Strategies for adaptive optimization with aggregation constraints using interior-point methods

Graeme J. Kennedy *

Abstract

Constraint-aggregation methods are used in engineering optimization problems to approximately impose a bound on a quantity of interest in a differentiable manner. In this paper, we present strategies to adaptively solve aggregation-constrained problems. These adaptive techniques achieve a tighter bound approximation while also reducing the computational cost of optimization. We focus on two aggregation techniques: Kreisselmeier–Steinhauser (KS) aggregation, and induced exponential aggregation. We demonstrate that the proposed adaptive technique achieves significant computational savings compared to fixed-aggregation methods for a series of stress-constrained mass-minimization problems.

1 Introduction

Engineering design optimization problems often require a constraint that a quantity of interest must not exceed an allowable value. For example, in structural optimization problems, a bound is often placed on a material failure criteria within the structure [1, 26, 18, 16], while in aerodynamic design problems, a constraint may be imposed on the maximum Mach number in the domain during a dive condition as a surrogate for a buffet criterion [6]. In the context of PDE-constrained optimization, this type of bound constraint is infinite dimensional in the sense that the bound must be enforced at all points in the problem domain. Constraint-aggregation methods remove this infinite-dimensional bound constraint and replace it with a differentiable approximation of the maximum value of the quantity of interest over the domain. Since the aggregation technique is differentiable, the resulting optimization problem can be solved using gradient-based design optimization.

Unfortunately, there are several issues associated with aggregation. First, the approximate bound produced by the aggregation constraint will not be exact, producing an optimal design with an often unknown performance penalty, or a bound violation. Second, aggregation functionals are highly nonlinear and can significantly increase the computational cost of optimization. Third, utilizing multiple aggregation functionals over separate aggregation sub-domains can reduce the number of optimization iterations, but also incurs additional gradient-evaluation costs. Finally, accuracy and optimization cost are related: aggregation methods utilize a parameter that controls the approximation, but increasing this parameter to produce a more accurate bound also increases

*Assistant Professor, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, email: graeme.kennedy@ae.gatech.edu, phone: 404-894-8911

the nonlinearity in the optimization problem, incurring additional optimization costs. These issues make it difficult to select the correct combination of aggregation technique, aggregation parameter, and aggregation domains *a priori*.

In this work, we present a method to address these issues. We adaptively control the aggregation parameter and the aggregation domains to control optimization cost and achieve an accurate, optimized design. To accomplish this, we use the properties of the Kreisselmeier–Steinhauser (KS) functional [17, 1] and the induced exponential functional [14]. The properties of these aggregation methods enable adaptation during an optimization with corrective steps to accommodate the adaptation in a computationally efficient manner.

1.1 Constraint aggregation methods

Two of the most common constraint-aggregation functions are the discrete KS function and the p-norm function. The discrete KS function was originally presented in the context of control systems design [17] and has subsequently been applied to constrained optimization methods [34], chemical engineering problems [3, 28], structural engineering design [1, 26, 18, 13, 25, 15], and aircraft design [20, 21, 16]. Variants of the discrete KS function have also been used for large-scale design problems based on a generalized exponential function [27]. The p-norm function has been used to impose stress constraints in many structural topology optimization problems [35, 18, 11].

Much of the literature on constraint aggregation focuses on discrete aggregation techniques which operate on a discrete set of constraints. These discrete aggregation methods can be utilized for PDE-constrained problems by performing discrete aggregation over a set of values obtained by evaluating the physical quantity of interest at trial locations within the problem domain. However, this approach is not well-suited for higher-order methods, such as isogeometric analysis or higher-order isoparametric analysis, since no bound is enforced between trial point locations. In addition, these discrete aggregation methods exhibit mesh-dependence [14]. Optimization problems formulated utilizing higher-order analysis techniques should utilize rigorous constraint-aggregation formulations instead of *ad hoc* approaches.

Constraint-aggregation methods also have a strong impact on the computational cost of optimization. Akgun et al. [1] presented an adjoint-based derivative evaluation method for stress and buckling constraints using both the discrete KS function and KS functional. They demonstrated that aggregation methods were more efficient when employed with adjoint-based derivative evaluation methods than disaggregated techniques. Martins et al. [20] used the discrete KS function for aggregating stress constraints for aerostructural design optimization of a supersonic business jet. These results demonstrate that the discrete KS function is well-suited for adjoint-based derivative methods for structural and multidisciplinary optimization.

Other authors have proposed adaptive strategies to control the accuracy of the constraint bound. Poon and Martins [26] proposed a discrete KS-function-type constraint where the aggregation parameter is selected adaptively for each evaluation to ensure sufficient accuracy. Poon and Martins used the complex step method [30, 19] to evaluate the gradient of this adaptive selection. Le et al. [18] used an adaptive method for p-norm aggregation by sorting the element stress values at different optimization iterations. These changes were non-differentiable, but were used within the method of moving asymptotes (MMA) [31].

1.2 Problem statement and definitions

In this work, we consider PDE-constrained optimization problems of the form:

$$\begin{aligned} \text{minimize} \quad & f(x,u) \\ \text{with respect to} \quad & x \in \mathbb{R}^n,\, u \in \mathbb{R}^m \\ \text{such that} \quad & c(g(\xi,x,u),\Omega_k,\rho) \leq 1 \qquad \cup_{k=1}^{M} \Omega_k = \Omega \\ \text{governed by} \quad & R(x,u) = 0 \end{aligned} \tag{Opt(ρ)}$$

where $x \in \mathbb{R}^n$ is a vector of design variables and $u \in \mathbb{R}^m$ is a vector of the state variables which is a finite-dimensional approximation of the solution to a PDE. We use a reduced-space formulation where the state variables will be considered implicit functions of the design variables, such that $u(x)$, via the discretized PDE, $R(x,u) = 0$, where $R : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is the residual.

We impose an approximate bound on the maximum value of the point-wise constraint

$$g(\xi,x,u) \leq 1, \qquad \xi \in \Omega,$$

by enforcing a series of aggregation constraints. We split the problem domain, Ω, into M non-overlapping sub-domains, Ω_k, and impose an aggregation constraint over each sub-domain, Ω_k, written as

$$c(g(\xi,x,u),\Omega_k,\rho) \leq 1. \tag{1}$$

Here c is the aggregation constraint and ρ is the aggregation parameter which controls the bound approximation. Each aggregation constraint is designed to approximately impose the constraint:

$$\max_{\xi \in \Omega_k} g(\xi,x,u) \leq 1. \tag{2}$$

However, the constraint (2) is not differentiable, and is not well-suited for gradient-based design optimization. Instead, the aggregation constraint (1) imposes an approximate bound in a differentiable manner so that gradient-based optimization techniques can be used.

The accuracy of the aggregation constraint (1) is controlled through the aggregation parameter ρ. The selection of ρ must balance several factors. Increasing ρ generally leads to a more accurate aggregation constraint, yielding a result that is closer to the true, desired bound (2). However, larger values of ρ also increases the nonlinearity in the aggregation constraint, producing optimization problems that are more difficult to solve. In addition, the selection of the aggregation domains Ω_k has an impact both on the accuracy of the aggregation constraint and the cost of the design optimization: more aggregation constraints require more design gradient evaluations. However, increasing the number of domains can also improve constraint-aggregation accuracy.

In this paper, we adaptively select ρ and Ω_k to achieve sufficient accuracy for the optimization problem Opt(ρ). The proposed technique uses an interior-point method with additional corrective steps taken after each adaptive step. To guide these adaptive steps, we derive relationships for two aggregation functionals subject to either changes in the aggregation domains, or changes to the aggregation parameter ρ. The remainder of the paper is organized as follows: In section 2 we describe the aggregation functionals, and in section 3 we derive properties of the aggregation functionals related to the adaptation steps. We then present an analysis of the aggregation steps in section 4, and in section 5 we describe the adaptation algorithm. Finally, in section 6 we present results from the proposed adaptive optimization algorithm.

2 Aggregation functionals

In this work, we consider two aggregation techniques: the KS functional and induced exponential functional. In this section, we describe the properties of these functionals that make them well-suited for constraint aggregation. For ease of presentation, throughout the remainder of the paper we drop the arguments to the point-wise constraint, $g(\xi, x, u)$ and simply write g.

The KS functional is related to the discrete KS function [17], where the discrete sum is replaced with an integral giving the functional mesh-independent convergence properties [1, 14]. The KS functional over the domain Ω takes the form

$$c_{\mathrm{KS}}(g,\Omega,\rho) = \frac{1}{\rho}\ln\left[\frac{1}{\alpha}\int_{\Omega} e^{\rho g}\, d\Omega\right] = m + \frac{1}{\rho}\ln\left[\frac{1}{\alpha}\int_{\Omega} e^{\rho(g-m)}\, d\Omega\right], \tag{3}$$

where g is the bound function, $\alpha > 0$ is a scaling parameter, and m can be chosen arbitrarily. The second, mathematically equivalent form should be used in all computations to avoid numerical issues due to finite-precision arithmetic. The KS functional has the property that

$$\lim_{\rho\to\infty} c_{\mathrm{KS}}(g,\Omega,\rho) = \max g. \tag{4}$$

For large but finite ρ, $c_{\mathrm{KS}}(g,\Omega,\rho)$ approximates $\max g$ in a differentiable manner.

The induced exponential functional is a specific form of the general approach of induced aggregation [14]. In particular, the expression for the induced exponential functional can be obtained by applying l'Hôspital's rule to the KS functional to determine its limiting behavior [14]. The induced exponential functional can be written as follows:

$$c_{\mathrm{IE}}(g,\Omega,\rho) = \frac{\int_{\Omega} g e^{\rho g}\, d\Omega}{\int_{\Omega} e^{\rho g}\, d\Omega} = \frac{\int_{\Omega} g e^{\rho(g-m)}\, d\Omega}{\int_{\Omega} e^{\rho(g-m)}\, d\Omega} \tag{5}$$

where again m can be chosen arbitrarily and the second mathematically equivalent form should be used for all computations. The induced exponential functional also has the property that

$$\lim_{\rho\to\infty} c_{\mathrm{IE}}(g,\Omega,\rho) = \max g. \tag{6}$$

Again, a large but finite ρ can be utilized so that $c_{\mathrm{IE}}(g,\Omega,\rho)$ approximates $\max g$ in a differentiable manner.

An aggregation functional is conservative when it has the property that the estimate converges from above, such that

$$c(g,\Omega,\rho) = \max g + r(g,\Omega,\rho)$$

with $r > 0$ for all $\rho > \rho^*$. In practice, the KS functional is only conservative when the parameter $\alpha > 0$ can be chosen to satisfy the following property [14]:

$$\alpha \leq |\Omega_{\max}| \qquad \Omega_{\max} = \{\xi \in \Omega \mid g = \max g\}. \tag{7}$$

That is α must be chosen at least as small as the area or volume in which g attains its maximum value. Frequently, g attains its maximum value on a set of zero measure in Ω — that is a point or line in an area, or a point, line or area in a volume. In such cases, α cannot be chosen to make the

KS functional conservative. However, depending on the value of α, this non-conservative behavior may occur in an asymptotic region for very large values of ρ. Large values of ρ, say $\rho > 10^3$, are computationally inaccessible due to the high cost of optimization as a result of aggregation nonlinearity. These factors complicate determining a good value for ρ.

While the KS functional may or may not be conservative depending on the behavior of g, the induced aggregation functional is always non-conservative due to the following inequality:

$$c_{\mathrm{IE}}(g,\Omega,\rho) \leq \frac{\int_\Omega (\max g) e^{\rho g}\, d\Omega}{\int_\Omega e^{\rho g}\, d\Omega} = \max g. \tag{8}$$

Furthermore, we show below, that the induced exponential function has a non-negative derivative with respect to the aggregation parameter ρ, and therefore converges monotonically from below for increasing ρ.

In practice, the integrals in (3) and (5) cannot be computed exactly and instead are replaced with numerical quadrature approximations. In conjunction with the quadrature scheme, we set $m = \max_i g(\xi_i)$, where ξ_i, $i = 1,\ldots,n_q$, are the quadrature points. For both the KS and induced exponential functionals, the limiting behavior of their quadrature approximations is $\max_i g(\xi_i)$ as $\rho \to \infty$, for a fixed mesh [14]. In light of this property, we adjust the quadrature scheme to reflect the expected behavior of the bound variable. For instance, when utilizing isoparametric elements, we use Gauss–Lobatto quadrature formula that include the end points of the integration interval [12]. These end points often have the largest values of the function g when material failure criteria are used, but this property is not guaranteed.

3 Accuracy analysis

The adaptive strategy proposed in this paper adjusts the domains of aggregation and the aggregation parameter ρ. To guide when we apply these adaptive steps, we present an analysis of the behavior of $c_{\mathrm{KS}}(g,\Omega,\rho)$ and $c_{\mathrm{IE}}(g,\Omega,\rho)$ as the aggregation parameter ρ increases and the domain changes.

3.1 The KS functional

The behavior of the KS functional for increasing ρ depends on the characteristics of g and the value of α. We consider two cases: first when $0 \leq |\Omega_{\max}| < \alpha$, and second when $0 < \alpha \leq |\Omega_{\max}|$. We note that it is common for the set $\Omega_{\max}$ to have zero measure in Ω such that $|\Omega_{\max}| = 0$. This is frequently the case since g is often constructed such that only locally uniform solutions, produce a locally uniform maximum g leading to $|\Omega_{\max}| \neq 0$. As a result, any design modification, or local solution change produces non-uniform g with $|\Omega_{\max}| = 0$.

Consider the aggregation parameters $\rho_1 > \rho_2 > 1$ such that:

$$\begin{aligned} 0 < e^{\rho_1(g-\max g)} &\leq e^{\rho_2(g-\max g)}, \\ 0 < \frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\, d\Omega &\leq \frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\, d\Omega. \end{aligned} \tag{9}$$

Next, consider the limit

$$\lim_{\rho\to\infty} \frac{1}{\alpha}\int_\Omega e^{\rho(g-\max g)}\, d\Omega = \frac{|\Omega_{\max}|}{\alpha} \tag{10}$$

therefore, when $\alpha > |\Omega_{\max}|$, for some ρ^* sufficiently large, with $\rho_1 > \rho_2 > \rho^* > 1$, the following integrals are bounded from above by unity

$$\frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\,d\Omega < 1, \qquad \frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\,d\Omega < 1. \tag{11}$$

As a result, utilizing both (9) and (11)

$$\left[\frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\,d\Omega\right]^{\frac{1}{\rho_1}} \geq \left[\frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\,d\Omega\right]^{\frac{1}{\rho_2}},$$

from the fact that for $0 < a \leq b < 1$ the inequality $a^{\frac{1}{p}} \geq b^{\frac{1}{q}}$ holds for $p > q > 1$. Now, taking the natural logarithm of both sides, we find:

$$\begin{aligned}
\frac{1}{\rho_1}\ln\left[\frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\,d\Omega\right] &\geq \frac{1}{\rho_2}\ln\left[\frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\,d\Omega\right], \\
\max g + \frac{1}{\rho_1}\ln\left[\frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\,d\Omega\right] &\geq \max g + \frac{1}{\rho_2}\ln\left[\frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\,d\Omega\right],
\end{aligned}$$

which yields the result:

$$c_{\mathrm{KS}}(g,\Omega,\rho_1) \geq c_{\mathrm{KS}}(g,\Omega,\rho_2) \tag{12}$$

for $\rho_1 > \rho_2 > \rho^* > 1$. Therefore, if we are in the asymptotic regime, with $\rho > \rho^*$, the value of $c_{\mathrm{KS}}(g,\Omega,\rho)$ will increase when ρ increases and the KS functional will converge from below.

On the other hand, when $0 < \alpha \leq |\Omega_{\max}|$, then the limit (10) will converge to $|\Omega_{\max}|/\alpha \geq 1$, and the integrals (11) will be bounded from below by $|\Omega_{\max}|/\alpha$ so that

$$\left[\frac{1}{\alpha}\int_\Omega e^{\rho_1(g-\max g)}\,d\Omega\right]^{\frac{1}{\rho_1}} \leq \left[\frac{1}{\alpha}\int_\Omega e^{\rho_2(g-\max g)}\,d\Omega\right]^{\frac{1}{\rho_2}},$$

from the fact that for $1 \leq a \leq b$ the inequality $a^{\frac{1}{p}} \leq b^{\frac{1}{q}}$ holds for $p > q > 1$. Now, again taking the natural logarithm of both sides as above, we find:

$$c_{\mathrm{KS}}(g,\Omega,\rho_1) \leq c_{\mathrm{KS}}(g,\Omega,\rho_2) \tag{13}$$

for $\rho_1 > \rho_2 > \rho^* > 1$. As a result, when $0 < \alpha \leq |\Omega_{\max}|$ the KS functional is conservative, and will converge from above.

Next, to illustrate the effect of a change in the aggregation domain, we consider the set $\Omega_0 \subset \Omega$. We consider two cases: first, when Ω_0 is arbitrary, and second, when Ω_0 has a special construction such that:

$$\Omega_0 = \Omega_\varepsilon, \qquad \Omega_\varepsilon = \{\xi \in \Omega \mid g \geq \max g - \varepsilon\},$$

for $\varepsilon > 0$. Based on the set Ω_0, we define the following ratio of integrals:

$$\beta = \left[\int_\Omega e^{\rho g}\,d\Omega\right]^{-1}\left(\int_{\Omega_0} e^{\rho g}\,d\Omega\right). \tag{14}$$

Note that this ratio satisfies the bounds $0 \leq \beta \leq 1$. For the case $\Omega_0 = \Omega_\varepsilon$ the ratio (14) reaches the upper bound in the limit as $\rho \to \infty$

$$\lim_{\rho \to \infty} \beta_\varepsilon = 1, \tag{15}$$

where we apply the label β_ε when $\Omega_0 = \Omega_\varepsilon$.

The relationship between the KS functional and changes to the domain can be obtained by relating the KS functional to the parameter β. The KS functional computed over the entire domain Ω, can be expressed in terms of β and the KS functional over the domain Ω_0 as follows:

$$\begin{aligned} c_{\mathrm{KS}}(g,\Omega,\rho) &= \frac{1}{\rho}\ln\left[\frac{1}{\alpha}\int_\Omega e^{\rho g}\,d\Omega\right] \\ &= \frac{1}{\rho}\ln\left[\frac{1}{\alpha}\frac{1}{\beta}\int_{\Omega_0} e^{\rho g}\,d\Omega\right] \\ &= \frac{1}{\rho}\ln\left[\frac{1}{\alpha}\int_{\Omega_0} e^{\rho g}\,d\Omega\right] - \frac{1}{\rho}\ln\beta \\ &= c_{\mathrm{KS}}(g,\Omega_0,\rho) - \frac{1}{\rho}\ln\beta \end{aligned} \tag{16}$$

Since $\beta \leq 1$, and $\rho > 0$, the final term in the expression above satisfies $-\frac{1}{\rho}\ln\beta \geq 0$. As a result, we have the inequality:

$$c_{\mathrm{KS}}(g,\Omega,\rho) \geq c_{\mathrm{KS}}(g,\Omega_0,\rho). \tag{17}$$

The equality (16) can be used to inform the construction of new aggregation sub-domains, $\Omega_k \subset \Omega$, for $k = 1,\ldots,M$, during optimization. Optimization will tend to push the point-wise quantity g towards its maximum allowable value. Therefore, it is common for g to be close to its bound at multiple points in the domain. In the aggregation refinement step, it is desirable that the KS functional take on values over each sub-domain that are as close as possible to the KS functional over the original domain. To achieve this, the new aggregation sub-domains can be constructed such that β is approximately equal for each Ω_k, $k = 1,\ldots,M$.

3.2 Induced exponential aggregation

While the behavior of the KS functional depends on g, the induced exponential functional converges from below uniformly for increasing ρ. We show this behavior by demonstrating that the induced exponential functional has a non-negative derivative with respect to the aggregation parameter ρ. Taking the derivative of $c_{\mathrm{IE}}(g,\Omega,\rho)$ with respect to the aggregation parameter gives

$$\frac{d}{d\rho}c_{\mathrm{IE}}(g,\Omega,\rho) = \frac{\int_\Omega g^2 e^{\rho g}\,d\Omega}{\int_\Omega e^{\rho g}\,d\Omega} - \left(\frac{\int_\Omega g e^{\rho g}\,d\Omega}{\int_\Omega e^{\rho g}\,d\Omega}\right)^2.$$

Rearranging this expression, we note that the derivative will be positive if

$$\left(\int_\Omega g e^{\rho g}\,d\Omega\right)^2 \leq \left(\int_\Omega g^2 e^{\rho g}\,d\Omega\right)\left(\int_\Omega e^{\rho g}\,d\Omega\right).$$

This inequality can be established using the Cauchy–Schwarz inequality by observing that $|\langle a,b\rangle|^2 \leq \langle a,a\rangle\langle b,b\rangle$ with $a = g\sqrt{e^{\rho g}}$ and $b = \sqrt{e^{\rho g}}$ where $\langle a,b\rangle = \int_\Omega ab\,d\Omega$. Therefore, the derivative of $c_{\mathrm{IE}}(g,\Omega,\rho)$ with respect to ρ is non-negative for all ρ.

In a similar manner to the KS functional, we can also obtain a relationship between the induced exponential functional and changes in the aggregation domain through the parameter β (14). The induced exponential functional can be written as follows:

$$\begin{aligned} c_{\mathrm{IE}}(g,\Omega,\rho) &= \frac{\int_{\Omega_0} g e^{\rho g}\,d\Omega}{\int_\Omega e^{\rho g}\,d\Omega} + \frac{\int_{\Omega\setminus\Omega_0} g e^{\rho g}\,d\Omega}{\int_\Omega e^{\rho g}\,d\Omega} \\ &= \beta\frac{\int_{\Omega_0} g e^{\rho g}\,d\Omega}{\int_{\Omega_0} e^{\rho g}\,d\Omega} + (1-\beta)\frac{\int_{\Omega\setminus\Omega_0} g e^{\rho g}\,d\Omega}{\int_{\Omega\setminus\Omega_0} e^{\rho g}\,d\Omega} \\ &= \beta c_{\mathrm{IE}}(g,\Omega_0,\rho) + (1-\beta)\,c_{\mathrm{IE}}(g,\Omega\setminus\Omega_0,\rho) \end{aligned} \tag{18}$$

where we use the expression $\int_{\Omega\setminus\Omega_0} e^{\rho g}\,d\Omega = (1-\beta)\int_\Omega e^{\rho g}\,d\Omega$. Now, if the following inequality holds:

$$c_{\mathrm{IE}}(g,\Omega_0,\rho) \geq c_{\mathrm{IE}}(g,\Omega\setminus\Omega_0,\rho), \tag{19}$$

then we have the following:

$$\begin{aligned} (1-\beta)c_{\mathrm{IE}}(g,\Omega_0,\rho) &\geq (1-\beta)c_{\mathrm{IE}}(g,\Omega\setminus\Omega_0,\rho), \\ c_{\mathrm{IE}}(g,\Omega_0,\rho) &\geq \beta c_{\mathrm{IE}}(g,\Omega_0,\rho) + (1-\beta)c_{\mathrm{IE}}(g,\Omega\setminus\Omega_0,\rho) \\ c_{\mathrm{IE}}(g,\Omega_0,\rho) &\geq c_{\mathrm{IE}}(g,\Omega,\rho). \end{aligned} \tag{20}$$

where we use $1-\beta > 0$. For the case $\Omega_0 = \Omega_\varepsilon$, the inequality (19) will hold for sufficiently large ρ since in the limit as $\rho\to\infty$, $\max_\Omega g > \max_\Omega - \varepsilon \geq \max_{\Omega\setminus\Omega_\varepsilon} g$. The inequality (20) has implications for the convergence of induced exponential functional. Since $c_{\mathrm{IE}}(g,\Omega,\rho)$ is non-conservative (converges from below), restricting the domain to Ω_0 improves the accuracy of the bound as long as the inequality (19) holds.

The equality (18) can also inform the construction of new aggregation sub-domains, $\Omega_k \subset \Omega$, for $k = 1,\ldots,M$. Again, it is desirable to create new sub-domains such that the induced exponential aggregate takes on approximately the same value over the whole domain as it does over each new sub-domain. This can be achieved by constructing each new aggregation sub-domain such that the value of β is approximately equal for each Ω_k, $k = 1,\ldots,M$.

4 Adaptive techniques

Based on the analysis presented above, increasing the parameter ρ has a beneficial impact on aggregation accuracy. Furthermore, creating new aggregation domains can reduce the number of optimization iterations. Therefore, at points during the optimization, we perform one of two steps:

1. We increase the aggregation parameter ρ; or

2. We sub-divide the existing aggregation domain, Ω, into non-overlapping sub-domains, Ω_k.

Both of these changes modify the design problem in a non-differentiable manner. Therefore, conventional gradient-based methods cannot handle these modifications without a partial restart, incurring additional computational costs. Instead of a restarting, we propose two corrective steps to effectively handle both of these modifications:

1. After increasing ρ, we obtain a new estimate for the design variables and Lagrange multipliers using a path-following step; or

2. After sub-dividing the aggregation domains, we obtain new Lagrange multiplier estimates and slack variables for each new aggregation constraint.

These corrective steps can be used within any optimization method, but are naturally compatible with interior-point methods. These corrective techniques allow the optimization to continue with the modified design problem without a restart by utilizing the properties of the induced exponential and KS functionals.

In the following section, we present adaptive techniques for both the KS functional and induced exponential aggregation. We first present notation and definitions for the aggregation sub-domains, we then briefly describe the interior-point method used within this work, and finally, we present the corrective steps performed after each adaptive design problem modification.

4.1 Domain splitting

During a domain modification step, we create M non-overlapping sub-domains Ω_k, that satisfy the following property:

$$\Omega = \bigcup_{k=1}^{M} \Omega_k$$

where $\Omega_i \cap \Omega_j = \partial\Omega_i \cap \partial\Omega_j$ for $i \neq j$. Following [18], we obtain these sub-domains by using a heuristic designed to obtain β-ratios (14) that are roughly equal for each sub-domain. In this work, we utilize a finite-element discretization and denote the domain of each element by $\Omega_i^{(e)}$, for $i = 1, \ldots, N_e$. We sort the element indices such that:

$$\int_{\Omega_1^{(e)}} e^{\rho(g-m)}\, d\Omega \geq \int_{\Omega_2^{(e)}} e^{\rho(g-m)}\, d\Omega \geq \ldots \geq \int_{\Omega_{N_e}^{(e)}} e^{\rho(g-m)}\, d\Omega. \tag{21}$$

We then form the domains by taking every M-th element as follows

$$\Omega_k = \bigcup_{i=k \bmod M} \Omega_i^{(e)}, \tag{22}$$

yielding a set of non-overlapping domains with approximately equal values of β assuming the values of the integrals (21) are well-distributed.

To evaluate the KS and induced exponential functionals and their gradients with respect to the design variables over each sub-domain, we introduce the following notation:

$$a_k = \int_{\Omega_k} e^{\rho(g-m)}\, d\Omega \qquad b_k = \int_{\Omega_k} g e^{\rho(g-m)}\, d\Omega \qquad c_k = \int_{\Omega_k} g^2 e^{\rho(g-m)}\, d\Omega$$

$$\mathbf{g}_k = \frac{1}{a_k}\int_{\Omega_k} e^{\rho(g-m)} \nabla_x g\, d\Omega \quad \mathbf{h}_k = \frac{1}{a_k}\int_{\Omega_k} g e^{\rho(g-m)} \nabla_x g\, d\Omega \quad \mathbf{p}_k = \frac{1}{a_k}\int_{\Omega_k} g^2 e^{\rho(g-m)} \nabla_x g\, d\Omega$$

Note that $\nabla_x g$ denotes the gradient with respect to the design variables. In addition, we introduce the weighting parameter

$$\eta_k = a_k \left[\sum_{i=1}^{M} a_i \right]^{-1}. \tag{23}$$

Based on these definitions, the KS functional and KS functional gradient are obtained as follows:

$$\begin{aligned} c_{\mathrm{KS}}(g,\Omega,\rho) &= m + \frac{1}{\rho} \ln \left[\sum_{k=1}^{M} a_k \right], \\ \nabla_x c_{\mathrm{KS}}(g,\Omega,\rho) &= \left[\sum_{k=1}^{M} a_k \right]^{-1} \left(\sum_{k=1}^{M} a_k \mathbf{g}_k \right), \\ &= \sum_{k=1}^{M} \eta_k \mathbf{g}_k. \end{aligned}$$

In turn, for the KS functional restricted to the domain Ω_k, we obtain the following functional value and derivative:

$$\begin{aligned} c_{\mathrm{KS}}(g,\Omega_k,\rho) &= m + \frac{1}{\rho} \ln a_k, \\ \nabla_x c_{\mathrm{KS}}(g,\Omega_k,\rho) &= \mathbf{g}_k. \end{aligned}$$

Note that the KS functional has a nested property such that the KS functional over the entire domain is equal to the KS function of the KS functionals over each sub-domain:

$$c_{\mathrm{KS}}(g,\Omega,\rho) = \frac{1}{\rho} \ln \left[\sum_{k=1}^{M} e^{\rho c_{\mathrm{KS}}(g,\Omega_k,\rho)} \right].$$

In addition, the KS functional gradient is a weighted combination of the gradients on all sub-domains:

$$\nabla_x c_{\mathrm{KS}}(g,\Omega,\rho) = \sum_{k=1}^{M} \eta_k \nabla_x c_{\mathrm{KS}}(g,\Omega_k,\rho). \tag{24}$$

Raspanti et al. [28] used this property to obtain Lagrange multiplier estimates for the discrete KS function for a series of constraints. In a similar manner, we use this property to estimate the Lagrange multipliers for a series of KS functionals on a new set of aggregation sub-domains.

The value and gradient of the induced exponential functional can also be obtained from the definitions above as follows:

$$\begin{aligned} c_{\mathrm{IE}}(g,\Omega,\rho) &= \left[\sum_{k=1}^{M} a_k \right]^{-1} \left(\sum_{k=1}^{M} b_k \right), \\ \nabla_x c_{\mathrm{IE}}(g,\Omega,\rho) &= \left[\sum_{k=1}^{M} a_k \right]^{-1} \sum_{k=1}^{M} a_k \left((1 - \rho c_{\mathrm{IE}}(g,\Omega,\rho)) \mathbf{g}_k + \rho \mathbf{h}_k \right) \\ &= \sum_{k=1}^{M} \eta_k \left((1 - \rho c_{\mathrm{IE}}(g,\Omega,\rho)) \mathbf{g}_k + \rho \mathbf{h}_k \right). \end{aligned}$$

The function value and gradient of the induced exponential functional for each domain Ω_k are:

$$\begin{aligned} c_{\mathrm{IE}}(g,\Omega_k,\rho) &= a_k^{-1} b_k, \\ \nabla_x c_{\mathrm{IE}}(g,\Omega_k,\rho) &= (1-\rho c_{\mathrm{IE}}(g,\Omega_k,\rho))\mathbf{g}_k + \rho \mathbf{h}_k. \end{aligned}$$

Unlike the KS functional, the induced exponential functional does not have a nested property. However, if the constraint-aggregation functionals over each domain are equal such that $c_{\mathrm{IE}}(g,\Omega_k,\rho) = c_{\mathrm{IE}}(g,\Omega,\rho)$, then we have the following equality:

$$\nabla_x c_{\mathrm{IE}}(g,\Omega,\rho) = \sum_{k=1}^{M} \eta_k \nabla_x c_{\mathrm{IE}}(g,\Omega_k,\rho). \tag{25}$$

We stress that this equality does not hold in general. However, through the selection of the aggregation sub-domains (22), we maintain values of $c_{\mathrm{IE}}(g,\Omega_k,\rho)$ that are approximately equal. Therefore, we approximate new Lagrange multipliers using (25), in an analogous manner as (24).

Finally, when increasing the aggregation parameter ρ between optimization iterations, we evaluate the derivative of the constraint-aggregation functionals with respect ρ to obtain a first-order estimate of the resulting design. This process is described in section 4.3.1.

The derivative of the KS functional with respect to the aggregation parameter ρ is

$$\frac{d}{d\rho} c_{\mathrm{KS}}(g,\Omega,\rho) = \frac{1}{\rho}\left(c_{\mathrm{IE}}(g,\Omega,\rho) - c_{\mathrm{KS}}(g,\Omega,\rho)\right). \tag{26}$$

Note that the derivative converges to zero as $\rho \to \infty$. For the induced exponential functional the derivative is

$$\frac{d}{d\rho} c_{\mathrm{IE}}(g,\Omega,\rho) = c_{\mathrm{IE2}}(g,\Omega,\rho) - \left(c_{\mathrm{IE}}(g,\Omega,\rho)\right)^2. \tag{27}$$

The functional $c_{\mathrm{IE2}}(g,\Omega,\rho)$ and its gradient are defined as follows:

$$\begin{aligned} c_{\mathrm{IE2}}(g,\Omega,\rho) &= \frac{\int_\Omega g^2 e^{\rho g}\, d\Omega}{\int_\Omega e^{\rho g}\, d\Omega} = \frac{\sum_{k=1}^{M} c_k}{\sum_{k=1}^{M} a_k}, \\ \nabla_x c_{\mathrm{IE2}}(g,\Omega,\rho) &= \left[\sum_{k=1}^{M} a_k\right]^{-1} \left(\sum_{k=1}^{M} (\rho a_k \mathbf{p}_k + 2 a_k \mathbf{h}_k) - c_{\mathrm{IE2}}(g,\Omega,\rho) \sum_{k=1}^{M} \rho a_k \mathbf{g}_k\right). \end{aligned} \tag{28}$$

For a single domain, the functional takes the following form

$$\begin{aligned} c_{\mathrm{IE2}}(g,\Omega_k,\rho) &= a_k^{-1} c_k, \\ \nabla_x c_{\mathrm{IE2}}(g,\Omega_k,\rho) &= \rho \mathbf{p}_k + 2\mathbf{h}_k - \rho c_{\mathrm{IE2}}(g,\Omega_k,\rho)\mathbf{g}_k. \end{aligned} \tag{29}$$

Note that the derivative of the induced exponential functional (27) also converges to zero as $\rho \to \infty$ since $\lim_{\rho\to\infty} c_{\mathrm{IE2}}(g,\Omega,\rho) = (\max g)^2$.

4.2 An interior-point method

In this work, we use an interior-point method to solve optimization problems with aggregation constraints. The interior point technique is well suited to the adaptive refinement steps that we

utilize during optimization. In this section, we present the essential aspects of the interior-point method for the purposes of this paper and omit details that can be found in the literature; see for instance [29, 8, 32].

We consider the nonlinear optimization problem:

$$\begin{aligned} \min_{x} \;& f(x,u(x)) \\ \text{s.t. } & 1 - c_k(g,\Omega_k,\rho) = s_k, \qquad k = 1,\ldots,M \\ & s \geq 0 \end{aligned} \tag{30}$$

where $s \in \mathbb{R}^M$ are non-negative slack variables. In our implementation, we also impose bound constraints on x such that $l \leq x \leq u$, but we omit these here to simplify the presentation. In the remainder of this section, we collect the aggregation constraints $c_k(g,\Omega_k,\rho)$, into a vector $c(x) : \mathbb{R}^n \to \mathbb{R}^M$ and write the objective function simply as $f(x)$. These simplifications are justified based on the use of the reduced-space method.

The perturbed Karush–Kuhn–Tucker (KKT) conditions for this optimization problem (30) can be written as follows:

$$\begin{aligned} \nabla_x f(x) + \nabla_x c(x)\,\lambda &= 0 \\ e - c(x) - s &= 0 \\ S\lambda &= \mu e \end{aligned} \tag{31}$$

where $\lambda \in \mathbb{R}^M$ is a vector of Lagrange multipliers, $e \in \mathbb{R}^M$ is a vector with unit entries, and $S \in \mathbb{R}^{M\times M}$ is a diagonal matrix such that $S = \text{diag}\{s_1,\ldots,s_M\}$. This system of equations can be interpreted either as perturbed KKT conditions, which recover the original KKT conditions when $\mu = 0$, or equivalently as the KKT conditions of a related logarithmic barrier problem. In the interior point method, we approximately solve a sequence of problems for a decreasing barrier parameter $\{\mu_i\}$. We use a monotone approach in which the barrier parameter is decreased by a constant fraction after each barrier problem is approximately solved [9]. We use a stopping criterion for each barrier problem and the overall optimization based on the following measure

$$E(x,\lambda,s;\mu) = \max\{\, ||\nabla_x f(x) + \nabla_x c(x)\,\lambda||_\infty,\; ||e - c(x) - s||_\infty,\; ||S\lambda - \mu_i e||_\infty \,\}. \tag{32}$$

We solve each subsequent barrier problem using a line search method. At each iteration, we compute an update to the variables and Lagrange multipliers, p_x, p_λ, and p_s, based on a solution of the following linear system:

$$\begin{bmatrix} B & A^T & 0 \\ -A & 0 & -1 \\ 0 & S & \Lambda \end{bmatrix} \begin{bmatrix} p_x \\ p_\lambda \\ p_s \end{bmatrix} = - \begin{bmatrix} \nabla_x f(x) + \nabla_x c(x)\,\lambda \\ e - c(x) - s \\ S\lambda - \mu_i e \end{bmatrix}, \tag{33}$$

where B is a limited-memory quasi-Newton approximation of the Hessian of the Lagrangian [22, 7], such that $B \approx \nabla_{xx} f + \sum_k \lambda_k \nabla_{xx} c_k$, and $A^T = \nabla_x c$ is the constraint Jacobian. We then compute α by performing a backtracking line search along the direction

$$(x,\lambda,s) + \alpha(\alpha_s p_x, \alpha_\lambda p_\lambda, \alpha_s p_s),$$

where α_s and α_λ are defined as follows:

$$\begin{aligned} \alpha_s &= \max\{\alpha \in (0,1] \mid s + \alpha p_s \geq (1-\tau)s\}, \\ \alpha_\lambda &= \max\{\alpha \in (0,1] \mid \lambda + \alpha p_\lambda \geq (1-\tau)\lambda\}, \end{aligned} \tag{34}$$

for $\tau = 0.995$ which ensures that $s > 0$ and $\lambda > 0$. For the line search we use the following merit function:

$$\phi(x, \lambda, s; \gamma) = f(x) - \mu \sum_{k=1}^{M} \ln s_k + \gamma ||e - c(x) - s||_2. \tag{35}$$

where we select γ large enough to ensure (35) decreases along (p_x, p_λ, p_s). We also ensure a descent direction of (35) is possible by monitoring the inertia of a symmetric version of the matrix (33) combined with a regularization technique [23, Appendix B]. We implement a second-order correction in conjunction with the line search to alleviate the Maratos effect [23].

4.3 Adaptive steps

The optimization problem changes in a non-differentiable manner after an adaptive step which increases ρ, or a change in the aggregation domains. In this section, we use the properties of the aggregation functionals to construct corrective steps after either of these adaptive steps are taken.

4.3.1 Steps increasing ρ

In an analogous manner to the barrier problem, we can use path-following steps to approximately follow the path of solutions of problem (30), $x^*(\rho)$, for increasing ρ. These steps are first-order predictions of the change in x^* corresponding to the change in aggregation parameter $\Delta\rho$. Taking the derivative of the perturbed KKT conditions (31) with respect to ρ gives the following equation for the ρ-adaptive corrective step:

$$\begin{bmatrix} B & A^T & 0 \\ -A & 0 & -1 \\ 0 & S & \Lambda \end{bmatrix} \begin{bmatrix} p_x^{\mathrm{cr}} \\ p_\lambda^{\mathrm{cr}} \\ p_s^{\mathrm{cr}} \end{bmatrix} = -\Delta\rho \, \frac{d}{d\rho} \begin{bmatrix} \nabla_x c(x)\, \lambda \\ -c(x) \\ 0 \end{bmatrix}. \tag{36}$$

For the KS functional, the update (36) can be computed as follows:

$$\begin{bmatrix} B & A^T & 0 \\ -A & 0 & -1 \\ 0 & S & \Lambda \end{bmatrix} \begin{bmatrix} p_x^{\mathrm{KS}} \\ p_\lambda^{\mathrm{KS}} \\ p_s^{\mathrm{KS}} \end{bmatrix} = -\frac{\Delta\rho}{\rho} \begin{bmatrix} (\nabla_x c_{\mathrm{IE}} - \nabla_x c_{\mathrm{KS}})\, \lambda \\ -(c_{\mathrm{IE}} - c_{\mathrm{KS}}) \\ 0 \end{bmatrix}, \tag{37}$$

where we use the expression (26). For the induced exponential aggregation functional, the update (36) can be computed as follows:

$$\begin{bmatrix} B & A^T & 0 \\ -A & 0 & -1 \\ 0 & S & \Lambda \end{bmatrix} \begin{bmatrix} p_x^{\mathrm{IE}} \\ p_\lambda^{\mathrm{IE}} \\ p_s^{\mathrm{IE}} \end{bmatrix} = -\Delta\rho \begin{bmatrix} (\nabla_x c_{\mathrm{IE2}} - 2\nabla_x c_{\mathrm{IE}} C)\, \lambda \\ -(c_{\mathrm{IE2}} - C\, c_{\mathrm{IE}}) \\ 0 \end{bmatrix}, \tag{38}$$

where $C = \mathrm{diag}\{c_{\mathrm{IE}}\}$ and we have used the expression (27). After we have computed either of the steps (37) or (38), we apply the fraction to the barrier rule (34) so that the non-negativity of the Lagrange multipliers, $\lambda > 0$, and the slack variables, $s > 0$, is not violated.

Note that we use the most recent quasi-Newton Hessian approximation from the optimization algorithm for the approximate linearization of the KKT conditions (36). Alternatively, we

could compute a more accurate step by solving (36) using Hessian-vector products computed using second-order adjoints [33, 10]. While this step is more accurate, it is also more costly to compute. We have found that the additional accuracy obtained by the second-order-adjoint step is not outweighed by the additional computational costs. We attribute this to the highly nonlinear behavior of the constraint aggregation methods, which favors a larger number of inexpensive smaller steps rather than fewer, more expensive, larger steps.

4.3.2 Steps splitting the domain

For adaptive steps that form new aggregation domains, we obtain new Lagrange multiplier estimates and slack variables using the weighting parameters η_k (23). We denote the old number of domains $\tilde{M}$, the old weighting parameters $\tilde{\eta}_k$ and the old Lagrange multipliers $\tilde{\lambda} \in \mathbb{R}^{\tilde{M}}$, and compute the k-th component of the new Lagrange multipliers $\lambda \in \mathbb{R}^M$ and slack variables $s \in \mathbb{R}^M$ as follows:

$$\tilde{\lambda}_0 = \sum_{k=1}^{\tilde{M}} \tilde{\eta}_k \tilde{\lambda}_k, \qquad \lambda_k = \eta_k \tilde{\lambda}_0, \qquad s_k = \frac{\mu_k}{\lambda_k}. \tag{39}$$

Note that since we have $\eta_k > 0$, the Lagrange multipliers and slack variables remain strictly positive as required for interior-point methods.

For the KS functional, the Lagrange multiplier update (39) yields the following expression for the first KKT condition:

$$\begin{aligned}
\nabla_x f + \sum_{k=1}^{\tilde{M}} \tilde{\lambda}_k \nabla_x c_{\mathrm{KS}}(g, \tilde{\Omega}_k, \rho) &= \nabla_x f + \tilde{\lambda}_0 \nabla_x c_{\mathrm{KS}}(g, \Omega, \rho) \\
&= \nabla_x f + \sum_{k=1}^{M} \eta_k \tilde{\lambda}_0 \nabla_x c_{\mathrm{KS}}(g, \Omega_k, \rho) \\
&= \nabla_x f + \sum_{k=1}^{M} \lambda_k \nabla_x c_{\mathrm{KS}}(g, \Omega_k, \rho)
\end{aligned}$$

where we use the expression for the gradient of the KS functional (24) and the update formula (39). As a result, the first KKT condition remains unchanged after a change in the domain for fixed ρ. Furthermore, the new slack variables satisfy the third perturbed KKT condition (31). Note, however, that the values of the new KS functionals are not fixed so the value of $E(x, \lambda, s; \mu)$ changes.

For the induced exponential functional, we again use the Lagrange multiplier estimates (39), however, in this case the expression is not exact. For the induced exponential functional, the first KKT condition can be approximated as follows:

$$\begin{aligned}
\nabla_x f + \sum_{k=1}^{\tilde{M}} \tilde{\lambda}_k \nabla_x c_{\mathrm{IE}}(g, \tilde{\Omega}_k, \rho) &\approx \nabla_x f + \tilde{\lambda}_0 \nabla_x c_{\mathrm{IE}}(g, \Omega, \rho) \\
&= \nabla_x f + \sum_{k=1}^{M} \tilde{\lambda}_0 \eta_k \left((1 - \rho c_{\mathrm{IE}}(g, \Omega, \rho)) \mathbf{g}_k + \rho \mathbf{h}_k \right) \\
&\approx \nabla_x f + \sum_{k=1}^{M} \lambda_k \nabla_x c_{\mathrm{IE}}(g, \Omega_k, \rho)
\end{aligned}$$

where we have applied both the update formula (39) and the approximation (25) relying on the approximate relationships $c_{\text{IE}}(g, \tilde{\Omega}_k, \rho) \approx c_{\text{IE}}(g, \Omega, \rho)$, followed by $c_{\text{IE}}(g, \Omega, \rho) \approx c_{\text{IE}}(g, \Omega_k, \rho)$.

5 Adaptive optimization algorithm

In this section, we present an algorithm for adaptive optimization with aggregation constraints. The algorithm described here is a simple application of the adaptive steps outlined above in which we sub-divide the aggregation domains only one time and increment ρ until the aggregation constraint is sufficiently accurate such that $c(g, \Omega, \rho) \approx \max g$ and the sensitivity of the aggregation constraints with respect to a change in ρ is small.

The adaptive algorithm for aggregation-constrained optimization is:

Given the initial aggregation parameter, ρ, the aggregation increment parameter, $\frac{1}{2} \geq \tau > 0$, the aggregation sensitivity tolerance, $\delta > 0$, the aggregation error estimate, $\delta_{\max}$, the number of sub-domains to create, M_{refine}, the monotone barrier fraction $\theta \in (0, 1)$ and the stopping tolerance, ε:
Set $i \leftarrow 1$, $M \leftarrow 1$, *refined* $\leftarrow$ False, *converged* $\leftarrow$ False, $\mu_1 \leftarrow 0.1$ and $E_0 \leftarrow E(x, \lambda, s; \mu_1)$
while $E(x, \lambda, s; \mu_i) > \varepsilon E_0$ and *converged* is False **do**
    Set *converged* $\leftarrow$ False
    while $E(x, \lambda, s; \mu_i) > 10\mu_i$ **do** ▷ Solve the barrier problem
        Compute a step (p_x, p_λ, p_s) using (33)
        Compute α_s and α_λ using (34)
        Perform a backtracking line search using the merit function (35)
    end while
    if $10\mu_i > \varepsilon$ **then** ▷ Decrease μ_i before the refinement steps
        Set $\mu_{i+1} \leftarrow \theta\mu_i$
        Set $i \leftarrow i + 1$ and **continue**
    end if
    if $||dc/d\rho||_\infty > \delta$ or $|\,||c(x)||_\infty - |\max_i g(\xi_i)|\,| > \delta_{\max}$ **then** ▷ Increase ρ for accuracy
        Compute correction step with $\Delta\rho = \tau\rho$ using (37) or (38)
        Compute α_s and α_λ using (34)
        Set $(x, \lambda, s) \leftarrow (x, \lambda, s) + (\alpha_s p_x^{\text{cr}}, \alpha_\lambda p_\lambda^{\text{cr}}, \alpha_s p_s^{\text{cr}})$
        Set $\rho \leftarrow (1 + \tau)\rho$
        Set $i \leftarrow i + 1$ and **continue**
    else
        Set *converged* $\leftarrow$ True
    end if
    if *refined* is False **then** ▷ Split the domain into sub-domains
        Set $M \leftarrow M_{\text{refine}}$ and refine the domains using (22)
        Compute new Lagrange multiplier estimates and slack variables using (39)
        Set *refined* $\leftarrow$ True, *converged* $\leftarrow$ False
        Set $i \leftarrow i + 1$ and **continue**
    end if
end while

In this algorithm, the solution of each subsequent barrier problem is performed using conventional interior-point techniques. The modifications for the adaptive steps occur outside the inner loop that solves the barrier problem itself. The proposed algorithm has the following features:

1. We solve the optimization problem to a tolerance of ε, without any adaptation steps. Therefore, all adaption steps are performed once the design is optimized with an initially fixed value of ρ.

2. We increment the aggregation parameter, ρ, until the aggregation constraints satisfy the criteria:

$$||dc/d\rho||_\infty \leq \delta, \qquad \text{and}$$

$$\left| \, ||c(x)||_\infty - |\max_i g(\xi_i)| \, \right| \leq \delta_{\max}.$$

   These conditions are designed to ensure that the aggregation constraint is not too sensitive to the aggregation parameter, and that its quadrature approximation is sufficiently close to $\max_i g(\xi_i)$ — the bound evaluated at all quadrature points.

3. We split the existing aggregation domain into sub-domains once ρ is sufficiently large so that the aggregation constraint is accurate and less-sensitive to changes in the parameter ρ.

Enhancements of this method could be developed that might further improve performance. These enhancements could include multiple refinement steps, an adaptive selection of the increment to ρ, or interleaving aggregation and barrier parameter updates. In addition, we have written the algorithm using a monotone technique in which the barrier parameter is decreased by fraction $\theta \in (0,1)$ after the approximate solution of each barrier problem [9]. Further refinements of this monotone strategy are possible [24]. Nevertheless, this algorithm contains the essential components of the proposed adaptive techniques and exhibits better performance than non-adaptive approaches.

6 Results

In this section, we apply the proposed adaptive optimization technique to three test cases: a variable-thickness sheet problem subject to in-plane loads, a plate thickness optimization subject to out-of-plane pressure loads, and a wingbox design problem, subject to fixed aerodynamic loads. In all cases, we minimize the mass of the structure subject to a bound on the von Mises stress. To demonstrate the flexibility of the aggregation and adaptation techniques, we model the in-plane variable-thickness sheet problem using conventional bilinear finite-elements, the plate-bending problem with isogeometric elements, and the wingbox problem with third-order MITC9 isoparametric shell elements [5]. For all cases, we perform the finite-element analysis using the Toolkit for the Analysis of Composite Structures (TACS) [15], a parallel finite-element code implemented for gradient-based design optimization. TACS implements the adjoint method and uses an optimized implementation for calculating the gradients of multiple constraint functions simultaneously [15].

In the following studies, we fix a number of the parameters within the adaptive optimization algorithm presented above. We set the overall stopping tolerance to a value of $\varepsilon = 10^{-5}$. We set

the barrier reduction parameter $\theta = 0.2$, and increment the aggregation parameter using a value of $\tau = 0.15$. To control aggregation accuracy, we set a sensitivity tolerance of $\delta = 10^{-5}$ and an aggregation error tolerance of $\delta_{\max} = 0.01$. We set the number of sub-domains to create at the domain splitting step to $M_{\text{refine}} = 16$. For the KS aggregation cases, we set the initial aggregation parameter to $\rho = 50$, and for induced exponential aggregation, we set an initial aggregation parameter to $\rho = 25$. Note that we use a smaller initial value of ρ for induced exponential aggregation since it tends to exhibit greater nonlinearity compared to KS aggregation for the same value of the aggregation parameter.

In the following section, we compare the computational cost of different methods by recording the number of function and gradient evaluations during optimization. In this work, we add a function evaluation each time the algorithm computes the objective and constraint values, and a gradient evaluation each time the algorithm computes the objective and constraint gradients. We use this technique, instead of counting each function and gradient call individually, because the most computational time is incurred in evaluating a single objective or constraint function, while additional constraint functions are significantly less expensive. Likewise, for our adjoint implementation [15], small numbers of additional constraint gradients, say less than 20, are relatively inexpensive to compute relative to a single gradient evaluation. These computational times, however, will depend on the implementation of the analysis and gradient evaluation technique.

6.1 Plane stress results using isoparametric elements

For the first study, we compare the computational cost of the adaptive constraint-aggregation method with a fixed-aggregation approach for the design of a planar structure subject to in-plane loads with thickness variables. In this problem, each element is assigned a spatially uniform thickness. To smooth the distribution of the thickness variables, we use a spatial filter [4] that is an intermediate layer between the thickness values and the design variables. The spatial filter is defined as follows:

$$t^{(i)} = \frac{\sum_{j \in \mathscr{J}_i} (r_0 - r_{ij}) x^{(j)}}{\sum_{j \in \mathscr{J}_i} (r_0 - r_{ij})} \qquad \text{where} \qquad \mathscr{J}_i = \left\{ \; j \in \mathbb{Z} \mid r_{ij} < r_0 \; \right\} \tag{40}$$

where $t^{(i)}$ is the thickness variable associated with element i, r_{ij} is the spatial distance between the centroids of elements i and j, r_0 is the filter radius, and $x^{(j)}$ is the design variable associated with element j.

Figure 1 illustrates the design domain and loading for the plane stress problems. For this problem, we use normalized material properties, loads and geometry. The problem consists of a $2L \times L$ domain where we use $L = 250$ with a load of $P = 2500$. For the material properties we use a Young's modulus of $E = 70\,000$, and a Poisson's ratio of $\nu = 0.3$, and apply a design allowable stress of $\sigma_d = 200$. We restrict the design variables to the range $[2, 40]$, using bound constraints. We model the structure using 96×48 bilinear plane stress finite-elements and treat the stress and strain distributions as linearly varying over each element. For constraint-aggregation, we use a 3×3 tensor-product Gauss–Lobatto quadrature scheme to match the integration accuracy of the 2×2 Gauss quadrature scheme for the element itself. Note that the Gauss–Lobatto scheme includes the element corner points.

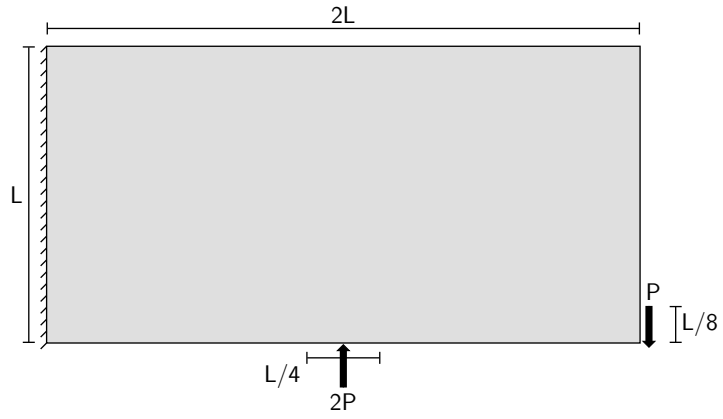
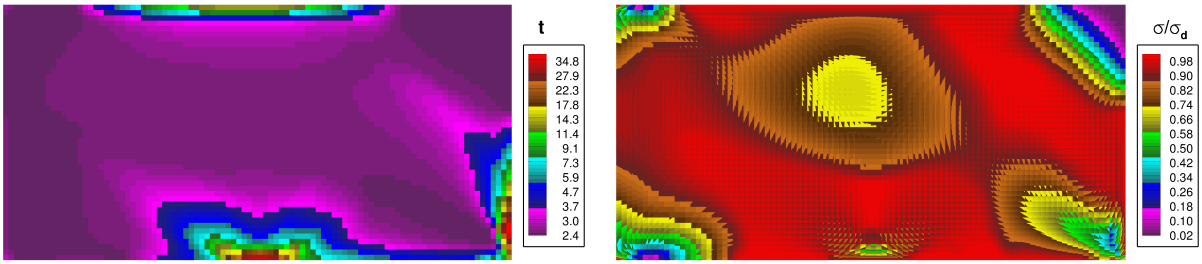


Figure 1: The design domain for the variable-thickness sheet minimum mass design.



(a) Optimized thickness

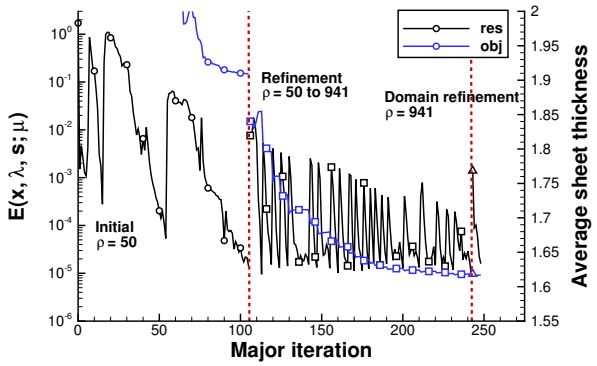
(b) Optimized von Mises stress

Figure 2: The design domain, boundary conditions, and loading for the variable-thickness sheet design problem.

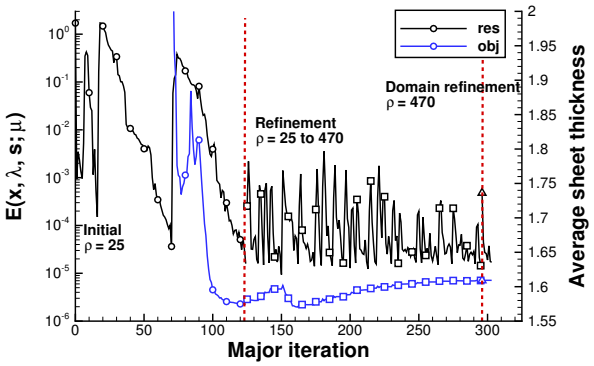
Figure 2 shows the optimized thickness distribution and the resulting von Mises stress distribution. The largest thicknesses are located in the vicinity of the applied loads and at the top of the design domain to maximize the overall bending stiffness of the structure. Figure 2b shows that the von Mises stress is near the design allowable over a large portion of the domain.

Figure 3 shows the optimality error (32) and average thickness for both fixed- and adaptive-aggregation optimization. Note that the average plate thickness is proportional to the mass of the structure since we use a constant material density. The vertical dashed lines indicate the transition between the different phases of the adaptive algorithm: the initial optimization at fixed aggregation parameter; the adaptive steps for increasing aggregation parameter; and the adaptive steps for increasing aggregation parameter after sub-domain refinement, respectively. For the adaptive KS aggregation method, the mass decreases as the aggregation parameter increases, while for the adaptive induced exponential aggregation method, the mass generally increases as the aggregation parameter increases. Furthermore, the induced exponential aggregation provides a better estimate of the final optimized mass for smaller values of the aggregation parameter. Note that the large jumps in the optimality error measure (32) correspond to barrier parameter updates.

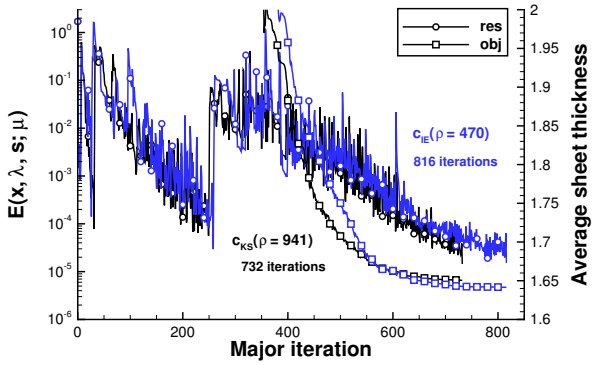
Figure 3a shows the KS aggregation results from the adaptive optimization algorithm. The overall optimization requires 282 function and 249 gradient evaluations. The final average thickness is 1.6168 units. The initial optimization with an aggregation parameter of $\rho = 50$, requires



(a) Adaptive KS aggregation



(b) Adaptive induced exponential aggregation



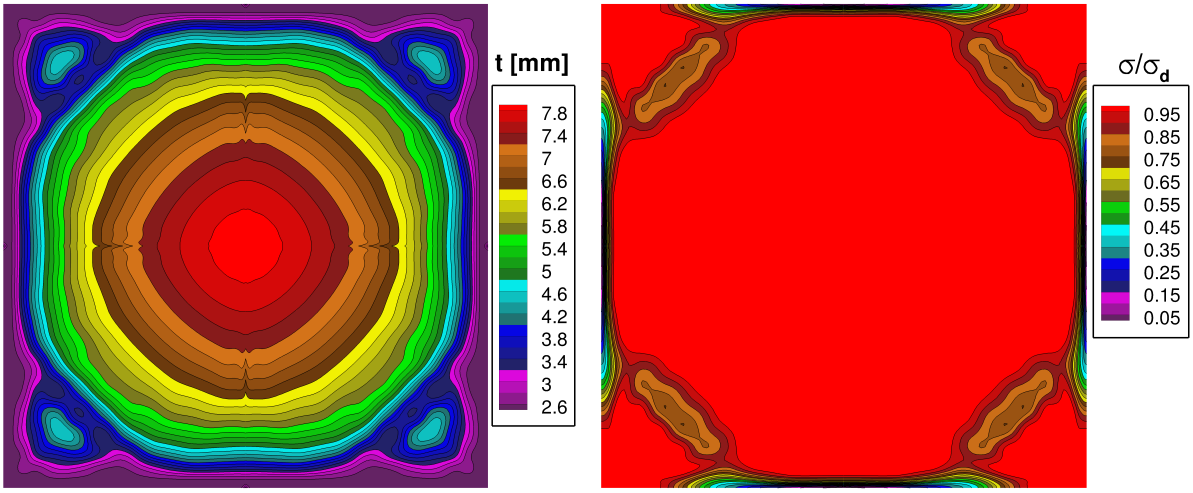
(c) Fixed aggregation

Figure 3: Design optimization histories for the variable-thickness-sheet design problem. The vertical dashed lines indicate the transition between the different phases of the adaptive algorithm.

134 function and 106 gradient evaluations. The ρ-adaptive stage of the optimization requires 21 adaptive steps in which the aggregation parameter is incremented from a value of $\rho = 50$ to a value of $\rho = 941.1$. The final ρ-adaptive optimization steps, subsequent to the domain refinement, require 7 function and 6 gradient evaluations before the full optimization tolerance is achieved.

Figure 3b shows the induced exponential aggregation results with the adaptive algorithm. The overall optimization requires 351 function and 304 gradient evaluations. The final average thickness is 1.6092 units. Note that the relative difference between the optimized objective values of the KS and induced exponential optimizations is less than 0.5%. The initial optimization with an aggregation parameter of $\rho = 25$ requires 166 function and 125 gradient evaluations. The ρ-adaptive stage of the adaptive optimization takes 21-steps where ρ is incremented from a value of $\rho = 25$ to a value of $\rho = 470.5$. The final stage of the adaptive optimization, following the domain refinement step, requires 8 function and 8 gradient evaluations.

Figure 3c shows the results of fixed-aggregation design optimization using KS and induced exponential aggregation with the parameters obtained from the final adaptive-aggregation optimization. For the fixed-aggregation optimization, KS aggregation requires 975 function calls and 733 gradient evaluations, and induced exponential aggregation requires 1057 function and 817 gradient evaluations, respectively. Therefore, the KS and induced exponential adaptive optimization methods require 71% and 67% fewer function evaluations and 61% and 62% fewer gradient eval-



(a) Optimized thickness

(b) Optimized von Mises stress

Figure 4: The design domain for the variable-thickness sheet minimum mass design.

uations than the fixed-aggregation approach. The adaptive aggregation method enables significant computational savings compared to the fixed-aggregation method.

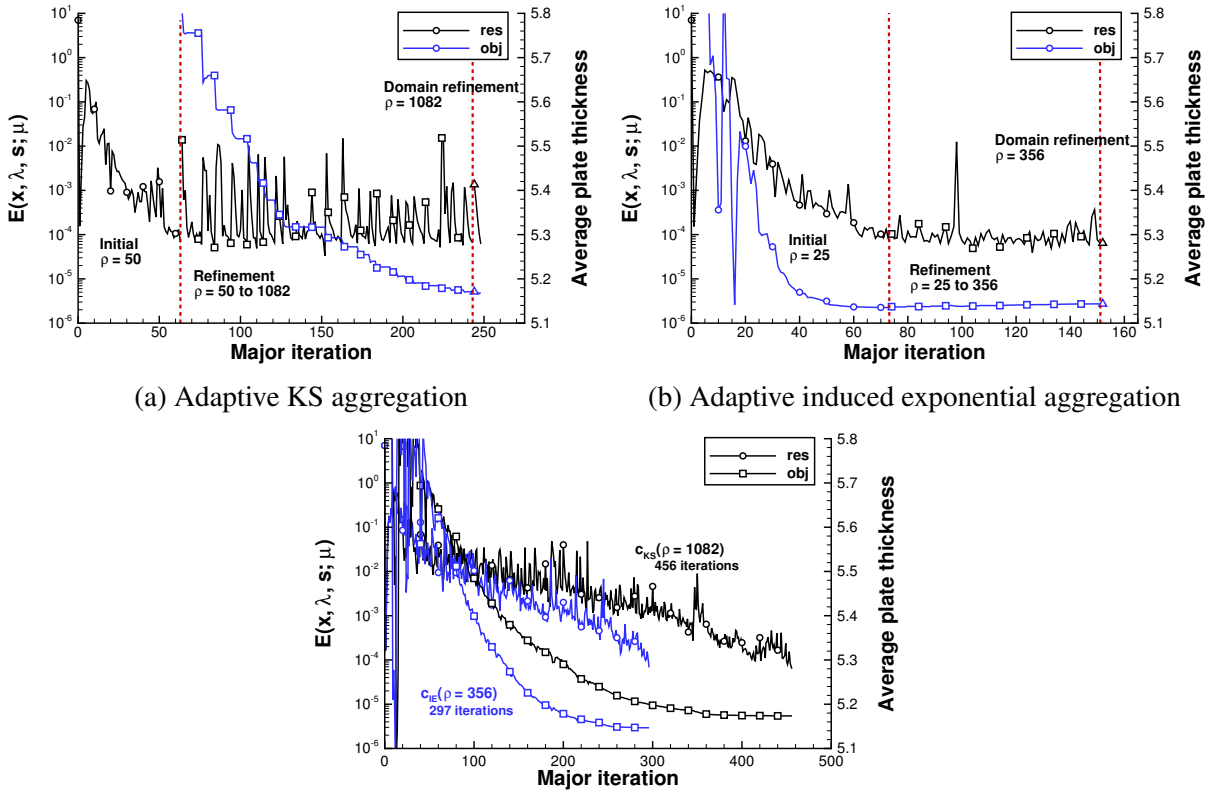
6.2 Plate bending results using isogeometric elements

In this study, we perform mass-minimization of a simply supported square plate subject to a uniform pressure load. The plate is 1000×1000 mm, the pressure load is 100 kPa and the material has a Young's modulus of $E = 70$ GPa, and a Poisson's ratio of $\nu = 0.3$. We use a design allowable stress of 350 MPa. Note that for this problem, we evaluate the von Mises stress at the top and bottom surface of the plate. We impose a lower thickness bound of 2.5 mm and an upper bound of 10 mm.

We model the plate using isogeometric Reissner–Mindlin shell elements [2] with 2D tensor-product cubic B-spline basis functions. Note that we use no special treatment to alleviate shear locking and integrate the finite-elements with a 4×4 tensor-product Gauss quadrature scheme. For this problem, we use the same set of quadrature points for both analysis and constraint aggregation. We exploit the symmetry of the problem and model only one quarter of the square plate, imposing symmetry conditions. We use a mesh consisting of 33×33 nodes, resulting in a discretization with 30×30 finite-elements. We distribute the nodes using a $1 - \cos$ distribution so that the solution near the support conditions and near the center of the plate are better-resolved.

In conjunction with the isogeometric elements, we interpolate the thickness distribution over the plate using a 2D tensor-product quadratic B-spline. To ensure smoothness of the stress distribution, we use a thickness interpolation constructed such that each knot interval from the thickness interpolation spans two knot intervals from the finite-element mesh. As a result, there are 15×15 knot intervals for the thickness interpolation resulting in 289 thickness variables distributed on a regular 17×17 mesh.

Figure 4 shows the optimized plate thickness distribution and the resulting von Mises stress.



(a) Adaptive KS aggregation

(b) Adaptive induced exponential aggregation

(c) Fixed aggregation

Figure 5: Design optimization histories for the plate bending design problem. The vertical dashed lines indicate the transition between the different phases of the adaptive algorithm.

The largest thickness is located at the center of the plate, and the smallest thicknesses are located around the edges. Figure 4b shows that, in a similar manner to the variable-thickness sheet problem, the von Mises stress is near the design allowable over a large portion of the domain.

Figure 5 shows the optimality error (32) and average plate thickness for both fixed- and adaptive-aggregation optimization. The average plate thickness is again proportional to the mass of the structure since we use a constant material density. For the adaptive KS aggregation method, the mass decreases gradually as the aggregation parameter increases, while for the induced exponential aggregation method, the mass increases slightly as the aggregation parameter increases. The adaptive induced exponential aggregation method converges to a value close to the final mass after the initial optimization step. The adaptive KS aggregation method requires a much larger aggregation parameter value to achieve similar accuracy.

Figure 5a shows the adaptive KS aggregation results. The entire adaptive optimization requires 283 function and 249 gradient evaluations. The final average plate thickness is 5.168 mm. The initial optimization with an aggregation parameter of $\rho = 50$, requires 68 function and 64 gradient evaluations. The ρ-adaptive stage of the optimization requires 22 adaptive steps in which the aggregation parameter is increased from a value of $\rho = 50$ to a value of $\rho = 1082.2$. After domain splitting, the final optimization requires 5 function and 5 gradient evaluations. The total number of function evaluations and adaptive optimization steps are similar to the adaptive KS aggregation

results for the variable-thickness sheet problem presented above.

Figure 5b shows the induced exponential aggregation results for the adaptive algorithm. In this case, the adaptive optimization requires 167 function and 153 gradient evaluations. The final average plate thickness is 5.144 mm. The relative difference between the average plate thicknesses for the KS and induced exponential aggregates is less than 0.5%. The initial optimization with an aggregation parameter of $\rho = 25$ requires 88 function and 74 gradient evaluations. The ρ-adaptive stage of the adaptive algorithm requires 19 adaptive steps in which ρ is increased from a value of $\rho = 25$ to a value of $\rho = 355.8$. The stopping criteria for the final stage of the adaptive optimization is satisfied immediately, and only requires one additional function and gradient evaluation.

Figure 5c shows the results of fixed-aggregation design optimizations using the KS and induced exponential aggregation with the aggregation parameters obtained from the final adaptive-aggregation iteration. For fixed-aggregation optimization, the KS aggregation technique requires 486 function and 457 gradient evaluations. The fixed induced exponential aggregation technique requires 340 function and 297 gradient evaluations. Comparing the performance of the fixed- and adaptive-optimization methods, the adaptive KS and induced exponential optimization methods require 42% and 55% fewer function evaluations and 45% and 48% fewer gradient evaluations than the fixed-aggregation approach.

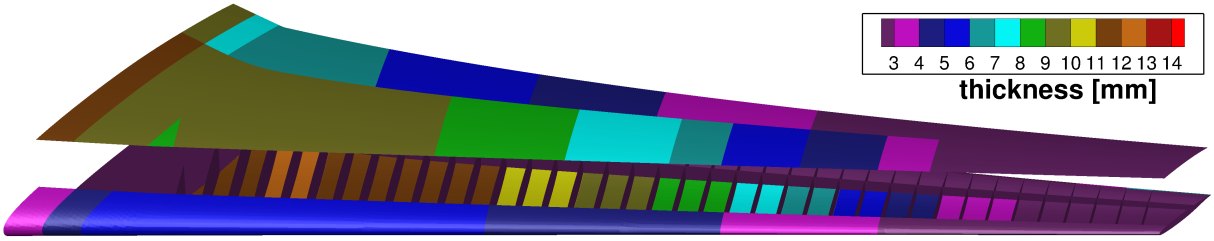
6.3 Wing design using isoparametric elements

In this section, we present the results of a wingbox optimization study with and without the adaptive optimization strategy. The wingbox geometry is based on a Boeing 777-200ER size aircraft. The wing has a root chord of 12 m, a tip chord of 2.75 m, a semi-span of 30 m, and a quarter-chord sweep of 35°. The wing structure consists of two main spars and 42 ribs. The finite-element mesh consists of 23 707 third-order MITC9 shell elements, with 90 792 nodes, resulting in just over 544 000 degrees of freedom. The loads on the wing are obtained from a 2.5 g maneuver condition evaluated using a three-dimensional panel method Kennedy and Martins [16]. In this problem, we set a Young's modulus of $E = 70$ GPa, a Poisson ratio of $\nu = 0.3$, and a design allowable stress of 350 MPa. We impose a minimum thickness of 2 mm in all structural components. Figure 6 shows the optimized thickness distribution and the resulting stress distribution under the aerodynamic loads.

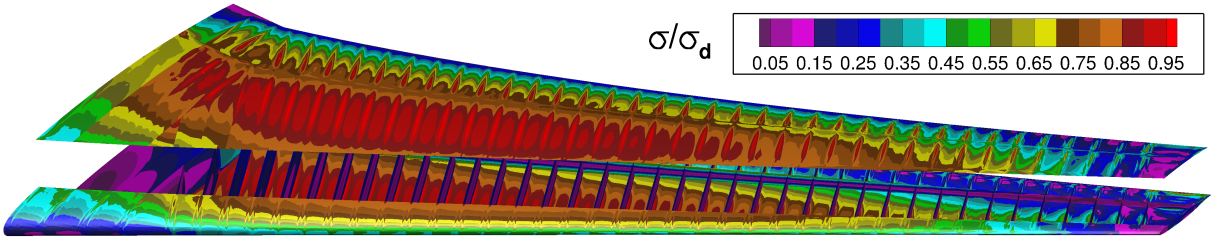
In the wingbox design problem, we group the thickness variables into common structural patches formed by the intersections of the spars and ribs, resulting in a total of 446 thickness design variables. We scale the thickness variables to work in units of mm. For constraint aggregation, we use a 4×4 tensor-product Gauss–Lobatto quadrature scheme.

Figure 7 shows the optimality error (32) and wing mass for both fixed and adaptive optimization using the KS and induced exponential aggregation. Unlike in the previous cases, the mass for both the adaptive KS and induced exponential methods increase as the aggregation parameter increases. Furthermore, the mass from the initial KS aggregation optimization is closer to the final mass than the initial induced exponential result. This change in behavior is due to the difference in the von Mises stress distributions. For the variable-thickness sheet and plate problems, the von Mises stress was near the stress allowable over a large portion of the domain, while for the wingbox case, the von Mises stress is close to the maximum allowable value over a much smaller portion of the domain.

Figure 7a shows the optimization history for the adaptive KS aggregation. The full optimization

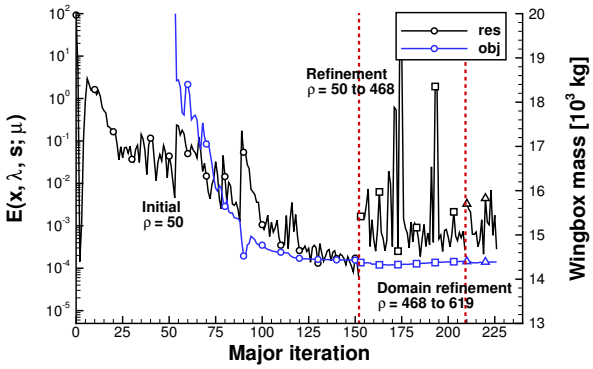


(a) Optimized thickness

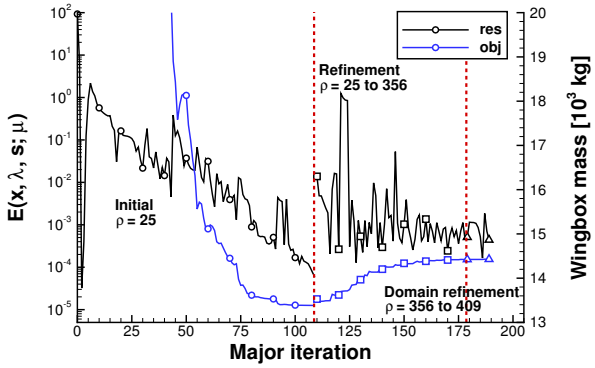


(b) Optimized von Mises stress

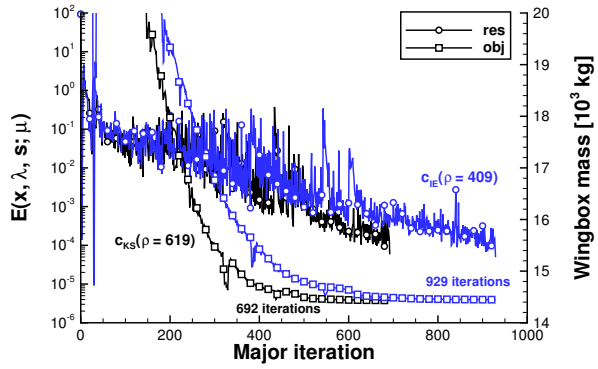
Figure 6: The optimized thicknesses and von Mises stress for the wingbox design optimization study.



(a) Adaptive KS aggregation



(b) Adaptive induced exponential aggregation



(c) Fixed aggregation

Figure 7: Design optimization histories for the wingbox design problem. The vertical dashed lines indicate the transition between the different phases of the adaptive algorithm.

requires 290 function and 227 gradient evaluations. The final wing mass is 14 387 kg. The initial optimization with an aggregation parameter of $\rho = 50$, requires 200 function and 153 gradient evaluations. The first ρ-adaptive stage of the algorithm requires 16 adaptive steps in which the aggregation parameter is incremented from a value of $\rho = 50$ to a value of $\rho = 467.9$. After the domain refinement step, 2 additional adaptive steps are taken to achieve the full tolerance, yielding a final aggregation parameter value of $\rho = 618.8$. These final steps require an additional 18 function and 17 gradient evaluations.

Figure 7b shows the induced exponential aggregation results with the adaptive algorithm. The overall optimization requires 258 function calls and 190 gradient evaluations. The final wing mass is 14 428 kg. The relative difference between the wingbox mass for the KS and induced exponential aggregates is less than 0.3%. The initial optimization with an aggregation parameter of $\rho = 25$, requires 159 function and 110 gradient evaluations. The first ρ-adaptive stage of the algorithm requires 19 adaptive steps in which the aggregation parameter is incremented from a value of $\rho = 25$ to a value of $\rho = 355.8$. After the domain refinement step, 3 additional adaptive steps are taken to achieve the full tolerance, yielding a final aggregation parameter value of $\rho = 409.2$. These final steps require an additional 15 function and 11 gradient evaluations.

Figure 7c shows the results of fixed-aggregation optimization using the KS and induced exponential functionals. In these cases, we use the final value of the aggregation parameters obtained from the adaptive-aggregation methods. For the fixed optimization approach, KS aggregation requires 869 function and 693 gradient evaluations, and induced exponential aggregation requires 1056 function and 930 gradient evaluations. As a result, the adaptive KS and induced exponential aggregation optimization techniques require 67% and 83% fewer function and 67% and 79% fewer gradient evaluations, respectively, than the fixed-aggregation methods.

7 Conclusions

In this paper, we have presented adaptive strategies for design optimization problems with aggregation constraints using interior-point methods. Using these strategies, we constructed an algorithm which we applied to a series of structural optimization problems with bounds on the allowable stress. The results demonstrate significant computational savings compared to fixed-aggregation methods where the adaptive approach required between 41% and 83% fewer function evaluations, and between 45% and 79% fewer gradient evaluations to achieve equivalent accuracy. The proposed adaptive strategies were constructed based on the properties of the KS functional and the induced exponential functional which enable corrective steps after an adaptation step which either increases the aggregation parameter, or modifies the aggregation domains. The proposed adaptive-aggregation strategy is well-suited to higher-order analysis methods such as isogeometric or isoparametric finite-element analysis. Future work will include extending the proposed adaptive optimization method to multidisciplinary applications.

Acknowledgment

The research reported in this paper was supported by the Georgia Institute of Technology. Computations within this work were performed using the Extreme Science and Engineering Discovery

Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

References

[1] M. A. Akgun, R. T. Haftka, K. C. Wu, J. L. Walsh, and J. H. Garcelon. Efficient structural optimization for multiple load cases using adjoint sensitivities. *AIAA Journal*, 39(3):511–516, 2001. doi:10.2514/2.1336.

[2] D. Benson, Y. Bazilevs, M. Hsu, and T. Hughes. Isogeometric shell analysis: The Reissner–Mindlin shell. *Computer Methods in Applied Mechanics and Engineering*, 199(58):276 – 289, 2010. ISSN 0045-7825. doi:10.1016/j.cma.2009.05.011. Computational Geometry and Analysis.

[3] K. F. Bloss, L. T. Biegler, and W. E. Schiesser. Dynamic process optimization through adjoint formulations and constraint aggregation. *Industrial & Engineering Chemistry Research*, 38 (2):421–432, 1999. doi:10.1021/ie9804733.

[4] 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 (2627):3443 – 3459, 2001. ISSN 0045-7825. doi:10.1016/S0045-7825(00)00278-4.

[5] M. L. Bucalem and K.-J. Bathe. Higher-order MITC general shell elements. *International Journal for Numerical Methods in Engineering*, 36:3729–3754, 1993. ISSN 1097-0207. doi:10.1002/nme.1620362109.

[6] H. P. Buckley, B. Y. Zhou, and D. W. Zingg. Airfoil optimization using practical aerodynamic design requirements. *Journal of Aircraft*, 47(5):1707–1719, 2013/01/29 2010. doi:10.2514/1.C000256.

[7] R. H. Byrd, J. Nocedal, and R. B. Schnabel. Representations of quasi-Newton matrices and their use in limited memory methods. *Mathematical Programming*, 63(1-3):129–156, 1994. ISSN 0025-5610. doi:10.1007/BF01582063.

[8] R. H. Byrd, J. C. Gilbert, and J. Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1):149–185, 2000. ISSN 0025-5610. doi:10.1007/PL00011391.

[9] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming*. Society for Industrial and Applied Mathematics, 1990. doi:10.1137/1.9781611971316.

[10] J. E. Hicken. Inexact Hessian-vector products in reduced-space differential-equation constrained optimization. *Optimization and Engineering*, pages 1–34, 2014. doi:10.1007/s11081-014-9258-6. URL http://dx.doi.org/10.1007/s11081-014-9258-6.

[11] E. Holmberg, B. Torstenfelt, and A. Klarbring. Stress constrained topology optimization. *Structural and Multidisciplinary Optimization*, 48(1):33–47, 2013. ISSN 1615-147X. doi:10.1007/s00158-012-0880-7.

[12] T. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover Civil and Mechanical Engineering. Dover Publications, 2012.

[13] K. A. James, J. S. Hansen, and J. R. Martins. Structural topology optimization for multiple load cases using a dynamic aggregation technique. *Engineering Optimization*, 41(12):1103–1118, 2009. doi:10.1080/03052150902926827.

[14] G. J. Kennedy and J. E. Hicken. Improved constraint-aggregation methods. *Computer Methods in Applied Mechanics and Engineering*, 2014. doi:10.1016/j.cma.2015.02.017. In press.

[15] G. J. Kennedy and J. R. R. A. Martins. A parallel finite-element framework for large-scale gradient-based design optimization of high-performance structures. *Finite Elements in Analysis and Design*, 87(0):56 – 73, 2014. ISSN 0168-874X. doi:10.1016/j.finel.2014.04.011.

[16] G. J. Kennedy and J. R. R. A. Martins. A parallel aerostructural optimization framework for aircraft design studies. *Structural and Multidisciplinary Optimization*, pages 1–23, 2014. ISSN 1615-147X. doi:10.1007/s00158-014-1108-9.

[17] G. Kreisselmeier and R. Steinhauser. Systematic control design by optimizing a vector performance index. In *International Federation of Active Controls Symposium on Computer-Aided Design of Control Systems*, Zurich, Switzerland, 1979.

[18] C. Le, J. Norato, T. Bruns, C. Ha, and D. Tortorelli. Stress-based topology optimization for continua. *Structural and Multidisciplinary Optimization*, 41:605–620, 2010. ISSN 1615-147X. doi:10.1007/s00158-009-0440-y.

[19] J. R. R. A. Martins, P. Sturdza, and J. J. Alonso. The complex-step derivative approximation. *ACM Transactions on Mathematical Software*, 29(3):245–262, Sept. 2003. doi:10.1145/838250.838251.

[20] J. R. R. A. Martins, J. J. Alonso, and J. J. Reuther. High-fidelity aerostructural design optimization of a supersonic business jet. *Journal of Aircraft*, 41(3):523–530, 2004. doi:10.2514/1.11478.

[21] J. R. R. A. Martins, J. J. Alonso, and J. J. Reuther. A coupled–adjoint sensitivity analysis method for high–fidelity aero–structural design. *Optimization and Engineering*, 6:33–62, 2005. doi:10.1023/B:OPTE.0000048536.47956.62.

[22] J. Nocedal. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation*, 35:773–782, 1980. doi:10.1090/S0025-5718-1980-0572855-7.

[23] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, 2006.

[24] J. Nocedal, A. Wächter, and R. Waltz. Adaptive barrier update strategies for nonlinear interior methods. *SIAM Journal on Optimization*, 19(4):1674–1693, 2009. doi:10.1137/060649513.

[25] J. París, F. Navarrina, I. Colominas, and M. Casteleiro. Topology optimization of continuum structures with local and global stress constraints. *Structural and Multidisciplinary Optimization*, 39(4):419–437, 2009. ISSN 1615-147X. doi:10.1007/s00158-008-0336-2.

[26] N. Poon and J. R. R. A. Martins. An adaptive approach to constraint aggregation using adjoint sensitivity analysis. *Structural and Multidisciplinary Optimization*, 34:61–73, 2007. ISSN 1615-147X. doi:10.1007/s00158-006-0061-7.

[27] J. Qin and D. Nguyen. Generalized exponential penalty function for nonlinear programming. *Computers & Structures*, 50(4):509 – 513, 1994. ISSN 0045-7949. doi:10.1016/0045-7949(94)90021-3.

[28] C. Raspanti, J. Bandoni, and L. Biegler. New strategies for flexibility analysis and design under uncertainty. *Computers & Chemical Engineering*, 24(910):2193 – 2209, 2000. ISSN 0098-1354. doi:10.1016/S0098-1354(00)00591-3.

[29] M. H. R.H. Byrd and J. Nocedal. An interior point method for large-scale nonlinear programming. *SIAM Journal on Optimization*, 9(4):877–900, 1999. doi:10.1137/S1052623497325107.

[30] W. Squire and G. Trapp. Using complex variables to estimate derivatives of real functions. *SIAM Review*, 40(1):110–112, 1998. doi:10.1137/S003614459631241X.

[31] K. Svanberg. The method of moving asymptotes - a new method for structural optimization. *International Journal for Numerical Methods in Engineering*, 24(2):359–373, 1987. ISSN 1097-0207. doi:10.1002/nme.1620240207.

[32] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*, 106(1):25–57, May 2006. ISSN 0025-5610. doi:10.1007/s10107-004-0559-y.

[33] Z. Wang, I. M. Navon, F. X. Dimet, and X. Zou. The second order adjoint analysis: Theory and applications. *Meteorology and Atmospheric Physics*, 50:3–20, 1992. doi:10.1007/BF01025501. URL `http://dx.doi.org/10.1007/BF01025501`.

[34] G. Wrenn. An indirect method for numerical optimization using the Kreisselmeier-Steinhauser function. NASA Technical Report CR-4220, 1989.

[35] R. Yang and C. Chen. Stress-based topology optimization. *Structural optimization*, 12(2-3): 98–105, 1996. ISSN 0934-4373. doi:10.1007/BF01196941.