

# Approximate Riesz Representatives of Shape Gradients

Alberto Paganini, Ralf Hiptmair

► **To cite this version:**

Alberto Paganini, Ralf Hiptmair. Approximate Riesz Representatives of Shape Gradients. 27th IFIP Conference on System Modeling and Optimization (CSMO), Jun 2015, Sophia Antipolis, France. pp.399-409, 10.1007/978-3-319-55795-3\_38 . hal-01626913

**HAL Id: hal-01626913**

**<https://hal.inria.fr/hal-01626913>**

Submitted on 31 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Approximate Riesz Representatives of Shape Gradients

Alberto Paganini and Ralf Hiptmair

Seminar for Applied Mathematics, ETH Zurich

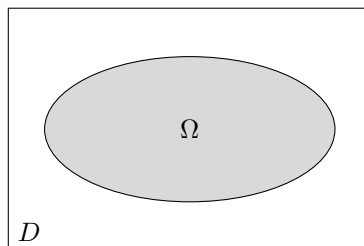
**Abstract.** We study finite element approximations of Riesz representatives of shape gradients. First, we provide a general perspective on its error analysis. Then, we focus on shape functionals constrained by elliptic boundary value problems and  $H^1$ -representatives of shape gradients. We prove linear convergence in the energy norm for linear Lagrangian finite element approximations. This theoretical result is confirmed by several numerical experiments.

**Keywords:** Shape Gradients, Finite Element Approximations.

## 1 Introduction

A *shape functional* is a real map defined on a set of admissible shapes. The goal of *shape optimization* is to modify an initial shape so that a shape functional attains an extremal value. A common approach is to employ steepest descent algorithms [8, Ch. 3.4]. Shapes may be parameterized by  $C^1$ -mappings acting on reference configurations. Then the shape gradient is a linear continuous operator on the non-reflexive Banach space  $C^1$ , and the concept of steepest descent may not be well-defined; see [7, Page 103]. A compromise is to replace “steepest descents” with Riesz representatives of shape gradients with respect to a Hilbert space  $X$ . Henceforth, we refer to these representatives as  *$X$ -representatives*.

After recalling basic definitions of shape calculus, we provide a general perspective on error analysis in the energy norm for finite element approximations of  $X$ -representatives of shape gradients. Then, we zero in on shape functionals constrained to elliptic boundary value problems. For this case, insight into shape Hessians [12, 14] suggests to select representatives of shape gradients with respect to  $X = H_0^1(D)$ , where  $D$  is a hold-all domain that encloses the initial guess  $\Omega$ ; see Figure 1. For the choice  $X = H_0^1(D)$ , it is natural to consider discretization by means of linear Lagrangian finite elements [2, 14, 15]. We show that linear Lagrangian finite element approximations of  $H^1$ -representatives of shape gradients converge linearly with respect to the mesh width. Additionally, this convergence rate does not deteriorate when state and adjoint variables are replaced by linear Lagrangian finite elements solutions. This is an improvement on the result presented in [2], which involves approximations of state and adjoint variables with quadratic finite elements. Finally, we provide numerical evidence of the linear convergence rate predicted.



**Fig. 1.** The hold-all domain  $D$  encloses the domain  $\Omega$ .

## 2 Shape Functionals and 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 shape sensitivity, which quantifies the impact of small perturbations of  $\partial\Omega$  on the value  $\mathcal{J}(\Omega)$ .

We model perturbations of the domain  $\Omega$  through maps of the form

$$T_{\mathcal{V}}(\mathbf{x}) := \mathbf{x} + \mathcal{V}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (1)$$

where  $\mathcal{V}$  is a vector field in  $C^1(\mathbb{R}^d; \mathbb{R}^d)$ . It can easily be proved that the map (1) is a diffeomorphism for  $\|\mathcal{V}\|_{C^1} < 1$  [8, Lemma 6.13].

The value  $\mathcal{J}(\Omega)$  is interpreted as the realization of a *shape functional*, a real map

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

defined on the ball  $\{\mathcal{V} \in C^1(\mathbb{R}^d; \mathbb{R}^d); \|\mathcal{V}\|_{C^1} < 1\}$ . Clearly,  $\mathcal{J}(\Omega) = \mathcal{J}(T_0(\Omega))$ .

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

$$d\mathcal{J}(\Omega; \mathcal{V}) := \lim_{s \searrow 0} \frac{\mathcal{J}(T_{s \cdot \mathcal{V}}(\Omega)) - \mathcal{J}(\Omega)}{s}. \quad (2)$$

We say that the shape functional is *shape differentiable* if Formula (2) defines a linear and bounded operator  $\mathcal{V} \mapsto d\mathcal{J}(\Omega; \mathcal{V})$ . In literature, this operator is called *shape gradient* [9, Ch. 9, Sect. 3.4]. As mentioned in the introduction,  $X$ -representatives of shape gradients can be employed to solve shape optimization problems, that is, to find

$$\Omega^* \in \operatorname{argmin}_{\Omega \in U_{\text{ad}}} \mathcal{J}(\Omega),$$

where  $U_{\text{ad}}$  denotes a set of admissible shapes.

Often, the quantity of interest takes the form

$$\mathcal{J}(\Omega) = \int_B \alpha \nabla(u - g) \cdot \nabla(u - g) + \beta(u - g)^2 \, d\mathbf{x}, \quad (3)$$

where the *state function*  $u$  is the solution of a boundary value problem stated on  $\Omega$ ,  $B \subset \Omega$ ,  $\alpha$  and  $\beta$  are two real constants and  $g$  is a sufficiently smooth target function. In this work,  $u \in H_0^1(\Omega)$  is the (weak) solution of the elliptic boundary value problem with homogeneous Dirichlet boundary conditions

$$-\Delta u + u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{4}$$

that is,

$$\int_{\Omega} \nabla u \cdot \nabla v + uv \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} \quad \text{for all } v \in H_0^1(\Omega), \tag{5}$$

where  $f \in H^1(\Omega)$ . For the sake of brevity, we set  $g = 0$ . Then, the shape gradient of the shape functional associated to (3) and constrained to (5) reads [4, Formula (2.9)]

$$d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\Omega} \left( (\nabla f \cdot \mathcal{V})p + \nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p + \operatorname{div} \mathcal{V}(fp + \chi_B(\alpha \nabla u \cdot \nabla u + \beta u^2) - \nabla u \cdot \nabla p - up) \right) d\mathbf{x}, \tag{6}$$

where the *adjoint function*  $p \in H_0^1(\Omega)$  is the solution of

$$\int_{\Omega} \nabla p \cdot \nabla v + pv \, d\mathbf{x} = \int_B \alpha \nabla u \cdot \nabla v + \beta uv \, d\mathbf{x} \quad \text{for all } v \in H_0^1(\Omega). \tag{7}$$

Formula (3) is a prototypical PDE-constrained shape functional. In this work, Formula (6) is used as test case for proving convergence estimates and performing numerical experiments.

*Remark 1.* Formula (6) holds even if homogeneous Dirichlet boundary conditions in (4) are replaced by homogeneous Neumann boundary conditions, in which case the test and the trial spaces in (5) and (7) are replaced with  $H^1(\Omega)$ .

*Remark 2.* For the sake of simplicity, we restrict our considerations to homogeneous boundary conditions. However, we expect that the results of this work hold true for (sufficiently regular) inhomogeneous boundary conditions, too. Note that Formula (6) should be adjusted accordingly; see [4, Section 2].

### 3 Error Analysis for Finite Element Representatives

#### 3.1 The General Case

Let  $(\cdot, \cdot)_X$  denote the inner product of a Hilbert space  $X$ , and let us assume that the shape gradient  $d\mathcal{J}$  is well-defined on  $X$ . The  $X$ -representative  $\mathcal{V}^X$  of  $d\mathcal{J}$  can be computed by solving

$$(\mathcal{V}^X, \mathcal{W})_X = d\mathcal{J}(\Omega; \mathcal{W}) \quad \text{for all } \mathcal{W} \in X.$$

Next, for an index set  $\mathcal{N}$ , we introduce a family  $\{X_n\}_{n \in \mathcal{N}}$  of finite-dimensional subspaces of  $X$ . Let  $\{\mathcal{V}^{X_n}\}_{n \in \mathcal{N}}$  be a sequence of approximate  $X$ -representatives of  $d\mathcal{J}$  defined by

$$(\mathcal{V}^{X_n}, \mathcal{W}_n)_X = d\mathcal{J}(\Omega; \mathcal{W}_n) \quad \text{for all } \mathcal{W}_n \in X_n.$$

By Cea's Lemma [11, Thm 2.4.1], there exists a constant  $C > 0$  independent of  $n$  such that

$$\|\mathcal{V}^X - \mathcal{V}^{X_n}\|_X \leq C \inf_{\mathcal{W}_n \in X_n} \|\mathcal{V}^X - \mathcal{W}_n\|_X. \quad (8)$$

By and large, the shape gradient of a PDE-constrained shape functional depends also on the state and the adjoint variables  $u$  and  $p$ . These functions are solutions of boundary value problem. Usually, only numerical approximations  $u_h$  and  $p_h$  are available. In that case, the approximate  $X$ -representative  $\mathcal{V}^{X_n}$  has to be replaced with the solution  $\mathcal{V}_h^{X_n}$  of

$$(\mathcal{V}_h^{X_n}, \mathcal{W}_n)_X = d\mathcal{J}_h(\Omega; \mathcal{W}_n) \quad \text{for all } \mathcal{W}_n \in X_n, \quad (9)$$

where  $d\mathcal{J}_h$  is an approximation of the operator  $d\mathcal{J}$  obtained by replacing the functions  $u$  and  $p$  with their numerical approximations  $u_h$  and  $p_h$ .

By Strang Lemma [11, Thm 4.1.1], the estimate (8) should be corrected by adding a consistency term, that is,

$$\|\mathcal{V}^X - \mathcal{V}_h^{X_n}\|_X \leq C \left( \inf_{\mathcal{W}_n \in X_n} \|\mathcal{V}^X - \mathcal{W}_n\|_X + \sup_{\mathcal{W}_n \in X_n} \frac{|d\mathcal{J}(\Omega; \mathcal{W}_n) - d\mathcal{J}_h(\Omega; \mathcal{W}_n)|}{\|\mathcal{W}_n\|_X} \right) \quad (10)$$

for a constant  $C > 0$  independent of  $n$  and  $h$ .

### 3.2 $H^1$ -representatives and Linear Lagrangian Finite Elements

A popular approach in shape optimization consists of replacing the initial domain  $\Omega$  with a polygon/polyhedron equipped with a finite element mesh  $\Omega_h$ . This mesh is used to compute linear Lagrangian finite element approximations of the functions  $u$  and  $p$ . Then, the coordinates of the mesh nodes are (iteratively) updated according to the shape gradient [8, Ch. 6.5]. This is equivalent to extending  $\Omega_h$  to a mesh  $D_h$  that covers a hold-all domain  $D$  and choosing linear Lagrangian finite elements to construct the finite-dimensional subspace  $X_n$ . Formula (10), standard finite element estimates, and Proposition 1 readily imply that, for this discretization, the approximate  $H^1$ -representative of (6) satisfies

$$\|\mathcal{V}^X - \mathcal{V}_h^{X_n}\|_{H^1(D)} = \mathcal{O}(h), \quad (11)$$

which is the main result of this work.

**Proposition 1.** *Let  $\Omega \subset \mathbb{R}^d$  be a polyhedral domain, let  $f \in W^{1,4}(\Omega)$  in (5), and let us assume that the solution  $u$  of (5) satisfies*

$$\|u\|_{W^{2,4}(\Omega)} \leq C\|f\|_{L^4(\Omega)}. \tag{12}$$

Let  $(V_h)_{h \in (0,1]}$  be a family of  $H_0^1(\Omega)$ -conforming piecewise linear Lagrangian finite element spaces built on a quasi-uniform family of simplicial meshes  $(\mathcal{T}^h)_{h \in (0,1]}$ , that is, a family of meshes such that

$$\max\{\text{diam}(T) : T \in \mathcal{T}^h\} \leq h \text{diam}(\Omega)$$

and

$$\min\{\text{diam}(B_T) : T \in \mathcal{T}^h\} \geq \rho h \text{diam}(\Omega) \quad \text{for all } h \in (0,1],$$

for a  $\rho > 0$ , where  $B_T$  is the largest ball contained in the simplex  $T$  [10, Def. 4.4.13]. Let  $u_h, p_h \in V_h$  be solutions of

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h + u_h v_h \, d\mathbf{x} = \int_{\Omega} f v_h \, d\mathbf{x} \quad \text{for all } v_h \in V_h, \tag{13}$$

$$\int_{\Omega} \nabla p_h \cdot \nabla v_h + p_h v_h \, d\mathbf{x} = \int_B \alpha \nabla u_h \cdot \nabla v_h + \beta u_h v_h \, d\mathbf{x} \quad \text{for all } v_h \in V_h, \tag{14}$$

where  $\alpha, \beta \in \mathbb{R}$ ,  $B \subset \Omega$ , and  $\alpha = 0$  or  $B = \Omega$ , if  $d = 3$ . Let  $d\mathcal{J}_h(\Omega; \mathcal{W}_n)$  denote the operator defined by Formula (6) with  $u$  and  $p$  replaced by  $u_h$  and  $p_h$ , respectively. Then,

$$\sup_{\mathcal{W}_n \in X_n} \frac{|d\mathcal{J}(\Omega; \mathcal{W}_n) - d\mathcal{J}_h(\Omega; \mathcal{W}_n)|}{\|\mathcal{W}_n\|_{H^1(D)}} \leq C(\Omega, f, u, p)h \tag{15}$$

for a constant  $C(\Omega, f, u, p) > 0$  independent of  $n$  and  $h$ .

*Proof.* First of all, note that

$$\begin{aligned} d\mathcal{J}(\Omega; \mathcal{W}_n) - d\mathcal{J}_h(\Omega; \mathcal{W}_n) &= \int_{\Omega} (\nabla f \cdot \mathcal{W}_n + f \text{div } \mathcal{W}_n)(p - p_h) \, d\mathbf{x} \\ &\quad + \int_{\Omega} \nabla u \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla p - \nabla u_h \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla p_h \, d\mathbf{x} \\ &\quad + \int_{\Omega} \text{div } \mathcal{W}_n (\nabla u_h \cdot \nabla p_h + u_h p_h - \nabla u \cdot \nabla p - up) \, d\mathbf{x} \\ &\quad + \int_B \text{div } \mathcal{W}_n (\alpha(\nabla u \cdot \nabla u - \nabla u_h \cdot \nabla u_h) + \beta(u^2 - u_h^2)) \, d\mathbf{x}. \end{aligned} \tag{16}$$

We recall that, for generic functions  $q_0 \in L^2(\Omega)$  and  $q_1, q_2 \in L^4(\Omega)$ , the Cauchy-Schwarz inequality implies

$$\|q_0 q_1 q_2\|_{L^1(\Omega)} \leq \|q_0\|_{L^2(\Omega)} \|q_1 q_2\|_{L^2(\Omega)} \leq \|q_0\|_{L^2(\Omega)} \|q_1\|_{L^4(\Omega)} \|q_2\|_{L^4(\Omega)}. \tag{17}$$

Thus, the first integral in (16) may be estimated as follows<sup>1</sup>

$$\left| \int_{\Omega} (\nabla f \cdot \mathcal{W}_n + f \operatorname{div} \mathcal{W}_n)(p - p_h) \, d\mathbf{x} \right| \leq C \|\mathcal{W}_n\|_{H^1(\Omega)} \|f\|_{W^{1,4}(\Omega)} \|p - p_h\|_{L^4(\Omega)}.$$

The second integral in (16) may be estimated as follows

$$\begin{aligned} & \left| \int_{\Omega} \nabla u \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla p - \nabla u_h \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla p_h \, d\mathbf{x} \right| \\ &= \left| \int_{\Omega} \nabla(u - u_h) \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla p + \nabla u_h \cdot (\mathbf{D}\mathcal{W}_n + \mathbf{D}\mathcal{W}_n^T) \nabla(p - p_h) \, d\mathbf{x} \right| \\ &\leq C \|\mathcal{W}_n\|_{H^1(\Omega)} (\|u - u_h\|_{W^{1,4}(\Omega)} \|p\|_{W^{1,4}(\Omega)} + \|u_h\|_{W^{1,4}(\Omega)} \|p - p_h\|_{W^{1,4}(\Omega)}). \end{aligned}$$

The third and the fourth integral in (16) may be estimated similarly.

Stability of the Ritz projection with respect to  $W^{1,4}(\Omega)$  [3]<sup>2</sup>

$$\|u_h\|_{W^{1,4}(\Omega)} \leq C \|u\|_{W^{1,4}(\Omega)} \tag{18}$$

implies  $\|u - u_h\|_{W^{1,4}(\Omega)} = \mathcal{O}(h)$ . To show

$$\|p_h\|_{W^{1,4}(\Omega)} \leq C \|p\|_{W^{1,4}(\Omega)}, \tag{19}$$

which in turn implies  $\|p - p_h\|_{W^{1,4}(\Omega)} = \mathcal{O}(h)$ , it is necessary to repeat the proof of (18) given in [3] tracking the consistency term

$$\int_{\Omega} \nabla(p - p_h) \cdot \nabla g_h^z + (p - p_h) g_h^z \, d\mathbf{x} = \int_B \alpha \nabla(u - u_h) \cdot \nabla g_h^z + \beta(u - u_h) g_h^z \, d\mathbf{x}. \tag{20}$$

The discrete Green's function  $g_h^z \in V_h$  is given in [3] and satisfies  $\|g_h^z\|_{H^1(\Omega)} = \mathcal{O}(h^{-d/2})$ . By the Cauchy-Schwarz inequality and standard finite element estimates,

$$\left| \int_B \alpha \nabla(u - u_h) \cdot \nabla g_h^z + \beta(u - u_h) g_h^z \, d\mathbf{x} \right| = \mathcal{O} \left( (|\alpha|h + |\beta|h^2) h^{-d/2} \right). \tag{21}$$

The stability result (19) holds if (21) is bounded independently of  $h$ . For this reason, we need to set  $\alpha = 0$  when  $d = 3$ , unless  $B = \Omega$ . In this latter case, by Galerkin orthogonality, (20) is bounded by  $\|(\beta - \alpha)(u - u_h) g_h^z\|_{L^1(\Omega)}$ .  $\square$

*Remark 3.* In Proposition 1, we assume  $W^{2,4}$ -regularity of the solution  $u$  of (5). This assumption is made to achieve linear convergence with respect to  $h$  in the estimate (15). However, a three-dimensional polyhedral domain must satisfy tight geometric conditions for  $u$  to be in  $W^{2,4}$  [6, Thm 7.1]. Nevertheless, in [5] the authors show  $W^{1,\infty}$ -stability of the Ritz projection for general convex polyhedral domains. Therefore, we expect that (in the latter case) the right-hand side of (15) can be replaced with a term of order  $\mathcal{O}(h^\alpha)$ , where the rate  $\alpha$  depends on the regularity of  $u$  and satisfies  $0 < \alpha \leq 1$ .

<sup>1</sup> Henceforth,  $C$  denotes a positive generic constant independent of  $n$  and  $h$ .

<sup>2</sup> The assumption  $\Omega \subset \mathbb{R}^2$  made in [3] can be replaced by  $\Omega \subset \mathbb{R}^3$ ; cf. [10, Ch. 8].

*Remark 4.* In [4, 13], the authors show that one can expect superconvergence in the approximation of the shape gradient  $d\mathcal{J}$ . In particular, they show that

$$|d\mathcal{J}(\Omega; \mathcal{W}) - d\mathcal{J}_h(\Omega; \mathcal{W})| \leq C\|\mathcal{W}\|_{W^{2,4}(\Omega)}h^2. \tag{22}$$

However, in the right-hand side of (22) appears the  $W^{2,4}(\Omega)$ -norm of  $\mathcal{W}$ . But to prove convergence in the approximation of a  $H^1$ -representative of  $d\mathcal{J}$ , the upper bound of

$$|d\mathcal{J}(\Omega; \mathcal{W}) - d\mathcal{J}_h(\Omega; \mathcal{W})|$$

cannot involve a norm stronger than the  $H^1$ -norm; see Equation (10).

*Remark 5.* By the *Hadamard structure theorem* [9, Ch. 9, Thm 3.6], most shape gradients admit representatives  $\mathbf{g}(\Omega)$  in the space of distributions  $\mathcal{D}^k(\partial\Omega)$ , that is,

$$d\mathcal{J}(\Omega; \mathcal{V}) = \langle \mathbf{g}(\Omega), \gamma_{\partial\Omega}\mathcal{V} \cdot \mathbf{n} \rangle_{\mathcal{D}^k(\partial\Omega)}, \tag{23}$$

where  $\gamma_{\partial\Omega}\mathcal{V} \cdot \mathbf{n}$  is the normal component of  $\mathcal{V}$  on the boundary  $\partial\Omega$ . For instance, if  $u, p \in H^2(\Omega)$ , Formula (6) is equivalent to [4, Formula (2.10)]

$$d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} (\mathcal{V} \cdot \mathbf{n}) (\alpha \nabla u \cdot \nabla u + \beta u^2 + pu) \, dS. \tag{24}$$

We advise against the use of  $\mathbf{g}(\Omega)$  (which corresponds to the  $L^2(\partial\Omega)$ -representative of  $d\mathcal{J}$ ) to define descent directions because  $L^2$ -representatives might bristle with undesirable oscillations [1].

## 4 Numerical Experiments

We provide numerical evidence of the estimate (11). We employ linear Lagrangian finite elements on quasi-uniform triangular meshes. The experiments are performed in MATLAB and are partly based on the library LehrFEM developed at ETHZ. Mesh generation and uniform refinement are performed with the functions `initmesh` and `refinemesh` of the MATLAB PDE Toolbox [16]. The boundary of computational domains is approximated by a polygon, which is generally believed not to affect the convergence of linear finite elements [10, Sect. 10.2]. For domains with curved boundaries, the refined mesh is always adjusted to fit the boundary. Integrals in the domain are computed with a 3-point quadrature rule of order 3 in each triangle and line integrals with a 6-point Gauss quadrature on each segment.

We consider three different geometries for the domain  $\Omega$  (see Figure 2):

1. A disc of radius  $\sqrt{6/5}$  centered in (0.01,0.02).
2. A triangle with corners located at

$$(-\sqrt{6/5}, -\sqrt{6/5}), (\sqrt{6/5}, -\sqrt{6/5}), (-\sqrt{6/5}, \sqrt{6/5}).$$

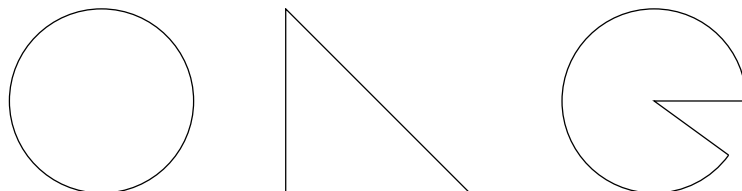
3. A circular sector of radius  $\sqrt{6/5}$  centered in (0.01, 0.02) of angle  $0.9 \cdot 2\pi$ .



The source function in (5) is

$$f(x, y) = \cos(x + \pi/4).$$

The hold-all domain  $D$  is a square with edges of length 3 centered in the origin. The region of interest  $B$  is the whole domain  $\Omega$ . We set  $\alpha = 0$  and  $\beta = 1$  in (3).



**Fig. 2.** The domain  $\Omega$  is chosen to be either a disc or a triangle or a sector.

The reference value  $\mathcal{V}^X$  is approximated by computing  $\mathcal{V}_h^{X_n}$  on a mesh with an extra level of refinement. In light of Remark 5, we employ both Formula (6) and Formula (24) to evaluate the right-hand side  $d\mathcal{J}_h$  in (9). To avoid biased results, we display convergence history of  $\|\mathcal{V}^X - \mathcal{V}_h^{X_n}\|_{H^1(D)}$  both with self- and cross-comparison.

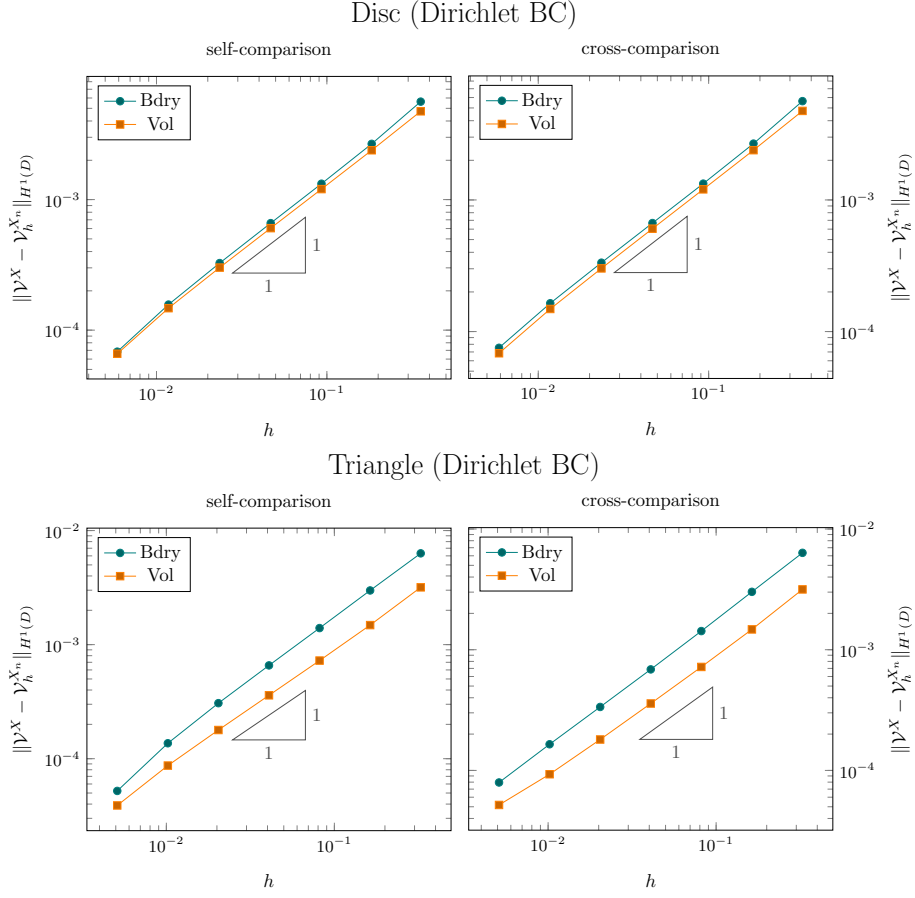
In Figure 3, we plot the convergence history when the domain  $\Omega$  is either a disc (first row) or a triangle (second row). As predicted by (11), we observe linear convergence when the right-hand side in (9) is evaluated according to (6). Interestingly, using Formula (24) seems not to affect the convergence rate. The same behavior is observed when homogeneous Dirichlet boundary conditions are replaced by homogeneous Neumann boundary conditions. Note that, in this latter case, the boundary-integral counterpart of Formula (6) reads [4]

$$d\mathcal{J}(\Omega; \mathcal{V}) = \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} (\nabla u \cdot \nabla(\alpha u - p) + u(\beta u - p) + fp) \, dS. \quad (25)$$

For the sake of brevity, we omit these plots.

In Figure 4 (first row), we plot the convergence history when the domain  $\Omega$  is a sector. This domain does not guarantee that  $u$  and  $p$  are in  $H^2(\Omega)$  because it has a re-entrant corner. We observe that the convergence rates decrease to fractional values. This is a consequence of the lower regularity of the functions  $u$  and  $p$ . Additionally, the convergence rates depend on the formula used to evaluate  $d\mathcal{J}_h$ . In particular, in the cross-comparison, the convergence line saturates when Formula (6) is used. This may be due to a poor accuracy of the reference solution. However, we point out that Formulas (6) and (24) may not be equivalent due to the lack of regularity of the functions  $u$  and  $p$ ; cf. Remark 5. Curiously, for homogeneous Neumann boundary conditions, the presence of the re-entrant corner seems to have a milder impact on convergence rates; see Figure 4 (second row). However, note that the approximate algebraic convergence rates of

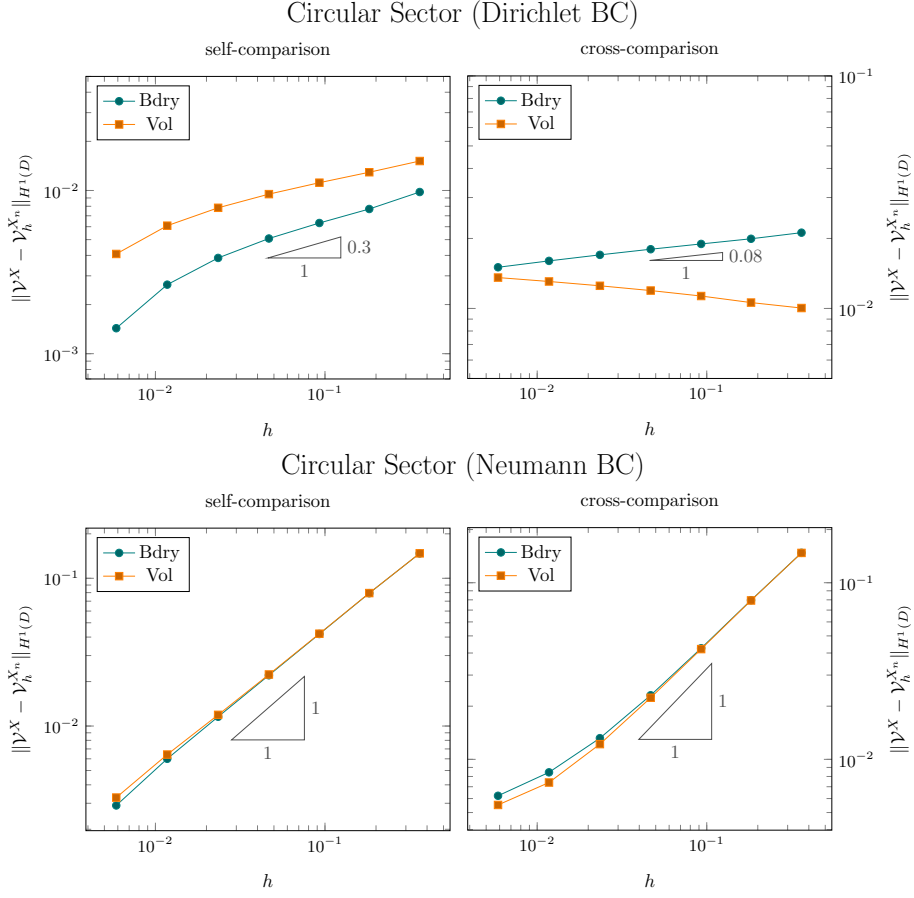
$$\|u - u_h\|_{H^1(\Omega)} \quad \text{and} \quad \|p - p_h\|_{H^1(\Omega)}$$



**Fig. 3.** Convergence history when  $\Omega$  is a disc (*first row*) and a triangle (*second row*). Line  $\text{---}\blacksquare\text{---}$  refers to evaluation of  $d\mathcal{J}_h$  according to Formula (6); line  $\text{---}\bullet\text{---}$  to Formula (24). We observe the linear convergence rate predicted by (11).

with respect to  $h$  drop to 0.67 and 0.62, respectively.

By the Hadamard structure theorem (see Remark 5), vector fields  $\mathcal{W}_n$  associated to interior nodes of the mesh  $\Omega_h$  lie in the kernel of  $d\mathcal{J}$ . However, these vector fields are not in the kernel of  $d\mathcal{J}_h$  because  $u$  and  $p$  are replaced by finite element approximations. Schulz et al. [14] report that this numerical error might largely affect the computation of the Riesz representative. Although we have not experienced this issue, we have repeated the numerical experiments by setting to zero the values of  $d\mathcal{J}_h(\Omega; \mathcal{W}_n)$  for all  $\mathcal{W}_n$  associated to interior nodes of  $\Omega_h$ . We have not observed any significant difference in the results. Thus, we acknowledge that computational resources might be saved by dropping the evaluation of  $d\mathcal{J}_h(\Omega; \mathcal{W}_n)$  for vector fields associated to interior nodes of  $\Omega_h$ .



**Fig. 4.** Convergence history when  $\Omega$  is a sector. Line  $\text{---}\blacksquare\text{---}$  refers to evaluation of  $d\mathcal{J}_h$  according to Formula (6); line  $\text{---}\bullet\text{---}$  to Formula (24) (in the *first row*) and to Formula (25) (in the *second row*). For Dirichlet boundary conditions, convergence rates decay to fractional values.

## 5 Conclusion

Most shape optimization algorithms rely on Riesz representatives of shape gradients with respect to a chosen Hilbert space. Numerical discretization is inevitable when the shape functional is constrained to a boundary value problem. Formula (10) indicates how to estimate the discretization error when the Riesz representative is computed on a finite-dimensional trial space and the shape gradient can be evaluated only approximately.

For linear Lagrangian approximations of  $H^1$ -representatives, Proposition 1 implies that the discretization error decays linearly with respect to the mesh width  $h$ . This convergence behavior is observed in several numerical experiments.

As a consequence of the Hadamard structure theorem, most shape gradients can be equivalently formulated as boundary or volume integrals. Although Proposition 1 relies on the volume formulation of the shape gradient, we have observed linear convergence independently of the formula employed to evaluate  $d\mathcal{J}$ . However, we advise to rely on the volume-based formula because it imposes lower regularity assumptions on the state and the adjoint variables [4, 9, 15].

**Acknowledgments.** The work of A. Paganini was partly supported by ETH Grant CH1-02 11-1.

## References

1. Hiptmair, R., Paganini, A.: Shape optimization by pursuing diffeomorphisms. *Comput. Methods Appl. Math.* 15(3), 291–305 (2015)
2. Murai, D., Azegami, H.: Error analysis of H1 gradient method for shape-optimization problems of continua. *JSIAM Letters* 5, pp. 29–32 (2013)
3. Rannacher, R., Scott, R.: Some optimal error estimates for piecewise linear finite element approximations. *Math. Comp.* 38, 437–445 (1982)
4. Hiptmair, R., Paganini, A., Sargheini, S.: Comparison of approximate shape gradients. *BIT Numerical Mathematics* 55(2), 459–485 (2015)
5. Guzmán, J., Leykekhman, D., Rossmann, J., Schatz, A. H.: Hölder estimates for Green’s functions on convex polyhedral domains and their applications to finite element methods. *Numerische Mathematik* 112(2), 221–243 (2009)
6. Maz’ya, V.G., Romann, J.: Weighted Lp estimates of solutions to boundary value problems for second order elliptic systems in polyhedral domains. *Z. Angew. Math. Mech.* 83(7), 435–467 (2003)
7. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints. Springer, New York (2009)
8. Allaire, G.: Conception optimale de structures. Springer-Verlag, Berlin (2007)
9. Delfour, M.C., Zolésio, J.P.: Shapes and geometries. Metrics, analysis, differential calculus, and optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2011)
10. Brenner, S.C., Ridgway Scott, L.R.: The mathematical theory of finite element methods. Springer, New York (2008)
11. Ciarlet, P.G.: The finite element method for elliptic problems. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2002).
12. Eppler, K., Harbrecht, H.: Shape optimization for free boundary problems—analysis and numerics. In: Constrained optimization and optimal control for partial differential equations, volume 160 of *Internat. Ser. Numer. Math.*, pp 277–288. Birkhäuser/Springer, Basel (2012)
13. Paganini, A.: Approximate shape gradients for interface problems. In: *New Trends in Shape Optimization*, volume 166 of *Internat. Ser. Numer. Math.*, pp 217–227. Springer International Publishing (2015)
14. Schulz, V., Siebenborn, M., Welker, K.: A novel Steklov-Poincaré type metric for efficient PDE constrained optimization in shape spaces. <http://arxiv.org/abs/1506.02244v4> (2015)
15. Laurain, A., Sturm, K.: Domain expression of the shape derivative and application to electrical impedance tomography. *WIAS Preprint No. 1863* (2013)

16. MATLAB and Partial Differential Equation Toolbox (R2015a), The MathWorks Inc., Natick, Massachusetts, United States (2015)