test A (left), and test B (right).*

9.3. Multigrid optimization and solution convergence. The multigrid optimization procedure is described in Algorithm 9.1. The subscript k indicates the current multigrid iteration. On each mesh level, we keep iterating until there is a sufficient reduction in the cost functional $\mathcal{J}_h$. A practical stopping criterion for the multigrid iterations can be based on the value of the target functional $\mathcal{J}^h$ at the current solution, or some reduced gradient norm.

The algorithm requires the constant relative and absolute tolerance vectors ATOL and RTOL, which decide if there has been sufficient decrease in the cost functional value on the current mesh level. In practice, one may also want to limit the number of optimization iterations on a given mesh, to prevent excessively long run times.

Algorithm 7.1 is used for the error estimation at step 9 of Algorithm 9.1. The mesh adaptation is performed at step 12, which get executed once per multigrid iteration. The exact choice of error functional is problem dependent. We will choose E to be a weighted average of the optimal solution over Ω:

$$E(\mathbf{x}) = \int_{\Omega} w\, \mathbf{q}\, \mathrm{d}\mathbf{x}\,. \tag{9.6}$$

The weight function $w \in \mathcal{H}^2(\Omega)$ is positive *a.e.* on Ω.

Figure 9.3 shows the convergence of multigrid optimization algorithm on several mesh levels. We plot both the decrease in the cost functional, and in the optimal solution error for the problem (9.1)–(9.2). For this experiment the mesh refinement process is guided by a simple element–wise error estimator based on the scaled gradient $\nabla \mathbf{q}^h$. The order of convergence is consistent with the theoretical expectations: the optimal solution shows cubic convergence with the mesh size h (for $p = r = 2$). We chose ATOL $= 10^{-10}$, and RTOL $= 10^{-2}$ on all iteration levels.

Remark. The meshes $\mathcal{T}_h^q$ and $\mathcal{T}^h$ may also be locally coarsened in step 12 of Algorithm 9.1. The coarsening operation is analogous to that of refinement. Care must be taken to maintain the mesh nesting property for $\mathcal{T}_h^q$ and $\mathcal{T}^h$.

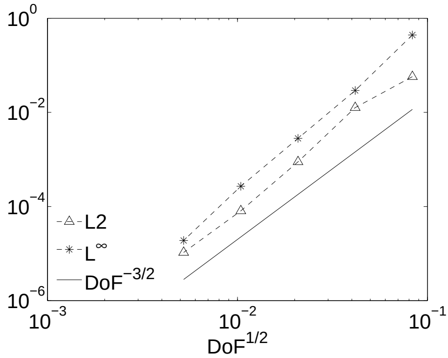

FIG. 9.3. *Convergence of the multigrid optimization algorithm: test A. The optimal solution error is plotted versus* $h \sim \mathrm{DoF}^{-1/2}$. *The errors correspond to the optimal solutions on each mesh level.*

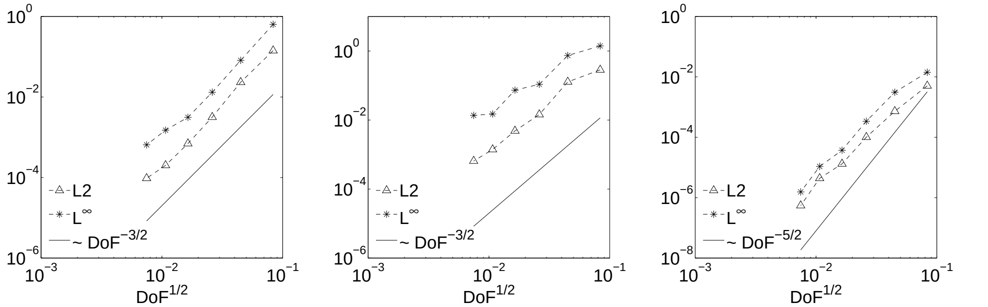

FIG. 9.4. *Convergence of the discrete optimal (left), primal (center), and dual (right) solutions for test B. The errors correspond to the converged solutions on each mesh level, and are plotted versus* $h \sim \mathrm{DoF}^{-1/2}$.

Algorithm 9.1 Multigrid optimization with error–driven AMR

Require: ATOL[maxit], RTOL[maxit].

Define the initial guess $\mathbf{q}^h$ on $(\mathcal{T}^h_{\mathrm{q}})_0$, and calculate $\mathbf{u}^h$, $\boldsymbol{\lambda}^h$ on $(\mathcal{T}^h)_0$.

Define the error functional $E[\mathbf{q}]$.

For example, $E[\mathbf{q}] := \int_\Omega w(\mathbf{x})\mathbf{q}(\mathbf{x})\,\mathrm{d}\mathbf{x}$, for some positive weight function w.

for $k = 0$ to maxit $- 1$ **do**

  Solve the primal model on $(\mathcal{T}^h)_k$ to get $\mathcal{J}^h_0 = \mathcal{J}^h(\mathbf{u}^h, \mathbf{q}^h)$.

  while $\mathcal{J}^h/\mathcal{J}^h_0 >$ RTOL[k] **and** $\mathcal{J}^h >$ ATOL[k] **do**

    Run the optimization iterations on $(\mathcal{T}^h)_k$ and $(\mathcal{T}^h_{\mathrm{q}})_k$, and update the current approximation $\mathbf{q}^h$, and cost functional $\mathcal{J}^h$.

    Construct approximation to the reduced Hessian $\nabla^2_{\mathbf{q}^h}\mathcal{J}^h$ from the optimization iterates $\mathbf{q}^h$ and gradients $\nabla_{\mathbf{q}^h}\mathcal{J}^h$.

  end while

  Calculate the element-wise error indicators for all $\kappa^q \in \mathcal{T}^h_{\mathrm{q}}$ using Algorithm 7.1.

  Flag the cells with highest estimated error on $(\mathcal{T}^q_h)_k$ for refinement.

  For any element $\kappa^q \in (\mathcal{T}^h_{\mathrm{q}})_k$ flagged for refinement, flag for refinement all elements $\kappa \in (\mathcal{T}^h)_k$ such that $\kappa \subseteq \kappa^q$. This is required to maintain the mesh nesting property.

  Refine the triangulations to obtain $(\mathcal{T}^h)_{k+1}$ and $(\mathcal{T}^q_h)_{k+1}$.

  Transfer the current optimal solution approximation $\mathbf{q}^h$ to the new parameter

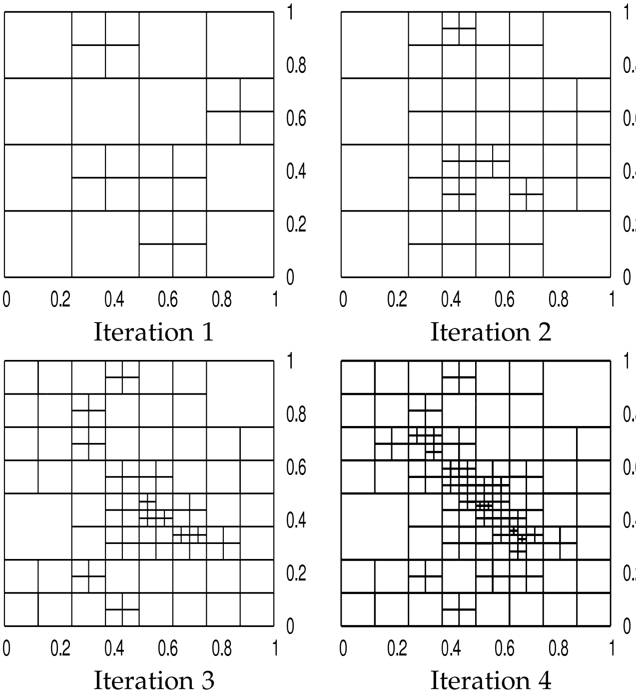

FIG. 9.5. *Optimization meshes generated by algorithm 9.1 on numerical test B.*

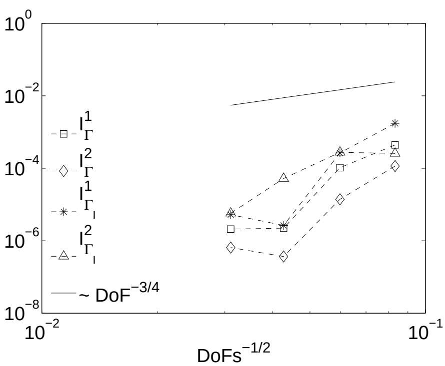

FIG. 9.6. *Asymptotic behavior of the integrals (4.29a)–(4.29b) for the a posteriori numerical experiments with test B.*

Figure 9.5 illustrates the sequence of meshes generated by algorithm 9.1 with the numerical test B. The observed numerical orders of convergence for $\mathbf{q}^h$, $\mathbf{u}^h$, and $\boldsymbol{\lambda}^h$, are consistent with the theoretical bounds. Note that the perturbation integrals (4.29a)–(4.29b) - shown in figure 9.3 - are sufficiently small to not affect the convergence of the optimization process (figure 9.7). However, if the optimal solution error is reduced further, the perturbations may impact the convergence process - see equation (5.6).

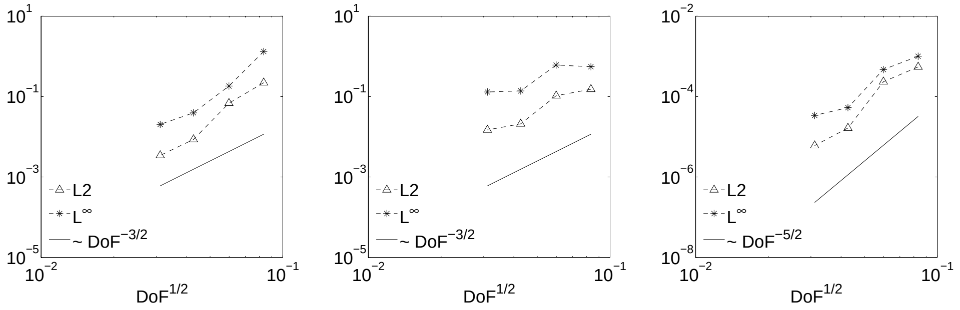

FIG. 9.7. *Solution convergence for the multigrid optimization algorithm 9.1 using the a posteriori estimation procedure 7.1: optimal solution (left), primal solution (center), and dual solution convergence (right).*

10. Conclusions. This work considers consistency and mesh adaptivity for discrete parameter estimation problems. The discrete, or *discretize – first* approach is very attractive in practice, because the dual and optimality equations in the KKT set, as well as the reduced gradient of the target functional, can be generated with low effort by automatic differentiation. However, the discrete approach introduces several complications in the inverse solution algorithm. A consistency analysis of the full set of KKT equations is needed before convergence of the discrete optimal solution can be established. Investigating consistency and stability of the discrete KKT system is one of the main objectives of this paper. While the discretization of the primal equation is *a priori* assumed to be convergent, the consistency of the dual and optimality equations does not automatically follow. We extend the dual consistency framework given in [16] to cover the third equation in the KKT set, namely, the discrete optimality condition. While consistency and stability may not hold *a priori* for this equation, the analysis shows that they can be restored through suitable consistent modifications of the discrete target functional. The theoretical results are supported by the numerical experiments with a symmetric DG discretization of the primal problem. Discontinuous Galerkin is chosen as the discretization method because of its amenability to parallel computations, and h/p-adaptivity.

A crucial ingredient in any adaptive algorithm is the error estimation step, that guides the mesh refinement process. Previous results in error estimation are either of limited practical use (because of their unfavorable dependence on the regularization parameter, or on unknown stability constants), or are valid under the assumption that the control parameter space is finite dimensional. We explain the construction of a practical error estimator for parameter estimation problems, in the more general case of infinite dimensional controls. The estimation process is based on dual–weighted residuals of first and second order sensitivity equations. The Hessian of the reduced cost functional is replaced by a low–order BFGS approximation. The use of a BFGS Hessian, obtained virtually at no cost from the optimization algorithm, keeps computational costs of the error estimation process sufficiently low to make it feasible in practice. The practical computations of the *a posteriori* estimator is demonstrated on an elliptical problem. We will also explore other choices of error functionals, and their impact on the quality of the inversion process.

Acknowledgments. The authors thank Wolfgang Bangerth and the `deal.II` community for their help with the `deal.II` software. This work was supported in part by the US National Science Foundation, through the following awards: NSF-CCF-0635194, NSF OCI-0904397, NSF CCF-0916493, and NSF DMS-0915047.

REFERENCES

[1] M. Alexe and A. Sandu, *Space-time adaptive solution of inverse problems with the discrete adjoint method*, Tech. Report TR-10-14, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, 2010.

[2] W. Bangerth, *A framework for the adaptive finite element solution of large-scale inverse problems*, SIAM J. Sci. Comput., 30 (2008), pp. 2965–2989.

[3] W. Bangerth, R. Hartmann, and G. Kanschat, `deal.II` *Differential Equations Analysis Library, Technical Reference*. `http://www.dealii.org`.

[4] W. Bangerth and A. Joshi, *Adaptive finite element methods for the solution of inverse problems in optical tomography*, Inverse Problems, 24 (2008), p. 034011 (22pp).

[5] R. Becker, *Estimating the control error in discretized PDE-constrained optimization*, J. Numer. Math., 14 (2006), pp. 163–185.

[6] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica, 10 (2001), pp. 1–102.

[7] R. Becker and B. Vexler, *A posteriori error estimation for finite element discretization of parameter identification problems*, Numer. Math., 96 (2004), pp. 435–459.

[8] ———, *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations*, J. Comput. Phys., 206 (2005), pp. 95–110.

[9] P. G. Ciarlet, *The finite element method for elliptic problems*, North Holland Publishing Company, New York, NY, USA, 2002.

[10] T. Feng, M. Gulliksson, and W. Liu, *Adaptive finite element methods for the identification of elastic constants*, J. Sci. Comput., 26 (2006), pp. 217–235.

[11] T. Feng, N. Yan, and W. Liu, *Adaptive finite element methods for the identification of distributed parameters in elliptic equation*, Advances in Computational Mathematics, 29 (2008), pp. 27–53. 10.1007/s10444-007-9035-6.

[12] A. Griewank, *Evaluating derivatives: principles and techniques of algorithmic differentiation*, SIAM, Philadelphia, PA, USA, 2000.

[13] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and studies in mathematics, Pitman Advanced Pub. Program, Boston, MA, USA, 1985.

[14] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[15] P. C. Hansen, *Discrete Inverse Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2010.

[16] R. Hartmann, *Adjoint consistency analysis of discontinuous Galerkin discretizations*, SIAM J. Numer. Anal., 45 (2007), pp. 2671–2696.

[17] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, Springer Verlag, 2007.

[18] A. Jameson, *Aerodynamic design via control theory*, J. Sci. Comput., 3 (1988), pp. 233–260.

[19] K. Kunisch and J. Zou, *Iterative choices of regularization parameters in linear inverse problems*, Inverse Problems, 14 (1998), p. 1247.

[20] R. M. Lewis and S. G. Nash, *Model problems for the multigrid optimization of systems governed by differential equations*, SIAM J. Sci. Comput., 26 (2005), pp. 1811–1837.

[21] J. L. Lions, *Optimal control of systems governed by partial differential equations*, Springer Verlag, New York, NY, USA, 1971.

[22] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, 2nd ed., 2006.

[23] R. Rannacher and B. Vexler, *Adaptive finite element discretization in PDE-based optimization*, GAMM-Mitteilungen, 33 (2010), pp. 177–193.

[24] M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations*, vol. 13 of Texts in Applied Mathematics, Springer Verlag, 2nd ed., 2004.

[25] B. Riviere, *Discontinuous Galerkin Methods For Solving Elliptic And parabolic Equations: Theory and Implementation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[26] A. Sandu, *On the properties of Runge - Kutta discrete adjoints*, in International Conference on Computational Science (4), 2006, pp. 550–557.

[27] C. R. Vogel, *Computational Methods for Inverse Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[28] Z. Wang, I. M. Navon, F. X. Le Dimet, and X. Zou, *The second order adjoint analysis: Theory and applications*, Meteorology and Atmospheric Physics, 50 (1992), pp. 3–20. 10.1007/BF01025501.

[29] J. Xie and J. Zou, *An improved model function method for choosing regularization parameters in linear inverse problems*, Inverse Problems, 18 (2002), p. 631.