CrossMark

BIT

Comparison of approximate shape gradients

R. Hiptmair · A. Paganini · S. Sargheini

Abstract Shape gradients of PDE constrained shape functionals can be stated in two equivalent ways. Both rely on the solutions of two boundary value problems (BVPs), but one involves integrating their traces on the boundary of the domain, while the other evaluates integrals in the volume. Usually, the two BVPs can only be solved approximately, for instance, by finite element methods. However, when used with finite element solutions, the equivalence of the two formulas breaks down. By means of a comprehensive convergence analysis, we establish that the volume based expression for the shape gradient generally offers better accuracy in a finite element setting. The results are confirmed by several numerical experiments.

R. Hiptmair · A. Paganini (✉) · S. Sargheini
Seminar for Applied Mathematics, ETH Zurich, Zurich, Switzerland
e-mail: alberto.paganini@sam.math.ethz.ch

R. Hiptmair
e-mail: hiptmair@sam.math.ethz.ch

S. Sargheini
e-mail: sahar.sargheini@sam.math.ethz.ch

🙂 Springer

1 Introduction

Shape calculus studies the "differentiation of shape functionals with respect to the variation of a domain they depend upon". Over the last three decades this notion has been made rigorous, notably by the introduction of the velocity method by Zolesio [10,28] and the domain perturbation method by Simon [24,25,27] and Eppler [11,12]. Shape calculus has also become important as a key tool in the field of optimization, where it supplies the so-called shape gradient, that is, the first derivative of a functional with respect to a shape, for use in the framework of descent methods. Since this article will not directly discuss methods for shape optimization we refer the reader to the monographs [2,8,15,18,19,26,28].

Shape optimization entails the approximate numerical computation of shape gradients. This step will be the focus of this article. Of course, many different shape functionals are conceivable, leading to vastly different types of shape gradients. Thus, we have to adopt a "case study approach" and restrict our study to a special, albeit important, class of shape functionals.

The shape functionals under scrutiny are least squares output functionals for solutions of scalar second-order elliptic boundary value problems. They belong to the category of PDE constrained shape functionals and have widely been considered in articles on shape optimization [3,16].

In [2], for instance, formulas have been derived for the associated shape gradients. They are based on solutions u and p of two boundary value problems, called state and adjoint problem. Starting point for our investigations was the insight that the formulas can be stated in two equivalent ways, (i) as expressions involving traces of u and p on the boundary of the domain, and (ii) by means of volume integrals on the domain, see [5, Sect. 6].

The situation resembles that faced for quite a few common output functionals depending on solutions of BVPs for second-order elliptic PDEs. Examples are the total heat flux in heat conduction, lift functionals for potential flow [16], far field functionals [22,23], and electromagnetic force functionals [21]. All these functionals can be stated as integrals either over boundaries or over parts of the domain, and the same value is obtained when inserting exact solutions of the BVPs. Both kinds of formulas can also be used in the context of finite element approximation, but when applied to discrete solutions, they fail to give the same answer. More strikingly, the volume integrals often display much faster convergence and provide superior accuracy compared to their boundary based counterparts. An explanation is that the expressions featuring volume integrals enjoy continuity in energy norm, whereas integrals of traces are not well-defined on the natural variational spaces. This makes a crucial difference, because we can benefit from superconvergence, when evaluating continuous functionals for Galerkin solutions [4, Sect. 2].

This made us suspect that similar effects could be observed for the different expressions for shape gradients and their use with finite element solutions. The analysis and numerical experiments of this article largely confirm our expectation that volume based expressions for the shape gradient often offer better accuracy than the use of formulas involving traces on boundaries. This is the message of both the a priori convergence

estimates developed in Sect. 3, see Theorems 3.1 and 3.2, and of the numerical tests reported in Sect. 4.

What compounds the difficulties of gauging the quality of formulas for shape gradients is the fact that they must be viewed as linear functionals on spaces of infinitesimal deformations. Of course, one can switch back to functions via the Riesz representation theorem, but the choice of the underlying inner product is somewhat arbitrary and might bias the outcome. Thus, we have decided to study the errors of shape gradients directly in the relevant dual norms.

2 Shape gradients

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be an open bounded domain with piecewise smooth boundary $\partial \Omega$, and let $\mathcal{J}(\Omega) \in \mathbb{R}$ be a real-valued quantity of interest associated to it. One is often interested in its shape sensitivity, which quantifies the impact of small perturbations of $\partial \Omega$ on the value $\mathcal{J}(\Omega)$.

For this purpose, we model perturbations of the domain Ω through maps of the form

$$T_\mathcal{V} := \mathcal{I} + \mathcal{V}, \tag{2.1}$$

where $\mathcal{I}$ is the identity operator and $\mathcal{V}$ is a vector field in $C^1(\mathbb{R}^d; \mathbb{R}^d)$. It can easily be proven that the map (2.1) is a diffeomorphism for $\|\mathcal{V}\|_{C^1} < 1$ [2, Lemma 6.13]. Therefore, it is natural to consider $\mathcal{J}(\Omega)$ as the realization of a shape functional, a real map

$$\mathcal{J} : \mathcal{A} \to \mathbb{R}$$

defined on the family of admissible domains

$$\mathcal{A} := \left\{ T_\mathcal{V}(\Omega) \, ; \mathcal{V} \in C^1(\mathbb{R}^d; \mathbb{R}^d) \, , \|\mathcal{V}\|_{C^1} < 1 \right\} .$$

The sensitivity of $\mathcal{J}(\Omega)$ with respect to the perturbation direction $\mathcal{V}$ can be expressed through the *Eulerian derivative* of the shape functional $\mathcal{J}$ in the direction $\mathcal{V}$, that is,

$$d\mathcal{J}(\Omega; \mathcal{V}) := \lim_{s \searrow 0} \frac{\mathcal{J}(T_{s \cdot \mathcal{V}}(\Omega)) - \mathcal{J}(\Omega)}{s} . \tag{2.2}$$

It goes without saying that it is desirable that (2.2) exists for all possible perturbation directions $\mathcal{V}$. It is therefore natural to define a shape functional $\mathcal{J}$ to be *shape differentiable* at Ω if the mapping

$$d\mathcal{J}(\Omega; \cdot) : C^1(\mathbb{R}^d; \mathbb{R}^d) \to \mathbb{R}, \qquad \mathcal{V} \mapsto d\mathcal{J}(\Omega; \mathcal{V}) . \tag{2.3}$$

defined by (2.2) is linear and bounded on $C^1(\mathbb{R}^d; \mathbb{R}^d)$. In literature, the mapping $d\mathcal{J}(\Omega; \mathcal{V})$ is called *shape gradient* of $\mathcal{J}$ at Ω, as it is the Gâteaux derivative in

$0 \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ of the map

$$\mathcal{V} \mapsto \mathcal{J}(T_{\mathcal{V}}(\Omega)) \,,$$

see [10, Ch. 9, Def. 2.2]. Note that Formula (2.2) is well-defined for any vector field in the Banach space $C^1(\mathbb{R}^d; \mathbb{R}^d)$, and the shape gradient is an element of its dual space.

Remark 2.1 In literature, perturbations as in (2.1) are known as *perturbations of the identity*. From a differential geometry point of view, this approach is less general than the so called *velocity method*, which is, for instance, introduced in [10, Ch. 4]. However, both methods lead to the same formula for the shape gradient, which merely takes into account first order perturbations of the shape functional $\mathcal{J}$ [10, Ch. 9, Thm 3.2].

An interesting property of shape gradients is expressed in the *Hadamard structure theorem* [10, Ch. 9, Thm 3.6]: If $\partial\Omega$ is smooth, $d\mathcal{J}(\Omega; \cdot)$ admits a representative $\mathfrak{g}(\Omega)$ in the space of distributions $\mathcal{D}^k(\partial\Omega)$

$$d\mathcal{J}(\Omega; \mathcal{V}) = \langle \mathfrak{g}(\Omega), \gamma_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \rangle_{\mathcal{D}^k(\partial\Omega)} \,, \tag{2.4}$$

where $\gamma_{\partial\Omega} \mathcal{V} \cdot \mathbf{n}$ is the normal component of $\mathcal{V}$ on the boundary $\partial\Omega$. This implies that only normal displacements of the boundary have an impact on the value of $\mathcal{J}(\Omega)$. However, we should take into account that this is no longer true, if the boundary $\partial\Omega$ is only piecewise smooth.

We are particularly interested in PDE constrained shape functionals of the form

$$\mathcal{J}(\Omega) = \int_{\Omega} j(u) \, d\mathbf{x} \,, \tag{2.5}$$

where $j : \mathbb{R} \to \mathbb{R}$ possesses a locally Lipschitz continuous derivative j' and u is the solution of the *state problem*, a scalar elliptic equation with Neumann or Dirichlet boundary conditions

$$\begin{cases} \mathcal{L}(u) = f & \text{in } \Omega \,, \\ u = g \text{ or } \frac{\partial u}{\partial \mathbf{n}} = g & \text{on } \partial\Omega \,. \end{cases} \tag{2.6}$$

The functions f and g are assumed to belong to $L^2(\mathbb{R}^d)$ $(H^1(\mathbb{R}^d)$ in the case of the Neumann BVP) and $H^2(\mathbb{R}^d)$, respectively, and they are identified with their restrictions onto Ω and $\partial\Omega$.

Explicit formulas for $d\mathcal{J}(\Omega)$ can easily be derived both for unconstrained and PDE constrained shape functionals, cf. [10, Ch. 9, Sect. 4.3, and Ch. 10, Sect. 2.5]. In the case of PDE constrained shape functionals, the formulas involve the integrals of u, the solution of (2.6), and of p, the solution of the adjoint problem[1]

$$\begin{cases} \mathcal{L}(p) = j'(u) & \text{in } \Omega \,, \\ p = 0 \text{ or } \frac{\partial p}{\partial \mathbf{n}} = 0 & \text{on } \partial\Omega \,. \end{cases} \tag{2.7}$$

---

[1] For simplicity, we assume that the operator $\mathcal{L}$ is self-adjoint.

As different $\mathcal{L}$ lead to different formulas for the Eulerian derivative, from now on we consider only the model elliptic operator

$$\mathcal{L}(u) = -\Delta u + u, \tag{2.8}$$

which should be regarded as a representative for the class of scalar elliptic differential operators of second order.

As mentioned in the introduction, $d\mathcal{J}(\Omega; \mathcal{V})$ can be formulated as an integral over a volume, as well as an integral on the boundary. For example, the formula for the PDE constrained shape functional (2.5) with elliptic operator (2.8) and Dirichlet boundary conditions $u = g$ on $\partial\Omega$ reads (see the Appendix for the derivation)

$$\begin{aligned}
d\mathcal{J}(\Omega; \mathcal{V}) = \int_\Omega \Big(& \nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p - f\mathcal{V} \cdot \nabla p \\
& + \operatorname{div} \mathcal{V}(j(u) - \nabla u \cdot \nabla p - up) \\
& + (j'(u) - p)(\nabla g \cdot \mathcal{V}) - \nabla p \cdot \nabla(\nabla g \cdot \mathcal{V}) \Big) \, d\mathbf{x},
\end{aligned} \tag{2.9}$$

and can be recast as

$$d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} (\mathcal{V} \cdot \mathbf{n}) \left(j(u) + \frac{\partial p}{\partial \mathbf{n}} \frac{\partial(u - g)}{\partial \mathbf{n}} \right) dS. \tag{2.10}$$

The volume integral (2.9) and the boundary integral (2.10) are equivalent representations of the shape gradient $d\mathcal{J}(\Omega; \mathcal{V})$. They can be converted into each other by means of integration by parts on $\partial\Omega$ [28, Sect. 3.8] and Gauss's theorem. However, the bulk of literature mainly considers (2.10) and does not pay attention to (2.9), probably because the former better matches the Hadamard structure theorem (2.4). Only recently it has been realized that the volume representation (2.9) may be better suited for computations, see [5] and [10, Ch. 10, Remark 2.3].

Remark 2.2 In the case of Neumann boundary conditions on smooth domains, the counterparts of Formulas (2.9) and (2.10) read

$$\begin{aligned}
d\mathcal{J}(\Omega; \mathcal{V}) = \int_\Omega \Big(& (\nabla f \cdot \mathcal{V})p + \nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p \\
& + \operatorname{div} \mathcal{V}(fp + j(u) - \nabla u \cdot \nabla p - up) \Big) \, d\mathbf{x} \\
& + \int_{\partial\Omega} (\nabla g \cdot \mathcal{V})p + gp \operatorname{div}_\Gamma \mathcal{V} \, dS,
\end{aligned} \tag{2.11}$$

where $\operatorname{div}_\Gamma$ denotes the tangential divergence on $\partial\Omega$, and

$$d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) - \nabla u \cdot \nabla p - up + fp + \frac{\partial gp}{\partial \mathbf{n}} + \mathrm{K}gp \right) dS, \tag{2.12}$$

where K is the mean curvature of $\partial\Omega$.

Remark 2.3 In general, the shape gradient does not feature the Hadamard structure (2.4) if the boundary is piecewise smooth only. For instance, in the presence of corners in 2D, Formula (2.12) has to be corrected by adding the term

$$\sum_i p(\mathbf{a}_i) g(\mathbf{a}_i) \mathcal{V}(\mathbf{a}_i) \cdot [[\tau(\mathbf{a}_i)]], \tag{2.13}$$

where the $\mathbf{a}_i$ denote the corner points and $[[\tau(\mathbf{a}_i)]]$ is the jump of the tangential unit vector field in the corner $\mathbf{a}_i$ [28, Ch. 3.8]. On the other hand, no correction has to be made to formula (2.10).

3 Approximation of shape gradients

In this section we investigate the approximation of the shape gradient $d\mathcal{J}$. For the sake of readability, we perform the analysis for the elliptic operator (2.8) with Dirichlet boundary conditions only. The results can easily be extended to general elliptic operators in divergence form with both Dirichlet and Neumann boundary conditions.

To highlight the dependence of $d\mathcal{J}$ on the solution of the state and adjoint problem u and p, as well as to distinguish between formulas (2.9) and (2.10), we introduce the notations

$$d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Vol}} := \int_\Omega \left(\nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p - f\mathcal{V} \cdot \nabla p \right.$$
$$+ \operatorname{div} \mathcal{V}(j(u) - \nabla u \cdot \nabla p - up)$$
$$\left. + (j'(u) - p)(\nabla g \cdot \mathcal{V}) - \nabla p \cdot \nabla(\nabla g \cdot \mathcal{V}) \right) d\mathbf{x}, \tag{3.1}$$

$$d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Bdry}} := \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) + \frac{\partial p}{\partial\mathbf{n}} \frac{\partial(u-g)}{\partial\mathbf{n}} \right) dS. \tag{3.2}$$

Note that, provided u and p are exact solutions of (2.6) and (2.7),

$$d\mathcal{J}(\Omega; \mathcal{V}) = d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Vol}} = d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Bdry}}. \tag{3.3}$$

The operator $d\mathcal{J}(\Omega; \cdot)$ can be approximated by replacing the functions u and p with Ritz–Galerkin Lagrangian finite element solutions of (2.6) and (2.7) respectively. We consider approximations based on discretization with finite elements, as this approach is very popular in shape optimization due to its flexibility for engineering applications. Approximations based on boundary element methods are also possible, cf. [13,17,29].

Equality (3.3) certainly breaks down when the functions u and p are approximated with finite elements [5]. Thus, a natural question is, which formula, (3.1) or (3.2), should be preferred for an approximation of $d\mathcal{J}(\Omega; \cdot)$ in the operator norm. The answer is provided by Theorems 3.1 and 3.2. Next we state a few assumptions necessary for a precise statement of the theorems.

Assumption 1 The Dirichlet BVP for the Laplacian is H^2-regular [6, Ch. II, Def. 7.1], that is, if a function $w \in H_0^1(\Omega)$ is the (unique) weak solution of the elliptic BVP

$$\begin{cases} -\Delta w + w = \rho & \text{in } \Omega , \\ w = 0 & \text{on } \partial\Omega , \end{cases}$$

for a function $\rho \in L^2(\Omega)$, then $w \in H^2(\Omega)$, and there is a constant C_r, depending only on Ω, so that

$$\|w\|_{H^2(\Omega)} \leq C_r \|\rho\|_{L^2(\Omega)} .$$

Remark 3.1 Assumption 1 holds for convex Lipschitz domains and (possibly non-convex) domains with C^2 boundary [6, Ch. II, Thm 7.2].

Assumption 2 The source function f and the boundary data g in (2.6) are restrictions of functions in $H^1(\mathbb{R}^d)$ and $H^3(\mathbb{R}^d)$ to Ω and $\partial\Omega$, respectively.

Next, for an index set $\mathbb{H}$, we introduce a family $(V_h)_{h \in \mathbb{H}}$ of finite-dimensional subspaces of $H_0^1(\Omega)$ and define $u_h \in g + V_h$, $p_h \in V_h$ as Ritz–Galerkin solutions[2] of (2.6) and (2.7), respectively, that is,

$$\int_\Omega \nabla u_h \cdot \nabla v_h + u_h v_h \, d\mathbf{x} = \int_\Omega f v_h \, d\mathbf{x} \qquad \forall v_h \in V_h , \tag{3.4}$$

$$\int_\Omega \nabla p_h \cdot \nabla v_h + p_h v_h \, d\mathbf{x} = \int_\Omega j(u_h) v_h \, d\mathbf{x} \qquad \forall v_h \in V_h . \tag{3.5}$$

In particular, let $(V_h)_{h \in \mathbb{H}}$ be a family of H^1-conforming piecewise linear Lagrangian finite element spaces built on a shape-regular and quasi-uniform family of simplicial meshes [6, Ch. II, Def. 5.1], and let h designate the meshwidth. We recall that the associated family of nodal interpolation operators

$$\mathcal{I}_h : H^2(\Omega) \cap H_0^1(\Omega) \to V_h$$

satisfies[3] [6, Ch. II, Thm 6.4]

$$\|w - \mathcal{I}_h w\|_{H^1(\Omega)} \leq Ch |w|_{H^2(\Omega)} \quad \forall h \in \mathbb{H}. \tag{3.6}$$

Theorem 3.1 *Let u and p be the solutions of (2.6) and (2.7), and let u_h and p_h be their Ritz–Galerkin approximations in the sense of (3.4) and (3.5) by piecewise linear Lagrangian finite elements. Furthermore, let Assumptions 1 and 2 be satisfied. Then*[4]

$$|d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\mathrm{Vol}}| \leq C(\Omega, u, p, f, g) h^2 \|\mathcal{V}\|_{W^{2,4}(\Omega)} ,$$

---

[2] Note that p_h is not a proper Ritz–Galerkin solution of (2.7), because the right-hand side is perturbed.

[3] We write C for generic constants, whose value may differ between different occurrences. They may depend only on Ω, shape-regularity and quasi-uniformity of the meshes.

[4] For the sake of readability, we use the same notation for scalar and vectorial Sobolev norms.

where the constant $C(\Omega, u, p, f, g)$ depends on the domain Ω and its discretization, $\|u\|_{H^2(\Omega)}$, $\|p\|_{H^2(\Omega)}$, $\|f\|_{H^1(\Omega)}$, *and* $\|g\|_{H^3(\Omega)}$.

Proof The proof heavily relies on duality techniques that are repeatedly used to obtain estimates for the various terms in (3.1). The impatient reader may skip the proof after (3.14) and will get main ideas nevertheless.

From the equality $d\mathcal{J}(\Omega; \mathcal{V}) = d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Vol}}$, we immediately get by the triangle inequality

$$
\begin{aligned}
&|d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}| \\
&\leq \left(\left| \int_\Omega \nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p - \nabla u_h \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p_h \, d\mathbf{x} \right| \right. \\
&\quad + \left| \int_\Omega f\mathcal{V} \cdot \nabla(p - p_h) \, d\mathbf{x} \right| \\
&\quad + \left| \int_\Omega \operatorname{div} \mathcal{V}(j(u) - j(u_h) - \nabla u \cdot \nabla p - up + \nabla u_h \cdot \nabla p_h + u_h p_h) \, d\mathbf{x} \right| \\
&\quad + \left| \int_\Omega (j'(u) - j'(u_h) - p + p_h)(\nabla g \cdot \mathcal{V}) \, d\mathbf{x} \right| \\
&\quad + \left. \left| \int_\Omega \nabla(p - p_h) \cdot \nabla(\nabla g \cdot \mathcal{V}) \, d\mathbf{x} \right| \right) .
\end{aligned} \tag{3.7}
$$

The proof boils down to bounding each integral in the previous inequality and applying standard finite element convergence and interpolation estimates. To begin with, we split the first integral into

$$
\begin{aligned}
&\int_\Omega (\nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p - \nabla u_h \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p_h \, d\mathbf{x} \\
&= \int_\Omega \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p \, d\mathbf{x} \\
&\quad + \int_\Omega \nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla(p - p_h) \, d\mathbf{x} \\
&\quad - \int_\Omega \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla(p - p_h) \, d\mathbf{x} .
\end{aligned} \tag{3.8}
$$

To bound the first and the second integral on the right-hand side of (3.8) we make use of standard duality techniques. For the first one we introduce the function w as weak solution of the adjoint BVP[5]

$$
\begin{cases}
-\Delta w + w = -\operatorname{div}\left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p\right) & \text{in } \Omega , \\
w = 0 & \text{on } \partial\Omega ,
\end{cases} \tag{3.9}
$$

---

[5] Many bounds in this proof rely on duality techniques, which introduce so-called adjoint BVPs. For the sake of readability, we abuse the notation and we always denote by w the solutions of these BVPs.

that is,

$$\int_\Omega \nabla w \cdot \nabla v + wv\, d\mathbf{x} = \int_\Omega \left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p \right) \cdot \nabla v \quad \forall\, v \in H_0^1(\Omega). \quad (3.10)$$

We recall that for two generic functions $q_1, q_2 \in L^4(\Omega)$ the Cauchy–Schwarz inequality implies

$$\|q_1 q_2\|_{L^2(\Omega)} \le \|q_1\|_{L^4(\Omega)} \|q_2\|_{L^4(\Omega)}. \quad (3.11)$$

By the triangle inequality, (3.11) and the Sobolev Imbedding Theorem [1, Thm 4.12], we bound the source function in (3.9) by

$$\| \operatorname{div} \left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p \right) \|_{L^2(\Omega)}$$
$$\le C \left(\|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{W^{1,4}(\Omega)} + \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|p\|_{H^2(\Omega)} \right),$$
$$\le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{H^2(\Omega)}.$$

By Assumption 1, w is in $H^2(\Omega)$ and it satisfies

$$\|w\|_{H^2(\Omega)} \le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{H^2(\Omega)}. \quad (3.12)$$

By exploiting the Galerkin orthogonality of $u - u_h$ to the finite dimensional trial space $V_h \subset H_0^1(\Omega)$, we derive the bound

$$\left| \int_\Omega \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p\, d\mathbf{x} \right|$$
$$= \left| \int_\Omega \nabla(u - u_h) \cdot \nabla w + (u - u_h)w\, d\mathbf{x} \right|,$$
$$= \left| \int_\Omega \nabla(u - u_h) \cdot \nabla(w - \mathcal{I}_h w) + (u - u_h)(w - \mathcal{I}_h w)\, d\mathbf{x} \right|,$$
$$\le \|u - u_h\|_{H^1(\Omega)} \|w - \mathcal{I}_h w\|_{H^1(\Omega)}. \quad (3.13)$$

Then by (3.6) and the standard finite element convergence estimate [6, Ch. II, Sect. 7]

$$\|u - u_h\|_{H^1(\Omega)} \le Ch \|u\|_{H^2(\Omega)}, \quad (3.14)$$

we conclude from (3.12)

$$\left| \int_\Omega \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p\, d\mathbf{x} \right| \le Ch^2 \|u\|_{H^2(\Omega)} \|p\|_{H^2(\Omega)} \|\mathcal{V}\|_{W^{2,4}(\Omega)}.$$

Similarly, for the second integral on the right-hand side of (3.8) we introduce the function w as weak solution of the adjoint BVP

$$\begin{cases} -\Delta w + w = -\operatorname{div} \left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla u \right) & \text{in } \Omega, \\ w = 0 & \text{on } \partial\Omega, \end{cases} \quad (3.15)$$

that is,

$$\int_\Omega \nabla w \cdot \nabla v + wv \, d\mathbf{x} = \int_\Omega \left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla u \right) \cdot \nabla v \quad \forall v \in H_0^1(\Omega) \,. \quad (3.16)$$

Assumption 1 and the bound

$$\| \operatorname{div} \left((\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla u \right) \|_{L^2(\Omega)} \leq C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)}$$

imply that $w \in H^2(\Omega)$ and that it satisfies

$$\|w\|_{H^2(\Omega)} \leq C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)} \,. \quad (3.17)$$

Next, we note that for every $v_h \in V_h$

$$\int_\Omega \nabla(p - p_h) \cdot \nabla v_h + (p - p_h)v_h \, d\mathbf{x} = \int_\Omega (j(u) - j(u_h))v_h \, d\mathbf{x}, \quad (3.18)$$

which implies

$$\left| \int_\Omega \nabla(p - p_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla u \, d\mathbf{x} \right| = \left| \int_\Omega \nabla(p - p_h) \cdot \nabla w + (p - p_h)w \, d\mathbf{x} \right|,$$

$$\leq \left| \int_\Omega \nabla(p - p_h) \cdot \nabla(w - \mathcal{I}_h w) + (p - p_h)(w - \mathcal{I}_h w) \, d\mathbf{x} \right|$$

$$+ \left| \int_\Omega (j(u) - j(u_h)) \mathcal{I}_h w \, d\mathbf{x} \right|,$$

$$\leq \|p - p_h\|_{H^1(\Omega)} \|w - \mathcal{I}_h w\|_{H^1(\Omega)} + \|\mathcal{I}_h w\|_{L^2(\Omega)} \|j(u) - j(u_h)\|_{L^2(\Omega)} \,.$$

$$(3.19)$$

For the concrete BVP considered the state solution u will belong to $C^0(\overline{\Omega})$. Further, $L^\infty(\Omega)$-estimates for finite element solutions [6, Ch. II, Sect. 7] ensure that $\|u - u_h\|_{L^\infty(\Omega)} \to 0$ as $h \to 0$. Hence, we can take for granted that there are h-independent bounds $\underline{u}$ and $\overline{u}$

$$-\infty < \underline{u} \leq u(\mathbf{x}), u_h(\mathbf{x}) \leq \overline{u} < \infty \quad \forall \mathbf{x} \in \Omega \,. \quad (3.20)$$

We write $I := [\underline{u}, \overline{u}]$ and point out that j' is bounded on I. Thus the standard finite element convergence estimate [6, Ch. II, Sect. 7]

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \|u\|_{H^2(\Omega)} \,, \quad (3.21)$$

gives

$$\|j(u) - j(u_h)\|_{L^2(\Omega)} \leq \|j'\|_{C^0(I)} \|u - u_h\|_{L^2(\Omega)} \,,$$

$$\leq Ch^2 \|j'\|_{C^0(I)} \|u\|_{H^2(\Omega)} \,. \quad (3.22)$$

In order to establish a bound for (3.19), we follow the arguments in the proof of Strang's first lemma [6, Ch. III, Thm. 1.1]. We note that for every $v_h \in V_h$

$$
\begin{aligned}
\|p_h - v_h\|_{H^1(\Omega)}^2 &= \int_\Omega \nabla(p_h - p) \cdot \nabla(p_h - v_h) + (p_h - p)(p_h - v_h)\, d\mathbf{x} \\
&\quad + \int_\Omega \nabla(p - v_h) \cdot \nabla(p_h - v_h) + (p - v_h)(p_h - v_h)\, d\mathbf{x} \\
&\leq \left(\|j(u_h) - j(u)\|_{L^2(\Omega)} + \|p - v_h\|_{H^1(\Omega)} \right) \|p_h - v_h\|_{H^1(\Omega)},
\end{aligned}
\tag{3.23}
$$

where in the last step we used (3.18) and the Cauchy–Schwarz inequality. Then by the triangle inequality, (3.23) and (3.6),

$$
\begin{aligned}
\|p - p_h\|_{H^1(\Omega)} &\leq \|p - \mathcal{I}_h p\|_{H^1(\Omega)} + \|\mathcal{I}_h p - p_h\|_{H^1(\Omega)}, \\
&\leq 2\|p - \mathcal{I}_h p\|_{H^1(\Omega)} + \|j(u_h) - j(u)\|_{L^2(\Omega)}, \\
&\leq Ch\|p\|_{H^2(\Omega)} + Ch^2 \|j'\|_{C^0(I)} \|u\|_{H^2(\Omega)},
\end{aligned}
\tag{3.24}
$$

which implies

$$
\left| \int_\Omega \nabla(p - p_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla u\, d\mathbf{x} \right|
$$
$$
\leq Ch^2 \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)} (\|p\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)} \|j'\|_{C^0(I)}).
$$

Finally, by the Cauchy–Schwarz inequality, (3.14) and (3.24), the following bound for the third integral on the right-hand side of (3.8) holds.

$$
\left| \int_\Omega \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla(p - p_h)\, d\mathbf{x} \right|
$$
$$
\leq \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|u - u_h\|_{H^1(\Omega)} \|p - p_h\|_{H^1(\Omega)},
$$
$$
\leq Ch^2 \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|u\|_{H^2(\Omega)} \|p\|_{H^2(\Omega)}.
$$

To bound the second integral on the right-hand side of (3.7) we introduce the function w as weak solution of the adjoint BVP

$$
\begin{cases}
-\Delta w + w = -\operatorname{div}(f\mathcal{V}) & \text{in } \Omega, \\
w = 0 & \text{on } \partial\Omega,
\end{cases}
\tag{3.25}
$$

that is,

$$
\int_\Omega \nabla w \cdot \nabla v + wv\, d\mathbf{x} = \int_\Omega f\mathcal{V} \cdot \nabla v \quad \forall v \in H_0^1(\Omega).
\tag{3.26}
$$

Note that

$$
\| -\operatorname{div}(f\mathcal{V}) \|_{L^2(\Omega)} \leq \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|f\|_{H^1(\Omega)},
$$

which implies that w is in $H^2(\Omega)$ and that it satisfies

$$\|w\|_{H^2(\Omega)} \leq C\|\mathcal{V}\|_{W^{1,\infty}(\Omega)}\|f\|_{H^1(\Omega)} . \tag{3.27}$$

Then by (3.26), (3.18), (3.24), (3.6), and (3.21),

$$\left| \int_\Omega f\mathcal{V} \cdot \nabla(p - p_h)\, d\mathbf{x} \right| = \left| \int_\Omega \nabla(p - p_h) \cdot \nabla w + (p - p_h)w\, d\mathbf{x} \right| ,$$

$$\leq \|p - p_h\|_{H^1(\Omega)}\|w - \mathcal{I}_h w\|_{H^1(\Omega)} + \|\mathcal{I}_h w\|_{L^2(\Omega)}\|j'\|_{C^0(I)}\|u - u_h\|_{L^2(\Omega)} ,$$

$$\leq Ch^2\|\mathcal{V}\|_{W^{1,\infty}(\Omega)}\|f\|_{H^1(\Omega)} \left(\|p\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)}\|j'\|_{C^0(I)} \right) .$$

To bound the third integral on the right-hand side of (3.7), we first apply the triangle inequality

$$\left| \int_\Omega \operatorname{div} \mathcal{V}(j(u) - j(u_h) - \nabla u \cdot \nabla p - up + \nabla u_h \cdot \nabla p_h + u_h p_h)\, d\mathbf{x} \right|$$

$$\leq \left| \int_\Omega \operatorname{div} \mathcal{V}(j(u) - j(u_h))\, d\mathbf{x} \right|$$

$$+ \left| \int_\Omega \operatorname{div} \mathcal{V}(\nabla u \cdot \nabla p + up - \nabla u_h \cdot \nabla p_h - u_h p_h)\, d\mathbf{x} \right| . \tag{3.28}$$

The first integral on the right-hand side of (3.28) can be bounded by

$$\left| \int_\Omega \operatorname{div} \mathcal{V}(j(u) - j(u_h))\, d\mathbf{x} \right| \leq C\|\mathcal{V}\|_{W^{1,\infty}}\|j'\|_{C^0(I)}\|u - u_h\|_{L^2(\Omega)} ,$$

$$\leq Ch^2\|\mathcal{V}\|_{W^{1,\infty}}\|j'\|_{C^0(I)}\|u\|_{H^2(\Omega)} , \tag{3.29}$$

whereas the second one can conveniently be rewritten as

$$\left| \int_\Omega \operatorname{div} \mathcal{V}(\nabla u \cdot \nabla p + up - \nabla u_h \cdot \nabla p_h - u_h p_h)\, d\mathbf{x} \right|$$

$$= \left| \int_\Omega \operatorname{div} \mathcal{V} \left(\nabla(u - u_h) \cdot \nabla p + (u - u_h)p \right)\, d\mathbf{x} \right.$$

$$+ \int_\Omega \operatorname{div} \mathcal{V} \left(\nabla u \cdot \nabla(p - p_h) + u(p - p_h) \right)\, d\mathbf{x}$$

$$\left. - \int_\Omega \operatorname{div} \mathcal{V} \left(\nabla(u - u_h) \cdot \nabla(p - p_h) + (u - u_h)(p - p_h) \right)\, d\mathbf{x} \right| . \tag{3.30}$$

Again, the first two integrals on the right-hand side of (3.30) can be bounded with standard duality techniques. For the first one we introduce the function w as weak solution of the adjoint BVP

$$\begin{cases} -\Delta w + w = -\operatorname{div}(\operatorname{div}(\mathcal{V})\nabla p) + \operatorname{div}(\mathcal{V})p & \text{in } \Omega , \\ w = 0 & \text{on } \partial\Omega , \end{cases} \tag{3.31}$$

that is,

$$\int_\Omega \nabla w \cdot \nabla v + wv \, d\mathbf{x} = \int_\Omega \operatorname{div}(\mathcal{V}) \, (\nabla p \cdot \nabla v + pv) \, d\mathbf{x} \quad \forall \, v \in H_0^1(\Omega). \quad (3.32)$$

Assumption 1 and the bound

$$\| \operatorname{div}(\operatorname{div}(\mathcal{V})\nabla p) + \operatorname{div}(\mathcal{V})p \|_{L^2(\Omega)} \le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{H^2(\Omega)}$$

imply that w is in $H^2(\Omega)$ and that it satisfies

$$\|w\|_{H^2(\Omega)} \le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{H^2(\Omega)}. \quad (3.33)$$

Then by (3.32), Galerkin orthogonality of $u - u_h$ to V_h, the Cauchy–Schwarz inequality, (3.14), and (3.6),

$$\left| \int_\Omega \operatorname{div} \mathcal{V} \, (\nabla p \cdot \nabla (u - u_h) + p(u - u_h)) \, d\mathbf{x} \right|$$
$$= \int_\Omega \nabla w \cdot \nabla (u - u_h) + w(u - u_h) \, d\mathbf{x} \le \|u - u_h\|_{H^1(\Omega)} \|w - \mathcal{I}_h w\|_{H^1(\Omega)},$$
$$\le C h^2 \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|p\|_{H^2(\Omega)} \|u\|_{H^2(\Omega)}.$$

For the second integral on the right-hand side of (3.30) we introduce the function w as weak solution of the adjoint BVP

$$\begin{cases} -\Delta w + w = -\operatorname{div}(\operatorname{div}(\mathcal{V})\nabla u) + \operatorname{div}(\mathcal{V})u & \text{in } \Omega, \\ w = 0 & \text{on } \partial\Omega, \end{cases} \quad (3.34)$$

that is,

$$\int_\Omega \nabla w \cdot \nabla v + wv \, d\mathbf{x} = \int_\Omega \operatorname{div}(\mathcal{V}) \, (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} \quad \forall \, v \in H_0^1(\Omega). \quad (3.35)$$

Assumption 1 and the bound

$$\| \operatorname{div}(\operatorname{div}(\mathcal{V})\nabla u) + \operatorname{div}(\mathcal{V})u \|_{L^2(\Omega)} \le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)}$$

imply that w is in $H^2(\Omega)$ and that it satisfies

$$\|w\|_{H^2(\Omega)} \le C \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)}. \quad (3.36)$$

Then, by (3.35), (3.18), the Cauchy–Schwarz inequality, (3.24), (3.6), and (3.21),

$$\left| \int_\Omega \operatorname{div} \mathcal{V} \left(\nabla u \cdot \nabla (p - p_h) + u(p - p_h) \right) d\mathbf{x} \right|$$
$$= \left| \int_\Omega \nabla (p - p_h) \cdot \nabla w + (p - p_h) w \, d\mathbf{x} \right|,$$
$$\le \|p - p_h\|_{H^1(\Omega)} \|w - \mathcal{I}_h w\|_{H^1(\Omega)} + \|\mathcal{I}_h w\|_{L^2(\Omega)} \|j'\|_{C^0(I)} \|u - u_h\|_{L^2(\Omega)},$$
$$\le C h^2 \|\mathcal{V}\|_{W^{2,4}(\Omega)} \|u\|_{H^2(\Omega)} \left(\|p\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)} \|j'\|_{C^0(I)} \right).$$

By the Cauchy–Schwarz inequality, (3.14), and (3.24), we obtain the following bound for the third integral on the right-hand side of (3.30):

$$\left| \int_\Omega \operatorname{div} \mathcal{V} \left(\nabla (u - u_h) \cdot \nabla (p - p_h) + (u - u_h)(p - p_h) \right) d\mathbf{x} \right|$$
$$\le \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|u - u_h\|_{H^1(\Omega)} \|p - p_h\|_{H^1(\Omega)},$$
$$\le C h^2 \|\mathcal{V}\|_{W^{1,\infty}(\Omega)} \|u\|_{H^2(\Omega)} \|p\|_{H^2(\Omega)}.$$

The fourth integral on the right-hand side of (3.7) can be bounded similarly as in (3.29), relying on $L^\infty(\Omega)$-estimates for finite element solutions. Now, we make use of the uniform Lipschitz continuity of j' on the compact interval I, which yields

$$\left| \int_\Omega (j'(u) - j'(u_h) - p + p_h)(\nabla g \cdot \mathcal{V}) \, d\mathbf{x} \right|$$
$$\le \|\mathcal{V}\|_{L^\infty(\Omega)} \|g\|_{H^1(\Omega)} \left(\|j'\|_{C^{0,1}(I)} \|u - u_h\|_{L^2(\Omega)} + \|p - p_h\|_{L^2(\Omega)} \right),$$

and since (3.18), (3.22), and (3.24) imply [6, Ch. III, Sect. 1]

$$\|p - p_h\|_{L^2(\Omega)} \le C h^2 \|j'\|_{C^0(I)} \|p\|_{H^2(\Omega)}, \tag{3.37}$$

we conclude

$$\left| \int_\Omega (j'(u) - j'(u_h) - p + p_h)(\nabla g \cdot \mathcal{V}) \, d\mathbf{x} \right|$$
$$\le C h^2 \|j'\|_{C^{0,1}(I)} \|\mathcal{V}\|_{L^\infty(\Omega)} \|g\|_{H^1(\Omega)} \left(\|u\|_{H^2(\Omega)} + \|p\|_{H^2(\Omega)} \right).$$

Finally, the fifth integral on the right-hand side of (3.7) can be bounded with standard duality techniques by introducing the function w as weak solution of the adjoint BVP

$$\begin{cases} -\Delta w + w = -\Delta (\nabla g \cdot \mathcal{V}) & \text{in } \Omega, \\ w = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.38}$$

that is,

$$\int_\Omega \nabla w \cdot \nabla v + w v \, d\mathbf{x} = \int_\Omega \nabla (\nabla g \cdot \mathcal{V}) \cdot \nabla v \, d\mathbf{x} \quad \forall v \in H_0^1(\Omega). \tag{3.39}$$

Assumption 1 and the bound

$$\|\Delta(\nabla g \cdot \mathcal{V})\|_{L^2(\Omega)}$$
$$\leq C\left(\|\mathcal{V}\|_{L^\infty(\Omega)}\|g\|_{H^3(\Omega)} + \|\mathcal{V}\|_{W^{1,\infty}(\Omega)}\|g\|_{H^2(\Omega)} + \|\mathcal{V}\|_{H^2(\Omega)}\|g\|_{W^{1,\infty}(\Omega)}\right)$$
$$\leq C\|\mathcal{V}\|_{W^{2,4}(\Omega)}\|g\|_{H^3(\Omega)}$$

imply that w is in $H^2(\Omega)$ and that it satisfies

$$\|w\|_{H^2(\Omega)} \leq C\|\mathcal{V}\|_{W^{2,4}(\Omega)}\|g\|_{H^3(\Omega)} . \tag{3.40}$$

Then by (3.39), (3.18), (3.24), (3.6), and (3.21),

$$\left|\int_\Omega \nabla(\nabla g \cdot \mathcal{V}) \cdot \nabla(p - p_h)\, d\mathbf{x}\right|$$
$$= \int_\Omega \nabla w \cdot \nabla(p - p_h) + w(p - p_h)\, d\mathbf{x}$$
$$\leq \|p - p_h\|_{H^1(\Omega)}\|w - \mathcal{I}_h w\|_{H^1(\Omega)} + \|\mathcal{I}_h w\|_{L^2(\Omega)}\|j\|_{C^{0,1}(I)}\|u - u_h\|_{L^2(\Omega)},$$
$$\leq Ch^2\|\mathcal{V}\|_{W^{2,4}(\Omega)}\|g\|_{H^3(\Omega)}\left(\|p\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)}\|j\|_{C^{0,1}(I)}\right) .$$

$\square$

Remark 3.2 The shape gradient formula (2.9) clearly represents a linear continuous operator on $W^{1,\infty}(\mathbb{R}^d)$. Nevertheless, to exploit finite element superconvergence as in Theorem (3.1), we have to restrict ourselves to vector fields in $W^{2,\infty}(\mathbb{R}^d)$. If this condition is violated, only first order convergence of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}$ to $d\mathcal{J}(\Omega; \mathcal{V})$ as $h \to 0$ can be shown, because two key duality estimates in the proof of Theorem 3.1 are no longer available.

Remark 3.3 The quadratic rate of convergence in Theorem 3.1 depends on the regularity of the functions u and p. If the assumption on the H^2-regularity of (2.6) is not fulfilled, the provable rate of convergence deteriorates to $O(h^\alpha)$ with fractional $\alpha < 2$, but the formula (3.1) remains meaningful, as long as a weak solutions in $H^1(\Omega)$ exist. On the other hand, if the functions u and p enjoy higher smoothness, the convergence may be improved by increasing the polynomial degree of the finite element space.

Remark 3.4 Theorem 3.1 holds true for Dirichlet boundary conditions only. However, a similar result can be achieved for Neumann boundary conditions. The proof follows the same lines as for the Dirichlet case and relies on H^2-regularity of the state problem and regularity assumptions on the source function f and the boundary data g. In particular, convergence for the boundary term in (2.11) can be conclude either via duality techniques or by continuity of the Dirichlet trace operator with respect to $H^1(\Omega)$.

For Formula (3.2), the following holds:

Theorem 3.2 *Let u_h and p_h be Ritz–Galerkin linear Lagrangian finite element approximations of the solutions u and p of (2.6) and (2.7). In addition to the hypothesis of Theorem* 3.1, *let us assume that*

$$\|u\|_{W^{2,p}(\Omega)} \leq C\|f\|_{L^p(\Omega)} \tag{3.41}$$

for some $p > d$, where d is the space dimension. Then

$$|d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}| \leq Ch\|\mathcal{V} \cdot \mathbf{n}\|_{L^\infty(\partial\Omega)},$$

where h stands for the meshwidth, and $C > 0$ does not depend on h.

Proof By the equality $d\mathcal{J}(\Omega; \mathcal{V}) = d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Bdry}}$, we immediately deduce from (3.2)

$$
\begin{aligned}
&|d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}| \\
&\leq \|\mathcal{V} \cdot \mathbf{n}\|_{L^\infty(\Omega)} \int_{\partial\Omega} \left| j(u) - j(u_h) + \frac{\partial p}{\partial \mathbf{n}}\frac{\partial(u-g)}{\partial \mathbf{n}} - \frac{\partial p_h}{\partial \mathbf{n}}\frac{\partial(u_h-g)}{\partial \mathbf{n}} \right| dS.
\end{aligned}
\tag{3.42}
$$

By linearity, and similarly as in (3.8), we find

$$
\begin{aligned}
&\frac{\partial p}{\partial \mathbf{n}}\frac{\partial(u-g)}{\partial \mathbf{n}} - \frac{\partial p_h}{\partial \mathbf{n}}\frac{\partial(u_h-g)}{\partial \mathbf{n}} \\
&= \frac{\partial p}{\partial \mathbf{n}}\frac{\partial u}{\partial \mathbf{n}} - \frac{\partial p_h}{\partial \mathbf{n}}\frac{\partial u_h}{\partial \mathbf{n}} + \frac{\partial p_h}{\partial \mathbf{n}}\frac{\partial g}{\partial \mathbf{n}} - \frac{\partial p}{\partial \mathbf{n}}\frac{\partial g}{\partial \mathbf{n}} \\
&= \frac{\partial p}{\partial \mathbf{n}}\frac{\partial(u-u_h)}{\partial \mathbf{n}} + \frac{\partial(p-p_h)}{\partial \mathbf{n}}\frac{\partial u}{\partial \mathbf{n}} - \frac{\partial(p-p_h)}{\partial \mathbf{n}}\frac{\partial(u-u_h)}{\partial \mathbf{n}} + \frac{\partial(p_h-p)}{\partial \mathbf{n}}\frac{\partial g}{\partial \mathbf{n}}.
\end{aligned}
$$

Therefore, applying the triangle inequality on the right-hand side of (3.42), the estimate of the theorem follows straightforwardly from finite element error estimates in $W^{1,\infty}(\Omega)$:

$$\|u - u_h\|_{W^{1,\infty}(\Omega)} \leq Ch \quad \text{and} \quad \|p - p_h\|_{W^{1,\infty}(\Omega)} \leq Ch,$$

cf. [7, Corollary 8.1.12], which requires the assumption (3.41).   □

Remark 3.5 For $d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\text{Bdry}}$ to be well-defined, the functions u and p must be smoother than merely belonging to $H^1(\Omega)$.

4 Numerical experiments

We numerically study the approximation of the shape gradient for the quadratic shape functional

$$\mathcal{J}(\Omega) = \int_\Omega u^2 \, d\mathbf{x},$$

for $\Omega \subset \mathbb{R}^2$, under the scalar PDE constraint

$$\begin{cases} -\Delta u + u = f & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{4.1}$$

It is challenging to investigate convergence rates in the $C^1(\mathbb{R}^d; \mathbb{R}^d)$ dual norm numerically. Therefore, we consider only an operator norm over a finite dimensional space of vector fields in $\mathcal{P}_{3,3}(\mathbb{R}^2)$, whose components are multivariate product polynomials of degree three. Moreover, the $C^1(\mathbb{R}^d; \mathbb{R}^d)$-norm is replaced with the $H^1(\Omega)$-norm, which is more tractable computationally. The convergence studies are performed monitoring the approximate dual norms

$$\text{err}^{\text{Vol}} := \left(\max_{\mathcal{V} \in \mathcal{P}_{3,3}} \frac{1}{\|\mathcal{V}\|^2_{H^1(\Omega)}} |d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}|^2 \right)^{1/2}$$

and

$$\text{err}^{\text{Bdry}} := \left(\max_{\mathcal{V} \in \mathcal{P}_{3,3}} \frac{1}{\|\mathcal{V}\|^2_{H^1(\Omega)}} |d\mathcal{J}(\Omega; \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}|^2 \right)^{1/2}$$

on different meshes generated through uniform refinement.[6]

To compute the values err^{Vol} and err^{Bdry}, we introduce a basis $\{\mathcal{V}_i\}_{i=1}^m$, $m = 20$, of $\mathcal{P}_{3,3}(\mathbb{R}^2)$, and define the column vectors

$$\mathbf{z}^{\text{Vol}} := \left(d\mathcal{J}(\Omega; \mathcal{V}_i) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V}_i)^{\text{Vol}} \right)_{i=1}^m,$$
$$\mathbf{z}^{\text{Bdry}} := \left(d\mathcal{J}(\Omega; \mathcal{V}_i) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V}_i)^{\text{Bdry}} \right)_{i=1}^m.$$

Let $\mathbf{M}$ be the Gramian matrix of $\{\mathcal{V}_i\}_{i=1}^{20}$ with respect to the $H^1(\Omega)$ inner product, and consider the matrices $\mathbf{A}^{\text{Vol}}$ and $\mathbf{A}^{\text{Bdry}}$ defined by

$$\{\mathbf{A}^{\text{Vol}}\}_{i,j=1}^{20} = \mathbf{z}^{\text{Vol}}(\mathbf{z}^{\text{Vol}})^T \quad \text{and} \quad \{\mathbf{A}^{\text{Bdry}}\}_{i,j=1}^{20} = \mathbf{z}^{\text{Bdry}}(\mathbf{z}^{\text{Bdry}})^T,$$

respectively. Then, err^{Vol} and err^{Bdry} can be obtained as the square roots of the maximal eigenvalues of $\mathbf{M}^{-1}\mathbf{A}^{\text{Vol}}$ and $\mathbf{M}^{-1}\mathbf{A}^{\text{Bdry}}$, which can be computed by

$$(\mathbf{z}^{\text{Vol}})^T \mathbf{M}^{-1} \mathbf{z}^{\text{Vol}} \quad \text{and} \quad (\mathbf{z}^{\text{Bdry}})^T \mathbf{M}^{-1} \mathbf{z}^{\text{Bdry}},$$

respectively.

Although analytical values are in some cases computable, the reference values $d\mathcal{J}(\Omega; \mathcal{V})$ are approximated by evaluating $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}$ on a mesh with an

---

[6] In experiments 1 and 4 we consider domains with curved boundaries. In this case the refined mesh is always adjusted to fit the boundary.
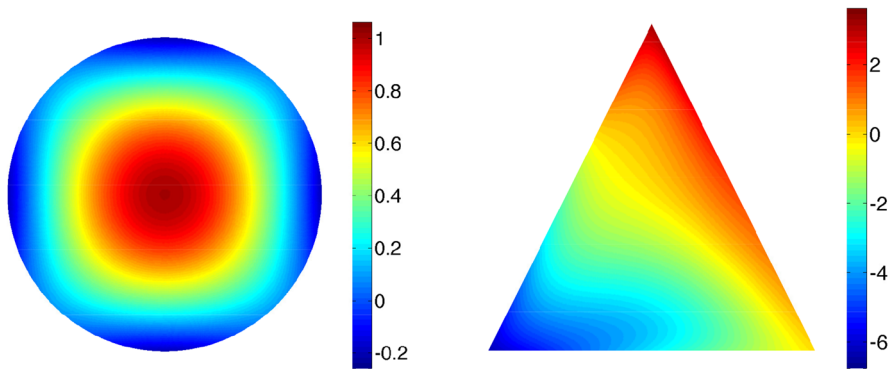
Fig. 1 Plot of the solution *u* of the state problem in the computational domain Ω for the **first** (*left*) and the **second** (*right*) **numerical experiment**

extra level of refinement. This gives us much flexibility in the selection of test cases (the same code can be used for different geometries Ω, source functions f and g, and vector fields $\mathcal{V}$). Agreement with the theoretical predictions of Theorem 3.1 and a numerical study in the third numerical experiment confirm the viability of this approach.

In the implementation, we opt for linear Lagrangian finite elements on quasi-uniform triangular meshes.[7] Integrals in the domain are computed by a 7-point quadrature rule in each triangle and line integrals with a 6-point Gauss quadrature on each segment. The boundary of the computational domains is approximated by a polygon, which is generally believed not to affect the convergence of linear finite elements [7, Sect. 10.2].

The **first numerical experiment** is constructed starting from the solution

$$u(x, y) = \cos(x) \cos(y)$$

and setting f and g accordingly. The computational domain is a disc with radius $\sqrt{\pi}$ (see Fig. 1, left). The predicted quadratic and linear convergence with respect to the meshwidth h for, respectively, Formulas (3.1) and (3.2) are evident in Fig. 2 (left).

The **second experiment** is performed on a triangle with corners located at $(-\pi, -\pi)$, $(\pi, -\pi)$, and $(0, \pi)$ (see Fig. 1, right). The source function and the boundary data are chosen as follows:

$$f(x, y) = x^2 - y^2, \quad g(x, y) = x + y.$$

Again, the rates of convergence predicted in Theorems 3.1 and 3.2 are confirmed by the experiment, see Fig. 2 (right).

The **third numerical experiment** is conducted on a domain which does not guarantee H^2-regularity of the state problem (2.6), see Fig. 3 (left). The source and the boundary functions are, in polar coordinates, $f(\mathbf{x}) = r^{2/3} \cos(2\varphi/3)$ and $g(\mathbf{x}) = 0$ respectively. As expected, the convergence rates deteriorate to fractional values due

---

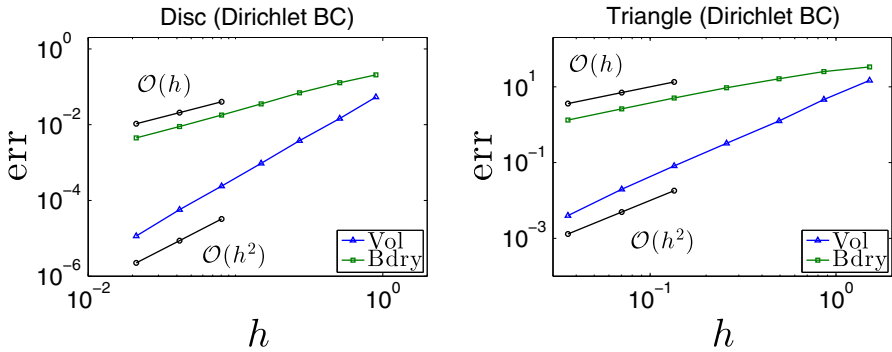[7] The experiments are performed in MATLAB and are based on the library LehrFEM developed at ETHZ.

Fig. 2 Convergence study for the **first** (*left*) and the **second** (*right*) **numerical experiment**. Obviously, Formula (3.1) is better suited for a finite element approximation of the Eulerian derivative $d\mathcal{J}(\Omega; \mathcal{V})$ than Formula (3.2)
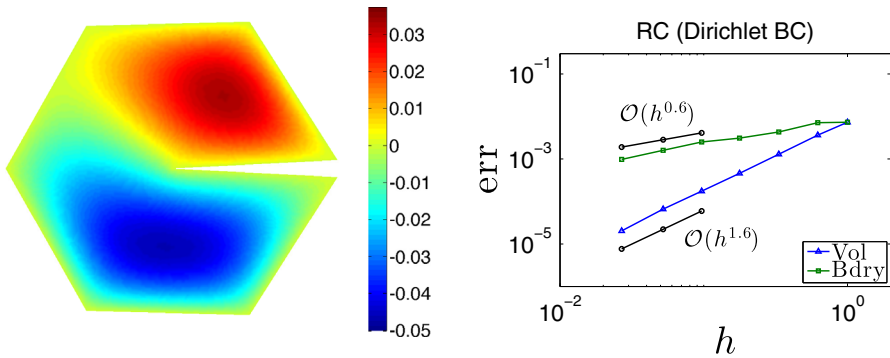

Fig. 3 Plot of the solution u of the state problem in the computational domain Ω (*left*) for the **third numerical experiment**, and corresponding convergence study (*right*). Due to the poor regularity of the functions u and p, the convergence rate of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}$ and $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}$ deteriorate

to the presence of a reentrant corner which, with an interior angle of size $2\pi \cdot 60/61$, affects the regularity of the functions u and p.

In the **fourth numerical experiment**, we investigate the *Neumann problem* and the accuracy of Formulas (2.11) and (2.12), for which we expect results similar to the Dirichlet case. We consider the solution

$$u(x, y) = \cos(x - 1) \cos(y + 1)$$

and we choose f and g accordingly. The computational domain is a disc with radius $\sqrt{\pi}$ (see Fig. 4, left). Surprisingly, we observe that Formula (2.12) performs as well as Formula (2.11), showing quadratic convergence in the meshwidth h, too (see Fig. 5, left).

This surprising observation is not confined to smooth domains, as will be demonstrated by our **fifth numerical experiment**. It investigates the convergence for the *Neumann case* on a triangle with corners located at $(-\pi, -\pi)$, $(\pi, -\pi)$, and $(0, \pi)$
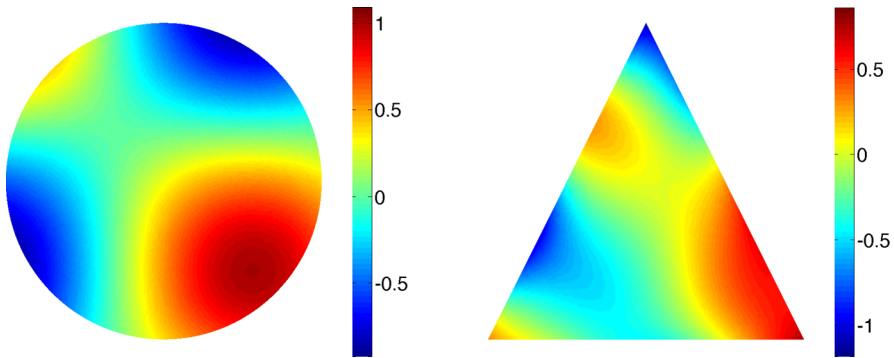
Fig. 4 Plot of the solution u of the state problem in the computational domain Ω for the **fourth** (*left*) and the **fifth** (*right*) **numerical experiment**
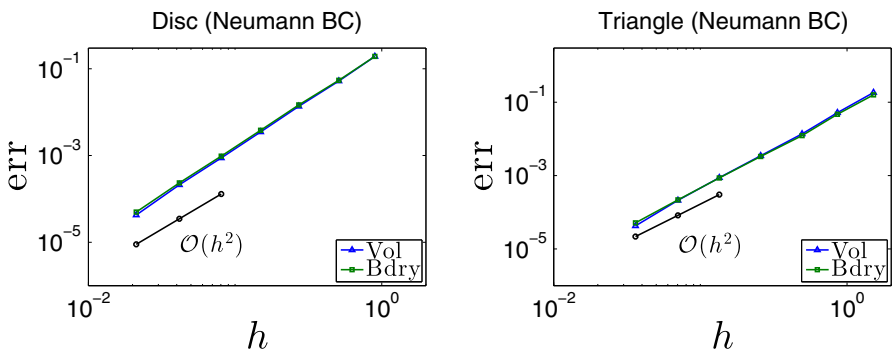

Fig. 5 Convergence study for the **fourth** (*left*) and **fifth** (*right*) **numerical experiment**. The quadratic convergence of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}$ is unexpected

(see Fig. 4, right). The source function and the boundary data are set as follows:

$$f(x, y) = \cos(x + 1)\cos(y - 1), \quad g(x, y) = \cos(x - 1)\cos(y + 1).$$

Again, we observe that Formula (2.12), corrected according to Remark 2.3, converges quadratically in the meshwidth h (see Fig. 5, right).

Nevertheless, the **sixth numerical experiment**, which studies the Neumann boundary value problem again, shows that Formula (2.11) is superior to (2.12) in terms of accuracy and convergence in case of domains which do not guarantee H^2-regularity, see Fig. 6. The source and the boundary functions are chosen as in the third numerical experiment.

Remark 4.1 The superconvergence observed in the fourth and in the fifth numerical experiments may be of interest for practical applications. For instance, in shape optimization it is common to arbitrarily restrict the choice of descent directions to vector fields which vanish on subregions of the computational domain, so that the optimization task is limited to the complement of these subregions [2,18,26]. At the same time,
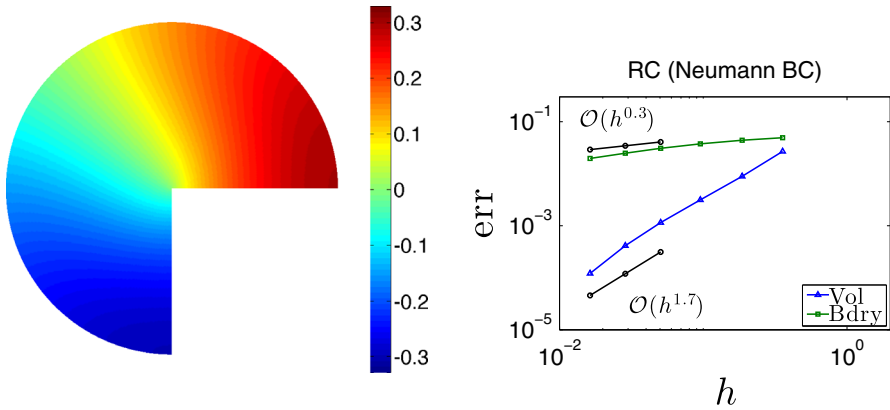
Fig. 6 Plot of the solution u of the state problem in the computational domain Ω (*left*) for the **sixth numerical experiment**, and corresponding convergence study (*right*). Due to the poor regularity of the functions u and p, the convergence rate of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}$ and $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}$ deteriorate

the formation of reentrant corners during the optimization routine is prevented by the use of regularization techniques such as filtering [14,20].

A closer look at Formula (2.12) reveals a cancellation of the normal derivatives of u and p, so that the formula is equivalent to

$$
d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) - \nabla_\Gamma u \nabla_\Gamma p - up + fp + Kgp \right) dS
$$
$$
+ \sum_{i=1}^{3} p(\mathbf{a}_i) g(\mathbf{a}_i) \mathcal{V}(\mathbf{a}_i) \cdot [[\tau(\mathbf{a}_i)]], \tag{4.2}
$$

where ∇_Γ stands for the tangential derivative. To elucidate the behavior of different contributions, we split Formula (4.2) according to

$$
d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) - up + fp + Kgp \right) dS \tag{4.3a}
$$

$$
+ \sum_{i=1}^{3} p(\mathbf{a}_i) g(\mathbf{a}_i) \mathcal{V}(\mathbf{a}_i) \cdot [[\tau(\mathbf{a}_i)]] \tag{4.3b}
$$

$$
- \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(\nabla_\Gamma u \nabla_\Gamma p \right) dS. \tag{4.3c}
$$

An approximation of the first integral (4.3a) by finite elements converges quadratically in h. This can be shown as in the proof of Theorem 3.1, since the Dirichlet trace operator is bounded on $H^1(\Omega)$. Quadratic convergence is also expected for the approximation of (4.3b), due to the convergence properties of finite element solutions in L^∞ [7, Ch.
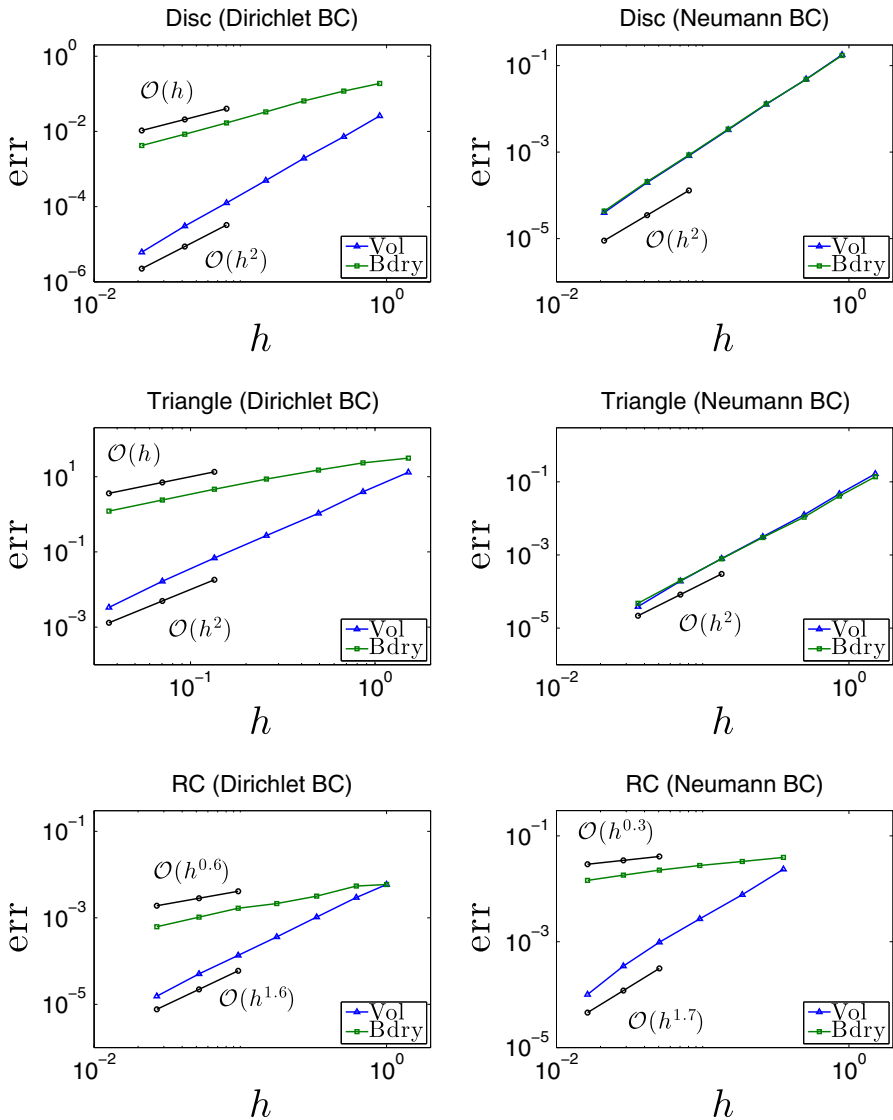
Fig. 7 Convergence study for the **first** (*left*, *up*), the **second** (*left*, *middle*), the **third** (*left*, *down*), the **fourth** (*right*, *up*), the **fifth** (*right*, *middle*) and the **sixth** (*right*, *down*) **numerical experiment**, when considering the operator norm *on the subspace of multivariate polynomials of degree two*. The results agree with those obtained with cubic polynomials

8]. On the other hand, the good approximation of the tangential derivative of *u* and *p* in (4.3c) still defies a theoretical explanation.

Finally, all experiments are repeated considering the operator norm on the subspace of multivariate polynomials of degree two instead of three. The measured errors well agree with those reported above, see Fig. 7. Thus, the arbitrary choice of computing the

operator norm on the finite dimensional subspace of multivariate polynomial vector fields of degree three does not seem to compromise our observations.

5 Conclusion

The shape gradient of shape differentiable PDE constrained shape functionals is an element of the dual space of $C^1(\mathbb{R}^d; \mathbb{R}^d)$, and it can be expressed either as an integration in volume or as an integration on the boundary. Theorems in Sect. 3 and numerical experiments in Sect. 4 confirm that it is advisable to evaluate the shape gradient through volume integrals, when the finite element method is used.

This observation might be of relevance for shape optimization, because, in the words of M. Berggren, "the sensitivity information - directional derivatives of objective functions and constraints - needs to be very accurately computed in order for the optimization algorithms to fully converge" [5]. However, shape optimization techniques usually rely on function representatives of the shape gradient on the boundary. If volume based formulas are used, it takes an extension of boundary deformations into the interior of the domain, in order to obtain those. It remains to be seen whether the superiority of volume based formulas persists under these conditions.

Appendix

Closely following [10, Ch. 10, Sect. 6], we give a detailed derivation of Formulas (2.9)–(2.13). Let u be the weak solution in $H^1(\Omega)$ of the following state problem:

$$
\begin{cases}
-\Delta u + u = f & \text{in } \Omega, \\
u = g & \text{on } \partial\Omega.
\end{cases}
\tag{5.1}
$$

It is assumed that the Dirichlet problem (5.1) is H^2-regular, so that its solution u is at least in $H^2(\Omega)$ for $f \in L^2(\Omega)$. We consider the shape functional

$$
\mathcal{J}(\Omega) = \int_\Omega j(u)\, d\mathbf{x},
$$

and we introduce the Lagrangian

$$
\mathcal{L}(\Omega, v, q, \lambda) := \int_\Omega j(v) + (\Delta v - v + f)q\, d\mathbf{x} + \int_{\partial\Omega} \lambda(g - v)\, dS,
\tag{5.2}
$$

where the functions v, q and λ are in $H^2(\mathbb{R}^d)$. Performing integration by parts, the Lagrangian can be rewritten as

$$
\mathcal{L}(\Omega, v, q, \lambda) = \int_\Omega j(v) - \nabla v \cdot \nabla q - vq + fq\, d\mathbf{x} + \int_{\partial\Omega} \frac{\partial v}{\partial \mathbf{n}} q + \lambda(g - v)\, dS,
$$

$$
= \int_\Omega j(v) + (\Delta q - q)v + fq\, d\mathbf{x} + \int_{\partial\Omega} \frac{\partial v}{\partial \mathbf{n}} q - \frac{\partial q}{\partial \mathbf{n}} v + \lambda(g - v)\, dS.
$$

The saddle point of $\mathcal{L}(\Omega, \cdot, \cdot, \cdot)$ is characterized by

$$\left\langle \frac{\partial \mathcal{L}(\Omega, v, q, \lambda)}{\partial v}, \phi \right\rangle_\Omega = \left\langle \frac{\partial \mathcal{L}(\Omega, v, q, \lambda)}{\partial q}, \phi \right\rangle_\Omega = \left\langle \frac{\partial \mathcal{L}(\Omega, v, q, \lambda)}{\partial \lambda}, \phi \right\rangle_{\partial\Omega} = 0$$

for all $\phi \in H^2(\mathbb{R}^d)$, which, by density, leads to

$$\begin{cases} -\Delta v + v = f & \text{in } \Omega, \\ v = g & \text{on } \partial\Omega, \end{cases} \tag{5.3a}$$

$$\begin{cases} -\Delta q + q = j'(v) & \text{in } \Omega, \\ q = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.3b}$$

$$\lambda = -\frac{\partial q}{\partial \mathbf{n}} \quad \text{on } \partial\Omega, \tag{5.3c}$$

weakly in $H^1(\mathbb{R}^d)$. Thus, for Ω fixed,

$$\mathcal{J}(\Omega) = \inf_{v \in H^2(\mathbb{R}^d)} \sup_{q, \lambda \in H^2(\mathbb{R}^d)} \mathcal{L}(\Omega, v, q, \lambda), \tag{5.4}$$

because

$$\mathcal{J}(\Omega) = \mathcal{L}(\Omega, u, q, \lambda) \quad \forall q, \lambda \in H^2(\mathbb{R}^d).$$

Recall that the material derivative of a generic function f with respect to the deformation $T_\mathcal{V}$ is defined as

$$\dot{f} := \lim_{s \searrow 0} \frac{f \circ T_{s.\mathcal{V}} - f}{s}.$$

Note that, if f is independent of Ω, $\dot{f} \in H^1(\mathbb{R}^d)$ for $f \in H^2(\mathbb{R}^d)$.

To compute the Eulerian derivative of $\mathcal{J}(\Omega)$, the Correa–Seeger theorem can be applied on the right-hand side of (5.4) [10, Ch. 10, Sect. 6.3], so that a formula for $d\mathcal{J}(\Omega)$ can be obtained by evaluating the Eulerian derivative of the Lagrangian (5.2) in its saddle point. For $T_\mathcal{V}(\mathbf{x}) := \mathbf{x} + \mathcal{V}(\mathbf{x})$, the Eulerian derivative of (5.2) reads

$$\lim_{s \searrow 0} \frac{\mathcal{L}(T_{s.\mathcal{V}}(\Omega), v, q, \lambda) - \mathcal{L}(\Omega, v, q, \lambda)}{s}$$

$$= \int_\Omega \Big(j'(v)\dot{v} - \nabla\dot{v} \cdot \nabla q - \nabla v \cdot \nabla\dot{q} + \nabla v \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla q$$

$$- \dot{v} q - v \dot{q} + \dot{f} q + f \dot{q} + \text{div}(\mathcal{V})\left(j(v) - \nabla v \cdot \nabla q - v q + f q\right)\Big)\, d\mathbf{x}$$

$$+ \int_{\partial\Omega} \frac{\partial v}{\partial \mathbf{n}} q + \frac{\partial v}{\partial \mathbf{n}}\dot{q} + \lambda(\dot{g} - \dot{v}) + \dot{\lambda}(g - v) + \text{div}_\Gamma(\mathcal{V})\left(\frac{\partial v}{\partial \mathbf{n}} q + \lambda(g - v)\right) dS$$

$$= \int_\Omega j'(v)\dot{v} + \Delta q\, \dot{v} - q\, \dot{v}\, d\mathbf{x} + \int_\Omega \Delta v\, \dot{q} - v \dot{q} + f \dot{q}\, d\mathbf{x}$$

$$+ \int_{\partial \Omega} \frac{\partial \dot{v}}{\partial \mathbf{n}} q + \dot{\lambda}(g - v) + \mathrm{div}_\Gamma(\mathcal{V}) \left(\frac{\partial v}{\partial \mathbf{n}} q + \lambda(g - v) \right) dS$$

$$+ \int_{\Omega} \nabla v \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla q + \dot{f} q + \mathrm{div}(\mathcal{V}) \left(j(v) - \nabla v \cdot \nabla q - v\, q + f q \right) d\mathbf{x}$$

$$+ \int_{\partial \Omega} \lambda(\dot{g} - \dot{v}) - \frac{\partial q}{\partial \mathbf{n}} \dot{v}\, dS.$$

So, in the saddle point defined by (5.3), we have

$$\lim_{s \searrow 0} \frac{\mathscr{L}(T_{s \cdot \mathcal{V}}(\Omega), v, q, \lambda) - \mathscr{L}(\Omega, v, q, \lambda)}{s} =$$

$$= \int_{\Omega} \nabla v \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla q + \dot{f} q + \mathrm{div}(\mathcal{V}) \left(j(v) - \nabla v \cdot \nabla q - v\, q + f q \right) d\mathbf{x}$$

$$+ \int_{\partial \Omega} -\frac{\partial q}{\partial \mathbf{n}} \dot{g}\, dS$$

$$= \int_{\Omega} \Big(\nabla v \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla q + \dot{f} q + (j'(v) - q)\dot{g} - \nabla q \cdot \nabla \dot{g}$$

$$+ \mathrm{div}(\mathcal{V}) \left(j(v) - \nabla v \cdot \nabla q - v\, q + f q \right) \Big) d\mathbf{x},$$

which, after an additional integration by parts on the term $\dot{f} q = \nabla f \cdot \mathcal{V} q$, corresponds to Formula (2.9). Formula (2.10) is obtained performing additional integrations by parts and exploiting the vector calculus identity

$$\mathcal{V} \cdot \nabla (\nabla v \cdot \nabla q) + \nabla v \cdot \left(\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T \right) \nabla q = \nabla (\mathcal{V} \cdot \nabla v) \cdot \nabla q + \nabla v \cdot \nabla (\mathcal{V} \cdot \nabla v).$$

We refer to [5, Sect. 6] for a detailed derivation. Alternatively, (2.10) can be derived with the so-called "fast derivation" method of Céa, which, formally, does not rely on the concept of material derivative, cf. [9] and [2, Ch. 6.4.3].

Similarly, Formula (2.11) can be derived considering the Lagrangian

$$\mathcal{L}(\Omega, v, q) := \int_{\Omega} j(v) + (\Delta v - v + f)q\, d\mathbf{x} + \int_{\partial \Omega} gq - \frac{\partial v}{\partial \mathbf{n}} q\, dS, \qquad (5.5)$$

$$= \int_{\Omega} j(v) - \nabla v \cdot \nabla q - vq + f q\, d\mathbf{x} + \int_{\partial \Omega} gq\, dS,$$

$$= \int_{\Omega} j(v) + (\Delta q - q)v + f q\, d\mathbf{x} + \int_{\partial \Omega} gq - \frac{\partial q}{\partial \mathbf{n}} v\, dS.$$

Its saddle point is characterized by

$$\begin{cases} -\Delta v + v = f & \text{in } \Omega, \\ \frac{\partial v}{\partial \mathbf{n}} = g & \text{on } \partial \Omega, \end{cases} \qquad (5.6a)$$

$$\begin{cases} -\Delta q + q = j'(v) & \text{in } \Omega, \\ \frac{\partial q}{\partial \mathbf{n}} = 0 & \text{on } \partial \Omega. \end{cases} \qquad (5.6b)$$

Thus, the Eulerian derivative of (5.5) in (5.6) reads

$$
\lim_{s \searrow 0} \frac{\mathscr{L}(T_{s.}\mathcal{V}(\Omega), v, q, \lambda) - \mathscr{L}(\Omega, v, q, \lambda)}{s} =
$$
$$
= \int_{\Omega} \nabla v \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla q + \dot{f}q + \mathrm{div}(\mathcal{V})\,(j(v) - \nabla v \cdot \nabla q - v\,q + f\,q)\,d\mathbf{x}
$$
$$
+ \int_{\partial\Omega} \dot{g}q + \mathrm{div}_{\Gamma}(\mathcal{V})\,(gq)\,dS\,.
$$

In this case, the term $\mathrm{div}_{\Gamma}(\mathcal{V})$ does not vanish, and to recover Formula (2.12) it is necessary to perform an integration by parts on the boundary, from which stems the mean curvature term. For piecewise smooth boundaries, this step has to be performed carefully, because, as in Remark 2.3, additional contributions of corner points appear.

References

1. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
2. Allaire, G.: Conception optimale de structures. Springer, New York (2007)
3. Allaire, G., de Gournay, F., Jouve, F., Toader, A.M.: Structural optimization using topological and shape sensitivity via a level set method. Control Cybernet. **34**(1), 59–80 (2005)
4. Becker, R., Rannacher, R.: An optimal control approach to a posteriori error estimation in finite element methods. Acta Numer. **10**, 1–102 (2001)
5. Berggren, M.: A unified discrete-continuous sensitivity analysis method for shape optimization. In: Applied and Numerical Partial Differential Equations, pp. 25–39. Springer, New York (2010)
6. Braess, D.: Finite Elements. Theory, Fast Solvers, and Applications in Elasticity Theory, 3rd edn. Cambridge University Press, Cambridge (2007)
7. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer, New York (2008)
8. Bucur, D., Buttazzo, G.: Variational Methods in Shape Optimization Problems. Birkhäuser Boston Inc., Basel (2005)
9. Céa, J.: Conception optimale ou identification de formes: calcul rapide de la dérivée directionnelle de la fonction coût. RAIRO Modél. Math. Anal. Numér. **20**(3), 371–402 (1986)
10. Delfour, M.C., Zolésio, J.P.: Shapes and Geometries. Metrics, Analysis, Differential Calculus, and Optimization. 2nd edn. Society for Industrial and Applied Mathematics (SIAM) (2011)
11. Eppler, K.: Boundary integral representations of second derivatives in shape optimization. Discuss. Math. Differ. Incl. Control Optim. **20**(1), 63–78 (2000)
12. Eppler, K.: Second derivatives and sufficient optimality conditions for shape functionals. Control Cybernet. **29**(2), 485–511 (2000)
13. Eppler, K., Harbrecht, H.: Coupling of FEM and BEM in shape optimization. Numer. Math. **104**(1), 47–68 (2006)
14. Firl, M., Wüchner, R., Bletzinger, K.U.: Regularization of shape optimization problems using FE-based parametrization. Struct. Multidiscip. Optim. **47**(4), 507–521 (2013)
15. Gunzburger, M.D.: Perspectives in flow control and optimization. Society for Industrial and Applied Mathematics (SIAM) (2003)
16. Harbrecht, H.: On output functionals of boundary value problems on stochastic domains. Math. Methods Appl. Sci. **33**(1), 91–102 (2010)
17. Harbrecht, H., Tausch, J.: On the numerical solution of a shape optimization problem for the heat equation. SIAM J. Sci. Comput. **35**(1), A104–A121 (2013)
18. Haslinger, J., Mäkinen, R.A.E.: Introduction to Shape Optimization. Theory, Approximation, and Computation. Society for Industrial and Applied Mathematics (SIAM) (2003)
19. Laporte, E., Le Tallec, P.: Numerical Methods in Sensitivity Analysis and Shape Optimization. Birkhäuser Boston Inc., Basel (2003)

20. Le, C., Bruns, T., Tortorelli, D.: A gradient-based, parameter-free approach to shape optimization. Comput. Methods Appl. Mech. Eng. **200**(9–12), 985–996 (2011)
21. McFee, S., Webb, J., Lowther, D.: A tunable volume integration formulation for force calculation in finite-element based computational magnetostatics. IEEE Trans. Magnetics **24**(1), 439–442 (1988)
22. Monk, P.: Finite Element Methods for Maxwell's Equations. Clarendon Press (2003)
23. Monk, P., Süli, E.: The adaptive computation of far-field patterns by a posteriori error estimation of linear functionals. SIAM J. Numer. Anal. **36**(1), 251–274 (1999)
24. Murat, F., Simon, J.: Etude de problèmes d'optimal design. In: Optimization Techniques Modeling and Optimization in the Service of Man Part 2, pp. 54–62. Springer, Berlin (1976)
25. Murat, F., Simon, J.: Sur le contrôle par un domaine géométrique. Internal Report No 76 015, Laboratoire d'Analyse Numérique de l'Université Paris 6 (1976)
26. Pironneau, O.: Optimal Shape Design for Elliptic Systems. Springer, Berlin (1984)
27. Simon, J.: Differentiation with respect to the domain in boundary value problems. Numer. Funct. Anal. Optim. **2**(7–8), 649–687 (1981)
28. Sokołowski, J., Zolésio, J.P.: Introduction to Shape Optimization. Shape Sensitivity Analysis. Springer, New York (1992)
29. Udawalpola, R., Wadbro, E., Berggren, M.: Optimization of a variable mouth acoustic horn. Int. J. Numer. Methods Eng. **85**(5), 591–606 (2011)