ON DISCRETE SHAPE GRADIENTS OF BOUNDARY TYPE FOR PDE-CONSTRAINED SHAPE OPTIMIZATION∗

WEI GONG† AND SHENGFENG ZHU‡

Abstract. Shape gradients have been widely used in numerical shape gradient descent algorithms for shape optimization. The two types of shape gradients, i.e., the distributed one and the boundary type, are equivalent at the continuous level but exhibit different numerical behaviors after finite element discretization. To be more specific, the boundary type shape gradient is more popular in practice due to its concise formulation and convenience to combine with shape optimization algorithms but has lower numerical accuracy. In this paper we provide a simple yet useful boundary correction for the normal derivatives of the state and adjoint equations, motivated by their continuous variational forms, to increase the accuracy and possible effectiveness of the boundary shape gradient in PDE-constrained shape optimization. We consider particularly the state equation with Dirichlet boundary conditions and provide a preliminary error estimate for the correction. Numerical results show that the corrected boundary type shape gradient has comparable accuracy to that of the distributed one. Moreover, we give a theoretical explanation for the comparable numerical accuracy of the boundary type shape gradient with that of the distributed shape gradient for Neumann boundary value problems.

Key words. Shape optimization, shape gradient, boundary formulation, boundary correction, a priori error estimate, finite element

AMS subject classifications. 49Q12, 65D15, 65K10, 65N30

1. Introduction. Shape optimization has wide applications in various fields in computational science and engineering [2, 3, 7, 17, 25, 27]. For optimal shape design of complex systems, numerical methods and techniques are useful to find “approximate” optimal shapes with the help of computer simulations [27]. Gradient-type optimization methods are widely used shape evolution algorithms to iteratively obtain the approximate optimal domain shapes. To perform a gradient-type algorithm in shape optimization, the so-called shape gradient, derived from Eulerian derivative, is indispensable. Eulerian derivatives can be derived by shape sensitivity analysis [17, 28, 34, 41], which is a classic mathematical tool in shape optimization to measure the variations of an objective functional (called “shape functional” in shape optimization) with respect to shape variations of some domain.

Hadamard-Zol´esio structure theorem [17] shows that the shape gradient for a general shape functional can be obtained by computing the Eulerian derivative in the form of a boundary integral assuming that the domain is smooth enough. Most existing research works on numerical algorithms for shape optimization rely on this structure theorem by using the boundary formulation, due to its attractive concise representation, see e.g., [11, 16, 25]. On the discrete level this type of Eulerian derivative should be discretized by, e.g., finite element method, which requires the according discretization of the state and (possible) adjoint partial differential equations (PDEs).

Finite element methods [10] are among the most popular approaches for discretizing PDEs in shape optimization (see e.g., [26]). One main reason is that the finite element method can solve PDEs on arbitrary domains and thus is flexible to domain changes in shape optimization. The accuracy of approximate shape gradients could be essential for the implementation of numerical optimization algorithms as stated in [17]. It was further pointed out in [5] that the sensitivity information needs to be very accurately computed in order for the optimization algorithms to fully converge. Recently, it is found that the finite element approximations of shape gradients in boundary formulations of Eulerian derivatives have unsatisfactory accuracy and low convergence rate with respect to the mesh size compared to those in volume formulations, see e.g., [29] for elliptic problems, [45] for eigenvalue problems, and [44] for Stokes flows. Numerical investigations even show that the approximate shape gradient flow computed with the traditional boundary formulation may fail to converge to the optimal target domain [33, 43], possibly because the poor accuracy of the approximate shape gradients does not ensure the descent property of the gradient at the discrete level. When it comes to using distributed shape gradients, i.e., the shape gradients associated with volume integrals of Eulerian derivatives in the context of numerical algorithms, we also refer to [9, 32, 40].

Aiming to enhance the accuracy of the finite element approximations of boundary type shape gradients, we propose in this paper a simple yet useful boundary correction for the normal derivatives of the state and

∗The work was supported in part by the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDB 41000000), the National Key Basic Research Program (Grant No. 2018YFB0704304), the National Natural Science Foundation of China (Grants 11671281 and 12071149), the Science and Technology Commission of Shanghai Municipality (No. 18dz2271000), the Natural Science Foundation of Shanghai (Grant No. 19ZR1411400), and the Fundamental Research Funds for the Central Universities.

†The State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics & National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100190 Beijing, China. E-mail address: wgong@lsec.cc.ac.cn.

‡School of Mathematical Sciences & Shanghai Key Laboratory of Pure Mathematics and Mathematical Practice, East China Normal University, Shanghai, 200241, China. E-mail address: sfzhms@math.ecnu.edu.cn.
adjoint equations with Dirichlet boundary conditions, motivated by their continuous variational forms. The cost of this correction requires only to inverse a boundary mass matrix for the state and adjoint state variables, respectively, and thus is negligible compared to the cost of shape gradient descent algorithm. We remark that this kind of boundary equation is extensively used to approximate Dirichlet boundary control of PDEs, see e.g., [4, 12, 15, 21, 22]. We present preliminary a priori error estimates for the finite element approximations of shape gradients with such boundary correction. The second part of this section is devoted to an error analysis for the approximate shape gradient of boundary type for distributed type shape gradients exhibit similar numerical accuracy for Neumann boundary value problems. We propose a boundary correction for the normal derivative of this kind of boundary equation is extensively used to approximate Dirichlet boundary control of PDEs, see e.g., 

\[ J(\Omega) = \int \Omega j(u)dx, \]  

(2.1)

where \( j : \mathbb{R} \to \mathbb{R} \) is \( C^1 \) and we assume further that the derivative \( j' \) is locally Lipschitz continuous (i.e., \( j' \in C^{0,1}(I) \) for any compact set \( I \subset \mathbb{R} \) ), \( u \) is the solution of the following linear elliptic state equation with either Dirichlet or Neumann boundary condition

\[ \begin{cases} -\Delta u + u = f & \text{in } \Omega, \\ u = g \text{ or } \partial_n u = g & \text{on } \Gamma := \partial \Omega, \end{cases} \]  

(2.2)

where the functions \( f \) and \( g \) are assumed to be smooth enough which will be specified later, and they are identified with their restrictions onto \( \Omega \) and \( \partial \Omega \). Here \( \partial_n \) denotes the normal derivative operator with \( n \) being the unit outward normal on the boundary.

Throughout the paper, we make the following assumptions on the data \( f \) and \( g \), the function \( j(u) \) and the domain \( \Omega \).

**Assumption 2.1.**

- \( \Omega \subset \mathbb{R}^n \) is a bounded domain with a \( C^3 \) boundary \( \partial \Omega \).
- \( f \in W^{1,r}(\mathbb{R}^n) \) \( (r > n) \).
- \( g \in W^{3,r}(\mathbb{R}^n) \) \( (r > n) \) for the Dirichlet case and \( g \in H^2(\mathbb{R}^n) \) for the Neumann case.
- \( j(u) \) is \( C^1 \) with its derivative \( j' \) being locally Lipschitz continuous, i.e., \( j' \in C^{0,1}(I) \) for any compact set \( I \subset \mathbb{R} \).

Let us briefly introduce some basic ingredients of the velocity method [17, 41] for shape calculus. Let \( D \) \( (\Omega \subset D) \) be a hold-all domain in \( \mathbb{R}^n \) whose boundary \( \partial D \) is piecewise \( C^1 \). Denote

\[ V^1(D) = \{ \mathbf{V} \in D^1(\mathbb{R}^n; \mathbb{R}^n) \mid \mathbf{V}(\bar{x}) = 0 \text{ on } \partial D \text{ except for the singular points } \hat{x} \text{ of } \partial D, \]  

\[ V(\hat{x}) = 0 \text{ for all singular points } \hat{x} \}. \]

For a variable \( t \in [0, \tau] \) with \( \tau > 0 \), let a vector field \( \mathbf{V} \in C(0, \tau; V^1(D)) \) and \( T_t(\mathbf{V}) \) be the associated one-to-one transformation from \( D \) onto \( \hat{D} \). Let \( x = x(t, X) \) denote the solution to the system of ordinary differential equations

\[ \frac{dx}{dt}(t, X) = \mathbf{V}(t, x(t, X)), \quad x(0, X) = X. \]  

(2.3)

Denote \( \Omega_t = T_t(\mathbf{V})(\Omega) \). The Eulerian derivative of \( J(\Omega) \) at \( \Omega \) in the direction \( \mathbf{V} \) can be expressed as the following limit

\[ dJ(\Omega; \mathbf{V}) := \lim_{t \searrow 0} \frac{J(\Omega_t) - J(\Omega)}{t}. \]  

(2.4)
The Eulerian derivative $dJ(\Omega, \mathbf{V})$ can be formulated as an integral over the volume $\Omega$ for the so-called distributed shape gradient, as well as an integral on the boundary for the so-called boundary type shape gradient. An explicit formula for the Eulerian derivative of the shape functional $J(\Omega)$ relies on the solution $p$ of the following adjoint problem

\[
\begin{aligned}
-\Delta p + p &= j'(u) \quad \text{in } \Omega, \\
p &= 0 \text{ or } \partial_n p = 0 \quad \text{on } \Gamma.
\end{aligned}
\]  

(2.5)

We summarize the known results in the following lemma dedicated to the case of Dirichlet boundary conditions.

**Lemma 2.2.** Let $u$ and $p$ be the solutions of the primal equation (2.2) and the adjoint state equation (2.5) with Dirichlet boundary conditions, respectively. There holds

\[
d^D(\Omega, u, p; \mathbf{V}) = \int_\Omega \left( \nabla u \cdot (D\mathbf{V} + D\mathbf{V}^T) \nabla p - f \mathbf{V} \cdot \nabla p + \text{div} \mathbf{V} (j(u) - \nabla u \cdot \nabla p - up) + (j'(u) - p)(\nabla g \cdot \mathbf{V}) - \nabla p \cdot \nabla (\nabla g \cdot \mathbf{V}) \right) dx.
\]  

(2.6)

The Eulerian derivative can be equivalently written as

\[
d^D(\Gamma, u, p; \mathbf{V}) = \int_\Gamma \mathbf{V}_n \left( j(u) + \partial_n \partial_n (u - g) \right) ds,
\]  

(2.7)

where $\mathbf{V}_n := \mathbf{V} \cdot \mathbf{n}$.

**Proof.** For the proof we refer to [29], see also [25] and [27]. \[Q.E.D.\]

The distributed integral (2.6) and the boundary integral (2.7) are equivalent representations of the Eulerian derivative $d^D(\Omega, \mathbf{V})$, and can be derived from each other by means of Gauss’s theorem, see [29] for the detailed derivation. However, the formula (2.7) is more popular in the literature compared to (2.6), possibly because the boundary type admits a representative $g(\Omega)$ in the space of distributions $\mathcal{D}^k(\Gamma)$ according to the structure theorem, i.e., if $\Gamma$ is smooth, there holds

\[
d^D(\Gamma, u, p; \mathbf{V}) = (g(\Omega), \gamma_\Gamma \mathbf{V} \cdot \mathbf{n})_{\mathcal{D}^k(\Gamma)},
\]  

(2.8)

where $\gamma_\Gamma \mathbf{V} \cdot \mathbf{n}$ is the trace for the normal component of $\mathbf{V}$ on the boundary $\Gamma$. This implies that only normal displacements of the boundary have an impact on the value of $J(\Omega)$. This expression comes in handy in many numerical methods, where the values of the velocity field $\mathbf{V}$ on the boundary of the shape are the only ones needed.

The next result is dedicated to the case of Neumann boundary conditions.

**Lemma 2.3.** [29, Remark 2.2] Let $u$ and $p$ be the solutions of the primal equation (2.2) and the adjoint state equation (2.5) with Neumann boundary conditions, respectively. The Eulerian derivatives read

\[
d^N(\Omega, u, p; \mathbf{V}) = \int_\Omega \left( (\nabla f \cdot \mathbf{V}) p + \nabla u \cdot (D\mathbf{V} + D\mathbf{V}^T) \nabla p + \text{div} \mathbf{V} (fp + j(u) - \nabla u \cdot \nabla p - up) \right) dx + \int_\Gamma (\nabla g \cdot \mathbf{V}) p + gp \text{div}_\Gamma \mathbf{V} ds,
\]  

(2.9)

where $\text{div}_\Gamma := \text{div} - \mathbf{n}^T D\mathbf{V} \mathbf{n}$ denotes the tangential divergence on $\Gamma$, and

\[
d^N(\Gamma, u, p; \mathbf{V}) = \int_\Gamma \mathbf{V}_n \left( j(u) - \nabla u \cdot \nabla p - up + fp + \partial_n (gp) + Kgp \right) ds,
\]  

(2.10)

where $K := \text{div} \mathbf{n}$ is the mean curvature of $\Gamma$.

We remark that, by using the fact that

\[
\nabla u = \nabla_\Gamma u + \partial_n u \mathbf{n} = \nabla_\Gamma u + g\mathbf{n} \quad \text{and} \quad \nabla p = \nabla_\Gamma p + \partial_n p \mathbf{n} = \nabla_\Gamma p \quad \text{on } \Gamma,
\]

where one defines the tangential gradient $\nabla_\Gamma u = P\nabla u$ with $P = I - \mathbf{nn}^T$ and $I$ being an identity matrix, we can rewrite formula (2.10) as

\[
d^N(\Gamma, u, p; \mathbf{V}) = \int_\Gamma \mathbf{V}_n \left( j(u) - \nabla_\Gamma u \cdot \nabla_\Gamma p - up + fp + \partial_n (gp) + Kgp \right) ds.
\]  

(2.11)

Below are a few remarks concerning the conditions in Assumption 2.1.

**Remark 2.4.** In this paper we assume that $\partial \Omega$ is $C^3$, the reasons are as follows.

- As noted in Lemma 2.2, the smoothness assumption on $\Omega$ allows the derivation of the shape gradient under boundary formulation that is equivalent to the volume one.
• This assumption allows for higher regularity of the solutions to the state and adjoint equations which is essential for the proof of improved convergence rates of boundary shape gradients.

• As pointed out in [29], for the Neumann case the shape gradient has to be corrected at the corners if the boundary is only piecewise smooth. Taking Ω to be a 2D polygonal domain as example, formula (2.10) has to be corrected by adding the term (cf. [29])

\[ \sum_i p(x_i)g(x_i)\mathbf{v}(x_i) \cdot \left[\nu(x_i)\right], \]

where the \( x_i \) denote the corner points and \([\nu(x_i)]\) is the jump of the tangential unit vector field in the corner \( x_i \) ([41, Ch. 3.8]). On the other hand, no correction has to be made to formula (2.7).

• When the boundary \( \Gamma \) is only piecewise smooth, the Hadamard-Zolésio structure theorem (2.8) does not hold for the Eulerian derivatives (2.7) and (2.10) for respective Dirichlet and Neumann problems.

When \( \Omega \) is only a convex polygonal/polyhedral domain, the boundary correction can still be used, but we can not expect the same accuracy as that of the volume formulation, see Remark 3.11 for more details.

3. Discretized shape gradients for PDE-constrained shape optimization. For function spaces and norms we are going to use the notations of [1, 10]. Now we consider the finite element approximations. Let \( \{\mathcal{T}_h\}_{h>0} \) be a quasi-uniform and shape regular family of triangulations of domains \( \Omega_h \) approximating \( \Omega \). Define \( \Gamma_h := \partial \Omega_h \). For the domain with curved boundary there are two types of finite element approximations, one is relied on the parametric finite elements (cf. [10, Section 10.4]) while the other is based on polygonal boundary approximation (cf. [8, 30]). Here we consider the latter case since \( \Omega_h \neq \Omega \), we assume further that all boundary vertices of \( \Gamma_h \) also lie on \( \Gamma \) and that at most one edge or face of a simplex belongs to \( \Gamma_h \). The latter assumption is used to avoid trivial finite element solutions near the boundary for problems with homogeneous Dirichlet boundary conditions. We denote by \( \langle \cdot, \cdot \rangle_\Gamma \) and \( \langle \cdot, \cdot \rangle_{\Gamma_h} \) the inner products on \( \Gamma \) and \( \Gamma_h \), respectively, associated with the norms \( \|\cdot\|_{L^2(\Gamma)} \) for \( L^2(\Gamma) \) and \( \|\cdot\|_{L^2(\Gamma_h)} \) for \( L^2(\Gamma_h) \) that are defined in the usual way. Let \( \langle \cdot, \cdot \rangle_h \) denote the inner product in \( L^2(\Omega_h) \). Associated with \( \mathcal{T}_h \) we construct the finite element space \( V_h \) which consists of piecewise polynomials of first order such that \( V_h \subset H^1(\Omega_h) \). We denote by \( V_h^0 := V_h \cap H^1_0(\Omega_h) \) and \( V_h(\Gamma_h) \) the restriction of \( V_h \) to the boundary \( \Gamma_h \). The functions in \( V_h^0 \) are globally continuous. In the following, the constants \( C \) appearing at different circumstances may be different, but are independent of \( h \).

To deal with variational problems defined on different domains \( \Omega \) and \( \Omega_h \), we need to introduce some extension operator. We define the usual Sobolev extension \( E : W^{r,p}(\Omega) \to W^{r,p}(\mathbb{R}^n) \), with \( r \geq 0 \), \( 1 \leq p \leq \infty \), satisfying \( Ev|_\Omega = v \) for \( v \in W^{r,p}(\Omega) \) and (cf. [8, 23, 30])

\[ \|Ev\|_{W^{r,p}(\mathbb{R}^n)} \leq C\|v\|_{W^{r,p}(\Omega)}. \]

In the following, we use the same notation for a function \( v \in W^{r,p}(\Omega) \) and its extension to \( \mathbb{R}^n \).

3.1. Dirichlet boundary value problem. In this subsection we consider the state equation (2.2) with Dirichlet boundary condition. The variational formulation of (2.2) is to find \( u \in H^1(\Omega) \) such that \( u|_\Gamma = g \) and

\[ (\nabla u, \nabla v) + (u, v) = (f, v) \quad \forall v \in H^1_0(\Omega). \]

On the other hand, the variational formulation of (2.5) reads: find \( p \in H^1_0(\Omega) \) such that

\[ (\nabla p, \nabla v) + (p, v) = (j'(u), v) \quad \forall v \in H^1_0(\Omega). \]

Now we are in the position to formulate the finite element approximations to the state and adjoint equations. For the state equation it is to find \( u_h \in V_h \) such that \( u_h|_{\Gamma_h} = Q_hg \) where \( Q_h \) is defined later (cf. (3.21)), and

\[ (\nabla u_h, \nabla v_h) + (u_h, v_h)_{\Gamma_h} = (f, v_h)_{\Gamma_h} \quad \forall v_h \in V_h^0. \]

For the discretization of the adjoint state equation it is to find \( p_h \in V_h^0 \) such that

\[ (\nabla p_h, \nabla v_h) + (p_h, v_h)_{\Gamma_h} = (j'(u_h), v_h)_{\Gamma_h} \quad \forall v_h \in V_h^0. \]

We remark that the right hand sides of (3.4) and (3.5) are usually evaluated by quadrature formulas. However, in the current paper we do not consider the quadrature errors but we only focus on the finite element discretization errors. Then we have the error estimates [8, Theorem 1]

\[ \|u - u_h\|_{L^2(\Omega_h)} + h\|u - u_h\|_{H^1(\Omega_h)} \leq Ch^2(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}), \]

\[ \|p - p_h\|_{L^2(\Omega_h)} + h\|p - p_h\|_{H^1(\Omega_h)} \leq Ch^2(\|j'(u)\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|g\|_{H^2(\Omega)}). \]

The following stability estimates also hold

\[ \|u_h\|_{H^1(\Omega_h)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{H^1(\Omega)}), \quad \|p_h\|_{H^1(\Omega_h)} \leq C(\|j'(u)\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|g\|_{H^1(\Omega)}). \]
Moreover, we assume the following $W^{1,\infty}$-norm error estimate

$$
\|u - u_h\|_{W^{1,\infty}(\Omega_h)} + \|p - p_h\|_{W^{1,\infty}(\Omega_h)} \leq C h (\|u\|_{W^{2,\infty}(\Omega)} + \|p\|_{W^{2,\infty}(\Omega)}).
$$

(3.8)

**Remark 3.1.** The $L^2$-norm and $H^1$-norm error estimates (3.6) have been proved in [8, Theorem 1] in the two-dimensional case. However, it is possible to extend the results to the three-dimensional case, because all the necessary geometric errors for the approximation of a three-dimensional curved domain by a polygonal domain have already been proved in [30, 31], see for instance, Section 2.2 of [30]. A different proof for the estimates (3.6) from that of [8] can be found in [15] for both 2D and 3D curved domains.

**Remark 3.2.** The $W^{1,\infty}$-norm error estimate (3.8) for the finite element approximation of the solution to an elliptic equation with Dirichlet boundary condition based on boundary polygonal approximation is not precisely available in the literature. However, we can find in [30] such an error estimate for the corresponding Neumann boundary value problem. We expect that the methods of proof in [30] can be applied to elliptic Dirichlet problem to derive (3.8). This is partially confirmed by a private communication with Professor Tomoya Kemmochi. In the special case that $\Omega$ is convex so that $\Omega_h \subset \Omega$, the equations (2.2) and (2.5) also hold in $\Omega_h$, we can expect that the standard $W^{1,\infty}$-norm error estimate (cf. [10, Chapter 8]) for elliptic Dirichlet problems implies (3.8).

**Remark 3.3.** Since (3.5) is not a direct Galerkin approximation to (3.3), in order to derive the error estimates (3.6) and (3.8) involving $p$ we have to introduce the intermediate discrete quantity $p_h(u) \in V_h^0$ such that

$$
(\nabla p_h(u), \nabla v_h) + (p_h(u), v_h)_h = (j(u), v_h)_h \quad \forall v_h \in V_h^0.
$$

Now $p_h(u)$ is the standard Galerkin discretization of $p$ and the following error estimates hold

$$
\|p - p_h(u)\|_{L^2(\Omega_h)} + h \|p - p_h(u)\|_{H^1(\Omega_h)} \leq C h^2 \|j(u)\|_{L^2(\Omega)}, \\
\|p - p_h(u)\|_{W^{1,\infty}(\Omega_h)} \leq C h \|p\|_{W^{2,\infty}(\Omega)}.
$$

By using the triangle inequality we only need to estimate

$$
\|p_h(u) - p_h\|_{L^2(\Omega_h)} + \|p_h(u) - p\|_{H^1(\Omega_h)} \leq C \|j(u) - j(u_h)\|_{L^2(\Omega_h)} + C \|j(u) - j(u_h)\|_{H^1(\Omega_h)} \leq C \|j(u) - j(u_h)\|_{L^\infty(\Omega_h)} \leq C \|j(u) - j(u_h)\|_{W^{1,\infty}(\Omega_h)},
$$

where we used the stability (3.7) and Theorem 8.5.3 in [10]. Here $I := [\bar{u}, \bar{u}]$, $\bar{u}$ and $\bar{u}$ are chosen such that $-\infty < \bar{u} < \bar{u}$ and $\bar{u}$ are bounded on $I$, we refer to [29, pp. 468, eq. (3.30)] for more details. Combining the above estimates we can derive the error estimates for the adjoint state equation.

With the discrete state $u_h$ and adjoint state $p_h$ at hand, the discretized boundary type Eulerian derivative for the shape optimization problem reads

$$
d_h J^D (\Gamma_h, u_h, p_h; \mathbf{V}) = \int_{\Gamma_h} V_n \left( j(u_h) + \partial_n p_h (\partial_n u_h - \partial_n g) \right) ds.
$$

(3.9)

This approach is extensively used in the literature as it is easy to implement. In [29] the authors compared the numerical performance of two discretized shape gradients, one with a volumetric formulation while the other with a boundary formulation. It is claimed that the volume type discrete shape gradient has better approximation property for state equations with Dirichlet boundary condition compared to the counterpart boundary integrals, which is then confirmed by extensive numerical examples (cf. [29, Section 4]).

**Proposition 3.1.** [35, Theorem 2.2.4 and 2.2.9] Let $u_h$ and $p_h$ be Ritz-Galerkin linear Lagrange finite element approximations of the solutions $u$ and $p$ of (2.2) and (2.5) associated with the Dirichlet boundary conditions. Assume that $f \in H^1(\mathbb{R}^n)$, $g \in H^3(\mathbb{R}^n)$, $\Gamma_h = \Gamma$, and that there holds

$$
\|u\|_{W^{2,q}(\Omega)} \leq C_q \|f\|_{L^\infty(\Omega)} \quad \text{for all } 1 < q < \infty,
$$

where $C_q \sim \frac{1}{q^{-1}}$ for $q \to 1$ and $C_q \sim q$ for $q \to \infty$ (cf. [29, Theorem 9.8 and 9.9]). We remark that the above $W^{2,q}$-elliptic regularity holds if $\partial \Omega \in C^{1,1}$, $f \in L^q(\Omega)$ and $g \in W^{2,q}(\Omega)$ (cf. [23, Theorem 2.4.2.5]). Then

$$
|d_j^D(\Gamma, u, p; \mathbf{V}) - d_h J^D (\Gamma, u_h, p_h; \mathbf{V})| \leq C h |\log h| \|\mathbf{V} \cdot n\|_{L^\infty(\partial \Omega)}.
$$

Moreover, for the discrete shape gradient of the volume type we have

$$
|d_h J^D (\Omega, u, p; \mathbf{V}) - d_h J^D (\Omega, u_h, p_h; \mathbf{V})| \leq C h^2 \|\mathbf{V}\|_{W^{2,\infty}(\Omega)}.
$$

In this paper we propose a modified discrete shape gradient on $\Gamma_h$:

$$
\tilde{d}_h J^D (\Gamma_h, u_h, p_h; \mathbf{V}) = \int_{\Gamma_h} V_n \circ a_h \left( j(u_h) + \partial_n p_h (\partial_n u_h - \partial_n g \circ a_h) \right) ds_h.
$$

(3.10)
where \( a_h \) is defined as in (3.15), \( \partial_h^h u_h \in V_h(\Gamma_h) \) and \( \partial_h^h p_h \in V_h(\Gamma_h) \) solve the following discrete variational problems:

\[
(\partial_h^h u_h, v_h)_{\Gamma_h} = (\nabla u_h, \nabla v_h)_h + (u_h, v_h)_h - (f, v_h)_h \quad \forall v_h \in V_h
\]

and

\[
(\partial_h^h p_h, v_h)_{\Gamma_h} = (\nabla p_h, \nabla v_h)_h + (p_h, v_h)_h - (j'(u_h), v_h)_h \quad \forall v_h \in V_h.
\]  

We remark that an additional work in the modified discrete shape gradient (3.10) is paid to the solution of two boundary equations (3.11) and (3.12), which only requires the solution of linear systems with small numbers of degrees of freedom are required to be solved. Moreover, the evaluation of the right-hand sides in (3.11) and (3.12) is only performed in the neighboring elements of the boundary nodes because of the discrete state equation (3.4) and adjoint equation (3.5).

**Remark 3.4.** Note that in the classical approximation (3.9) one has one order loss for the approximations \( \partial_h u_h \) of \( \partial u \), this is because we use Optimize-then-Discretize approach ([5]) for the optimization problem, i.e., we derive the first order optimality system and the Eulerian derivative of the objective functional at the continuous level. Then we discretize the optimality system and define the discrete shape gradient at the discrete level by inserting the discrete state and adjoint state. While for the modified scheme (3.10), the approximations \( \partial_h^h u_h \) and \( \partial_h^h p_h \) have the same polynomial degree as that of \( u_h \) and \( p_h \), thus these corrections mimic the procedure of Discretize-then-Optimize approach by using discrete counterparts of integration by parts. Moreover, compared to the classical discrete shape gradient (3.9) which is discontinuous when \( P_1 \) Lagrange finite elements are used for \( u_h \) and \( p_h \), our modified formula is continuous on the boundary, and thus can be used for continuous extensions in the shape optimization algorithms (cf. [11]).

**Remark 3.5.** Instead of the boundary polygonal approximation \( \Omega_h \) of \( \Omega \) utilized in the present paper, we can also use the isoparametric finite element method to solve partial differential equations posed on curved domain (cf. [10, Section 10.4]). This is particularly useful if we try to use higher order finite element methods. In this case, we can expect a better boundary approximation \( \Gamma_h \) to \( \Gamma \) than the polygonal approximation, along with a better outward normal approximation \( \mathbf{n}_h \) to \( \mathbf{n} \). Therefore, we can also define the modified shape gradient (3.10) based on the recovered solutions to (3.11) and (3.12) defined on \( \Gamma_h \). The following analysis can be carried over to the isoparametric finite element approximation, but a rigorous justification is still necessary.

In the following we will show in Theorem 3.10 that the modified discrete shape gradient (3.10) has improved convergence rate compared to the classical one (3.9). The proof relies on an improved approximation property of \( \partial_h^h u_h \) to \( \partial u \) and \( \partial_h^h p_h \) to \( \partial p \), presented in Lemma 3.8, compared to that of \( \partial_h u_h \) and \( \partial_h p_h \). Also we need to introduce an auxiliary problem (cf. eq. (3.43)) and use a duality argument to improve the convergence rate, compared to the approach of Theorem 3.2 in [29].

Before giving the main results we present some preliminary results. Firstly, it follows that for each \( u \in H^{2+\varepsilon}(\Omega) \) \((\varepsilon > 0)\) we can define the outward normal trace \( \partial_n u \in L^2(\Gamma) \) for a Lipschitz domain \( \Omega \) (cf. [23, Theorem 1.5.1.2]). Let \( u \in H^1(\Omega) \) be the solution to the elliptic equation (2.2) with \( f \in L^2(\mathbb{R}^n) \) and \( g \in H^2(\mathbb{R}^n) \), we conclude that \( u \in \{ u \in H^1(\Omega) : -\Delta u + u \in L^2(\Omega) \} \subset H^{2+\varepsilon}(\Omega) \) for some \( \varepsilon \in (0, \frac{1}{2}] \) if \( \Omega \) is a Lipschitz polygonal domain (cf. [24], Theorem 2.4.3 for the two-dimensional case and Corollary 2.6.7 for the three-dimensional case). In addition, there holds \( \partial_n u \in H^{\varepsilon}(\Gamma) \) and this quantity satisfies (cf. [6, Theorem 3.2])

\[
(\partial_n u, v)_\Gamma = (\nabla u, \nabla v) + (u, v) - (f, v) \quad \forall v \in H^1(\Omega).
\]  

Under Assumption 2.1, we have for each \( f \in W^{1,r}(\Omega) \), \( g \in W^{3,r}(\Omega) \) \((r > n)\) that \( u \in W^{3,r}(\Omega) \) (cf. [15]), and thus \( u \in W^{2,\infty}(\Omega) \) by the Sobolev embedding theorem (cf. [1]) and \( \partial_n u \in H^{\frac{3}{2}}(\Gamma) \) by the trace theorem (cf. [23, Section 1.5.1]).

Similarly, we have the variational problem for the outward normal derivative of the solution to the adjoint equation

\[
(\partial_n p, v)_\Gamma = (\nabla p, \nabla v) + (p, v) - (j'(u), v) \quad \forall v \in H^1(\Omega).
\]  

The above regularity results also hold for the adjoint problem under the regularity assumption that \( j'(u) \in W^{1,r}(\Omega) \) \((r > n)\).

In the following we present some results on the approximation property between \( \Omega_h \) and \( \Omega \). Note that with the boundary polygonal approximation we have \( \Omega_h \neq \Omega \). We define the projection \( a_h : \Gamma_h \to \Gamma \) as (cf. [14], [8, 13] for 2D and [30, 42] for 3D)

\[
a_h(x) := x + \delta_h(x)\mathbf{n}_h(x) \quad \text{for } x \in e \subset \Gamma_h,
\]  

where \( \mathbf{n}_h \) is the constant normal to \( \Gamma_h \) on \( e \) and \( \delta_h(x) \) (which could be negative) is chosen in such a way that \( a_h(x) \in \Gamma \) (cf. Figure 3.1 for an illustration in 2D). Then it can be shown that \( a_h \) is locally bijective for
sufficiently small $h$ and $\text{dist}(\Gamma, \Gamma_h) \leq Ch^2$, $|\Omega \setminus \Omega_h| \leq Ch^2$, $|\Omega_h \setminus \Omega| \leq Ch^2$ (cf. [8, 30]). Furthermore, for any $v \in L^2(\Gamma_h)$ we have ([15, 13, 31])

$$
\int_{\Gamma_h} v \, ds_h = \int_{\Gamma} v \circ a_h^{-1} \, t_h \, ds, \quad \text{with } t_h = \frac{ds_h}{ds},
$$

(3.16)

where $ds$ and $ds_h$ denote the smooth and discrete surface measures, respectively.

The above derivation implies the following error estimate (cf. [13, eq. (4.3)] for 2D and [30, eq. (2.1)], [31, Section 8] for 3D)

$$
||| \int_{\Gamma} v \, ds - \int_{\Gamma_h} v \circ a_h \, ds_h ||| \leq Ch^2 \int_{\Gamma} |v| \, ds \quad \forall v \in L^1(\Gamma)
$$

(3.17)

and the stability (cf. [8, eq. (2.5)] for 2D and [31, Section 8] for 3D)

$$
\frac{1}{C} ||v||_{L^2(\Gamma)} \leq ||v \circ a_h||_{L^2(\Gamma_h)} \leq C ||v||_{L^2(\Gamma)} \quad \forall v \in L^2(\Gamma).
$$

(3.18)

On the other hand, we associate with $V_h$ an analogous space $\tilde{V}_h$ in $\Omega$. To do so we follow the ideas of [8] where the 2D case was considered, and extend the results to 3D. For each $v_h \in V_h$ we denote by $\tilde{v}_h : \Omega \to \mathbb{R}$ as follows: if $\Omega_e$ is the subset of $\Omega \setminus \Omega_h$ bounded by the boundary face $e \subset T \cap \Gamma_h$ and the curved surface $\tilde{e} \subset \partial \Omega$, $\tilde{v}_h|_{\Omega_e}$ is defined as the linear extension of $v_h$ from $T$. That is, for any $v_h|_T := \sum_{i=1}^{n+1} v_T^i \phi_T^i(x)$, where $\phi_T^i$ denotes the nodal basis function associated with the nodal $x_i$ of $T$ and $v_T^i$ is the nodal value of $v_h$ at $x_i$, we define $\tilde{v}_h|_{\Omega_e} := \sum_{i=1}^{n+1} v_T^i \phi_T^i(x)$ for any $x \in \Omega_e$. Then we define

$$
\tilde{V}_h := \{ \tilde{v}_h : v_h \in V_h \}
$$

and there holds (cf. [8, eq. (2.6)] and [30, Section 2.3])

$$
||\tilde{v}_h||_{H^1(\Omega)} \leq ||v_h||_{H^1(\Omega_h)}.
$$

(3.19)

On $\Gamma$ we can define the discrete piecewise linear finite element space

$$
\tilde{V}_h := \{ v_h \circ a_h^{-1} : v_h \in V_h(\Gamma_h) \}.
$$

With this definition we remark that for $v_h \in V_h$ the identity $\tilde{v}_h|_{\Gamma} = v_h|_{\Gamma_h} \circ a_h^{-1}$ is not necessarily valid, but there holds the following estimate.

**Lemma 3.6.** For each $v_h \in V_h$ let $\tilde{v}_h$ be its linear extension to $\Omega$ defined above. Then there holds

$$
||\tilde{v}_h - v_h \circ a_h^{-1}||_{L^2(\Gamma)} \leq Ch^2 ||\nabla v_h||_{L^2(\Omega_h)}.
$$

(3.20)
Proof. In fact, it follows from (3.18) and Taylor’s expansion that
\[
\|\tilde{v}_h - v_h \circ a_h^{-1}\|_{L^2(\Gamma_h)}^2 \leq C\|\tilde{v}_h \circ a_h - v_h\|_{L^2(\Gamma_h)}^2 \\
= C \sum_{e \subset \Gamma_h, e \subset \partial T} \int_e \left( \sum_{i=1}^{n+1} v_i^T (\phi_i^T (a_h(x)) - \phi_i^T (x)) \right)^2 dx \\
= C \sum_{e \subset \Gamma_h, e \subset \partial T} \int_e \left( \sum_{i=1}^{n+1} v_i^T \nabla \phi_i^T (x) \cdot (a_h(x) - x) \right)^2 dx \\
\leq Ch^4 \sum_{e \subset \Gamma_h, e \subset \partial T} \int_e \left( \sum_{i=1}^{n+1} v_i^T \nabla \phi_i^T \right)^2 dx \\
= Ch^4 \|\nabla v_h\|_{L^2(\Gamma_h)}^2 \\
\leq Ch^3 \int_T \left[ \sum_{i=1}^{n+1} v_i^T \nabla \phi_i^T \right]^2 dx \\
= Ch^3 \|\nabla v_h\|_{L^2(\Gamma_h)}^2,
\]
where we used the facts that \(|a_h(x) - x| \leq Ch^2| (\text{cf. } [8, \text{Section 2}])\) and and
\[
\int_T \left[ \sum_{i=1}^{n+1} v_i^T \nabla \phi_i^T \right]^2 dx \leq Ch^{-1} \int_T \left[ \sum_{i=1}^{n+1} v_i^T \nabla \phi_i^T \right]^2 dx
\]
when \(T\) is an element with face \(e\), because \(\nabla \phi_i^T\) is a constant on each element \(T\).

Now we can define some projection operators. Firstly, we consider the \(L^2\)-projection operator \(Q_h : L^2(\Gamma_h) \rightarrow V_h(\Gamma_h)\) defined by:
\[
\langle Q_h v, v_h \rangle_{\Gamma_h} = \langle v, v_h \rangle_{\Gamma_h} \quad \forall v_h \in V_h(\Gamma_h).
\]
(3.21)
Similarly, we can define the \(L^2\)-projection operator \(\tilde{Q}_h : L^2(\Gamma) \rightarrow \tilde{V}_h\) such that for any \(v \in L^2(\Gamma)\) there holds
\[
\langle v, \tilde{v}_h \rangle_\Gamma = \langle Q_h v, \tilde{v}_h \rangle_\Gamma \quad \forall \tilde{v}_h \in \tilde{V}_h.
\]
(3.22)
According to [8, eq. (2.20)] one has
\[
\| (I - \tilde{Q}_h) v \|_{L^2(\Gamma)} \leq Ch^s |v|_{H^s(\Gamma)}, \quad 0 \leq s \leq 2.
\]
(3.23)
Note that \(Q_h(v \circ a_h) \neq (\tilde{Q}_h v) \circ a_h\) for \(v \in L^2(\Gamma)\), but we have the following approximation result (cf. [8, Lemma 4])
\[
\|Q_h(v \circ a_h) \circ a_h^{-1} - \tilde{Q}_h v\|_{L^2(\Gamma)} \leq C\|Q_h(v \circ a_h) - (\tilde{Q}_h v) \circ a_h\|_{L^2(\Gamma)} \\
\leq Ch^2\|v\|_{L^2(\Gamma)}.
\]
(3.24)

Remark 3.7. The error estimate (3.23) for the \(L^2(\Gamma)\) projection \(\tilde{Q}_h\) is proved in [8] for 2D case. However, with a closer look at the proof we find that only the approximation property of the finite element space \(\tilde{V}_h\) and the stability of the projection \(\tilde{Q}_h\) are used. Therefore, the extension to 3D is possible. For a slightly different definition of the projection \(\tilde{Q}_h\) in both 2D and 3D we refer to [15, eq. (3.7) and (3.8)]. Moreover, (3.24) is also proved in [8, Lemma 4] for 2D, and the proof relies only on the mapping property (3.16) and the fact \(|1 - t_h| \leq Ch^2\) which are also valid in 3D. Therefore, the estimate (3.24) holds also in the 3D case.

Converting the boundary formula (3.10) to \(\Gamma\) we have
\[
\tilde{a}_h, j_D(\Gamma, u_h, p_h; \mathbf{V}) = \int_{\Gamma} \mathbf{V}_n\left( j(u_h \circ a_h^{-1}) + \partial_{n, p_h}^h (\partial_{n, u_h}^h - \partial_{n, g}) \right) t_h ds,
\]
(3.25)
where \(\partial_{n, u_h}^h := \partial_{n, u_h}^h - \partial_{n, a_h}^h\) and \(\partial_{n, p_h}^h := \partial_{n, p_h}^h - \partial_{n, a_h}^h\), \(t_h := \frac{ds}{d\Gamma}\) with \(ds\) and \(d\Gamma\) denoting the smooth and discrete surface measures, respectively (cf. (3.16)). We remark that in the above formula (3.10) we use \(\mathbf{V}_n \circ a_h\) instead of \(\mathbf{V} \cdot \mathbf{n}_h\), because for the latter case we have one order lost when we convert the integral on \(\Gamma_h\) to that on \(\Gamma\), since one can only expect \(|\mathbf{n}_h \circ a_h^{-1} - \mathbf{n}| \leq Ch\) (cf. [13, eq. (4.1)] and [31, Section 8]).

Before proving the main result we first investigate the approximation property of \(\partial_{n, u_h}^h\) to \(\partial_{n, u}\) and \(\partial_{n, p_h}^h\) to \(\partial_{n, p}\).

Lemma 3.8. Let \(\partial_{n, u}\) and \(\partial_{n, p}\) be the continuous outward normal derivatives defined in (3.13) and (3.14), respectively. Let \(\partial_{n, u_h}^h\) and \(\partial_{n, p_h}^h\) be the approximations defined in (3.11) and (3.12), respectively. Assume that
Assumption 2.1 holds. Then we have
\[ \| \partial_n u - \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma)} \leq C h (\| u \|_{W^{3,r}(\Omega)} + \| f \|_{H^{1}(\Omega)}), \]  
(3.26)
for some \( r > n \).

Proof. We first estimate \( \| \partial_n u - \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma)} \). Note that by the triangle inequality

\[ \| \partial_n u - \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma)} \leq \| \partial_n u - \hat{Q}^{h}_{n} \partial_n u \|_{L^2(\Gamma)} + \| \hat{Q}^{h}_{n} \partial_n u - \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma)}. \]
(3.27)

Now the proof can be divided into the following steps.

**Step 1:** For the first term, it follows from the standard projection error estimate (3.23) that

\[ \| \partial_n u - \hat{Q}^{h}_{n} \partial_n u \|_{L^2(\Gamma)} \leq C h \| \partial_n u \|_{H^1(\Omega)} \leq C h \| u \|_{H^{\gamma/2}(\Omega)}. \]
(3.28)

**Step 2:** Now it remains to estimate the second term \( \| \hat{Q}^{h}_{n} \partial_n u - \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma)} \). This step can be further divided into the following several steps.

**Substep 2.1:** To begin with, we first derive the error equation for \( \partial_n u - \tilde{\partial}^{h}_{n} u_h \). For each \( \xi_h \in V_h(\Gamma_h) \) let \( S_h \xi_h \) be an extension of \( \xi_h \) into \( \Omega_h \) such that \( S_h \xi_h \in V_h \), \( S_h \xi_h = \xi_h \) on \( \Gamma_h \) and satisfies

\[ (\nabla S_h \xi_h, \nabla v_h) + (S_h \xi_h, v_h)_h = 0 \quad \forall v_h \in V_h^0. \]
(3.29)

We denote by \( \hat{S}_h \xi_h \in V_h \) the extension of \( \xi_h \) to the interior of the domain \( \Omega_h \) by setting \( \hat{S}_h \xi_h = 0 \) at the interior nodes of the triangulation. Using the norm equivalence in finite-dimensional spaces we have for any \( q \in [1, \infty) \) (cf. [4, eq. (29)] and [21, Lemma 5.3]):

\[ \| S_h \xi_h \|_{L^q(\Omega_h)} + h \| \nabla \hat{S}_h \xi_h \|_{L^q(\Gamma_h)} \leq C h^{\frac{1}{q}} \| \xi_h \|_{L^q(\Gamma_h)}. \]
(3.30)

Note that for any \( \xi_h \in V_h(\Gamma_h) \) there holds

\[ (\tilde{\partial}^{h}_{n} u_h, \xi_h)_{\Gamma_h} = (\nabla u_h, \nabla \xi_h)_{\Gamma_h} + (u_h, \xi_h)_h - (f, \xi_h)_h = (\nabla u_h, \nabla \xi_h)_{\Gamma_h} + (u_h, \xi_h)_h - (f, \xi_h)_h \]
by recalling the definition (3.11).

On the other hand, (3.11) is equivalent to

\[ (\tilde{\partial}^{h}_{n} u_h, v_h)_{\Gamma_h} = (\nabla u_h, \nabla v_h) + (u_h, v_h)_h - (f, v_h)_h + (\tilde{\partial}^{h}_{n} u_h, v_h - v_h \circ a_h^{-1} t_h)_\Gamma \quad \forall v_h \in V_h \]
(3.31)
by using (3.16) and the fact that \( (\tilde{\partial}^{h}_{n} u_h, v_h)_{\Gamma_h} = (\tilde{\partial}^{h}_{n} u_h, v_h - v_h \circ a_h^{-1} t_h)_\Gamma \). Denote by \( v_h = \hat{S}_h \xi_h \) and its linear extension by \( \hat{v}_h = \hat{S}_h \xi_h \). Then we obtain from (3.13) and (3.31) the error equation for \( \partial_n u - \tilde{\partial}^{h}_{n} u_h \):

\[ \begin{align*}
(\partial_n u - \tilde{\partial}^{h}_{n} u_h, \hat{v}_h)_{\Gamma} = & (\nabla u, \nabla \hat{v}_h) + (u, \hat{v}_h) - (f, \hat{v}_h) - (\tilde{\partial}^{h}_{n} u_h, \hat{v}_h - \xi_h \circ a_h^{-1} t_h)_{\Gamma} \\
& - (\nabla u_h, \nabla v_h)_h - (u_h, v_h)_h + (f, v_h)_h
\end{align*} \]
(3.32)

where

\[ M_h (u, \hat{v}_h) = \int_{\Omega \setminus \Gamma_h} (\nabla u \cdot \nabla \hat{v}_h + u \hat{v}_h - f \hat{v}_h) dx - \int_{\Omega \setminus \Gamma} (\nabla u \cdot \nabla v_h + u v_h - f v_h) dx. \]

**Substep 2.2:** Then, we give some estimates for the last two terms on the right-hand side of the above identity (3.32). The remaining term \( (\nabla (u - u_h), \nabla v_h) + (u - u_h, v_h)_h \) will be estimated in Step 3. Using \( |1 - t_h| \leq C h^2 \), (3.18), Lemma 3.6 and (3.30) we have

\[ \| \tilde{\partial}^{h}_{n} u_h - v_h \circ a_h^{-1} t_h \|_{L^2(\Gamma)} \leq \| \tilde{\partial}^{h}_{n} u_h - v_h \circ a_h^{-1} t_h \|_{L^2(\Gamma)} + \| \tilde{\partial}^{h}_{n} u_h - v_h \circ a_h^{-1} (1 - t_h) \|_{L^2(\Gamma)} \]
\[ \leq C \| \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma_h)} + \| v_h \circ a_h^{-1} t_h \|_{L^2(\Gamma_h)} + C h^2 \| v_h \|_{L^2(\Gamma_h)} \]
\[ \leq C \| \tilde{\partial}^{h}_{n} u_h \|_{L^2(\Gamma_h)} + C h^2 \| v_h \|_{L^2(\Gamma_h)}. \]
(3.33)
Now it remains to estimate $M_h(u, \tilde{v}_h)$. Using [8, Lemma 2] or [30, eq. (2.1)] we have for $v \in H^1(\Omega)$
\begin{equation}
\|v\|_{L^2(\Omega\setminus\Omega_h)} \leq C(h\|v\|_{L^2(\Gamma)} + h^2\|\nabla v\|_{L^2(\Omega)}) \leq C(h\|v\|_{H^1(\Omega)}).
\end{equation}
This combined with (3.18), (3.19), Lemma 3.6 implies that for any $v_h \in V_h$,
\begin{align*}
\|\tilde{v}_h\|_{L^2(\Omega\setminus\Omega_h)} &\leq C(h\|\tilde{v}_h\|_{L^2(\Gamma)} + h^2\|\nabla \tilde{v}_h\|_{L^2(\Omega)}) \\
&\leq C(h\|\tilde{v}_h - v_h \circ a_h^{-1}\|_{L^2(\Gamma)} + h\|v_h \circ a_h^{-1}\|_{L^2(\Gamma)} + h^2\|\nabla \tilde{v}_h\|_{L^2(\Omega)}) \\
&\leq C(h\|v_h\|_{L^2(\Gamma_h)} + h^2\|\nabla v_h\|_{L^2(\Omega_h)}).
\end{align*}
Assume that $\tilde{\Omega}\setminus\Omega_h = \sum \tilde{\Omega}_h^\delta$ where $\tilde{\Omega}_h^\delta$ is the curved region bounded by $\Gamma_h^\delta \subset \Gamma_h$ and $\Gamma^\delta \subset \Gamma$. For the readers’ convenience, we first recall some known estimates for functions near the boundary region, presented in [8, eq. (2.9)] for the 2D case:
\begin{equation}
\|\varphi\|^2_{L^2(\tilde{\Omega}_h^\delta)} \leq C\left(h^2\|\varphi\|^2_{L^2(\Gamma^\delta)} + h^4\left|\frac{\partial \varphi}{\partial x}\right|^2_{L^2(\tilde{\Omega}_h^\delta)}\right),
\end{equation}
where to illustrate the idea a special piece $\Gamma_h^\delta$ is chosen as $\Gamma_h^\delta = \{(x_1, x_2) : x_2 = 0, 0 \leq x_1 < C_1 h\}$ so that $\Gamma^\delta = \{(x_1, x_2) : x_2 = \delta_h(x_1) \geq 0, 0 \leq x_1 < C_1 h\}$, see also Lemma 2 and Lemma 3 in [8] for the general case. A similar result is presented in [30, eq. (2.1)] for 2D and 3D cases:
\begin{equation}
\|f\|_{L^2(\Gamma(\delta))} \leq C(\delta^{1/2}\|f\|_{L^2(\Gamma)} + \delta\|\nabla f\|_{L^2(\Gamma(\delta))}), \quad \delta \in [1, \infty],
\end{equation}
where $\Gamma(\delta)$ denotes the tubular neighborhood of $\Gamma$ with distance $\delta$ that is the so-called boundary region. Now we can derive using the above results that
\begin{align*}
\|\nabla \tilde{v}_h\|_{L^2(\tilde{\Omega}_h^\delta)} &\leq C(h\|\nabla \tilde{v}_h\|_{L^2(\Gamma^\delta)} + h^2\|\nabla \tilde{v}_h\|_{L^2(\Omega_h)})
\end{align*}
where we used the fact that $\tilde{v}_h$ is a linear polynomial in $\Omega_h$ so that $\|\nabla \tilde{v}_h\|_{H^1(\Omega_h)} = \|\nabla \tilde{v}_h\|_{L^2(\Omega_h^\delta)}$. That is, for $h$ sufficiently small, the second term on the right-hand side of the above inequality can be absorbed by the left-hand side. Then, we obtain the fact that $\nabla \tilde{v}_h$ in the curved region $\tilde{\Omega}_h^\delta$ is exactly the same piecewise constant vector as $\nabla v_h$ in the boundary element related to $\Gamma_h^\delta$ and (3.18)
\begin{align*}
\|\nabla \tilde{v}_h\|_{L^2(\Omega\setminus\Omega_h)} &\leq Ch\|v_h\|_{L^2(\Gamma_h)} = Ch\left(\sum_i \|\nabla v_h\|^2_{L^2(\Gamma_i)}\right)^{1/2} \\
&\leq Ch\left(\sum_i \|\nabla v_h\|^2_{L^2(\Gamma_i)}\right)^{1/2} \\
&= Ch\|\nabla v_h\|_{L^2(\Gamma_h)} \\
&\leq C\|v_h\|_{L^2(\Omega_h)}.
\end{align*}
where the standard inverse estimate has been used in the last inequality above. Recalling $v_h = \hat{S}_h \xi_h$, the above estimates combined with (3.30) with $q = 2$ imply that
\begin{equation}
\int_{\Omega\setminus\Omega_h} (\nabla u \cdot \nabla \tilde{v}_h + u \tilde{v}_h - f \tilde{v}_h) dx \\
\leq \|\nabla u\|_{L^2(\Omega\setminus\Omega_h)} \|\nabla \tilde{v}_h\|_{L^2(\Omega\setminus\Omega_h)} + \left|\int_{\Omega\setminus\Omega_h} (u - f) \tilde{v}_h dx\right| \leq C(h\|v_h\|_{L^2(\Gamma_h)} + h^2\|\nabla v_h\|_{L^2(\Omega_h)}).
\end{equation}
Similarly, using [8, Lemma 3] or [30, eq. (2.2)] we have for $v \in H^1(\Omega_h)$
\begin{equation}
\|v\|_{L^2(\Omega_h\setminus\Omega)} \leq C(h\|v\|_{L^2(\Gamma_h)} + h^2\|\nabla v\|_{L^2(\Omega_h)}) \leq C(h\|v\|_{H^1(\Omega_h)}).
\end{equation}
This implies that for any $v_h \in V_h \subset H^1(\Omega_h)$,
\begin{align*}
\|v_h\|_{L^2(\Omega_h\setminus\Omega)} &\leq C(h\|v_h\|_{L^2(\Gamma_h)} + h^2\|\nabla v_h\|_{L^2(\Omega_h)}).
\end{align*}
With abuse of notation, we assume again that $\Omega_h\setminus\Omega = \sum \tilde{\Omega}_h^\delta$ where $\tilde{\Omega}_h^\delta$ is the curved region bounded by $\Gamma_h^\delta \subset \Gamma_h$ and $\Gamma^\delta \subset \Gamma$. Similar to the above estimates, it follows from [8, eq. (2.10)] or [30, eq. (2.2)] that
\begin{equation}
\|\nabla v_h\|_{L^2(\tilde{\Omega}_h^\delta)} \leq C(h\|\nabla v_h\|_{L^2(\Gamma_h^\delta)} + h^2\|\nabla v_h\|_{L^2(\Omega_h^\delta)}).
\end{equation}
therefore,
\[ \| \nabla v_h \|_{L^2(\Omega_h \setminus \Omega)} \leq C h \| \nabla v_h \|_{L^2(\Gamma_h)} \leq C \| v_h \|_{L^2(\Gamma_h)}. \]

Recalling again \( v_h = \tilde{S}_h \xi_h \) and using (3.30) with \( q = 2 \), we are now ready to prove that
\[
\int_{\Gamma_h} (\nabla u \cdot \nabla v_h + u v_h - f v_h) \, dx
\leq \| \nabla u \|_{L^2(\Omega_h \setminus \Omega)} \| \nabla v_h \|_{L^2(\Omega_h \setminus \Omega)} + \left| \int_{\Omega_h \setminus \Omega} (u - f) v_h \, dx \right|
\leq C \| \nabla u \|_{W^{1,\infty}(\Omega)} \| \nabla v_h \|_{L^2(\Omega_h \setminus \Omega)} + h^2 \| \nabla v_h \|_{L^2(\Gamma_h)} + h \| \nabla v_h \|_{L^2(\Omega_h \setminus \Omega)}
\leq C h \| u \|_{W^{1,\infty}(\Omega)} \| \nabla v_h \|_{L^2(\Gamma_h)} + C h \| f \|_{H^1(\Omega_h \setminus \Omega)} \| \nabla v_h \|_{L^2(\Gamma_h)} + h \| \nabla v_h \|_{L^2(\Omega_h \setminus \Omega)}
\leq C h \| u \|_{W^{1,\infty}(\Omega)} + \| f \|_{H^1(\Omega)} \| \nabla v_h \|_{L^2(\Gamma_h)}.
\] (3.37)

where we have used (3.1) because \( u \) has been continuously extended from \( \Omega \) to \( \Omega_h \setminus \Omega \). Combining the estimates (3.35) and (3.37) we obtain that
\[ M_h(u, \tilde{v}_h) \leq C h \| u \|_{W^{1,\infty}(\Omega)} \| f \|_{H^1(\Omega)} \| v_h \|_{L^2(\Gamma_h)}. \] (3.38)

Step 3: Finally, we turn to the estimate of the second term in (3.27). Let \( z_h \in V_h \) be the solution to the following discrete variational problem
\[ z_h|_{\Gamma_h} = Q_h (\partial_n u) \circ a_h - \partial_n^h u_h \quad \text{and} \quad (\nabla z_h, \nabla v_h)_h + (z_h, v_h)_h = 0 \quad \forall v_h \in V^0_h. \] (3.39)

That is, \( z_h := S_h \tilde{Q}_h (\partial_n u) \circ a_h - \partial_n^h u_h \). Further, we define \( \bar{z}_h := \tilde{S}_h (\tilde{Q}_h (\partial_n u) \circ a_h - \partial_n^h u_h) \) and \( \tilde{z}_h \) as its linear extension to \( \Omega \).

Using the definition of \( \tilde{Q}_h \) we have
\[ \| \tilde{Q}_h \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)}^2 = (\partial_n u - \partial_n^h u_h, \tilde{Q}_h \partial_n u - \partial_n^h u_h)_{\Gamma} \]
\[ = (\partial_n u - \partial_n^h u_h, \bar{z}_h)_{\Gamma} + (\partial_n u - \partial_n^h u_h, \tilde{Q}_h \partial_n u - \partial_n^h u_h - \bar{z}_h)_{\Gamma} \]
\[ := I_1 + I_2. \]

For the first term \( I_1 \) we use the error equation (3.32) by setting \( \bar{v}_h := \tilde{z}_h \). We conclude from (3.18), (3.33), (3.38), the \( L^2 \) error estimate (3.6) and the \( W^{1,\infty} \) error estimate (3.8) that
\[ (\partial_n u - \partial_n^h u_h, \bar{z}_h)_{\Gamma} = (\nabla(u - u_h), \nabla \tilde{z}_h)_h + (u - u_h, \bar{z}_h)_h - (\partial_n^h u_h, \tilde{z}_h - \bar{z}_h)_{\Gamma} + h \| u - u_h \|_{L^2(\Omega_h \setminus \Omega)} \| \tilde{z}_h \|_{L^2(\Omega_h \setminus \Omega)}
\leq \| \nabla(u - u_h) \|_{L^2(\Omega)} \| \nabla \tilde{z}_h \|_{L^1(\Omega_h \setminus \Omega)} + \| u - u_h \|_{L^2(\Omega_h \setminus \Omega)} \| \tilde{z}_h \|_{L^2(\Omega_h \setminus \Omega)}
\leq C h \| u \|_{W^{2,\infty}(\Omega)} + h^2 \| u \|_{H^2(\Omega)} \| \tilde{Q}_h (\partial_n u) \circ a_h - \partial_n^h u_h \|_{L^2(\Gamma_h)} + C h \| u \|_{W^{2,\infty}(\Omega)} \| f \|_{H^1(\Omega)} \| \tilde{z}_h \|_{L^2(\Gamma_h)}
\leq C h \| u \|_{W^{2,\infty}(\Omega)} + \| f \|_{H^1(\Omega)} \| \tilde{Q}_h \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)}.
\]

where in the second last step we used (3.30) with \( q = 1 \) and the embedding \( L^2(\Gamma_h) \hookrightarrow L^1(\Gamma_h) \). For the second term \( I_2 \) the estimates (3.18), Lemma 3.6 and (3.30) yield
\[ (\partial_n u - \partial_n^h u_h, \tilde{Q}_h \partial_n u - \partial_n^h u_h - \bar{z}_h)_{\Gamma} = (\partial_n u - \partial_n^h u_h, \tilde{z}_h - a_h^{-1} \bar{z}_h)_{\Gamma}
\leq \| \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma)} \| \tilde{z}_h - a_h^{-1} \bar{z}_h \|_{L^2(\Gamma_h)}
\leq C h \| u \|_{W^{2,\infty}(\Omega)} + \| f \|_{H^1(\Omega)} \| \tilde{Q}_h \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)}.
\]

This implies that
\[ \| \tilde{Q}_h \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)} \leq C h \| u \|_{W^{2,\infty}(\Omega)} + \| \partial_n^h u_h \|_{L^2(\Gamma_h)} + \| f \|_{H^1(\Omega)} + \| \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)} \]
\leq C h \| u \|_{W^{2,\infty}(\Omega)} + \| \partial_n^h u_h \|_{L^2(\Gamma_h)} + \| f \|_{H^1(\Omega)} + \| \partial_n u \|_{L^2(\Gamma_h)}. \] (3.40)

Combining (3.28) and (3.40) we complete the estimate of \( \| \partial_n u - \partial_n^h u_h \|_{L^2(\Gamma_h)}, \)
The estimate for \( \| \partial_n p - \overline{\partial_n p_h} \|_{L^2(\Gamma)} \) can be proved in a similar way. For any \( \xi_h \in V_h(\Gamma_h) \) we still denote by \( v_h = \widehat{S}_h \xi_h \) and its linear extension by \( \tilde{v}_h = \widehat{S}_h \xi_h \). We obtain from (3.14) and (3.12) the error equation for \( \overline{\partial_n p_h} - \overline{\partial_n p_h} \):

\[
(\partial_n p - \overline{\partial_n p_h}, \tilde{v}_h)_\Gamma = (\nabla p - \nabla \tilde{v}_h) + (p, \tilde{v}_h) - (j'(u), \tilde{v}_h) - (\nabla p_h - \nabla \tilde{v}_h, \tilde{v}_h - \xi_h \circ a_h^{-1} t_h)_\Gamma
+ (j'(u_h), v_h) + (j'(u), v_h) + (\nabla (p - p_h), \nabla v_h) + (p - p_h, v_h)
- (\overline{\partial_n p_h}, \tilde{v}_h - \xi_h \circ a_h^{-1} t_h)_\Gamma + N_h(p, \tilde{v}_h),
\]

where

\[
N_h(p, \tilde{v}_h) = \int_{\Omega \setminus \Omega_h} (\nabla p \cdot \nabla \tilde{v}_h + p \tilde{v}_h - j'(u) \tilde{v}_h) dx - \int_{\Omega \setminus \Omega_h} (\nabla p \cdot \nabla v_h + p v_h - j'(u) v_h) dx.
\]

A similar error estimate for \( (\overline{\partial_n p_h}, \tilde{v}_h - \xi_h \circ a_h^{-1} t_h)_\Gamma \) and \( N_h(p, \tilde{v}_h) \) can be derived as in (3.33) and (3.38), respectively. The remaining estimates are very similar as above, except for the following term which can be estimated as

\[
(j'(u_h) - j'(u), v_h)_h \leq \| j'(u) - j'(u_h) \|_{L^2(\Omega_h)} \| v_h \|_{L^2(\Omega_h)} \leq C \| j' \|_{C^0(\overline{\Omega})} \| u - u_h \|_{L^2(\Omega_h)} h^\frac{1}{2} \| \xi_h \|_{L^2(\Gamma_h)}
\]

by considering (3.6) and (3.30). Then we finish the proof. \( \square \)

**Remark 3.9.** There is another way to prove Lemma 3.8 by avoiding the \( W^{1,\infty} \)-norm error estimate (3.8) for finite element approximation to elliptic equation with Dirichlet boundary condition, but using instead the \( W^{1,\infty} \)-norm error estimate for Lagrange interpolation (cf. [15, Theorem 4.1]). In this case we use the extension operator \( S_h \) instead of \( \widehat{S}_h \) and define the error equation (3.32) in \( \Omega \).

To illustrate the main idea we follow the notations in [15, Section 3]. For each element \( T \in T_h \) there exists an affine and bijective mapping

\[ F_T : \hat{T} \subset \mathbb{R}^n \rightarrow T \subset \mathbb{R}^n, \quad F_T(\hat{x}) = A T \hat{x} + b_T, \]

for a fixed reference element \( \hat{T} \). We denote by \( \hat{T}_h \) an exact triangulation of \( \Omega \), which can be constructed by replacing the boundary elements of \( T_h \) with curved elements. For each \( T \in T_h \) there exists a mapping \( \Phi_T \in C^1(\hat{T}, \mathbb{R}^n) \) such that \( F_T := F_T \circ \Phi_T \) maps \( \hat{T} \) onto a curved \( n \)-simplex \( T \). Then we can define the mapping \( G_h \) locally by \( G_h|_T := F_T \circ \Phi_T^{-1} \) such that it is a homeomorphism between \( \Omega_h \) and \( \Omega \).

Now for any function \( u_h \in V_h \) we extend it to \( \Omega \), denoting by \( \hat{u}_h \) this extension, by composing \( u_h \) with the mapping function \( G_h : \Omega_h \rightarrow \Omega \) (cf. [15, Section 3]), i.e., \( \hat{u}_h = u_h \circ G_h \). Define \( \hat{V}_h := \{ \hat{v}_h : v_h \in V_h \} \). For any \( \xi_h \in V_h(\Gamma_h) \) we denote by \( v_h = \widehat{S}_h \xi_h \) and by \( \tilde{v}_h \) its extension. Let \( I_h \) be the Lagrange interpolation operator corresponding to \( V_h \), we denote by \( I_h : C^0(\Omega) \rightarrow V_h \) the Lagrange interpolation operator associated with \( \hat{V}_h \), such that \( (I_h \hat{u}) \circ G_h = I_h(u \circ G_h) \) on \( \Gamma_h \) (cf. [15, p. 2802]).

Analogously to the derivation of (3.32), we obtain by using the definition of \( S_h \) that

\[
(\overline{\partial_n u} - \overline{\partial_n u_h}, \tilde{v}_h)_\Gamma = (\nabla u, \nabla \tilde{v}_h) + (u, \tilde{v}_h) - (f, \tilde{v}_h) - (\overline{\partial_n u_h}, \tilde{v}_h - \xi_h \circ G_h^{-1} t_h)_\Gamma
- (\nabla u_h, \nabla v_h) - (u_h, v_h) + (f, v_h)
= (\nabla u - \nabla \tilde{v}_h, \nabla v_h) - (\nabla I_h u_h, \tilde{v}_h - \xi_h \circ G_h^{-1} t_h)_\Gamma
- (\nabla I_h u, \nabla v_h) - (I_h u_h, v_h) + (f, v_h)
= (\nabla (u - \tilde{v}_h), \nabla v_h) + (u - \tilde{v}_h, v_h)
- (\overline{\partial_n u_h}, \tilde{v}_h - \xi_h \circ G_h^{-1} t_h)_\Gamma + M_h(I_h u, \tilde{v}_h),
\]

where

\[
M_h(I_h u, \tilde{v}_h) = \int_{\Omega \setminus \Omega_h} (\nabla I_h u \cdot \nabla \tilde{v}_h + I_h u \tilde{v}_h - f \tilde{v}_h) dx - \int_{\Omega_h \setminus \Omega} (\nabla I_h u \cdot \nabla v_h + I_h u v_h - f v_h) dx.
\]

The last two terms can be estimated in the conventional way, cf. [15, eq. (3.4)] for the estimate of \( M_h(I_h u, \tilde{v}_h) \) and (3.33) for the estimate of the second last term. Following the arguments of [15, eq. (4.9)] and using the \( W^{1,\infty} \)-norm error estimate for the Lagrange interpolation \( I_h u \), the main difference is the following estimate
Let the variational formulation be to find \( u_h \) such that
\[
\langle \nabla (u - I_h u), \nabla v_h \rangle + \langle (u - I_h u), \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} v_h \rangle \leq C \left( \int_{\Omega} \omega^{-1} |\nabla (u - I_h u)|^2 \right)^{\frac{1}{2}} \left( \int_{\Omega} \omega |\nabla \tilde{v}_h|^2 \right)^{\frac{1}{2}} + \|u - I_h u\|_{L^2(\Omega)} \|\tilde{v}_h\|_{L^2(\Omega)}
\]
where we used [8, Lemma 1] or [30, eq. (2.1)] in the last step to estimate the errors. Assume that Assumption 2.1 holds. Then we have
\[
\|\tilde{v}_h\|_{L^2(\Gamma)} \leq C \sqrt{\log h} \|\tilde{v}_h\|_{L^2(\Gamma)},
\]
where \( \omega \) is defined as in [15, Lemma 3.2] and we used the following estimate in [15, Lemma 3.2]
\[
\int_{\Omega} (|\tilde{v}_h|^2 + \omega |\nabla \tilde{v}_h|^2) \leq C \|\xi_h\|_{L^2(\Gamma)}.
\]
The above error estimate could be used in the proof of Lemma 3.8 with \( v_h \) replaced by \( z_h \) defined in (3.39). Now we are ready to prove our main result in this subsection.

**Theorem 3.10.** Let \( u \) and \( p \) be the continuous solutions to the state equation (2.2) and the adjoint equation (2.5) with Dirichlet boundary conditions, while \( u_h \) and \( p_h \) denote the discrete solutions to (3.4) and (3.5), respectively. Assume that Assumption 2.1 holds. Then we have
\[
|dJ^D(\Gamma, u, p; V) - \partial_{\nu_{\Gamma}} dJ^D(\Gamma, u_h, p_h; V)| \leq C \left( \frac{h}{n} \right)^{\alpha} \|u\|_{W^{3,5}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^2(\Omega)} + \|p\|_{W^{3,5}(\Omega)} \|V_n\|_{H^{3/2}(\Gamma)}
\]
for some \( r > n \).

**Proof.** Recalling the definitions of \( dJ^D(\Gamma, u, p; V) \) in (2.7) and \( \partial_{\nu_{\Gamma}} dJ^D(\Gamma, u_h, p_h; V) \) in (3.25), we have
\[
dJ^D(\Gamma, u, p; V) - \partial_{\nu_{\Gamma}} dJ^D(\Gamma, u_h, p_h; V) = \int_{\Gamma} V_n \left[ j(u) - j(u_h \circ a_h^{-1}) + \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g \right] ds + \int_{\Gamma} V_n \left[ j(u) - j(u_h \circ a_h^{-1}) + \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g \right] (1 - t_h) ds
\]
where
\[
|J_2| \leq C \left( \frac{h}{n} \right)^{\alpha} \|u\|_{W^{3,5}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^2(\Omega)}
\]
because \( |1 - t_h| \leq C \frac{h}{n} \). Now we estimate \( J_1 \). The first and second terms of \( J_1 \) can be estimated as follows by using (3.23), the stability of \( \tilde{Q}_h \) and (3.24):
\[
\int_{\Gamma} V_n \left[ j(u) - j(u_h \circ a_h^{-1}) \right] ds \leq \|V_n\|_{L^2(\Gamma)} \|j(u) - j(u_h \circ a_h^{-1})\|_{L^2(\Gamma)}
\]
where we used [8, Lemma 1] or [30, eq. (2.1)] in the last step to estimate \( \|g\|_{H^1(\Gamma)} \). For the remaining terms we have
\[
V_n \left[ \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g \right] = V_n \left[ \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g + \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g \right] - \partial_{\nu_{\Gamma}} \partial_{\nu_{\Gamma}} g
\]
Let \( w \) be the solution to the following problem
\[
\begin{cases}
- \Delta w + w = 0 & \text{in } \Omega \\
w = V_n \partial_{\nu_{\Gamma}} p & \text{on } \Gamma.
\end{cases}
\]
The variational formulation is to find \( w \in H^1(\Omega) \) and \( w = V_n \partial_{\nu_{\Gamma}} p \) on \( \Gamma \) such that
\[
(\nabla w, \nabla v) + (w, v) = 0 \quad \forall v \in H^1_0(\Omega),
\]
where
while its finite element approximation can be written as: find $w_h \in V_h$ such that

$$
(\nabla w_h, \nabla v_h)_h + (w_h, v_h)_h = 0 \quad \forall v_h \in V_h^0, \quad w_h = Q_h(\nabla v_o \partial_p \circ a_h) \text{ on } \Gamma_h.
$$

(3.45)

Since $p \in H^3(\Omega) \cap H^2_0(\Gamma)$, we have $\partial_n p \in H^2(\Gamma)$ by the trace theorem (cf. [23, Section 1.5.1]) and thus $\mathbf{V}_n \partial_n p \in H^2(\Gamma)$ if $\mathbf{V}$ is sufficiently smooth, this implies $w \in H^2(\Omega)$. Then it follows from Theorem 1 in [8] and Remark 3.1 that

$$
\|w - w_h\|_{H^s(\Omega)} \leq Ch^{s-\frac{3}{2}} \|\mathbf{V}_n \partial_n p\|_{H^s(\Gamma)}, \quad \frac{1}{2} \leq s \leq \frac{3}{2}.
$$

(3.46)

Denote by $\tilde{w}_h$ the linear extension of $w_h$ to $\Omega$. Firstly, we have

$$
(\mathbf{V}_n \partial_n p, \partial_n u - \partial_n \tilde{w}_h)_h = (\mathbf{V}_n \partial_n p - \tilde{w}_h, \partial_n u - \partial_n \tilde{w}_h)_h + \langle \tilde{w}_h, \partial_n u - \partial_n \tilde{w}_h \rangle_h
$$

$$
= (\mathbf{V}_n \partial_n p - \tilde{w}_h, \partial_n u - \partial_n \tilde{w}_h)_h - (\partial_n \tilde{w}_h, \tilde{w}_h - w_h \circ a_h^{-1} t_h)_h
$$

$$
= (\mathbf{V}_n \partial_n p - \tilde{w}_h, \partial_n u - \partial_n \tilde{w}_h)_h + (\nabla(u - u_h), \nabla w_h)_h + (u - u_h, w_h)_h + M_h(u, \tilde{w}_h)
$$

$$
= (\mathbf{V}_n \partial_n p - \tilde{w}_h, \partial_n u - \partial_n \tilde{w}_h)_h - (\partial_n \tilde{w}_h, \tilde{w}_h - w_h \circ a_h^{-1} t_h)_h
$$

$$
+ (\nabla(u - u_h), \nabla(w_h - w))_h + (u - u_h, w_h - w)_h + (u - u_h, \partial_n w)_{\Gamma_h}
$$

(3.47)

where (3.43) is used in the last equality by testing $u - u_h$ and integration by parts on $\Omega_h$. In fact, by using (3.43) we have

$$
0 = (-\Delta w + w, u - u_h)_{\Omega_h \setminus \Gamma_h}
$$

$$
= (-\Delta w + w, u - u_h)_{\Omega_h \setminus \Gamma_h} - (-\Delta w + w, u - u_h)_{\Omega_h \setminus \Omega}
$$

$$
= (\nabla w, \nabla(u - u_h))_{\Omega_h \setminus \Gamma_h} + (u - u_h, w_h - w)_{\Gamma_h} - (-\Delta w + w, u - u_h)_{\Omega_h \setminus \Omega}
$$

by using the fact that $w_h$ have been extended to $\Omega_h \setminus \Omega$.

Now it remains to estimate the right-hand side of (3.47) terms by terms. An application of (3.6) and (3.46) yields

$$
(\nabla(u - u_h), \nabla(w_h - w))_h + (u - u_h, w_h - w)_h \leq \|u - u_h\|_{H^1(\Omega_h)} \|w - w_h\|_{H^1(\Omega_h)}
$$

$$
\leq Ch^2(\|f\|_{L^2(\Omega)} + \|g\|_{H^1(\Omega)}) \|\mathbf{V}_n \partial_n p\|_{H^2(\Gamma)}.
$$

(3.48)

On the other hand, using (3.6) we obtain

$$
(-\Delta w + w, u - u_h)_{\Omega_h \setminus \Omega} \leq \|w\|_{H^2(\Omega_h)} \|u - u_h\|_{L^2(\Omega_h)} \leq Ch^2\|u\|_{H^2(\Omega)} \|w\|_{H^2(\Omega)}.
$$

(3.49)

A direct use of (3.33) gives only first order convergence for $(\partial_n \tilde{w}_h, \tilde{w}_h - w_h \circ a_h^{-1} t_h)_\Gamma$. However, proceeding as in (3.33) and taking into account the proof of Lemma 3.6 give

$$
|\partial_n \tilde{w}_h, \tilde{w}_h - w_h \circ a_h^{-1} t_h|_\Gamma \leq C(\|\partial_n \tilde{w}_h\|_{L^2(\Gamma_h)} + \|\nabla \tilde{w}_h\|_{L^2(\Gamma_h)} + \|\tilde{w}_h\|_{L^2(\Gamma_h)})
$$

$$
\leq C\|\partial_n \tilde{w}_h\|_{L^2(\Gamma_h)}(\|\tilde{w}_h\|_{L^2(\Gamma_h)} + \|\nabla \tilde{w}_h\|_{L^2(\Gamma_h)} + \|w_h\|_{L^2(\Gamma_h)})
$$

$$
\leq C\|\partial_n \tilde{w}_h\|_{L^2(\Gamma_h)}(\|\tilde{w}_h\|_{L^2(\Gamma_h)} + \|\nabla \tilde{w}_h\|_{L^2(\Gamma_h)} + \|w_h\|_{L^2(\Gamma_h)})
$$

$$
\leq C\|\partial_n \tilde{w}_h\|_{L^2(\Gamma_h)}(\|\tilde{w}_h\|_{L^2(\Gamma_h)} + \|\nabla \tilde{w}_h\|_{L^2(\Gamma_h)} + \|w_h\|_{L^2(\Gamma_h)})
$$

(3.50)

where we have used the third last inequality in the proof of Lemma 3.6, the trace inequality

$$
\|\nabla \tilde{w}_h\|_{L^2(\Gamma_h)} \leq C\|\nabla \tilde{w}_h\|_{L^2(\Omega_h)} \|w_h\|_{L^2(\Omega_h)} \leq C\|\nabla \tilde{w}_h\|_{L^2(\Omega_h)}
$$

and the fact that $\tilde{w}_h$ is linear in $\Omega_h$, while in this case we assume $(\Omega \setminus \Omega_h) \cup (\Omega_h \setminus \Omega) = \sum \Omega_h^i$ such that $\Omega_h^i$ is the curved region bounded by $\Gamma^i_h \subset \Gamma_h$ and $\Gamma^i \subset \Gamma$. Similarly, a careful calculation implies

$$
M_h(u, \tilde{w}_h) = \int_{\Omega \setminus \Omega_h} (\nabla u \cdot \nabla \tilde{w}_h + u \tilde{w}_h - f \tilde{w}_h) \text{d}x - \int_{\Omega_h \setminus \Omega} (\nabla u \cdot \nabla w_h + u w_h - f w_h) \text{d}x
$$

$$
\leq \|\nabla u\|_{L^2(\Omega \setminus \Omega_h)}(\|\nabla (\tilde{w}_h - w)_h\|_{L^2(\Omega \setminus \Omega_h)} + \|\nabla (w_h - w)\|_{L^2(\Omega \setminus \Omega_h)}) + \int_{\Omega \setminus \Omega_h} (u - f) \tilde{w}_h \text{d}x
$$

$$
+ \|\nabla u\|_{L^2(\Omega \setminus \Omega_h)}(\|\nabla (w_h - w)\|_{L^2(\Omega \setminus \Omega_h)} + \|\nabla (w_h - w)\|_{L^2(\Omega \setminus \Omega_h)}) + \int_{\Omega_h \setminus \Omega} (u - f) w_h \text{d}x
$$

$$
\leq C h^2(\|u\|_{H^2(\Omega)} \|w\|_{H^2(\Omega)} + h \|w\|_{H^2(\Omega)}) + C h^2(\|u\|_{H^2(\Omega)} + \|f\|_{H^1(\Omega)}) \|\tilde{w}_h\|_{H^2(\Omega)}
$$

$$
+ C h^2(\|u\|_{H^2(\Omega)} \|w\|_{H^2(\Omega)} + h \|w\|_{H^2(\Omega)}) + C h^2(\|u\|_{H^2(\Omega)} + \|f\|_{H^1(\Omega)}) \|w_h\|_{H^2(\Omega)}
$$

$$
\leq C h^2(\|u\|_{H^2(\Omega)} + \|f\|_{H^1(\Omega)}) \|w\|_{H^2(\Omega)}.
$$

(3.51)
where we have used (3.46), and the facts that for \( w \in H^1(\Omega), v \in H^1(\Omega_n) \) there hold (cf. [13, eq. (5.16)], [8, Lemma 2 and Lemma 3] and [30, eq. (2.1) and (2.2)])

\[
\|w\|_{L^2(\Omega, \Omega_n)} \leq C h \|w\|_{H^1(\Omega)}, \quad \|v\|_{L^2(\Omega_n, \Omega)} \leq C h \|v\|_{H^1(\Omega_n)}.
\]

Furthermore, by using Lemma 3.6, (3.23), (3.24) and Lemma 3.8 we have

\[
(\mathbf{V}_n \partial_n p - \hat{w}_h, \partial_n u - \hat{u}_h)_{\Gamma_h} = (\mathbf{V}_n \partial_n p - \hat{q}_h(\mathbf{V}_n \partial_n p), \partial_n u - \hat{u}_h)_{\Gamma_h} + (\hat{q}_h(\mathbf{V}_n \partial_n p) - Q_h(\mathbf{V}_n \partial_n p \circ a_h) \circ a_h^{-1}, \partial_n u - \hat{u}_h)_{\Gamma_h} + (w_h \circ a_h^{-1} - \hat{w}_h, \partial_n u - \hat{u}_h)_{\Gamma_h} \\
\leq C h^2 \|\mathbf{V}_n \partial_n p\|_{H^2(\Gamma)}(\|u\|_{W^{3,6}(\Omega)} + \|f\|_{H^1(\Omega)}).
\]

It remains to estimate \( (u - u_h, \partial_n w)_{\Gamma_h} \). We conclude from (3.23) and (3.24) that

\[
(\mathbf{V}_n \partial_n p - \hat{w}_h, \partial_n u - \hat{u}_h)_{\Gamma_h} \leq \|g\|_{\Gamma_h} \leq C l(\|g\|_{\Gamma_h} \circ a_h^{-1} \circ a_h^{-1} \|L^2(\Gamma)\|) \leq C \|\mathbf{V}_n \partial_n p\|_{H^2(\Gamma)}(\|u\|_{W^{3,6}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^3(\Omega)}).
\]

Similarly, we can prove

\[
(\mathbf{V}_n \partial_n u, \partial_n p - \hat{p}_h)_{\Gamma_h} \leq C h^2 \|\mathbf{V}_n \partial_n u\|_{H^2(\Gamma)}(\|u\|_{W^{3,6}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^3(\Omega)} + \|p\|_{W^{3,6}(\Omega)}).
\]

and

\[
(\mathbf{V}_n \partial_n g, \partial_n p - \hat{p}_h)_{\Gamma_h} \leq C h^2 \|\mathbf{V}_n \partial_n g\|_{H^2(\Gamma)}(\|u\|_{W^{3,6}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^3(\Omega)} + \|p\|_{W^{3,6}(\Omega)}).
\]

Moreover,

\[
(\mathbf{V}_n (\partial_n u - \hat{u}_h), \partial_n p - \hat{p}_h)_{\Gamma_h} \leq \|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|\partial_n p - \hat{p}_h\|_{L^2(\Gamma)} \|\partial_n u - \hat{u}_h\|_{L^2(\Gamma)} \leq C h^2 \|\mathbf{V}_n\|_{H^2(\Gamma)}(\|u\|_{W^{3,6}(\Omega)} + \|f\|_{H^1(\Omega)} + \|g\|_{H^3(\Omega)} + \|p\|_{W^{3,6}(\Omega)})^2.
\]

Combining the above results we have \(|J_1| \leq C h^2\), this together with \(|J_2| \leq C h^2\) gives the result. \(\square\)

In the following we discuss the possible extension of the results in Theorem 3.10 to the convex polygonal domains.

**Remark 3.11.** First we recall the regularity results of the solution to the elliptic equations (2.2) with Dirichlet boundary conditions. Let \(\Omega\) be a convex polygonal domain with Lipschitz boundary \(\Gamma\) and \(f, g\) be smooth enough in \(\mathbb{R}^n\). We denote by \(\frac{2}{n} \leq \omega_1 < \pi\) the largest interior angle of \(\Gamma\), and by

\[
s_\Omega = \frac{2}{2 - \min\{2, \frac{\pi}{\omega_1}\}} > 2, \quad r_\Omega = 1 + \pi/\omega_1 \in (2, 4]
\]

the exponents giving the maximal regularity to the solution of (2.2) in \(W^{2,s}(\Omega)\) for \(s < s_\Omega\) ([23, Theorem 4.4.3.7]) and \(H^r(\Omega)\) for \(r < r_\Omega\) ([23, Theorem 5.1.1.4]) if \(f\) and \(g\) are smooth enough. In the case \(n = 3\), \(s_\Omega\) and \(r_\Omega\) have more complicated expressions depending on interior angles of both edges and corners of the polyhedron; see [18, Corollary 3.9 and Section 4.c]. The above regularity also applies to the adjoint state equation (2.5) if \(j'(u)\) is sufficiently smooth.

In particular, the formulas (3.13) and (3.14) are well-defined. In fact, taking the state \(u\) as an example, if \(\Omega\) is a convex polygonal domain and \(g = 0\) we have \(u \in W^{2,s}(\Omega) \cap H^s_0(\Omega)\) and thus \(\partial_n u \in W^{1-1/s,s}(\Gamma)\) (We refer to [12, Theorem 3.4] where \(u \in H^s_0(\Omega)\) plays a crucial role in deriving \(\partial_n u \in W^{1-1/s,s}(\Gamma)\)). If \(g \neq 0\) we may need some compatibility condition for \(g\) to derive \(\partial_n u \in W^{1-1/s,s}(\Gamma)\).

We remark that the proof for the improved convergence rate of the discrete shape gradients of boundary type by using the corrected outward normal derivatives relies on the \(L^2\)- and \(H^1\)-norm error estimates (3.6), the
$W^{1,\infty}$-norm error estimate (3.8), the first order convergence of the outward norm derivative approximations in Lemma 3.8, the regularity $V_p \partial_n^0 p, V_p \partial_n^0 u, V_p \partial_n^0 g \in H^2(\Gamma)$ which in turn ensures the $H^2(\Omega)$ regularity for the solution $w$ to problem (3.43) and the error estimates (3.46).

Specifically, (3.6) can be achieved for a convex polygonal domain, while the $W^{2,\infty}$ regularity for $u$ and $p$ imposes restriction to the maximal interior angle $\omega_i$, i.e., $\omega_i \leq \frac{\pi}{3}$. Further, if $u, p \in H^3(\Omega) \cap H^1_0(\Omega)$ and $g = 0$ we can expect $V_p \partial_n^0 p, V_p \partial_n^0 u \in H^{2-\varepsilon}(\Gamma)$, because $V_p \partial_n^0 p, V_p \partial_n^0 u$ are piecewise $H^{2-\varepsilon}$ on $\Gamma$ and $\partial_n^0 p$ and $\partial_n^0 u$ vanish at the corners. For the approximation of the outward normal derivative in general polygonal domain we can only expect ([12])

$$\|\partial_n^0 u - \partial_n^0 u_h\|_{L^2(\Gamma)} \leq Ch^{1-\frac{\varepsilon}{2}}\|u\|_{W^{2-\varepsilon}(\Omega)}.$$ 

However, the above error estimate can be improved to be $O(h^{1-\varepsilon})$ for any $\varepsilon > 0$ if the maximum angle of $\Omega$ is smaller than $\frac{\pi}{3}$ (cf. [4, Figure 1 and Theorem 4.1]). Furthermore, the order can be improved to be $O(h^{2-\varepsilon})$ (cf. [4, Figure 1 and Theorem 4.1]) if the maximum angle of $\Omega$ is smaller than $\frac{\pi}{3}$ and the mesh satisfies the super-convergence property, i.e., is of $O(h^{2\varepsilon})$ (cf. [15]).

In conclusion, it seems that we cannot prove second order convergence for the modified shape gradient in a general polygonal domain. However, in the case that $f, g$ are smooth enough and $\omega_i \leq \frac{\pi}{3}$, we can expect almost second order convergence rate. The first example (Example 1) in Section 4 shows this optimal second order convergence rate, while the third example (Example 3) shows a better result than that we expect.

3.2. Neumann boundary value problem. In this subsection we consider the Neumann boundary value problem. For the state equation (2.2) with Neumann boundary condition, the weak formulation is to find $u \in H^1(\Omega)$ such that

$$(\nabla u, \nabla v) + (u, v) = (f, v) + (g, v)_{\Gamma} \quad \forall v \in H^1(\Omega).$$

(3.54)

The finite element approximation of (3.54) reads: find $u_h \in V_h \subset H^1(\Omega_h)$ such that

$$(\nabla u_h, \nabla v_h)_h + (u_h, v_h)_h = (f, v_h)_h + (g, v_h)_{\Gamma_h} \quad \forall v_h \in V_h.$$ 

(3.55)

Similarly, the variational weak formulation of (2.5) is to find $p \in H^1(\Omega)$ such that

$$(\nabla p, \nabla v) + (p, v) = (j'(u), v) \quad \forall v \in H^1(\Omega),$$ 

(3.56)

while its finite element approximation reads: find $p_h \in V_h \subset H^1(\Omega_h)$ such that

$$(\nabla p_h, \nabla v_h)_h + (p_h, v_h)_h = (j'(u_h), v_h)_h \quad \forall v_h \in V_h.$$ 

(3.57)

To use known results in the literature we use a slightly different lifting operator (cf. [19, 30]) from that used in Section 3.1. As an embedded, compact hypersurface in $\mathbb{R}^n$, $\Gamma$ is orientable with a unit normal field $n$ pointing to the outside of the domain $\Omega$ and hence can be represented by the zero level set of the signed distance function $d$ such that (cf. [17])

$$|d(x)| = \text{dist}(x, \Gamma) \quad \text{and} \quad n(x) = \frac{\nabla d(x)}{||\nabla d(x)||} \quad \text{for } x \in \Gamma.$$ 

Furthermore, one can find a neighborhood $U \subset \mathbb{R}^n$ of $\Gamma$, such that $d$ is also of class $C^3$ on $U$ and the projection $a : \Gamma_h \to \Gamma$ is well defined by

$$a(x) := x - d(x)\nabla d(x), \quad x \in \Gamma_h$$

such that $\nabla d(x) = n(a(x))$, and $a$ can be extended as a mapping from $U$ to $\Gamma$. Then one can show that $|d|_{L^\infty(\Gamma_h)} \leq Ch^2$ and $|1 - \delta_d|_{L^\infty(\Gamma_h)} \leq Ch^2$ (cf. [19, Lemma 4.1]) where $\delta_d$ denotes the quotient between the smooth and discrete surface measures $ds$ and $ds_h$, i.e., $\delta_d ds_h = ds$. The main difference between the projection mapping $a(x)$ defined in this subsection with the one $a_h(x)$ defined in (3.15) lies in the normal directions. Specifically, in the definition (3.15) the projection $a_h(x)$ is along the normal direction of $\Gamma_h$, while the projection $a(x)$ used in this subsection is along the normal direction of $\Gamma$.

For any function $v_h$ defined on $\Gamma_h$ we can define the lifting $v^\Gamma_h$ on $\Gamma$ as $v^\Gamma_h(a(x)) = v_h(x)$. Similarly, for any function $v$ defined on $\Gamma$ we can define the inverse lifting $v^{-\Gamma}$ on $\Gamma_h$ as $v^{-\Gamma}(x) = v \circ a(x)$. Then there holds $(v^{-\Gamma})^\Gamma = v$. On $\Gamma_h$ we define for $v_h \in V_h(\Gamma_h)$ the tangential operator $\nabla_{\Gamma_h} v_h := \nabla v_h - (\nabla v_h \cdot n_h)n_h = P_h n_h \nabla v_h$, where $(P_h)_i,k = \delta_{i,k} - n_h,i n_h,k, i, k = 1, 2$. From the definition of $v^\Gamma_h$ we have $\nabla v_h = (P - dH)\nabla(v^\Gamma_h) \circ a$ with $P_h = \delta_2 n_h, n_h$ and $H_{i,j} = \partial_{x_i}\partial_{x_j} d = \partial_{x_i} n_h = \partial_{x_j} n_h$. Using the property $PH = HP = H$ we obtain (cf. [19, Lemma 4.2])

$$\nabla_{\Gamma_h} v_h(x) = P_h(x)(I - d(x)H(x))P(x)\nabla v^\Gamma_h(a(x)) = P_h(x)(I - d(x)H(x))\nabla_{\Gamma} v^\Gamma_h(a(x)) \quad \text{on } \Gamma_h.$$ 

(3.58)
Then one has the following stabilities for the lifting operator (cf. [19, Lemma 4.2]):

\[ \frac{1}{C} \| v \|_{L^2(\Gamma_h)} \leq \| v' \|_{L^2(\Gamma)} \leq C \| v \|_{L^2(\Gamma_h)}, \]

\[ \frac{1}{C} \| \nabla_{\Gamma_h} v \|_{L^2(\Gamma_h)} \leq \| \nabla_{\Gamma} v \|_{L^2(\Gamma)} \leq C \| \nabla_{\Gamma_h} v \|_{L^2(\Gamma_h)}. \]  

(3.59)

Now we are ready to define the approximate shape gradient to (2.11):

\[ d_h J^N(\Gamma_h, u_h, p_h; \mathbf{V}) = \int_{\Gamma_h} \mathbf{V}_n \cdot \partial_a \left( j(u_h) - \nabla_{\Gamma_h} u_h \cdot \nabla_{\Gamma_h} p_h - u_h p_h + (f + \partial_n g \circ a + K \circ ag)p_h \right) ds_h, \]

(3.60)

where \( \nabla_{\Gamma_h} \) denotes the tangential gradient on \( \Gamma_h \).

**Remark 3.12.** The definition (3.60) involves the mapping \( a \) between \( \Gamma_h \) and \( \Gamma \), which is not trivial to compute explicitly. In the special case of implicitly defined domain (cf. [19, Sections 7 and 8]), the boundary \( \Gamma \) is the zero level set of the signed distance function \( d \) and thus the mapping \( a \) can be explicitly computed. In the general case, we usually compute the integral (3.60) by using quadrature formula, thus we only need to evaluate \( \partial_n \) explicitly. In the special case of implicitly defined domain (cf. [19, Sections 7 and 8]), the boundary \( \Gamma \) is the zero level set of the signed distance function \( d \) and thus the mapping \( a \) can be explicitly computed. In the general case, we usually compute the integral (3.60) by using quadrature formula, thus we only need to evaluate \( \partial_n \) explicitly.

It was numerically observed in [29] that the approximate shape gradient in the boundary formulation of the Eulerian derivative for the Neumann boundary value problem has the same convergence rate and accuracy as for the volume formulation on convex polygon or smooth domain having \( C^2 \) boundary. For such domains, however, the theoretical analysis therein shows that the discrete boundary formulation has lower convergence rate than the volumetric formulation when the data is smooth. In this subsection, we prove that the boundary shape gradient has nearly the same convergence rate as in the volume case for the Neumann boundary value problem, when the solution is sufficiently regular.

Converting (3.60) to \( \Gamma \) we have (cf. [19, Lemma 4.7])

\[ \tilde{d}_h J^N(\Gamma, u_h, p_h; \mathbf{V}) = \int_{\Gamma} \mathbf{V}_n \left( j(u_h) - \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h - u_h p_h + (f + \partial_n g + K g)p_h \right) \frac{1}{\delta_h} ds \]

\[ + \int_{\Gamma} \mathbf{V}_n (1 - B_h) \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h ds + \int_{\Gamma} \mathbf{V}_n \left( \frac{1}{\delta_h} - 1 \right) (j(u_h) - u_h p_h + (f + \partial_n g + K g)p_h) ds, \]

where (cf. [19, P. 317, proof of Lemma 4.7])

\[ B_h = \frac{1}{\delta_h} (1 - dH) P_h P (\| - dH) = \frac{1}{\delta_h} P (1 - dH) P_h (\| - dH) P \]

since \( P \) is a projection.

**Theorem 3.13.** Let \( u \) and \( p \) be the continuous solutions to the state equation (3.54) and adjoint equation (3.56) with Neumann boundary conditions, while \( u_h \) and \( p_h \) denote the discrete solutions to (3.55) and (3.57), respectively. Assume that Assumption 2.1 holds and \( u, p \in W^{3+\gamma}(\Omega) \) for some \( r > n \), we have

\[ |d J^N(\Gamma, u, p; \mathbf{V}) - \tilde{d}_h J^N(\Gamma, u_h, p_h; \mathbf{V})| \leq C h^2 \| \mathbf{V}_n \|_{W^{1,\infty}(\Gamma)} (\| u \|_{W^{2,\infty}(\Omega)} + \| p \|_{W^{2,\infty}(\Omega)} + \| f \|_{H^1(\Omega)} + \| g \|_{H^2(\Omega)}). \]  

(3.61)

**Proof.** From the definition we have

\[ d J^N(\Gamma, u, p; \mathbf{V}) - \tilde{d}_h J^N(\Gamma, u_h, p_h; \mathbf{V}) \]

\[ = \int_{\Gamma} \mathbf{V}_n (j(u) - j(u_h)) - (u - u_h)p_h + (f + \partial_n g + K g)(p - p_h) ds \]

\[ - \int_{\Gamma} \mathbf{V}_n (\nabla_{\Gamma} u \cdot \nabla_{\Gamma} p - \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h) ds - \int_{\Gamma} \mathbf{V}_n (1 - B_h) \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h ds \]

\[ - \int_{\Gamma} \mathbf{V}_n \left( \frac{1}{\delta_h} - 1 \right) (j(u_h) - u_h p_h + (f + \partial_n g + K g)p_h) ds. \]

Now it remains to estimate the right-hand side. First, it follows from Lemma 4.7 in [19] that

\[ \int_{\Gamma} \mathbf{V}_n (1 - B_h) \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h ds = \int_{\Gamma} \mathbf{V}_n (P - B_h) \nabla_{\Gamma} u_h \cdot \nabla_{\Gamma} p_h ds \]

\[ \leq C h^2 \| \mathbf{V}_n \|_{L^\infty(\Gamma)} (\| u \|_{W^{1,\infty}(\Omega)} + \| p \|_{W^{1,\infty}(\Omega)}). \]  

(3.63)
where we used the fact that $|P - B_h| \leq Ch^2$ on $\Gamma$. Second, using $\|1 - \delta_h\|_{L^\infty(\Gamma_h)} \leq Ch^2$ and the stability results (3.59) one can derive

$$
\int_\Gamma \mathbf{V}_n \left( \frac{1}{\delta_h} - 1 \right) \left( j(u_h) - u_h p_h + (f + \delta_n g + Kg)p_h \right) ds \leq Ch^2 \| \mathbf{V}_n \|_{L^\infty(\Gamma)} \left( \|u\|_{H^1(\Omega)} + \|g\|_{H^2(\Omega)} \right) + \|f\|_{H^1(\Omega)} + \|g\|_{H^2(\Omega)} \|p\|_{H^1(\Omega)}.
$$

(3.64)

Note that for our case of $C^3$ domain $\Omega$ and sufficiently smooth data $f$ and $g$, we have the $W^{2,\infty}(\Omega)$ regularity for the solutions $u$ and $p$ to the state and adjoint equations. Denote by $\tilde{u}, \tilde{p}$ their continuous extensions to $\mathbb{R}^n$ satisfying (3.1), we thus have $\tilde{u}, \tilde{p} \in W^{2,\infty}(\Omega_h)$. In the following we will derive the estimates for $\|u - \tilde{u}\|_{L^2(\Gamma)}$ and $\|p - \tilde{p}\|_{L^2(\Gamma)}$.

It follows from (3.59) and [30, eq. (2.1) with $p = 2$] that

$$
\|u - \tilde{u}\|_{L^2(\Gamma)} \leq C\|u - \tilde{u}\|_{L^2(\Gamma)} + \|\tilde{u} - u_h\|_{L^2(\Gamma)} \leq C\|\tilde{u}\|_{L^2(\Gamma)} + h^2 \|\nabla \tilde{u}\|_{L^2(\Gamma)},
$$

(3.65)

which together with (3.1), (3.59) and the $L^\infty$-norm error estimate [30, Theorem 3.1] gives

$$
\|u - u_h\|_{L^2(\Gamma)} \leq C\|u - \tilde{u}\|_{L^2(\Gamma)} + \|\tilde{u} - u_h\|_{L^2(\Gamma)} \leq C\|\tilde{u}\|_{L^2(\Gamma)} + C\|\tilde{u} - u_h\|_{L^\infty(\Omega_h)} \leq C\|\tilde{u}\|_{L^2(\Gamma)} + C\|\tilde{u} - u_h\|_{L^\infty(\Omega_h)} \leq C\|\tilde{u}\|_{L^2(\Gamma)} + Ch^2 \|\log h\|_{\Omega} \|\tilde{u}\|_{W^{2,\infty}(\Omega)}.
$$

(3.66)

As for the adjoint state variable $p$, we can proceed similarly as above

$$
\|p - p_h\|_{L^2(\Gamma)} \leq \|p - \tilde{p}\|_{L^2(\Gamma)} + \|\tilde{p} - p_h\|_{L^2(\Gamma)} \leq C\|\tilde{p}\|_{L^2(\Gamma)} + C\|\tilde{p} - p_h\|_{L^2(\Gamma)} \leq C\|\tilde{p}\|_{L^2(\Gamma)} + C\|\tilde{p} - p_h\|_{L^\infty(\Omega_h)} \leq C\|\tilde{p}\|_{L^2(\Gamma)} + C\|\tilde{p} - p_h\|_{L^\infty(\Omega_h)} \leq C\|\tilde{p}\|_{L^2(\Gamma)} + Ch^2 \|\log h\|_{\Omega} \|p\|_{W^{2,\infty}(\Omega)},
$$

(3.67)

where $p_h(u) \in V_h$ is defined similarly as in Remark 3.3, and $I$ is also defined as in Remark 3.3.

Then we have

$$
\int_\Gamma \mathbf{V}_n \left( j(u) - j(u_h) \right) ds \leq C\|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|j\|_{C^{0,1}(I)} \|u - u_h\|_{L^1(\Gamma)} \leq C\|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|j\|_{C^{0,1}(I)} \|u - u_h\|_{L^2(\Gamma)} \leq C\|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|j\|_{C^{0,1}(I)} \|u - u_h\|_{W^{2,\infty}(\Omega)}.
$$

Moreover, it holds

$$
\int_\Gamma \mathbf{V}_n \left( u(p - u_h') + u(p - p_h) - (u - u_h')(p - p_h) \right) ds \leq \|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|p\|_{L^\infty(\Gamma)} \|u - u_h\|_{L^2(\Gamma)} + \|u\|_{L^2(\Gamma)} \|p - p_h\|_{L^2(\Gamma)} + \|u - u_h'\|_{L^2(\Gamma)} \|p - p_h\|_{L^2(\Gamma)} + Ch^2 \|\log h\|_{\Omega} \|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|u\|_{W^{2,\infty}(\Omega)} + \|p\|_{W^{2,\infty}(\Omega)}
$$

$$
\int_\Gamma \mathbf{V}_n (f + Kg)(p - p_h) ds \leq \|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|f + Kg\|_{L^2(\Gamma)} \|p - p_h\|_{L^2(\Gamma)} \leq Ch^2 \|\log h\|_{\Omega} \|\mathbf{V}_n\|_{L^\infty(\Gamma)} \|f\|_{H^1(\Omega)} + \|g\|_{H^2(\Omega)}.
$$

(3.68)

By the tangential Green’s formula [19, Theorem 2.14]

$$
\int_\Gamma (w \text{div}_\Gamma v + \nabla_\Gamma w \cdot v) ds = \int_\Gamma Kw \cdot nds
$$

(3.69)
for a function \( w \in H^1(\Omega) \) and vector \( \mathbf{v} \in H^1(\Omega)^n \), we have
\[
\int_{\Gamma} \mathbf{v}_n (\nabla_G u \cdot \nabla_G p \cdots \nabla_G u \cdot \nabla_G p_h) ds \\
= \int_{\Gamma} \mathbf{v}_n (\nabla_G u \cdot \nabla_G (p - p_h) + \nabla_G p \cdot \nabla_G (u - u_h) - \nabla_G (u - u_h) \cdot \nabla_G (p - p_h)) ds \\
= -\int_{\Gamma} \text{div}_G(\mathbf{v}_n \nabla_G u)(p - p_h) ds - \int_{\Gamma} \text{div}_G(\mathbf{v}_n \nabla_G p)(u - u_h) ds \\
- \int_{\Gamma} \mathbf{v}_n \nabla_G (u - u_h) \cdot \nabla_G (p - p_h) ds.
\]

Then, we estimate the three terms on the right-hand side. First, using (3.67) we have
\[
\int_{\Gamma} \text{div}_G(\mathbf{v}_n \nabla_G u)(p - p_h) ds \leq \|\text{div}_G(\mathbf{v}_n \nabla_G u)\|_{L^2(\Gamma)}\|p - p_h\|_{L^2(\Gamma)} \\
\leq C h^2 \log h \|\mathbf{v}_n\|_{W^{1,\infty}(\Gamma)} \|u\|_{H^{1/2}(\Omega)} (\|u\|_{W^{2,\infty}(\Omega)} + \|p\|_{W^{2,\infty}(\Omega)}).
\]

Similarly, it follows from (3.66) that
\[
\int_{\Gamma} \text{div}_G(\mathbf{v}_n \nabla_G p)(u - u_h) ds \leq \|\text{div}_G(\mathbf{v}_n \nabla_G p)\|_{L^2(\Gamma)}\|u - u_h\|_{L^2(\Gamma)} \\
\leq C h^2 \log h \|\mathbf{v}_n\|_{W^{1,\infty}(\Gamma)} \|p\|_{H^{1/2}(\Omega)} \|u\|_{W^{2,\infty}(\Omega)}.
\]

Furthermore, it follows from the triangle inequality, the definition of \( \nabla_G v \), the stability results (3.1) and (3.59), the estimate (3.65) and an \( W^{1,\infty} \) error estimate (cf. [30, Theorem 3.1]) that
\[
\|\nabla_G (u - u_h)\|_{L^2(\Gamma)} \leq \|\nabla_G (u - \bar{u})\|_{L^2(\Gamma)} + \|\nabla_G (\bar{u} - u_h)\|_{L^2(\Gamma)} \\
\leq C \|\nabla_G (u - \bar{u})\|_{L^2(\Gamma_h)} + C \|\nabla_G (\bar{u} - u_h)\|_{L^2(\Gamma_h)} \\
\leq C h \|\nabla_G \bar{u}\|_{L^2(\Gamma_h)} + C h \|u\|_{W^{2,\infty}(\Omega)} \\
\leq C h \|u\|_{W^{2,\infty}(\Omega)}.
\]

A similar argument as above and proceeding as in the estimate of \( \| p - p_h \|_{L^2(\Gamma)} \) yield
\[
\|\nabla_G (p - p_h)\|_{L^2(\Gamma)} \leq C h \|u\|_{W^{2,\infty}(\Omega)}.
\]

Finally, we have
\[
\int_{\Gamma} \mathbf{v}_n \nabla_G (u - u_h) \cdot \nabla_G (p - p_h) ds \leq C \|\mathbf{v}_n\|_{L^\infty(\Gamma)} \|\nabla_G (u - u_h)\|_{L^2(\Gamma)} \|\nabla_G (p - p_h)\|_{L^2(\Gamma)} \\
\leq C h^2 \|\mathbf{v}_n\|_{L^\infty(\Gamma)} \|u\|_{W^{2,\infty}(\Omega)} \|p\|_{W^{2,\infty}(\Omega)}.
\]

Combining the above estimates we finish the proof.

4. Numerical results. In this section we present some numerical results to verify the convergence rate of finite element approximations to shape gradients with boundary corrections.

We numerically verify the theoretical results in Theorem 3.10. The shape gradient for shape functional is a linear continuous operator on \( H^{3/2}(\mathbb{R}^n; \mathbb{R}^n) \) and belongs to its dual space in either the volume or the boundary type Eulerian derivative. It is challenging to compute numerically the continuous infinite-dimensional operator norm for the approximate shape gradients. This norm can be approximately replaced by a tractable one on a finite-dimensional subspace. More precisely, given a positive integer \( \gamma \) as in [29], we consider an approximate operator norm on a finite-dimensional space consisting of vector fields in \( P_\gamma(\mathbb{R}^n; \mathbb{R}^n) \subset H^{3/2}(\mathbb{R}^n; \mathbb{R}^n) \), whose components are multivariate polynomials of degree up to \( \gamma \). Based on the equivalence of norms over finite-dimensional spaces, we replace the \( H^{3/2} \)-norm with a more tractable \( H^1 \)-norm. Finally, we compute the approximate dual norms
\[
\mathcal{E} := \left( \max_{\mathbf{v} \in P_\gamma(\mathbb{R}^n; \mathbb{R}^n)} \frac{|dJ^b(\Gamma, u, p; \mathbf{v}) - \delta_h J^b(\Gamma, u_h, p_h; \mathbf{v})|^2}{\|\mathbf{v}\|^2_{H^1(\Omega)}} \right)^{1/2}.
\]
We take a global basis \( \{ \mathbf{V}_i \}^q_{i=1} \) of vector fields in \( \mathcal{P}_q(\mathbb{R}^n;\mathbb{R}^n) \), where \( q = nC^n_{\gamma+n} \) with \( C^n_{\gamma+n} \) denoting the combination coefficient. More precisely, we choose
\[
\{ \mathbf{V}_i \}^q_{i=1} = \{ (\Pi_{i=1}^n x_i^{\beta_i}, 0, \cdots, 0), \cdots, [0, \cdots, 0, \Pi_{i=1}^n x_i^{\beta_i}] \} \text{ with } \sum_{i=1}^n \beta_i \leq \gamma,
\]
where \( \beta_i \) (\( i = 1, \cdots, n \)) are non-negative integers. We denote the Gramian matrix associated with the \( H^1(\Omega) \) inner product by \( \mathbb{K} = [ (\mathbf{V}_i, \mathbf{V}_k)_{H^1(\Omega)} ]^q_{i,k=1} \in \mathbb{R}^{q \times q} \). The errors \( (4.1) \) can be obtained by
\[
\mathcal{E} := (w^T \mathbb{K}^{-1} w)^{1/2},
\]
where
\[
w := |dJ^D(\Gamma, u, p; \mathbf{V}_i) - \delta_h J^D(\Gamma, u_h, p_h; \mathbf{V}_i)|^q_{i=1}.
\]

In the following examples we set \( u_d = 0 \) and use the cost functional \( J(\Omega) = \frac{1}{2} \int\Omega u^2 dx \), so that \( j(u) := \frac{1}{2} u^2 \). We construct **Example 1** with
\[
u(x_1, x_2) = \cos \left( \frac{\pi}{2} x_1 \right) \cos \left( \frac{\pi}{2} x_2 \right), \quad f = \left( \frac{1}{2} \pi^2 + 1 \right) \cos \left( \frac{\pi}{2} x_1 \right) \cos \left( \frac{\pi}{2} x_2 \right)
\]
on \( \Omega = (-1,1)^2 \) and set \( g \) accordingly. The adjoint state has exact solution
\[
p(x_1, x_2) = \frac{1}{2\pi^2 + 4} \cos \left( \frac{\pi}{2} x_1 \right) \cos \left( \frac{\pi}{2} x_2 \right).
\]

In the following, we use a numerical solution on a very fine mesh as reference for computing numerical errors. In **Example 2**, we set \( \Omega \) to be a disk of radius \( \sqrt{3} \) centered at the origin. We choose
\[
u(x_1, x_2) = \cos(x_1) \cos(x_2), \quad f = 3 \cos(x_1) \cos(x_2)
\]
and set inhomogeneous Dirichlet boundary condition \( g = \cos(x_1) \cos(x_2) \). In this example with a curved computational domain, the vector field \( \mathbf{V}_n \) was not composed with \( a_h \) when using the formula \( (3.10) \), however, we still observe second order convergence. In **Example 3**, we choose \( \Omega \) to be an irregular convex polygon intersected by four lines: \( x_1 = 0 \), \( x_2 = 0 \), \( x_2 = -\sqrt{3} x_1 + \sqrt{3} \), and \( x_2 = (\sqrt{3} - 2) x_1 + 1 \), such that the maximum interior angle is larger than 90°. We set
\[
f = e^{-x_1^2 x_2} \cos(x_1) \sin(x_2)
\]
and \( g = 0 \). Both solutions to the state and adjoint equation are unknown. In **Example 4**, we choose a domain that does not guarantee \( H^2 \)-regularity of the state problem. We set a L-shaped domain with a reentrant corner \( \Omega = (-1,1)^2 \setminus \{ [0,1] \times (-1,0) \} \). The source function is, in polar coordinates, \( f(x) = r^{2/3} \sin(2\theta/3) \) with \( r = \sqrt{x_1^2 + x_2^2} \) and \( \theta \) being the polar angle. Choose \( g = 0 \). In Fig. 4.1, we show the finite element approximations of the state equations for domains associated with the three examples.

For **Examples 1-3**, we see from Fig. 4.2 that the errors committed by shape gradients with boundary correction converge quadratically, while the errors of those without boundary correction converge linearly. When the domain is nonconvex, so that the solution is not sufficiently regular (not in \( H^2(\Omega) \)) as for **Example 4**, Fig. 4.2 (the third subplot) shows that the shape gradient with boundary correction converges faster than the classical one, but with a rate less than 2, this is in agreement with that of the volume one (cf. [29]). The parameter \( \gamma \) has no obvious effect on convergence as shown in Fig. 4.3.

**Acknowledgements:** The authors would like to thank the two anonymous referees for their careful reading of the manuscript and many valuable and constructive comments, which helped to improve the quality of the paper greatly. We also would like to thank Professor Tomoya Kemmochi for the helpful discussions on the \( L^\infty \)-norm error estimates of elliptic Dirichlet problems in curved domains.

**REFERENCES**

Four computational domains and the according finite element solutions for Example 1, Example 2, Example 3, and Example 4, respectively.

Fig. 4.2. The convergence history of approximate shape gradients with $\gamma = 3$.

Fig. 4.3. The convergence history of approximate shape gradients with $\gamma = 2$: square (left) and disk (right).

459-485.


