# SHAPE OPTIMIZATION OF OPTICAL MICROSCALE INCLUSIONS

MANASWINEE BEZBARUAH*, MATTHIAS MAIER*, AND WINNIFRIED WOLLNER†

**Abstract.** This paper describes a class of shape optimization problems for optical metamaterials comprised of periodic microscale inclusions composed of a dielectric, low-dimensional material suspended in a non-magnetic bulk dielectric. The shape optimization approach is based on a homogenization theory for time-harmonic Maxwell's equations that describes effective material parameters for the propagation of electromagnetic waves through the metamaterial. The control parameter of the optimization is a deformation field representing the deviation of the microscale geometry from a reference configuration of the cell problem. This allows for describing the homogenized effective permittivity tensor as a function of the deformation field. We show that the underlying deformed cell problem is well-posed and regular. This, in turn, proves that the shape optimization problem is well-posed. In addition, a numerical scheme is formulated that utilizes an adjoint formulation with either gradient descent or BFGS as optimization algorithms. The developed algorithm is tested numerically on a number of prototypical shape optimization problems with a prescribed effective permittivity tensor as the target.

**Key words.** Shape optimization, adjoint formulation, inverse homogenization, plasmonic crystals, time-harmonic Maxwell's equations

**AMS subject classifications.** 35Q60, 49M41, 65N21, 65N30

## 1. Introduction.

Plasmonic crystals consisting of 2D dielectric inclusion have given way to a wide range of striking optical phenomena that are unlike the classical behavior of electromagnetic waves. Some examples of such optical phenomenon are optical cloaking, negative-refraction, subwavelength focusing, and the *epsilon-near-zero* effect [1, 9, 21, 28, 30, 33, 41]. A well-studied example of a plasmonic crystal is graphene nanosheets that are arranged periodically with subwavelength spacing and suspended in a bulk non-magnetic dielectric host [1, 9]. The wide wealth of applications motivates the need to describe, design, and tune optical properties of plasmonic crystals [24, 31]. This, in turn, motivates the current work presented here. Analytic homogenization approaches for time-harmonic Maxwell equations [45–47] have been extended successfully to the setting of plasmonic crystals with lower-dimensional interfaces [2, 26, 27]. Here, the periodic microstructure is replaced by an effective permittivity tensor that is determined by a weighted average over a so-called cell problem. This setup can serve as an efficient computational tool for computing the optical macroscale response [23, 27]. In this manuscript, we introduce a shape-optimization approach that is based on this notion of computing effective material parameters with the help of microscopic cell problems. Specifically, we introduce a control variable describing the deformation of a given reference geometry of the cell problem with the target of achieving a preset macroscopic permittivity tensor. Our key objectives with the present paper are to

- Extend the two-scale homogenization result by introducing a deformation field. We show, in particular, that the resulting system is well-posed and show regularity of the corrector solution of the modified cell problem, provided that the deformation field is of suitable regularity.
- Formulate a shape-optimization problem based on the homogenization procedure that uses the deformation field as the control variable. We also discuss

*Department of Mathematics, Texas A&M University, 3368 TAMU, College Station, TX 77843, USA ({bezba004,maier}@tamu.edu)

† Fachbereich Mathematik, MIN Fakultät, Universität Hamburg, Bundesstr. 55, 20146 Hamburg, Germany (winnifried.wollner@uni-hamburg.de)

penalization strategies for ensuring mesh regularity of the cell problem.
– Discuss a number of prototypical numerical results highlighting the algorithmic approach. We in particular demonstrate that our shape optimization formulation works for manipulating optical microstructures to achieve an epsilon-near-zero plasmonic crystal.

**1.1. Microscopic model and homogenization theory.** Figure 1 shows the geometry of a three-dimensional plasmonic crystal consisting of periodic copies of a scaled representative volume element $Y$ containing a curved surface $\Sigma$ representing a 2D material with surface conductivity $\sigma$ and otherwise filled with a dielectric host material. The periodicity is denoted by $d$ and $\Sigma^d$ is the union of all scaled copies of $\Sigma$; see Figure 1. We assume that the dielectric host has a uniform and isotropic

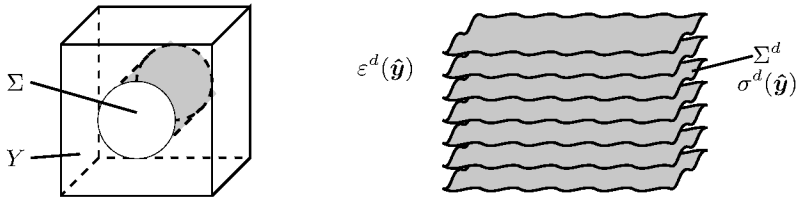



Figure 1: (a) The unit cell $Y = [0,1]^3$ consisting of 2D graphene inclusions $\Sigma$ with surface conductivity $\sigma$ in an ambient host material with permittivity $\varepsilon$; (b) the plasmonic crystal formed by many scaled and repeated copies of $Y$ in all space directions.

relative permittivity $\varepsilon$ and permeability $\mu$. The (scaled) surface conductivity $\sigma^d$ in the numerical tests will be given by a simple Drude model; see Section 2.1. The time-harmonic response of the plasmonic crystal is described by the time-harmonic Maxwell equations governing the scattering of an electromagnetic wave $(\boldsymbol{E}^d, \boldsymbol{H}^d)$, viz.

$$\nabla \times \boldsymbol{E}^d = i\omega\mu \boldsymbol{H}^d, \qquad \nabla \times \boldsymbol{H}^d = -i\omega\varepsilon \boldsymbol{E}^d + \boldsymbol{J}_a^d.$$

Here, $\omega$ is the angular frequency, $\mu$ denotes the relative permeability and $\boldsymbol{J}_a^d$ is a prescribed current density. This system of equations is now closed with jump conditions over the interface $\Sigma^d$ that arise from the surface conductivity $\sigma^d(\boldsymbol{x})$; see Section 2.1. In the limit $d \to 0$ for the periodicity, the time-harmonic system reduces to a *homogenized system*,

$$\nabla \times \boldsymbol{\mathcal{E}} = i\omega\mu \boldsymbol{\mathcal{H}}, \qquad \nabla \times \boldsymbol{\mathcal{H}} = -i\omega\varepsilon^{\text{eff}} \boldsymbol{\mathcal{E}} + \boldsymbol{J}_a^d,$$

where the heterogeneous permittivity and microscale inclusions have been replaced by an effective permittivity tensor $\varepsilon^{\text{eff}}(\boldsymbol{x})$ that is given by a *weighted* average [2, 26] over the unit cell $Y$. The weight is due to a *corrector* $\boldsymbol{\chi}$ that is described by an auxiliary cell problem incorporating the effects of the microscale inclusions $\Sigma$; see Section 2.2.

**1.2. Shape optimization approach.** We show in Section 2.4 that the above homogenization result establishes a smooth functional relationship between the effective permittivity tensor $\varepsilon^{\text{eff}}$ and the material parameters $\varepsilon$, $\mu$, $\sigma^d$, as well as the hypersurface $\Sigma$. We exploit this functional relationship to formulate a shape optimization problem

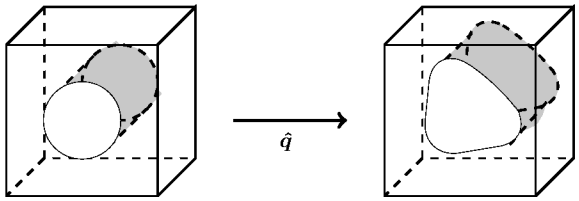



Figure 2: A reference unit cell $\hat{Y}$ with a 2D inclusion $\hat{\Sigma}$ that is deformed by a deformation vector field $\hat{\boldsymbol{q}}(\boldsymbol{x})$.

in which we deform $\Sigma$ for minimizing the misfit of $\varepsilon^{\text{eff}}$ and a preset target permittivity tensor $\varepsilon^{\text{target}}$. To this end, we introduce a deformation field $\hat{\boldsymbol{q}} : \hat{Y} \to \mathbb{R}^3$ that acts on a reference configuration $\hat{Y}$ with inclusion $\hat{\Sigma}$, see Figure 2:

$$\hat{Y} \ni \hat{\boldsymbol{y}} \quad \leftrightarrow \quad \boldsymbol{y}(\hat{\boldsymbol{y}}) := \hat{\boldsymbol{y}} + \hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}) \in Y.$$

The numerical value of the effective permittivity tensor, $\varepsilon^{\text{eff}}(\hat{\boldsymbol{\chi}}; \hat{\boldsymbol{q}})$, now depends on the *control* $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}})$ as well as the *state* $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{y}})$; see 2.3. We are now in a position to introduce a cost function

$$C(\hat{\boldsymbol{\chi}}; \hat{\boldsymbol{q}}) \;:=\; \frac{1}{2} \left\| \varepsilon^{\text{eff}}[\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{q}}] - \varepsilon^{\text{target}} \right\|_{\text{Fr.}}^2 + \frac{\alpha}{2} \left\| \nabla \hat{\boldsymbol{q}} \right\|^2 + \text{penalization},$$

where $\alpha$ is a scaling factor for a Tikhonov regularization term and the remaining penalization terms for ensuring mesh regularity are discussed in detail in Section 3. Here, $\| \cdot \|$ denotes the $L^2$-norm over $\hat{Y}$, while $\| \cdot \|_{\text{Fr.}}$ denotes the Frobenius-norm.

**1.3. Past works.** The use of nanoscale 2D dielectric inclusions for engineering and manipulating an effective optical response has been discussed extensively in the physics literature [1, 9, 21, 28, 30, 33, 41]. For some recent advances in optical materials and photonics, we refer the reader to [16, 42, 48].

The following shape optimization approach is based on a well-established periodic homogenization theory for time-harmonic Maxwell equations [45–47] extended to inclusions composed of lower-dimensional interfaces [2, 26, 27]. A number of applications of the homogenization approach have been discussed for simulating finite-sized geometries [25] and describing Lorentz resonances [23] of plasmonic crystals.

The minimization of the cost-functional is a particular case of PDE-constrained optimization, where we utilize adjoint calculus to obtain efficient formulas for the derivatives of the cost-functional with respect to the deformations $\hat{\boldsymbol{q}}$, see, e.g., [5, 8, 17, 20, 43].

Using homogenization principles for optimal material configurations has been explored before. For example, in [11] the authors use upscaling techniques to optimize microscale fiber orientation. More generally, the optimization of optical nanostructures [29, 35] and related inverse design problems [34, 40] have been explored extensively.

Our computational approach uses a *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) *quasi-Newton scheme* [10]. Such quasi-Newton methods have been generalized for optimization problems over Riemannian manifolds, particularly with emphasis on shape optimization, see, e.g., [38, 39, 44]. To assert positive definiteness of the iterates these approaches typically rely on a Powell-Wolfe type step-length selection. As such the step-length selection relies on a potential increase of the step length which can lead to conflicts with barrier method we use for the penalization. We thus utilize a damped BFGS update suggested by [36] for the inverse Hessian as proposed by [18].

**1.4. Paper organization.** The remainder of the paper is organized as follows. In Section 2, we summarize the underlying homogenization result for plasmonic crystals, reformulate the cell problem and averaging for the case of a deformed geometry, and establish well-posedness and regularity of the cell problem for the deformed configuration. The shape optimization problem is described in Section 3. A computational approach and numerical illustration are discussed in Section 4, specifically an optimization algorithm based on gradient descent and an improved version using an inverse Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method is discussed. We summarize our work and conclude in Section 5. Detailed proofs of the lemmas and theorems in Section 2 are given in Appendices A and B.

**2. Problem Description.** The manuscript is concerned with a class of shape optimization problems involving optical metamaterials comprised of periodic microscale inclusions. In this section, we summarize the underlying microscale model and a corresponding homogenization theory that will serve as a basis for the shape optimization problem. We then introduce a deformed formulation of the cell problems and formulate an optimization problem.

**2.1. Background: Heterogeneous Maxwell equations.** Consider a three-dimensional geometry $\Omega$ consisting of scaled and periodic copies of a *unit cell* $Y$, which incorporates microscale inclusions given by 2D material surfaces $\Sigma$; see Figure 1 [26, 27]. The time-harmonic response of an electromagnetic wave $(\boldsymbol{E}^d, \boldsymbol{H}^d)$ to the scattering configuration is described by the time-harmonic Maxwell equations valid on the domain $\Omega\backslash\Sigma^d$, viz., [32]

$$\nabla \times \boldsymbol{E}^d = i\omega\mu \boldsymbol{H}^d, \qquad \nabla \times \boldsymbol{H}^d = -i\omega\varepsilon^d \boldsymbol{E}^d + \boldsymbol{J}_a^d. \tag{2.1}$$

Here, $\boldsymbol{E}^d(\boldsymbol{x})$ (and $\boldsymbol{H}^d(\boldsymbol{x})$) denote the time-harmonic, complex-valued electric (and magnetic) field component, $\varepsilon^d$ is the electric permittivity of the ambient host, $\boldsymbol{J}_a^d$ is a given electric current density, and $\omega$ denotes the angular frequency. In (2.1) we used the convention that fields exhibit a $e^{-\mathrm{i}\omega t}$ time dependence, meaning, the time-dependent real-valued physical field $\boldsymbol{F}(\boldsymbol{x}, t)$ is reconstructed from its time-harmonic counterpart $\boldsymbol{F}(\boldsymbol{x})$ by setting $\boldsymbol{F}(\boldsymbol{x}, t) = \mathrm{Re}\,(\boldsymbol{F}(\boldsymbol{x})e^{-\mathrm{i}\omega t})$. The parameter $d$ denotes a *scaling parameter* for the microscale inclusions:

$$\Sigma^d = \big\{ d\,\boldsymbol{z} + d\,\boldsymbol{\varsigma} \;\; : \;\; \boldsymbol{z} \in \mathbb{Z}^3,\ \boldsymbol{\varsigma} \in \Sigma \big\}.$$

Equation (2.1) is now furnished with jump conditions over the interface $\Sigma^d$ that arise due to the presence of a current density $\boldsymbol{J}_{\Sigma^d} = \delta_{\Sigma^d}\sigma^d \boldsymbol{E}_{\mathbf{t}}^d$ caused by *surface conductivity* $\sigma^d$ of the microscale inclusions [27]. Here, $\delta_{\Sigma^d}$ is the surface measure of $\Sigma^d$. The current density leads to a jump in the tangential component of the magnetic field. In summary,

$$\Big[\boldsymbol{\nu} \times \boldsymbol{E}^d\Big]_{\Sigma^d} = 0, \qquad \Big[\boldsymbol{\nu} \times \boldsymbol{H}^d\Big]_{\Sigma^d} = (\sigma^d \boldsymbol{E}^d)_{\mathbf{t}}, \tag{2.2}$$

where the subscript $\mathbf{t}$ denotes projection onto the tangential plane of $\Sigma$ and $[\,.\,]_{\Sigma^d}$ denotes the jump over $\Sigma^d$ with respect to a chosen normal field $\boldsymbol{\nu}$,

$$[\boldsymbol{F}]_{\Sigma^d} := \lim_{\alpha \searrow 0} \Big( \boldsymbol{F}(\boldsymbol{x} + \alpha\boldsymbol{\nu}) - \boldsymbol{F}(\boldsymbol{x} - \alpha\boldsymbol{\nu}) \Big) \qquad \boldsymbol{x} \in \Sigma^d.$$

Note that numerical value of the jumps in (2.2) is independent of the orientation of the chosen normal field $\boldsymbol{\nu}$ because of the cross product of the fields with $\boldsymbol{\nu}$. We will

allow the permittivity $\varepsilon^d$ and surface conductivity $\sigma^d$ to be tensor-valued. We make the additional scaling assumption on (bulk) permittivity and surface conductivity that they depend on a slow scale ($\boldsymbol{x} \in \Omega$) and are periodic and rapidly oscillating on a fast scale ($\boldsymbol{y} \in Y$) proportional to $d$, namely

$$\varepsilon^d(\boldsymbol{x}) = \varepsilon(\boldsymbol{x}, \boldsymbol{x}/d), \qquad \sigma^d(\boldsymbol{x}) \;=\; d\,\sigma(\boldsymbol{x}, \boldsymbol{x}/d).$$

Concretely, in the following we will use a simple Drude model [15] for all numerical tests,

$$\sigma(\boldsymbol{x}, \boldsymbol{y}) \;=\; \tilde{\sigma}(\omega)\,\boldsymbol{I}, \qquad \tilde{\sigma}(\omega) = \frac{\mathrm{i}\omega_p}{\omega + \mathrm{i}/\tau}, \qquad \omega_p,\, \tau \in \mathbb{R}_{\geq 0}. \tag{2.3}$$

Equation (2.3) is a common model for the surface conductivity of graphene [15]. After a suitable nondimensionalization and rescaling typical values for the constants are $\omega_p \approx 4/137$ and $\tau \approx 100$ for frequencies $\omega \approx 1$ [23].

**2.2. Background: Homogenization results.** System (2.1) & (2.2) furnished with the Drude model (2.3) exhibits a pronounced two-scale character with an ambient vacuum wavelength $2\pi/k_0 \sim 1$ and a small scale associated with the microscale inclusions and nanoscale resonances scaling with $d$; typically, $d \ll 2\pi/k_0$ [23, 26]. In the limit $d \to 0$, system (2.1) & (2.2) reduces to a *homogenized system* [2, 23, 26, 27, 45–47]:

$$\nabla \times \boldsymbol{\mathcal{E}} = i\omega\mu\boldsymbol{\mathcal{H}}, \qquad \nabla \times \boldsymbol{\mathcal{H}} = -i\omega\varepsilon^{\mathrm{eff}}\boldsymbol{\mathcal{E}} + \boldsymbol{J}_a^d, \tag{2.4}$$

where the heterogeneous permittivity and microscale inclusions have been replaced by an effective permittivity tensor

$$\begin{aligned}\varepsilon_{ij}^{\mathrm{eff}}(\boldsymbol{x}) = \int_Y \varepsilon(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{e}_j + \nabla\chi_j^{\,T}) \cdot (\boldsymbol{e}_i + \overline{\nabla\chi_i^{\,T}})\,\mathrm{d}y \\ - \frac{1}{i\omega}\int_\Sigma \sigma(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{e}_{j\mathbf{t}} + \nabla_{\mathbf{t}}\chi_j^{\,T}) \cdot (\boldsymbol{e}_{i\mathbf{t}} + \overline{\nabla_{\mathbf{t}}\chi_i^{\,T}})\,\mathrm{d}o_y.\end{aligned} \tag{2.5}$$

Here, we have adopted the convention that the gradient is a row vector and $\boldsymbol{e}_i$ denotes the i-th (column) unit vector. The subscript $\mathbf{t}$ denotes projection onto the tangential plane of $\Sigma$, and $\bar{\cdot}$ denotes the complex conjugate of a complex-valued quantity. The corrector $\chi_i \in H := \left\{\varphi \in H^1_{\mathrm{per}}(Y, \mathbb{C}),\ \nabla_{\mathbf{t}}\varphi \in \boldsymbol{L}^2(\Sigma, \mathbb{C})\right\}$, $i = 1, 2, 3$ is given by an associated cell problem

$$\begin{aligned}\int_Y \varepsilon(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{e}_i + \nabla\chi_i^{\,T}) \cdot \overline{\nabla\varphi^{\,T}}\,\mathrm{d}y \\ - \frac{1}{i\omega}\int_\Sigma \sigma(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{e}_{i\mathbf{t}} + \nabla_{\mathbf{t}}\chi_i^{\,T}) \cdot \overline{\nabla_{\mathbf{t}}\varphi^{\,T}}\,\mathrm{d}o_y \;=\; 0 \quad \text{for all } \varphi \in H.\end{aligned} \tag{2.6}$$

$H^1_{\mathrm{per}}(Y, \mathbb{C})$ denotes the Sobolev-space of all complex-valued square-integrable function with square-integrable partial derivatives defined on the unit cell $Y$ that are periodic in all three space directions, $\varphi(\boldsymbol{y}) = \varphi(\boldsymbol{y} + \boldsymbol{e}_i)$ for $i = 1, 2, 3$ and for all $\boldsymbol{y} \in Y$. It has been established in [2, 26] that (2.1) & (2.2), (2.4), and (2.6) are well-posed provided some assumptions on the material constants are satisfied. We summarize:

THEOREM 2.1 (Well-posedness and two-scale convergence [2, 26]). *Let $\mu \in \mathbb{R}_{>0}$ and let $\varepsilon^d, \sigma^d \in L^\infty(\Omega, \mathbb{C}^{d\times d})$ be bounded, complex and tensor-valued functions such*

*that* $\mathrm{Re}\,(\varepsilon^d(\boldsymbol{x}))$ *and* $\mathrm{Im}\,(\sigma^d(\boldsymbol{x}))$ *are symmetric, and* $Im(\varepsilon^d(\boldsymbol{x}))$ *and* $Re(\sigma^d(\boldsymbol{x}))$ *are symmetric and uniformly elliptic. Then, provided* $\Sigma$ *is sufficiently smooth [2, 26],* (2.1) *&* (2.2)*,* (2.4)*, and* (2.6) *are well-posed, the solution* $(\boldsymbol{E}^d, \boldsymbol{H}^d)$ two-scale converges *to the solution* $(\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{H}})$ *of the homogenized problem* (2.4)*.*

*Remark* 2.2. The above theorem is, strictly speaking, a generalized version of what has been shown in [2, 26] where the material parameters $\varepsilon^d$ and $\sigma^d$ had been assumed to be scalar (and not tensor-valued coefficients). Proving the generalized statement, however, only requires minor adjustments to the proofs.

**2.3. Deformed cell problem.** We now introduce a deformation of the cell problem. For this we let $\hat{Y}$ and $\hat{\Sigma}$ denote a given reference configuration consisting of the unit cell $\hat{Y} = [0,1]^3$ with given (two dimensional) inclusions $\hat{\Sigma}$. Let $Y$ and $\Sigma$ denote the deformed volume and inclusions. We assume that the deformed geometry is given by a deformation vector field $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}) \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ of the reference configuration [37, pp. 70-79], such that

$$\hat{Y} \ni \hat{\boldsymbol{y}} \;\leftrightarrow\; \boldsymbol{y}(\hat{\boldsymbol{y}}) := \hat{\boldsymbol{y}} + \hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}) \in Y \tag{2.7}$$

is a bijection that also maps $\hat{\Sigma}$ onto $\Sigma$ bijectively. We adopt the notation that $\hat{\cdot}$ indicates that a function, coordinate or differential operator is on the reference configuration [37]. Here, in analogy to the function space $H$ we have introduced

$$\boldsymbol{D}(\hat{Y}, \hat{\Sigma}) := \big\{\hat{\boldsymbol{\varphi}} \in \boldsymbol{W}_0^{1,\infty}(\hat{Y}) :\; \hat{\boldsymbol{\tau}}_k \cdot \hat{\nabla}\hat{\boldsymbol{\varphi}} \in L^\infty(\hat{\Sigma}),\, k = 1, 2\big\}, \tag{2.8}$$

where $W_0^{1,\infty}(\hat{Y})$ shall denotes the Sobolev space of all functions $\varphi \in W_0^{1,\infty}(\hat{Y})$ with vanishing trace on $\partial\hat{Y}$.

*Remark* 2.3. We note that $\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ implies that $\hat{\boldsymbol{q}}$ is already Lipschitz continuous; see, for example, [**?** , Ex. 8, p. 156]. This implies that the trace of $\hat{\boldsymbol{q}}$ onto $\Sigma$ and $\partial\Omega$ is well defined and also Lipschitz continuous. Strictly speaking, this renders the condition $\hat{\boldsymbol{\tau}}_k \cdot \hat{\nabla}\hat{\boldsymbol{\varphi}} \in L^\infty(\hat{\Sigma})$ superfluous.

Definition 2.4. *The* transformation gradient $\hat{F}(\hat{\boldsymbol{y}})$ *and determinant* $\hat{J}$ *are given by*

$$\hat{F}(\hat{\boldsymbol{y}}) \;:=\; \boldsymbol{I} + \hat{\nabla}\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}), \qquad \hat{J}(\hat{\boldsymbol{y}}) \;:=\; \det\big(\hat{F}(\hat{\boldsymbol{y}})\big). \tag{2.9}$$

Here, we have adopted the convention that $\big(\hat{\nabla}\hat{\boldsymbol{q}}\big)_{ij} = \frac{\partial}{\partial\hat{y}_j}\hat{q}_i$ and $\boldsymbol{I}$ denotes the unit matrix. We then have,

Lemma 2.5 (Transformation). *Let* $\hat{\varphi}(\hat{\boldsymbol{y}})$ *be a differentiable function defined on* $\hat{Y}$ *and let* $\varphi(\boldsymbol{y}) \;:\; Y \to \mathbb{R}$ *be defined by setting* $\varphi(\boldsymbol{y}) := \hat{\varphi}(\hat{\boldsymbol{y}}(\boldsymbol{y}))$*. Then,*

$$\nabla\varphi(\boldsymbol{y})^T \;=\; \hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\varphi}(\hat{\boldsymbol{y}})^T. \tag{2.10}$$

*Let* $\hat{\boldsymbol{\nu}}$ *be a (unit) normal field on* $\hat{\Sigma}$ *and let* $\hat{\boldsymbol{\tau}}_1$*,* $\hat{\boldsymbol{\tau}}_2$ *be two orthonormal (unit) tangential fields. Then,*

$$\boldsymbol{\nu} = \hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}/\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}, \quad \boldsymbol{\tau}_i = \hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_i/\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_i\|_{\ell^2},\; i = 1, 2.$$

*are a (unit) normal field and orthonormal tangential fields on* $\Sigma$*. Moreover,*

$$\mathrm{d}y \;=\; \hat{J}\,\mathrm{d}\hat{y}, \qquad\qquad \mathrm{d}o_y \;=\; \|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}\,\mathrm{d}\hat{o}_{\hat{y}}. \tag{2.11}$$

*Proof.* We refer to [37, § 2.1.2] for a detailed discussion of the transformation identities. The transformation formula for $\boldsymbol{\nu}$ and $\mathrm{d}o_y$ is a direct consequence of *Nanson's formula*. ∎

LEMMA 2.6. *In the setting of Lemma 2.5 we have the identity*

$$\nabla_{\mathbf{t}}\varphi \;=\; \sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T(\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\varphi}) \;=\; \sum_k \frac{\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_k}{\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_k\|_{\ell^2}^2}\,\hat{\boldsymbol{\tau}}_k\cdot\hat{\nabla}\hat{\varphi}.$$

*Proof.* The statement is a direct consequence of Lemma 2.5 and the fact that

$$\nabla_{\mathbf{t}}\varphi = \sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T\nabla\varphi.$$ ∎

We are now in a position to recast (2.5) and (2.6) in reference coordinates. We introduce

$$\boldsymbol{H}(\hat{Y},\hat{\Sigma}) := \big\{\hat{\boldsymbol{\varphi}}\in\boldsymbol{H}^1_{\mathrm{per}}(\hat{Y},\mathbb{C}) \;:\; \hat{\boldsymbol{\tau}}_k\cdot\hat{\nabla}\hat{\boldsymbol{\varphi}}\in L^2(\hat{\Sigma},\mathbb{C}^{d\times d}),\ k=1,2\big\}$$

The corrector $\hat{\boldsymbol{\chi}} = \big(\hat{\chi}_i\big)_{i=1}^3 \in \boldsymbol{H}(\hat{Y},\hat{\Sigma})$ is determined by

$$E(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{\varphi}};\hat{\boldsymbol{q}}) \;=\; 0 \quad \text{for all } \hat{\boldsymbol{\varphi}}\in\boldsymbol{H}(\hat{Y},\hat{\Sigma}), \tag{2.12}$$

where,

$$\begin{aligned} E(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{\varphi}};\hat{\boldsymbol{q}}) := {}& \int_{\hat{Y}} \varepsilon(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))(\boldsymbol{I}+\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T})\cdot(\hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}})\hat{J}\,\mathrm{d}\hat{y} \\ & - \frac{1}{i\omega}\int_{\hat{\Sigma}} \sigma(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))\Big(\sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T\big(\boldsymbol{I}+\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T}\big)\Big) \\ & \qquad\cdot\Big(\sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T\hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}}\Big)\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}\,\mathrm{d}\hat{o}_{\hat{y}}. \end{aligned} \tag{2.13}$$

Here, $\boldsymbol{I}$ is the unit matrix. We note that $\sigma(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))$ and $\varepsilon(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))$, still denote tensors acting on gradients in transformed (non-reference) coordinates. Similarly, the effective permittivity tensor is given by

$$\begin{aligned} \varepsilon_{ij}^{\mathrm{eff}}(\hat{\boldsymbol{\chi}};\hat{\boldsymbol{q}}) = {}& \int_{\hat{Y}} \varepsilon(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))(\boldsymbol{e}_j+\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\chi}_j^{\,T})\cdot(\boldsymbol{e}_i+\hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\chi}_i^{\,T}})\hat{J}\,\mathrm{d}\hat{y} \\ & - \frac{1}{i\omega}\int_{\hat{\Sigma}} \sigma(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))\Big(\boldsymbol{e}_{j\mathbf{t}}+\sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\boldsymbol{\chi}}_j^{\,T}\Big) \\ & \qquad\cdot\Big(\boldsymbol{e}_{i\mathbf{t}}+\sum_k \boldsymbol{\tau}_k\boldsymbol{\tau}_k^T\hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\boldsymbol{\chi}}_i^{\,T}}\Big)\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}\,\mathrm{d}\hat{o}_{\hat{y}}. \end{aligned} \tag{2.14}$$

**2.4. Well-posedness, regularity and dependence on deformation field.** We now examine the well-posedness and regularity of problem (2.12). To this end we introduce two transformed tensors:

$$\hat{\varepsilon}(\boldsymbol{x},\hat{\boldsymbol{y}}) := \hat{F}(\hat{\boldsymbol{y}})^{-1}\varepsilon(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{J}, \quad \hat{\sigma}(\boldsymbol{x},\hat{\boldsymbol{y}}) := \sum_{mn}\hat{\sigma}_{mn}\,\hat{\boldsymbol{\tau}}_m\hat{\boldsymbol{\tau}}_n^T, \tag{2.15}$$

$$\text{where}\quad \hat{\sigma}_{mn} := \frac{\sigma_{mn}(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}}))\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}}{\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_m\|_{\ell^2}^2\|\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_n\|_{\ell^2}^2},$$

and $\sigma_{mn}$ is defined by $\sigma(\boldsymbol{x},\boldsymbol{y}(\hat{\boldsymbol{y}})) =: \sum_{mn}\sigma_{mn}\,\boldsymbol{\tau}_m\boldsymbol{\tau}_n^T$. We make the following observations:

LEMMA 2.7. *Assume that $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}) \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ and $0 < \delta \leq \hat{J}(\hat{\boldsymbol{y}})$ is uniformly bounded. Then, the tensors $\hat{\varepsilon}(\hat{\boldsymbol{y}})$ and $\hat{\sigma}(\hat{\boldsymbol{y}})$ are bounded, complex and tensor-valued functions and $\operatorname{Re}(\hat{\varepsilon}(\hat{\boldsymbol{y}}))$ and $Im(\hat{\sigma}(\hat{\boldsymbol{y}}))$ are symmetric, $Im(\hat{\varepsilon}(\hat{\boldsymbol{y}}))$ and $Re(\hat{\sigma}(\hat{\boldsymbol{y}}))$ are symmetric and uniformly elliptic with a constant depending on $\delta$.*

LEMMA 2.8. *The bilinear term (2.13) can be equivalently written as follows:*

$$E(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{\varphi}}; \hat{\boldsymbol{q}}) := \int_{\hat{Y}} \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}}) \big(\hat{F}(\hat{\boldsymbol{y}})^T + \hat{\nabla} \hat{\boldsymbol{\chi}}^T\big) \cdot \overline{\hat{\nabla} \hat{\boldsymbol{\varphi}}^T} \, \mathrm{d}\hat{y} - \frac{1}{i\omega} \int_{\hat{\Sigma}} \hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}}) \big((\hat{F}(\hat{\boldsymbol{y}})^T)_{\hat{\mathbf{t}}} + \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}^T\big) \cdot \overline{\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\varphi}}^T} \, \mathrm{d}\hat{o}_{\hat{y}}. \tag{2.16}$$

*Similarly, (2.14) takes the form:*

$$\varepsilon_{ij}^{\mathit{eff}}(\hat{\boldsymbol{\chi}}; \hat{\boldsymbol{q}}) = \int_{\hat{Y}} \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}}) \big(\hat{F}(\hat{\boldsymbol{y}})^T \boldsymbol{e}_j + \hat{\nabla} \hat{\boldsymbol{\chi}}_j^T\big) \cdot \big(\hat{F}(\hat{\boldsymbol{y}})^T \boldsymbol{e}_i + \overline{\hat{\nabla} \hat{\boldsymbol{\chi}}_i^T}\big) \, \mathrm{d}\hat{y} - \frac{1}{i\omega} \int_{\hat{\Sigma}} \hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}}) \big((\hat{F}(\hat{\boldsymbol{y}})^T \boldsymbol{e}_j)_{\hat{\mathbf{t}}} + \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}_j^T\big) \cdot \big((\hat{F}(\hat{\boldsymbol{y}})^T \boldsymbol{e}_i)_{\hat{\mathbf{t}}} + \overline{\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}_i^T}\big) \, \mathrm{d}\hat{o}_{\hat{y}}. \tag{2.17}$$

Detailed proofs of Lemmas 2.7 and 2.8 are given in Appendix A. Lemma 2.8 implies, in particular, that the deformed cell problem (2.12) has the same structure as (2.6) with a slightly modified forcing. We can thus summarize:

THEOREM 2.9. *Under the assumptions on the deformation field $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}})$ as stated in Lemma 2.7 problem (2.12) is well-posed.*

*Proof.* Lemma 2.6 ensures that equation (2.12) is the same as (2.6) with modified material tensors and a modified forcing. Lemma 2.7 ensures that all assumptions on the material tensors stated in Theorem 2.1 hold true. Well-posedness is thus an immediate consequence of Theorem 2.1. ∎

We finish the discussion by introducing a number of robustness and regularity results that will be used later to justify the optimization approach.

THEOREM 2.10. *For the unique solution $\hat{\boldsymbol{\chi}} \in \boldsymbol{H}(\hat{Y}, \hat{\Sigma})$ to (2.12), we have the following a priori estimate:*

$$\|\hat{\nabla} \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega} \|\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})} \leq C \left( \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{L^\infty(\hat{Y})} \|\hat{F}^T\|^2_{L^2(\hat{Y})} + \frac{1}{\omega} \|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{L^\infty(\hat{\Sigma})} \|\hat{F}^T\|^2_{L^2(\hat{\Sigma})} \right)$$

*for a constant $C > 0$ only depending on $\hat{\Sigma}$ and the lower bound $\delta$ of $\hat{J}(\hat{\boldsymbol{x}})$ as defined in Lemma 2.7.*

*Proof.* The statement is a consequence of Lemmas 2.7 and 2.8. We start by testing (2.16) with $\hat{\boldsymbol{\varphi}} = \overline{\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})}$ and taking the imaginary part. Recalling that $\operatorname{Re} \hat{\varepsilon}$, $\operatorname{Im} \hat{\varepsilon}$, $\operatorname{Re} \hat{\sigma}$, $\operatorname{Im} \hat{\sigma}$ are symmetric by virtue of Lemma 2.7 we arrive at

$$\int_{\hat{Y}} \operatorname{Im} \hat{\varepsilon} \, \hat{\nabla} \hat{\boldsymbol{\chi}}^T \cdot \overline{\hat{\nabla} \hat{\boldsymbol{\chi}}^T} \, \mathrm{d}\hat{y} + \frac{1}{\omega} \int_{\hat{\Sigma}} \operatorname{Re} \hat{\sigma} \, \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}^T \cdot \overline{\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}^T} \, \mathrm{d}\hat{o}_{\hat{y}} = -\operatorname{Im} \Big\{ \int_{\hat{Y}} \hat{\varepsilon} \, \big(\hat{F}(\hat{\boldsymbol{y}})^T\big) \cdot \overline{\hat{\nabla} \hat{\boldsymbol{\chi}}^T} \, \mathrm{d}\hat{y} + \frac{1}{\mathrm{i}\omega} \int_{\hat{\Sigma}} \hat{\sigma} \, \big(\hat{F}(\hat{\boldsymbol{y}})^T\big)_{\hat{\mathbf{t}}} \cdot \overline{\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}^T} \, \mathrm{d}\hat{o}_{\hat{y}} \Big\}.$$

The statement now follows from using Young's inequality for both terms on the right side and uniform ellipticity of $\operatorname{Im}\hat{\varepsilon}$ and $\operatorname{Re}\hat{\sigma}$ with a $\delta$-dependent constant that was established in Lemma 2.7. ∎

THEOREM 2.11. *Under the assumptions on the deformation field $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}})$ as stated in Lemma 2.7, the corrector $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ given by the cell problem (2.12) depends at least Lipschitz-continuously on $\hat{\boldsymbol{q}}$. More precisely, let $\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2 \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ be arbitrarily chosen such that the assumptions of Lemma 2.7 are satisfied. Let $\hat{\boldsymbol{\chi}}_1, \hat{\boldsymbol{\chi}}_2 \in \boldsymbol{H}(\hat{Y}, \hat{\Sigma})$ denote the solutions to (2.12) for deformation fields $\hat{\boldsymbol{q}}_1$ and $\hat{\boldsymbol{q}}_2$, respectively. Then, assuming that $\hat{\boldsymbol{q}}_1$ is suitably close to $\hat{\boldsymbol{q}}_2$, we have:*

$$
\begin{aligned}
(2.18)\quad & \|\hat{\nabla}(\hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2)\|^2_{L^2(\hat{Y})} + \frac{1}{\omega}\|\hat{\nabla}_{\hat{\mathbf{t}}}(\hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2)\|^2_{L^2(\hat{\Sigma})} \\
&\leq\ C(\hat{\boldsymbol{q}}) \left\{ \|\hat{\varepsilon}_1(\boldsymbol{x}, \hat{\boldsymbol{y}}) - \hat{\varepsilon}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{L^\infty(\hat{Y})} + \frac{1}{\omega}\,\|\hat{\sigma}_1(\boldsymbol{x}, \hat{\boldsymbol{y}}) - \hat{\sigma}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{L^\infty(\hat{\Sigma})} \right\} \\
&\leq\ C(\hat{\boldsymbol{q}}_2) \left\{ \|\hat{\boldsymbol{q}}_1 - \hat{\boldsymbol{q}}_2\|^2_{L^\infty(\hat{Y})} + \|\hat{\boldsymbol{q}}_1 - \hat{\boldsymbol{q}}_2\|^2_{L^\infty(\hat{\Sigma})} \right\},
\end{aligned}
$$

*where the constant $C$ only depends on $\hat{F}_2(\hat{\boldsymbol{y}})$ and thus on $\hat{\boldsymbol{q}}_2$.*

A detailed proof of Theorem 2.11 is given in Appendix A.

*Remark* 2.12. The Lipschitz continuity of $\hat{\boldsymbol{q}} \mapsto \hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ implies that this mapping is already Gâteaux differentiable almost everywhere, see, e.g., [7]. This justifies to take Gâteaux derivatives with respect to the control $\hat{\boldsymbol{q}}$ in the adjoint formulation; see Section 3.

THEOREM 2.13. *Let $\hat{\Sigma}$ be a smooth, closed hypersurface, i. e., $\partial\hat{\Sigma} = 0$. Suppose there exist a smooth, $\hat{Y}$-periodic extensions $\hat{\boldsymbol{\tau}}_i(\boldsymbol{x}) : \hat{Y} \to \mathbb{R}^n$ of the tangential fields $\hat{\boldsymbol{\tau}}_i$ with $|\hat{\boldsymbol{\tau}}_i(\boldsymbol{x})| \leq 1$ for all $\boldsymbol{x} \in \hat{Y}$. Then, under the assumptions stated in Theorem 2.10 and provided that $\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ and $\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ are sufficiently regular the following stability estimate holds true:*

$$
\begin{aligned}
(2.19)\quad & \|\hat{\nabla}\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega}\|\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})} \\
&\qquad\leq\ C\ \max\left\{ \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})}, \frac{1}{\omega}\|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})} \right\} \\
&\quad\times \left\{ \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})} \|\hat{F}^T \boldsymbol{e}_i\|^2_{H^1(\hat{Y})} + \frac{1}{\omega}\|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})} \|\hat{F}^T \boldsymbol{e}_i\|^2_{H^1(\hat{\Sigma})} \right\}.
\end{aligned}
$$

*Here, the constant $C$ only depends on $\hat{\Sigma}$ and the chosen extension of $\hat{\boldsymbol{\tau}}$.*

A detailed proof of Theorem 2.13 is given in Appendix B.

**3. Shape optimization problem and adjoint formulation.** The previous section establishes that the effective permittivity tensor $\varepsilon^{\text{eff}}(\hat{\boldsymbol{\chi}}; \hat{\boldsymbol{q}})$ given by (2.17) enjoys a sufficient degree of regular dependence on $\hat{\boldsymbol{q}}$ to formulate an optimization problem; and solve it by means of derivative based optimization methods. To this end, we introduce a cost functional with the target to minimize the Frobenius norm between $\varepsilon^{\text{eff}}(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{q}})$ and a given target permittivity tensor. A particular difficulty that has to be addressed is the necessity to maintain a lower bound, $\hat{J} \geq \delta > 0$, on the transformation determinant. We thus set:

DEFINITION 3.1. *For a given target tensor $\varepsilon^{target}$, introduce a cost function*

$$C(\hat{\boldsymbol{\chi}};\hat{\boldsymbol{q}}) := \frac{1}{2}\left\|\varepsilon^{\mathit{eff}}(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{q}}) - \varepsilon^{\mathit{tgt.}}\right\|_{Fr.}^2 + \frac{\alpha}{2}\|\nabla\hat{\boldsymbol{q}}\|_w^2 + \beta\int_{\hat{Y}} P(\hat{\boldsymbol{y}};\hat{\boldsymbol{q}})\,\mathrm{d}\hat{y}, \tag{3.1}$$

$$\|\nabla\hat{\boldsymbol{q}}\|_w^2 := \int_Y w(\hat{\boldsymbol{y}})|\nabla\hat{\boldsymbol{q}}|^2\,\mathrm{d}\hat{y}, \quad P(\hat{\boldsymbol{y}};\hat{\boldsymbol{q}}) := \begin{cases} \frac{1}{2}\frac{(\hat{J}(\hat{\boldsymbol{y}})-1)^2}{|\hat{J}(\hat{\boldsymbol{y}})|+\hat{J}(\hat{\boldsymbol{y}})} & \text{if } \hat{J}(\hat{\boldsymbol{y}}) < 1, \\ \frac{1}{2}\,(\hat{J}(\hat{\boldsymbol{y}})-1)^2 & \text{if } \hat{J}(\hat{\boldsymbol{y}}) \geq 1. \end{cases}$$

*Here, $\alpha > 0$ is an appropriately chosen Tikhonov regularization parameter, and the coefficients $\beta > 0$ control a penalty on the deviation of transformation determinant from $\hat{J} = 1$. Moreover, $w(\hat{\boldsymbol{y}}) > 0$ is a weight function that will be chosen later.*

The penalty term $P(\hat{\boldsymbol{y}};\hat{\boldsymbol{q}})$ is chosen to provide a barrier that enforces positivity of the transformation gradient $\hat{J}$ and penalizes a deviation away from 1. Strictly speaking, this penalty only enforces positivity of the transformation gradient, $\hat{J} > 0$, but not a uniform lower bound. This is not a problem for the discretized setting of our numerical computations (see Section 4) because the finite dimensionality will ensure that $\hat{J}$ remains bounded away from 0, though with a possibly discretization dependent constant $\delta_h$. Nevertheless, if a guaranteed lower bound $\delta$ is desired then $P(\hat{\boldsymbol{y}};\hat{\boldsymbol{q}})$ can be easily modified to accommodate this by substituting $\hat{J}$ by $\hat{J} - \delta$.

We now seek solutions $(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{q}}) \in X := \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \times \boldsymbol{D}(\hat{Y},\hat{\Sigma})$ of the optimization problem

$$\min_{(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{q}})\in X} C(\hat{\boldsymbol{\chi}};\hat{\boldsymbol{q}}) \quad \text{subject to} \quad E(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{\varphi}};\hat{\boldsymbol{q}}) = 0 \ \text{ for all } \ \hat{\boldsymbol{\varphi}} \in \boldsymbol{H}(\hat{Y},\hat{\Sigma}). \tag{3.2}$$

**3.1. Adjoint formulation.** In order to derive the optimality condition (3.2), we would have to compute a partial derivative in multiple directions, which is inconvenient. Therefore, we use an adjoint formulation, see, e.g., [5, 20, 43].

DEFINITION 3.2. *Define a Lagrangian*

$$\mathcal{L} : \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \times \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \times \boldsymbol{D}(\hat{Y},\hat{\Sigma}) \to \mathbb{C},$$
$$\mathcal{L}(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{z}};\hat{\boldsymbol{q}}) = C(\hat{\boldsymbol{\chi}};\hat{\boldsymbol{q}}) - E(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{z}};\hat{\boldsymbol{q}}),$$

*where we have introduced a Lagrange multiplier $\hat{\boldsymbol{z}}$ for the PDE constraint in (3.2). For a given deformation field $\hat{\boldsymbol{q}} \in H_0^1(\hat{Y},\mathbb{C})^3$ we further introduce a state equation:*

$$\text{find } \hat{\boldsymbol{\chi}} \in \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \text{ s.t. } \mathcal{L}'_{\hat{\boldsymbol{z}}}(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{z}};\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{z}}] = 0 \ \ \forall\delta\hat{\boldsymbol{z}} \in \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \tag{3.3}$$

*and denote the solution of the state equation by $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$. Here, $\mathcal{F}'_{\hat{\boldsymbol{z}}}[\delta\hat{\boldsymbol{z}}]$ denotes the Gâteaux derivative of a functional $F$ with respect to $\hat{\boldsymbol{z}}$ in direction $\delta\hat{\boldsymbol{z}}$. We then introduce an adjoint equation:*

$$\text{find } \hat{\boldsymbol{z}} \in \boldsymbol{H}(\hat{Y},\hat{\Sigma}) \text{ s.t. } \mathcal{L}'_{\hat{\boldsymbol{\chi}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}),\hat{\boldsymbol{z}};\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{\chi}}] = 0 \ \ \forall\delta\hat{\boldsymbol{\chi}} \in \boldsymbol{H}(\hat{Y},\hat{\Sigma}). \tag{3.4}$$

The central observation is the fact that a solution to (3.2) is a critical point of $\mathcal{L}(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{z}};\hat{\boldsymbol{q}})$; see [5, 20, 43]. For the sake of completeness we summarize:

LEMMA 3.3 (First order necessary conditions [5, 20, 43]). *The solution $(\hat{\boldsymbol{\chi}}^*,\hat{\boldsymbol{q}}^*)$ of (3.2) coincides with a critical point $(\hat{\boldsymbol{\chi}}^*,\hat{\boldsymbol{z}}^*,\hat{\boldsymbol{q}}^*)$ of the Lagrangian $\mathcal{L}(\hat{\boldsymbol{\chi}},\hat{\boldsymbol{z}};\hat{\boldsymbol{q}})$.*

*Proof.* Let $(\hat{\boldsymbol{\chi}}^*,\hat{\boldsymbol{q}}^*)$ be a solution to (3.2) and let $\hat{\boldsymbol{z}}^*$ be the solution to the adjoint equation (3.4). We then have $\mathcal{L}'_{\hat{\boldsymbol{z}}}(\hat{\boldsymbol{\chi}}^*,\hat{\boldsymbol{z}}^*;\hat{\boldsymbol{q}}^*)[\delta\hat{\boldsymbol{z}}] = 0$ and $\mathcal{L}'_{\hat{\boldsymbol{\chi}}}(\hat{\boldsymbol{\chi}}^*,\hat{\boldsymbol{z}}^*;\hat{\boldsymbol{q}}^*)[\delta\hat{\boldsymbol{\chi}}] = 0$ by virtue of (3.3) and (3.4).

**Algorithm 3.1** Computing the $H^1$-gradient $\boldsymbol{\delta c}(\hat{\boldsymbol{q}})$ of the cost functional $c(\hat{\boldsymbol{q}})$ by means of the adjoint formulation; see Lemma 3.3.

Given $\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$

a) compute a solution $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}) \in \boldsymbol{H}(\hat{Y}, \hat{\Sigma})$ of the state equation (3.3),

$$\mathcal{L}'_{\hat{\boldsymbol{z}}}(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{z}}; \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{z}}] = -E(\hat{\boldsymbol{\chi}}, \delta\hat{\boldsymbol{z}}; \hat{\boldsymbol{q}}) = 0 \ \ \forall \delta\hat{\boldsymbol{z}} \in \boldsymbol{H}(\hat{Y}, \hat{\Sigma});$$

b) compute a solution $\hat{\boldsymbol{z}}(\hat{\boldsymbol{q}})$ of the adjoint equation (3.4),

$$\mathcal{L}'_{\hat{\boldsymbol{\chi}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}; \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{\chi}}] = 0 \ \ \forall \delta\hat{\boldsymbol{\chi}} \in \boldsymbol{H}(\hat{Y}, \hat{\Sigma});$$

c) solve the gradient equation (3.6),

$$\int_{\hat{Y}} \nabla \boldsymbol{\delta c}(\hat{\boldsymbol{q}}) \cdot \nabla \delta\hat{\boldsymbol{q}} \, \mathrm{d}\hat{y} \ = \ \mathcal{L}'_{h,\hat{\boldsymbol{q}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] \qquad \forall \delta\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma}).$$

Return $\boldsymbol{\delta c}(\hat{\boldsymbol{q}})$.

For an arbitrary deformation $\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ let $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ denote the unique solution to the state equation (3.3), as well as $\hat{\boldsymbol{z}}(\hat{\boldsymbol{q}})$ denote the unique solution to the adjoint equation (3.4). We now introduce the functional $c(\hat{\boldsymbol{q}}) := C(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}}) = \mathcal{L}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})$ and make the observation that $\hat{\boldsymbol{q}}$, by virtue of being an optimum, is necessarily a critical point of $c(\hat{\boldsymbol{q}})$, i. e., $c'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] = 0$ for all $\delta\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$. Using above identity and the chain rule we get for all $\delta\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$:

$$\begin{aligned} 0 \ &= \ c'(\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] \\ &= \ \mathcal{L}'_{\hat{\boldsymbol{\chi}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})\big[\hat{\boldsymbol{\chi}}'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}]\big] + \mathcal{L}'_{\hat{\boldsymbol{z}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\hat{\boldsymbol{z}}'(\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}]] \\ &\quad + \mathcal{L}'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] \\ &= \ \mathcal{L}'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}], \end{aligned} \tag{3.5}$$

where for the last equality we have exploited the fact that the first two terms vanish due to $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ and $\hat{\boldsymbol{z}}(\hat{\boldsymbol{q}})$ solving the state and adjoint equations, respectively. $(\hat{\boldsymbol{\chi}}^*, \hat{\boldsymbol{z}}^*, \hat{\boldsymbol{q}}^*)$ is thus a critical point of the Lagrangian $\mathcal{L}(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{z}}; \hat{\boldsymbol{q}})$. ∎

The previous lemma is based on the fact the Lagrangian $\mathcal{L}$ can be used to describe the derivative of $c(\hat{\boldsymbol{q}})$ [5, 20, 43]; for all $\delta\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$:

$$c'(\hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] \ = \ \mathcal{L}'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}].$$

We will make use of this result in the numerical algorithm to compute a gradient direction for the deformation. Noting that $c'(\hat{\boldsymbol{q}})$ is an element of the dual space of $H^1_0(\hat{Y}, \mathbb{C})^3$, we find the Riesz-representation, or gradient, as follows:

Definition 3.4 (Gradient equation). *Given a deformation field $\hat{\boldsymbol{q}}$ and corresponding $\hat{\boldsymbol{z}}(\hat{\boldsymbol{q}})$, $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ given by (3.3) and (3.4), we find $\boldsymbol{\delta c}(\hat{\boldsymbol{q}}) \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ by solving the* gradient equation

$$\int_{\hat{Y}} \nabla \boldsymbol{\delta c}(\hat{\boldsymbol{q}}) \cdot \nabla \delta\hat{\boldsymbol{q}} \, \mathrm{d}\hat{y} \ = \ \mathcal{L}'_{\hat{\boldsymbol{q}}}(\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}}), \hat{\boldsymbol{z}}(\hat{\boldsymbol{q}}); \hat{\boldsymbol{q}})[\delta\hat{\boldsymbol{q}}] \qquad \forall \delta\hat{\boldsymbol{q}} \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma}). \tag{3.6}$$

**Algorithm 3.2** The inverse damped BFGS algorithm for finding an approximate solution $\hat{\boldsymbol{q}}_h^*$ of the optimization problem (3.2).

Given an initial guess $\hat{\boldsymbol{q}}_h^0 \in \boldsymbol{D}_h$ and an initial approximate inverse Hessian operator $B^0 \in \mathcal{L}(\boldsymbol{D}_h, \boldsymbol{D}_h)$ iterate:

a) Compute the gradient $\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n) \in \boldsymbol{D}_h$ with Algorithm 3.1, and obtain a search direction $\boldsymbol{p}^n \in \boldsymbol{D}_h$ by setting

$$\boldsymbol{p}^n = -B^n \boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n).$$

b) Perform an Armijo backtracking line search with parameters $\beta \in (0,1)$ and $\gamma \in (0, 1/2)$ to find the maximal $\lambda^n \in \{1, \beta, \beta^2, \dots\}$ satisfying the *Armijo condition* $\boldsymbol{c}_h(\hat{\boldsymbol{q}}_h^n + \lambda^n \boldsymbol{p}^n) \leq \boldsymbol{c}_h(\hat{\boldsymbol{q}}_h^n) + \gamma \lambda^n (\nabla \boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n), \nabla p^n)$. Then, update

$$\boldsymbol{s}^n := \lambda^n \boldsymbol{p}^n, \quad \hat{\boldsymbol{q}}_h^{n+1} = \hat{\boldsymbol{q}}_h^n + \boldsymbol{s}^n.$$

c) Update the approximate inverse Hessian matrix using a damped inverse BFGS update formula derived in [19] to assert positive definiteness of the operator $B^{n+1}$. To this end, compute

$$\boldsymbol{y}^n = \boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^{n+1}) - \boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n).$$

Define a scaling parameter

$$\theta_n = \begin{cases} 1 & (\nabla \boldsymbol{y}^n, \nabla \boldsymbol{s}^n) \geq 0.2 (\nabla \boldsymbol{y}^n, \nabla B^n \boldsymbol{y}^n), \\ 0.8 \frac{(\nabla \boldsymbol{y}^n, \nabla B^n \boldsymbol{y}^n)}{(\nabla \boldsymbol{y}^n, \nabla B^n \boldsymbol{y}^n) - (\nabla \boldsymbol{y}^n, \nabla \boldsymbol{s}^n)} & \text{otherwise,} \end{cases}$$

and with $\widehat{\boldsymbol{s}}^n = \theta_n \boldsymbol{s}^n + (1-\theta_n) B^n \boldsymbol{y}^n$ set

$$\begin{aligned} B^{n+1} \;=\; B^n + \frac{(\widehat{\boldsymbol{s}}^n - B^n \boldsymbol{y}^n)(\nabla \widehat{\boldsymbol{s}}^n, \nabla \cdot) + \widehat{\boldsymbol{s}}^n (\nabla (\widehat{\boldsymbol{s}}^n - B^n \boldsymbol{y}^n), \nabla \cdot)}{(\nabla \boldsymbol{y}^n, \nabla \widehat{\boldsymbol{s}}^n)} \\ - \frac{(\nabla (\widehat{\boldsymbol{s}}^n - B^n \boldsymbol{y}^n), \nabla \boldsymbol{y}^n)}{(\nabla \boldsymbol{y}^n, \nabla \widehat{\boldsymbol{s}}^n)^2} \widehat{\boldsymbol{s}}^n (\nabla \widehat{\boldsymbol{s}}^n, \nabla \cdot), \end{aligned}$$

where, of course, the matrix corresponding to the operator $B^{n+1}$ is never constructed directly. Instead, the application of $B^{n+1}$ to the direction $\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n)$ is computed by storing the vectors $\widehat{\boldsymbol{s}}^n$ and $B^n \boldsymbol{y}^n$.

d) If the stopping criterion was reached, return $\hat{\boldsymbol{q}}_h^{n+1}$, otherwise continue at (a).

**3.2. Finite element discretization and optimization framework.** For our numerical tests we use the optimization toolkit `DOpElib` [13] which is based on the finite element library `deal.II` [3, 4]. The library supports a variety of finite element formulations based on quadrilateral (in 2d) and hexahedrical (in 3d) meshes.

Let $\mathcal{T}_h$ be a partition of $\hat{Y}$ into shape-regular (quadrilateral or) hexahedral elements that are *fitted* to the hypersurface $\hat{\Sigma}$. This is to say, we make the assumption that every element that is intersected by $\hat{\Sigma}$ has a face for which all four vertices of the face are located on $\hat{\Sigma}$. We denote by $\hat{\Sigma}_h$ the set of all faces for which all vertices of the face are located on $\hat{\Sigma}$. By slight abuse of notation we will interpret $\hat{\Sigma}_h$ either as a collection of faces or as the polyhedral hypersurface created by the union of all faces.

We denote by $\{\phi_i^h\}_{i=1}^{\mathcal{N}}$ the Lagrange basis of $\mathbb{Q}_1(\mathcal{T}_h)$, the space of piecewise (bilinear) trilinear finite elements defined on $\mathcal{T}_h$. Note that the tangential derivative $\hat{\nabla}_{\hat{\mathbf{t}}}\varphi$ on a face $f \in \hat{\Sigma}_h$ of a finite element function $\varphi \in \mathbb{Q}_1(\mathcal{T}_h)$ is single valued. We can thus introduce a discrete bilinear form corresponding to (2.16)

$$E_h(\hat{\boldsymbol{\chi}}_h, \hat{\boldsymbol{\varphi}}_h; \hat{\boldsymbol{q}}_h) := \int_{\hat{Y}} \hat{\varepsilon}_h(\boldsymbol{x}, \hat{\boldsymbol{y}})\big(\hat{F}_h(\hat{\boldsymbol{y}})^T + \hat{\nabla}\hat{\boldsymbol{\chi}}_h^T\big) \cdot \hat{\nabla}\hat{\boldsymbol{\varphi}}_h^T \,\mathrm{d}\hat{y} - \frac{1}{i\omega}\int_{\hat{\Sigma}_h} \hat{\sigma}_h(\boldsymbol{x}, \hat{\boldsymbol{y}})\big((\hat{F}_h(\hat{\boldsymbol{y}})^T)_{\hat{\mathbf{t}}} + \hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}_h^T\big) \cdot \hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\varphi}}_h^T \,\mathrm{d}\hat{o}_{\hat{y}}, \tag{3.7}$$

for $\hat{\boldsymbol{\chi}}_h, \hat{\boldsymbol{\varphi}}_h \in \boldsymbol{H}_h := \mathbb{Q}_1(\mathcal{T})^{2\times 3}$ and $\hat{\boldsymbol{q}}_h \in \boldsymbol{D}_h := \mathbb{Q}_1(\mathcal{T})^3$. Here, $\hat{\varepsilon}_h$, $\hat{\sigma}_h$ and $\hat{F}_h$ are computed with respect to the discrete objects $\hat{\boldsymbol{q}}_h$ and $\hat{\Sigma}_h$. Similarly, we introduce a discrete counterpart $\varepsilon_h^{\mathrm{eff}}$ of the effective permittivity tensor given by (2.17) and we set

$$C_h(\hat{\boldsymbol{\chi}}_h; \hat{\boldsymbol{q}}_h) := \frac{1}{2}\left\|\varepsilon_h^{\mathrm{eff}}(\hat{\boldsymbol{\chi}}_h; \hat{\boldsymbol{q}}_h) - \varepsilon^{\mathrm{trgt}}\right\|_{\mathrm{Fr.}}^2 + \frac{\alpha}{2}\left\|\nabla\hat{\boldsymbol{q}}_h\right\|_{w_h}^2 + \beta\int_{\hat{Y}} P(\hat{\boldsymbol{y}}; \hat{\boldsymbol{q}}_h)\,\mathrm{d}\hat{y}. \tag{3.8}$$

Here, we choose the following weight function in order to penalize more the deformation gradient on mesh cells at the interface that ensures that the discrete interface $\Sigma_h$ retains a sufficient degree of smoothness:

$$w_h(\hat{\boldsymbol{y}}) = \begin{cases} 1 + \alpha_\Sigma/\mathrm{diam}K & \text{for } \hat{\boldsymbol{y}} \in K \text{ with } \partial K \cap \delta\Sigma_h \neq \emptyset, \\ 1 & \text{otherwise.} \end{cases}$$

By introducing a discrete Lagrangian

$$\mathcal{L}_h \;:\; \boldsymbol{H}_h \times \boldsymbol{H}_h \times \boldsymbol{D}_h \to \mathbb{C},$$
$$\mathcal{L}(\hat{\boldsymbol{\chi}}_h, \hat{\boldsymbol{z}}_h; \hat{\boldsymbol{q}}_h) = C_h(\hat{\boldsymbol{\chi}}; \hat{\boldsymbol{q}}) - E_h(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{z}}; \hat{\boldsymbol{q}}),$$

we are now in a position to use the adjoint formulation to compute a discrete gradient [13]; see Algorithm 3.1. We use an inverse BFGS algorithm with damping as proposed in [19] to find an approximate minimum of $C_h(\hat{\boldsymbol{q}}_h)$; see Algorithm 3.2. The $\lambda^n$ in the above algorithm is the step size it is chosen by an Armijo backtracking linesearch and changes at every iteration. As mentioned, $B^n$ is an approximate inverse Hessian matrix. The update procedure in Algorithm 3.2(c) ensures that $B^{n+1}$ remains symmetric and positive definite. The initial value $B^0$ is chosen as

$$B_0 = \frac{-1}{\alpha}\Delta_h^{-1} \colon \boldsymbol{D}_h \to \boldsymbol{D}_h$$

in order to exactly recover the inverse Hessian of the control cost as suggested by the local convergence theory outlined in [14, 22].

**4. Numerical illustrations.** In this section, we discuss a number of numerical results to illustrate the performance of the shape optimization algorithm. We implemented the algorithm in a small C++ code using the optimization toolkit `DOpElib` [13] which is based on the finite element library `deal.II` [3, 4]. We have made our source code publicly available on Zenodo[1] [6]. For the sake of simplicity we restrict our numerical computations to 2D by assuming translation invariance in the $z$-direction. Then,

[1] https://zenodo.org/records/10459309

| $\boldsymbol{\beta}$ | $\boldsymbol{\alpha}$ | $\boldsymbol{\alpha_\Sigma}$ | $\varepsilon^{\text{eff}}_{xx}$ | $\varepsilon^{\text{eff}}_{xy}$ | deviation | optimality |
|---|---|---|---|---|---|---|
| 0.2 | 0.01 | 10 | $0.8244 + 0.0200\text{i}$ | $0.0444 - 0.0174\text{i}$ | 4.63% | 0.00503 |
| 0.2 | 0.001 | 100 | $0.8233 + 0.0202\text{i}$ | $0.0441 - 0.0170\text{i}$ | 4.51% | 0.00739 |
| 0.2 | 0.001 | 10 | $0.8057 + 0.0133\text{i}$ | $0.0484 - 0.0090\text{i}$ | 1.70% | 0.00890 |
| 0.1 | 0.01 | 10 | $0.8239 + 0.0204\text{i}$ | $0.0440 - 0.0174\text{i}$ | 4.61% | 0.02082 |
| 0.1 | 0.001 | 100 | $0.8230 + 0.0206\text{i}$ | $0.0438 - 0.0171\text{i}$ | 4.51% | 0.02598 |
| 0.1 | 0.001 | 10 | $0.8052 + 0.0140\text{i}$ | $0.0481 - 0.0090\text{i}$ | 1.72% | 0.01366 |
| 0.05 | 0.01 | 10 | $0.8235 + 0.0207\text{i}$ | $0.0438 - 0.0173\text{i}$ | 4.59% | 0.02360 |
| 0.05 | 0.001 | 100 | $0.8228 + 0.0209\text{i}$ | $0.0436 - 0.0170\text{i}$ | 4.51% | 0.07232 |
| 0.05 | 0.001 | 10 | $0.8049 + 0.0145\text{i}$ | $0.0480 - 0.0090\text{i}$ | 1.73% | 0.00472 |

Table 1: Diagonal and off-diagonal components of $\varepsilon^{\text{eff}}(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{q}})$, relative deviation $\left\|\varepsilon^{\text{eff}} - \varepsilon^{\text{target}}\right\|_{\text{Fr.}} / \left\|\varepsilon^{\text{target}}\right\|_{\text{Fr.}}$, optimality, i. e., relative norm of the reduced gradient, for different values of stabilization parameters $\beta$, $\alpha$ and $\alpha_\Sigma$ (rows). Results are shown for a fixed number of 200 iterations of the BFGS algorithm 3.2. The initial deviation was 10.97%.

| $\boldsymbol{\beta}$ | $\boldsymbol{\alpha}$ | $\boldsymbol{\alpha_\Sigma}$ | $\varepsilon^{\text{eff}}_{xx}$ | $\varepsilon^{\text{eff}}_{xy}$ | deviation | optimality |
|---|---|---|---|---|---|---|
| 0.2 | 0.001 | 10 | $0.8057 + 0.0135\text{i}$ | $0.0483 - 0.0092\text{i}$ | 1.73% | 0.00667 |
| 0.1 | 0.001 | 10 | $0.8052 + 0.0141\text{i}$ | $0.0481 - 0.0092\text{i}$ | 1.74% | 0.00747 |
| 0.05 | 0.001 | 10 | $0.8048 + 0.0145\text{i}$ | $0.0480 - 0.0091\text{i}$ | 1.75% | 0.00533 |

Table 2: Subset of the parameter study reported in Table 1 but with 5 instead of 6 global refinement steps resulting in 13 312 quadrilaterals which amounts to 108 040 degrees of freedom for the (complex tensor-valued) state problem and in 27 010 for the (vector-valued) control problem.

the cell problem (2.12) and the averaging (2.14) reduce to a 2D problem. Throughout the section we have chosen $\varepsilon = \boldsymbol{I}$ and the surface conductivity to be given by (2.3) with a fixed frequency of either $\omega = 0.3$, or $\omega = 0.5$. The reference geometry consists of an inscribed circle $\hat{\Sigma}$ at the center of $\hat{Y} = [0, 1]^2$ with a radius of $r = 0.3$; see Figure 2. We have chosen a fixed spatial discretization $\mathcal{T}_h$ of 53 248 quadrilaterals (fitted to the hypersurface $\hat{\Sigma}$) which amounts to 429 062 degrees of freedom for the (complex tensor-valued) state problem and 107 266 for the (vector-valued) control problem.

**4.1. Influence of the regularization parameters.** We first present a parameter study to assess the influence of the regularization parameters $\alpha$ and $\beta$ found in (3.1) on the target tensor $\varepsilon^{\text{eff}}$ and the deformation field $\hat{\boldsymbol{q}}$. The optimization problem without stabilization terms in (3.1) is highly ill-posed; the main reason being the fact that the deformation vector $\hat{\boldsymbol{q}}$ has no influence on the target functional away from the interface $\hat{\Sigma}$. Thus, a reasonable amount of penalization is required to (a) ensure consistent mesh regularity (i. e. $\hat{J}$ being reasonably close to 1), and (b) allow the geometry to deform sufficiently to actually obtain an effective tensor $\varepsilon^{\text{eff}}(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{q}})$ close to the target $\varepsilon^{\text{target}}$. For the parameter study we choose a target permittivity tensor of

$$\varepsilon^{\text{target}} = \begin{pmatrix} 0.8 + 0.008\text{i} & 0.05 \\ 0.05 & 0.8 + 0.008\text{i} \end{pmatrix},$$

which has a moderate initial relative deviation $\|\varepsilon^{\text{eff}} - \varepsilon^{\text{target}}\|_{\text{Fr.}} / \|\varepsilon^{\text{target}}\|_{\text{Fr.}}$ of around 10.97%. We choose to perform a fixed number of 200 steps of Algorithm 3.2 without an active stopping criterion. Results for varying degrees of regularization $\alpha = 0.01$, 0.001, $\alpha_\Sigma = 10$, 100, and $\beta = 0.2$, 0.1, 0.05 are reported in Table 1. We see in the above experiments that the combination of $\alpha$ and $\alpha_\Sigma$ have the largest influence on the achieved deviation from the target tensor. Lower values of the stabilization parameters result in smaller deviations; see Table 1. As a last figure of merit we also report the achieved *optimality*, i. e., the norm of the reduced gradient normalized over the initial value: $\|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n)\| / \|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^0)\|$ for the final step $n = 200$. Here, we observe that the highest reduction after 200 steps with an optimality of around 0.005 is achieved for the choice $\alpha = 0.001$, $\alpha_\Sigma = 10$, $\beta = 0.05$. However, if the stabilization parameters are chosen too small, the mesh can degrade, in particular near the edges of the interface $\Sigma_h$. Thus, in order to balance both these factors, we make a conservative choice of $\alpha = 0.001$, $\alpha_\Sigma = 10$ and $\beta = 0.1$ for all subsequent numerical tests.

As a final test we examine the influence of mesh refinement on the numerical result and rerun the case of $\alpha = 0.001$, $\alpha_\Sigma = 10$ and $\beta = 0.2$, 0.1, 0.05 with a lower resolution of 13 312 quadrilaterals resulting in 108 040 degrees of freedom for the (complex tensor-valued) state problem and in 27 010 for the (vector-valued) control problem; see Table 2. The final $\varepsilon^{\text{eff}}$ values after 200 iterations are very close to the results obtained for 6 global refinement steps; cf Table 1. We conclude that the chosen resolution of 53 248 quadrilaterals is appropriate with minimal influence on the optimization result.

**4.2. Optimizing an epsilon-near-zero material.** We now illustrate the shape optimization procedure for three different target permittivity tensors given by

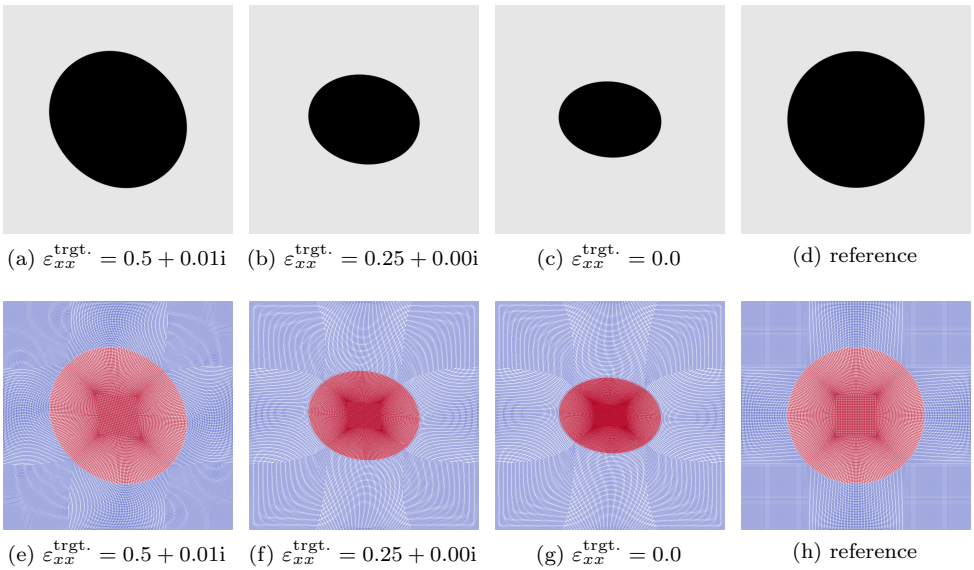
$$\varepsilon^{\text{target}} = \begin{pmatrix} * & 0.05 \\ 0.05 & 0.5 + 0.01\text{i} \end{pmatrix},$$

where we vary the $\varepsilon_{xx}^{\text{target}}$ component from (a) $0.5 + 0.01\text{i}$, (b) $0.25 + 0.005\text{i}$, to (c) 0.0. The target tensor has been chosen close to the initial permittivity tensor of

$$\varepsilon_{\text{ref}}^{\text{eff}} = \begin{pmatrix} 0.50304 + 0.01114\text{i} & 0.0 \\ 0.0 & 0.50304 + 0.01114\text{i} \end{pmatrix},$$

obtained for the reference configuration with frequency $\omega = 0.3$. As the target vector gradually gets closer to an epsilon-near-zero material [27, 28] an increasingly larger mesh deformation is required to achieve an optimal configuration. We chose to add an off-diagonal value of 0.05 in the target permittivity tensor to force an increased interaction between the $x$- and $y$-directions. For our numerical computation we use the stabilization parameters discussed in Section 4.1 and a stopping criterion to achieve a reduction of the reduced gradient, viz. $\|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n)\| / \|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^0)\|$, of better than $10^{-4}$. The initial and final geometry is illustrated in Figure 3. It can be seen that proceeding from case (a) to (c) an increasingly larger mesh deformation is required, the shapes remain largely elliptic. The final permittivity tensor values and deviation are given in Table 3. With the chosen stabilization parameters we were able to improve the initial deviation to our target tensor consistently by an order of magnitude.

*Remark* 4.1. We point out that it does not seem possible in general to achieve a deviation of zero for arbitrary target permittivities [27]. This is largely due to the fact that the effective permittivity tensor $\varepsilon^{\text{eff}}$ as a function of shape and frequency possesses a well defined structure that does not allow to tune all components and the real and imaginary part arbitrarily; see [27].



(a) $\varepsilon_{xx}^{\text{trgt.}} = 0.5 + 0.01\text{i}$ (b) $\varepsilon_{xx}^{\text{trgt.}} = 0.25 + 0.00\text{i}$ (c) $\varepsilon_{xx}^{\text{trgt.}} = 0.0$ (d) reference

(e) $\varepsilon_{xx}^{\text{trgt.}} = 0.5 + 0.01\text{i}$ (f) $\varepsilon_{xx}^{\text{trgt.}} = 0.25 + 0.00\text{i}$ (g) $\varepsilon_{xx}^{\text{trgt.}} = 0.0$ (h) reference

Figure 3: Epsilon-near-zero testcase: Final geometry obtained for target cases (a), (b) and (c) with increasingly smaller $\varepsilon_{xx}^{\text{target}}$ component. The corresponding deformed (and initial) meshes are shown in (e)-(h). The black region in (a)-(d), as well as the red region in (e)-(h) show the volume surrounded by the interface $\Sigma_h$.

| | $\varepsilon_{xx}^{\text{eff}}$ | $\varepsilon_{xy}^{\text{eff}}$ | $\varepsilon_{yy}^{\text{eff}}$ | initial | final | steps |
|---|---|---|---|---|---|---|
| ref. | $0.5030 + 0.011\text{i}$ | $0.0 - 0.0\text{i}$ | $0.5030 + 0.011\text{i}$ | | | |
| (a) | $0.5041 + 0.011\text{i}$ | $0.0488 - 0.001\text{i}$ | $0.5041 + 0.011\text{i}$ | 7.09% | 0.65% | 263 |
| (b) | $0.2897 + 0.027\text{i}$ | $0.0486 - 0.002\text{i}$ | $0.5285 + 0.016\text{i}$ | 26.3% | 5.40% | 2047 |
| (c) | $0.0271 + 0.054\text{i}$ | $0.0493 - 0.003\text{i}$ | $0.5161 + 0.021\text{i}$ | 50.8% | 6.36% | 2977 |

Table 3: Epsilon-near-zero testcase: Final permittivity tensors obtained for target cases (a), (b), (c), and the starting value for the undeformed reference configuration. In addition, we report the initial and final deviation $\left\|\varepsilon^{\text{eff}} - \varepsilon^{\text{target}}\right\|_{\text{Fr.}} / \left\|\varepsilon^{\text{target}}\right\|_{\text{Fr.}}$, as well as the number of BFGS iterations in algorithm 3.2 needed to achieve convergence.

In Figure 4 we report the evolution of the optimality, $\|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n)\| / \|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^0)\|$, during the solution process. We note that with the larger deformation necessary for cases (b) and (c) a significant increase in the number of required steps to achieve convergence can be observed. As a final figure of merit we examine the evolution of the deviation, $\left\|\varepsilon^{\text{eff}} - \varepsilon^{\text{target}}\right\|_{\text{Fr.}} / \left\|\varepsilon^{\text{target}}\right\|_{\text{Fr.}}$, during the solution process; see Figure 5. We observe that the deviation converges to its final value very fast (in between 10 to 100 steps) which corresponds to the initial steep slope for the optimality; cf. Figure 4. The remaining steps are then spend on further minimizing the penalty and thus improving the overall mesh quality.

**4.3. Large mesh deformations.** As a final test, we demonstrate that our optimization algorithm can handle larger mesh deformations. For this, we use again

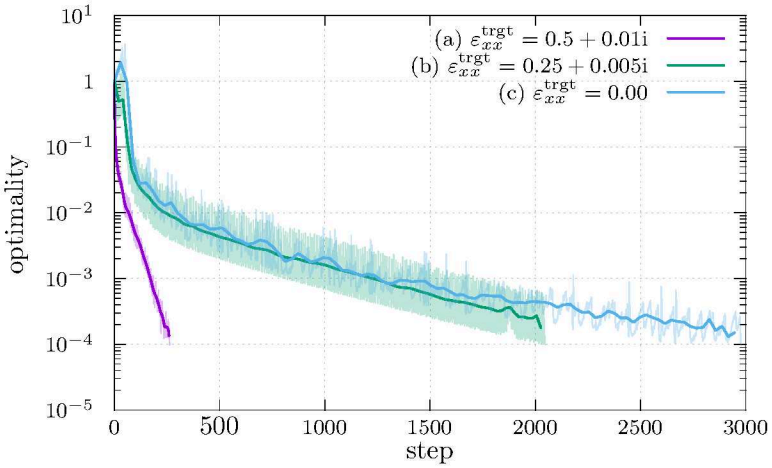



Figure 4: Evolution of the optimality, $\|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^n)\|/\|\boldsymbol{\delta c}_h(\hat{\boldsymbol{q}}_h^0)\|$, during the BFGS solution process for the three cases (a) $0.5 + 0.01\text{i}$, (b) $0.25 + 0.005\text{i}$, to (c) 0.0. The thick line for each case is a smoothed Bezier curve (gnuplot builtin) that is overlayed over the actual, oscillatory value.

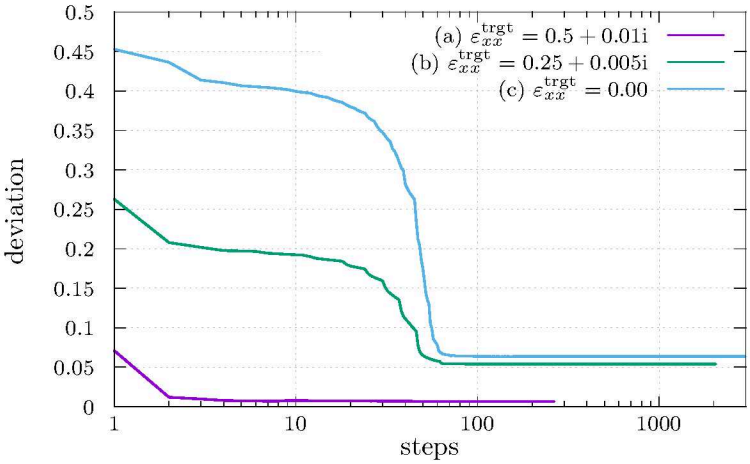
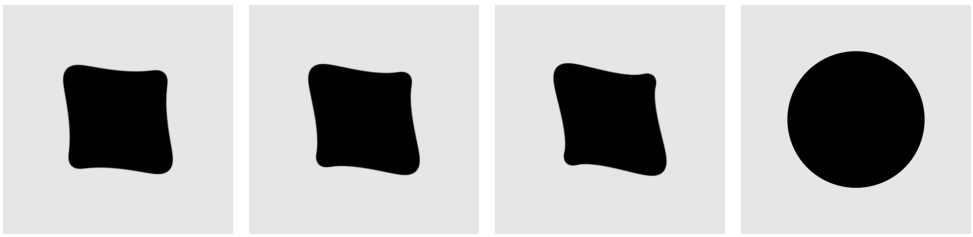



Figure 5: Evolution of the deviation, $\left\|\varepsilon^{\text{eff}} - \varepsilon^{\text{target}}\right\|_{\text{Fr.}} / \left\|\varepsilon^{\text{target}}\right\|_{\text{Fr.}}$, during the BFGS solution process for the three cases (a) $0.5 + 0.01\text{i}$, (b) $0.25 + 0.005\text{i}$, to (c) 0.0.

the target tensor introduced in Section 4.2 but vary the off-diagonal elements instead of the $xx$-component:

$$\varepsilon^{\text{target}} = \begin{pmatrix} 0.5 + 0.01 & * \\ * & 0.5 + 0.01\text{i} \end{pmatrix},$$

where we vary the $\varepsilon_{xy}^{\text{target}}$, $\varepsilon_{yx}^{\text{target}}$ components from (a) 0.10, (b) 0.15, to (c) 0.20. In addition, we set the angular frequency to $\omega = 0.4$. For visualization purposes, we also chose to run this set of computations with a lower resolution of 13 312 quadrilaterals (see also the discussion about mesh resolution in Section 4.1). The mesh deformation



(a) $\varepsilon_{xy}^{\text{trgt.}} = 0.10$ (b) $\varepsilon_{xy}^{\text{trgt.}} = 0.15$ (c) $\varepsilon_{xy}^{\text{trgt.}} = 0.20$ (d) reference

Figure 6: Large deformation testcase: Final geometry obtained for target cases (a), (b) and (c) with increasingly larger $\{\varepsilon_{xy}^{\text{target}}, \varepsilon_{yx}^{\text{target}}\}$ components.

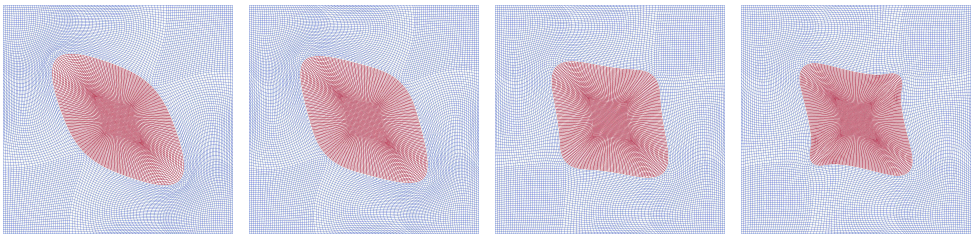
| | $\varepsilon_{xx}^{\text{eff}}$ | $\varepsilon_{xy}^{\text{eff}}$ | initial | final | steps |
|---|---|---|---|---|---|
| ref. | $0.7783 + 0.003\text{i}$ | $0.0 - 0.0\text{i}$ | | | |
| (a) | $0.5254 + 0.063\text{i}$ | $0.0953 - 0.013\text{i}$ | 41.8% | 8.58% | $0 + 512$ |
| (b) | $0.5256 + 0.063\text{i}$ | $0.1429 - 0.020\text{i}$ | 44.7% | 8.89% | $100 + 435$ |
| (c) | $0.5258 + 0.063\text{i}$ | $0.1904 - 0.027\text{i}$ | 48.4% | 9.30% | $100 + 709$ |

Table 4: Large deformation testcase: Final permittivity tensors obtained for target cases (a), (b), (c), and the starting value for the undeformed reference configuration. In addition, we report the initial and final deviation, as well as the number of BFGS iterations in Algorithm 3.2 needed to achieve convergence.

for cases (b) and (c) is so large that it would require to increase the values of the stabilization parameters significantly. This would result in a significantly higher deviation of the obtained permittivity tensor from the target tensor possibly beyond what would be deemed acceptable. We thus employ a slightly more sophisticated strategy: In cases (b) and (c) we first run 100 steps of the BFGS algorithm with a large penalty of $\beta = 0.4$ for (b) and $\beta = 0.8$ for (c) which ensures that the mesh does not degrade too much and the regular solutions shown in Figure 6 are obtained. Afterwards we fall back to the original stabilization value of $\beta = 0.1$ and run the BFGS algorithm until the stopping criterion is reached.

As reported in Figure 6 with the modified choice of target tensor and angular frequency a much more dramatic shape deformation away from a simple ellipse was achieved. The corresponding final permittivity tensors, final deviation and number of BFGS steps are reported in Table 4. We reach a final deviation of around 9% for all three cases. We report in Figure 7 the mesh evolution observed during the solution process for case (c) with $\varepsilon_{xy}^{\text{target}} = 0.2$. The most crucial part in the solution process is between steps 30 and 100: if we lower $\beta$ too much below 0.8 then the algorithm tries to approximate an eight-figure shape (and becoming singular in the process) instead of reaching the shape shown in Figure 6(c).

**5. Conclusion and outlook.** In this paper, a shape optimization problem for plasmonic crystals consisting of dielectric inclusions was discussed. The objective was to modify the shape of microscale inclusions so that the effective permittivity tensor attained a set target. To achieve this goal, a mesh deformation technique was introduced



(a) step 10, dev. 23% (b) step 30, dev. 21 (c) step 100, dev. 16% (d) st. 100+50, dev. 9%

Figure 7: Large deformation testcase: Mesh evolution for Case (c) with $\varepsilon_{xy}^{\text{target}} = 0.2$ after (a) 10, (b) 30, and (c) 100 steps with a large regularization parameter $\beta = 0.8$, and (d) after running another 50 steps with a smaller penalty of $\beta = 0.1$. The deviation value reported is the normalized difference of the effective permittivity tensor to target tensor.

in the cell problem and the underlying homogenization process. Furthermore, well-posedness and regularity of the deformed cell problem were established.

Next, an optimization algorithm based on the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton method was presented. Through a series of numerical experiments, the effectiveness of the algorithm was demonstrated. Importantly, it was shown that the formulation and choice of penalization effectively handled large mesh deformations, as evidenced in the experimental results.

*Outlook.* The current approach is limited to optimizing a microstructure for a single fixed frequency $\omega$. This seems somewhat limiting as in practice a target permittivity that is valid over a (large) frequency interval is desired. In this regard we will explore how the optimization approach can be adapted to a result reported in [27]: The frequency response of $\varepsilon_{ij}^{\text{eff}}(\omega)$ is described by a purely algebraic expression

$$\varepsilon_{ij}^{\text{eff}}(\omega) \;=\; \varepsilon\,\delta_{ij} \;-\; \eta(\omega)\,N_{ij} \;-\; \sum_{n=1}^{\infty} \frac{\lambda_n\,\eta^2(\omega)}{\varepsilon - \lambda_n\,\eta(\omega)}\; M_{jn}\,M_{in},$$

where $\eta(\omega) = \frac{\sigma(\omega)}{\mathrm{i}\omega}$, the coefficient $N_{ij}$ is a weight only depending on the geometry, and $M_{in}$, $M_{jn}$ are weights depending on eigenfunctions $\varphi_n$ with corresponding eigenvalues $\lambda_n$ that are characterized by a purely geometric eigenvalue problem:

$$\begin{cases} \Delta\varphi_n(\boldsymbol{x})\big) = 0 & \text{in } Y \setminus \Sigma\,, \\ [\varphi_n(\boldsymbol{x}))]_\Sigma = 0 & \text{on } \Sigma\,, \\ \lambda_n\,[\boldsymbol{\nu}\cdot\nabla\varphi_n(\boldsymbol{x})]_\Sigma = \nabla_T\cdot\nabla_T\varphi_n(\boldsymbol{x}) & \text{on } \Sigma\,. \end{cases}$$

Many of the ideas developed in this publication can be adapted to this eigenvalue problem. The stated goal is then to optimize the deviation of $\varepsilon_{ij}^{\text{eff}}(\omega)$ to a target permittivity over a suitable frequency band.

**Acknowledgments.** M.B. and M.M. acknowledge partial support from the National Science Foundation under grants DMS-1912846 and DMS-2045636.

## Appendix A. Proofs of Lemmas 2.7 and 2.8, and Theorem 2.11.

*Proof of Lemma 2.7.* We know that $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}}) \in \boldsymbol{D}(\hat{Y}, \hat{\Sigma})$ implying that $\hat{J}$ and $F(\hat{\boldsymbol{y}})$ are bounded with respect to $\|.\|_{L^\infty}$. Moreover, $\varepsilon(\boldsymbol{x}, \boldsymbol{y}(\hat{\boldsymbol{y}})), \sigma(\boldsymbol{x}, \boldsymbol{y}(\hat{\boldsymbol{y}}))$ are bounded, complex and tensor valued by assumption. Therefore, $\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ and $\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ are by construction also bounded, complex and tensor valued. Simultaneous multiplication by a tensor from the left and its transpose from the right preserves symmetry, and so does scaling and multiplication by $\hat{J}$. Therefore, $\operatorname{Re}\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\operatorname{Im}\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\operatorname{Re}\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$, and $\operatorname{Im}\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ are symmetric owing to the symmetry of $\varepsilon(\boldsymbol{x}, \boldsymbol{y})$ and $\sigma(\boldsymbol{x}, \boldsymbol{y})$. Similarly, the uniform ellipticity of $\operatorname{Im}\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ and $\operatorname{Re}\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ is a direct consequence of the uniform ellipticity of $\operatorname{Im}\varepsilon(\boldsymbol{x}, \boldsymbol{y})$ and $\operatorname{Re}\sigma(\boldsymbol{x}, \boldsymbol{y})$ and the uniform lower bound on the determinant of the deformation gradient: $0 < \delta < \hat{J}(\hat{\boldsymbol{y}})$. ∎

*Proof of Lemma 2.8.* We start with (2.13). Substituting $\sigma_{mn}$ given in (2.15) and using Lemma 2.6 for transforming $\boldsymbol{\tau}_m$ and $\boldsymbol{\tau}_n$ gives:

$$\begin{aligned} E(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{\varphi}}; \hat{\boldsymbol{q}}) = &\int_{\hat{Y}} \varepsilon(\boldsymbol{x}, \boldsymbol{y}(\hat{\boldsymbol{y}}))(\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{F}(\hat{\boldsymbol{y}})^{T} + \hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T}) \cdot (\hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}})\hat{J}\,\mathrm{d}\hat{y} \\ &- \frac{1}{i\omega}\sum_{m,n}\int_{\hat{\Sigma}} \sigma_{mn}\Big(\frac{\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_m}{\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_m\|} \cdot \big(\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{F}(\hat{\boldsymbol{y}})^{T} + \hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T}\big)\Big) \\ &\qquad \cdot \Big(\frac{\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_n}{\|\hat{F}(\hat{\boldsymbol{y}})\hat{\boldsymbol{\tau}}_n\|} \cdot \hat{F}(\hat{\boldsymbol{y}})^{-T}\overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}}\Big)\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}\,\mathrm{d}\hat{o}_{\hat{y}}. \end{aligned}$$

After some more simplifications and using the definitions of $\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ and $\hat{\sigma}_{mn}$ from equation (2.15):

$$\begin{aligned} E(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{\varphi}}; \hat{\boldsymbol{q}}) = &\int_{\hat{Y}} \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})(\hat{F}(\hat{\boldsymbol{y}})^{T} + \hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T}) \cdot (\overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}})\hat{J}\,\mathrm{d}\hat{y} \\ &- \frac{1}{i\omega}\sum_{m,n}\int_{\hat{\Sigma}} \hat{\sigma}_{mn}\Big(\hat{\boldsymbol{\tau}}_m \cdot (\hat{F}(\hat{\boldsymbol{y}})^{T} + \hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T})\Big) \\ &\qquad \Big(\hat{\boldsymbol{\tau}}_n \cdot \overline{\hat{\nabla}\hat{\boldsymbol{\varphi}}^{\,T}}\Big)\|\hat{F}(\hat{\boldsymbol{y}})^{-T}\hat{\boldsymbol{\nu}}\|_{\ell^2}\hat{J}\,\mathrm{d}\hat{o}_{\hat{y}} = 0. \end{aligned}$$

Now using the definition for $\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})$ given in (2.15) yields (2.16). Equation (2.14) can be transformed into (2.17) in a similar fashion. ∎

*Proof of Theorem 2.11..* We begin with $E(\hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2, \hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2; \hat{\boldsymbol{q}}_1)$. Setting $\hat{F}_1(\hat{\boldsymbol{y}}) := \boldsymbol{I} + \hat{\nabla}\hat{\boldsymbol{q}}_1(\hat{\boldsymbol{y}})$, $\hat{F}_2(\hat{\boldsymbol{y}}) := \boldsymbol{I} + \hat{\nabla}\hat{\boldsymbol{q}}_2(\hat{\boldsymbol{y}})$ and correspondingly, $\hat{J}_1$, $\hat{J}_2$, $\hat{\varepsilon}_1(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\hat{\varepsilon}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\hat{\sigma}_1(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\hat{\sigma}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})$, we observe that

$$\begin{aligned} E(\hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2, \hat{\boldsymbol{\chi}}_1 - \hat{\boldsymbol{\chi}}_2; \hat{\boldsymbol{q}}_1) \;&=\; E(\hat{\boldsymbol{\chi}}_2, \hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_1; \hat{\boldsymbol{q}}_1) - E(\hat{\boldsymbol{\chi}}_2, \hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_1; \hat{\boldsymbol{q}}_2) \\ &= \int_{\hat{Y}} \{\hat{\varepsilon}_1(\boldsymbol{x}, \hat{\boldsymbol{y}}) - \hat{\varepsilon}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})\}\hat{\nabla}\hat{\boldsymbol{\chi}}_2^{\,T} \cdot \overline{\hat{\nabla}(\hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_1)^{\,T}}\,\mathrm{d}\hat{y} \\ &\quad - \frac{1}{i\omega}\int_{\hat{\Sigma}} \{\hat{\sigma}_1(\boldsymbol{x}, \hat{\boldsymbol{y}}) - \hat{\sigma}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})\}\overline{\hat{\nabla}_{\hat{\mathbf{t}}}(\hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_2)^{\,T} \cdot \hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}_2^{\,T}}\,\mathrm{d}\hat{o}_{\hat{y}} \\ &\quad + \int_{\hat{Y}} \{\hat{\varepsilon}_1(\boldsymbol{x}, \hat{\boldsymbol{y}})\hat{F}_1(\hat{\boldsymbol{y}})^{T} - \hat{\varepsilon}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})\hat{F}_2(\hat{\boldsymbol{y}})^{T}\} \cdot \overline{\hat{\nabla}(\hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_1)^{\,T}}\,\mathrm{d}\hat{y} \\ &\quad - \frac{1}{i\omega}\int_{\hat{\Sigma}} \{\hat{\sigma}_1(\boldsymbol{x}, \hat{\boldsymbol{y}})(\hat{F}_1(\hat{\boldsymbol{y}}))_{\mathbf{t}}^{T} - \hat{\sigma}_2(\boldsymbol{x}, \hat{\boldsymbol{y}})(\hat{F}_2(\hat{\boldsymbol{y}}))_{\mathbf{t}}^{T}\} \cdot \overline{\hat{\nabla}_{\hat{\mathbf{t}}}(\hat{\boldsymbol{\chi}}_2 - \hat{\boldsymbol{\chi}}_1)^{\,T}}\,\mathrm{d}\hat{o}_{\hat{y}}. \end{aligned}$$

Now, proceeding as in the proof of Theorem 2.10

$$\begin{aligned}
&\|\hat\nabla(\hat{\boldsymbol\chi}_1-\hat{\boldsymbol\chi}_2)\|^2_{L^2(\hat Y)}+\frac{1}{\omega}\|\hat\nabla_{\hat{\mathbf t}}(\hat{\boldsymbol\chi}_1-\hat{\boldsymbol\chi}_2)\|^2_{L^2(\hat\Sigma)}\\
&\qquad\le C\Big\{\|\hat\varepsilon_1(\boldsymbol x,\hat{\boldsymbol y})-\hat\varepsilon_2(\boldsymbol x,\hat{\boldsymbol y})\|_{L^\infty(\hat Y)}\|\hat\nabla\hat{\boldsymbol\chi}_2\|_{L^2(\hat Y)}\|\hat\nabla(\hat{\boldsymbol\chi}_2-\hat{\boldsymbol\chi}_1)\|_{L^2(\hat Y)}\\
&\qquad\qquad+\frac{1}{\omega}\|\hat\sigma_1(\boldsymbol x,\hat{\boldsymbol y})-\hat\sigma_2(\boldsymbol x,\hat{\boldsymbol y})\|_{L^\infty(\hat\Sigma)}\|\hat\nabla_{\hat{\mathbf t}}\hat{\boldsymbol\chi}_2\|_{L^2(\hat\Sigma)}\|\hat\nabla_{\hat{\mathbf t}}(\hat{\boldsymbol\chi}_2-\hat{\boldsymbol\chi}_1)\|_{L^2(\hat\Sigma)}\\
&\qquad\qquad+\|\hat\varepsilon_1(\boldsymbol x,\hat{\boldsymbol y})\hat F_1(\hat{\boldsymbol y})^T-\hat\varepsilon_2(\boldsymbol x,\hat{\boldsymbol y})\hat F_2(\hat{\boldsymbol y})^T\|_{L^2(\hat Y)}\|\hat\nabla(\hat{\boldsymbol\chi}_2-\hat{\boldsymbol\chi}_1)\|_{L^2(\hat Y)}\\
&\qquad\qquad+\frac{1}{\omega}\|\hat\sigma_1(\boldsymbol x,\hat{\boldsymbol y})\hat F_1(\hat{\boldsymbol y})^T-\hat\sigma_2(\boldsymbol x,\hat{\boldsymbol y})\hat F_2(\hat{\boldsymbol y})^T\|_{L^2(\hat\Sigma}\|\hat\nabla_{\hat{\mathbf t}}(\hat{\boldsymbol\chi}_2-\hat{\boldsymbol\chi}_1)\|_{L^2(\hat\Sigma)}\Big\}
\end{aligned}$$

Young's inequality allows to absorb all factors with differences $\hat{\boldsymbol\chi}_2-\hat{\boldsymbol\chi}_1$ into the left hand side, and the factors $\|\hat\nabla\hat{\boldsymbol\chi}_2\|_{L^2(\hat Y)}$ and $\|\hat\nabla_{\hat{\mathbf t}}\hat{\boldsymbol\chi}_2\|_{L^2(\hat\Sigma)}$ can be bounded by a constant $C(\hat{\boldsymbol q}_2)$ only depending on $\hat{\boldsymbol q}_2$ (and the shape $\hat\Sigma$); see Theorem 2.10. In summary this implies that

$$\begin{aligned}
&\|\hat\nabla(\hat{\boldsymbol\chi}_1-\hat{\boldsymbol\chi}_2)\|^2_{L^2(\hat Y)}+\frac{1}{\omega}\|\hat\nabla_{\hat{\mathbf t}}(\hat{\boldsymbol\chi}_1-\hat{\boldsymbol\chi}_2)\|^2_{L^2(\hat\Sigma)}\\
&\qquad\le C(\hat{\boldsymbol q})\Big\{\|\hat\varepsilon_1(\boldsymbol x,\hat{\boldsymbol y})-\hat\varepsilon_2(\boldsymbol x,\hat{\boldsymbol y})\|^2_{L^\infty(\hat Y)}+\frac{1}{\omega}\|\hat\sigma_1(\boldsymbol x,\hat{\boldsymbol y})-\hat\sigma_2(\boldsymbol x,\hat{\boldsymbol y})\|^2_{L^\infty(\hat\Sigma)}\Big\}.
\end{aligned}$$

The final inequality now follows from the fact that for a fixed coordinate $\hat{\boldsymbol y}$ the tensors $\varepsilon$ and $\sigma$ depend analytically on $\hat{\boldsymbol q}$; see (2.15). ∎

**Appendix B. Proof of Theorem 2.13.** In order to show regularity we follow the well-established strategy of introducing a difference quotient and then passing to the limit; see for example [12]. For the sake of completeness, we restate and generalize the argument here so that it can be applied to our case of a periodic domain with a curved hypersurface. A number of preparatory steps are in order.

DEFINITION B.1. *Let $\boldsymbol\eta(x):\hat Y\to\mathbb{R}^n$ be a smooth, $\hat Y$-periodic vector field with $\boldsymbol\eta(\hat{\boldsymbol y})\cdot\boldsymbol n(\hat{\boldsymbol y})=0$ for $\hat{\boldsymbol y}\in\hat\Sigma$, where $\boldsymbol n$ denotes a fixed normal field on $\hat\Sigma$, and assume that $\boldsymbol\eta(\hat{\boldsymbol y})=\mathbf 0$ for $\hat{\boldsymbol y}\in\partial\hat\Sigma$.*

*a) For $\hat{\boldsymbol y}\in\hat Y$ let $\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}:\mathbb{R}\to\hat Y$ be the integral curve of $\boldsymbol\eta$ with $\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}(0)=\hat{\boldsymbol y}$ and $\frac{d}{ds}\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}(s)=\boldsymbol\eta(\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}(s))$ for all $s\in\mathbb{R}$. Here, by slight abuse of notation, we equip $\hat Y$ with a toroidal topology by identifying opposing periodic boundaries, so that $\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}$ admits $\mathbb{R}$ as domain of definition.*

*b) Introduce a transformation, $T_h:\hat Y\to\hat Y$ characterized by $T_h(\hat{\boldsymbol y}):=\boldsymbol\xi_{\boldsymbol\eta,\hat{\boldsymbol y}}(h)$.*

*c) For a given $f\in C^\infty_{per}(\hat Y)$ and $h>0$ introduce two difference operators:*

$$\hat\nabla^h_\eta f(\hat{\boldsymbol y}):=\frac{f\big(T_h(\hat{\boldsymbol y})\big)-f(\hat{\boldsymbol y})}{h},\qquad \hat\nabla^{-h}_\eta f(\hat{\boldsymbol y}):=\frac{f(\hat{\boldsymbol y})-f\big(T_{-h}(\hat{\boldsymbol y})\big)\kappa(\hat{\boldsymbol y})}{h}.$$

*Here, $\kappa(\hat{\boldsymbol y}):=det\big(\nabla T_{-h}(\hat{\boldsymbol y})\big)$.*

We have the following results at hand:

LEMMA B.2. *Let $\boldsymbol\eta$ be a vector field as characterized in definition B.1. Then, for all $f,g\in C^\infty_{per}(\hat Y)$ and under the assumption that $h>0$ is sufficiently small it holds*

*that*

*(1.)* $T_h$ *and* $T_{-h}$ *are automorphisms on* $\hat{Y}$ *and* $\hat{\Sigma}$, *with* $T_{-h} \circ T_h = Id$,

*(2.)* $|\nabla\kappa(\hat{\boldsymbol{y}})| \leq c\,h\,\max\big(|\nabla\boldsymbol{\eta}|_\infty(\hat{\boldsymbol{y}}), |\nabla^2\boldsymbol{\eta}(\hat{\boldsymbol{y}})|_\infty\big)$ *for all* $\hat{\boldsymbol{y}} \in \hat{Y}$,

*(3.)* $$\int_{\hat{Y}} (\hat{\nabla}^h_\eta f(\hat{\boldsymbol{y}}))g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y} = -\int_{\hat{Y}} f(\hat{\boldsymbol{y}})(\hat{\nabla}^{-h}_\eta g(\hat{\boldsymbol{y}}))\,\mathrm{d}\hat{y},$$

*(4.)* $$\int_{\hat{\Sigma}} (\hat{\nabla}^h_\eta f(\hat{\boldsymbol{y}}))g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{o}_{\hat{y}} = -\int_{\hat{\Sigma}} f(\hat{\boldsymbol{y}})(\hat{\nabla}^{-h}_\eta g(\hat{\boldsymbol{y}}))\,\mathrm{d}\hat{o}_{\hat{y}},$$

*(5.)* $$\hat{\nabla}^h_\eta(fg)(\hat{\boldsymbol{y}}) = f(T_h(\hat{\boldsymbol{y}}))(\hat{\nabla}^h_\eta g)(\hat{\boldsymbol{y}}) + (\hat{\nabla}^h_\eta f)(\hat{\boldsymbol{y}})g(\hat{\boldsymbol{y}}),$$

*(6.)* $$\lim_{h\to 0} \hat{\nabla}^h_\eta f(\hat{\boldsymbol{y}}) = \boldsymbol{\eta}(\hat{\boldsymbol{y}}) \cdot \hat{\nabla} f(\hat{\boldsymbol{y}}) =: \hat{\nabla}_\eta f(\hat{\boldsymbol{y}}).$$

*Proof.* (1.) The integral curve $\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s)$ is described by an initial value problem with smooth right hand side on $\hat{Y}$. Thus, by virtue of the Picard-Lindelöf theorem a unique solution $\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s)$ exists. Moreover, $\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s)$ depends smoothly on the initial data $\hat{\boldsymbol{y}}$ and the differential $\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s)$ is given by an initial value problem

$$\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(0) = \boldsymbol{I}, \qquad \frac{\mathrm{d}}{\mathrm{d}s}\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s) = \nabla\boldsymbol{\eta}(\xi_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s))\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s). \tag{B.1}$$

This shows that the transformation $T_h(\hat{\boldsymbol{y}}) = \boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)$ is a well-defined differentiable function. Moreover, owing to the compactness of $\hat{Y}$ there exists an $h_0$ such that $\|\nabla T_h(\hat{\boldsymbol{y}}) - \boldsymbol{I}\| \leq \delta < 1$ for all $\hat{\boldsymbol{y}} \in \hat{Y}$ and $|h| \leq h_0$. This implies that $T_h$ is injective for $|h| \leq h_0$. A consequence of the initial value problem of the integral curves is the fact that $\boldsymbol{\xi}_{\boldsymbol{\eta},\tilde{\boldsymbol{y}}}(-h) = \hat{\boldsymbol{y}}$ for $\tilde{\boldsymbol{y}} = \boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)$ for all $\hat{\boldsymbol{y}} \in \hat{Y}$. This implies that $T_{-h} \circ T_h = Id$.

(2.) We first observe that $\nabla\kappa(\hat{\boldsymbol{y}})$ is a sum of products of $(d-1)$ entries of $\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)$ and one entry of the tensor of second derivatives $\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)$. This implies that

$$|\nabla\kappa(\hat{\boldsymbol{y}})|_\infty \leq c\,|\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)|^{d-1}_\infty\,|\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)|_\infty \leq c\,|\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)|_\infty,$$

where for the second inequality we increased the constant $c$ with a uniform bound on $\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)$ that was established for $|h| \leq h_0$ in (1.). We now observe that $\nabla^2_{\hat{\boldsymbol{y}}}\hat{\boldsymbol{y}}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(0) = \boldsymbol{0}$. The tensor of second derivatives obeys an initial value problem similarly to (B.1) but with a right hand side involving $\nabla\boldsymbol{\eta}(\hat{\boldsymbol{y}})$ and $\nabla^2\boldsymbol{\eta}(\hat{\boldsymbol{y}})$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s) = \big(\nabla^2\boldsymbol{\eta}(\xi_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s))\nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s)\big) \cdot \nabla_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s) + \nabla\boldsymbol{\eta}(\xi_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s))\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s).$$

Integrating the differential equation and using the initial condition we get

$$|\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)|_\infty \;\leq\; c\int_0^h \max\Big(|\nabla\boldsymbol{\eta}|, |\nabla^2\boldsymbol{\eta}|\Big)(\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(s))\mathrm{d}s.$$

possibly shrinking $h_0$ again with a compactness argument now establishes

$$|\nabla^2_{\hat{\boldsymbol{y}}}\boldsymbol{\xi}_{\boldsymbol{\eta},\hat{\boldsymbol{y}}}(h)|_\infty \;\leq\; ch\,\max\big(|\nabla\boldsymbol{\eta}|_\infty(\hat{\boldsymbol{y}}), |\nabla^2\boldsymbol{\eta}(\hat{\boldsymbol{y}})|_\infty\big).$$

(3.) By definition,

$$\int_{\hat{Y}} (\hat{\nabla}^h_\eta f(\hat{\boldsymbol{y}}))g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y} = \int_{\hat{Y}} \frac{f(T_h(\hat{\boldsymbol{y}}))}{h} g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y} - \int_{\hat{Y}} \frac{f(\hat{\boldsymbol{y}})}{h} g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y}$$

Applying the transformation $\hat{\boldsymbol{y}} \mapsto T_{-h}(\hat{\boldsymbol{y}})$ and exploiting $\hat{Y}$-periodicity gives:

$$\begin{aligned}&= \int_{\hat{Y}} \frac{f(\hat{\boldsymbol{y}})}{h} g(T_{-h}(\hat{\boldsymbol{y}}))\kappa(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y} - \int_{\hat{Y}} \frac{f(\hat{\boldsymbol{y}})}{h} g(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{y}\\ &= -\int_{\hat{Y}} f(\hat{\boldsymbol{y}})(\hat{\nabla}_\eta^{-h} g(\hat{\boldsymbol{y}}))\,\mathrm{d}\hat{y}.\end{aligned}$$

(4.) The proof of this statement is similar to (1.) with the important detail that we have to establish that the transformed surface element is given by $\kappa(\hat{\boldsymbol{y}})\,\mathrm{d}\hat{o}_{\hat{y}}$. By definition of $\boldsymbol{\eta}$ we have $\boldsymbol{\eta}(\hat{\boldsymbol{y}}) \cdot \boldsymbol{n}(\hat{\boldsymbol{y}}) = 0$ for $\hat{\boldsymbol{y}} \in \hat{\Sigma}$. This implies that $\hat{\Sigma}$ is *parallel* to integral curves and as a consequence we have that

$$\hat{\nabla} T_h(\hat{\boldsymbol{y}}) \simeq \begin{pmatrix} \partial_{\tau_1}(T_h \cdot \boldsymbol{\tau}_1) & \partial_{\tau_2}(T_h \cdot \boldsymbol{\tau}_1) & \partial_n(T_h \cdot \boldsymbol{\tau}_1) \\ \partial_{\tau_1}(T_h \cdot \boldsymbol{\tau}_2) & \partial_{\tau_2}(T_h \cdot \boldsymbol{\tau}_2) & \partial_n(T_h \cdot \boldsymbol{\tau}_2) \\ 0 & 0 & 1 \end{pmatrix},$$

when expressing the Jacobian of $T_h(\hat{\boldsymbol{y}})$ for $\hat{\boldsymbol{y}} \in \hat{\Sigma}$ in a local coordinate system spanned by $(\boldsymbol{n}, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$.

(5.) An elementary calculation shows:

$$\begin{aligned}\hat{\nabla}_\eta^h (fg)(\hat{\boldsymbol{y}}) &= \frac{f(T_h(\hat{\boldsymbol{y}}))g(T_h(\hat{\boldsymbol{y}})) - f(T_h(\hat{\boldsymbol{y}}))g(\hat{\boldsymbol{y}})}{h}\\ &\qquad + \frac{f(T_h(\hat{\boldsymbol{y}}))g(\hat{\boldsymbol{y}}) - f(\hat{\boldsymbol{y}})g(\hat{\boldsymbol{y}})}{h}\\ &= f(T_h(\hat{\boldsymbol{y}}))(\hat{\nabla}_\eta^h g)(\hat{\boldsymbol{y}}) + g(\hat{\boldsymbol{y}})(\hat{\nabla}_\eta^h f)(\hat{\boldsymbol{y}}).\end{aligned}$$

(6.) This is an immediate consequence of the differentiability of $f(\hat{\boldsymbol{y}})$ and the L'Hôpital theorem. ∎

With this definition and lemma, we are in a position to prove the following intermediate result:

PROPOSITION B.3. *Let $\boldsymbol{\eta}(x) : \hat{Y} \to \mathbb{R}^n$ be a smooth, $\hat{Y}$-periodic vector field with $\boldsymbol{\eta}(\hat{\boldsymbol{y}}) \cdot \boldsymbol{n}(\hat{\boldsymbol{y}}) = 0$ for $\hat{\boldsymbol{y}} \in \hat{\Sigma}$, as well as $\boldsymbol{\eta}(\hat{\boldsymbol{y}}) = \boldsymbol{0}$ for all $\hat{\boldsymbol{y}} \in \partial\hat{\Sigma}$. Let $\hat{\boldsymbol{\chi}} \in H$ be the solution to (2.6). Then, provided that $\varepsilon(\boldsymbol{x}, \hat{\boldsymbol{y}})$, $\sigma(\boldsymbol{x}, \hat{\boldsymbol{y}})$ and $\hat{\boldsymbol{q}}(\hat{\boldsymbol{y}})$ are sufficiently regular:*

$$\begin{aligned}\text{(B.2)}\quad &\|\hat{\nabla}\hat{\nabla}_\eta \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega}\|\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\nabla}_\eta \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})} \le\ C\ \max\Big\{1, \|\hat{\nabla}\boldsymbol{\eta}\|_{L^\infty(\hat{Y})}, \|\hat{\nabla}^2\boldsymbol{\eta}\|_{L^\infty(\hat{Y})}\Big\}\\ &\quad\times \Big\{\|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})}\|\hat{\nabla}\hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega}\|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})}\|\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})}\\ &\qquad + \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})}\|\hat{F}^T \boldsymbol{e}_i\|^2_{H^1(\hat{Y})} + \frac{1}{\omega}\|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})}\|\hat{F}^T \boldsymbol{e}_i\|^2_{H^1(\hat{\Sigma})}\Big\}.\end{aligned}$$

*Proof.* Above assumptions on $\boldsymbol{\eta}(\hat{\boldsymbol{y}})$ ensure that $\hat{\boldsymbol{\varphi}} := \hat{\nabla}_\eta^{-h}\hat{\nabla}_\eta^h\ \hat{\boldsymbol{\chi}} \in H$. Testing (2.12) with this choice of test function $\hat{\boldsymbol{\varphi}}$ and taking the real part:

$$\begin{aligned}\text{(B.3)}\quad &\int_{\hat{Y}} \operatorname{Re}\hat{\varepsilon}\,\hat{\nabla}\hat{\boldsymbol{\chi}}^{\,T} \cdot \hat{\nabla}\hat{\nabla}_\eta^{-h}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^{\,T}\,\mathrm{d}\hat{y} + \frac{1}{\omega}\int_{\hat{\Sigma}} \operatorname{Im}\hat{\sigma}\,\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}^{\,T} \cdot \hat{\nabla}_{\hat{\mathbf{t}}}\hat{\nabla}_\eta^{-h}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^{\,T}\,\mathrm{d}\hat{o}_{\hat{y}}\\ &\qquad = -\operatorname{Re}\Big\{\int_{\hat{Y}} \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})(\hat{F}(\hat{\boldsymbol{y}})^T) \cdot \hat{\nabla}\hat{\nabla}_\eta^{-h}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^{\,T}\,\mathrm{d}\hat{y}\\ &\qquad\qquad + \frac{1}{i\omega}\int_{\hat{\Sigma}} \hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})(\hat{F}(\hat{\boldsymbol{y}})^T)_{\hat{\mathbf{t}}} \cdot \hat{\nabla}_{\hat{\mathbf{t}}}\hat{\nabla}_\eta^{-h}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^{\,T}\,\mathrm{d}\hat{o}_{\hat{y}}\Big\}.\end{aligned}$$

We now wish to move the difference operator $\hat{\nabla}_\eta^{-h}$ from the testfunction over to $\hat{\boldsymbol{\chi}}(\hat{\boldsymbol{q}})$ and the forcing terms but we are faced with the issue that due to the curvature encoded in $\boldsymbol{\eta}(\hat{\boldsymbol{y}})$ the operators $\hat{\nabla}$, $\hat{\nabla}_{\hat{\mathbf{t}}}$, $\hat{\nabla}_\eta^{-h}$, $\hat{\nabla}_\eta^h$ do not necessarily commute. Observe, for example, that

$$[\hat{\nabla}, \hat{\nabla}_\eta^{-h}] f(\hat{\boldsymbol{y}}) \;=\; \hat{\nabla} f\big(T_{-h}(\hat{\boldsymbol{y}})\big) \hat{\nabla} \boldsymbol{\eta}(\hat{\boldsymbol{y}}) \kappa(\hat{\boldsymbol{y}}) + f\big(T_{-h}(\hat{\boldsymbol{y}})\big) \frac{\hat{\nabla}\kappa(\hat{\boldsymbol{y}})}{h}, \tag{B.4}$$

where $[.,.]$ denotes the commutator. Lemma B.2(2.) establishes that $|\kappa(\hat{\boldsymbol{y}})/h|_\infty \leq c \max\big(|\nabla^2 \boldsymbol{\eta}(\hat{\boldsymbol{y}})|_\infty, |\nabla^2 \boldsymbol{\eta}(\hat{\boldsymbol{y}})|_\infty\big)$. Applying this result to (B.4) and taking the limit $h \to 0$ on the right hand side yields

$$\big\| [\hat{\nabla}, \hat{\nabla}_\eta^{-h}] f(\hat{\boldsymbol{y}}) \big\|_{L^2(\hat{Y})} \leq C \max\big\{1, \|\hat{\nabla}\boldsymbol{\eta}\|_{L^\infty(\hat{Y})}, \|\hat{\nabla}^2 \boldsymbol{\eta}\|_{L^\infty(\hat{Y})}\big\} \big\| \hat{\nabla} f(\hat{\boldsymbol{y}}) \|_{L^2(\hat{Y})}, \tag{B.5}$$

where we have used a Poincaré inequality valid for periodic $f(\hat{\boldsymbol{y}})$. A similar result holds for all other commutator pairings. Then, moving the difference operator and applying the integration by parts formula and product rule formula established in Lemma B.2 we arrive at

$$\begin{aligned}
&\int_{\hat{Y}} \operatorname{Re} \hat{\varepsilon}(\boldsymbol{x}, T_h(\hat{\boldsymbol{y}}))\, \hat{\nabla}\hat{\nabla}_\eta^h \hat{\boldsymbol{\chi}}^T \cdot \hat{\nabla}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T + (\hat{\nabla}_\eta^h \operatorname{Re}\hat{\varepsilon})\, \hat{\nabla}\hat{\boldsymbol{\chi}}^T \cdot \hat{\nabla}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{y} \\
&\quad + \frac{1}{\omega} \int_{\hat{\Sigma}} \operatorname{Im} \hat{\sigma}(\boldsymbol{x}, T_h(\hat{\boldsymbol{y}})) \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\nabla}_\eta^h \hat{\boldsymbol{\chi}}^T \cdot \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T + (\hat{\nabla}_\eta^h \operatorname{Im}\hat{\sigma}) \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}^T \cdot \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{o}_{\hat{y}} \\
&\qquad = -\operatorname{Re}\Big\{ \int_{\hat{Y}} \hat{\nabla}_\eta^h \big\{ \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}}) (\hat{F}(\hat{\boldsymbol{y}})^T) \big\} \cdot \hat{\nabla}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{y} \\
&\qquad\quad + \frac{1}{i\omega} \int_{\hat{\Sigma}} \hat{\nabla}_\eta^h \big\{ \hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}}) (\hat{F}(\hat{\boldsymbol{y}})^T)_{\hat{\mathbf{t}}} \big\} \cdot \hat{\nabla}_{\hat{\mathbf{t}}} \hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{o}_{\hat{y}} \Big\} \\
&\qquad\qquad + \big\{ \text{commutator terms} \big\},
\end{aligned} \tag{B.6}$$

where we have collected all commutator terms in the last term (and discuss them further down below). Proceeding again as in the proof for Theorem 2.10 by using Young's inequality for all terms on the right hand side and uniform ellipticity of $\operatorname{Re}\hat{\varepsilon}$ and $\operatorname{Im}\hat{\sigma}$ yields:

$$\begin{aligned}
&\|\hat{\nabla}\hat{\nabla}_\eta^h \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega} \|\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\nabla}_\eta^h \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})} \\
&\quad \leq C \Big\{ \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})} \|\hat{\nabla}\hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{Y})} + \frac{1}{\omega} \|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})} \|\hat{\nabla}_{\hat{\mathbf{t}}} \hat{\boldsymbol{\chi}}\|^2_{L^2(\hat{\Sigma})} \\
&\qquad + \|\hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{Y})} \|\hat{F}^T\|^2_{H^1(\hat{Y})} + \frac{1}{\omega} \|\hat{\sigma}(\boldsymbol{x}, \hat{\boldsymbol{y}})\|^2_{W^{1,\infty}(\hat{\Sigma})} \|\hat{F}^T\|^2_{H^1(\hat{\Sigma})} \\
&\qquad\qquad + \big| \text{commutator terms} \big| \Big\}
\end{aligned} \tag{B.7}$$

In the estimate above we have passed to the limit, $h \to 0$, on the right hand side.

As a last step we will discuss the commutator terms. The volume integral over $\hat{Y}$, for example, gives rise to the following terms:

$$\begin{aligned}
\text{(a)} \;=\; &\int_{\hat{Y}} \operatorname{Re} \hat{\varepsilon}\, \hat{\nabla}\hat{\boldsymbol{\chi}}^T \cdot [\hat{\nabla}, \hat{\nabla}_\eta^{-h}] \hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{y} \\
&+ \int_{\hat{Y}} \operatorname{Re} \hat{\varepsilon}(T_h(\hat{\boldsymbol{y}}))\, [\hat{\nabla}, \hat{\nabla}_\eta^h] \hat{\boldsymbol{\chi}}^T \cdot \hat{\nabla}\hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{y} \\
&+ \operatorname{Re}\Big\{ \int_{\hat{Y}} \hat{\varepsilon}(\boldsymbol{x}, \hat{\boldsymbol{y}}) (\hat{F}(\hat{\boldsymbol{y}})^T) \cdot [\hat{\nabla}, \hat{\nabla}_\eta^{-h}] \hat{\nabla}_\eta^h \overline{\hat{\boldsymbol{\chi}}}^T \,\mathrm{d}\hat{y} \Big\}.
\end{aligned}$$

Applying (B.5) allows us to estimate (a) by

$$\big|(a)\big| \;\le\; C\,\|\hat{\varepsilon}(\boldsymbol{x},\hat{\boldsymbol{y}})\|_{L^\infty(\hat{Y})}\max\big\{1,\|\hat{\nabla}\boldsymbol{\eta}\|_{L^\infty(\hat{Y})},\|\hat{\nabla}^2\boldsymbol{\eta}\|_{L^\infty(\hat{Y})}\big\} \\ \|\hat{\nabla}\hat{\nabla}_\eta^h\hat{\boldsymbol{\chi}}\|_{L^2(\hat{Y})}\Big\{\|\hat{\nabla}\hat{\boldsymbol{\chi}}\|_{L^2(\hat{Y})}+\|\hat{F}^T\|_{H^1(\hat{Y})}\Big\}.$$

This expression can again be absorbed into the remaining terms of (B.7) by changing the constant $C$ to $C\max\{1,\|\hat{\nabla}\boldsymbol{\eta}\|_{L^\infty(\hat{Y})},\|\hat{\nabla}^2\boldsymbol{\eta}\|_{L^\infty(\hat{Y})}\}$. An analogous result holds for all terms arising from the surface integral over $\hat{\Sigma}$. Passing to the limit $h\to 0$ on the left side [12] now shows the final estimate given in (B.2). ∎

*Proof of Theorem 2.13.* Estimate (2.19) is an immediate consequence of Proposition B.3 by setting $\boldsymbol{\eta}=\hat{\boldsymbol{\tau}}_i$ and using Theorem 2.10 to estimate the $\|\hat{\nabla}\hat{\boldsymbol{\chi}}\|_{L^2(\hat{Y})}$ and $\|\hat{\nabla}_{\hat{\mathbf{t}}}\hat{\boldsymbol{\chi}}\|_{L^2(\hat{\Sigma})}$ terms on the right hand side of (B.2). ∎

**References.**


[1] A. Alù and N. Engheta, *Achieving transparency with plasmonic and metamaterial coatings*, Physical Review E, 72 (2005), p. 016623.

[2] Y. Amirat and V. V. Shelukhin, *Homogenization of time harmonic Maxwell equations: the effect of interfacial currents*, Mathematical Methods in the Applied Sciences, 40 (2017), pp. 3140–3162.

[3] D. Arndt, W. Bangerth, D. Davydov, T. Heister, L. Heltai, M. Kronbichler, M. Maier, J.-P. Pelteret, B. Turcksin, and D. Wells, *The deal.II finite element library: design, features, and insights*, Computers & Mathematics with Applications, 81 (2021), pp. 407–422.

[4] D. Arndt, W. Bangerth, M. Feder, M. Fehling, R. Gassmöller, T. Heister, L. Heltai, M. Kronbichler, M. Maier, P. Munch, J.-P. Pelteret, S. Sticko, B. Turcksin, and D. Wells, *The deal.II Library, Version 9.4*, journal of Numerical Mathematics, (2022).

[5] R. Becker, D. Meidner, and B. Vexler, *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optimization Methods and Software, 22 (2007), pp. 813–833.

[6] M. Bezbaruah, M. Maier, and W. Wollner, *InverseHomogenization: shape optimization solver for optical metamaterials with low-dimensional interfaces*, 2024, https://doi.org/10.5281/zenodo.10459309, https://zenodo.org/records/10459309.

[7] V. I. Bogachev and E. Mayer-Wolf, *Some remarks on Rademacher's theorem in infinite dimensions*, Potential Analysis, 5 (1996), pp. 23–30.

[8] C. Brandenburg, F. Lindemann, M. Ulbrich, and S. Ulbrich, *A continuous adjoint approach to shape optimization for navier stokes flow*, in Optimal control of coupled systems of partial differential equations, Springer, 2009, pp. 35–56.

[9] J. Cheng, W. L. Wang, H. Mosallaei, and E. Kaxiras, *Surface plasmon engineering in graphene functionalized with organic molecules: A multiscale theoretical investigation*, Nano Letters, 14 (2014), pp. 50–56.

[10] R. Fletcher, *Practical Methods of Optimization*, Wiley, 2 ed., 1988.

[11] S. Frei, H. Andrä, R. Pinnau, and O. Tse, *Optimizing fiber orientation in fiber-reinforced materials using efficient upscaling*, Computational Optimization and Applications, 62 (2015), pp. 111–129.

[12] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, vol. 224 of Classics in Mathematics, Springer, 2 ed., 2001.

[13] C. Goll, T. Wick, and W. Wollner, *DOpElib: Differential equations and optimization environment; a goal oriented software library for solving pdes and optimization problems with pdes*, Archive of Numerical Software, 5 (2017).
[14] A. Griewank, *The local convergence of Broyden-like methods on Lipschitzian problems in Hilbert spaces*, SIAM Journal of Numerical Analysis, 24 (1987), pp. 684–705.
[15] A. N. Grigorenko, M. Polini, and K. S. Novoselov, *Graphene plasmonics*, Nature Photonics, 6 (2012), pp. 749–758.
[16] E. Hassan and A. C. Lesina, *Time-domain topology optimization of wideband dispersive plasmonic nanostructures*, in 2022 Photonics North (PN), 2022, pp. 1–1.
[17] J. Haubner, M. Ulbrich, and S. Ulbrich, *Analysis of shape optimization problems for unsteady fluid-structure interaction*, Inverse Problems, 36 (2020), p. 034001.
[18] C. Herter, S. Schöps, and W. Wollner, *Eigenvalue optimization with respect to shape-variations in electromagnetic cavities*, PAMM, 22 (2022).
[19] C. Herter and W. Wollner, *Numerical eigenvalue optimization by shape-variations for electromagnetic cavities.* in preparation, 2023.
[20] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, vol. 23 of Mathematical Modelling: Theory and Applications, Springer, 2009.
[21] X. Huang, Y. Lai, Z. H. Hang, H. Zheng, and C. T. Chan, *Dirac cones induced by accidental degeneracy in photonic crystals and zero-refractive-index materials*, Nature Materials, 10 (2011), pp. 582–586.
[22] C. T. Kelley and E. W. Sachs, *A new proof of superlinear convergence for Broyden's method in Hilbert space*, SIAM Journal on Optimization, 1 (1991), pp. 146–150.
[23] W. Li, R. Lipton, and M. Maier, *Lorentz resonance in the homogenization of plasmonic crystals*, Proceedings of the Royal Society A: Mathematical, Physical, and Engineering Sciences, 477 (2021), p. 20210609.
[24] I. Liberal and N. Engheta, *Near-zero refractive index photonics*, Nature Photonics, 11 (2017), pp. 149–158.
[25] M. Maier, D. Margetis, and M. Luskin, *Finite-size effects in wave transmission through plasmonic crystals: A tale of two scales*, Physical Review B, 102 (2020), p. 075308.
[26] M. Maier, D. Margetis, and A. Mellet, *Homogenization of Maxwell's equations in nonhomogeneous plasmonic structures*, Journal of Computational and Applied Mathematics, 377 (2020).
[27] M. Maier, M. Mattheakis, E. Kaxiras, M. Luskin, and D. Margetis, *Homogenization of plasmonic crystals: Seeking the epsilon-near-zero effect*, Proceedings of the Royal Society A: Mathematical, Physical, and Engineering Sciences, 475 (2019).
[28] M. Mattheakis, C. A. Valagiannopoulos, and E. Kaxiras, *Epsilon-near-zero behavior from plasmonic dirac point: Theory and realization using two-dimensional materials*, Physical Review B, 94 (2016), p. 201404(R).
[29] N. Mohammadi Estakhri, C. Argyropoulos, and A. Alù, *Graded metascreens to enable a new degree of nanoscale light management*, Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 373 (2015), p. 20140351.
[30] P. Moitra, Y. Yang, Z. Anderson, I. I. Kravchenko, D. P. Briggs, and J. Valentine, *Realization of an all-dielectric zero-index optical metamaterial*,

Nature Photonics, 7 (2013), pp. 791–795.
[31] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, *Inverse design in nanophotonics*, Nature Photon, 12 (2018), pp. 659–670.
[32] P. Monk, *Finite Element Methods for Maxwell's Equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, 2003.
[33] A. Nemilentsau, T. Low, and G. Hanson, *Anisotropic 2D materials for tunable hyperbolic plasmonics*, Physical Review Letters, 116 (2016), p. 066804.
[34] V. Nikkhah, M. J. Mencagli, and N. Engheta, *Reconfigurable nonlinear optical element using tunable couplers and inverse-designed structure*, Nanophotonics, (2023).
[35] A. N. Papadimopoulos, A. Duspayev, N. L. Tsitsas, N. V. Kantartzis, and C. Valagiannopoulos, *Wavefront engineering with optimally loaded absorbing metamirrors*, Phys. Rev. B, 103 (2021), p. 165307.
[36] M. J. D. Powell, *A fast algorithm for nonlinearly constrained optimization calculations*, in Numerical analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977), vol. 630 of Lecture Notes in Math., 1978.
[37] T. Richter, *Fluid-structure interactions: models, analysis and finite elements*, vol. 118, Springer, 2017.
[38] W. Ring and B. Wirth, *Optimization methods on riemannian manifolds and their application to shape space*, SIAM J. Optim., 22 (2012), pp. 596–627.
[39] V. H. Schulz, M. Siebenborn, and K. Welker, *Efficient pde constrained shape optimization based on steklov–poincaré-type metrics*, SIAM J. Optim., 26 (2016), pp. 2800–2819.
[40] A. Silva, F. Monticone, G. Castaldi, V. Galdi, A. Alù, and N. Engheta, *Performing mathematical operations with metamaterials*, Science, 343 (2014), pp. 160–163.
[41] M. Silveirinha and N. Engheta, *Tunneling of electromagnetic energy through subwavelength channels and bends using epsilon-near-zero materials*, Phus. Rev. Lett., 97 (2006), p. 157403.
[42] T. Tan, X. Jiang, C. Wang, B. Yao, and H. Zhang, *2D material optoelectronics for information functional device applications: Status and challenges*, Advanced Science, 7 (2020), p. 2000058.
[43] F. Tröltzsch, *Optimal control of partial differential equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010.
[44] K. Welker, *Efficient PDE Constrained Shape Optimization in Shape Spaces*, PhD thesis, 2017.
[45] N. Wellander, *Homogenization of the Maxwell equations: Case i. linear theory*, Applications of Mathematics, 46 (2001), pp. 29–51.
[46] N. Wellander, *Homogenization of the Maxwell equations: Case ii. nonlinear conductivity*, Applications of Mathematics, 47 (2002), pp. 255–283.
[47] N. Wellander and G. Kristensson, *Homogenization of the Maxwell equations at fixed frequency*, SIAM Journal on Applied Mathematics, 64 (2003), pp. 170–195.
[48] H. Yang, Y. Wang, Z. C. Tiu, S. J. Tan, L. Yuan, and H. Zhang, *All-optical modulation technology based on 2D layered materials*, Micromachines, 13 (2022), p. 92.