

Nonlinear filters in topology optimization: Existence of solutions and efficient implementation for minimal compliance problems

Linus Hägg^{*} and Eddie Wadbro[†]

Department of Computing Science, Umeå University, SE-901 87 Umeå, Sweden

Abstract

It is well known that material distribution topology optimization problems often are ill-posed if no restriction or regularization method is used. A drawback with the standard linear density filter is that the resulting designs have large areas of intermediate densities, so-called gray areas, especially when large filter radii are used. To produce final designs with less gray areas, several different methods have been proposed; for example, projecting the densities after the filtering or using a nonlinear filtering procedure. In a recent paper, we presented a framework that encompasses a vast majority of currently available density filters. In this paper, we show that all these nonlinear filters ensure existence of solutions to a continuous version of the minimal compliance problem. In addition, we provide a detailed description on how to efficiently compute sensitivities for the case when multiple of these nonlinear filters are applied in sequence. Finally, we present a numerical experiment that illustrates that these cascaded nonlinear filters can be used to obtain independent size control of both void and material regions in a large-scale setting.

Keywords: Topology optimization, Regularization, Nonlinear filters, Existence of solutions, Large-scale problems

1 Introduction

Since the seminal paper by Bendsøe & Kikuchi [3] in 1988 regarding topology optimization of linearly elastic continuum structures, the field of topology optimization has been subject to intense research. Today, material distribution topology optimization is applied to a range of different disciplines, such as linear and nonlinear elasticity [9, 16, 18], acoustics [23, 8, 17], electromagnetics [11, 12, 15, 24], and fluid–structure interaction [2, 26]. A comprehensive account on topology optimization and its various applications can be found in the monograph by Bendsøe & Sigmund [4], as well as in the more recent reviews by Sigmund & Maute [21] and Deaton & Grandhi [10].

^{*}linush@cs.umu.se

[†]eddiew@cs.umu.se

In material distribution topology optimization, a material indicator function $\rho : \Omega_D \subset \mathbb{R}^d \rightarrow \{0, 1\}$ indicates the presence ($\rho = 1$) or absence ($\rho = 0$) of material within the design domain Ω_D [4]. To numerically solve the topology optimization problem, the domain Ω_D is typically discretized into n elements. The aim of the optimization is to determine element values $\rho_i \in \{0, 1\}$, $i \in \{1, \dots, n\}$, that is, to determine if the region corresponding to a given element contains material or not. The resulting nonlinear integer optimization problem is typically relaxed by allowing $\rho_i \in [0, 1]$, $i \in \{1, \dots, n\}$. This relaxation enables the use of gradient based optimization algorithms that are suitable for solving large-scale problems involving millions of design variables. In order to promote binary designs penalization techniques are used. However the optimized designs, resulting from the relaxed and penalized problem, are typically mesh-dependent. Several strategies have been proposed to resolve the issue of mesh-dependence; Borvall [5] presents a systematic investigation of several common techniques.

Amongst the most popular techniques to achieve mesh-independent designs is to use a filtering procedure. Filtering procedures are commonly categorized as either *sensitivity filtering* [19] or *density filtering* [6, 7], where the derivatives of the objective function are filtered or where the design variables are filtered, respectively. When using a density filtering procedure the design variables are no longer the physical “density”, that is, the coefficients that enter the governing equation. In a classic paper, Bourdin [6] established existence of solutions to a continuous version of the linearly filtered minimal compliance problem. Since the introduction of the linear density filter by Bruns and Tortorelli [7], a whole range of *nonlinear* filters has been presented [14, 20, 13, 22]. We have recently introduced the class of generalized *fW*-mean filters that include the vast majority of filters used in topology optimization [25], and provide a common framework for analyzing and evaluating various filters.

The remainder of this paper is organized as follows. Section 2 provides a short discussion about which properties that are desired for the filtering procedure as well as a short summary on the *fW*-mean filters. In Section 3, we prove that there exists a solution to a continuous version of the *fW*-mean filtered minimum compliance topology optimization problem. Section 4 discusses various aspects on the fast evaluation of filtered densities and their sensitivities. Finally, Section 5 presents numerical experiments that illustrate that by using a fast algorithm to perform the nonlinear filtering, it is possible to cascade filters to achieve independent size control of structural and void regions for large-scale problems.

2 Background

2.1 Typical requirements on filters and their implications

The discussion below treats the discretized case, where the design domain Ω_D is partitioned into n elements and the aim of the optimization is to determine the design vector $\boldsymbol{\rho} = (\rho_1, \dots, \rho_n)^T \in [0, 1]^n$. A *general* filter is any function $\mathbf{F} : [0, 1]^n \rightarrow [0, 1]^n$, and the filtered design is $\mathbf{F}(\boldsymbol{\rho})$. Below, we discuss typical requirements on such filters.

The first requirement is already included in the definition of \mathbf{F} , namely that we require that the range must be conforming, that is

$$\mathbf{F}(\boldsymbol{\rho}) \in [0, 1]^n. \quad (1)$$

In addition, we also require that the function \mathbf{F} is coordinate- and component-wise non-decreasing; that is, for any i, j and any $\delta \geq 0$,

$$F_i(\boldsymbol{\rho} + \delta \mathbf{e}_j) \geq F_i(\boldsymbol{\rho}), \quad (2)$$

where \mathbf{e}_j denotes the j th basis vector of \mathbb{R}^n . We note that (1) and (2) imply that

$$F_i(\mathbf{0}_n) \leq F_i(\boldsymbol{\rho}) \leq F_i(\mathbf{1}_n), \quad (3)$$

where, $\mathbf{0}_n = (0, \dots, 0)^T \in \mathbb{R}^n$ and $\mathbf{1}_n = (1, \dots, 1)^T \in \mathbb{R}^n$. Expression (3) shows that if we want each element in the filtered design to be able to attain the values 0 or 1 then we must require that

$$\begin{aligned} \mathbf{F}(\mathbf{0}_n) &= \mathbf{0}_n, \\ \mathbf{F}(\mathbf{1}_n) &= \mathbf{1}_n. \end{aligned} \quad (4)$$

It is natural to require that the filtered density in element i is strictly increasing with the density in that element. A weaker assumption would be to require that for each $\boldsymbol{\rho} \in [0, 1]^n$ there exists an i and a j such that $F_i(\boldsymbol{\rho})$ is strictly increasing in ρ_j in the vicinity of $\boldsymbol{\rho}$. That is there exists $\epsilon > 0$ such that

$$F_i(\boldsymbol{\rho} + \delta \mathbf{e}_j) > F_i(\boldsymbol{\rho}) \text{ for all } 0 < \delta < \epsilon. \quad (5)$$

If we want to use gradient based optimization algorithms we would also need to require that \mathbf{F} is differentiable. In this case requirement (2) translates to

$$\frac{\partial F_i}{\partial \rho_j} \geq 0. \quad (6)$$

Requirement (5) is often replaced by the more restrictive condition that for each $\boldsymbol{\rho}$, there exists an i and a j such that

$$\frac{\partial F_i}{\partial \rho_j}(\boldsymbol{\rho}) > 0. \quad (7)$$

Another property that often is mentioned is *volume preservation*, which mathematically can be expressed by

$$\mathbf{1}_n^T \mathbf{F}(\boldsymbol{\rho}) = \mathbf{1}_n^T \boldsymbol{\rho} \text{ for all } \boldsymbol{\rho} \in [0, 1]^n. \quad (8)$$

The obvious benefit of using a volume preserving filter is that the volume constraint can be left unaltered.

For a linear filter $\mathbf{F}(\boldsymbol{\rho}) = \mathbf{A}\boldsymbol{\rho}$ where $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$ it can be shown that conditions (1), (2) and (4) are equivalent to

$$a_{ij} \geq 0 \text{ for all } i, j \text{ and } \mathbf{A}\mathbf{1}_n = \mathbf{1}_n. \quad (9)$$

Furthermore we find that volume preservation is equivalent to

$$\mathbf{1}_n^T \mathbf{A} = \mathbf{1}_n^T. \quad (10)$$

Thus, a linear filter is both volume preserving and has a conforming range if and only if \mathbf{A} is a so-called doubly stochastic matrix.

2.2 fW -mean filters

Next, we present a short summary on fW -mean filters [25]. For a given smooth and invertible function $f : [0, 1] \rightarrow \mathbb{R}$ with nonzero derivative the fW -mean filter of a vector $\boldsymbol{\rho} \in [0, 1]^n$ is defined as

$$\mathbf{F}(\boldsymbol{\rho}) = \mathbf{f}^{-1}(\mathbf{W}\mathbf{f}(\boldsymbol{\rho})), \quad (11)$$

where $\mathbf{f}(\boldsymbol{\rho}) = (f(\rho_1), f(\rho_2), \dots, f(\rho_n))^T \in \mathbb{R}^n$, and $\mathbf{W} = [w_{ij}] \in \mathbb{R}^{n \times n}$ is a weight matrix with non-negative entries such that $\mathbf{W}\mathbf{1}_n = \mathbf{1}_n$.

We note that the neighborhood $\mathcal{N}_i \subset \{1, \dots, n\}$ of element i is implicitly defined by the weight matrix \mathbf{W} ,

$$w_{ij} > 0 \text{ if and only if } j \in \mathcal{N}_i. \quad (12)$$

In topology optimization, it is common to define a neighborhood shape $\mathcal{N} \subset \mathbb{R}^d$ centered at the origin and use \mathcal{N} to define the neighborhoods

$$\mathcal{N}_i = \{j : x_j - x_i \in \mathcal{N}\}, i \in \{1, \dots, n\}, \quad (13)$$

where $x_i \in \mathbb{R}^d$, $i \in \{1, \dots, n\}$ are the element centroids. When equal weights are used within neighborhoods, $\mathbf{W} = \mathbf{D}^{-1}\mathbf{G}$, where $\mathbf{D} = \text{diag}(|\mathcal{N}_1|, \dots, |\mathcal{N}_n|)^T$ and \mathbf{G} is the neighborhood matrix with entries $g_{ij} = 1$ if and only if $j \in \mathcal{N}_i$ and $g_{ij} = 0$ otherwise.

The fW -mean filter maps a vector with equal entries to itself, that is, if $c \in [0, 1]$ then $\mathbf{F}(c\mathbf{1}_n) = c\mathbf{1}_n$. However the fW -mean filter is in general *not* volume preserving; that is, there exists a $\boldsymbol{\rho} \in [0, 1]^n$ such that

$$\mathbf{1}_n^T \mathbf{F}(\boldsymbol{\rho}) \neq \mathbf{1}_n^T \boldsymbol{\rho}. \quad (14)$$

The ij entry of the Jacobian $\nabla \mathbf{F}$ of the fW -mean filter is given by

$$\frac{\partial F_i}{\partial \rho_j} = w_{ij} \frac{f'(\rho_j)}{f'(F_i(\boldsymbol{\rho}))} \geq 0, \quad (15)$$

where the inequality is strict as long as $w_{ij} > 0$, that is j is in the neighborhood of i . Note that, since i and j are free indices, summation is *not* implied in expression (15).

The fW -mean filter framework contains many filter types, such as the harmonic erode filter introduced by Svanberg & Svård [22], but not all: one example is the Heaviside filter introduced by Guest et al. [14]. However, the non- fW -mean filters can often be fitted into the *generalized fW -mean filter framework* [25], in which the function f^{-1} , used in definition (11), is replaced by a smooth function $g : f([0, 1]) \rightarrow [0, 1]$. That is, the generalized fW -mean filters are of the form

$$\tilde{\mathbf{F}}(\boldsymbol{\rho}) = \mathbf{g}(\mathbf{W}\mathbf{f}(\boldsymbol{\rho})). \quad (16)$$

3 Existence of solutions for the fW -mean filtered continuous minimal compliance problem

In this section, we show that there exists a solution to the fW -mean filtered version of a penalized continuous minimal compliance problem.

Let $\Omega \subset \mathbb{R}^d$ be a bounded and connected domain in which we want to place our structure. Assume that the boundary $\partial\Omega$ is Lipschitz and that the structure is fixed at an open boundary portion $\Gamma_D \subset \partial\Omega$, with $|\Gamma_D| > 0$. The kinematically admissible displacements are

$$\mathcal{U} = \{u \in H^1(\Omega)^d \mid u|_{\Gamma_D} \equiv 0\}. \quad (17)$$

The displacement of the structure is the solution to the following variational problem:

$$\text{Find } u \in \mathcal{U} \text{ such that } a(\rho; u, v) = \ell(v) \quad \forall v \in \mathcal{U}. \quad (18)$$

The energy bilinear form a and the load linear form ℓ are defined as

$$a(\rho; u, v) = \int_{\Omega} \tilde{\rho}(\rho) E \epsilon(u) : \epsilon(v), \quad (19)$$

$$\ell(v) = \int_{\Omega} b \cdot v + \int_{\Gamma_L} t \cdot v, \quad (20)$$

where $b \in L^2(\Omega)^d$ and $t \in L^2(\Gamma_L)^d$ represent the internal force in Ω and surface traction densities on the boundary portion $\Gamma_L = \partial\Omega \setminus \bar{\Gamma}_D$, respectively, $\epsilon(u) = (\nabla u + \nabla u^T)/2$ is the strain tensor (or the symmetrized gradient) of u , the colon “:” denotes the scalar product of the two matrices, E is a constant forth-order elasticity tensor, and $\tilde{\rho}(\rho)$ is the physical density.

Remark 3.1. Throughout the article we do not explicitly specify the measure symbol (such as $d\Omega$, for instance) in the integrals, whenever there is no risk for confusion. The type of measure will be clear from the domain of integration.

We define the physical density as

$$\tilde{\rho}(\rho) = \underline{\rho} + (1 - \underline{\rho})P(F(\rho)), \quad (21)$$

where $\underline{\rho} > 0$, $F(\rho)$ is a continuous version of the fW -mean filter, and $P : [0, 1] \rightarrow [0, 1]$ is a smooth and invertible penalty function. The above formulation includes the case when the problem is penalized using SIMP [4], that is, to use $P(x) = x^p$ in (21) for some $p > 1$. The addition of a minimal physical density $\underline{\rho} > 0$ ensures that the bilinear form $a(\cdot; \cdot, \cdot)$ is coercive, that is there exists a constant $C > 0$ such that

$$a(\rho; u, u) \geq C \|u\|_{H^1(\Omega)^d}^2. \quad (22)$$

The continuous fW -mean filtered density is, for $x \in \Omega$, given by

$$(F(\rho))(x) = f^{-1} \left(\frac{1}{|\mathcal{N}_x|} \int_{\mathcal{N}_x} (f \circ \rho)(y) dy \right), \quad (23)$$

where f is a smooth and invertible function $f : [0, 1] \rightarrow [f_{\min}, f_{\max}] \subset \mathbb{R}$, \mathcal{N}_x is the neighborhood of x , and $|\mathcal{N}_x| > 0$ is the measure (area or volume) of \mathcal{N}_x . We define the set of admissible designs $\mathcal{A} \subset L^\infty(\Omega)$ as

$$\mathcal{A} = \left\{ \rho \mid 0 \leq \rho \leq 1 \text{ almost everywhere on } \Omega \text{ and } \int_{\Omega} F(\rho) \leq V \right\}. \quad (24)$$

We remark that, to be able to apply the filter F on any design in \mathcal{A} , we use a continuous and bijective extension of f , in this paragraph denoted by $\hat{f} : \mathbb{R} \rightarrow \mathbb{R}$, when evaluating expression (23). Furthermore for any Lebesgue measurable function $\rho : \Omega \rightarrow \mathbb{R}$ the composition $\hat{f} \circ \rho$ is Lebesgue measurable. To simplify the notation below, whenever we write f it should be interpreted as a continuous and bijective extension of f .

Theorem 3.2. *If $|\mathcal{N}_x| > 0$ for all $x \in \Omega$, then there exists a solution to the following variation of the minimal compliance problem*

$$\inf_{u \in \mathcal{U}^*} \ell(u), \quad (25)$$

where

$$\mathcal{U}^* = \{u \in \mathcal{U} \mid \exists \rho \in \mathcal{A} \text{ such that } a(\rho; u, v) = \ell(v) \quad \forall v \in \mathcal{U}\}. \quad (26)$$

Proof. Let (u_m) , $u_m \in \mathcal{U}^*$ for all $m \in \mathbb{N}$, be a minimizing sequence for ℓ ; without loss of generality, we stipulate that $(\ell(u_m))$ is non-increasing. By the definition of \mathcal{U}^* , there exists a sequence of designs (ρ_m) such that, for each $m \in \mathbb{N}$, $a(\rho_m; u_m, v) = \ell(v)$ for all $v \in \mathcal{U}$. Since the bilinear form (19) is coercive, we have from (22) that

$$C \|u_m\|_{H^1(\Omega)^d}^2 \leq a(\rho_m; u_m, u_m) = \ell(u_m) \leq \ell(u_1). \quad (27)$$

That is, (u_m) is uniformly bounded in $H^1(\Omega)^d$ and thus there exists an element $u^* \in \mathcal{U}$ such that u_m converges weakly to u^* in $H^1(\Omega)^d$ as $m \rightarrow \infty$.

For each $m \in \mathbb{N}$, we define $\tau_m = f \circ \rho_m$; by construction, we have that $f_{\min} \leq \tau_m \leq f_{\max}$ almost everywhere in Ω . Since (τ_m) is bounded in $L^\infty(\Omega)$, we can according to the sequential Banach–Alaoglu theorem find a subsequence, still denoted (τ_m) and a limit element $\tau^* \in L^\infty(\Omega)$, so that τ_m converges weak* to τ^* in $L^\infty(\Omega)$ as $m \rightarrow \infty$. As a direct consequence of the weak star convergence we have that for all $x \in \Omega$

$$\begin{aligned} \frac{1}{|\mathcal{N}_x|} \int_{\mathcal{N}_x} \tau_m(y) \, dy &= \frac{1}{|\mathcal{N}_x|} \int_{\Omega} \tau_m(y) \mathbb{1}_{\mathcal{N}_x}(y) \, dy \xrightarrow{m \rightarrow \infty} \\ &= \frac{1}{|\mathcal{N}_x|} \int_{\Omega} \tau^*(y) \mathbb{1}_{\mathcal{N}_x}(y) \, dy = \frac{1}{|\mathcal{N}_x|} \int_{\mathcal{N}_x} \tau^*(y) \, dy, \end{aligned} \quad (28)$$

where $\mathbb{1}_{\mathcal{N}_x} \in L^1(\Omega)$ is the characteristic function of \mathcal{N}_x .

The sequential Banach–Alaoglu theorem also guarantees that $f_{\min} \leq \tau^* \leq f_{\max}$ almost everywhere in Ω . We can thus define $\rho^* = f^{-1} \circ \tau^*$ and by construction $0 \leq \rho^* \leq 1$ almost everywhere in Ω . Since, f^{-1} is continuous we have that

$$\begin{aligned} F(\rho_m)(x) &= f^{-1} \left(\frac{1}{|\mathcal{N}_x|} \int_{\mathcal{N}_x} \tau_m(y) \, dy \right) \xrightarrow{m \rightarrow \infty} \\ &= f^{-1} \left(\frac{1}{|\mathcal{N}_x|} \int_{\mathcal{N}_x} \tau^*(y) \, dy \right) = F(\rho^*)(x), \end{aligned} \quad (29)$$

that is, the filtered design converges pointwise. Moreover for all $x \in \Omega$, we have that $0 \leq F(\rho^*)(x) \leq 1$ and $0 \leq F(\rho_m)(x) \leq 1$ for all m . Thus, by Lebesgue's dominated

convergence theorem,

$$\int_{\Omega} F(\rho^*) = \lim_{m \rightarrow \infty} \int_{\Omega} F(\rho_m) \leq V, \quad (30)$$

or in words, the design obtained by filtering ρ^* satisfies the volume constraint, so $\rho^* \in \mathcal{A}$. We note that the pointwise convergence of the filtered design implies that the physical design converges pointwise, and that for all $x \in \Omega$, $\underline{\rho} \leq \tilde{\rho}(\rho^*)(x) \leq 1$ and $\underline{\rho} \leq \tilde{\rho}(\rho_m)(x) \leq 1$ for all m .

Let v be an arbitrary function in \mathcal{U} , then

$$\begin{aligned} \underbrace{a(\rho_m; u_m, v)}_{\ell(v)} - a(\rho^*; u^*, v) &= \int_{\Omega} (\tilde{\rho}(\rho_m) - \tilde{\rho}(\rho^*)) E \epsilon(u_m) : \epsilon(v) \\ &+ \int_{\Omega} \tilde{\rho}(\rho^*) E (\epsilon(u_m) - \epsilon(u^*)) : \epsilon(v). \end{aligned} \quad (31)$$

We have that $\tilde{\rho}(\rho^*) \epsilon(v) \in L^2(\Omega)^{d \times d}$ and since u_m converges weakly to u^* in $H^1(\Omega)^d$ as $m \rightarrow \infty$, we have that $\epsilon(u_m)$ converges weakly to $\epsilon(u^*)$ in $L^2(\Omega)^{d \times d}$ as $m \rightarrow \infty$. Thus, the second term on the right hand side of expression (31) tends to 0 as $m \rightarrow \infty$. The absolute value of the first term on the right hand side of expression (31) is bounded by

$$\sum_{i,j,k,l} \int_{\Omega} |(\tilde{\rho}(\rho_m) - \tilde{\rho}(\rho^*)) E_{ijkl} \epsilon_{ij}(u_m) \epsilon_{kl}(v)|. \quad (32)$$

Let A_{ijkl} denote an arbitrary term in expression (32). We note that for a fixed $v \in H^1(\Omega)^d$, we have that $L^2(\Omega) \ni |\tilde{\rho}(\rho_m) - \tilde{\rho}(\rho^*)| |\epsilon_{kl}(v)| \rightarrow 0$ almost everywhere on Ω when $m \rightarrow \infty$ since $|\epsilon_{kl}(v)|$ is finite almost everywhere on Ω . Moreover, $|\tilde{\rho}(\rho_m) - \tilde{\rho}(\rho^*)| |\epsilon_{kl}(v)| \leq |\epsilon_{kl}(v)|$. Hence, by using Cauchy–Schwartz’s inequality and Lebesgue’s dominated convergence theorem, we find that

$$A_{ijkl}^2 \leq |E_{ijkl}|^2 \|\epsilon_{ij}(u_m)\|_{L^2(\Omega)}^2 \int_{\Omega} |\tilde{\rho}(\rho_m) - \tilde{\rho}(\rho^*)|^2 |\epsilon_{kl}(v)|^2 \xrightarrow{m \rightarrow \infty} 0. \quad (33)$$

Thus, $a(\rho^*; u^*, v) = \ell(v) \quad \forall v \in \mathcal{U}$ so $u^* = \mathcal{U}^*$. Moreover, because u_m converges weakly to u^* in $H^1(\Omega)$ as $m \rightarrow \infty$ and ℓ is a bounded linear functional on $H^1(\Omega)$, we have that $\ell(u_m) \rightarrow \ell(u^*)$ as $m \rightarrow \infty$. Since u_m is a minimizing sequence for $\ell(\cdot)$, we have that

$$\ell(u^*) = \inf_{v \in \mathcal{U}^*} \ell(v), \quad (34)$$

that is, we have existence of a minimizer to the nonlinearly filtered minimal compliance problem. \square

Remark 3.3. We note that all steps in the above proof also holds true if we would replace the function f^{-1} in definition (24) of the continuous version of the fW -mean filter by another smooth function $g : [f_{\min}, f_{\max}] \rightarrow [0, 1]$. In particular, this holds if we replace f^{-1} by a projected version $h \circ f^{-1}$ provided that $h : [0, 1] \rightarrow [0, 1]$ is smooth. That is the proof also holds for the generalized fW -mean filters.

Remark 3.4. The proof also holds in the case with normalized but non-uniform weights within the neighborhoods. The only change required is to replace $\mathbb{1}_{\mathcal{N}_x}(y)/|\mathcal{N}_x|$ in equation (28) by an $L^1(\Omega)$ function that describes the non-uniform weights.

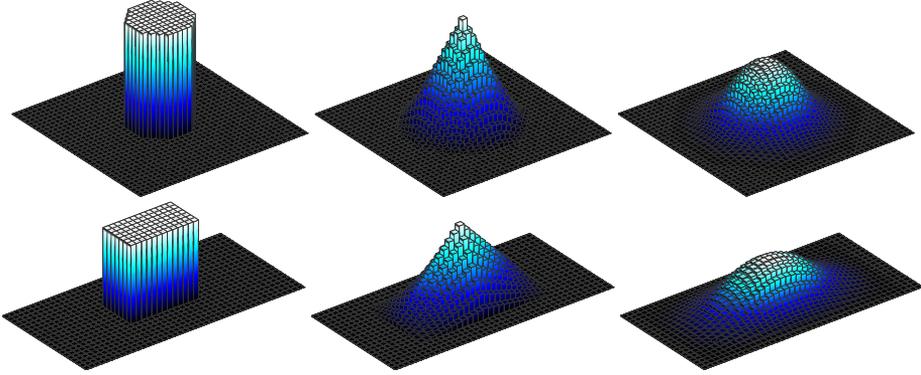


Figure 1: Illustration of the first three powers (from left to right) of the weight matrix, $\mathbf{W} = \mathbf{D}^{-1}\mathbf{G}$, \mathbf{W}^2 , \mathbf{W}^3 , for an octagonal and a rectangular neighborhood (top and bottom row, respectively).

4 Aspects of fast evaluation of filtered densities and sensitivities

4.1 On the computational complexity

In our previous paper [25], we showed that if the computational domain is discretized into n elements in a regular grid, the neighborhood shape is a polytope, and equal weighting within each neighborhood is used, then the fW -mean filter can be applied with computational complexity $O(n)$. Moreover the computational complexity is essentially independent of the size of the neighborhoods.

Non-equal weighting within neighborhoods can be achieved by sequentially applying the same equally weighted fW -mean filter twice (or more):

$$\begin{aligned} \mathbf{F}(\mathbf{F}(\boldsymbol{\rho})) &= \mathbf{f}^{-1}(\mathbf{W}^2\mathbf{f}(\boldsymbol{\rho})), \\ \mathbf{W} &= \mathbf{D}^{-1}\mathbf{G}. \end{aligned} \tag{35}$$

If the neighborhood shape is a *convex* polytope $\mathcal{P} \subset \mathbb{R}^d$ then it can be shown that a neighborhood shape corresponding to \mathbf{W}^2 is $2\mathcal{P}$. Figure 1 illustrates the weights corresponding to the first three powers of $\mathbf{W} = \mathbf{D}^{-1}\mathbf{G}$, from left to right \mathbf{W} , \mathbf{W}^2 , and \mathbf{W}^3 , for an octagonal (top row) and a rectangular neighborhood (bottom row).

We remark that, in general, the computational cost of the filter application is dominated by the cost of evaluating the sums $\mathbf{G}\mathbf{f}$. Moreover, the computational effort required to evaluate these sums grows with the complexity of the neighborhood polytope. Thus, if one wishes to apply a filter with weights that decay with the distance from the neighborhood, particularly in three space dimensions, one could save a great portion of the computational time by selecting a simple neighborhood. For example, the computational complexity for filtering over a box shaped neighborhood is approximately 10 times lower than the complexity for filtering over the significantly more complex rhombicuboctahedron (a polytope with 26 faces—twelve rectangular, six square, and eight triangular faces) neighborhood [25].

On the other hand, if one wishes to use one of the nonlinear filters that are designed to mimic min or max operators over the neighborhood, then the neighborhood shape

is important. In this latter case one cannot use the weighted version with a square or box shaped neighborhood to approximate a circular or spherical neighborhood, since the nonlinearity of the filtering process will essentially pick out the maximum/minimum of the elements within the support of the filter. However, by using neighborhoods of different but still relatively simple shapes, for example two square shaped neighborhoods (with a relative rotation of 45 degrees), one can get a filtering procedure where the neighborhood around each element is an octagon. A similar approach can also be used for the three dimensional case, where one can use four cubic neighborhoods (the scaled unit cube plus three version rotated 45 degrees in the x_1x_2 , x_1x_3 , and x_2x_3 planes, respectively). The use of a sequence of filter applications using simple neighborhood shapes streamlines the implementation. Moreover, the resulting memory access pattern can be made very regular, which paves the way for future highly efficient parallel implementations.

4.2 Sensitivity evaluation

By examining expression (15), we see that in order to evaluate sensitivities, that is to compute $\mathbf{v}^T \nabla \mathbf{F}(\boldsymbol{\rho})$ for some vector $\mathbf{v} \in \mathbb{R}^n$, we need to carry out matrix multiplication by \mathbf{W}^T . In practice, \mathbf{v} is the gradient of the objective or a constraint function with respect to the filtered (physical) densities. In the special case of equal weighting within neighborhoods, multiplication by \mathbf{W}^T translates to multiplication by \mathbf{G}^T ; or expressed differently to perform summation over the *transposed neighborhoods* $\mathcal{N}_i^T = \{j : i \in \mathcal{N}_j\} = \{j : g_{ji} = 1\}$. If the neighborhoods are symmetric then $\mathbf{G}^T = \mathbf{G}$ and the same summation algorithm can be used for both filtering and sensitivity calculation, which facilitates the implementation.

Assume that the neighborhoods are defined by a neighborhood shape \mathcal{N} so that $\mathcal{N}_i = \{j : x_j - x_i \in \mathcal{N}\}$, where x_i and x_j denote the centroid of elements i and j , respectively. For each element $j \in \mathcal{N}_i^T$ we have that $i \in \mathcal{N}_j$ and by definition $x_i - x_j \in \mathcal{N}$, that is, there exists $y \in \mathcal{N}$ so that $x_i - x_j = y$ or equivalently $x_j - x_i = -y$. Since all steps above are bidirectional, we have that $\mathcal{N}_i^T = \{j : x_j - x_i \in -\mathcal{N}\}$; hence a neighborhood shape which defines the transposed neighborhoods is found by inversion of \mathcal{N} in the origin. Since \mathcal{P} and $-\mathcal{P}$ are essentially the same polytope, this means that implementing the fast summation algorithm over $-\mathcal{P}$ requires the same amount of work as implementing it over \mathcal{P} .

If we work with a Cartesian grid and the design variables are stored using a standard slice/fiberwise numbering (row- or column-wise in the two dimensional case), then we can use that $\mathbf{G}^T = \mathbf{P}\mathbf{G}\mathbf{P}$, where \mathbf{P} is the flip or exchange matrix. That is, \mathbf{P} is the matrix with ones along the anti-diagonal and zeros elsewhere.

4.3 Cascaded fW -mean filters

It has already been established that sequential application of filters is a means to arrive at filters with desirable properties, see for instance the open–close and close–open filters introduced by Sigmund [20].

```

1 oneVec = ones (nelem, 1) ;
2 for K = 1:N
3     Ni{K} = G{K}(oneVec) ;
4 end

```

Listing 1: Matlab code that computes the neighborhood sizes.

Assume that we are given $\{\mathbf{F}^{(K)}\}$, a family of fW -mean filters,

$$\mathbf{F}^{(K)}(\boldsymbol{\rho}) = \mathbf{f}_K^{-1}(\mathbf{W}^{(K)}\mathbf{f}_K(\boldsymbol{\rho})). \quad (36)$$

For $N \geq 1$, we define the cascaded filter function $\mathbf{C}^{(N)} : [0, 1]^n \rightarrow [0, 1]^n$ to be the composition

$$\mathbf{C}^{(N)} = \mathbf{F}^{(N)} \circ \mathbf{F}^{(N-1)} \circ \dots \circ \mathbf{F}^{(1)}. \quad (37)$$

We let $\boldsymbol{\rho}^{(0)} = \boldsymbol{\rho}$ and for each $N \geq 1$, we define

$$\boldsymbol{\rho}^{(N)} = \mathbf{F}^{(N)}(\boldsymbol{\rho}^{(N-1)}) = \dots = \mathbf{C}^{(N)}(\boldsymbol{\rho}). \quad (38)$$

We proceed to show how the sensitivity analysis for the cascaded filter can be performed in a way which will be suitable for fast evaluation. The setting is just as before; given a vector $\mathbf{v} \in \mathbb{R}^n$, we want to compute $\mathbf{v}^T \nabla \mathbf{C}^{(N)}(\boldsymbol{\rho})$. We let $\mathbf{v}^{(N)} = \mathbf{v}$, by evaluating $(\mathbf{v}^{(N)})^T \nabla \boldsymbol{\rho}^{(N)}$, using definition (38), and the chain rule, we find that

$$\sum_{i=1}^n v_i^{(N)} \frac{\partial}{\partial \rho_j} \rho_i^{(N)} \Big|_{\boldsymbol{\rho}} = \sum_{l=1}^n v_l^{(N-1)} \frac{\partial}{\partial \rho_j} \rho_l^{(N-1)} \Big|_{\boldsymbol{\rho}}, \quad (39)$$

where

$$v_l^{(N-1)} = \sum_{i=1}^n v_i^{(N)} \frac{\partial F_i^{(N)}}{\partial \rho_l} \Big|_{\boldsymbol{\rho}^{(N-1)}}. \quad (40)$$

The recursion given by expressions (39) and (40) also holds if we replace N by K for $K = N, N-1, \dots, 1$. Thus by using equation (15) we get

$$v_l^{(K-1)} = \sum_{i=1}^n v_i^{(K)} w_{il}^{(K)} \frac{f'_K(\rho_i^{(K-1)})}{f'_K(\rho_i^{(K)})} = f'_K(\rho_l^{(K-1)}) \sum_{i=1}^n g_{il}^{(K)} \frac{v_i^{(K)}}{|\mathcal{N}_i^{(K)}| f'_K(\rho_i^{(K)})}. \quad (41)$$

We remark that the rightmost sum above is a summation over transposed neighborhoods, that is multiplying a vector by $(\mathbf{G}^{(K)})^T$, the transpose of the neighborhood matrix. Hence, if we store the intermediate values $\boldsymbol{\rho}^{(K)}$, $K \in \{1, \dots, N\}$ when applying the filter, then we can modify the sensitivities by using $O(Nn)$ operations.

Remark 4.1. Existence of solutions to the continuous minimal compliance problem in the case when a cascade of fW -mean filters is applied can be proven by following the same reasoning as in the proof in Section 3.

In the following section, we present numerical experiments performed in Matlab using a modified version of the 2D multigrid-CG topology optimization code by Amir et al. [1].

<i>Field</i>	<i>Explanation</i>
N	Number of individual filters $N, K \in \{1, \dots, N\}$
f{K}	Function handle to compute $\mathbf{f}_K(\boldsymbol{\rho})$
g{K}	Function handle to compute $\mathbf{g}_K(\boldsymbol{\rho}) (= \mathbf{f}_K^{-1}(\boldsymbol{\rho}))$
df{K}	Function handle to compute $\mathbf{f}'_K(\boldsymbol{\rho})$
dg{K}	Function handle to compute $\mathbf{g}'_K(\boldsymbol{\rho})$
G{K}	Function handle to compute $\mathbf{G}^{(K)}\boldsymbol{\rho}$
GT{K}	Function handle to compute $(\mathbf{G}^{(K)})^T\boldsymbol{\rho}$

Table 1: The fields in the filterParam struct.

```

1  Dinvs{1} = G{1}(f{1}(rho)) ./ Ni{1};
2  for K = 2:N
3      Dinvs{K} = G{K}(f{K}(g{K-1}(Dinvs{K-1}))) ./ ...
4          Ni{K};
5  end
6  rhoPhys = g{N}(Dinvs{N});

```

Listing 2: Matlab code that filters the design variables by using a cascade of generalized fW -mean filters.

Below, we describe the major changes done to the code. First, we introduce a filter struct `filterParam` to hold all information needed to perform filtering and sensitivity calculation, see Table 1. In the following code excerpts we have suppressed the struct name `filterParam` to increase the readability, for instance instead of writing `filterParam.N` we simply write `N`. Listings 1–3 includes the new parts of code that needs to be added in order to use a filtering procedure composed of a cascade of generalized fW -mean filters. The Matlab code in Listing 1 computes the neighborhood sizes. The Matlab code in Listing 2 filters the vector `rho` by using the `filterParam` struct and the procedure outlined in Section 4.3. The observant reader notices that in fact it is not $\boldsymbol{\rho}^{(K)}$ that is saved in the filtering but $(\mathbf{D}^{(K)})^{-1}\mathbf{G}^{(K)}\mathbf{f}_K(\boldsymbol{\rho}^{(K-1)})$ (denoted by `Dinvs{K}` in the Matlab code) since this enables the use of *generalized* fW -mean filters where $g_K \neq f_K^{-1}$. The Matlab code in Listing 3 computes $\mathbf{v}^T\nabla\boldsymbol{\rho}^{(N)}$, as outlined in Section 4.3. We remark that filtering of densities must be moved inside the optimality criteria update whenever a non volume preserving filter is used.

5 Numerical experiment

Figure 2 shows the final physical design (not post processed nor sharpened!) of a cantilever beam optimized with aim of minimizing its compliance when the load is distributed over the middle 10 % of the beam's right side; a standard test problem in topology optimization.

```

1 for K = N:-1:2
2     v = df{K}(g{K-1}(DinvS{K-1})).*...
3         GT{K}(v.*dg{K}(DinvS{K}))./Ni{K});
4 end
5 v = df{1}(rho).*...
6     GT{1}(dc(:)).*dg{1}(DinvS{1}))./Ni{1});

```

Listing 3: Matlab code that modifies the sensitivities with respect to the physical design following the description in Section 4.3.

Here, we use an open–close (open followed by close) filtering strategy over octagonal shaped neighborhoods, which is expected to impose bounds on the minimum sizes of structural members and void regions [20]. However, instead of using exponential averaging to define the open–close we are using harmonic averaging as introduced in topology optimization by Svanberg and Svärd [22]. More precisely we have used $f_1(x) = f_4(x) = (x + \alpha)^{-1}$, $f_2(x) = f_3(x) = f_1(1 - x)$, $g_K = f_K^{-1}$ with the fixed parameter $\alpha = 10^{-4}$. The volume fraction used was 0.5, and the measure of non discreteness of the optimized cantilever beam (of the physical design) was 0.26 %. To the best of our knowledge, previously no contribution has used an open–close or close–open filtering strategy to solve problems with more than a few tens of thousands degrees of freedom. Here, we capitalize on the fast filtering strategy by Wadbro & Hägg [25], to solve a design problem with approximately 3.11 million degrees of freedom; the size of the filter neighborhoods are 1,981 elements for the *open step* and 145 elements for the *close step*. In total, the solution process required 91 iterations and took two hours, the filtering and modification of sensitivities accounted for around 25 % of this time, on a standard laptop equipped with an Intel i7-3740QM CPU and 16 GB RAM. It should be noted that the fast summation algorithm was executed almost 20,000 times. By examination of figure 2, we conclude that the physical design of the optimized cantilever beam exhibits minimum size control on both material and void regions. To illustrate that the bounds on the minimum sizes are dictated by the size of the neighborhoods used in the harmonic open–close filter, we present in figure 3 a series of optimized cantilever beams using different neighborhood sizes. In each sub-figure, the upper neighborhood corresponds to the open step that should impose a minimum size on the material regions while the lower neighborhood corresponds to the close step that should impose a minimum size on the void regions. We remark that the upper-left cantilever beam in figure 3 is a “coarse” version of the cantilever beam in figure 2 included to illustrate mesh-independence.

6 Concluding summary

In this paper, we have proven the existence of solutions to a fW -mean filtered, penalized continuous version of the minimal compliance problem. The existence of solutions is in accordance with previous experimental experience on mesh-independence gained by using nonlinear filters [20, 22]. As was pointed out by Svanberg and Svärd [22]; due to the non-convexity of the problem, a different filter is likely to give rise to a different solution. To facilitate switching between different filters, we recommend using a data structure

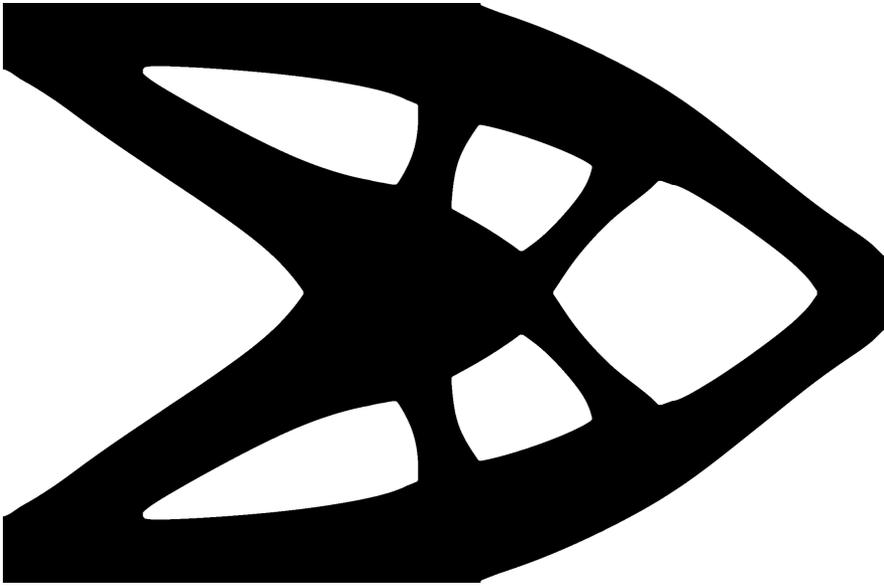


Figure 2: Filtered densities for an optimized cantilever beam using 2160×1440 elements, and a harmonic open filter over a “large” octagonal neighborhood followed by a harmonic close filter over a “small” octagonal neighborhood. The neighborhoods are indicated in the upper-right corner.

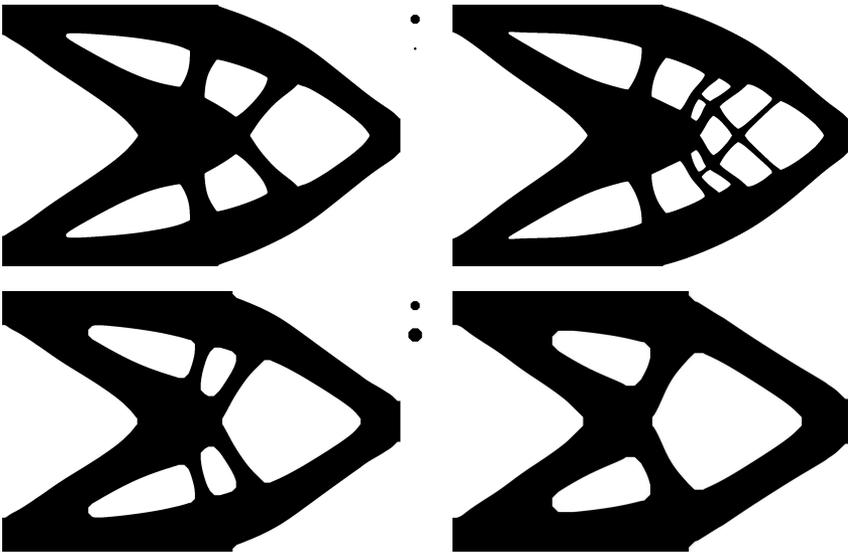


Figure 3: Optimized cantilever beams using 1440×960 elements, and different harmonic open–close filters. The octagonal neighborhoods are indicated in the upper-right corner of each sub-figure, and the upper neighborhood corresponds to the open step while the lower corresponds to the close step.

similar to that found in Table 1. We have presented results from a large-scale topology optimization of a two dimensional cantilever beam. By using a cascade of four fW -mean filters we are able to impose minimum feature sizes on both material and void regions independently, and the nonlinear nature of the filter results in a physical design that is almost black and white. One of the keys to enable solutions of large-scale problems is the fast filtering algorithm presented in [25], which enables us to filter densities and modify sensitivities with a computational cost proportional to the number of design variables. It should also be remarked that no continuation over penalty or filter parameters was used. Filtering over complex neighborhood shapes can be achieved by cascading filters over simple neighborhood shapes, and doing so might considerably simplify the implementation of the filtering procedure. We have explained that uniform weighting within neighborhoods is the preferred choice when using filters designed to mimic max or min operators.

7 Acknowledgements

This work is financially supported by the Swedish Foundation for Strategic Research (No. AM13-0029) and by the Swedish Research Council (No. 621-3706). The authors thank Martin Berggren at the Department of Computing Science, Umeå University for valuable feedback.

References

- [1] O. Amir, N. Aage, and B. S. Lazarov. On multigrid-CG for efficient topology optimization. *Structural and Multidisciplinary Optimization*, 49(5):815–829, 2014.
- [2] C. Andreasen and O. Sigmund. Topology optimization of fluid–structure-interaction problems in poroelasticity. *Computer Methods in Applied Mechanics and Engineering*, 258:55–62, 2013.
- [3] M. P. Bendsøe and N. Kikuchi. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71:197–224, 1988.
- [4] M. P. Bendsøe and O. Sigmund. *Topology Optimization. Theory, Methods, and Applications*. Springer, 2003.
- [5] T. Borrvall. Topology optimization of elastic continua using restriction. *Archives of Computational Methods in Engineering*, 8(4):351–385, 2001.
- [6] B. Bourdin. Filters in topology optimization. *International Journal for Numerical Methods in Engineering*, 50:2143–2158, 2001.
- [7] T. E. Bruns and D. A. Tortorelli. Topology optimization of non-linear elastic structures and compliant mechanisms. *Computer Methods in Applied Mechanics and Engineering*, 190:3443–3459, 2001.

- [8] R. E. Christiansen, B. S. Lazarov, J. S. Jensen, and Ole Sigmund. Creating geometrically robust designs for highly sensitive problems using topology optimization: Acoustic cavity design. *Structural and Multidisciplinary Optimization*, In press:18 pages, 2015.
- [9] A. Clausen, N. Aage, and O. Sigmund. Topology optimization of coated structures and material interface problems. *Computer Methods in Applied Mechanics and Engineering*, 290:524–541, 2015.
- [10] J. D. Deaton and R. V. Grandhi. A survey of structural and multidisciplinary continuum topology optimization: post 2000. *Structural and Multidisciplinary Optimization*, 49(1):1–38, 2014.
- [11] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund. Time domain topology optimization of 3d nanophotonic devices. *Photonics and Nanostructures - Fundamentals and Applications*, 12(1):23–33, 2014.
- [12] A. Erentok and O. Sigmund. Topology optimization of sub-wavelength antennas. *IEEE Transactions on Antennas and Propagation*, 59(1):58–69, 2011.
- [13] J. K. Guest, A. Asadpoure, and S.-H. Ha. Eliminating beta-continuation from heaviside projection and density filter algorithms. *Structural and Multidisciplinary Optimization*, 44(4):443–453, 2011.
- [14] J. K. Guest, J. H. Provost, and T. Belytschko. Achieving minimum length scale in topology optimization using nodal design variables and projection functions. *International Journal for Numerical Methods in Engineering*, 61(2):238–254, 2004.
- [15] E. Hassan, E. Wadbro, and M. Berggren. Topology optimization of metallic antennas. *IEEE Transactions on Antennas and Propagation*, 63(5):2488–2500, 2014.
- [16] A. Klarbring and N. Strömberg. Topology optimization of hyperelastic bodies including non-zero prescribed displacements. *Structural and Multidisciplinary Optimization*, 47(1):37–48, 2013.
- [17] J. Kook, K. Koo, J. Hyun, J.S. Jensen, and S. Wang. Acoustical topology optimization for Zwicker’s loudness model — Application to noise barriers. *Computer Methods in Applied Mechanics and Engineering*, 237–240:130–151, 2012.
- [18] J. Park and A. Sutradhar. A multi-resolution method for 3d multi-material topology optimization. *Computer Methods in Applied Mechanics and Engineering*, 285:571–586, 2015.
- [19] O. Sigmund. *Design of material structures using topology optimization*. PhD thesis, Technical University of Denmark, 1994.
- [20] O. Sigmund. Morphology-based black and white filters for topology optimization. *Structural and Multidisciplinary Optimization*, 33(4–5):401–424, 2007.
- [21] O. Sigmund and A. Maute. Topology optimization approaches. *Structural and Multidisciplinary Optimization*, 48(6):1031–1055, 2013.

- [22] K. Svanberg and H. Svärd. Density filters for topology optimization based on the pythagorean means. *Structural and Multidisciplinary Optimization*, 48(5):859–875, 2013.
- [23] E. Wadbro. Analysis and design of acoustic transition sections for impedance matching and mode conversion. *Structural and Multidisciplinary Optimization*, 50(3):395–408, 2014.
- [24] E. Wadbro and C. Engström. Topology and shape optimization of plasmonic nano-antennas. *Computer Methods in Applied Mechanics and Engineering*, 293:155–169, 2015.
- [25] E. Wadbro and L. Hägg. On quasi-arithmetic mean based filters and their fast evaluation for large-scale topology optimization. *Structural and Multidisciplinary Optimization*, 52(5):879–888, 2015.
- [26] G. H. Yoon. Topology optimization for stationary fluid–structure interaction problems using a new monolithic formulation. *International Journal for Numerical Methods in Engineering*, 82:591–616, 2010.