Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Two-norm discrepancy and convergence of the stochastic gradient method with application to shape optimization

Marc Dambrine[1], Caroline Geiersbach[2], Helmut Harbrecht[3]

submitted: August 20, 2024

[1] Université de Pau et des Pays de l'Adour
IPRA-LMA, UMR CNRS 5142
Avenue de l'Université
64000 Pau
France
E-Mail: marc.dambrine@univ-pau.fr

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: caroline.geiersbach@wias-berlin.de

[3] Department of Mathematics and Computer Science
University of Basel
Spiegelgasse 1
4051 Basel
Switzerland
E-Mail: helmut.harbrecht@unibas.ch

No. 3121
Berlin 2024

2020 *Mathematics Subject Classification.* 35R60, 60H30, 35R35, 60H35, 90C15.

Key words and phrases. Stochastic gradient method, shape optimization, free boundary problem, two-norm discrepancy, optimization under uncertainty.

This research has been in part performed while H.H. was visiting the Laboratory of Mathematics and its Applications of PAU – UMR CNRS 5142. The hospitality and the support are gratefully acknowledged.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Two-norm discrepancy and convergence of the stochastic gradient method with application to shape optimization

Marc Dambrine, Caroline Geiersbach, Helmut Harbrecht

Abstract

The present article is dedicated to proving convergence of the stochastic gradient method in case of random shape optimization problems. To that end, we consider Bernoulli's exterior free boundary problem with a random interior boundary. We recast this problem into a shape optimization problem by means of the minimization of the expected Dirichlet energy. By restricting ourselves to the class of convex, sufficiently smooth domains of bounded curvature, the shape optimization problem becomes strongly convex with respect to an appropriate norm. Since this norm is weaker than the differentiability norm, we are confronted with the so-called two-norm discrepancy, a well-known phenomenon from optimal control. We therefore need to adapt the convergence theory of the stochastic gradient method to this specific setting correspondingly. The theoretical findings are supported and validated by numerical experiments.

1 Introduction

Shape optimization under uncertainty is a topic of growing interest, see for example [1, 2, 11, 13, 15, 39] and the references therein. The most common approach is the minimization of the expectation of the shape functional. In specific cases, this problem can be reformulated as a deterministic one, see e.g. [17, 19]. However, this is not possible in general, which makes the shape optimization algorithm quite costly. One popular approach for the minimization of the expectation is offered by the stochastic gradient method, which originated in [41] and has been used in recent years in the optimal control of partial differential equations involving uncertain inputs or parameters; see, e.g., [27, 38]. In the present article, we intend to verify the convergence of this method in case of *Bernoulli's exterior free boundary problem* in case of a *random interior boundary*. Bernoulli's exterior free boundary problem is an overdetermined boundary value problem for the Laplacian, where one has an inclusion with Dirichlet boundary condition and an exterior, *free boundary* with Dirichlet and Neumann boundary condition. This free boundary problem becomes random when we assume that the interior boundary is random.

The aforementioned random free boundary problem has already been considered in several articles in different settings by some of the authors of this article. Bernoulli's free boundary problem can be seen as a "fruit fly" of shape optimization, see [18, 20, 34]. In particular, much is known about existence and regularity of the solution to the free boundary problem in the deterministic setting, see e.g. [3, 7, 22, 36, 44] for some of such results. If we restrict ourselves, for example, to starlike domains and the interior boundary Σ_1 lies within the boundary Σ_2, then the solution Γ_1 of the free boundary problem for Σ_1 lies within the solution Γ_2 for Σ_2. This important monotonicity property helps to ensure well-posedness in case of randomness.

The mathematical formulation of Bernoulli's free boundary problem with a random interior boundary is given in Section 2. We choose the Dirichlet boundary value problem as the state equation and reformu-

DOI 10.20347/WIAS.PREPRINT.3121 Berlin 2024

late the problem under consideration as a shape optimization problem for the state's Dirichlet energy. That way, a variational formulation of the desired Neumann boundary condition at the free boundary is derived. We then intend to minimize the mean of the energy functional. Although the problem under consideration is well-posed in the deterministic case (see e.g. [22, 23]), the present random shape optimization problem is not, in general. This fact is motivated in Section 3 by an analytical example with circular boundaries. Indeed, the random interior boundary has to lie almost surely within some sufficiently narrow concentric annulus to ensure that the sought free boundary does not intersect this annulus, which would imply a degenerated situation. For the sake of simplicity, we will consider circular annuli throughout the rest of this article.

Section 4 is then concerned with shape calculus in the case of the deterministic free boundary problem. We provide the shape gradient and shape Hessian of the energy functional under consideration for general boundaries. Then, we study the convexity of the shape optimization problem under consideration. We are able to prove $H^{1/2}$-convexity for all convex exterior boundaries that are sufficiently smooth, lie in a fixed annulus, and have a uniformly bounded curvature. This is the one of the main results of our article and the key to verifying convergence of the iterates of stochastic gradient method. Additionally, to the best of our knowledge, global convexity has never been derived in shape optimization for a specific problem before.

In Section 5, we prove convergence of the stochastic gradient method in a novel setting, namely, one involving the so-called *two-norm discrepancy*. The two-norm discrepancy is a well-known phenomenon in optimal control and may occur in the infinite-dimensional setting since not all norms are equivalent; see in particular [12, 37]. We note that, for shape optimization, convergence of approximation (deterministic) solutions with the two-norm discrepancy was already established in [23]. There, second-order sufficient conditions were used to ensure stability around a local optimum. Since we have in fact strong convexity for the free boundary problem, we are able to prove convergence to the unique minimum, even in the presence of uncertainty. This is a stronger result than can be expected in a typical shape optimization problem under uncertainty; we note that convergence of the stochastic gradient method was shown in the context of Riemannian manifolds in [31]. Due to the (geodesic) non-convexity of the unconstrained problem studied there, one can at most expect that the corresponding Riemannian gradient vanishes in the limit. The main difficulty in the analysis here is that the convexity for the energy functional is with respect to a weaker space than the one to which the exterior boundaries belong. We provide a complete proof of convergence of iterates to the unique solution with respect to the weaker norm in the almost sure sense. We explain why the typical convergence rates in expectation cannot be derived in the function space setting due to the two-norm discrepancy. On the other hand, the discretized sequence will yield the expected rates for strongly convex functions.

Numerical experiments are presented in Section 6 in order to validate the theoretical findings. For a random starlike interior boundary, we compute the solution of the present random version of Bernoulli's free boundary problem. We observe very fast convergence towards the correct shape of the sought free boundary, which is a huge improvement over previously studied methods such as the use of sampling methods to compute the expected shape functional and its gradient or the direct computation of an appropriate expectation of the free boundary. In all, we observe a rate of convergence with respect to necessary optimality condition that is inverse proportional to the square root of the number of iterations. The rate of convergence with respect to the objective function values is inverse proportional to the number of iterations, as predicted by the theory.

Throughout this article, for $D \subset \mathbb{R}^2$ being a sufficiently smooth domain, we denote the space of square integrable functions by $L^2(D)$. For a nonnegative real number $s \geq 0$, the associated Sobolev spaces are labelled by $H^s(D) \subset L^2(D)$. Especially, there holds $H^0(D) = L^2(D)$. Moreover, when $s \geq$

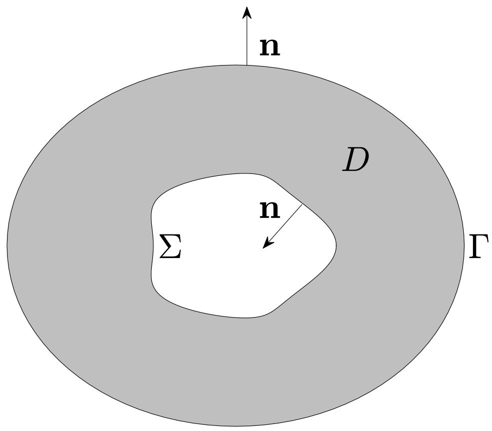

Figure 1: The geometrical setup: The annular domain D with the given interior boundary Σ and the free exterior boundary Γ.

$1/2$, the respective Sobolev spaces on the boundary ∂D are defined as the traces $H^{s-1/2}(\partial D) := \gamma(H^s(D))$ of the Sobolev spaces $H^s(D)$. The set of k-times differentiable functions is denoted by $\mathcal{C}^k$ and $\mathcal{C}^{k,\alpha}$ denotes the set of functions in $\mathcal{C}^k$ whose k-th order partial derivatives are additionally α-Hölder continuous.

2 Problem setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. In this article, we consider Bernoulli's free boundary problem when the interior boundary is random. The precise meaning of the random boundary will be specified later on. Given an event $\omega \in \Omega$, we are thus looking for an annular domain $D = D(\omega) \subset \mathbb{R}^2$ with interior boundary $\Sigma = \Sigma(\omega)$ and unknown deterministic exterior boundary Γ such that the function $u = u(\omega) \in H^1\big(D(\omega)\big)$ satisfies the following Dirichlet boundary value problem for the Laplacian

$$\begin{aligned} \Delta u &= 0 \quad \text{in} \quad D(\omega), \\ u &= 1 \quad \text{on} \quad \Sigma(\omega), \\ u &= 0 \quad \text{on} \quad \Gamma, \end{aligned} \tag{1}$$

with the additional flux condition

$$-\frac{\partial u}{\partial \boldsymbol{n}} = \lambda \quad \text{on} \quad \Gamma, \tag{2}$$

where $\lambda > 0$ is a given constant. Here and in the following, $\boldsymbol{n} = \boldsymbol{n}(\omega)$ denotes the exterior unit normal to $D(\omega)$. Note that the geometrical setup is illustrated in Figure 1.

Bernoulli's free boundary problem arises in many applications, for example in ideal fluid dynamics, optimal design, electrochemistry, or electrostatics. Generally speaking, Bernoulli's free boundary problem is an overdetermined partial differential problem since, for fixed boundaries $\Sigma(\omega)$ and Γ, the unknown harmonic function $u = u(\omega)$ has to vanish at the outer boundary and also to satisfy a flux condition (2). However, it becomes solvable when the *free boundary* Γ is also considered as an unknown. We refer the reader to e.g. [3, 24, 25] and the references therein for further details.

For any fixed realization $\Sigma(\omega)$ of the interior boundary, where $\omega \in \Omega$, it is well-known that a variational formulation for the sought boundary Γ such that the overdetermined (deterministic) boundary value problem (1) and (2) admits a solution is given by

$$\underset{\Gamma\subset\mathbb{R}^2}{\text{minimize}} \quad J(\Gamma,\Sigma(\omega)) = \int_{D(\omega)} \|\nabla u(\omega)\|_2^2 + \lambda^2 \,\mathrm{d}x = \int_{\Sigma(\omega)} \frac{\partial u}{\partial \boldsymbol{n}} \,\mathrm{d}s + \lambda^2 |D(\omega)|. \tag{3}$$

where the state $u = u(\omega)$ is the solution to (1). Uniqueness and existence of solutions to this free boundary problem follows from the seminal work [3].

The free boundary defined as the minimizer of the shape optimization problem (3) depends on the particular random event $\omega \in \Omega$. Therefore, in order to get a deterministic free boundary Γ while accounting for all possibilities of $\Sigma(\cdot)$, we shall consider the minimization (with respect to the boundary Γ) of the expected functional

$$\underset{\Gamma\subset\mathbb{R}^2}{\text{minimize}} \quad \mathbb{E}[J(\Gamma,\Sigma(\cdot))] = \int_\Omega \int_{D(\omega)} \|\nabla u(\omega)\|_2^2 + \lambda^2 \,\mathrm{d}x \,\mathrm{d}\mathbb{P}(\omega) \tag{4}$$

with the state $u = u(\omega)$ given by (1). Note that this is a free boundary problem, where the underlying domain D is random. Although the pointwise solution (3) is well-defined, this does not necessarily hold true for (4). We motivate this fact in the next section.

3 Analytical computations in the case of concentric annuli

Calculations can be performed analytically if the interior boundary Σ is a circle around the origin with radius r_Σ. Then, due to symmetry, the free boundary Γ will also be a circle around the origin with unknown radius r_Γ.

Using polar coordinates and making the ansatz $u(r,\theta) = y(r)$, we find $\Delta u(r,\theta) = y''(r) + y'(r)/r$. Hence, the solution with respect to the prescribed Dirichlet boundary condition of (1) in the case of dimension two is given by

$$y(r) = \frac{\log\left(\frac{r}{r_\Gamma}\right)}{\log\left(\frac{r_\Sigma}{r_\Gamma}\right)}.$$

The desired Neumann boundary condition at the free boundary r_Γ yields the equation

$$-y'(r_\Gamma) = \frac{1}{r_\Gamma \log\left(\frac{r_\Gamma}{r_\Sigma}\right)} = \lambda,$$

which can be solved by means of Lambert's W-function:

$$r_\Gamma = F(r_\Sigma) := \frac{1}{\lambda W\left(\frac{1}{\lambda r_\Sigma}\right)}.$$

Let us recall that Lambert's W-function is the inverse of $x \mapsto xe^x$. It is a non-decreasing function on $(0,+\infty)$ which, however, provides a non-analytic expression.

Since the Neumann data of u on the interior free boundary r_Σ are given by

$$-y'(r_\Sigma) = \frac{\partial u}{\partial \boldsymbol{n}}\Big|_\Sigma = \frac{1}{r_\Sigma \log\left(\frac{r_\Gamma}{r_\Sigma}\right)},$$

we conclude

$$J(r_\Gamma, r_\Sigma) = \frac{2\pi}{\log\left(\frac{r_\Gamma}{r_\Sigma}\right)} + \pi\lambda^2(r_\Gamma^2 - r_\Sigma^2),$$

compare (3). One readily verifies that this functional has indeed the unique minimizer $r_\Gamma = F(r_\Sigma)$.

Let us next consider the case where r_Σ switches randomly between $r_{\Sigma,1}$ and $r_{\Sigma,2}$ with the probability $\mathbb{P}(r_\Sigma = r_{\Sigma,1}) = p$ and $\mathbb{P}(r_\Sigma = r_{\Sigma,2}) = 1 - p$, where $p \in [0, 1]$. If we choose $r_{\Sigma,2}$ such that it satisfies $r_{\Sigma,2} > F(r_{\Sigma,1})$, then we obviously obtain the inequality chain

$$0 < r_{\Sigma,1} < F(r_{\Sigma,1}) < r_{\Sigma,2} < F(r_{\Sigma,2}) < \infty. \tag{5}$$

The expected functional reads

$$\mathbb{E}[J(r_\Gamma, r_\Sigma(\cdot))] = \frac{2\pi p}{\log\left(\frac{r_\Gamma}{r_{\Sigma,1}}\right)} + \frac{2\pi(1-p)}{\log\left(\frac{r_\Gamma}{r_{\Sigma,2}}\right)} + \pi\lambda^2\big(r_\Gamma^2 - p r_{\Sigma,1}^2 - (1-p) r_{\Sigma,2}^2\big).$$

Its unique minimizer is $r_\Gamma = F(r_{\Sigma,1})$ if $p = 1$ while it is $r_\Gamma = F(r_{\Sigma,2})$ if $p = 0$. In view of (5), this means that r_Γ has to cross $r_{\Sigma,2}$ during the transition from $p = 0$ to $p = 1$. However, this is impossible since then the domain $D(\omega)$ is not well-defined anymore as $r_\Sigma(\omega) < r_\Gamma$ is violated. Therefore, it is required to impose an inequality constraints to the sought boundary r_Γ, demanding that $r_\Gamma \geq \delta + r_\Sigma(\omega)$ for $\mathbb{P}$-almost all $\omega \in \Omega$ ("almost surely") and some $\delta > 0$.

The above observations motivate the assumption that $\Sigma(\omega)$ lies inside some annulus such that

$$B(0, \underline{r}_\Sigma) \subset \Sigma(\omega) \subset B(0, \overline{r}_\Sigma) \quad \text{almost surely.} \tag{6}$$

Thus, it follows from [7, 44] that the resulting free exterior boundary $\Gamma = \Gamma(\omega)$ satisfies

$$B(0, \underline{r}_\Gamma) \subset \Gamma(\omega) \subset B(0, \overline{r}_\Gamma) \quad \text{almost surely,}$$

provided that the interior domain surrounded by $\Sigma(\omega)$ is starlike. In order to ensure well-posedness in our subsequent analysis, we restrict ourselves to starlike interior boundaries satisfying (6), where $\underline{r}_\Sigma$ and $\overline{r}_\Sigma$ are such that $\overline{r}_\Sigma \leq \underline{r}_\Gamma$.

4 Properties of the objective with respect to shape variations

We shall now focus on the particular case where the free boundary Γ is the boundary of a convex domain. Indeed, it is known that the solution to Bernoulli's free boundary problem is starlike if the interior boundary is; see [44]. If the interior boundary is even convex, then the exterior one is also convex [36]. However, the exterior boundary can also be convex although the interior is not.

Our main result in this section is the convexity of the objective with respect to an appropriate norm. Of course, since shape spaces are not linear, convexity has to be understood in terms of a parameterization of the boundary Γ. Among the many ways to parameterize such a curve, we discuss the case of the parameterization with respect to a point, then with the support function. Since they do not recover exactly the same geometric perturbations, the respective second-order derivative has different properties.

4.1 Shape sensitivity analysis

We first consider general geometries and perturbations and we compute the first and second order shape derivatives of the objective around a given boundary Γ for general perturbations. To this end, we choose the interior boundary Σ arbitrary but fixed and suppress its explicit dependence in the objective J. Throughout this section, the outer boundary Γ is always such that it encloses the interior boundary Σ to ensure that the annular domain D in between is well-defined.

We first study the dependence with respect to the outer boundary Γ and consider sufficiently regular deformation fields $\boldsymbol{V}$ that are defined in the neighborhood of the exterior boundary Γ. As we need regularity on the shapes, let us assume for convenience that Γ is of class $\mathcal{C}^2$ and that the deformation field $\boldsymbol{V}$ has the same regularity. While the first-order derivative of the shape functional under consideration has already been calculated and used many times in the literature, the second-order derivative has only been studied at a critical point for stability analyses (see [22], for example). Here, we will study its expression for domains D where the gradient does not vanish. For a comprehensive introduction to shape calculus, we refer the reader to [21, 35, 43].

Lemma 1. *Let the boundaries Γ and Σ be of class $\mathcal{C}^2$. Then, the first- and second-order shape derivatives for of the objective are given by*

$$DJ(\Gamma)[\boldsymbol{V}] = \int_\Gamma \boldsymbol{V}_{\boldsymbol{n}} \left[\lambda^2 - \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\right] \mathrm{d}s \tag{7}$$

and

$$D^2J(\Gamma)[\boldsymbol{V},\boldsymbol{V}] = \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}} u' + H\lambda^2 \boldsymbol{V}_{\boldsymbol{n}}^2 + \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2 \boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}\boldsymbol{V}_{\boldsymbol{n}} \,\mathrm{d}s, \tag{8}$$

where we used the abbreviation $\boldsymbol{V}_{\boldsymbol{n}} = \boldsymbol{V}\cdot\boldsymbol{n}$ and where $\nabla_{\boldsymbol{\tau}}$ denotes the surface gradient with respect to the boundary Γ.

Proof. By usual arguments given in e.g. [21, 35, 43], the solution u of (1) has the following first- and second-order derivatives u' and u'' that are characterized by differentiating the boundary condition on the boundary Γ:

$$\begin{aligned} \Delta u' &= 0 \quad \text{in} \quad D,\\ u' &= 0 \quad \text{on} \quad \Sigma,\\ u' &= -\frac{\partial u}{\partial \boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{n}} \quad \text{on} \quad \Gamma, \end{aligned}$$

and

$$\begin{aligned} \Delta u'' &= 0 \quad \text{in} \quad D,\\ u'' &= 0 \quad \text{on} \quad \Sigma,\\ u'' &= -\left[\frac{\partial u'}{\partial \boldsymbol{n}} + \nabla u\cdot\boldsymbol{n}'\right]\boldsymbol{V}_{\boldsymbol{n}} - \frac{\partial u}{\partial \boldsymbol{n}}\boldsymbol{V}\cdot\boldsymbol{n}'\\ &= -\frac{\partial u'}{\partial \boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{n}} + \frac{\partial u}{\partial \boldsymbol{n}}\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}(\boldsymbol{V}_{\boldsymbol{n}}) \quad \text{on} \quad \Gamma, \end{aligned}$$

where we used that $\boldsymbol{n}' = -\nabla_{\boldsymbol{\tau}}(\boldsymbol{V}_{\boldsymbol{n}})$, hence $\nabla u\cdot\boldsymbol{n}' = 0$. Therefore, we immediately arrive at

$$DJ(\Gamma)[\boldsymbol{V}] = \int_\Gamma \boldsymbol{V}_{\boldsymbol{n}} \left[\lambda^2 - \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\right] \mathrm{d}s$$

and

$$\begin{aligned} D^2J(\Gamma)[\boldsymbol{V},\boldsymbol{V}] &= \int_\Gamma u''\frac{\partial u}{\partial \boldsymbol{n}} + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 \,\mathrm{d}s \\ &= \int_\Gamma \left(-\frac{\partial u'}{\partial \boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{n}} + \frac{\partial u}{\partial \boldsymbol{n}}\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}(\boldsymbol{V}_{\boldsymbol{n}})\right)\frac{\partial u}{\partial \boldsymbol{n}} + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 \,\mathrm{d}s \\ &= \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}}\left(-\frac{\partial u}{\partial \boldsymbol{n}}\boldsymbol{V}_{\boldsymbol{n}}\right) + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 + \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}(\boldsymbol{V}_{\boldsymbol{n}}) \,\mathrm{d}s \\ &= \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}}u' + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 \,\mathrm{d}s + \int_\Gamma \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}\boldsymbol{V}_{\boldsymbol{n}} \,\mathrm{d}s. \end{aligned}$$

□

4.2 On the sign of the shape Hessian

In order to the study the sign of the shape Hessian, we split it into two terms

$$D^2J(\Gamma)[\boldsymbol{V},\boldsymbol{V}] = I_1(\boldsymbol{V}) + I_2(\boldsymbol{V}),$$

where we set

$$I_1(\boldsymbol{V}) = \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}}u' + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 \,\mathrm{d}s \quad \text{and} \quad I_2(\boldsymbol{V}) = \int_\Gamma \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}\boldsymbol{V}_{\boldsymbol{n}} \,\mathrm{d}s.$$

Notice that the term $\boldsymbol{V}\cdot\nabla_{\boldsymbol{\tau}}\boldsymbol{V}_{\boldsymbol{n}}$ appears in the former expressions as expected by the structure theorems of second order shape derivatives (see [35, Theorem 5-9-2, page 220] and [16, Theorem 2-1]) since we are not at the optimum and because we do not restrict ourselves to normal perturbations. According to the structure of the shape Hessian, the first term I_1 is a quadratic form in $\boldsymbol{V}_{\boldsymbol{n}}$. The second term I_2 is a remainder of the shape gradient. It is bilinear in $\boldsymbol{V}$ but also involves tangential derivatives. Consequently, finding the sign of the shape Hessian requires studying the two terms separately.

4.2.1 On the sign of I_1

Recall the we assume that Γ is the boundary of a convex set. Hence, its curvature H is nonnegative and we obtain after integration by parts

$$I_1(\boldsymbol{V}) = \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}}u' + H\lambda^2\boldsymbol{V}_{\boldsymbol{n}}^2 \,\mathrm{d}s \geq \int_\Gamma \frac{\partial u'}{\partial \boldsymbol{n}}u' \,\mathrm{d}s = \int_D \|\nabla u'\|_2^2 \,\mathrm{d}x > 0,$$

i.e., the integral I_1 is clearly positive. Since $u' = 0$ on the component Σ of the boundary of D, we then get by Poincaré's inequality and the trace theorem that

$$I_1(\boldsymbol{V}) = \int_D \|\nabla u'\|_2^2 \,\mathrm{d}x \geq C_P(D)\|u'\|_{H^{1/2}(\Gamma)}^2, \tag{9}$$

where $C_P(D)$ is the Poincaré constant of the domain D with homogeneous boundary condition on Σ. Abbreviating $\partial_{\boldsymbol{n}}u = (\partial u)/(\partial\boldsymbol{n})$, an immediate first lower bound is thus

$$I_1(\boldsymbol{V}) \geq C_P(D)\|u'\|_{L^2(\Gamma)}^2 = C_P(D)\|(\partial_{\boldsymbol{n}}u)\boldsymbol{V}_{\boldsymbol{n}}\|_{L^2(\Gamma)}^2 \geq C_P(D)(\inf_\Gamma \partial_{\boldsymbol{n}}u)^2\|\boldsymbol{V}_{\boldsymbol{n}}\|_{L^2(\Gamma)}^2.$$

Here, we have used the strong maximum principle to ensure that $\inf_\Gamma \partial_{\boldsymbol{n}}u > 0$.

In fact, we can have a more precise lower bound in the Sobolev norm $H^{1/2}(\Gamma)$. To that end, we use the following lemma to estimate the Sobolev norm of the product $u' = -(\partial_{\boldsymbol{n}}u)\boldsymbol{V}_{\boldsymbol{n}}$ from below.

Lemma 2. *If $f \in H^{1/2}(\Gamma)$ and $g \in H^{1/2+\epsilon}(\Gamma)$, then there exists some $C > 0$ such that*

$$\|fg\|_{H^{1/2}(\Gamma)} \leq C\|f\|_{H^{1/2}(\Gamma)}\, \|g\|_{H^{1/2+\epsilon}(\Gamma)}. \tag{10}$$

If there exists some $a > 0$ such that $g \geq a$ on Γ, then there exists some $C > 0$ such that

$$\|f\|_{H^{1/2}(\Gamma)} \leq C\|fg\|_{H^{1/2}(\Gamma)}\, \|1/g\|_{H^{1/2+\epsilon}(\Gamma)}. \tag{11}$$

Proof. The key ingredient is the following product estimate in Sobolev spaces taken from [6, Lemma 7-2]: If $F \in H^1(D)$ and $G \in H^{1+\epsilon}(D)$ for some $\epsilon > 0$, then the product satisfies $FG \in H^1(D)$ and we have

$$\|FG\|_{H^1(D)} \leq C\|F\|_{H^1(D)}\|G\|_{H^{1+\epsilon}(D)}. \tag{12}$$

We now translate this estimate to the trace space on the boundary Γ. To this end, set $f \in H^{1/2}(\Gamma)$ and $g \in H^{1/2+\epsilon}(\Gamma)$. Let F and G be harmonic extensions of f and g to D so that by Dirichlet's principle $\|f\|_{H^{1/2}(\Gamma)} = \|F\|_{H^1(D)}$ and $\|g\|_{H^{1/2+\epsilon}(\Gamma)} = \|G\|_{H^{1+\epsilon}(D)}$, respectively. Then, in view of (12), we first get

$$\|FG\|_{H^1(D)} \leq C\|F\|_{H^1(D)}\, \|G\|_{H^{1+\epsilon}(D)} = C\|f\|_{H^{1/2}(\Gamma)}\, \|g\|_{H^{1+\epsilon}(\Gamma)},$$

and then by the definition of the trace norm

$$\|fg\|_{H^{1/2}(\Gamma)} = \inf_{\substack{\phi\in H^1(D)\\ \phi=fg \text{ on } \Gamma}} \|\phi\|_{H^1(D)} \leq \|FG\|_{H^1(D)} \leq C\|f\|_{H^{1/2}(\Gamma)}\, \|g\|_{H^{1/2+\epsilon}(\Gamma)}.$$

Assume now that g satisfies the additional property that there exists some $a > 0$ such that $g \geq a$ on Γ. Consider the function I_a defined on $(0, +\infty) \to \mathbb{R}$ by $I_a(t) = 1/a$ if $t \leq a$ and by $I_a(t) = 1/t$ otherwise. Obviously, this is a bounded Lipschitz function. We notice that $1/g = I_a \circ g$ and hence $1/g$ belongs to $H^{1/2+\epsilon}(\Gamma)$. As a consequence, since $f = (fg)\,(1/g)$, we obtain by the product estimate (10) that

$$\|f\|_{H^{1/2}(\Gamma)} \leq C\|fg\|_{H^{1/2}(\Gamma)}\, \|1/g\|_{H^{1/2+\epsilon}(\Gamma)}.$$

□

With the help of this lemma, we obtain the following result.

Lemma 3. *There exists a constant $C > 0$ depending on Γ such that*

$$I_1(\boldsymbol{V}) \geq C\|\boldsymbol{V}_{\boldsymbol{n}}\|^2_{H^{1/2}(\Gamma)}.$$

Proof. Under our regularity assumptions on the boundaries Σ and Γ, we can apply (11) for $f = \boldsymbol{V}_{\boldsymbol{n}}$ and $g = \partial_{\boldsymbol{n}} u$ so that $u' = -fg$. The lower bound on g comes from the strong maximum principle and the compactness of Γ. Hence, we have

$$\|\boldsymbol{V}_{\boldsymbol{n}}\|_{H^{1/2}(\Gamma)} \leq C\|\boldsymbol{V}_{\boldsymbol{n}}\partial_{\boldsymbol{n}} u\|_{H^{1/2}(\Gamma)}\, \|1/(\partial_{\boldsymbol{n}} u)\|_{H^{1/2+\epsilon}(\Gamma)}.$$

The claim then follows from (9). □

4.2.2 On the sign of I_2

The sign of the second term

$$I_2(\boldsymbol{V}) = \int_\Gamma \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2 \boldsymbol{V} \cdot \nabla_{\boldsymbol{\tau}} \boldsymbol{V}_{\boldsymbol{n}} \, \mathrm{d}s$$

is less clear since it has the sign of the purely geometric term $\boldsymbol{V} \cdot \nabla_{\boldsymbol{\tau}} \boldsymbol{V}_{\boldsymbol{n}}$. Indeed, the sign of that term and hence of I_2 depends on the specific class of perturbations under consideration.

The natural parameterization of convex domains is the one using support functions. We restrict ourselves to perturbations of a convex domain that preserve convexity. We then check that in this situation I_2 takes only nonnegative values.

Recently, shape calculus for convex domains based on the Minkowski sum and therefore on support functions was developed in [9, 10]). For our purposes, however, it suffices to use simpler tools. To this end, let us recall the definition of the support function and its main properties. Convex sets $K \subset \mathbb{R}^d$ are parameterized by their support function h_K defined on $\mathbb{R}^d$ by

$$h_K(x) = \sup\{x \cdot y \mid y \in K\}.$$

The monotonicity property $K_1 \subset K_2 \Rightarrow h_{K_1} \leq h_{K_2}$ is clear from this definition. In particular, for nonnegative real numbers $a < b$, we have

$$B(0,a) \subset K \subset B(0,b) \Rightarrow a \leq h_K \leq b.$$

The support function of a convex set is homogeneous of degree one and hence can be restricted to the unit sphere $\mathbb{S}^{d-1}$ without loss of generality. To simplify notation, we are still abusively calling this restriction h_K.

Let us introduce the parameterization mapping $\boldsymbol{\Upsilon}$ defined over the set $\mathfrak{K}^d$ of convex domains in $\mathbb{R}^d$ by

$$\boldsymbol{\Upsilon} : \mathfrak{K}^d \to \mathcal{C}^0(\mathbb{S}^{d-1}), \quad K \mapsto h_K.$$

A crucial property is the isometric connection between the Hausdorff distance and the L^∞-norm on $\mathcal{C}^0(\mathbb{S}^{d-1}, \mathbb{R})$: for all $K_1, K_2 \in \mathfrak{K}^d$,

$$d_{\mathfrak{H}}(K_1, K_2) = \|h_{K_1} - h_{K_2}\|_{L^\infty(\mathbb{S}^{d-1})}. \tag{13}$$

Reconstructing a convex set from a support function can be performed using the envelope operator

$$\mathcal{E} \colon \mathcal{C}^1(\mathbb{S}^{d-1}, \mathbb{R}) \to \mathcal{C}^1(\mathbb{S}^{d-1}, \mathbb{R}^d), \quad h \mapsto \mathcal{E}[h],$$

defined for all $x \in \mathbb{S}^{d-1}$ by

$$\mathcal{E}[h](x) = h(x)x + \nabla_{\boldsymbol{\tau}} h(x).$$

This operator allows to reconstruct a convex set whose restricted support is h. Notice that this point has been investigated in the works of Antunes and Bogosel [4, 8].

In the planar case, one gets simply a periodic function $h \colon [0, 2\pi] \to \mathbb{R}$ and a parameterization of a set whose support function h is

$$\mathcal{E}[h] : \theta \mapsto h(\theta)\boldsymbol{e}_r(\theta) + h'(\theta)\boldsymbol{e}_\theta(\theta).$$

Here and in the following, $\boldsymbol{e}_r(\theta) = (\cos\theta, \sin\theta)$ denotes the radial direction and $\boldsymbol{e}_\theta(\theta) := \boldsymbol{e}_r(\theta)' = (-\sin\theta, \cos\theta) \perp \boldsymbol{e}_r$. Thus, we get

$$\begin{aligned} \mathcal{E}[h]'(\theta) &= h'(\theta)\boldsymbol{e}_r(\theta) + h(\theta)\boldsymbol{e}_\theta(\theta) + h''(\theta)\boldsymbol{e}_\theta(\theta) - h'(\theta)\boldsymbol{e}_r(\theta) \\ &= \big(h(\theta) + h''(\theta)\big)\boldsymbol{e}_\theta(\theta). \end{aligned} \tag{14}$$

Therefore, the unit tangent vector $\boldsymbol{\tau}$ at $\mathcal{E}[h](\theta)$ is $\boldsymbol{e}_\theta(\theta)$ and the unit outward normal vector $\boldsymbol{n}$ is then $\boldsymbol{e}_r(\theta)$. A perturbation q of the support function generates the support function $h+tq$ for $|t|$ sufficiently small and thus the parameterization

$$\mathcal{E}[h+q](\theta) = (h+tq)(\theta)\boldsymbol{e}_r(\theta) + (h+tq)'(\theta)\boldsymbol{e}_\theta(\theta).$$

Therefore, the deformation field is

$$\boldsymbol{V} = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[h+tq] = q\boldsymbol{e}_r + q'\boldsymbol{e}_\theta$$

and we conclude $\boldsymbol{V}_{\boldsymbol{n}} = q$. Notice that this expression also makes sense in the neighborhood of the curve Γ. We can next compute directly the gradient and observe that it is tangent to the curve

$$\nabla \boldsymbol{V}_{\boldsymbol{n}} = \frac{q'}{\sqrt{h^2 + (h')^2}}\boldsymbol{e}_\theta = \nabla_{\boldsymbol{\tau}} \boldsymbol{V}_{\boldsymbol{n}},$$

which implies

$$\boldsymbol{V} \cdot \nabla_{\boldsymbol{\tau}} \boldsymbol{V}_{\boldsymbol{n}} = \frac{(q')^2}{\sqrt{h^2 + (h')^2}} \geq 0.$$

Since $h > 0$ (indeed, we need a disk in the inner domain here that is uniformly greater than zero), we have herewith shown that $I_2[\boldsymbol{q}] \geq 0$.

Remark 4. *Under radial deformations of convex domains, I_2 has no sign. In the case of starlike domains, the outer boundary Γ can be parameterized by $\gamma\boldsymbol{e}_r$, where $\gamma : [0, 2\pi] \to (0, +\infty)$ denotes the radial function and $\boldsymbol{e}_r = (\cos\theta, \sin\theta)$ is the radial direction. Then, the unit tangent vector $\boldsymbol{\tau}$ and the unit outward normal vector $\boldsymbol{n}$ are given by the formulae*

$$\boldsymbol{\tau} = \frac{1}{\sqrt{\gamma^2 + (\gamma')^2}}\left(\gamma'\boldsymbol{e}_r + \gamma\boldsymbol{e}_\theta\right) \quad \textit{and} \quad \boldsymbol{n} = \frac{1}{\sqrt{\gamma^2 + (\gamma')^2}}\left(\gamma\boldsymbol{e}_r - \gamma'\boldsymbol{e}_\theta\right).$$

Thus, the normal component of any boundary deformation field of the type $\boldsymbol{V} = \varphi\boldsymbol{e}_r$ is

$$\boldsymbol{V}_{\boldsymbol{n}} = \boldsymbol{V} \cdot \boldsymbol{n} = \frac{\gamma\varphi}{\sqrt{\gamma^2 + (\gamma')^2}}.$$

Hence, we find

$$\nabla(\boldsymbol{V} \cdot \boldsymbol{n}) = \frac{1}{\gamma}\left(\frac{\gamma\varphi}{\sqrt{\gamma^2 + (\gamma')^2}}\right)'\boldsymbol{e}_\theta$$

and thus

$$\nabla_{\boldsymbol{\tau}}(\boldsymbol{V} \cdot \boldsymbol{n}) = \frac{1}{\gamma^2 + (\gamma')^2}\left(\frac{\gamma\varphi}{\sqrt{\gamma^2 + (\gamma')^2}}\right)'\left(\gamma'\boldsymbol{e}_r + \gamma\boldsymbol{e}_\theta\right).$$

Consequently, the term $\boldsymbol{V} \cdot \nabla_{\boldsymbol{\tau}} \boldsymbol{V}_{\boldsymbol{n}}$ is given by

$$\begin{aligned}\boldsymbol{V} \cdot \nabla_{\boldsymbol{\tau}}(\boldsymbol{V} \cdot \boldsymbol{n}) &= \frac{\gamma'\varphi}{\gamma^2 + (\gamma')^2}\left(\frac{\gamma\varphi}{\sqrt{\gamma^2 + (\gamma')^2}}\right)' \\ &= \frac{(\gamma'\varphi)^2}{\left(\gamma^2 + (\gamma')^2\right)^{3/2}} + \frac{(\gamma^2)'}{2\sqrt{\gamma^2 + (\gamma')^2}}\left(\frac{\varphi^2}{\gamma^2 + (\gamma')^2}\right)'.\end{aligned}$$

Obviously, this previous expression has no sign since φ is arbitrary.

4.2.3 Restricting the objective to a class of domains to get a strongly convex one

The previous investigation of the shape Hessian motivates the study of the free boundary problem under consideration in the class of convex domains, parameterized by means of the support function. To this end, we identify the boundary Γ with its support function and set $\mathcal{J}(h) = J\big(\mathcal{E}[h]\big)$. In view of the above results and translating them in terms of support function (see [9, 10]), we have proven that

$$D\mathcal{J}(h)[q] = \int_\Gamma q(\theta_{\boldsymbol{n}})\left[\lambda^2 - \left(\frac{\partial u}{\partial \boldsymbol{n}}\right)^2\right] \mathrm{d}s$$

while for the shape Hessian one gets

$$D^2\mathcal{J}(h)[q] \geq C(h)\|q(\theta_{\boldsymbol{n}})\|^2_{H^{1/2}(\Gamma)} \tag{15}$$

by combining the estimates on I_1 and I_2. Herein, $\theta_{\boldsymbol{n}}$ is the angle $\theta \in [0, 2\pi]$ that is imposed by the normal vector, i.e., $\boldsymbol{n} = (\cos\theta_{\boldsymbol{n}}, \sin\theta_{\boldsymbol{n}})$. This is a convexity result but a weak one. Its main weakness is its non-uniformity with respect to the design variable. Moreover, it is not formulated in a differentiation norm.

4.3 Uniform lower bounds of the Hessian

We shall next study when there exists a uniform bound that is independent of Γ. To that end, we proceed with the following strategy. First, we introduce a parameterization of the family of domains under consideration by restricting ourselves to starlike boundaries Σ and Γ. The boundary value problem is first transported to a fixed annulus, resulting in a family of parameterized problems on that annulus. The local inversion theorem is employed to demonstrate the regularity of the map associating the boundaries to the solution of the parameterized boundary value problem. Subsequently, the boundaries are restricted to a compact context for the parameterization, allowing us to obtain uniform bounds.

In that spirit, for given positive numbers $\alpha \in (0, 1/2)$, $0 < r_{\underline{\Gamma}} < r_{\overline{\Gamma}} < M_\Gamma$, we consider the class $\mathcal{S}_\Gamma$ of periodic functions defined on $[0, 2\pi]$ by

$$\mathcal{S}_\Gamma = \{h \in \mathcal{C}^{3,2\alpha}_{per} \mid \forall\theta \in [0, 2\pi],\ r_{\underline{\Gamma}} \leq h(\theta) \leq r_{\overline{\Gamma}},\ (h + h'')(\theta) \geq 0, \text{ and } \|h\|_{\mathcal{C}^{3,2\alpha}} \leq M_\Gamma\}. \tag{16}$$

This is a compact subset of $\mathcal{C}^{3,\alpha}_{per}$ that parameterizes through the support function the class

$$\mathcal{K}_\Gamma = \{K \subset \mathbb{R}^2 \mid \exists h \in \mathcal{S}_\Gamma,\ \Gamma = \partial K = \mathcal{E}[h]\}, \tag{17}$$

of convex subsets of $\mathbb{R}^2$ with a $\mathcal{C}^{2,2\alpha}_{per}$ boundary between the two concentric circles of radii $r_{\underline{\Gamma}}$ and $r_{\overline{\Gamma}}$, respectively. Note that the sets $\mathcal{S}_\Gamma$ and $\mathcal{K}_\Gamma$ are convex and closed, respectively, as for any pair of functions $h_1, h_2 \in \mathcal{S}_\Gamma$, the convex combination $\lambda h_1 + (1-\lambda)h_2$ is also a member of the class $\mathcal{S}_\Gamma$ for all $\lambda \in (0,1)$.

With this notation at hand, we are now in the position to state the main result of this section.

Proposition 5. *Given positive numbers $\alpha \in (0,1/2)$, $0 < r_{\underline{\Gamma}} < r_{\overline{\Gamma}} < M_\Gamma$, there exists a positive number C depending on $r_{\underline{\Gamma}}$, $r_{\overline{\Gamma}}$, and M_Γ such that for all $h \in \mathcal{S}_\Gamma$,*

$$D^2\mathcal{J}(h)[q] \geq C\|q(\theta_{\boldsymbol{n}})\|^2_{H^{1/2}(\Gamma)}. \tag{18}$$

To prove Proposition 5, we check the uniform behavior of each constant in the successive inequalities we used. These are

- the Poincaré inequality in (9). The uniform bound follows from the geometric bounds of Γ.
- the product inequality (12). The uniform bound follows from the existence of uniform (with respect to $D \in \mathcal{K}_\Gamma$) extension operators for the Sobolev spaces $H^1(D)$ and $H^{1+\epsilon}(D)$ to the whole $H^1(\mathbb{R}^2)$ and $H^{1+\epsilon}(\mathbb{R}^2)$, induced by the upper bound for $\|h\|_{C^{3,2\alpha}}$.
- the equivalence between the trace norm and the intrinsic Sobolev norm for fractional Sobolev spaces on a boundary for the upper bound of $1/g$ by composition. Gagliardo has shown in [26] that the two different norms on $H^{1/2}$ are equivalent if the domains are uniformly Lipschitz.
- finally, the lower bound for the normal derivative $\partial_{\boldsymbol{n}} u$.

The latter item is less standard, hence we shall elaborate on it. The main difficulty we face here is to get a uniform lower bound of the normal derivative. Clearly, it is nonnegative thanks to the maximum principle. Nevertheless, by its own, this argument cannot provide a uniform lower bound. We need an additional ingredient: continuity and compactness with respect to the inner and outer boundaries.

We transform the boundary value problem with variable boundaries to a boundary value problem with fixed boundary but variable coefficients. The boundaries Σ and Γ are parameterized by $\sigma(\theta)\mathbf{e}_r$ and $\gamma(\theta)\mathbf{e}_r$, respectively. Here,

$$\sigma : [0,2\pi] \to [\underline{r}_\Sigma, \overline{r}_\Sigma], \quad \gamma : [0,2\pi] \to [\underline{r}_\Gamma, \overline{r}_\Gamma] \tag{19}$$

denote the associated radial functions of the interior and exterior boundaries and $\mathbf{e}_r = \big(\cos(\theta), \sin(\theta)\big)$ is the radial vector. Consider the annulus $\circledcirc$ with bounds $\overline{r}_\Sigma < \underline{r}_\Gamma$. Then, the map

$$\Phi : \circledcirc \to \Omega, \quad (r,\theta) \mapsto \left[\frac{r-\overline{r}_\Sigma}{\underline{r}_\Gamma - \overline{r}_\Sigma}\gamma(\theta) + \frac{\underline{r}_\Gamma - r}{\underline{r}_\Gamma - \overline{r}_\Sigma}\sigma(\theta)\right]\mathbf{e}_r$$

maps the annulus $\circledcirc$ one-to-one to the annular domain Ω described by the boundaries Σ and Γ.

Lemma 6. *The singular values of the Jacobian $\Phi'(r,\theta)$ are uniformly bounded from above and below for all (r,θ) from the annulus $\circledcirc$ provided that the parameterizations γ and σ satisfy* (19) *with uniformly bounded derivatives.*

Proof. We shall compute the Jacobian of the map Φ. With $\mathbf{e}_\theta = \big(-\sin(\theta), \cos(\theta)\big)$, we find

$$\begin{aligned}
\Phi'(r,\theta) &= \frac{\partial\Phi(r,\theta)}{\partial r}\mathbf{e}_r^\top + \frac{1}{r}\frac{\partial\Phi(r,\theta)}{\partial\theta}\mathbf{e}_\theta^\top \\
&= \frac{1}{\underline{r}_\Gamma - \overline{r}_\Sigma}[\gamma(\theta) - \sigma(\theta)]\mathbf{e}_r\mathbf{e}_r^\top + \frac{1}{r}\left[\frac{r-\overline{r}_\Sigma}{\underline{r}_\Gamma - \overline{r}_\Sigma}\gamma'(\theta) + \frac{\underline{r}_\Gamma - r}{\underline{r}_\Gamma - \overline{r}_\Sigma}\sigma'(\theta)\right]\mathbf{e}_r\mathbf{e}_\theta^\top \\
&\quad + \frac{1}{r}\left[\frac{r-\overline{r}_\Sigma}{\underline{r}_\Gamma - \overline{r}_\Sigma}\gamma(\theta) + \frac{\underline{r}_\Gamma - r}{\underline{r}_\Gamma - \overline{r}_\Sigma}\sigma(\theta)\right]\mathbf{e}_\theta\mathbf{e}_\theta^\top.
\end{aligned}$$

The Jacobian $\Phi'(r,\theta)$ is hence triangular with diagonal entries

$$a(r,\theta) := \frac{\gamma(\theta)-\sigma(\theta)}{\underline{r}_\Gamma - \overline{r}_\Sigma}, \quad b(r,\theta) := \frac{1}{r}\left[\frac{r-\overline{r}_\Sigma}{\underline{r}_\Gamma - \overline{r}_\Sigma}\gamma(\theta) + \frac{\underline{r}_\Gamma - r}{\underline{r}_\Gamma - \overline{r}_\Sigma}\sigma(\theta)\right].$$

and the off-diagonal entry

$$c(r,\theta) := \frac{1}{r}\left[\frac{r-\overline{r}_\Sigma}{\underline{r}_\Gamma - \overline{r}_\Sigma}\gamma'(\theta) + \frac{\underline{r}_\Gamma - r}{\underline{r}_\Gamma - \overline{r}_\Sigma}\sigma'(\theta)\right].$$

In view of

$$\underline{r}_\Sigma < \overline{r}_\Sigma \le r \le \underline{r}_\Gamma < \overline{r}_\Gamma$$

and

$$0 < \underline{r}_\Gamma - \overline{r}_\Sigma \le \gamma(\theta)-\sigma(\theta) \le \overline{r}_\Gamma - \underline{r}_\Sigma < \infty$$

for all (r,θ) from the annulus $\circledcirc$, the diagonal entries $a(r,\theta)$ and $b(r,\theta)$ are uniformly bounded from above and below for all annular domains Ω with starlike boundaries such that (19) holds. In addition, the modulus $|c(r,\theta)|$ of the off-diagonal entry is uniformly bounded from above if $\gamma'(\theta)$ and $\sigma'(\theta)$ are. Straightforward calculation verifies that consequently the singular values of the Jacobian $\Phi'(r,\theta)$ are uniformly bounded from above and below. $\square$

The boundary value problem (1) posed on Ω can be transformed to a boundary value problem in $\circledcirc$ by using the map Φ. There holds

$$\begin{aligned} \operatorname{div}(A\nabla u) &= 0 \quad \text{in} \quad \circledcirc,\\ u &= 1 \quad \text{on} \quad \|x\|_2 = \underline{r}_\Gamma,\\ u &= 0 \quad \text{on} \quad \|x\|_2 = \overline{r}_\Sigma, \end{aligned} \tag{20}$$

with the diffusion matrix A is given by

$$A(r,\theta) = \big(\Phi'(r,\theta)\big)^{-1}\big(\Phi'(r,\theta)\big)^{-\top}\det\big(\Phi'(r,\theta)\big).$$

Therefore, the diffusion matrix A depends on the parameterizations γ and σ of the inner and outer boundaries and on their first order derivatives. If both γ and σ are in the Hölder class $\mathcal{C}^{2,\alpha}$ for some given $\alpha \in (0,1)$, then the diffusion matrix A is $\mathcal{C}^{1,\alpha}$-smooth.

By the assumption $h \in \mathcal{S}_\Gamma$, there exists a positive real number $M > 0$ such that the diffusion matrix distribution belongs to the subset K_M of $\mathcal{C}^{1,\alpha}(\mathbb{R}^{2\times 2}_{sym})$ given by

$$K_M = \left\{A \in \mathcal{C}^{1,\alpha}(\mathbb{R}^{2\times 2}_{sym}) \text{ with } \frac{1}{M}\|\xi\|_2^2 \le A\xi\cdot\xi \le M\|\xi\|_2^2 \text{ and } \|DA\|_\infty \le M\right\}.$$

Using standard arguments (local inverse theorem and a priori Hölder bounds up to the boundary, see [32, Theorem 6-6, page 98], the solution map $A \mapsto v_A$ of the boundary value problem (20) is smooth and in particular continuous on the subset K_M of $\mathcal{C}^{1,\alpha}(\mathbb{R}^{2\times 2}_{sym})$. The Neumann boundary data of the transformed solution are computed by $A\nabla u\cdot\boldsymbol{n}$. In view of the uniform bounds on the singular values of Φ', we conclude the $L^\infty(0,2\pi)$-bound

$$0 < c \le \partial_{\boldsymbol{n}} u(\theta) \le C < \infty$$

for the Neumann derivative at the boundary Γ.

4.4 Different interior boundaries

So far, we have shown uniform bounds of the shape Hessian in the case of a fixed interior boundary Σ. In particular, all the constants in these bounds depend on this specific Σ. However, compactness arguments similar to those used for Γ also apply to Σ. Hence, if we assume that Σ is starlike with periodic $\mathcal{C}^{2,2\alpha}$-smooth parameterization $\sigma(\theta)\mathbf{e}_r$ such that $r_{\underline{\Sigma}} \leq \sigma(\theta) \leq r_{\overline{\Sigma}}$ for all $\theta \in [0, 2\pi]$ and $\|\sigma\|_{\mathcal{C}^{2,2\alpha}} \leq M_\Sigma$, then the uniform bounds still hold. In other words, we shall consider parameterizations from the set

$$\mathcal{S}_\Sigma := \{\sigma \in \mathcal{C}^{2,2\alpha}_{per} \,|\, \forall \theta \in [0, 2\pi],\ r_{\underline{\Sigma}} \leq \sigma(\theta) \leq r_{\overline{\Sigma}} \text{ and } \|\sigma\|_{\mathcal{C}^{2,2\alpha}} \leq M_\Sigma\}.$$

It can easily be shown that this set is also convex and closed. It is then sufficient to repeat the argument from the previous subsection (transporting the state equation from D onto $\circledcirc$ and using explicit formulas as a function of the parameterization of the interior boundary) to show that the state and then the objective are continuous with respect to the interior boundary, and that the coercivity constant of the shape Hessian with respect to the support function of the outer boundary can be chosen to be uniform with respect to the interior boundary.

Consequently, the functional $\mathcal{J}$ defined on $\mathcal{K}_\Gamma \times \mathcal{S}_\Sigma$ by

$$\mathcal{J}(h, \sigma) = \int_D \|\nabla u\|_2^2 + \lambda^2 \,\mathsf{d}x,$$

where the boundary of D has two connected components, the interior one being parameterized by the distance σ to the origin and the outer one through the support function h. This functional $\mathcal{J}$ is continuous and there exists a constant $c_E > 0$ such that for all $\sigma \in \mathcal{S}_\Sigma$ and all $h \in \mathcal{K}_\Gamma$ one has the estimate

$$D^2_{h,h}\mathcal{J}(h, \sigma)[q, q] \geq c_E \|q\|^2_{H^{1/2}} \quad \forall q \in H^{1/2}_{per}([0, 2\pi]). \tag{21}$$

To conclude this section, we define the random boundary Σ as the image by the polar parameterization of a vector-valued random variable $\sigma \in L^\infty(\Omega, \mathcal{S}_\Sigma)$, where the latter set is comprised of (strongly) $\mathbb{P}$-measurable functions σ from Ω to the closed and convex set $\mathcal{S}_\Sigma$ and satisfying $\operatorname{ess\,sup}_{\omega \in \Omega} \|\sigma(\omega)\|_{\mathcal{C}^{2,2\alpha}} < \infty$ for all $\sigma \in \mathcal{S}_\Sigma$. The continuity of the map $\mathcal{J}$ on $\mathcal{K}_\Gamma \times \mathcal{S}_\Sigma$ implies the measurability of the map $\omega \mapsto \mathcal{J}(h, \sigma(\omega))$.

5 On the stochastic gradient method with two-norm discrepancy

In this section, we prove convergence of the projected stochastic gradient method for an abstract setting involving the two-norm discrepancy. This classical method dating back to Robbins and Monro [41] involves randomly sampling the otherwise intractable gradient and has been well-investigated in the literature. For the function space setting without this discrepancy, the stochastic gradient method and its variants have already been analyzed; see [5, 14, 33, 45] and more recent contributions in the context of PDE-constrained optimization under uncertainty [27, 28, 29, 30, 38]. The setting we present in section 5.1 is adapted from [23], where convergence of a deterministic Ritz–Galerkin-type method was proven. We show that this framework fits the free boundary problem investigated in the previous sections, where it was established that the energy functional is coercive in a weaker space than where it is continuous. In section 5.2, we present the method, which involves a modification of the typical projected stochastic gradient iteration whereby a stochastic gradient is computed on the weaker space and a projection is performed onto the stronger space. We provide a complete proof of almost sure convergence of iterates to the unique solution with respect to the weaker norm.

5.1 Abstract setting

In this section, we summarize our numerical approach to solving the free boundary problem (1)–(2). Let $X \subset H$ be two Hilbert spaces, which are dense in L^2, endowed with the inner products $(\cdot,\cdot)_X$ and $(\cdot,\cdot)_H$, respectively, and corresponding norms $\|\cdot\|_X$ and $\|\cdot\|_H$. The respective dual spaces are denoted by X^* and H^*, which yields the Gelfand chain $X \subset H \subset L^2 \subset H^* \subset X^*$. A ball centered at r in a space X' is denoted by $B_\delta^{X'}(r) = \{h \in X' \mid \|h-r\|_{X'} < \delta\}$. We assume that $X \subset H$ with continuous embedding and that $X_{ad} \subset X$ is a bounded, closed, convex, and nonempty admissible set.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\xi\colon \Omega \to \Xi$ be a random function mapping to a (real) complete separable metric space Ξ. Now, consider the problem

$$\underset{h \in X_{ad}}{\text{minimize}} \quad \{j(h) = \mathbb{E}[\mathcal{J}(h,\xi)]\}, \tag{22}$$

where we assume that $j\colon O \subset X \to \mathbb{R}$ is twice continuously differentiable on an open set $O \supset X_{ad}$. A point h^* is said to be a *local solution* of (22) in X' if $j(h^*) \leq j(h)$ for all $h \in X_{ad} \cap B_\delta^{X'}(h^*)$ for some $\delta > 0$. A necessary condition for $h^* \in X_{ad}$ to be a local solution (in X) to (22) is given by the variational inequality

$$Dj(h^*)[h - h^*] \geq 0 \quad \forall h \in X_{ad}. \tag{23}$$

In the event that j is convex on X_{ad}, this condition is also sufficient and h^* is even a global solution. Notice that since H is the weaker space, any local solution of (22) in H is also a local solution in X; i.e., local solutions in H also satisfy the condition (23). Our strategy of handling the two-norm discrepancy will be to show that our method converges in H to a local solution.

In our application, it is only possible to show continuity and coercivity of $D^2 j$ on the weaker space H. In the abstract setting, this translates to the following assumption, which is motivated by [23].

Assumption 7. *The functional $j\colon O \to \mathbb{R}$ is twice continuously differentiable. We assume that for every $h \in X_{ad}$, the second derivative $D^2 j(h) \in \mathcal{L}(X, X^*)$ extends continuously to a bilinear form on $H \times H$, i.e., there exists $C_S > 0$ such that*

$$|D^2 j(h)[q_1, q_2]| \leq C_S \|q_1\|_H \|q_2\|_H \; \forall q_1, q_2 \in H. \tag{24}$$

Additionally, we assume that for there exists $c_E > 0$ such that the following strong coercivity condition is satisfied for every $h \in X_{ad}$:

$$D^2 j(h)[q,q] \geq c_E \|q\|_H^2 \quad \forall q \in H. \tag{25}$$

We note that this assumption does not require that the objective j extends continuously from X to H, as this is not satisfied by our problem. Twice continuous differentiability of j provides Lipschitz continuity of the corresponding gradient with respect to the (X^*, X)-duality. We note that since the second derivative is continuously extendable to a bilinear form on $H \times H$, we have Lipschitizianity of Dj on the weaker space, as was shown in [23].

The inequality (25) implies strong convexity with respect to H. Indeed, a Taylor expansion around any point $h' \in X_{ad}$ gives

$$j(h'+q) = j(h') + Dj(h')[q] + \tfrac{1}{2} D^2 j(\eta)[q,q]$$

with η being a point on the segment between h' and $h'+q$. Therefore,

$$j(h'+q) - j(h') \geq Dj(h')[q] + \frac{c_E}{2}\|q\|_H^2 \tag{26}$$

provided that $h' + q \in X_{ad}$. If $h' = h^*$ is a (local) optimum, we have using (23) that

$$j(h) - j(h^*) \geq \frac{c_E}{2}\|h - h^*\|_H^2 \quad \forall h \in X_{ad}. \tag{27}$$

In particular, h^* is a strict minimizer in X since $j(h) > j(h^*)$ for $h \neq h^*$. Since the coercivity space H differs from the stronger space X over which j is continuous, we cannot expect to have strong convexity with respect to the stronger space. On the other hand, it is possible in certain cases to show that a strict minimizer h^* in X is also one with respect to the weaker topology; see [12].

The free boundary problem We shall now look in more detail at how the free boundary problem fits into this framework. The space H is the fractional Sobolev space $H^{1/2}_{per}([0, 2\pi])$, the energy space of the shape Hessian. The smaller space X is the Sobolev space $H^4_{per}([0, 2\pi])$, since it is continuously embedded into $C^{3,2\alpha}_{per}([0, 2\pi])$ for all $\alpha \in (0, 1/4)$ by the Sobolev embedding theorem. The set X_{ad} is

$$X_{ad} = \{h \in X \mid \forall\theta \in [0, 2\pi] : r_{\underline{\Gamma}} \leq h(\theta) \leq r_{\overline{\Gamma}},\ (h + h'')(\theta) \geq 0, \text{ and } \|h\|_X \leq M_\Gamma\}.$$

Concerning Assumption 7, we recall that the inner boundary is modeled as random with $\xi = \sigma$. The objective is therefore

$$j(h) = \mathbb{E}[\mathcal{J}(h, \sigma(\cdot)].$$

The continuity estimate (24) is obvious by (8) as $\mathcal{K}_\Gamma \times \mathcal{S}_\Sigma$ is a compact set, see also [22]. For (25), we have from (21) that

$$D^2\mathcal{J}(h, \sigma(\cdot))[q, q] \geq c_E\|q\|_H^2 \quad \forall q \in H \text{ a.s.} \tag{28}$$

Using the fact that $\operatorname{ess\,sup}_{\omega\in\Omega}\|\sigma(\omega)\|_{\mathcal{C}^{2,2\alpha}} < \infty$, it is straightforward to argue that $D^2 j(h) = \mathbb{E}[D^2\mathcal{J}(h, \sigma(\cdot))]$. Applying the expectation on both sides of (28) yields (25).

5.2 Stochastic gradient method

Let $\pi_{X_{ad}} \colon X \to X_{ad}$ be a projection onto the set X_{ad}, defined by

$$\pi_{X_{ad}}(h) = \operatorname*{arg\,min}_{w \in X_{ad}} \|h - w\|_X,$$

which is well-defined and single-valued since $X_{ad} \subset X$ is assumed to be nonempty, closed, and convex. We assume that it is possible to compute an approximation of the gradient in the form of a *stochastic gradient* $G \colon X \times \Xi \to X$, which is defined as the (parameterized) Riesz representative of the mapping $D\mathcal{J}(\cdot, \hat{\xi}) \colon X \to X^*$, i.e., we have for every $\hat{\xi} \in \Xi$ that

$$(G(h, \hat{\xi}), q)_X = \langle D\mathcal{J}(h, \hat{\xi}), q\rangle_{X^*,X} \quad \forall (h, q) \in X_{ad} \times X. \tag{29}$$

We use the notation ∇j for the gradient of j in X, i.e., $(\nabla j(h), q)_X = \langle Dj(h), q\rangle_{X^*,X}$, where $h, q \in X$. The projected stochastic gradient method relies on a recursion of the form

$$h_{n+1} := \pi_{X_{ad}}(h_n - t_n G(h_n, \xi_n)), \tag{30}$$

where $h_1 \in X_{ad}$ and ξ_n is randomly sampled from the law $\mathbb{P} \circ \xi^{-1}$ independently of previous samples $\xi_1, \ldots, \xi_{n-1}$. We require that the step sizes given in (30) satisfy the Robbins–Monro rule from the original paper [41] on stochastic approximation:

$$t_n \geq 0, \quad \sum_{n=1}^\infty t_n = \infty, \quad \sum_{n=1}^\infty t_n^2 < \infty. \tag{31}$$

We will show that the recursion (30) with the step sizes (31) converges using similar arguments to those used in [27]. Note that the convergence result there applies to problems formulated over a single Hilbert space (without the two-norm discrepancy) and so cannot be immediately used for our setting. Here, we also work with assumptions that are verifiable for our application. For completeness, therefore, we provide a proof.

First, we recall some concepts that will be of use in the proof. A filtration is a sequence $\{\mathcal{F}_n\}$ of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$. Given a Banach space Y, we define a discrete Y-valued stochastic process as a collection of Y-valued random variables indexed by n, in other words, the set $\{\beta_n \mid \Omega \to Y \mid n \in \mathbb{N}\}$. The stochastic process is said to be adapted to a filtration $\{\mathcal{F}_n\}$ if and only if β_n is $\mathcal{F}_n$-measurable for all n. Suppose $\mathcal{B}(Y)$ denotes the set of Borel sets of Y. The natural filtration is the filtration generated by the sequence $\{\beta_n\}$ and is given by $\mathcal{F}_n = \sigma(\{\beta_1, \ldots, \beta_n\}) = \{\beta_i^{-1}(B) \mid B \in \mathcal{B}(Y), i = 1, \ldots, n\}$. If for an event $F \in \mathcal{F}$ we have that $\mathbb{P}(F) = 1$, we say F occurs almost surely (a.s.). For an integrable random variable $\beta \colon \Omega \to \mathbb{R}$, the conditional expectation is denoted by $\mathbb{E}[\beta|\mathcal{F}_n]$, which is itself a random variable that is $\mathcal{F}_n$-measurable and which satisfies $\int_A \mathbb{E}[\beta|\mathcal{F}_n](\omega)\, \mathrm{d}\mathbb{P}(\omega) = \int_A \beta(\omega)\, \mathrm{d}\mathbb{P}(\omega)$ for all $A \in \mathcal{F}_n$.

To demonstrate convergence, we will apply the following lemma.

Lemma 8 (Robbins–Siegmund [42]). *Assume that $\{\mathcal{F}_n\}$ is a filtration and v_n, a_n, b_n, c_n nonnegative random variables adapted to $\{\mathcal{F}_n\}$. If*

$$\mathbb{E}[v_{n+1}|\mathcal{F}_n] \le v_n(1 + a_n) + b_n - c_n \text{ a.s.}$$

and $\sum_{n=1}^\infty a_n < \infty$, $\sum_{n=1}^\infty b_n < \infty$ a.s., then with probability one, $\{v_n\}$ is convergent and $\sum_{n=1}^\infty c_n < \infty$.

We will also need the following result.

Proposition 9 ([27]). *Let $\{\tau_n\}$ be a nonnegative deterministic sequence and $\{\beta_n\}$ a nonnegative random sequence in $\mathbb{R}$ adapted to $\{\mathcal{F}_n\}$. Assume that $\sum_{n=1}^\infty \tau_n = \infty$ and $\mathbb{E}[\sum_{n=1}^\infty \tau_n \beta_n] < \infty$. Moreover assume that $\beta_n - \mathbb{E}[\beta_{n+1}|\mathcal{F}_n] \le \gamma \tau_n$ a.s. for all n and some $\gamma > 0$. Then*

$$\beta_n \text{ converges to } 0 \text{ a.s.}$$

To ensure convergence of (30), we make the following assumptions, which is a slight modification of those used in [27, Theorem 3.6]. Since in our application, X_{ad} is bounded, we can reasonably impose a uniform bound on the second moment term in Assumption 10 (iii) instead of the growth condition used in [27].

Assumption 10. *Let $\{\mathcal{F}_n\}$ be an increasing sequence of σ-algebras. For each n, there exist b_n, w_n with*

$$b_n = \mathbb{E}[G(h_n, \xi_n)|\mathcal{F}_n] - \nabla j(h_n), \quad w_n = G(h_n, \xi_n) - \mathbb{E}[G(h_n, \xi_n)|\mathcal{F}_n],$$

which satisfy the following assumptions:
(i) h_n and b_n are $\mathcal{F}_n$-measurable; (ii) for $K_n := \operatorname{ess\,sup}_{\omega \in \Omega} \|b_n(\omega)\|_X$ we have that $\sum_{n=1}^\infty t_n K_n < \infty$ and $\sup_n K_n < \infty$; (iii) there exists $M \ge 0$ such that $\mathbb{E}[\|G(h, \xi)\|_X^2] \le M$ for all $r \in X_{ad}$.

The sequence b_n is a bias term that can be neglected if the stochastic gradient can be computed exactly for ξ_n. The following result follows using similar arguments to those made in[27, Theorem 3.6]. The main difference is that strong convergence occurs with respect to the weaker norm H, even though the iterates belong to X. Here, some arguments are simplified since we assume that X_{ad} is bounded.

Theorem 11. *Suppose that Assumption 7 and Assumption 10 hold. If the sequence of step sizes satisfy* (31)*, then for iterates defined by the recursion* (30)*, we have*

1 $\{j(h_n)\}$ *converges a.s. and* $\lim_{n\to\infty} j(h_n) = j(h^*)$*,*

2 $\{h_n\}$ *almost surely converges weakly in* X *to* h^**, and*

3 $\{h_n\}$ *almost surely converges strongly in* H *to* h^**.*

Proof. Recall that since j is strongly convex with respect to H, a unique minimum $h^* \in X_{ad}$ exists. We note that $\pi_{X_{ad}}(h^*) = h^*$. Now, with $g_n := G(h_n, \xi_n)$, we use the nonexpansivity of the projection to obtain

$$\begin{aligned} \|h_{n+1} - h^*\|_X^2 &= \|\pi_{X_{ad}}(h_n - t_n g_n) - \pi_{X_{ad}}(h^*)\|_X^2 \\ &\le \|h_n - t_n g_n - h^*\|_X^2 \\ &= \|h_n - h^*\|_X^2 - 2t_n(g_n, h_n - h^*)_X + t_n^2\|g_n\|_X^2. \end{aligned} \tag{32}$$

Due to (strong) convexity and the fact that $h_n, h^* \in X_{ad}$, we have

$$j(h^*) - j(h_n) \ge Dj(h_n)[h^* - h_n] + \frac{c_E}{2}\|h^* - h_n\|_H^2 \ge Dj(h_n)[h^* - h_n]$$

so that $-(j(h_n) - j(h^*)) \ge -Dj(h_n)[h_n - h^*] = -(\nabla j(h_n), h_n - h^*)_X$. Moreover, optimality of h^* gives (23). Taking the conditional expectation on both sides of (32) and applying Cauchy–Schwarz for the bias term, we have

$$\begin{aligned} \mathbb{E}[\|h_{n+1} - h^*\|_X^2 | \mathcal{F}_n] &\le \|h_n - h^*\|_X^2 - 2t_n(\nabla j(h_n) + b_n, h_n - h^*)_X + t_n^2\mathbb{E}[\|g_n\|_X^2|\mathcal{F}_n] \\ &\le (1 + 2t_n K_n)\|h_n - h^*\|_X^2 - 2t_n(j(h_n) - j(h^*)) + t_n^2 M. \end{aligned} \tag{33}$$

Lemma 8 implies that $\{\|h_n - h^*\|_X\}$ is a.s. convergent and $\sum_{n=1}^\infty t_n(j(h_n) - j(h^*)) < \infty$ a.s., from which we can conclude that $\liminf_{n\to\infty} j(h_n) = j(h^*)$ with probability one. To show that in fact $\lim_{n\to\infty} j(h_n) = j(h^*)$, can use a simpler argument than in [27, Theorem 3.6] since we assumed X_{ad} to be bounded. Indeed, applying expectation (33) and again using Lemma 8, we obtain that $\sum_{n=1}^\infty t_n\mathbb{E}[j(h_n) - j(h^*)] < \infty$ surely. Convexity of j implies that

$$\begin{aligned} j(h_n) - j(h_{n+1}) &\le (\nabla j(h_n), h_n - h_{n+1})_X \\ &\le \|\nabla j(h_n)\|_X\|h_n - h_{n+1}\|_X \\ &= \|\nabla j(h_n)\|_X\|\pi_{X_{ad}}(h_n) - \pi_{X_{ad}}(h_n - t_n g_n)\|_X \\ &\le \|\nabla j(h_n)\|_X\|t_n g_n\|_X. \end{aligned}$$

Since X_{ad} is bounded, so is $\{h_n\}$, and so there exists a $\tilde{M} > 0$ such that $\|\nabla j(h_n)\|_X \le \tilde{M}$ for all n. After applying the conditional expectation, we have that

$$j(h_n) - \mathbb{E}[j(h_{n+1})|\mathcal{F}_n] \le t_n\tilde{M}\mathbb{E}[\|g_n\|_X|\mathcal{F}_n].$$

From Jensen's inequality, we see that

$$(\mathbb{E}[\|g_n\|_X|\mathcal{F}_n])^2 \le \mathbb{E}[\|g_n\|_X^2|\mathcal{F}_n] \le M,$$

from which we can conclude that $j(h_n) - \mathbb{E}[j(h_{n+1})|\mathcal{F}_n] \le t_n\tilde{M}M$. Now, we can apply Proposition 9 to conclude that $\lim_{n\to\infty} j(h_n) = j(h^*) = 0$, which was the first claim.

For the second claim, we observe an arbitrary trajectory of the random sequence $\{h_n\}$. Since $\{\|h_n - h^*\|_X^2\}$ is convergent, it is also bounded. In particular, there exists a weak accumulation point $\bar{h} \in X_{ad}$ of the sequence $\{h_n\}$. Let $\{h_{n_k}\}$ be a subsequence such that $h_{n_k} \rightharpoonup_X \bar{h}$. By weak lower semicontinuity of j (which follows from the continuity and convexity of j in X), we have

$$j(\bar{h}) \leq \lim_{k\to\infty} j(h_{n_k}) = j(h^*),$$

where equality follows by the first part of this proof. Since h^* is the (unique) minimizer, it follows that $j(h^*) = j(\bar{h})$, from which we can conclude that $h^* = \bar{h}$. The accumulation point being unique gives in fact $h_n \rightharpoonup_X h^*$ with probability one.

The third claim follows now directly from (27), namely that $\|h_n - h^*\|_H^2 \to 0$ a.s. as $n \to \infty$. □

5.3 Discussion

The above result is stronger than it may seem at first glance. While j is strongly convex with respect to H, it is only convex with respect to X. On the other hand, the H-strong convexity makes j strictly convex in X, from which we can conclude that a unique solution exists. We proved that, at least with respect to the weaker norm H, we can expect (almost sure) strong convergence of method to this unique minimizer. We note that the dimension of the underlying random vector ξ appears to be immaterial in the original result from [27, Theorem 3.6].

Theorem 11 provides the argument for almost sure convergence of the projected stochastic gradient method. It is natural to ask whether convergence rates (in the mean square) can be derived as in [30]. Interestingly, because of the two-norm discrepancy, one cannot obtain the expected convergence rates for strongly convex problems. Let us investigate this further.

Note that strong convexity (26) of j implies

$$Dj(h_n)[h_n - h^*] - Dj(h^*)[h_n - h^*] \geq c_E \|h_n - h^*\|_H^2 \tag{34}$$

for all n. Picking up from the estimate in (32), we write

$$\begin{aligned}
&\mathbb{E}[\|h_{n+1} - h^*\|_X^2 | \mathcal{F}_n] \\
&\quad \leq \|h_n - h^*\|_X^2 - 2t_n(\mathbb{E}[g_n|\mathcal{F}_n], h_n - h^*)_X + t_n^2 \mathbb{E}[\|g_n\|_X^2 | \mathcal{F}_n] \\
&\quad \leq \|h_n - h^*\|_X^2 - 2t_n(\nabla j(h_n) + b_n, h_n - h^*)_X + t_n^2 \mathbb{E}[\|g_n\|_X^2 | \mathcal{F}_n] \\
&\quad \leq \|h_n - h^*\|_X^2 - 2t_n(\nabla j(h_n) - \nabla j(h^*) + b_n, h_n - h^*)_X + t_n^2 \mathbb{E}[\|g_n\|_X^2 | \mathcal{F}_n] \\
&\quad \leq (1 + 2t_n K_n)\|h_n - h^*\|_X^2 - 2t_n c_E \|h_n - h^*\|_H^2 + 2t_n K_n + t_n^2 M.
\end{aligned}$$

Due to the mixture of norms in the final line, we fail to produce the recursion necessary to prove the usual rate for strongly convex functions. If we tried to do the above computations, but in the H norm, we would fail because

$$(\nabla j(h_n), h_n - h^*)_H \neq Dj(h_n)[h_n - h^*],$$

and $\nabla j(h_n)$ is the Riesz representative with respect to X, not H.

On the other hand, in the numerical section, we will observe convergence rates that fit the theory for strongly convex functions. Once discretized, the underlying spaces are finite-dimensional, where all norms are equivalent. In the finite-dimensional case with norm $\|\cdot\|$, if step sizes are chosen such that $t_n = \theta/n$ with $\theta > 1/(2c_E)$, we can expect in the unbiased case (see [40]):

$$\mathbb{E}[\|h_n - h^*\|] \leq \sqrt{\frac{\rho}{n}}$$

with $\rho = \max\{\|h_1 - h^*\|^2, \theta^2 M(2c_E\theta - 1)^{-1}\}$. Moreover, if h^* satisfies $\nabla j(h^*) = 0$ (i.e., h^* is an interior point of j in finite dimensions), C_S-Lipschitz continuity of j gives the following rate of convergence for function values:

$$\mathbb{E}[j(h_n) - j(h^*)] \leq \frac{C_S\rho}{2n}.$$

We note that in the case where $\nabla j(h^*) = 0$, C_S-Lipschitz continuity of ∇j allows us to obtain a convergence rate of the expected norm of the gradient, since by Lipschitz continuity of ∇j,

$$\mathbb{E}[\|\nabla j(h_n)\|] \leq C_S \mathbb{E}[\|h_n - h^*\|] \leq C_S \sqrt{\frac{\rho}{n+\nu}}. \tag{35}$$

The convergence rate (35) will indeed be observed in the numerical simulation, even though we cannot show this in the appropriate function space.

As a final comment, we remark that the assumptions made on measurability in Assumption 10 are not too strong, as shown in the following lemma from [29]. We recall our assumption that the image space Ξ of the random vector ξ is a complete separable metric space.

Lemma 12. *Suppose X is also assumed to be separable and $\{\mathcal{F}_n\}$ is the natural filtration generated by the stochastic process $\{\xi_n\}$. Suppose $G\colon X \times \Xi \to X$ and $\nabla j\colon X \to X$ are continuous with respect to the X norm in $X_{ad} \times \Xi$ and X_{ad}, respectively. Then h_n, defined by the recursion* (30)*, as well as the functions b_n and w_n respectively, are adapted to $\mathcal{F}_n$ for all n.*

6 Numerical results

For our numerical setting, we assume that both, the interior and the exterior boundary, are starlike and use polar coordinates to parameterize them. The associated exterior radial function is represented by the finite Fourier series

$$r_\Gamma(\theta) = a_{0,\Gamma} + \sum_{\ell=1}^{N} a_{-\ell,\Gamma}\sin(\ell\theta) + a_{\ell,\Gamma}\cos(\ell\theta), \quad \theta \in [0, 2\pi], \tag{36}$$

and likewise the interior one by

$$r_\Sigma(\theta,\omega) - \overline{r}_\Sigma(\theta) = \xi_0(\omega) + \sum_{\ell=1}^{N} \xi_{-\ell}(\omega)\sin(\ell\theta) + \xi_\ell(\omega)\cos(\ell\theta), \quad \theta \in [0, 2\pi].$$

Here, $\overline{r}_\Sigma(\theta)$ is chosen as the radial function which describes the ellipse with semi-axes 0.4 and 0.2, while the random variables $\xi_\ell(\omega) \in \mathcal{U}([-0.5, 0.5])$ are uniformly distributed and independent. We thus have $\mathbb{E}[r_\Sigma(\theta)] = \overline{r}_\Sigma(\theta)$.

For our numerical experiments, we employ 17 degrees of freedom in (36), which corresponds to $N = 8$. Due to the use of finite dimensional Fourier series, both boundaries are always C^∞-smooth and of bounded curvature provided that the radial functions are uniformly bounded from above and below. Especially, the Riesz projection (29) of the discretized gradient is just the identity as the gradient is a member of X. Also the realization of the projection of the exterior boundaries onto the class of convex boundaries becomes obsolete as the exterior boundary is always convex during the runs of the stochastic gradient method.

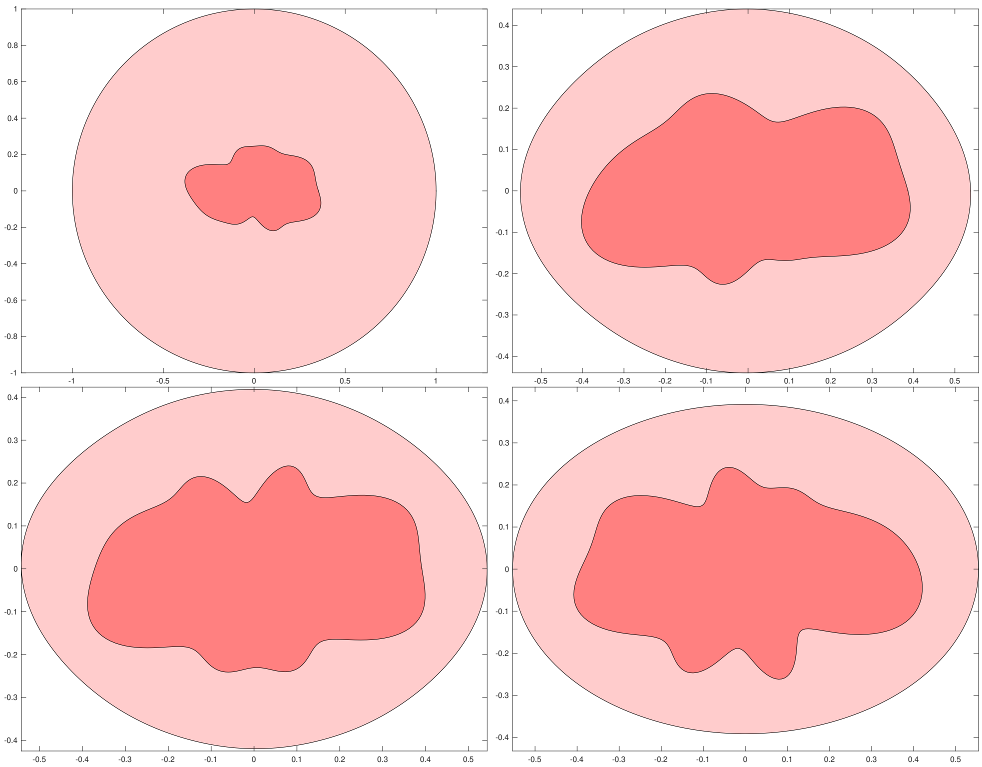

Figure 2: The initial (circular) exterior boundary (top left), the exterior boundary after 10 iterations (top right) after 20 iterations (bottom left), and at after 1000 iterations (bottom right). The interior boundaries represent different random samples.

The $H^{1/2}$-energy norm of the shape gradient is realized by applying an appropriate scaling of its Fourier coefficients. The initial guess for the exterior boundary is a circle of radius 0.75, which is centered in the origin (compare Figure 2 top left). It was not neceessary to impose constraints on the parameterization of the outer boundary, as it is of bounded curvature since the radial function consists only of a few terms. Moreover, we never observed difficulties in the numerical simulations which is in line with the observations made in [22] that the optimization problem under consideration is convex in the present setting despite of the non-convex boundaries. Note that all the details of the implementation, which is based on a boundary element method, can be found therein, too.

We apply K steps of the stochastic gradient method for different numbers of K, where the step size t_k is in any case chosen in accordance with $t_k = \frac{1}{400k}$. The factor $\frac{1}{400}$ is found to be necessary in order to avoid degeneration of the underlying domains during the course of iteration. We observe quite a fast convergence of stochastic gradient method towards the final ellipse-like outer boundary. After already 10 iterations, we get the result found in the top right plot of Figure 2, while after 20 iterations we get we get the result found in the bottom left plot of Figure 2. The boundary computed after $K = 10\,000$ iterations is found in the bottom right plot of Figure 2. The interior boundaries seen in Figure 2 represent different draws of the random interior boundary.

In Figure 3, we plot the error between the mean energy functional and its minimizer as well as the norm of the respective shape gradient for the K-th iterate versus the number K of iterations of the stochastic gradient method. Both expectations are computed by a quasi-Monte Carlo method using 1000 samples. Moreover, each particular data point reflects the mean of three runs of the stochastic gradient method.

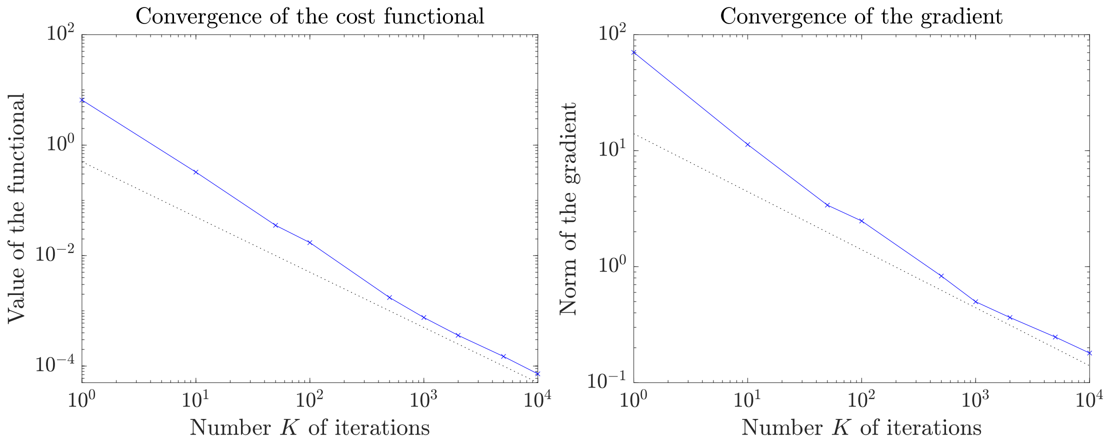

Figure 3: Convergence of the stochastic gradient method with respect to the number K of samples. The values of the cost functional are seen on the left, the norm of the gradient is seen on the right. We observe the rate of convergence K^{-1} for the cost functional and $K^{-1/2}$ for the gradient, indicated by the dotted black lines.

One can read from the right plot in Figure 3 that the norm of the initial mean gradient is approximately 70, while after $K = 10\,000$ iterations the norm of the mean gradient lies between 0.010 and 0.020, depending on the specific run. The cost functional converges towards the value $E_{\min} \approx 31.856$, which has been computed by using $K = 20\,000$ samples in the stochastic gradient method, compare the left plot in Figure 3. We observe the rate K^{-1} of convergence for the cost functional while it is $K^{-1/2}$ for the norm of the gradient. These rates are indicated by the dotted lines in Figure 3. Indeed, the rate of convergence seems to be a bit faster for the first few samples in the beginning.

7 Conclusion

In the present article, we developed the convergence theory of the stochastic gradient method in case of a problem which exhibits the two-norm discrepency. The two-norm discrepency is a well-known phenomenon in the optimal control of partial differential equations. We considered exemplarily Bernoulli's free boundary problem with a random interior boundary which can be seen as a fruit fly of a shape optimization problem under uncertainty. We have proven the strong convexity of the underlying shape optimization problem with respect to the $H^{1/2}$-norm, being weaker than the $C^{3,2\alpha}$-regularity required to ensure differentiability. Numerical results validate our theoretical findings.

References

[1] G. Allaire and C. Dapogny. A linearized approach to worst-case design in parametric and geometric shape optimization. *Math. Models Methods Appl. Sci.*, 24:2199–2257, 2014.

[2] G. Allaire and C. Dapogny. A deterministic approximation method in shape optimization under random uncertainties. *SMAI J. Comput. Math.*, 1:83–143, 2015.

[3] H. W. Alt and L. A. Caffarelli. Existence and regularity for a minimum problem with free boundary. *J. Reine Angew. Math.*, 325:105–144, 1981.

[4] P. R. S. Antunes and B. Bogosel. Parametric shape optimization using the support function. *Comput. Optim. Appl.*, 82(1):107–138, 2022. ISSN 0926-6003,1573-2894. doi: 10.1007/s10589-022-00360-4. URL https://doi.org/10.1007/s10589-022-00360-4.

[5] K. Barty, J.-S. Roy, and C. Strugarek. Hilbert-valued perturbed subgradient algorithms. *Math. Oper. Res.*, 32(3):551–562, 2007.

[6] A. Behzadan and M. Holst. Multiplication in Sobolev spaces, revisited. *Ark. Mat.*, 59(2):275–306, 2021.

[7] A. Beurling. On free boundary problems for the Laplace equation. *Seminars on Analytic functions, Institute for Advanced Study, Princeton, NJ*, 1:248–263, 1957.

[8] B. Bogosel. Numerical shape optimization among convex sets. *Appl. Math. Optim.*, 87: 1, 2023. ISSN 0095-4616,1432-0606. doi: 10.1007/s00245-022-09920-w. URL https://doi.org/10.1007/s00245-022-09920-w.

[9] A. Boulkhemair. On a shape derivative formula in the Brunn-Minkowski theory. *SIAM J. Control Optim.*, 55(1):156–171, 2017. ISSN 0363-0129,1095-7138. doi: 10.1137/15M1015844. URL https://doi.org/10.1137/15M1015844.

[10] A. Boulkhemair and A. Chakib. On a shape derivative formula with respect to convex domains. *J. Convex Anal.*, 21(1):67–87, 2014. ISSN 0944-6532,2363-6394.

[11] R. Brügger, R. Croce, and H. Harbrecht. Solving a Bernoulli type free boundary problem with random diffusion. *ESAIM Control Optim. Calc. Var.*, 26:56, 2020.

[12] E. Casas and F. Tröltzsch. Second order analysis for optimal control problems: Improving results expected from abstract theory. *SIAM J. Optim.*, 22(1):261–279, Jan. 2012. ISSN 1052-6234. doi: 10.1137/110840406. URL https://epubs.siam.org/doi/abs/10.1137/110840406. Publisher: Society for Industrial and Applied Mathematics.

[13] S. Conti, H. Held, M. Pach, M. Rumpf, and R. Schultz. Shape optimization under uncertainty. a stochastic programming perspective. *SIAM J. Optim.*, 19(4):1610–1632, 2009.

[14] J.-C. Culioli and G. Cohen. Decomposition/coordination algorithms in stochastic optimization. *SIAM J. Control Optim.*, 28(6):1372–1403, 1990.

[15] M. Dambrine and V. Karnaev. Robust obstacle reconstruction in an elastic medium. *Discrete Continuous Dyn. Syst. Ser. B.*, 29(1):124–150, 2024.

[16] M. Dambrine and J. Lamboley. Stability in shape optimization with second variation. *J. Differ. Equ.*, 267(5):3009–3045, 2019. ISSN 0022-0396,1090-2732. doi: 10.1016/j.jde.2019.03.033. URL https://doi.org/10.1016/j.jde.2019.03.033.

[17] M. Dambrine, C. Dapogny, and H. Harbrecht. Shape optimization for quadratic functionals and states with random right-hand sides. *SIAM J. Control Optim.*, 53(5):3081–3103, 2015.

[18] M. Dambrine, H. Harbrecht, M. Peters, and B. Puig. On Bernoulli's free boundary problem with a random boundary. *Int. J. Uncertain. Quantif.*, 7(4):335–353, 2017.

[19] M. Dambrine, H. Harbrecht, and B. Puig. Incorporating knowledge on the measurement noise in electrical impedance tomography. *ESAIM Control Optim. Calc. Var.*, 25:84, 2019.

[20] M. Dambrine, H. Harbrecht, and B. Puig. Bernoulli free boundary problems under uncertainty: the convex case. *Comput. Methods Appl. Math.*, 23(2):333–352, 2023. ISSN 1609-4840,1609-9389. doi: 10.1515/cmam-2022-0038. URL https://doi.org/10.1515/cmam-2022-0038.

[21] M. C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization, Second Edition*. Advances in Design and Control. Society for Industrial and Applied Mathematics, Philadelphia, 2011.

[22] K. Eppler and H. Harbrecht. Efficient treatment of stationary free boundary problems. *Appl. Numer. Math.*, 56(10-11):1326–1339, 2006.

[23] K. Eppler, H. Harbrecht, and R. Schneider. On convergence in elliptic shape optimization. *SIAM J. Control Optim.*, 46(1):61–83, 2007.

[24] M. Flucher and M. Rumpf. Bernoulli's free-boundary problem qualitative theory and numerical approximation. *J. Reine Angew. Math.*, 486:165–204, 1997.

[25] A. Friedman. *Variational Principles and Free Boundaries*. John Wiley & Sons Inc., New York, 1982.

[26] E. Gagliardo. Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili. *Rend. Sem. Mat. Univ. Padova*, 27:284–305, 1957. ISSN 0041-8994. URL http://www.numdam.org/item?id=RSMUP_1957__27__284_0.

[27] C. Geiersbach and G. C. Pflug. Projected stochastic gradients for convex constrained problems in Hilbert spaces. *SIAM J. Optim.*, 29(3):2079–2099, 2019. doi: https://doi.org/10.1137/18m1200208.

[28] C. Geiersbach and T. Scarinci. Stochastic proximal gradient methods for nonconvex problems in Hilbert spaces. *Comput. Optim. Appl.*, 3(78):705–740, 2021. doi: 10.1007/s10589-020-00259-y.

[29] C. Geiersbach and T. Scarinci. A stochastic gradient method for a class of nonlinear PDE-constrained optimal control problems under uncertainty. *J. Differ. Equ.*, 364:635–666, 2023. ISSN 0022-0396. doi: 10.1016/j.jde.2023.04.034. URL https://www.sciencedirect.com/science/article/pii/S0022039623003066.

[30] C. Geiersbach and W. Wollner. A stochastic gradient method with mesh refinement for PDE-constrained optimization under uncertainty. *SIAM J. Sci. Comput.*, 42(5):A2750–A2772, 2020. doi: https://doi.org/10.1137/19m1263297.

[31] C. Geiersbach, E. Loayza-Romero, and K. Welker. Stochastic approximation for optimization in shape spaces. *SIAM J. Optim.*, 31(1):348–376, 2021. doi: https://doi.org/10.1137/20M1316111.

[32] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, Berlin-Heidelberg, 2001.

[33] L. Goldstein. Minimizing noisy functionals in Hilbert space: An extension of the Kiefer-Wolfowitz procedure. *J. Theor. Probab.*, 1(2), 1988.

[34] H. Harbrecht and M. Peters. Solution of free boundary problems in the presence of geometric uncertainties. In M. B. et al., editor, *Topological Optimization and Optimal Transport in the Applied Sciences*, page 20–39. de Gruyter, Berlin-Bosten, 2017.

[35] A. Henrot and M. Pierre. *Variation et optimisation de formes. Une analyse géométrique*, volume 48 of *Mathématiques & Applications*. Springer, Berlin, 2005. ISBN 978-3-540-26211-4; 3-540-26211-3. doi: 10.1007/3-540-37689-5. URL https://doi.org/10.1007/3-540-37689-5.

[36] A. Henrot and H. Shahgholian. Convexity of free boundaries with Bernoulli type boundary condition. *Nonlinear Anal. Theory Methods Appl.*, 28(5):815–823, 1997.

[37] A. D. Ioffe. Necessary and sufficient conditions for a local minimum. 3: Second order conditions and augmented duality. *SIAM J. Control Optim.*, 17(2):266–288, 1979.

[38] M. Martin, S. Krumscheid, and F. Nobile. Complexity analysis of stochastic gradient methods for pde-constrained optimal control problems with uncertain parameters. *ESAIM: Math. Model. Numer. Anal.*, 55(4):1599–1633, 2021.

[39] J. Martínez-Frutos, D. Herrero-Pérez, M. Kessler, and F. Periago. Risk-averse structural topology optimization under random fields using stochastic expansion methods. *Comput. Methods Appl. Mech. Engrg.*, 330:180–206, 2018.

[40] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust Stochastic Approximation Approach to Stochastic Programming. *SIAM J. Optim.*, 19(4):1574–1609, 2009.

[41] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Statist.*, 22(3):400–407, 1951. doi: 10.1214/aoms/1177729586.

[42] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In *Optimizing Methods in Statistics*, pages 233–257. Academic Press, Cambridge, 1971.

[43] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization. Shape Sensitivity Analysis*. Springer Series in Computational Mathematics. Springer, Berlin-Heidelberg, 1992.

[44] D. E. Tepper. On a free boundary problem, the starlike case. *SIAM J. Math. Anal.*, 6(3):503–505, 1975.

[45] G. Yin and Y. M. Zhu. On H-valued Robbins-Monro processes. *J. Multivar. Anal.*, 34:116–140, 1990.