

Adjoint methods for electromagnetic shape optimization of the low-loss cavity for the International Linear Collider

Volkan Akçelik¹, George Biros², Omar Ghattas³, David Keyes⁴,
Kwok Ko⁵, Lie-Quan Lee⁵ and Esmond G. Ng⁶

¹ Ultrascale Simulation Lab, Carnegie Mellon University, Pittsburgh, PA, USA

² Institute for Computational Engineering and Sciences, and Departments of Geological Sciences, Mechanical Engineering, Computer Science, and Biomedical Engineering, University of Texas, Austin, TX, USA

³ Departments of Mechanical Engineering & Applied Mechanics and Computer & Information Science, University of Pennsylvania, Philadelphia, PA, USA

⁴ Applied Physics & Applied Mathematics, Columbia University, New York, NY, USA

⁵ Stanford Linear Accelerator Center, Menlo Park, CA, USA

⁶ Lawrence Berkeley National Laboratory, Berkeley, CA, USA

E-mail: volkan@cmu.edu, oghattas@cs.cmu.edu, david.keyes@columbia.edu,
kwok@slac.stanford.edu, liequan@slac.stanford.edu, EGNg@lbl.gov

Abstract. We formulate the problem of designing the low-loss cavity for the International Linear Collider (ILC) as an electromagnetic shape optimization problem involving a Maxwell eigenvalue problem. The objective is to maximize the stored energy of a trapped mode in the end cell while maintaining a specified frequency corresponding to the accelerating mode. A continuous adjoint method is presented for computation of the design gradient of the objective and constraint. The gradients are used within a nonlinear optimization scheme to compute the optimal shape for a simplified model of the ILC in a small multiple of the cost of solving the Maxwell eigenvalue problem.

1. Introduction

The International Linear Collider (ILC) is a proposed international particle accelerator for high energy physics research. The ILC will create high-energy particle collisions between electrons and positrons, their antimatter counterparts. It will allow scientists to address fundamental questions about dark matter, dark energy, extra dimensions and the nature of matter, energy, space and time. The multi-billion dollar facility is to be designed, funded, managed and operated as an international project involving physicists and engineers from Asia, Europe, and the Americas.

A critical issue in the design of the ILC is the presence of high-order dipole modes (HOMs) that are excited by the transiting beam in the accelerator cavity [1]. HOMs can interact strongly with the beam and lead to a deterioration of beam quality. One of the design goals is to optimize the shape of accelerator cavity to damp these HOMs as much as possible while maintaining a specified cavity accelerating frequency. This gives rise to a shape optimization

problem governed by a large-scale electromagnetic eigenvalue problem. A prototype ILC shape optimization problem will be formulated in the next section.

The large design space, considerable cost of solving the electromagnetic eigenvalue problem, and high accuracy requirements of the frequency constraint dictate the use of gradient-based methods to solve the optimization problem. These methods require the computation of the derivative of the objective function and constraints with respect to the design variables (typically CAD parameters) that describe the accelerator cavity shape. Since the shape is parameterized by as many as hundreds of design variables, a standard differencing approach to computing these derivatives would require several hundred electromagnetic eigenvalue solves per design iteration — an intractable proposition. Moreover, due to the iterative nature of the eigenvalue problem and the stringent accuracy requirements for eigenvalues, accurate numerical gradients would be difficult to obtain. On the other hand, while automatic differentiation (AD) methods have enjoyed many successes, the need to differentiate through complex, coupled meshing, CAD, eigenvalue solver, and adaptivity routines precludes their use here. Alternatively, whereas computing analytical derivatives via direct sensitivity methods avoids the numerical differencing problems, intractability associated with the large number of eigenvalue solves remains — mitigated only by the need now to solve just *linearized* eigenvalue problems. Furthermore, differentiating the mesh movement scheme (with respect to design variables) is still required.

Ultimately, the best approach to computing gradients for such large, complex eigenvalue shape optimization problems is provided by continuous adjoint methods. Like direct sensitivity methods, they avoid numerical differencing difficulties, but unlike direct methods, they scale only with the number of active constraints (which are usually many fewer than the number of design variables). Moreover, continuous adjoint methods avoid differentiating through the mesher for most functions of interest. This is significant, since unstructured mesh movement schemes usually incorporate discrete decisions and are thus non-differentiable.

Despite their compelling advantages over other gradient computation methods, a number of challenges must be addressed before continuous adjoints can be used for accelerator cavity shape optimization.

- (i) Although differentiating the *volume* mesh motion is avoided, computing the gradient still requires the derivative of the *surface* mesh with respect to the design variables, the so-called *design velocity field*.
- (ii) This in turn requires differentiating the CAD surface representation with respect to the design variables.
- (iii) Computation of the gradient of the objective and constraints requires solution of an additional system of partial differential equations — an adjoint eigenvalue problem — for each function(al) of interest, which has somewhat different structure from the original Maxwell eigenvalue problem and may require different discretizations, solvers, and preconditioners.
- (iv) The gradient expressions require evaluation of surface integrals that involve higher derivatives of the magnetic field and adjoint magnetic field eigenvectors. These integrals must be computed with care to avoid excessive loss of accuracy.
- (v) Since the problems we consider involve millions of variables, all components described above must run in parallel and scale to large numbers of processors.

The challenges described above are manifold: they involve such areas as accelerator physics; computational electromagnetics; high-order finite element discretizations; large scale eigensolvers, linear solvers, and preconditioners; mesh generation, movement, and adaptivity; solid modeling; continuous adjoint methods; and large-scale optimization, among others. Clearly, these challenges cannot be addressed from the vantage of any one discipline. To this end, and under the auspices of the Department of Energy's Scientific Discovery through Advanced

Computing (SciDAC) program, a collaboration among three SciDAC centers — the centers for Accelerator Science and Technology (AST), Terascale Optimal PDE Simulations (TOPS), and Terascale Simulation Tools and Technology (TSTT) — has been formed to tackle the grand challenge of accelerator design optimization. This collaboration merges expertise in parallel finite element electromagnetics codes [1]; parallel mesh generation [2], smoothing [3], and adaptivity [4]; advanced eigensolvers [5, 6]; and large-scale PDE-constrained optimization [7, 8].

In this article, we focus on the third and fourth issues itemized above, i.e. those related to adjoint-based gradient computation. Subsequent articles will describe all of the components of the end-to-end system. Section 2 presents a formulation of a design problem for the low-loss cavity of the ILC aimed at detuning the trapped mode while maintaining the accelerating frequency. Computation of the design gradient of the objective is discussed in Section 3, and that of the constraint in Section 4. Section 5 concludes with solution of a low-loss cavity shape optimization problem for an ILC-like geometry. The example demonstrates that through the use of continuous adjoints, the optimum can be computed at the cost of just a handful of electromagnetic eigenvalue solves.

2. A shape optimization problem for the ILC low-loss cavity

Figure 1 shows a model of the Low-Loss (LL) cavity for the ILC. It consists of 7 identical interior cells and 2 possibly different end cells connected to 3 coax couplers. Figure 2 shows the field pattern of a well-trapped dipole mode, which is localized to the middle 7 cells. A design problem can be formulated to optimize the shape of the end cells to damp the trapped modes (we regard the geometry of the interior cells as fixed). This can be approached by finding the shape of the end cells that maximizes the stored energy of the trapped modes in these cells, where the couplers will be able to leak greater energy. For simplicity of presentation, we shall formulate the optimization problem for a single trapped mode, which we seek to de-tune by optimizing its shape (the procedure is similar whenever multiple trapped modes are present). At the same time, the shape must be constrained so that it generates a desired accelerating frequency; in the case of the LL cavity model, this frequency is specified as 1.3 GHz, and the accelerating mode corresponds to the 9th mode. Shape optimization will be effected via modification of design variables describing the shape of an end cell, for example those displayed in Figure 3.

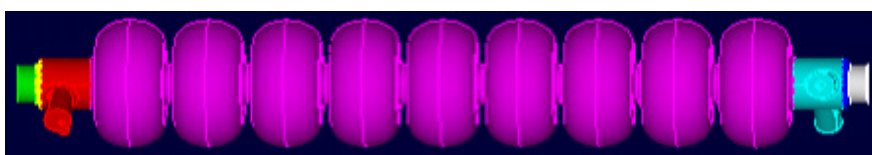


Figure 1. Low-Loss Cavity model.

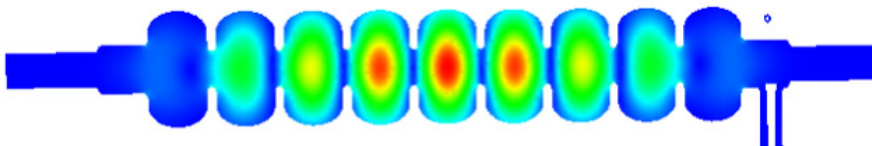


Figure 2. Field pattern of a well-trapped dipole mode.

The shape optimization problem for this design problem can be formulated using the following notation. Let $\mathbf{q} = \{q_1, q_2, \dots, q_n\}$ denote the vector of design variables, for example the

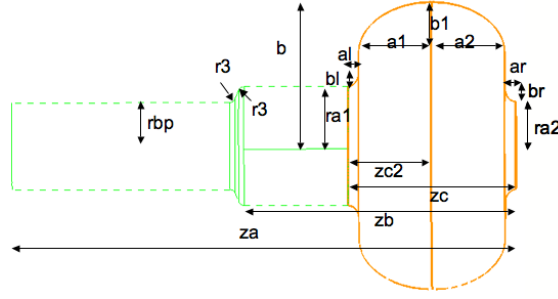


Figure 3. The dimensions of the end cell and its connected beam pipes.

variables shown in Figure 3. Let $\mathbf{H}_t$ and λ_t represent the magnetic field eigenvector and eigenvalue, respectively, corresponding to the trapped mode. The eigenvalue is given in terms of the frequency ω by $\lambda = \frac{\omega^2}{c^2}$, where c is the speed of light. We denote by $\mathbf{H}_a$ and λ_a the eigenvector and eigenvalue of the accelerating mode. Let λ^* represent the target eigenvalue (which here corresponds to a frequency of 1.3 GHz) for the accelerating mode. The constants ε and μ represent the relative permittivity and relative permeability. Finally, let Ω represent the entire cavity, Ω_e the end cell cavity, Γ_M the portion of the boundary on which magnetic field boundary conditions are prescribed, Γ_E the portion of the boundary on which electric field boundary conditions are prescribed, and $\mathbf{n}$ the outward unit normal to a surface. We assume that $\Gamma_E \cup \Gamma_M = \partial\Omega$, the entire boundary of Ω .

The eigenpairs $(\mathbf{H}_t, \lambda_t)$ and $(\mathbf{H}_a, \lambda_a)$ satisfy the the magnetic form of the Maxwell eigenvalue problem for each mode, i.e. for $i \in \{t, a\}$,

$$\begin{aligned}
 \nabla \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{H}_i \right) - \lambda \mu \mathbf{H}_i &= \mathbf{0} \text{ in } \Omega(\mathbf{q}), \\
 \nabla \cdot \mathbf{H}_i &= 0 \text{ in } \Omega(\mathbf{q}), \\
 \mathbf{n} \times \mathbf{H}_i &= \mathbf{0} \text{ on } \Gamma_M(\mathbf{q}), \\
 \mathbf{n} \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{H}_i \right) &= \mathbf{0} \text{ on } \Gamma_E(\mathbf{q}), \\
 \frac{1}{2} \int_{\Omega(\mathbf{q})} \mu \mathbf{H}_i \cdot \mathbf{H}_i \, d\mathbf{x} &= \frac{1}{2},
 \end{aligned} \tag{1}$$

where the dependence of the domain and boundary on the design variables is explicitly noted. We refer to the Maxwell eigenvalue problem as the *state eigenvalue problem* and the eigenpairs $(\mathbf{H}_i, \lambda_i), i \in \{t, a\}$ as the *state eigenvectors/eigenvalues* to distinguish them from their *adjoint* counterparts, to be introduced shortly. Note that we must explicitly write the orthonormalization condition on the magnetic eigenvector $\mathbf{H}_i$, which is enforced through the last equation in (1). Otherwise, it would not be revealed to the optimization problem. The factor of one-half on each side of the constraint is needed to produce a symmetric adjoint operator, as will be seen below.

The design problem is to maximize the stored energy of the trapped modes in an end cell, or equivalently, minimize the negative of the stored energy

$$\mathcal{J}(\mathbf{q}) \stackrel{\text{def}}{=} - \int_{\Omega_e(\mathbf{q})} \mu \mathbf{H}_t(\mathbf{q}) \cdot \mathbf{H}_t(\mathbf{q}) \, d\mathbf{x}, \tag{2}$$

subject to a constraint that the cavity shape generates a desired accelerating frequency, i.e.

$$\mathcal{C}(\mathbf{q}) \stackrel{\text{def}}{=} \lambda_a(\mathbf{q}) - \lambda^* = 0. \tag{3}$$

The stored energy objective $\mathcal{J}$ depends on the design variables $\mathbf{q}$ explicitly through the dependence of the domain integral on $\mathbf{q}$ and implicitly through the dependence of the eigenvector $\mathbf{H}_t$ on $\mathbf{q}$ via solution of the eigenvalue problem (1) for the trapped mode. On the other hand, the accelerating frequency constraint residual $\mathcal{C}$ depends on the design variables $\mathbf{q}$ only implicitly, through solution of the eigenvalue problem (1) for the accelerating mode. In summary, the shape optimization problem is to minimize (2) subject to (3), with the relation between the state eigenpairs and the design variables given implicitly by (1).

In the next two sections, we derive expressions for the gradient of the trapped energy objective function $\mathcal{J}(\mathbf{q})$ and the accelerating frequency constraint $\mathcal{C}(\mathbf{q})$ with respect to the design variable vector $\mathbf{q}$. Highly-efficient optimization methods such as Sequential Quadratic Programming (SQP) require such gradient information to solve the optimization problem.

3. Design gradient of the trapped mode stored energy objective

To derive an expression for the gradient of the objective $\mathcal{J}$ with respect to the design variables, we follow classical optimal control theory [9], and in particular its elaboration for shape optimization problems [10]. The dependence of the solution of the state eigenvalue problem (1) on the design variables $\mathbf{q}$ can be enforced by constructing a Lagrangian functional $\mathcal{L}_{\mathcal{J}}$ that augments $\mathcal{J}$ with inner products (formally, duality pairings) of so-called *adjoint variables* $\mathbf{T}_{\mathcal{J}}, \mathbf{\Lambda}_{M_t}, \mathbf{\Lambda}_{E_t}, \xi_{\mathcal{J}}$ (also known as Lagrange multipliers) with residuals of the the state eigenvalue equations corresponding to the trapped-mode eigenpair $(\mathbf{H}_t, \lambda_t)$:

$$\begin{aligned} \mathcal{L}_{\mathcal{J}}(\mathbf{H}_t, \lambda_t, \mathbf{q}, \mathbf{T}_{\mathcal{J}}, \mathbf{\Lambda}_{M_t}, \mathbf{\Lambda}_{E_t}, \xi_{\mathcal{J}}) = & -\frac{1}{2} \int_{\Omega_e} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} + \int_{\Omega} \mathbf{T}_{\mathcal{J}} \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t - \lambda_t \mu \mathbf{H}_t \right) d\mathbf{x} \\ & + \int_{\Gamma_M} \mathbf{\Lambda}_{M_t} \cdot (\mathbf{n} \times \mathbf{H}_t) \, d\mathbf{s} + \int_{\Gamma_E} \mathbf{\Lambda}_{E_t} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) d\mathbf{s} + \frac{\xi_{\mathcal{J}}}{2} \left(\int_{\Omega} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} - 1 \right), \end{aligned} \quad (4)$$

in which the dependence of Ω and Γ on $\mathbf{q}$ has been suppressed for simplicity of presentation, and the dependence of the state eigenpair $(\mathbf{H}_t, \lambda_t)$ on $\mathbf{q}$ has been removed since the state eigenvalue problem (1) is now enforced in the Lagrangian $\mathcal{L}_{\mathcal{J}}$ explicitly via Lagrange multipliers. The first term in (4) corresponds to the end-cell trapped energy, the second term corresponds to the domain equation of the state eigenvalue problem (1), the third term corresponds to the magnetic boundary condition, the fourth term corresponds to the electric boundary condition, and the fifth term corresponds to the orthonormalization constraint on $\mathbf{H}_t$. Accordingly, $\mathbf{T}_{\mathcal{J}}$ represents the adjoint magnetic eigenvector in Ω , $\mathbf{\Lambda}_{M_t}$ the adjoint magnetic eigenvector on Γ_{M_t} , $\mathbf{\Lambda}_{E_t}$ the adjoint magnetic eigenvector on Γ_E , and $\xi_{\mathcal{J}}$ the adjoint eigenvalue. Note that there is no need to enforce the gauge condition on $\mathbf{H}_t$; the modes of interest in the optimization problem correspond to nonzero eigenvalues, and thus are automatically divergence-free. The use of Nédélec edge elements guarantees that this property is inherited by the discretized problem [11]. In fact, this is not the full story; we must guarantee that the adjoint eigenvector is also divergence-free, and this remains to be shown below. We next derive the adjoint eigenvalue problem whose solution yields the adjoint variables $\mathbf{T}_{\mathcal{J}}, \mathbf{\Lambda}_{M_t}, \mathbf{\Lambda}_{E_t}, \xi_{\mathcal{J}}$, and then proceed to derive expressions for the design gradient of $\mathcal{J}$ that can be evaluated once the state and adjoint eigenvalue problems have been solved.

3.1. Adjoint eigenvalue problem

First define the space of three-dimensional vector functions with curl in L^2 by

$$H(\text{curl}; \Omega) = \{\mathbf{v} \in (L^2(\Omega))^3 \mid \nabla \times \mathbf{v} \in (L^2(\Omega))^3\}$$

To obtain the adjoint system, we take the variation of the Lagrangian (4) with respect to $\mathbf{H}_t$, and require that it vanish for all admissible $\hat{\mathbf{H}}_t$:

$$\begin{aligned} & - \int_{\Omega_e} \mu \mathbf{H}_t \cdot \hat{\mathbf{H}}_t \, d\mathbf{x} + \int_{\Omega} \mathbf{T}_{\mathcal{J}} \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t - \lambda_t \mu \hat{\mathbf{H}}_t \right) d\mathbf{x} + \int_{\Gamma_M} \boldsymbol{\Lambda}_{M_t} \cdot (\mathbf{n} \times \hat{\mathbf{H}}_t) \, d\mathbf{s} \\ & + \int_{\Gamma_E} \boldsymbol{\Lambda}_{E_t} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t \right) d\mathbf{s} + \xi_{\mathcal{J}} \int_{\Omega} \mu \mathbf{H}_t \cdot \hat{\mathbf{H}}_t \, d\mathbf{x} = 0 \quad \forall \hat{\mathbf{H}}_t \in H(\text{curl}; \Omega) \end{aligned}$$

Integration by parts twice with the integral identity

$$\int_{\Omega} [(\nabla \times \mathbf{a}) \cdot (\nabla \times \mathbf{b}) - \mathbf{a} \cdot (\nabla \times \nabla \times \mathbf{b})] \, d\mathbf{x} = \int_{\Gamma} (\mathbf{a} \times (\nabla \times \mathbf{b})) \cdot \mathbf{n} \, d\mathbf{s} \quad (5)$$

yields

$$\begin{aligned} & - \int_{\Omega_e} \mu \mathbf{H}_t \cdot \hat{\mathbf{H}}_t \, d\mathbf{x} + \int_{\Omega} \hat{\mathbf{H}}_t \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} - \lambda_t \mu \mathbf{T}_{\mathcal{J}} \right) d\mathbf{x} \\ & + \int_{\Gamma} \frac{1}{\varepsilon} (\hat{\mathbf{H}}_t \times \nabla \times \mathbf{T}_{\mathcal{J}} - \mathbf{T}_{\mathcal{J}} \times \nabla \times \hat{\mathbf{H}}_t) \cdot \mathbf{n} \, d\mathbf{s} + \int_{\Gamma_M} \boldsymbol{\Lambda}_{M_t} \cdot (\mathbf{n} \times \hat{\mathbf{H}}_t) \, d\mathbf{s} \\ & + \int_{\Gamma_E} \boldsymbol{\Lambda}_{E_t} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t \right) d\mathbf{s} + \xi_{\mathcal{J}} \int_{\Omega} \mu \mathbf{H}_t \cdot \hat{\mathbf{H}}_t \, d\mathbf{x} = 0 \quad \forall \hat{\mathbf{H}}_t \in H(\text{curl}; \Omega). \end{aligned}$$

Since $\hat{\mathbf{H}}_t$ is arbitrary, we set it to zero on $\partial\Omega$ to obtain

$$\nabla \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} \right) - \lambda_t \mu \mathbf{T}_{\mathcal{J}} + \xi_{\mathcal{J}} \mu \mathbf{H}_t - \chi_e \mu \mathbf{H}_t = \mathbf{0} \quad \text{in } \Omega$$

where the Heaviside function χ_e is equal to 1 inside Ω_e and zero otherwise. We then choose $\hat{\mathbf{H}}_t$ such that it vanishes in Ω , leaving

$$\begin{aligned} & \int_{\Gamma} \frac{1}{\varepsilon} (\hat{\mathbf{H}}_t \times \nabla \times \mathbf{T}_{\mathcal{J}} - \mathbf{T}_{\mathcal{J}} \times \nabla \times \hat{\mathbf{H}}_t) \cdot \mathbf{n} \, d\mathbf{s} + \int_{\Gamma_M} \boldsymbol{\Lambda}_{M_t} \cdot (\mathbf{n} \times \hat{\mathbf{H}}_t) \, d\mathbf{s} \\ & + \int_{\Gamma_E} \boldsymbol{\Lambda}_{E_t} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t \right) d\mathbf{s} = 0 \quad \forall \hat{\mathbf{H}}_t \in H(\text{curl}; \Omega). \end{aligned}$$

Next, setting $\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t = \mathbf{0}$, we obtain

$$\int_{\Gamma} \frac{1}{\varepsilon} \hat{\mathbf{H}}_t \times (\nabla \times \mathbf{T}_{\mathcal{J}}) \cdot \mathbf{n} \, d\mathbf{s} + \int_{\Gamma_M} \boldsymbol{\Lambda}_{M_t} \cdot (\mathbf{n} \times \hat{\mathbf{H}}_t) \, d\mathbf{s} = 0 \quad \forall \hat{\mathbf{H}}_t \in H(\text{curl}; \Omega),$$

which implies

$$\begin{aligned} \boldsymbol{\Lambda}_{M_t} &= -\frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} \quad \text{on } \Gamma_M \\ \mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} &= \mathbf{0} \quad \text{on } \Gamma_E. \end{aligned}$$

The remaining terms give

$$- \int_{\Gamma} \frac{1}{\varepsilon} \mathbf{T}_{\mathcal{J}} \times (\nabla \times \hat{\mathbf{H}}_t) \cdot \mathbf{n} \, d\mathbf{s} + \int_{\Gamma_E} \boldsymbol{\Lambda}_{E_t} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \hat{\mathbf{H}}_t \right) d\mathbf{s} = 0 \quad \forall \hat{\mathbf{H}}_t \in H(\text{curl}; \Omega),$$

which implies that $\mathbf{\Lambda}_E = -\mathbf{T}_{\mathcal{J}}$ on Γ_E and $\mathbf{n} \times \mathbf{T}_{\mathcal{J}} = \mathbf{0}$ on Γ_M .

Finally, the derivative of the Lagrangian with respect to λ_t yields

$$\int_{\Omega} \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t \, d\mathbf{x} = 0.$$

In summary, the (strong form of the) *adjoint eigenvalue problem for the objective* is given by:

$$\begin{aligned} \nabla \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} \right) - \lambda_t \mu \mathbf{T}_{\mathcal{J}} + \mu \xi_{\mathcal{J}} \mathbf{H}_t &= -\chi_e \mu \mathbf{H}_t \quad \text{in } \Omega \\ \mathbf{n} \times \mathbf{T}_{\mathcal{J}} &= \mathbf{0} \quad \text{on } \Gamma_M \\ \mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} &= \mathbf{0} \quad \text{on } \Gamma_E \\ \int_{\Omega} \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t \, d\mathbf{x} &= 0 \end{aligned} \tag{6}$$

The adjoint eigenvalue problem inherits (homogeneous) electric and magnetic boundary conditions from the state eigenvalue problem. Note that the adjoint eigenvalue problem is a *linear self-adjoint system* in the adjoint eigenvector/eigenvalue pair $(\mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}})$. That is, given a design $\mathbf{q}$, the state eigenvalue problem (1) is solved for the state eigenpair $(\mathbf{H}_t, \lambda_t)$, and the result substituted into the adjoint eigenvalue problem, yielding a linear system for $(\mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}})$. This is in contrast to the state eigenvalue problem, which is nonlinear in $(\mathbf{H}_t, \lambda_t)$. In fact, the operator for the adjoint eigenvalue problem is the adjoint of the linearized state operator (which is in turn self-adjoint). Note also that: enforcement of the orthonormalization condition on $\mathbf{H}$ (the last equation in (1)) has resulted in an orthogonality condition between $\mathbf{T}$ and $\mathbf{H}$; the adjoint eigenvalue $\xi_{\mathcal{J}}$ is a Lagrange multiplier for this orthogonality condition; and the adjoint eigenvalue problem is a saddle point system in $(\mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}})$.

The saddle point system corresponding to the adjoint eigenvalue problem is nonsingular provided the state eigenvalues corresponding to the trapped modes are distinct. This can be seen by recognizing that when λ_t is a distinct eigenvalue, the null space of the (1,1) adjoint operator $(\nabla \times \frac{1}{\varepsilon} \nabla \times - \lambda_t \mu)$ contains a single vector, given by the state eigenvector $\mathbf{H}_t$. The orthogonality condition $\int_{\Omega} \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t \, d\mathbf{x} = 0$ then insures that $\mathbf{T}$ lies in the range space of the (1,1) operator and therefore the saddle point system is nonsingular. When λ_t is a repeated eigenvalue, the null space of the (1,1) operator is of the dimension of the multiplicity of the repeated eigenvalue, and the system is no longer singular. In fact the gradient is no longer defined at a repeated eigenvalue, and the objective is only directionally differentiable, which may cause problems for a smooth optimization method. Fortunately, the frequencies corresponding to trapped modes in accelerator cavities are expected to be well-separated, and discontinuity of the gradient is not an issue for practical problems.

We comment finally on the role of the gauge condition $\nabla \cdot \mathbf{H}_t = 0$. Recall that we did not enforce this condition in the Lagrangian (4), in part because eigenvectors $\mathbf{H}_t$ corresponding to trapped modes (and hence nonzero eigenvalues) are guaranteed to be divergence-free. However, had we included the condition in the Lagrangian, the adjoint eigenvalue problem (6) would have ended up with a similar gauge condition on the adjoint eigenvector, i.e. $\nabla \cdot \mathbf{T}_{\mathcal{J}} = 0$. The adjoint eigenvalue problem would then have involved a Lagrange multiplier — a scalar potential-like field variable — to enforce the gauge condition. In general, we *do* need to enforce the gauge condition in the adjoint eigenvalue problem, even when $\nabla \cdot \mathbf{H}_t = 0$. However, for our particular objective function (2), it turns out that this is not necessary. This can be seen by taking the divergence of the first equation in (6); since $\nabla \cdot \mathbf{H}_t = 0$ and since the right side involves $\mathbf{H}_t$, we arrive at $\lambda_t \nabla \cdot \mathbf{T}_{\mathcal{J}} = 0$, and therefore $\mathbf{T}_{\mathcal{J}}$ is automatically divergence-free for nonzero λ_t . Were a different objective chosen so that the right-hand side of the first equation of (6) was

not divergence-free, then the divergence of the adjoint eigenvector would be nonzero, and the gauge condition would have to be enforced. Again, use of Nédélec elements guarantees that these properties of the continuous problem carry over to the discrete one.

3.2. Design gradient of $\mathcal{J}$

The gradient of the stored energy objective $\mathcal{J}$ with respect to the design variables $\mathbf{q}$ is found by taking partial derivatives of the Lagrangian (4) with respect to the q_i . Substituting into (4) the relations $\mathbf{\Lambda}_{E_t} = -\mathbf{T}_{\mathcal{J}}$ on Γ_E and $\mathbf{\Lambda}_{M_t} = \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}}$ on Γ_M we obtain

$$\begin{aligned} \mathcal{L}_{\mathcal{J}}(\mathbf{H}_t, \lambda_t, \mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}}) = & -\frac{1}{2} \int_{\Omega_e} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} + \int_{\Omega} \mathbf{T}_{\mathcal{J}} \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t - \lambda_t \mu \mathbf{H}_t \right) d\mathbf{x} \\ & - \int_{\Gamma_M} \frac{1}{\varepsilon} (\nabla \times \mathbf{T}_{\mathcal{J}}) \cdot (\mathbf{n} \times \mathbf{H}_t) \, ds - \int_{\Gamma_E} \mathbf{T}_{\mathcal{J}} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) ds + \frac{1}{2} \xi_{\mathcal{J}} \left(\int_{\Omega} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} - 1 \right). \end{aligned}$$

Using the integral identity (5),

$$\begin{aligned} \mathcal{L}_{\mathcal{J}}(\mathbf{H}_t, \lambda_t, \mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}}) = & -\frac{1}{2} \int_{\Omega_e} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} + \int_{\Omega} \left[\frac{1}{\varepsilon} (\nabla \times \mathbf{H}_t) \cdot (\nabla \times \mathbf{H}_t) - \lambda_t \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t \right] d\mathbf{x} \\ & + \int_{\Gamma} \mathbf{T}_{\mathcal{J}} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) ds - \int_{\Gamma_M} \frac{1}{\varepsilon} (\nabla \times \mathbf{T}_{\mathcal{J}}) \cdot (\mathbf{n} \times \mathbf{H}_t) ds - \int_{\Gamma_E} \mathbf{T}_{\mathcal{J}} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) ds \\ & + \frac{1}{2} \xi_{\mathcal{J}} \left(\int_{\Omega} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} - 1 \right), \end{aligned}$$

and simplifying, we obtain

$$\begin{aligned} \mathcal{L}_{\mathcal{J}}(\mathbf{H}_t, \lambda_t, \mathbf{T}_{\mathcal{J}}, \xi_{\mathcal{J}}) = & -\frac{1}{2} \int_{\Omega_e} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} + \int_{\Omega} \left[\frac{1}{\varepsilon} (\nabla \times \mathbf{H}_t) \cdot (\nabla \times \mathbf{T}_{\mathcal{J}}) - \lambda_t \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t \right] d\mathbf{x} \\ & + \int_{\Gamma_M} \left[\mathbf{T}_{\mathcal{J}} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) + \mathbf{H}_t \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} \right) \right] ds + \frac{1}{2} \xi_{\mathcal{J}} \left(\int_{\Omega} \mu \mathbf{H}_t \cdot \mathbf{H}_t \, d\mathbf{x} - 1 \right). \end{aligned}$$

The gradient of the objective $\mathcal{J}$ with respect to the design variables can now be obtained (under suitable smoothness conditions) by taking the partial derivatives of the $\mathcal{L}_{\mathcal{J}}$ with respect to the design variables. Since the domain and boundary depend on the design variables, the *material derivative* of volume and surface integrals is needed [10]. Thus, the i^{th} component of the gradient of $\mathcal{J}$ is given by

$$\begin{aligned} \frac{D\mathcal{J}}{Dq_i} \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}_{\mathcal{J}}}{\partial q_i} = & -\frac{1}{2} \int_{\Gamma_e} \mu \mathbf{H}_t \cdot \mathbf{H}_t (\mathbf{V}_i \cdot \mathbf{n}) \, ds \\ & + \int_{\Gamma} \left[\frac{1}{\varepsilon} (\nabla \times \mathbf{H}_t) \cdot (\nabla \times \mathbf{T}_{\mathcal{J}}) - \lambda_t \mu \mathbf{T}_{\mathcal{J}} \cdot \mathbf{H}_t + \frac{1}{2} \xi_{\mathcal{J}} \mu \mathbf{H}_t \cdot \mathbf{H}_t \right] (\mathbf{V}_i \cdot \mathbf{n}) \, ds \\ & + \int_{\Gamma_M} (\mathbf{V}_i \cdot \mathbf{n}) (\mathbf{n} \cdot \nabla + \nabla \cdot \mathbf{n}) \left[\mathbf{T}_{\mathcal{J}} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_t \right) + \mathbf{H}_t \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_{\mathcal{J}} \right) \right] ds. \end{aligned} \quad (7)$$

Here, $\nabla \cdot \mathbf{n}$ is the mean curvature and the normal derivative $\mathbf{n} \cdot \nabla$ operates on the term in square brackets. Γ_e denotes the boundary of the end cell, and $\mathbf{V}_i$ is the so-called *design velocity field*, that is, the derivative of the boundary coordinate with respect to the i^{th} design variable,

$$\mathbf{V}_i(\mathbf{x}) = \frac{\partial \mathbf{x}}{\partial q_i}.$$

This is computed by differentiating the surface model with respect to the design variables, i.e. at the discrete level by differentiating the CAD model (typically done by numerical differentiation or AD). The state eigenpair $(\mathbf{H}_t, \lambda_t)$ is found by solving the state eigenvalue problem (1), and the adjoint eigenpair $(\mathbf{T}_\mathcal{J}, \xi_\mathcal{J})$ by solving the adjoint eigenvalue problem (6). Once the $\mathbf{V}_i$ and the state and adjoint eigenpairs have been found, the surface integrals in (7) are evaluated to determine the objective gradient. One difficulty in computing the surface integrals (by numerical quadrature) is the need evaluate derivatives of the computed state and adjoint fields on the surface. The second term on the right side of (7) involves curls of the state and adjoint eigenvectors; more problematic is the third term, which involves the normal derivative of the tangential component of curls of these fields. A loss of accuracy will generally occur in numerically differentiating the computed field, and this can be severe enough to render the gradient insufficiently accurate for optimization purposes. There are a number of ways to mitigate this difficulty; our choice is to use high order Nédélec bases [11] to approximate the state and adjoint eigenvectors. Note however that the third integral in (7) is defined only on the magnetic boundary Γ_M . Since the magnetic boundary is usually a symmetry plane, $\mathbf{V} \cdot \mathbf{n}$ is typically zero on this boundary, and the third integral vanishes. Thus, only a single derivative of the state and adjoint eigenvector fields is usually needed. Care must also be taken at edges and corners of the domain that appear in (7), where the surface normal $\mathbf{n}$ is discontinuous. The mean curvature term then provides additional delta function contributions at these points.

4. Design gradient of the accelerating mode frequency constraint

The procedure for deriving an expression for the design gradient of the constraint $\mathcal{C}$ follows that for the design gradient of the objective $\mathcal{J}$ in Section 3. In this case the Lagrangian resembles (4), except the objective functional (2) is replaced by the constraint function (3), and the state eigenvalue problem (1) is written for $(\mathbf{H}_a, \lambda_a)$, the eigenpair corresponding to the accelerating mode,

$$\begin{aligned} \mathcal{L}_\mathcal{J}(\mathbf{H}_a, \lambda_a, \mathbf{q}, \mathbf{T}_c, \mathbf{\Lambda}_{M_c}, \mathbf{\Lambda}_{E_c}, \xi_c) &= \lambda_a - \lambda^* + \int_\Omega \mathbf{T}_c \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_a - \lambda_a \mu \mathbf{H}_a \right) d\mathbf{x} \\ &+ \int_{\Gamma_M} \mathbf{\Lambda}_{M_c} \cdot (\mathbf{n} \times \mathbf{H}_a) d\mathbf{s} + \int_{\Gamma_E} \mathbf{\Lambda}_{E_c} \cdot \left(\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_a \right) d\mathbf{s} + \frac{\xi_c}{2} \left(\int_\Omega \mu \mathbf{H}_a \cdot \mathbf{H}_a d\mathbf{x} - 1 \right), \end{aligned} \quad (8)$$

where $(\mathbf{T}_c, \mathbf{\Lambda}_{M_c}, \mathbf{\Lambda}_{E_c}, \xi_c)$ are adjoint variables for the constraint $\mathcal{C}$. Taking variations with respect to the state eigenpair $(\mathbf{H}_a, \lambda_a)$ and integrating by parts, we arrive at the *adjoint eigenvalue problem for the constraint*,

$$\begin{aligned} \nabla \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{T}_c \right) - \lambda_a \mu \mathbf{T}_c + \mu \xi_c \mathbf{H}_a &= \mathbf{0} \quad \text{in } \Omega \\ \mathbf{n} \times \mathbf{T}_c &= \mathbf{0} \quad \text{on } \Gamma_M \\ \mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_c &= \mathbf{0} \quad \text{on } \Gamma_E \\ \int_\Omega \mu \mathbf{T}_c \cdot \mathbf{H}_a d\mathbf{x} &= -1 \end{aligned} \quad (9)$$

The operator for this adjoint problem is the same as that for the objective adjoint eigenvalue problem (6). The right-hand side however is different; it vanishes for the first equation, and is nonzero for the last. This special structure permits us to write the solution directly. First, by taking the inner product of the left-hand side of the first equation in (9) with $\mathbf{H}_a$, integrating by parts using (5), and invoking the state and adjoint eigenvector boundary conditions $\mathbf{n} \times \mathbf{H}_a = \mathbf{n} \times \mathbf{T}_c = \mathbf{0}$ on Γ_M and $\mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_a = \mathbf{n} \times \frac{1}{\varepsilon} \nabla \times \mathbf{T}_c = \mathbf{0}$ on Γ_E we obtain

$$\int_\Omega \mathbf{T}_c \cdot \left(\nabla \times \frac{1}{\varepsilon} \nabla \times \mathbf{H}_a - \lambda_a \mu \mathbf{H}_a \right) d\mathbf{x} + \int_\Omega \mu \xi_c \mathbf{H}_a \cdot \mathbf{H}_a d\mathbf{x} = 0.$$

The first integral is zero because the eigenpair $(\mathbf{H}_a, \lambda_a)$ satisfies the state eigenvalue problem (1). Therefore the adjoint eigenvalue for the constraint, ξ_c , is zero. Substituting this result into the first equation of (9), we see that $\mathbf{T}_c$ must satisfy the state eigenvalue problem (1) for the accelerating mode; i.e. when λ_a is a simple eigenvalue, we must have $\mathbf{T}_c = \alpha \mathbf{H}_a$ for some α . Substituting this result into the last equation of (9) and making use of the orthonormalization condition in (1), we obtain $\alpha = -1$. Therefore, the adjoint eigenvector for the constraint $\mathbf{T}_c$ is just the negative of the state eigenvector for the accelerating mode $\mathbf{H}_a$. Thus, there is no need to solve the adjoint eigenvalue problem for the constraint, (9) — and this is true in general whenever we seek the gradient of a function that depends only on the state eigenvalue. Making use of this solution for $(\mathbf{T}_c, \xi_c)$ in the Lagrangian for the constraint (8), taking the partial derivative with respect to the i^{th} design variable, and invoking the magnetic boundary condition for the accelerating mode $\mathbf{n} \times \mathbf{H}_a = \mathbf{0}$, we arrive at the expression for the i^{th} component of the constraint gradient:

$$\frac{DC}{Dq_i} \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}_c}{\partial q_i} = - \int_{\Gamma} \left[\frac{1}{\varepsilon} (\nabla \times \mathbf{H}_t) \cdot (\nabla \times \mathbf{H}_t) - \lambda_t \mu \mathbf{H}_t \cdot \mathbf{H}_t \right] (\mathbf{V}_i \cdot \mathbf{n}) \, ds \quad (10)$$

5. Numerical example

We briefly present a simple example of optimizing an ILC-like structure in which the shape of the end cells is optimized according to the optimization problem (2)–(3). The surfaces of the cells are represented by analytic expressions, which permits readily-computed design velocity fields $\mathbf{V}_i$. Four variables are chosen to parameterize the shape. Implementation of full mesh movement is currently under way; as a temporary solution just the outer layer of elements is permitted to move with the changing surface. We bound the design variable change to keep the mesh from distorting. At each optimization iteration, the state eigenvalue problem (1) is solved by the Omega3P finite element electromagnetics code [1], and trapped-mode and accelerating mode eigenpairs are identified automatically. The adjoint eigenvalue problem for the objective (6) is solved by an adjoint extension to Omega3P, using the same elements, mesh, and many of the same numerical components used in the state eigenproblem solver. The objective (2), constraint (3), objective gradient (7), and constraint gradient (10) expressions involving boundary and domain integrals are also computed using an extension to Omega3P. At each iteration, the function and gradient information, along with bounds on the design variables, are passed to the sequential quadratic programming code fSQP [12], which solves a quadratic programming subproblem to generate a search direction, performs a line search, and returns a new design vector. The updated design variables generate a new surface, the mesh is deformed to its new location, and the process repeated until convergence.

We defer a detailed evaluation of the computational complexity, parallel implementation, and scalability to a subsequent article. Here, we give a brief discussion of the relative costs of a typical iteration. The dominant cost of a design iteration is in solving the Maxwell (state) eigenvalue problem for frequencies and mode shapes of interest, which overwhelms the costs of solving the linear adjoint eigenvalue problem and remeshing. The latter two in turn dominate the cost of evaluating the boundary integrals to compute the objective and constraint gradients, and the cost of differencing the geometric model to obtain the design velocity field, both of which depend on surface rather than volume computations.

Figure 4 shows a cross-section through the mesh corresponding to the initial shape (blue) and final shape (red, the change in which has been amplified for clarity). Convergence to the optimum occurs in 5 iterations, for which the value of the objective — the trapped energy in the end cell — is improved by 58%. Moreover, the initial design is not even feasible — it violates the accelerating frequency constraint by 17%. However, after 5 iterations, the optimum shape's accelerating frequency is brought to within 0.15% of the desired frequency of 1.3 GHz.

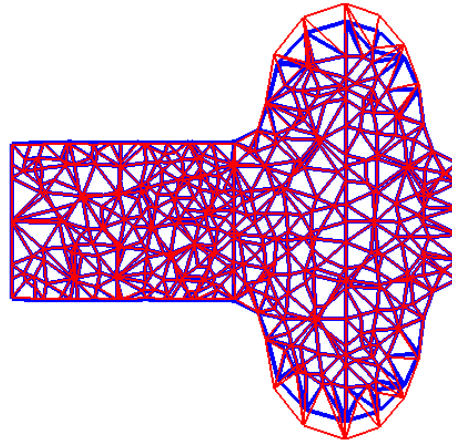


Figure 4. Meshes for initial and optimized designs of end cell.

Since each iteration is dominated by the (state) eigenvalue solve, the *optimum* is found in a small multiple of the cost of the *simulation*. While this example is relatively small, it serves to demonstrate the power of (adjoint-based) shape optimization: one is able to move beyond merely *evaluating* the performance of a proposed design, to actually *optimizing* it. Already this points to a revolution in the way accelerator structures will be designed.

References

- [1] Z. Li, N. Folwell, L. Ge, A. Guetz, V. Ivanov, M. Kowalski, L.-Q. Lee, C.-K. Ng, G. Schussman, R. Uplenchwar, M. Wolf, L. Xiao, and K. Ko. High performance computing in accelerator structure design and analysis. In *Proceedings of the 8th International Computational Accelerator Physics Conference*, 2004.
- [2] H.-J. Kim, T.J. Tautges, and J. Uicker. Unstructured parallel mesh generation by distributing and partitioning solid geometry model. In *SIAM Conference on Computational Science and Engineering*, Orlando, Florida, 2005.
- [3] M. Brewer, L. Freitag Diachin, P. Knupp, T. Leurent, and D. Melander. The Mesquite mesh quality improvement toolkit. In *Proceedings, 12th International Meshing Roundtable*, pages 239–250, September 2003.
- [4] L. Ge, L.-Q. Lee, Z. Li, C. Ng, K. Ko, Y. Luo, and M. Shephard. Adaptive mesh refinement for high accuracy wall loss determination in accelerating cavity design. In *Proceedings of the Eleventh Biennial IEEE Conference on Electromagnetic Field Computation*, 2004.
- [5] Y. Sun, N. Folwell, Z. Li, and G. Golub. High precision accelerator cavity design using the parallel eigensolver Omega3P. In *Proceedings of the 18th Annual Review of Progress in Applied Computational Electromagnetics*, 2002.
- [6] C. Yang, W. Gao, Z. Bai, X. Li, L.-Q. Lee, P. Husbands, and E. Ng. An algebraic sub-structuring method for large-scale eigenvalue calculation. Technical Report LBNL-55050, Lawrence Berkeley National Laboratory, 2004.
- [7] V. Akçelik, G. Biros, O. Ghattas, J. Hill, D. Keyes, and B. van Bloemen Waanders. Parallel algorithms for PDE-constrained optimization. In M. Heroux, P. Raghaven, and H. Simon, editors, *Frontiers of Parallel Computing*. SIAM, 2006. To appear.
- [8] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver. Part II: The Lagrange-Newton solver, and its application to optimal control of steady viscous flows. *SIAM Journal on Scientific Computing*, 2005. To appear.
- [9] J.-L. Lions. *Some Aspects of the Optimal Control of Distributed Parameter Systems*. SIAM, 1972.
- [10] O. Pironneau. *Optimal Shape Design for Elliptic Systems*. Springer-Verlag, 1983.
- [11] P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford, 2003.
- [12] C.T. Lawrence, J.L. Zhou, and A.L. Tits. *User's Guide for CFSQP Version 2.3: A C Code for Solving (Large Scale) Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality Constraints*. Institute for Systems Research, University of Maryland, 1995.