OPEN

A topology optimization algorithm for magnetic structures based on a hybrid FEM–BEM method utilizing the adjoint approach

Gregor Wautischer^{1✉}, Claas Abert^{1,2}, Florian Bruckner¹, Florian Slanovc¹ & Dieter Suess^{1,2}

A method to optimize the topology of hard as well as soft magnetic structures is implemented using the density approach for topology optimization. The stray field computation is performed by a hybrid finite element–boundary element method. Utilizing the adjoint approach the gradients necessary to perform the optimization can be calculated very efficiently. We derive the gradients using a “first optimize then discretize” scheme. Within this scheme, the stray field operator is self-adjoint allowing to solve the adjoint equation by the same means as the stray field calculation. The capabilities of the method are showcased by optimizing the topology of hard as well as soft magnetic thin film structures and the results are verified by comparison with an analytical solution.

In many applications like magnetic sensors, magnetic storage, electric machines and so on, magnetic fields with specific spatial profiles are exploited in order to achieve a specific goal. The properties of magnetic fields are defined by the properties of the involved magnetic structures. Most commonly, in order to obtain a certain magnetic field, the geometry of a magnetic structure is optimized. This can e.g. be the geometry of a permanent magnet producing a well defined magnetic field, or a soft magnetic structure shaping the field possibly generated by an electric current. The process of optimizing the geometry of these structures can be very time consuming. A common approach is called geometry parameterization where the structure's geometry is parameterized with relatively few parameters and the parameter region is investigated until an optimal solution is found. Geometry parameterization is one representative of a group of techniques summarized as shape optimization¹. Another approach is to directly optimize the distribution of material within a given domain. This approach is called topology optimization². It drastically increases the degrees of freedom (dofs) but allows for possibly better, geometric realizations of the considered system, independent of any introduced parameterizations. Successful applications of topology optimization of magnetic structures include the optimization of magnetic recording write heads³, of rotors of brushless DC motors⁴ and of electromagnetic actuators^{5,6}.

In this paper a method to optimize the topology of magnetic structures is presented. The presented method is based on a highly efficient hybrid finite element–boundary element method (FEM–BEM) approach⁷ solving the magnetostatic Maxwell's problem. In contrast to already presented finite element^{3,8} and finite differences⁹ algorithms using a FEM–BEM approach has the advantage that only the regions of interest need to be discretized reducing the dofs dramatically. Furthermore, in order to efficiently calculate gradients during optimization, the adjoint approach is utilized. Our method takes the B/H curve of a hard or soft magnetic material linearized at the working point as input and is able to include external fields generated by electric coils. After optimization, the result is an optimized magnetic structure and its stray field dependent magnetization state.

The algorithm is implemented on the basis of magnum.fe¹⁰ which in turn utilizes the finite element library FEniCS¹¹. In particular FEniCS is used for the definition of the objective functionals and to automatically differentiate the involved partial derivatives.

The paper is structured as follows. In the first section the solution of the forward problem—the calculation of the stray field of a magnetic structure—is presented. In the following section the density approach to topology optimization is introduced. After that the gradient necessary to perform the optimization and the adjoint equation are introduced. Thereafter details of the numerical optimization are discussed and numerical experiments

¹Faculty of Physics, University of Vienna, Vienna, Austria. ²Research Platform MMM Mathematics-Magnetism-Materials, University of Vienna, Vienna, Austria. ✉email: gregor.wautischer@univie.ac.at

showcasing the capabilities of the method are presented. A detailed derivation of the gradient using the adjoint approach within a “first optimize then discretize” scheme is given in the appendix.

Forward problem

The stray field of a magnetic body consisting of a magnetic material with linear magnetic susceptibility χ and possibly a remanence magnetization $\mathbf{M}_r$ within an external magnetic field $\mathbf{H}_{\text{ext}}$, can be calculated starting from the magnetostatic Maxwell's equation. Using the reduced scalar potential formulation the total magnetic field $\mathbf{H}$ is split into an external part $\mathbf{H}_{\text{ext}}$ (created via currents outside of Ω_m) and the curl-free induced magnetic field $\mathbf{H}_d$. Expressing the induced field as gradient of a scalar potential u , $\mathbf{H}_d = -\nabla u$ finally yields

$$\nabla \cdot [(1 + \chi)(\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r] = 0. \quad (1)$$

In order to use the hybrid FEM–BEM approach, next $\mathbb{R}^3$ is divided into a region Ω_m containing all magnetic material and a region $\mathbb{R}^3 \setminus \Omega_m$. Outside Ω_m Eq. (1) then reduces to

$$\Delta u = 0 \quad \text{in } \mathbb{R}^3 \setminus \Omega_m \quad (2)$$

and the following boundary and jump conditions apply

$$[u]_{\partial\Omega_m} = 0 \quad (3)$$

$$\left[(1 + \chi) \frac{\partial u}{\partial \mathbf{n}} \right]_{\partial\Omega_m} = [\chi]_{\partial\Omega_m} \mathbf{n} \cdot \mathbf{H}_{\text{ext}} + \mathbf{n} \cdot \mathbf{M}_r \quad (4)$$

$$u(\mathbf{r}) \rightarrow \mathcal{O}\left(\frac{1}{|\mathbf{r}|}\right) \quad \text{for } |\mathbf{r}| \rightarrow \infty, \quad (5)$$

where $\mathbf{n}$ is the outward pointing normal vector and $[\cdot]_{\partial\Omega}$ denotes the jump over the boundary $\partial\Omega$. The resulting system is solved using a hybrid FEM–BEM algorithm⁷.

Topology optimization

Topology optimization tries to find, within a predefined spatial domain Ω_{opt} , the optimal material distribution with respect to a design goal. This can be achieved using the density approach for topology optimization¹². Here, a scalar indicator function

$$\rho(\mathbf{x}) \in [0, 1], \quad \text{where } \rho(\mathbf{x}) = \begin{cases} 0: \text{no material} \\ 1: \text{material} \end{cases} \quad (6)$$

is introduced, transforming the remanence magnetization $\mathbf{M}_r \rightarrow \mathbf{M}_r(\rho) = \rho^p \mathbf{M}_r$ as well as the susceptibility $\chi \rightarrow \chi(\rho) = \rho^p \chi$. Here, p is a parameter introduced in order to penalize intermediate values of $0 \leq \rho \leq 1$ that are allowed during optimization¹². However, it turns out that the impact of p is more complex as is investigated for each presented numerical experiment (see “Numerical experiments” section).

The forward problem then reads

$$F(u, \rho) = \nabla \cdot [(1 + \rho^p \chi)(\nabla u - \mathbf{H}_{\text{ext}}) - \rho^p \mathbf{M}_r] = 0, \quad (7)$$

where u is an at least twice and ρ an at least once differentiable function. Note, that after transformation into its weak form and shifting of the divergence operator onto the test function by application of partial integration, these conditions are relaxed and after discretization piecewise linear basis function ($\mathcal{P}_1$) can be used for the scalar potential u and the scalar indicator function ρ only has to be constant within each element.

The numerical optimization is performed by casting the design goal into an objective functional $\hat{J}(\rho) = J(\mathbf{H}(\rho), \rho)$ and minimizing it using an iterative, gradient based optimization algorithm.

Adjoint approach

In order to efficiently minimize the reduced objective functional $\hat{J}(\rho)$ containing the design goal, its gradient with respect to the design variable ρ , $\nabla_\rho \hat{J}$ is needed. This is where the adjoint approach comes into play, offering a very efficient method to calculate the gradient. As derived in detail in the appendix (“Appendix” section) within a “first optimize then discretize” scheme, the gradient is given by

$$\nabla_\rho \hat{J} = (p\rho^{p-1}[\chi(\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]) \cdot \nabla \lambda + \mathcal{R}\left(\frac{\partial J}{\partial \rho}\right), \quad (8)$$

where λ is the adjoint variable that is obtained by solving the adjoint equation

$$\nabla \cdot \left((1 + \rho^p \chi) \nabla \lambda - \mathcal{R}\left(\frac{\partial J}{\partial \mathbf{H}}\right) \right) = 0 \quad (9)$$

with boundary condition

Figure 1. Mesh and geometry used for the numerical experiments. Dimensions are given in μm . The thickness (z -dimension) of the thin film is $12.5 \mu\text{m}$. The field box Ω_h consists of 91 cells and the optimization region Ω_{opt} consists of 16505 cells.

$$\lambda(\mathbf{r}) \rightarrow \mathcal{O}\left(\frac{1}{|\mathbf{r}|}\right) \quad \text{for } |\mathbf{r}| \rightarrow \infty. \quad (10)$$

Here $\mathcal{R}$ are the Riesz representers of the corresponding functionals as introduced in “Appendix” section. Note, that the adjoint equation has the same form as the forward problem. Therefore, the same algorithm can be utilized to solve the forward as well as the adjoint equation.

Numerical experiments

In the following, the topology of thin film magnetic flux guide concentrators is optimized towards maximization of the vertical (y -) component of the magnetic field H_y inside the target domain Ω_h . The objective functional that is to be minimized reads $\hat{J} = - \int_{\Omega_h} H_y \, dV$. The geometry and the mesh used is depicted in Fig. 1 where also the optimization domain (Ω_{opt}) and the domain in which the magnetic field is evaluated (Ω_h) is specified. Note that $\Omega_m = \Omega_{\text{opt}} \cup \Omega_h$. Note, that we use an irregular mesh in order to avoid any solution bias introduced by a given regularity. The presented numerical experiments include the usage of hard as well as soft magnetic material.

The potential u and the adjoint variable λ are calculated using piecewise linear basis functions ($\mathcal{P}_1$) and the scalar indicator function ρ as well as χ , $\mathbf{M}_r$, $\mathbf{H}_{\text{ext}}$ and the derived strayfield $-\nabla u$ and $\nabla \lambda$ are constant within each element.

The optimizations are performed using the Scipy implementation of the limited-memory Broyden-Fletcher-Goldfarb-Shanno minimization algorithm with bounds (L-BFGS-B)¹³. Note, that as shown in¹⁴, common ready-to-use optimization algorithms such as this, are constructed to only receive coefficient vectors as input and internally use the euclidean (ℓ^2) inner product to perform all necessary operations (e.g. approximation of the Hessian). Since here an optimization in a Hilbert spaces with the L^2 inner product is performed, this leads to mesh dependent convergence rates. To prevent this, there exists the python package Moola¹⁵ that internally uses the correct inner product during optimization. However, in order to perform topology optimization bounds on the optimization variable ρ have to be enforced since $0 \leq \rho \leq 1$ and so far, there is no optimization algorithm offered by Moola enabling bounds on the optimization variable. Furthermore, note that the L^2 Riesz representer of the gradient is independent of the flux concentrator’s geometrical dimensions. However, the objective functional is not since it includes an integration over the volume Ω_h . For the chosen geometry dimensions the gradient size is relatively large compared to the size of the objective functional. This leads to the L-BFGS-B algorithm terminating prematurely with problems during the line search, simply returning the initial distribution of ρ . This is since the Armijo rule, one of the Wolfe conditions necessary to end the line search is never met. The objective functional value exposed to the optimization algorithm, is therefore scaled to avoid this problem. This scaling does not alter the results since internally, the L-BFGS-B algorithm performs a quasi-Newton method on the gradient and the function value is only used during the line search.

For results of optimizations that terminate with dofs of ρ at intermediate values $0 \leq \rho \leq 1$ a brute force regularization is applied where all dofs with $\rho < 0.5$ are set to 0 and all dofs with $\rho \geq 0.5$ are set to 1.

Permanent magnet optimization. First, the topology for an ideal permanent magnet with susceptibility $\chi = 0$, remanence magnetization $M_r = 280000 \text{ A/m}$ and the anisotropy axis chosen as the y -axis is optimized towards maximization of the vertical (y -) component of the magnetic field H_y inside the domain Ω_h , as introduced above.

As presented in⁸ for a point like target domain, the optimal topology can be derived analytically by considering the magnetization inside Ω_{opt} as being comprised of single dipoles μ . A dipole at a vertical (y -) distance produces a positive flux $B_y(\mathbf{x})$ at a point $\mathbf{x}$ for an horizontal (x -) distance of

$$x < \sqrt{2}y. \quad (11)$$

The surface condition $x = \sqrt{2}y$ therefore shows that the analytical solution is a cone with an opening angle of $\alpha \approx 35.26^\circ$. Note that the assumption of a point like target domain is still valid, since the changes due to the finite

Figure 2. Optimized topology for the ideal permanent magnet and comparison with the analytical solution (green line).

Figure 3. Dependence of n_{fev} on γ and size of the gradient's smallest dof value. The black horizontal line represents a value of 1.

Figure 4. Optimized topology for the optimization with $\chi = 0.2$.

size of Ω_h are too small to be resolved by the chosen mesh. The analytical solution is shown in Fig. 2, where it is compared to the optimal numerical solution.

For the ideal permanent magnet, the optimized topology, as shown in Fig. 2, produces a mean field $\bar{H}_y = \int_{\Omega_h} H_y dV / V_{\Omega_h}$ of 22.83 mT. This solution is found independently of the initial distribution of the density functional ρ_{init} for a penalty parameter of $p \leq 4$. Furthermore, scaling the gradient has no influence on the obtained solution, but increasing the gradient size reduces the number of function evaluations n_{fev} needed. This is since the size of the gradient determines how fast the dofs of ρ are set to 0 or 1. This shows, that the optimization problem in the case of zero susceptibility is linear in ρ as expected since for $\chi = 0$ the only source term in the forward problem (7) is given by $\rho^p \mathbf{M}_r$ where $\mathbf{M}_r$ is constant. In Fig. 3 the value of the smallest dof in the gradient and n_{fev} are plotted against the gradient scaling factor γ for $p = 1$. As soon as the smallest dof is larger 1 for $\gamma = 10^3$, all elements are set to 0 or 1 during the optimization, leaving no dofs with intermediate values of ρ and n_{fev} reaches a minimum of 2.

Realistic hard magnetic material $\chi = 0.2$. Using a realistic hard magnetic material with a susceptibility of $\chi = 0.2$, the optimal solution found has a mean field of $\bar{H}_y$ of 22.35 mT. This smaller mean field value compared to the ideal permanent magnet is due the effect of demagnetization. The optimal topology found (Fig. 4) differs only slightly from the topology obtained from the ideal permanent magnet optimization.

The optimal solution is found independently of the initial distribution of the density functional ρ for $p = 1$. For larger p -values the solution gets worse showing, that the optimization is no longer independent of the parameter p . However, for $p \leq 4$ the difference in the mean field is below 0.5%. Scaling the gradient again reduces the number of necessary function evaluations to perform the optimization n_{fev} . However, since the optimization problem is no longer linear in ρ , the number of function evaluations n_{fev} needed is larger than for

Figure 5. Dependence of the number of necessary function evaluations n_{fev} on γ with $p = 1$ for the optimization with $\chi = 0.2$.

χ	$\bar{H}_y$ (mT)	Increase (%)
10^2	14.65	52.25
10^3	21.95	84.38
10^4	23.34	91.23

Table 1. Mean field values $\bar{H}_y$ and increase with respect to the full plate for the optimized topologies.

the ideal permanent magnet case. In Fig. 5 n_{fev} in dependence of the gradient scaling factor γ for $p = 1$ is shown. The optimization is fastest for a constant initial distribution of $\rho_{\text{init}} = 1$ where for $\gamma \geq 10$ $n_{\text{fev}} < 250$. Also, the optimization process terminates before all dofs of ρ are set to 0 or 1, although the number of dofs with an intermediate value of ρ is negligible for all simulations with $p = 1$. It has been verified that no better solution is found by scaling the gradient down to $\gamma = 10^{-5}$. Note, that the topology obtained from the optimization using an ideal permanent magnetic material, for a realistic hard magnetic material with $\chi = 0.2$ produces a magnetic field that is only negligibly smaller ($< 0.1\%$) than the optimal topology obtained using the realistic hard magnetic material during optimization. This confirms, that $\chi = 0$ is a good assumption for a permanent magnetic structure.

Optimization of a soft magnetic structure.

Next, the topology of a soft magnetic thin film is optimized to act as a magnetic flux guide concentrator maximizing the magnetic flux in vertical (y -)direction inside Ω_h . Such a structure can be used to e.g. improve the magnetic field detection limits of spin valve sensors¹⁶ by enhancing the external field at the sensor position.

The optimization is performed for three

different values of χ of 10^2 , 10^3 and 10^4 with an external field of 10 mT applied. The optimized topologies for the three different χ -values are shown in Fig. 6. The found structures show the same “tanga” shape as reported in¹⁶ as optimal design. While the optimal topology differs only slightly for $\chi = 10^3$ and $\chi = 10^4$, for $\chi = 10^2$ the topology is significantly broader. In fact, switching the optimal topologies of $\chi = 10^3$ and $\chi = 10^4$ changes the generated mean field by under 0.2%. The generated mean field and the increase with respect to the full plate ($\rho = 1$ in Ω_{opt}) is shown in Table 1.

As opposed to the optimization of hard magnetic materials, the optimization using a soft magnetic material depends much more on the initial distribution of the density functional ρ_{init} as well as on the p -parameter. In Fig. 7 the p -dependence of the mean field $\bar{H}_y$ for the regularized solutions for $\chi = 10^3$ is plotted for different initial distributions of ρ_{init} and for different values of the gradient scaling factor γ .

The solution found for a constant $\rho_{\text{init}} = 1$ and a randomly initialized ρ_{init} are clearly inferior to the solutions found with a constant $\rho_{\text{init}} = 0.5$. For $\rho_{\text{init}} = 0.5$ the mean field increases with increasing p until it plateaus. While decreasing γ does not lead to a higher mean field, the p -value at which the mean field plateaus decreases with decreasing γ . At $\gamma = 1$ the plateau starts at $p = 26$ while for $\gamma = 10^{-4}$ the plateau value is 14. Further decreasing γ does not alter the plateau value anymore. In fact, as shown in Fig. 8 also for single p -values, the mean field increases with decreasing γ until a maximum is found for 10^{-4} . Note, that for higher p -values than the plateau value, the mean field differs well under 1% from the mean field obtained by the optimization performed with the plateau value, except for individual outliers with very small mean field values for large p -values.

In Fig. 9, the amount of dofs with intermediate values of ρ after optimization is plotted in dependence of p . While the solutions with $p \leq 4$ show no dofs with intermediate values of ρ , with increasing p , their number increases and reaches over 60% for the optimal solutions. This is in contrast to the original purpose of the penalization factor p . However, since the changes of the mean field $\bar{H}_y$ due to regularization is under 1% for all optimizations, the effect of the regularization is still negligible. This is since the remaining intermediate values

Figure 6. Optimized topologies before (top row) and after regularization (bottom row) for (a) $\chi = 10^2$ (for $p = 5$) (b) $\chi = 10^3$ (for $p = 14$) and (c) $\chi = 10^4$ (for $p = 17$) after regularization for $\gamma = 10^{-4}$.

Figure 7. p -dependency for different initial distributions and gradient scaling factors γ for the optimization performed for $\chi = 10^3$.

Figure 8. Dependency of the mean field $\bar{H}_y$ on γ individual p -values.

are very close to 0 and 1 respectively. In Fig. 6 the optimized topologies are shown before and after regularization for comparison.

The dependence of the number of function evaluations necessary to perform the optimizations n_{fev} on p is plotted in Fig. 9 for two different values of the gradient scaling factor. It can be seen n_{fev} is lower with a smaller gradient scaling factor. With a gradient scaling factor of 10^{-4} , the optimization with $p = 14$ needs 325 n_{fev} while with a gradient scaling factor of 1 for the optimization with $p = 26$ $n_{\text{fev}} = 14631$. In Fig. 10 the iteration histories are plotted for selected values of p for a gradient scaling factor of 10^{-4} . Note that the different starting values of the optimizations are due to factorization of ρ by p .

Figure 9. p -dependency of n_{fev} and the percentage of dofs with intermediate values of ρ after optimization for values of γ for the optimization performed for $\chi = 10^3$.

Figure 10. Iteration histories for selected values of p for a gradient scaling factor of 10^{-4} .

For $\chi = 10^4$ the p -dependence of the mean field plateaus in the same way as for $\chi = 10^3$, although with a different plateau value of $p = 17$ for a gradient scaling factor of 10^{-4} . The dependency on the gradient scaling factor is also similar as shown in Fig. 8. For $\chi = 100$ the situation is slightly different. Here, for gradient scaling factors smaller 10^{-4} the mean field increases with increasing p -value until a maximum is found at $p = 5$ and the mean field decreases when further increasing p .

Conclusion

An algorithm to optimize the topology of magnetic structures based on a hybrid FEM–BEM method and the density approach for topology optimization has been presented. The gradients necessary to efficiently solve the optimization are obtained very efficiently using the adjoint approach. It has been shown that within a “first optimize then discretize” scheme the continuous forward operator is self-adjoint and therefore, the same method can be used to solve the forward as well as the adjoint problem.

The dependence on the p -parameter, originally introduced to penalize intermediate values of the density indicator function $0 \leq \rho \leq 1$ has been investigated for the presented numerical experiments. For ideal permanent magnetic structures ($\chi = 0$) the linearity of the optimization problem can be utilized by enlarging the gradient. Scaling the gradient by a factor $> 10^3$, with $p = 1$ the optimization needs only two function evaluations and the found solution is in accordance with analytical expectations and shows no intermediate values of the indicator function ρ . For a realistic permanent magnetic material with $\chi = 0.2$ the optimized topology produces a magnetic field that differs by $< 0.1\%$ from the topology found for the ideal permanent magnetic material. This shows that for thin films with negligible demagnetization, $\chi = 0$ is good approximation for a realistic permanent magnetic material with $\chi = 0.2$. Furthermore, the number of function evaluation needed using a realistic hard magnetic material is in the hundreds and hence much larger than the number of evaluations needed for the assumption of an ideal permanent magnetic material.

For the optimization of soft magnetic materials the optimal solutions are found for a gradient scaling factor $\leq 10^{-4}$. The solutions found get better with increasing value of the p -parameter. For $\chi = 10^2$ the best solution is found for $p = 5$ and gets worse for larger p -values. For $\chi = 10^3$ and $\chi = 10^4$ the solutions improve with p until it plateaus for $p \geq 14$ and $p \geq 17$ respectively. Note, that in contrast to its original purpose, the p -parameter does not decrease the number of dofs with intermediate values of ρ . On the contrary, the percentage of dofs with intermediate values of ρ increases with increasing p . However, using a brute force regularization forcing all dofs with $\rho < 0.5$ to 0 and all dofs with $\rho \geq 0.5$ to 1 only negligibly impacts the field value created by the structures. The optimized topologies have a shape similar to the optimal shape found experimentally in¹⁶.

Appendix

Deriving the gradient $\nabla_{\rho}\hat{J}$ using the adjoint approach. In the following the adjoint approach is applied to obtain the gradient $\nabla_{\rho}\hat{J}$ of the reduced objective functional $\hat{J}(\rho)$ with respect to the design variable ρ . A “first optimize then discretize” scheme¹⁷ is used, within which the gradient is derived using continuous operators. These operators are, as will turn out, self-adjoint where the discrete forward operator of the hybrid FEM–BEM algorithm is not. This simplifies the numerics involved in solving the adjoint equation.

Following¹⁷, in order to derive the gradient $\nabla_{\rho}\hat{J}$ we consider the derivative

$$\frac{d\hat{J}}{d\rho} = \frac{\partial J}{\partial u} \frac{du}{d\rho} + \frac{\partial J}{\partial \rho}. \quad (12)$$

Here, $\frac{du}{d\rho}$ is the numerically challenging term since u depends on ρ via the forward problem (7). In the discretized form, naive evaluation of $\frac{du}{d\rho}$ by means of finite difference would require n_{ρ} solutions of the forward problem, where n_{ρ} is the number of dofs of ρ . Looking at the derivative of the forward problem

$$\frac{dF}{d\rho} = \frac{\partial F}{\partial u} \frac{du}{d\rho} + \frac{\partial F}{\partial \rho} = 0 \quad (13)$$

the critical term can be expressed as

$$\frac{du}{d\rho} = -\left(\frac{\partial F}{\partial u}\right)^{-1} \frac{\partial F}{\partial \rho}. \quad (14)$$

Inserting this into Eq. (12) gives

$$\frac{d\hat{J}}{d\rho} = -\frac{\partial J}{\partial u} \left(\frac{\partial F}{\partial u}\right)^{-1} \frac{\partial F}{\partial \rho} + \frac{\partial J}{\partial \rho}, \quad (15)$$

an expression for the derivative where $\frac{du}{d\rho}$ does not appear explicitly anymore.

Consider now that the goal is to perform an optimization on ρ , an at least once differentiable element of a suitable Hilbert space R . This means, that an initial guess ρ_0 is updated iteratively using a suitable update candidate to obtain a new ρ_i that further reduces the objective functional. Since the derivative $\frac{d\hat{J}}{d\rho} : R \rightarrow \mathbb{R}$ is a functional and therefore an element of the dual space R^* of R , it is not a suitable candidate to be used as update during optimization. What is needed is an element of the primal space similar to ρ , the gradient $\nabla_{\rho}\hat{J} \in R$ ¹⁴.

Since within a finite element setting all function spaces are Hilbert spaces, the Riesz representation theorem can be used to obtain the gradient. The Riesz representation theorem states, that the dual space H^* of a Hilbert space H is isometric to itself. This means that the identification $H = H^*$ can be done and for a functional $I \in H^*$ applied to $u \in H$ a unique element w of the primal space H can be found such that

$$I(u) := (w, u)_H \quad \forall u \in H, \quad (16)$$

where $(\cdot, \cdot)_H$ is the inner product defined on H and $w \in H$ is the Riesz representer of I with respect to the inner product on H , also denoted as $w = \mathcal{R}_H(I)$. Note that, for the presented example the appropriate inner product is the L^2 inner product that arises naturally when the forward problem is brought into its weak form by multiplication with a suitable test function. The specific reference is therefore dropped from this point on an each inner product as well as the Riesz representers are assumed to be with respect to L^2 if not stated otherwise.

The gradient $\nabla_{\rho}\hat{J} \in R$ can now be identified as the Riesz representer of the derivative $\frac{d\hat{J}}{d\rho}$ ¹⁷, pg. 99)

$$\frac{d\hat{J}}{d\rho}(s) = (\nabla_{\rho}\hat{J}, s) = -\left(\mathcal{R}\left(\frac{\partial J}{\partial u}\right), \left(\frac{\partial F}{\partial u}\right)^{-1} \frac{\partial F}{\partial \rho} s\right) + \left(\mathcal{R}\left(\frac{\partial J}{\partial \rho}\right), s\right), \quad (17)$$

where also the Riesz representers of $\frac{\partial J}{\partial u}$ and $\frac{\partial J}{\partial \rho}$ have to be used. Using, that for an operator $A : H_1 \rightarrow H_2$ where H_1 and H_2 are Hilbert spaces, dual operator $A^* : H_2 \rightarrow H_1$ is defined as

$$(A^*(v), u) = (v, A(u)) \quad \forall u \in H_1, v \in H_2, \quad (18)$$

one can continue with

$$(\nabla_{\rho}\hat{J}, s) = -\left(\left(\frac{\partial F}{\partial u}\right)^* \left(\frac{\partial F}{\partial u}\right)^{-*} \mathcal{R}\left(\frac{\partial J}{\partial u}\right), s\right) + \left(\mathcal{R}\left(\frac{\partial J}{\partial \rho}\right), s\right) \quad (19)$$

and the gradient can be written as

$$\nabla_{\rho}\hat{J} = \left(\frac{\partial F}{\partial \rho}\right)^* \left(\frac{\partial F}{\partial u}\right)^{-*} \mathcal{R}\left(\frac{\partial J}{\partial u}\right) + \mathcal{R}\left(\frac{\partial J}{\partial \rho}\right). \quad (20)$$

The adjoint equation is now obtained by expressing the term $\left(\frac{\partial F}{\partial u}\right)^{-*} \mathcal{R}\left(\frac{\partial J}{\partial u}\right)$ as a new variable λ and rearranging to get

$$\left(\frac{\partial F}{\partial u}\right)^* \lambda = \mathcal{R}\left(\frac{\partial J}{\partial u}\right). \quad (21)$$

This is the adjoint equation that can be used to calculate the adjoint variable λ and obtain the gradient via

$$\nabla_{\rho} \hat{J} = \left(\frac{\partial F}{\partial \rho}\right)^* \lambda + \mathcal{R}\left(\frac{\partial J}{\partial \rho}\right). \quad (22)$$

Obtaining the dual operators. In order to calculate the gradient $\nabla_{\rho} \hat{J}$, the continuous dual operators $\left(\frac{\partial F}{\partial u}\right)^*$ and $\left(\frac{\partial F}{\partial \rho}\right)^*$ are needed. Starting with $\left(\frac{\partial F}{\partial u}\right)^*$, first the operator $\frac{\partial F}{\partial u}$ has to be calculated using the definition of the Fréchet derivative

$$\frac{\partial F}{\partial u} h = \lim_{\epsilon \rightarrow 0^+} \frac{F(u(\rho) + \epsilon h, \rho) - F(u(\rho), \rho)}{\epsilon} = \nabla \cdot [(1 + \rho^p \chi) \nabla h] \quad (23)$$

Using the definition of dual operators in Hilbert space (18) to calculate $\left(\frac{\partial F}{\partial u}\right)^*$ gives

$$\begin{aligned} \left(\frac{\partial F}{\partial u} w, z\right) &= \int_{\mathbb{R}^3} \nabla \cdot [(1 + \rho^p \chi) \nabla w] z \, dV \\ &= - \int_{\mathbb{R}^3} [(1 + \rho^p \chi) \nabla w] \nabla z \, dV = \int_{\mathbb{R}^3} w \nabla \cdot [(1 + \rho^p \chi) \nabla z] \, dV = \left(w, \left(\frac{\partial F}{\partial u}\right)^* z\right), \end{aligned} \quad (24)$$

where partial integration is applied twice and the boundary integrals vanish if the same boundary condition as valid for the scalar potential u is also demanded from w and z ¹⁸.

The adjoint equation becomes

$$\nabla \cdot [(1 + \rho^p \chi) \nabla \lambda] = \mathcal{R}\left(\frac{\partial J}{\partial u}\right) \quad (25)$$

with boundary condition

$$\lambda(r) \rightarrow \mathcal{O}\left(\frac{1}{|r|}\right) \quad \text{for } |r| \rightarrow \infty. \quad (26)$$

Note that this equation has the same form as the forward problem F itself. This is a consequence of the self-adjointness of the operator $\frac{\partial F}{\partial u}$ that is given since the above calculation shows that $\frac{\partial F}{\partial u} = \left(\frac{\partial F}{\partial u}\right)^*$. Therefore, λ also has to be twice differentiable and the adjoint equation can be solved by the same method as described in “[Forward problem](#)” section, as will be seen below in “[Solving the adjoint equation](#)” section.

The operator $\left(\frac{\partial F}{\partial \rho}\right)^*$ is obtained by again first deriving the operator $\frac{\partial F}{\partial \rho}$ using the Fréchet derivative

$$\begin{aligned} \frac{\partial F}{\partial \rho} r &= \lim_{\epsilon \rightarrow 0^+} \frac{F(u(\rho), \rho + \epsilon r) - F(u(\rho), \rho)}{\epsilon} \\ &= \nabla \cdot (p \rho^{p-1} r [\chi (\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]). \end{aligned} \quad (27)$$

Using Eq. (18) gives

$$\begin{aligned} \left(\frac{\partial F}{\partial \rho} r, s\right) &= \int_{\mathbb{R}^3} \nabla \cdot (p \rho^{p-1} r [\chi (\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]) s \, dV \\ &= - \int_{\mathbb{R}^3} (p \rho^{p-1} r [\chi (\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]) \cdot \nabla s \, dV \\ &= - \int_{\mathbb{R}^3} r (p \rho^{p-1} [\chi (\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]) \cdot \nabla s \, dV = \left(r, \left(\frac{\partial F}{\partial \rho}\right)^* s\right). \end{aligned} \quad (28)$$

Here again partial integration was applied and the boundary integral vanished for s satisfying the boundary condition (26), which is fulfilled since in the expression for the gradient $\left(\frac{\partial F}{\partial \rho}\right)^*$ is applied to λ that also satisfies this boundary condition.

Finally, the gradient can be written as

$$\nabla_{\rho} \hat{J} = (p \rho^{p-1} [\chi (\nabla u - \mathbf{H}_{\text{ext}}) - \mathbf{M}_r]) \cdot \nabla \lambda + \mathcal{R}\left(\frac{\partial J}{\partial \rho}\right) \in R. \quad (29)$$

Solving the adjoint equation. In order to solve the adjoint Eq. (25) we first have a closer look at it. On the right hand side is the partial derivative of the objective functional J with respect to the scalar potential u . However, we will not directly impose conditions on the potential u but on the magnetic field $\mathbf{H}$ that is defined as $\mathbf{H} = -\nabla u$. Consider the weak form of the adjoint equation

$$\int_{\mathbb{R}^3} \nabla \cdot [(1 + \rho^p \chi) \nabla \lambda] v \, dV = \int_{\mathbb{R}^3} \mathcal{R} \left(\frac{\partial J}{\partial u} \right) v \, dV \quad (30)$$

where v is a suitable test function. Performing partial integration on the left hand side and inserting $\frac{\partial J}{\partial u} v = \frac{\partial J}{\partial \mathbf{H}} \frac{\partial \mathbf{H}}{\partial u} v = -\frac{\partial J}{\partial \mathbf{H}} \nabla v$ on the right hand side we get

$$-\int_{\mathbb{R}^3} (1 + \rho^p \chi) \nabla \lambda \cdot \nabla v \, dV = -\int_{\mathbb{R}^3} \mathcal{R} \left(\frac{\partial J}{\partial \mathbf{H}} \right) \nabla v \, dV, \quad (31)$$

where the boundary term vanishes due to the boundary condition on λ . Assuming that the objective functional J is defined locally on an area Ω_{opt} , this is the weak form of

$$\nabla \cdot \left((1 + \rho^p \chi) \nabla \lambda - \mathcal{R} \left(\frac{\partial J}{\partial \mathbf{H}} \right) \right) = 0, \quad (32)$$

that is identical to the forward problem in Eq. (7) with λ instead of u , $\mathcal{R} \left(\frac{\partial J}{\partial \mathbf{H}} \right)$ replacing $\rho^p \mathbf{M}_r$ and $\mathbf{H}_{\text{ext}} = 0$. We can therefore use the same algorithm to solve the adjoint equation as is used to solve the forward problem. Note, that it is perfectly fine to assume the same jump conditions on λ as on u since they are contained in the weak form of the forward problem anyways as shown in¹⁹.

Obtaining the Riesz representers. To obtain the Riesz representers $\mathcal{R} \left(\frac{\partial J}{\partial \mathbf{H}} \right)$ and $\mathcal{R} \left(\frac{\partial J}{\partial \rho} \right)$ appearing inside the adjoint equation and the expression for the gradient, the coefficient vectors $\mathbf{J}_H$ and $\mathbf{J}_\rho$ belonging to the functionals $\frac{\partial J}{\partial \mathbf{H}}$ and $\frac{\partial J}{\partial \rho}$ respectively, are calculated using FEniCS¹¹. These coefficient vectors can be used to calculate the action of the functionals onto an element of the primal space r ¹⁴. For the functional $\frac{\partial J}{\partial \rho}$ this reads

$$\frac{\partial J}{\partial \rho}(r) = (\mathbf{J}_\rho, \mathbf{r})_{\ell^2} = \mathbf{J}_\rho^T \mathbf{r}, \quad (33)$$

where $(\cdot, \cdot)_{\ell^2}$ represents the euclidean inner product. On the other hand, using the Riesz representer $\mathcal{R} \left(\frac{\partial J}{\partial \rho} \right)$ the action can be written as

$$\begin{aligned} \frac{\partial J}{\partial \rho}(r) &= \left(\mathcal{R} \left(\frac{\partial J}{\partial \rho} \right), r \right)_{L^2} = \int_{\Omega} \sum_i \mathcal{R}_i \left(\frac{\partial J}{\partial \rho} \right) \psi_i(\mathbf{x}) \sum_j r_j \psi_j(\mathbf{x}) \, d\mathbf{x} \\ &= \sum_i \mathcal{R}_i \left(\frac{\partial J}{\partial \rho} \right) \sum_j r_j \int_{\Omega} \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) \, d\mathbf{x} = \mathbf{R}_{L^2}^T \left(\frac{\partial J}{\partial \rho} \right) \mathbf{M} \mathbf{r} \end{aligned} \quad (34)$$

with $\psi_i(\mathbf{x})$ being the discrete basis functions and $\mathbf{M} = \int_{\Omega} \psi_i \psi_j \, dV$ being the mass matrix. Since the two equations must deliver the same result and furthermore $\mathbf{M}^T = \mathbf{M}$, the Riesz representer with respect to the L^2 inner product can be obtained from the coefficient vector as

$$\mathcal{R}_{L^2} \left(\frac{\partial J}{\partial \rho} \right) = \mathbf{M}^{-1} \mathbf{J}_\rho. \quad (35)$$

The Riesz representer $\mathcal{R} \left(\frac{\partial J}{\partial \mathbf{H}} \right)$ is obtained in the same way.

Received: 17 June 2021; Accepted: 14 December 2021

Published online: 21 January 2022

References

- Kang, J. & Wang, S. Shape optimization of BLDC motor using 3-D finite element method. *IEEE Trans. Magn.* **36**(4), 1119–1123. <https://doi.org/10.1109/20.877637> (2000).
- Campelo, F. Rmírez, J. A. and Igarashi, H. A survey of topology optimization in electromagnetics: Considerations and current trends (2010).
- Okamoto, Y., Akiyama, K. & Takahashi, N. 3-D topology optimization of singlepole- type head by using design sensitivity analysis. *IEEE Trans. Magn.* **42**(4), 1087–1090. <https://doi.org/10.1109/TMAG.2006.871422> (2006).
- Wang, Semyung *et al.* Topology optimization of electromagnetic systems considering magnetization direction. *IEEE Trans. Magn.* **41**(5), 1808–1811. <https://doi.org/10.1109/TMAG.2005.846480> (2005).
- Wang, S. & Kang, J. Topology optimization of nonlinear magnetostatics. *IEEE Trans. Magn.* **38**(2), 1029–1032. <https://doi.org/10.1109/20.996264> (2002).
- Seebacher, P. *et al.* Optimization of an Electromagnetic Actuator using Topology- Optimization. In: 2019 22nd International Conference on the Computation of Electromagnetic Fields (COMPUMAG). 2019 22nd International Conference on the Computation of Electromagnetic Fields (COMPUMAG). Paris, France: IEEE, July 2019, pp. 1–4. <https://doi.org/10.1109/COMPUMAG45669.2019.9032821>. <https://ieeexplore.ieee.org/document/9032821/> (visited on 10/25/2021).
- Bruckner, Florian *et al.* 3D FEM-BEM-coupling method to solve magnetostatic Maxwell equations. *J. Magn. Magn. Mater.* **324**(10), 1862–1866. <https://doi.org/10.1016/j.jmmm.2012.01.016> (2012).
- Huber, C. *et al.* Topology optimized and 3D printed polymer-bonded permanent magnets for a predefined external field. *J. Appl. Phys.* **122**(5), 053904. <https://doi.org/10.1063/1.4997441> (2017).

9. Abert, Claas *et al.* A fast finite-difference algorithm for topology optimization of permanent magnets. *J. Appl. Phys.* **122**(11), 113904. <https://doi.org/10.1063/1.4998532> (2017).
10. Abert, C. *et al.* magnum.fe: A micromagnetic finite-element simulation code based on FEniCS. *J. Magn. Magn. Mater.* **345**, 29–35. <https://doi.org/10.1016/j.jmmm.2013.05.051> (2013).
11. Logg, A. *et al.* (eds) *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book. Lecture Notes in Computational Science and Engineering* 84. OCLC: 794507863 723 (Springer, 2012).
12. Bendsoe, M. P. & Sigmund, O. *Topology Optimization* (Springer, 2004). <https://doi.org/10.1007/978-3-662-05086-6>.
13. Zhu, Ciyu *et al.* Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.* **23**(4), 550–560. <https://doi.org/10.1145/279232.279236> (1997).
14. Schwedes, T. *et al.* *Mesh Dependence in PDE-Constrained Optimisation* (Springer, 2017). <https://doi.org/10.1007/978-3-319-59483-5>.
15. Nordaas, M. & Funke, S. W. The Moola optimisation package (2016).
16. Guedes, A. *et al.* Improving magnetic field detection limits of spin valve sensors using magnetic flux guide concentrators. *IEEE Trans. Magn.* **43**(6), 2376–2378. <https://doi.org/10.1109/TMAG.2007.893119> (2007).
17. Hinze, M. (ed.) *Optimization with PDE Constraints. Mathematical modelling* 23. OCLC: 255277417 270 (Springer, 2009).
18. Giles, M. B. & Pierce, N. A. An Introduction to the Adjoint Approach to Design. *Flow Turbulence Combust.* **65**(3), 393–415. <https://doi.org/10.1023/A:1011430410075> (2000).
19. Bossavit, A. *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*. OCLC: 162128760 (Academic Press, 1998).

Acknowledgements

The financial support of the FWF project I 4917 is acknowledged.

Author contributions

G.W. and F.S. applied the adjoint approach deriving the necessary expressions for the gradient with consultation from D.S., C.A. and F.B., G.W., F.B. and C.A. developed the numerical simulation code and G.W. conducted the numerical experiments. The manuscript was written by G.W. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022