Stress-constrained topology optimization of lattice-like structures using component-wise reduced order models

Sean McBane\textsuperscript{1}, Youngsoo Choi\textsuperscript{2}, and Karen Willcox\textsuperscript{1}

\textsuperscript{1}Oden Institute for Computational Engineering \& Sciences, University of Texas, Austin, TX 78712
\textsuperscript{2}Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA 94550

Abstract

Lattice-like structures can provide a combination of high stiffness with light weight that is useful in many applications, but a resolved finite element mesh of such structures results in a computationally expensive discretization. This computational expense may be particularly burdensome in many-query applications, such as optimization. We develop a stress-constrained topology optimization method for lattice-like structures that uses component-wise reduced order models as a cheap surrogate, providing accurate computation of stress fields while greatly reducing run time relative to a full order model. We demonstrate the ability of our method to produce large reductions in mass while respecting a constraint on the maximum stress in a pair of test problems. The ROM methodology provides a speedup of about 150x in forward solves compared to full order static condensation and provides a relative error of less than 5% in the relaxed stress.

Keywords— Topology optimization, model reduction, substructuring, static condensation, stress constraint, ground structure

1 Introduction

Lattice-like structures are advantageous in many applications due to their combination of high stiffness and light weight. Applications include aerospace structures \cite{26}, medical uses such as prostheses or implants \cite{41}, and the traditional use of trusses as supports in structural engineering. Interest in lattice-like structures has further increased as advances in additive manufacturing enable their production. Their analysis using a standard finite element method (FEM) requires a high-dimensional discretization to capture the complex geometry, especially when accurate computation of localized quantities – for example, stress – is needed. Optimization is then expensive, requiring many evaluations of this high-dimensional model. In this work, we present a stress-constrained topology optimization (TO) formulation for lattice-like structures that leverages component-wise reduced order models (ROMs) as a surrogate, providing accurate computation of the stress field while greatly reducing computational cost relative to a FEM analysis.

TO of lattice-like structures is often based on asymptotic homogenization. Such techniques are compelling when the final design should be composed of material with a periodic structure. These techniques may solve a macroscopic design problem in which the optimization variables control the geometry of unit cells throughout the domain \cite{4}; a microscopic design problem, designing a unit cell with particular properties as first seen in \cite{31}; or a combination of the two as in \cite{12}. Homogenization has been applied to stress-constrained problems, for example in \cite{7}. Other methods also assume a periodic structure, for example, the generalized and high-fidelity generalized method of cells \cite{1}. ROMs have been applied to this family of techniques \cite{28, 39, 40}: however, together with homogenization, these methods share limiting assumptions on length scale and periodicity. The length scale of unit cells should be much less than that of macroscopic features, and the approximation may not be accurate if material does not satisfy the periodicity assumption. In particular, the behavior of stress near the boundary of the material is not well understood \cite{7}.

\textsuperscript{*}Corresponding author; sean.mcbane@utexas.edu
Several other existing techniques in TO also apply to design of lattice-like structures. Ground structure approaches for truss optimization rely on a beam model for truss members and make the optimization variables the cross-sectional areas of truss members. This approximation is invalid when truss members become too thick. Deng and To in [8] develop a method that projects the ground structure onto a background finite element mesh, eliminating the thin beam assumption. In a similar vein, the family of methods related to moving morphable components (MMC) [15, 25] represent a structure as a collection of discrete geometric components whose shape and position are controlled by the optimization parameters, and model the resulting structure by projecting component geometries on a background mesh. Zhang, Gain, and Norato in [43] and Zhang et al. in [44] apply such projection-based methods to solve stress-constrained TO problems. The shared shortcoming of these approaches for our purposes is the requirement of a sufficiently fine background mesh to capture small-scale structural features accurately. Substructuring methods accelerate solution of a high-resolution model by rewriting equations in terms of a reduced set of degrees of freedom, alleviating the computational cost associated with a fine discretization. In [24], we apply substructuring-based model order reduction in the form of port-reduced static condensation (PRSC) to compliance minimization. Wu et al. solve the same problem using a related approach in [38]; [37] and [21] are additional examples of substructuring-based model order reduction to TO. Finally, there is an extensive body of work applying conventional (i.e., density-based or level set) TO methods to stress-constrained problems, and these approaches could be used to design lattice-like structures given a fine enough discretization. See, for example, [22, 16] for examples in density-based optimization to which we refer in our approach, and [15, 27] for examples of level set approaches to stress-based optimization. Methods based on augmented Lagrangian techniques to address stress constraints without aggregation are promising due to their ability to directly enforce a maximum stress constraint locally; see, for example, [30], and [32] for a comparison of the performance of aggregation vs. local strategies. Using PRSC in conjunction with an augmented Lagrangian optimization strategy is possible, but outside the scope of this work.

In this paper we leverage PRSC to efficiently solve stress-constrained TO problems, which are not addressed in the previous literature on substructuring for TO. Our approach is both a ground structure and a substructuring method. We constrain the design space by choosing a fixed arrangement of subdomains (“components”) of which all possible designs are a subset. We apply PRSC to construct a surrogate model parameterized by a component-wise density parameter, which penalizes stiffness using the SIMP scheme [5]. Then, we minimize the mass of the structure subject to a constraint on the maximum stress implemented using the $\text{qp}$-relaxation of Bruggi [6] to address the “singularity problem” [29], and stress aggregation in the form of the Kreisselmeir-Steinhauser (KS) functional to convert the infinite-dimensional stress constraint to a small number of differentiable constraints that approximate the max function. Constraints are defined by aggregating over non-overlapping aggregation domains, as in [22, 16]. Finally, after optimization we postprocess by removing components with densities less than a prescribed minimum value. Due to the component-wise formulation this postprocessing results in a well-defined geometry without additional postprocessing steps, unlike the result from an element-wise density based TO. Both during optimization and in the postprocessed design, we avoid stress singularities by carefully choosing components such that the domain boundary is smooth. Our original contributions are:

- The component-wise formulation of stress constraints, and in particular the heuristics used to achieve a conservative optimization result without adaptively modifying the constraint definition;
- Rewriting of stress constraints using component-wise ROM operators for an additional computational speedup;
- Sensitivity analysis of the stress constraints in the context of the component-wise model; and
- Our postprocessing method.

The remainder of this paper is organized as follows. In Section 2, we provide a brief overview of PRSC, introducing the notation required for the rest of our discussion. Section 3 describes our optimization formulation: the material model, constraint formulation, objective function, and postprocessing methodology. Following this description, in Section 4 we present mass minimization results for an L-bracket and a cantilever beam geometry, along with studies of ROM accuracy and performance. Finally, we give our conclusions.

2 An overview of port-reduced static condensation

Port-reduced static condensation (PRSC) is developed in [10] and further discussed in [2, 33, 34, 17, 24] and others. The discussion below is a high-level overview of the technique in which we seek to introduce only the
concepts and notation needed to describe the use of PRSC for component-wise TO and perform the required sensitivity analysis.

PRSC applies to solution of an elliptic PDE on a domain \( \Omega \subset \mathbb{R}^n \) (where \( n = 2 \) or \( 3 \) in general, but we focus on \( n = 2 \) here) given in weak form by

\[
a(u, v; \mu) = f(v; \mu), \quad \forall v \in X
\]

(1)

Here, \( a : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R} \) is a bilinear form, taken to be coercive so that the problem (1) is well-posed; \( f : H^1(\Omega) \rightarrow \mathbb{R} \) is a linear form; \( \mu \in M \) is a vector in \( M \subset \mathbb{R}^{n_m} \), parametrizing both the linear and bilinear forms; \( X \subset H^1(\Omega) \) is a finite dimensional function space arising from a finite element discretization of (1) and incorporating essential boundary conditions; and \( u, v \in X \) are the solution to (1) and a test function, both residing in \( X \).

PRSC reduces the number of degrees of freedom in (1) by applying static condensation, before further reducing the problem dimension via projection-based model reduction. \( \Omega \) is decomposed into \( n_c \) subdomains \( \Omega_i \), \( i \in \{1, \ldots, n_c\} \), termed “components”. The bilinear and linear forms may now be decomposed as

\[
a(u, v; \mu) = \sum_{i=1}^{n_c} a_i (u_i, v_i; \mu_i)
\]

(2)

\[
f(v; \mu) = \sum_{i=1}^{n_c} f_i (v_i; \mu_i)
\]

(3)

where \( u_i \) and \( v_i \) are the restrictions of \( u \) and \( v \) to \( \Omega_i \); \( a_i \) and \( f_i \) are the restrictions of \( a \) and \( f \) to act on functions in \( X|_{\Omega_i} \); and \( \mu_i \subset \mathbb{R}^{n_{m_i}} \), with \( n_{m_i} \) the dimension of the parameter space for component \( i \), is a parameter vector containing only those parameters in \( \mu \) to which \( a_i \) and \( f_i \) are sensitive.

Each component is assigned a set of \( n_i \) “ports”, \( \{\gamma_{i,j}, j \in 1, \ldots, n_i\} \), which are subsets of the boundary of \( \Omega_i \) where \( \Omega_i \) may (but is not required to) interface with another component \( \Omega_{i'} \). That is, if \( \Omega_i \cap \Omega_{i'} \) is non-empty, it corresponds to port \( \gamma_{i,j} \) and \( \gamma_{i',j'} \) respectively on \( \Omega_i \) and \( \Omega_{i'} \). Ports may also be unconnected to any neighboring component, or have a Dirichlet boundary applied. Ports on the same component are assumed to be disjoint.

Now \( \Omega \) is fully defined by specifying the component domains \( \Omega_i \), \( i \in \{1, \ldots, n_c\} \), and the connections between them in the form of port pairs \( \{\gamma_{i,j}, \gamma_{i',j'}\} \). Although the members of this port pair refer to subsets of \( \Omega_i \) and \( \Omega_{i'} \) (“local ports”) respectively, since they coincide in space we also refer to them as a single “global port”. A single, unconnected port also corresponds to a single global port. Global ports are denoted by \( \Gamma_p \), \( p \in \{1, \ldots, n^f\} \) and defined by an index set \( \pi_p = \{(i, j), (i', j')\} \), for connections of two local ports, or \( \pi_p = \{(i, j)\} \) for disconnected ports. The correspondence between global and local ports is illustrated in Fig. 1. We also define a mapping \( G_i \) such that given a local port \( \gamma_{i,j} \), \( G_i(j) \) maps \( j \) to the index \( p \) of the global port \( \Gamma_p \).

The restriction of \( X \) to \( \gamma_{i,j} \) (its “port space”) is denoted \( P_{i,j} \). Where a global port \( \Gamma_p \) is the coincidence of two local ports, \( \pi_p = \{(i, j), (i', j')\} \), their port spaces are identical: \( P_{i,j} = P_{i',j'} \). Letting the

\[
\Gamma_1; \quad \pi_1 = \{(1,1)\}
\]

\[
\Gamma_2; \quad \pi_2 = \{(1,2)\}
\]

\[
\Gamma_3; \quad \pi_3 = \{(1,3), (2,1)\}
\]

\[
\Gamma_4; \quad \pi_4 = \{(1,4)\}
\]

\[
\Gamma_5; \quad \pi_5 = \{(2,2)\}
\]

\[
\Omega_1 \quad \Omega_2
\]

\[
\Omega\]

\[
\text{Figure 1: Illustration of PRSC nomenclature in a two-component system; red segments of the boundary are ports, with their global index and index set shown.}
\]
dimension of \( P_{i,j} \) be \( n_i^7 \), define a basis \( \{ \chi_{i,j,k} \in P_{i,j}, k \in \{1, \ldots, n_i^7\} \} \) for \( P_{i,j} \), and a “lifted port basis” \( \{ \psi_{i,j,k} \in X|\Omega_i, k \in \{1, \ldots, n_i^7\} \} \), where \( \psi_{i,j,k}|_{\gamma_{i,j}} = \chi_{i,j,k} \) and \( \psi_{i,j,k}|_{\gamma_{i,j'}} = 0, j' \neq j \). We lift the port bases by solving

\[
\begin{align*}
  a_i(\psi_{i,j,k}, v; \mu_i) &= 0, \quad \forall v \in B_i \\
  \psi_{i,j,k} &= \chi_{i,j,k} \text{ on } \gamma_{i,j} \\
  \psi_{i,j,k} &= 0 \text{ on } \gamma_{i,j'}, j' \neq j.
\end{align*}
\]

(4)

The global lifted port basis associated to \( \Gamma_p \) is given by \( \Psi_{p,k} = \psi_{i,j,k} + \psi_{i',j',k} \) if \( \pi_p = \{(i,j),(i',j')\} \), or \( \Psi_{p,k} = \psi_{i,j,k} \) if \( \pi_p = \{(i,j)\} \), where port basis functions are extended by zero outside of their domain.

In addition to port spaces, each component has a “bubble space” defined as

\[
B_i \equiv \{ v \in X_i : v|_{\gamma_{i,j}} = 0, j \in \{1, \ldots, n_i^7\} \}.
\]

(5)

Define two classes of parameter dependent “bubble functions” by solution of the equations

\[
\begin{align*}
  a_i(b_{i,j,k}(\mu_i), v; \mu_i) &= -a_i(\psi_{i,j,k}, v; \mu_i), \quad \forall v \in B_i \\
  a_i(b_i^f(\mu_i), v; \mu_i) &= f_i(v; \mu_i), \quad \forall v \in B_i
\end{align*}
\]

(6)

(7)

where \( b_{i,j,k}, b_i^f \in B_i \). With these definitions, the solution to (1) on \( \Omega_i \) may be written as:

\[
u(\mu)|_{\Omega_i} = b_i^f(\mu_i) + \sum_{j=1}^{n_i^7} \sum_{k=1}^{n_i^7} \sum_{\gamma_{i,j,k} \in \pi_p} u_{G_{i,j,k}}(\mu)(b_{i,j,k}(\mu_i) + \psi_{i,j,k})
\]

(8)

with \( U_{p,k} \) unknown coefficients.

Defining interface functions \( \phi_{i,j,k}(\mu_i) \) by

\[\phi_{i,j,k}(\mu_i) \equiv \psi_{i,j,k} + b_{i,j,k}(\mu_i)\]

(9)

and global interface functions \( \Phi_{p,k}(\mu) \) by

\[
\Phi_{p,k}(\mu) \equiv \sum_{(i,j) \in \pi_p} \phi_{i,j,k}(\mu_i),
\]

(10)

the global solution to (1) may be written as

\[
u(\mu) = \sum_{i=1}^{n_v} b_i^f(\mu_i) + \sum_{p=1}^{n_p} \sum_{k=1}^{n_i^7} U_{p,k}(\mu) \Phi_{p,k}(\mu).
\]

(11)

Substituting (11) in (1) and restricting test functions to the space spanned by the interface functions leads to the condensed system of equations

\[
K(\mu)U(\mu) = F(\mu),
\]

(12)

where the entries of \( K \) and \( F \) are

\[
K_{(p,k),(p',k')} = a(\Phi_{p,k}(\mu), \Phi_{p',k'}(\mu) ; \mu)
\]

(13)

and

\[
F_{(p,k)} = f(\Phi_{p,k}(\mu) ; \mu) - \sum_{i=1}^{n_v} a(b_i^f(\mu_i), \Phi_{p,k}(\mu) ; \mu).
\]

(14)

A pair \( (p,k) \) identifies a single degree of freedom in the linear system (12): the coefficient of \( \Phi_{p,k} \) in (11). The Schur complement consists of local contributions from each component, analogous to element stiffness matrices in a FEM. These contributions are defined by

\[
K^i_{(j,k),(j',k')} = a_i(\phi_{i,j,k}(\mu_i), \phi_{i,j',k'}(\mu_i) ; \mu_i).
\]

(15)
We make use of this definition in sensitivity analysis. In this description we have referred only to a collection of \( n_c \) different subdomains \( \Omega \) and termed these “components”. Each instantiated component is associated to a corresponding reference component, where the number of reference components is small, but \( n_c \) may be very large. An offline data set needs only to be constructed for the small number of reference components, and is used to construct the linear system \( [\mathbf{L}] \) using efficient operations on the offline data set. To simplify the presentation of PRSC we have avoided introducing separate notation for instantiated and reference components and describing the correspondence between the two, but these details are key to the efficient implementation of the method.

This description of PRSC so far only includes static condensation. Model reduction is incorporated by using a reduced port basis \( \tilde{\chi}_{i,j,k} \in \tilde{P}_{i,j} \), with reduced port basis dimension \( \tilde{n}_{\gamma i,j} < n_{\gamma i,j} \), and corresponding lifted port basis functions \( \tilde{\psi}_{i,j,k} \). In presenting the Schur complement system, we have used the notation for the full order model; however, the reduced order system may be derived by substituting the reduced lifted port basis and its dimension wherever the full order lifted port basis and its dimension appear. Henceforth, ROM quantities are denoted by adding a tilde to their FOM counterpart; for example, \( \tilde{u} \) is the ROM solution for displacement. We determine an appropriate reduced port basis using the pairwise training procedure described in \cite{Ref10, Algorithm 2} and proper orthogonal decomposition with respect to the \( L_2^2 \) inner product on each port.

3 A component-wise formulation for topology optimization with stress constraints

Here we present the main contribution of this work: a method for stress-based TO that leverages PRSC to enable an efficient optimization, even when using a high-resolution discretization. With the use of PRSC, we also inherit its greatest advantage, which is the ability to reuse the component library constructed in the offline phase to model any arrangement of instantiated components. This allows the use of the same offline dataset to solve many TO problems in various geometries.

Below, we describe our methodology in detail: the formulation of the optimization problem, the material model and parameterization, the form of the aggregated stress constraints, and finally, postprocessing considerations.

3.1 Optimization formulation

Our goal is to solve the problem

\[
\begin{align*}
\text{minimize} \quad & J(\rho) = \sum_{i=1}^{n_c} \rho_i |\Omega_i| \\
\text{subject to} \quad & \max_{\Omega} \sigma(u) \leq \sigma_{\max} \\
& a(u,v;\rho) = f(v), \quad \forall v \in X, 
\end{align*}
\]

where \( \rho \in \{0,1\}^{n_c} \) is the parameter vector, containing a density for each component in the system with \( \rho_i = 0 \) meaning that component \( i \) is removed and \( \rho_i = 1 \) meaning that component \( i \) is fully solid; \( |\Omega_i| \) is the volume of \( \Omega_i \); \( J(\rho) \) gives the mass of the system; \( u \) is displacement; and \( \sigma \) is some yield criterion. The bilinear and linear forms are defined in the next section in a component-wise fashion, and the global bilinear and linear forms given by Eqs. 2 and 3. In this work, \( \sigma \) is the von Mises yield stress. \( \sigma_{\max} \) is an upper bound on the allowed value of the yield criterion \( \sigma \).

To render (16) solvable by a gradient-based optimization solver, we make two changes following common practice. The first is to allow the component densities to reside in a continuum: \( \rho \in [0,1]^{n_c} \). Material properties for intermediate densities are interpolated by a standard SIMP scheme. Second, because of the non-differentiable nature of the max function, we replace the constraint in (16) with a differentiable approximation. With these
modifications, the optimization formulation becomes:

$$\min_{\rho \in [0,1]^n_c} J(\rho) = \sum_{i=1}^{n_c} \rho_i |\Omega_i|$$

subject to \( g_m(u, \rho) \leq 1 \), \( 1 \leq m \leq n_{agg} \)

$$a(u, v; \rho) = f(v), \ \forall v \in X,$$

with the aggregate constraints \( g_m(u, \rho) \) defined in Section 3.3.2. This is the optimization formulation that we actually solve. A local minimum of (17) provides an approximation to solution of (16), but the challenging integer programming problem has been replaced by one solvable using gradient-based nonlinear optimization solvers.

### 3.2 Material model & component-based parameterization

We model a material described by isotropic linear elasticity. On each component, the bilinear form \( a_i \) is given by

$$a_i(u, v; \rho_i) = \int_{\Omega_i} s(\rho_i) C(u) \cdot \nabla v \ dx,$$

where \( s \) is the SIMP penalization term and \( C \) is the symmetric elasticity tensor given by

$$C(u) = \mu (\nabla u + \nabla u^T) + \lambda (\nabla \cdot u) I,$$

with \( \lambda \) and \( \mu \) the Lamé parameters. Note that \( \mu \) is a fixed material parameter and bears no relation to the vector of component parameters \( \mu_i \) from the discussion of PRSC. In our optimization formulation \( \mu_i = [\rho_i] \), so we use the scalar parameter \( \rho_i \) in its place. This form for \( a_i \) satisfies the necessary assumptions to implement the affine simplification of PRSC described in our previous work [24]. This simplification allows us to compute Schur complement contributions offline, making the model even more economical than PRSC without this simplification.

The linear form \( f_i \) is

$$f_i(v) = \int_{\Omega_i} F \cdot v \ dx,$$

where \( F \) is the vector of forces applied per unit volume. We assume \( F \) to be independent of parameter in this work; this assumption is not necessary, but simplifies the sensitivity analysis.

We aim to obtain a “black and white” solution, \( \rho_i \in \{0,1\} \). To drive the continuous density values to black and white solutions, we use the solid isotropic material (SIMP) parameterization [3, 46]. The Young’s modulus is penalized using a power law in the density: \( E_i \propto \rho_i^p \), with \( E_i \) the Young’s modulus of the material in \( \Omega_i \). In our optimizations, we choose \( p = 3 \). To ensure a well-posed problem, the SIMP parameterization is modified so that \( s(\rho_i) \) in (18) is given by

$$s(\rho_i) = [\rho_i + (1 - \rho_i) \rho_{min}]^3,$$

where \( \rho_{min} > 0 \) is chosen to be small \( (10^{-3}) \). This ensures that the linear system (12) possesses a solution, but the stiffness of components with \( \rho_i = 0 \) is negligible compared to components with \( \rho_i = 1 \).

This parameterization is the same as a typical topology optimization using element-wise density variables, but enables the use of a component-wise ROM to reduce the computational cost because each component ROM only depends on a single density parameter. In a conventional, element-based density TO, the large dimension of parameter space makes model reduction impractical. The component-wise scheme has an additional benefit: it obviates the need for a density filter (as in [22, 16, 30, 23] and others) to avoid checkerboarding. The imposition of a ground structure intrinsically imposes a length scale without the use of a filter.

### 3.3 Imposing stress constraints using the ROM

Our formulation of stress constraints consists in three parts: stress relaxation, aggregation using the KS functional, and rewriting of the aggregate constraints using ROM operators. Of these parts, the first two are standard practice; the rewriting using ROM operators is an original contribution.

We note here that while the discussion below assumes the use of the von Mises yield criterion, our methodology is not limited by this choice. Other yield criteria may be used in its place, with the appropriate corrections to the derivations in Sections 3.3.4 and 3.3.5 where the form of the von Mises criterion is used explicitly. Criteria which contain a non-differentiable function such as the absolute value may require rewriting as multiple constraints to make the constraint formulation differentiable.
3.3.1 Stress relaxation

It is well known that stress constraints in the absence of some relaxation result in an optimization problem where minima are contained in a degenerate subspace (the “singularity problem”) \[20\]. As remarked in \[9\], this occurs where material in some region must vanish to reach a local optimum, but the material remains strained so that solutions for intermediate densities in that region violate constraints. When material has vanished, however, these constraints should be ignored – the constraint is discontinuous. Stress relaxation smooths this discontinuity to allow a gradient-based optimization algorithm to reach solutions in the degenerate subspace.

We adopt the \(qp\)-relaxation proposed by Bruggi \[6\], with the particular choice of \(q = 2.5\) as in \[22, 16\].

Denoting the von Mises yield stress by \(\sigma_{vm}\), the relaxed stress is then given by

\[
\sigma_r(u, \rho)|_{\Omega_i} = s'(\rho_i)\sigma_{vm}(u_i),
\]

(22)

where \(\sigma_{vm}\) is the von Mises stress as computed using the stiffness of the base material (not the SIMP penalized stiffness), and

\[
s'(\rho_i) = \rho_i^{1/2}.
\]

(23)

We note that the optimization has the trivial solution \(\rho = 0\); in practice, this solution is not reached by the optimizer. Convergence to the trivial solution, if problematic, could be addressed by imposing a lower bound on \(s'(\rho_i)\) in the same fashion as in (21). Applying stress relaxation, the stress constraint becomes

\[
\max_{\Omega} \sigma_r(u, \rho) \leq \sigma_{max},
\]

(24)

which is identical to the constraint in \[10\] for black and white solutions.

3.3.2 Stress aggregation

We cannot use the constraint (24) in a gradient-based optimization because the max function is non-differentiable. To circumvent this difficulty, we use stress aggregation to provide a differentiable approximation of the max function. Aggregation strategies include the \(p\)-norm, the Kreisselmeier-Steinhauser (KS) functional, or an “induced aggregation” functional as described by Kennedy and Hicken \[19\]. We use a continuous KS aggregation, with stresses aggregated over multiple aggregation domains (as in \[16, 22\], and others). For numerical stability in finite precision arithmetic, we aggregate the ratio of the relaxed stress to the maximum stress, rather than the relaxed stress itself. The single constraint (24) becomes \(n_{agg}\) constraints given by

\[
g_m(u, \rho) = \frac{1}{p} \ln \left( \frac{1}{\alpha} \int_{\Omega_m^{agg}} \exp \left( p \frac{\sigma_r(u, \rho)}{\sigma_{max}} \right) dx \right) \leq 1, \ 1 \leq m \leq n_{agg},
\]

(25)

where \(p\) and \(\alpha\) are fixed parameters of the aggregation. Increasing \(p\) results in a closer approximation of the max function, but also a more difficult optimization problem. \(\alpha\) is a normalization, whose determination we address below. The aggregation domains \(\Omega_m^{agg}\) are not the same as component domains \(\Omega_i\), nor even spatially contiguous domains. Their determination is addressed in the next section. Ideal values of \(p\) and \(n_{agg}\) are problem dependent, and studied in our numerical experiments.

The normalization \(\alpha\) is computed based on the following observation from \[19\] Eq. 8: for given values of \(\rho\) and \(p\), we will have

\[
g_m(u, \rho) \geq \frac{\sigma_r(u, \rho)}{\sigma_{max}}
\]

if

\[
\alpha \leq \int_{\Omega_m^{agg}} \exp \left( \frac{p}{\sigma_{max}} \left[ \sigma_r - \max_{\Omega_m^{agg}} \sigma_r \right] \right) dx.
\]

(26)

Let the initial value of \(\rho\) in the optimization be \(\rho_0\), and the displacement solution for this density field be \(u^0\). Using (26), we compute values of \(\alpha\) for each \(\Omega_m^{agg}\) domain at the initial condition of the optimization as follows: compute

\[
\alpha_k = \exp \left( \frac{p}{\sigma_{max}} \left[ \max_{\Omega_m^{agg}} \sigma_r \left( u^0, \rho_0 \right) \right] \right)^{-1} \int_{\Omega_m^{agg}} \exp \left( \frac{p}{\sigma_{max}} \sigma_r \left( u^0, \rho_0 \right) \right) dx,
\]

(27)

then set \(\alpha\) for all aggregates to be equal to the minimum of all \(\alpha_k\).
Figure 2: Illustration of aggregation region assignment for a single component and five aggregation regions. Elements contained in the same aggregation region are assigned the same color.

3.3.3 Achieving a conservative optimization result

The aggregation described in the previous section is not necessarily conservative; for some values of $\rho$, the constraints (25) may be satisfied while the true max stress constraint (24) is not. To consider an optimization successful, we require satisfaction of the latter. In most cases determining $\alpha$ from Eq. (27) is sufficient to obtain a conservative optimization solution; where this is not the case, we use a simple heuristic: substitute $\hat{\sigma}_{\text{max}} < \sigma_{\text{max}}$ for $\sigma_{\text{max}}$ in (25). The choice of $\hat{\sigma}_{\text{max}}$ is problem dependent. Values used are reported in the numerical experiments – for most, $\hat{\sigma}_{\text{max}} = \sigma_{\text{max}}$.

This heuristic approach has mathematical justification. Eq. (25) may be rewritten as

$$g_m(u, \rho) = \frac{1}{p} \ln \left[ \int_{\Omega_{agg}} \exp \left( p \frac{\sigma_r(u, \rho)}{\sigma_{\text{max}}} \right) dx \right] - \frac{n \alpha}{p} - 1 \leq 0.$$  \hfill (28)

Thus if $\alpha < e$, the value of the constraint is increased by a constant factor. A decrease in the upper bound in the optimization likewise corresponds to a constant increase in the value of $g_m$; therefore the imposition of $\hat{\sigma}_{\text{max}} < \sigma_{\text{max}}$ may be viewed as a correction to the value of $\alpha$ estimated from (26), necessary because that value is conservative only for a particular value of $\rho$.

We also assign aggregation regions differently from previous works. The two main strategies are to either distribute stress evenly among aggregation regions [22], or to assign the highest stresses to a single aggregation region, resulting in a better approximation of the maximum in that region [16]. However, because we do not adaptively reassign regions and cannot know how stresses will be distributed at a given optimization iteration, neither of these approaches is possible here. Instead, we approximate the effect of an approach that computes an even distribution of stresses by assigning aggregation regions randomly. For each $\Omega$, we randomly assign an equal number of elements to each of the $n_{agg}$ aggregation domains $\Omega_{agg}^{m}$ (illustrated in Fig. 2). Ideally, this assignment of aggregation regions results in a problem where stresses are evenly distributed between the different aggregation regions. We show results on the same problem for different assignments of aggregation regions in Section 4.3.2.

3.3.4 Efficient computation of aggregates using the CWROM

The aggregation in Eq. (25) is nonlinear in $\sigma_{vm}$, which is in turn a nonlinear function of displacement. This nonlinearity prohibits the expression of the constraints in terms of $\hat{\bar{u}}(\rho)$ without reconstructing the displacement field $\bar{u}(\rho)$. However, computing the constraints by reconstructing $\bar{u}(\rho)$ and performing quadrature in a loop
over elements is inefficient because of the high dimension of the underlying discretization. For example, in our
smaller numerical example of Section 4.2, the mesh underlying the component-wise ROM has about 2.5 million
elements. Integration over each element requires retrieving coefficients of the finite element basis functions, which
are generally not contiguous in memory, and computing values of the non-linear integrand as a function of these
coefficients. While it is not possible to remove the dependence of the cost of aggregation on the FOM dimension,
we reduce the cost by expressing aggregation in terms of more cache- friendly operations using \( \tilde{\rho} \) instead of
\( \varrho(\rho) \).

We rewrite the constraints in terms of the ROM by noting that \( \sigma_{\text{com}} \), which is not linear in displacement,
may be written as a function of the Cauchy stress tensor components. Limiting ourselves to two dimensions, we define
vectors containing the values of stress tensor components at quadrature points for aggregation region \( \Omega_{\text{agg}} \). We assume
that there are no forces applied except on ports. Then, from the definition of the Cauchy stress tensor and
Eq. (11), the linear operators \( S_{\text{xx}}, S_{\text{yy}}, \) and \( S_{\text{xy}} \) are given by:

\[
S_{\text{xx}} = (2\mu + \lambda) \tilde{\Phi}^{\text{m}}_{\text{xx}} + \lambda \tilde{\Phi}^{\text{m}}_{\text{yy}}
\]

\[
S_{\text{yy}} = (2\mu + \lambda) \tilde{\Phi}^{\text{m}}_{\text{yy}} + \lambda \tilde{\Phi}^{\text{m}}_{\text{xx}}
\]

\[
S_{\text{xy}} = \mu \left( \tilde{\Phi}^{\text{m}}_{\text{xx,y}} + \tilde{\Phi}^{\text{m}}_{\text{y,x}} \right)
\]

with \( \tilde{\Phi}^{\text{m},k}_{x,x} \) a matrix such that \( \left[ \tilde{\Phi}^{\text{m},k}_{x,x} \right]_{j,k} \) contains the partial derivative with respect to \( x \) of the \( k \)-th interface function at the \( j \)-th quadrature point in \( \Omega_{\text{agg}} \), and \( \tilde{\Phi}^{\text{m}}_{\text{y,y}}, \tilde{\Phi}^{\text{m}}_{\text{y,x}}, \) and \( \tilde{\Phi}^{\text{m}}_{\text{x,x}} \) defined similarly.

In terms of the stress tensor component vectors \( \sigma^{\text{m}} \), the vector of von Mises stresses at quadrature points in
\( \Omega_{\text{agg}} \) is given by

\[
\sigma^{\text{m}}_{\text{VM}} \left( \tilde{\rho}(\rho) \right) = \left( \sigma^{\text{m}}_{\text{xx}} \odot \sigma^{\text{m}}_{\text{xx}} + \sigma^{\text{m}}_{\text{yy}} \odot \sigma^{\text{m}}_{\text{yy}} - \sigma^{\text{m}}_{\text{xx}} \odot \sigma^{\text{m}}_{\text{yy}} + 3\sigma^{\text{m}}_{\text{xy}} \odot \sigma^{\text{m}}_{\text{xy}} \right)^{1/2},
\]

where \( \odot \) denotes the Hadamard product and the exponentiation is applied element-wise. We omit the functional
dependence of \( \sigma^{\text{m}}_{\text{xx}} \), etc. on \( \tilde{\rho}(\rho) \) for brevity. The vector of relaxed stresses at quadrature points is then

\[
\sigma^{\text{r}} \left( \rho, \tilde{\rho}(\rho) \right) = s' \left( \rho^{\text{agg}} \right) \odot \sigma^{\text{m}}_{\text{VM}},
\]

where \( \rho^{\text{agg}} \) contains the value of the density parameter at each quadrature point and \( s' \) is the relaxed stress \( \sigma^{\text{r}} \),
applied elementwise.

Taking \( \mathbf{w} \) to be a vector containing the coefficients for quadrature over \( \Omega_{\text{agg}} \), we can rewrite the constraints
in Eq. (25):

\[
g_{\text{m}} \left( \rho, \tilde{\rho}(\rho) \right) = \frac{1}{p} \ln \left( \frac{1}{\alpha} \mathbf{w}^T \left[ \exp \left( \frac{p}{\sigma_{\text{max}}} \sigma^{\text{m}} \left( \rho, \tilde{\rho}(\rho) \right) \right) \right] - 1 \right),
\]

with exponentiation interpreted elementwise.

Although this expression improves the performance of aggregation by rewriting it in a more cache-friendly
manner, the size of the stress operators in Eq. (30) grows with the number of quadrature points and the dimension
of the ROM, until at some point it is more efficient to compute the aggregates in the obvious fashion. In our
numerical examples, we use a four-point quadrature rule exact for second-order polynomials. Also, because we
assign aggregation regions randomly, the stress operators in (30) cannot be computed offline; instead they are
computed once at the beginning of the optimization. The cost of this computation is insignificant relative to the
total cost of optimization.

To further reduce the cost of computing aggregates, we investigated the use of a second-order Taylor series
expansion to approximate the inner integral in Eq. (25) on each component. This approach is highly effective at
reducing the time to compute constraints, but the approximation is not accurate enough to effectively constrain the
maximum value of stress. Higher order Taylor series approximations quickly become impractical due to excessive
memory requirements. More sophisticated hyperreduction approaches are not investigated here, but could be
used in future work to harness the full potential of the component-wise ROM approach for stress-constrained
optimization.
3.3.5 Sensitivity analysis of stress constraints

To derive the gradient of the constraint $g_m$ with respect to $\rho$, we note that $g_m(\rho)$ depends both directly on $\rho$ and implicitly through $\tilde{U}(\rho)$, with the latter dependence defined by the forward model. Therefore, the gradient is given by a total derivative:

$$\nabla_\rho g_m(\rho, \tilde{U}(\rho)) = \frac{\partial g_m}{\partial \rho} + \left( \frac{d\tilde{U}}{d\rho} \right)^T \frac{\partial g_m}{\partial \tilde{U}},$$

where the notation $\frac{d\tilde{U}}{d\rho}$ indicates the Jacobian matrix with

$$\left[ \frac{d\tilde{U}}{d\rho} \right]_{i,j} = \frac{\partial \tilde{U}_i}{\partial \rho_j}.$$

The action of $\left( \frac{d\tilde{U}}{d\rho} \right)^T$ is computed using the adjoint method. From Eq. (12), we have

$$\frac{d\tilde{K}}{d\rho} + \tilde{K} \frac{d\tilde{U}}{d\rho} = \frac{d\tilde{F}}{d\rho}.$$  

In our problem settings the forcing is independent of parameter and the above yields

$$\frac{d\tilde{U}}{d\rho} = -\tilde{K}^{-1} \left[ \frac{d\tilde{K}}{d\rho} \tilde{U} \right].$$

By substitution in Eq. (34), we obtain the gradient:

$$\nabla_\rho g_m(\rho, \tilde{U}(\rho)) = \frac{\partial g_m}{\partial \rho} - \left[ \frac{d\tilde{K}}{d\rho} \tilde{U} \right]^T \lambda,$$

where, making use of the symmetry of $\tilde{K}(\rho)$, $\lambda$ is given by solution of the adjoint equation

$$\tilde{K}(\rho)\lambda(\rho) = \frac{\partial g_m}{\partial \tilde{U}}(\rho).$$

Note that $\frac{d\tilde{K}}{d\rho}$ is a third-order tensor; for this sensitivity analysis, its product with $\tilde{U}$ is defined by:

$$\left[ \frac{d\tilde{K}}{d\rho} \right]_{ij} = \sum_k \frac{d\tilde{K}_{ik}}{d\rho_j} \tilde{U}_k.$$

To close Eq. (34), we require expressions for the partial derivatives of $g_m$. These are found using the chain rule:

$$\frac{\partial g_m}{\partial \rho} = \left( \frac{\partial \sigma_r^m}{\partial \rho} \right)^T \frac{\partial g_m}{\partial \sigma_r^m},$$

$$\frac{\partial g_m}{\partial \tilde{U}} = \left( \frac{\partial \sigma_r^m}{\partial \tilde{U}} \right)^T \frac{\partial g_m}{\partial \sigma_r^m}.$$

Defining the vector

$$e \left( \rho_i, \tilde{U}(\rho) \right) = \exp \left[ \frac{P}{\sigma_{\max}} \sigma_r^m \left( \rho_i, \tilde{U}(\rho) \right) \right],$$

we find from Eq. (33) that $\frac{\partial g_m}{\partial \sigma_r^m}$ is

$$\frac{\partial g_m}{\partial \sigma_r^m} = \frac{1}{\sigma_{\max}} \sigma_r^m \sigma_r^m \sigma_r^m e \left( \rho, \tilde{U}(\rho) \right) \odot w.$$
Finally, we require the derivatives of $\sigma^m$ in Eqs. (40) and (41). With respect to $\rho$, obtain

$$\frac{\partial \sigma^m}{\partial \rho} = P^m \left( \frac{ds'}{d\rho} (\rho_{agg}^m) \odot \sigma^m_{VM} \right) \odot I,$$

with $P^m$ a prolongation operator taking values in $\rho_{agg}^m$, to their corresponding entry in $\rho$ and padding with zeros for entries in $\rho$ not corresponding to elements in $\Omega_{agg}^m$, and $\odot$ denoting a row-wise Kronecker product.

With respect to $\tilde{U}$, the Jacobian of $\sigma^m$ is given by

$$\frac{\partial \sigma^m}{\partial \tilde{U}} = s' (\rho_{agg}^m) \odot \frac{\partial \sigma^m_{VM}}{\partial \tilde{U}},$$

(45)

with $\frac{\partial \sigma^m_{VM}}{\partial \tilde{U}}$ given by

$$\frac{\partial \sigma^m_{VM}}{\partial \tilde{U}} = \frac{1}{2} (\sigma^m_{VM})^{-1} \otimes \left[ 2 \sigma^m_{zx} \otimes S^m_{zx} + 2 \sigma^m_{yy} \otimes S^m_{yy} - \sigma^m_{zx} \otimes S^m_{yy} - \sigma^m_{yy} \otimes S^m_{zx} + 6 \sigma^m_{xy} \otimes S^m_{xy} \right].$$

(46)

### 3.4 Postprocessing

SIMP penalization does not necessarily succeed at creating black and white solutions. Therefore, we postprocess by removing components with $\rho_i$ less than a prescribed minimum value $\rho_{min}$ from the domain entirely, and setting the densities of remaining components to unity. This reveals some amount of wasted mass in the structure where a component that was previously attached to another now has a “hanging” port not on a load-bearing path. In a post-processing step, we substitute streamlined versions of these components from a library of alternatives, as illustrated in Fig. 3. This procedure decreases the mass of the system from that at the local optimum, often substantially, and its implementation is simple in the component-wise framework.

### 4 Numerical results

We present numerical results for a pair of two-dimensional stress constrained problems. In the first, we minimize mass of an L-shaped bracket subject to a vertical load on its top right tip; in the second, we minimize the mass of a cantilever beam, also with a vertical load at the top right. Both examples use the same set of reference components, pictured in Fig. 3 to model different lattice structures as seen in Figures 6, 9, 10, and 11. In this way the component-wise model is an analogue to the use of beam elements in truss optimization. The examples share the same material properties: Young’s modulus of 113.8 GPa, Poisson’s ratio of 0.34, and a von Mises yield criterion of $\sigma_{max} = 880$ MPa. Linear elasticity is approximated in two dimensions assuming plane stress conditions. We demonstrate that our methodology succeeds in creating designs that respect a constraint on maximum stress while approximating that constraint using PRSC and stress aggregation.

#### 4.1 Problem setup

The reduced order model used for our primary optimization results is trained by collecting 500 snapshots for each pair of connected reference ports and constructing port bases that capture 99.9% of the total energy of each snapshot set, unless a different total energy percentage is specified. When training, we compute the POD for the X and Y components of displacement separately but using the same snapshot set, and apply the total energy criterion to the POD for each component of displacement separately to determine bases. The regularization parameter for pairwise training was set to $\eta = 10$ [10, Sec. 3.2.2]. The resulting bases contain four basis functions, including the constant function, for each component of displacement. An example reduced basis is pictured in Fig. 4. For our primary optimization results, the ROM is constructed using second-order triangular finite elements and the full basis for a port contains twenty-one basis functions for each component of displacement. We address the impact of using second-order elements vs. first-order in Section 4.2.3.

All optimizations are initialized from a ground structure composed of fully solid components, i.e., $\rho = 1$. When postprocessing, we validate the optimized design by performing analysis with the full-order model. Maximum stress values in postprocessing are computed using values at cell centers, and at the quadrature points of a nine-point quadrature rule in each element. We do not consider values of stress on the boundary of an element due to its discontinuity there.
Figure 3: Top: The set of reference components used for optimization in our numerical examples, not including the reflection of the diagonal strut component. The lattice structure is built from repetitions of the cell shown and its reflection. Bottom: Selected examples of substitutions made during postprocessing. The full component library contains a component to be substituted for every combination of hanging ports, and a total of 252 components.
The optimization problem is solved using an interior point method as implemented in Ipopt [36]. We found it was necessary to set the parameter \( \theta_{\text{max fact}} \) to a small value (0.5 was used in these results) to prevent large steps into infeasible regions from which it is difficult for the optimizer to escape. For the same reason, we disable the watchdog procedure that attempts to escape regions of slow convergence by disabling line search for one iteration. In all of the optimization runs presented, Ipopt’s convergence tolerance was set to \( 10^{-6} \).

The Schur complement system is solved using a Cholesky factorization as implemented in Eigen [14], and all optimizations are performed without use of parallelism, for reproducibility. Timings are performed on a Linux server with 2 AMD EPYC 7H12 processors with 64 physical cores and a base clock speed of 2.6 GHz, and 2 TB of memory.

4.2 Mass minimization of an L-bracket

Our first example addresses the design of an L-shaped bracket, a standard benchmark in stress-based TO. While in most TO approaches the L-bracket presents a challenge because of its reentrant corner, using the component-wise methodology, we eliminate this difficulty by a careful choice of the ground structure.

The problem setup for the L-bracket optimization is illustrated in Figure 5. The thickness of the bracket into the page is taken to be 5 cm. The ground structure contains 4,257 components and 2,427,362 finite elements; the reduced static condensation system contains 44,214 unknowns while the full order static condensation system has 278,544. The underlying finite element model contains more than 5 million unknowns (not accounting for those degrees of freedom on ports, which are counted twice). The bracket’s upper boundary is fixed, and a load of 67,500 N is applied as an evenly distributed pressure force over the rightmost two ports at the tip of the bracket.

4.2.1 Optimization results and comparison of aggregation parameters

We solve the optimization (17) for values of the aggregation multiplier \( p = 10, 15, 25, \) and 50, and with 1, 5, 10, 15, 20, and 25 aggregation regions. The results are summarized in Table 1, which records the number of constraint evaluations (\( N_{\text{cons}} \)), constraint Jacobian evaluations (\( N_{\text{jac}} \)), total optimization time, the mass fraction at convergence (\( m_{\text{frac}}^{\text{post}} \)), the mass fraction after postprocessing (\( m_{\text{frac}}^{\text{pp}} \)), the maximum relaxed stress at convergence.
Figure 5: Illustration of the L-bracket geometry and loading condition.
Table 1: Summary of optimization results for varying aggregation parameters. For each combination of $n_{agg}$ and $p$, reported are, from left to right: number of constraint evaluations, number of constraint Jacobian evaluations, total run time, mass fraction at the local optimum, mass fraction after postprocessing, maximum relaxed stress at the optimum, and maximum von Mises stress after postprocessing. Stresses are computed using the full order model in a postprocessing step.
Figure 6: Optimization result using 10 aggregation regions and \( p = 15 \). Left: locally optimal density field. Right: the resulting design after postprocessing. Maximum von Mises stress in the postprocessed design is 799 MPa, compared to \( \sigma_{\text{max}} = 880 \) MPa. Before postprocessing, the maximum relaxed stress was \( \sigma_r = 949 \) MPa.

Fig. 6 illustrates the result using 10 aggregation regions and \( p = 15 \), a representative optimization result. Many components at the local optimum have intermediate values of density; however, component densities are clearly divided between material and void components with densities being either greater than approximately \( \rho = 0.25 \), or close to 0. This motivates our choice of dropping tolerance. Postprocessing creates a fully solid design that satisfies the desired stress constraint, as well as reducing the mass from the value at the local optimum, since many components are replaced with lighter substitutes despite having their density increased to 1 from an intermediate value. The design uses square lattice cells both with and without diagonal reinforcement in different areas of the design, as well as long, solid struts and larger open cells. We highlight here the flexibility of the component-wise approach over a design method that assumes a functionally graded structure, with regions composed of periodic unit cells.

A note is in order here regarding the characteristics of our design as compared to designs created by TO methods based on element-wise SIMP or level set approaches. Those methods will round the concave corner of the L-bracket, eliminating the stress singularity in order to satisfy the constraint. In our case, however, the corner is already rounded in the ground structure by virtue of our choice of reference components. Fig. 6 illustrates the stress field near the corner in the postprocessed design; it turns out that this corner is not, in fact, the location of the maximum stress, which occurs in one of the concave corners in the bottom left part of the structure. This ability to build desirable characteristics into the ground structure is an advantage of the component-wise approach, but comes at the price of limiting the design space. Given a sufficiently fine discretization, we expect that an element-wise SIMP or level set approach will in general find a better design than our method when the same length scale is imposed, due to the much larger design space. We expect our methodology, on the other hand, to be capable of producing designs that have nearly the same performance while realizing a large computational
speedup in the optimization process. Performance of the design resulting from our method depends on the set of reference components; they must be chosen both to enforce desired characteristics such as smooth boundaries, and to provide a rich enough design space to create performant designs. We explore this point in Sec. 4.3.3.

Although we observe variability in the optimization’s performance, there are some general trends. First, increasing the aggregation multiplier \( p \) results in the expected behavior: designs are more conservative, and the optimization problem is more difficult. Increasing the number of aggregation regions, however, does not have a discernable effect on the value of the maximum stress attained in the final design, at least for a moderate number of aggregation regions as used here. It does appear that adding regions is beneficial for convergence of the algorithm up to 10-15 regions; for values of \( p = 15, 25, \) and \( 50 \), the number of constraint evaluations is reduced enough that overall run time is faster for ten aggregation regions than five, despite requiring twice the number of adjoint solves for each Jacobian evaluation.

In every case, the postprocessing step results in a mass that is less than or equal to the objective value at the local optimum. In a few runs, however, it does result in violating the stress constraint because the nonlinear optimization cannot account for the changes made to the shape of components. Removing components with a near-zero density has a negligible effect on the structural response, but the component substitution step may have a large effect. In the examples using one aggregation region and \( p = 15 \) and 25, we observe that component substitution results in increasing the maximum stress to a level that violates the max constraint, where before substitution it satisfied or nearly satisfied the constraint. In the example using 20 aggregation regions and \( p = 15 \), a large increase in maximum stress is observed; however, this is due to the choice of \( \rho_{\text{min}} = 0.2 \), which is too aggressive for this case and removes some components that have a smaller density but are structurally important. For other examples, the postprocessing step either results in a maximum stress that is less than the maximum relaxed stress at the local optimum, or increases it by a small amount such that the max constraint is still respected.

Despite variation across runs, it is clear that with a large enough choice of aggregation multiplier the proposed
methodology consistently converges to conservative designs with a large mass reduction.

### 4.2.2 Performance of the ROM

We compare the accuracy of the ROM used for the primary results to results using a ROM trained with a 99.99% criterion for total energy as well as one using 99% of total energy. We define the following error measures:

\[
e_{\sigma_{\text{max}}} = \frac{\max_{\Omega} \tilde{\sigma}_r - \max_{\Omega} \sigma_r}{\max_{\Omega} \sigma_r}\quad (47)
\]

\[
e_{\sigma} = \frac{\|\tilde{\sigma}_r - \sigma_r\|_{L^2}}{\|\sigma_r\|_{L^2}}\quad (48)
\]

\[
e_u = \frac{\|\tilde{u} - u\|_{L^2}}{\|u\|_{L^2}}\quad (49)
\]

where \(\tilde{\sigma}_r\) is the relaxed von Mises stress as computed from the ROM solution, \(\sigma_r\) is the relaxed von Mises stress computed from the full-order solution, \(\tilde{u}\) is the ROM displacement, and \(u\) is the full-order displacement. The \(L^2\) norm has its usual definition.

Reported in Table 2 are these errors for both the full system at the locally optimum parameter value, and for the postprocessed design. We report results for the designs using an aggregation multiplier of 15; the errors reported for these designs are representative of those for all designs. The exception is error of 13% in the postprocessed design for 5 aggregation regions, which is an outlier. In no other case was the error in maximum stress larger than 10%. The use of a 99.99% total energy criterion is sufficient to reduce the error in maximum stress to less than 5% in every case, and less than 1% for all but one; the effect on the displacement error is smaller. Errors when using a ROM capturing only 99% of total energy are large in every case; this criterion is clearly not strict enough to be useful.

Table 3 additionally reports the resulting number of degrees of freedom and the speedup relative to the full-order static condensation model for each total energy criterion. We report the speedup for the forward solve, which includes the cost of factoring the Schur complement matrix and is the largest component of optimization run time, and the speedup from use of our efficient aggregation scheme relative to the straightforward method using quadrature. Speedup is measured for the full system as used during the optimization procedure. The 99.9% total energy model achieves a speedup of more than a two orders of magnitude vs. a full-order static condensation solve, and the efficient aggregation scheme yields an almost 2.5x speedup vs. straightforward quadrature. In [24], we found that the full-order static condensation achieved approximately a 5x speedup compared to the underlying finite element model; thus, we estimate that the 99.9% total energy ROM used in these optimizations will provide approximately a 750x speedup over a finite element solve using the same mesh.

### 4.2.3 Comparison to results using first-order elements

We also inquire whether the use of second-order finite elements to construct the ROM is actually beneficial. We expect that the resulting stresses in element interiors are more accurate, but it is uncertain whether this increased accuracy has a beneficial effect in the optimization. To determine whether it does, we ran optimizations using a ROM constructed using first-order elements and a 99.9% total energy criterion to compare to the corresponding results using second-order elements. These optimizations use \(p = 15\) and 10, 15, 20 and 25 aggregation regions, respectively. The ROM constructed using first-order elements has exactly the same dimension as that constructed using second-order elements, and the quadrature rule used to compute stress aggregates is identical. The aggregation regions are also identical to those used for the optimizations in Section 4.2.1.

To compare the optimization results we examine the run time, mass fraction of the postprocessed design, and the max stresses at the optimum and in the postprocessed design. The latter are computed using both the full-order first-order finite element model and the full-order second-order finite element model to assess the impact of second-order elements on accuracy in the stress. The results are summarized in Table 4; the optimization run time and mass fraction may be compared to the second block in Table 1. We find that the first-order model underestimates the maximum stress significantly. As a result, the optimized mass fractions are smaller than the corresponding optimization results with second-order elements, but none of the designs found using first-order elements is conservative when analyzed using second-order elements. Additionally, in three out of four cases, the optimization with first-order elements as the model had a longer run time than the second-order optimization. This is due to the fact that the ROM constructed from first-order elements has the same dimension as that
Errors at local optimum

<table>
<thead>
<tr>
<th>n_{agg}</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.52%</td>
<td>4.12%</td>
<td>2.80%</td>
<td>-0.67%</td>
<td>2.96%</td>
<td>0.63%</td>
</tr>
<tr>
<td>5</td>
<td>3.17%</td>
<td>4.46%</td>
<td>4.18%</td>
<td>13.06%</td>
<td>2.93%</td>
<td>4.21%</td>
</tr>
<tr>
<td>10</td>
<td>-0.58%</td>
<td>3.67%</td>
<td>2.98%</td>
<td>-0.78%</td>
<td>2.84%</td>
<td>5.21%</td>
</tr>
<tr>
<td>15</td>
<td>6.42%</td>
<td>3.91%</td>
<td>4.04%</td>
<td>-0.98%</td>
<td>2.80%</td>
<td>1.96%</td>
</tr>
<tr>
<td>20</td>
<td>0.05%</td>
<td>4.17%</td>
<td>3.19%</td>
<td>0.09%</td>
<td>3.22%</td>
<td>3.35%</td>
</tr>
<tr>
<td>25</td>
<td>2.33%</td>
<td>3.78%</td>
<td>3.41%</td>
<td>8.66%</td>
<td>2.54%</td>
<td>4.44%</td>
</tr>
</tbody>
</table>

Errors in postprocessed design

<table>
<thead>
<tr>
<th>n_{agg}</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.54%</td>
<td>2.72%</td>
<td>1.83%</td>
<td>0.02%</td>
<td>2.30%</td>
<td>0.49%</td>
</tr>
<tr>
<td>5</td>
<td>0.08%</td>
<td>3.05%</td>
<td>2.83%</td>
<td>-0.90%</td>
<td>2.28%</td>
<td>4.09%</td>
</tr>
<tr>
<td>10</td>
<td>0.02%</td>
<td>2.33%</td>
<td>2.01%</td>
<td>0.09%</td>
<td>2.33%</td>
<td>5.04%</td>
</tr>
<tr>
<td>15</td>
<td>4.03%</td>
<td>2.62%</td>
<td>2.68%</td>
<td>-0.08%</td>
<td>2.26%</td>
<td>1.94%</td>
</tr>
<tr>
<td>20</td>
<td>-0.79%</td>
<td>2.77%</td>
<td>2.11%</td>
<td>-0.18%</td>
<td>2.71%</td>
<td>3.24%</td>
</tr>
<tr>
<td>25</td>
<td>-0.98%</td>
<td>2.50%</td>
<td>2.27%</td>
<td>0.24%</td>
<td>2.10%</td>
<td>4.34%</td>
</tr>
</tbody>
</table>

Errors using 99.9% of total energy

<table>
<thead>
<tr>
<th>n_{agg}</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>155.67%</td>
<td>48.38%</td>
<td>25.88%</td>
<td>108.61%</td>
<td>21.24%</td>
<td>8.65%</td>
</tr>
<tr>
<td>5</td>
<td>167.64%</td>
<td>42.61%</td>
<td>29.99%</td>
<td>134.32%</td>
<td>21.10%</td>
<td>47.11%</td>
</tr>
<tr>
<td>10</td>
<td>160.05%</td>
<td>33.56%</td>
<td>21.75%</td>
<td>149.29%</td>
<td>19.87%</td>
<td>55.57%</td>
</tr>
<tr>
<td>15</td>
<td>173.93%</td>
<td>37.32%</td>
<td>30.97%</td>
<td>112.66%</td>
<td>20.84%</td>
<td>23.10%</td>
</tr>
<tr>
<td>20</td>
<td>176.95%</td>
<td>43.40%</td>
<td>23.86%</td>
<td>71.62%</td>
<td>19.09%</td>
<td>38.91%</td>
</tr>
<tr>
<td>25</td>
<td>140.72%</td>
<td>37.28%</td>
<td>25.13%</td>
<td>197.87%</td>
<td>17.45%</td>
<td>46.89%</td>
</tr>
</tbody>
</table>

Errors using 99.99% of total energy

<table>
<thead>
<tr>
<th>n_{agg}</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
<th>e_{\sigma_{max}}</th>
<th>e_{\sigma}</th>
<th>e^u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13936</td>
<td>10.6%</td>
<td>1063</td>
<td>769</td>
<td>1215</td>
<td>944</td>
</tr>
<tr>
<td>15</td>
<td>3287</td>
<td>13.5%</td>
<td>1020</td>
<td>962</td>
<td>1195</td>
<td>1023</td>
</tr>
<tr>
<td>20</td>
<td>2743</td>
<td>11.8%</td>
<td>1001</td>
<td>899</td>
<td>1148</td>
<td>910</td>
</tr>
<tr>
<td>25</td>
<td>4699</td>
<td>11.1%</td>
<td>1026</td>
<td>1027</td>
<td>1153</td>
<td>1155</td>
</tr>
</tbody>
</table>
### 4.2.4 Effect of decreasing the maximum stress

To verify that the heuristic of imposing a reduced stress limit $\hat{\sigma}_{\text{max}}$ (Sec. 3.3.3) achieves the desired effect, we solve the optimization problems with $p = 10$ using $\hat{\sigma}_{\text{max}} = 800$ MPa. The corresponding optimizations with $\hat{\sigma}_{\text{max}} = 880$ MPa yielded postprocessed designs that did not satisfy the stress constraint. Results are reported in Table 5 in the same format as those in Table 1.

Setting $\hat{\sigma}_{\text{max}} < \sigma_{\text{max}}$ achieves a reduction in the maximum relaxed stress at the optimum in every case, as well as accelerating convergence in most cases. The effect after postprocessing is somewhat unpredictable. In the case with 25 aggregation regions, the maximum postprocessed stress using a lower upper bound for optimization is actually higher than that with the higher upper bound. In all other cases, the postprocessed stress was decreased. We expect that with enough reduction of $\hat{\sigma}_{\text{max}}$, optimization results will eventually lead to conservative postprocessed designs. Determining how much reduction is required, however, may require an iterative procedure and a choice that is too conservative will trade optimality for conservativeness. Instead, it may be preferable to choose a higher value for the aggregation multiplier $p$, and/or a larger number of aggregation regions. The choice of these parameters of the optimization is problem-dependent.

### 4.3 Mass minimization of a cantilever beam

Our second numerical example demonstrates how the component-wise approach solves a different structural optimization problem while using the same set of components. This illustrates a key advantage: the offline phase of constructing a component library need only be performed once, then the resulting dataset used to solve multiple design problems. We minimize the mass of a cantilever beam, fixed at one end and with a vertical load applied to the opposite tip. The problem setup is illustrated in Fig. 8; material properties are the same as for the L-bracket. The cantilever beam has a length of 1 m and height of approximately 25.8 cm (due to the lattice structure); a 30 kN load is applied to the rightmost two ports on the upper surface of the beam as an evenly distributed pressure force.

The underlying finite element mesh for the beam’s ground structure contains 3,750,278 second-order triangular elements for a total of more than 15 million degrees of freedom (not accounting for those duplicated on internal ports). The full-order static condensation system has 428,736 degrees of freedom, and the ROM constructed with 99.9% of total energy for port spaces contains 81,664.

<table>
<thead>
<tr>
<th>$n_{\text{agg}}$</th>
<th>$p$</th>
<th>$N_{\text{cons}}$</th>
<th>$N_{\text{jac}}$</th>
<th>Run time (s)</th>
<th>$m_{f,\text{frac}}^{\text{opt}}$</th>
<th>$m_{f,\text{frac}}^{\text{pp}}$</th>
<th>$\max \sigma_r$ (MPa)</th>
<th>$\max \sigma_{\text{vm}}$ (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>2399</td>
<td>1251</td>
<td>4572</td>
<td>15.2%</td>
<td>12.9%</td>
<td>1159</td>
<td>910</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>827</td>
<td>517</td>
<td>1866</td>
<td>16.2%</td>
<td>14.9%</td>
<td>996</td>
<td>775</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>803</td>
<td>458</td>
<td>2047</td>
<td>16.2%</td>
<td>14.8%</td>
<td>984</td>
<td>698</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>702</td>
<td>384</td>
<td>1887</td>
<td>16.3%</td>
<td>15.4%</td>
<td>1029</td>
<td>704</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>806</td>
<td>484</td>
<td>2393</td>
<td>15.5%</td>
<td>13.6%</td>
<td>994</td>
<td>885</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>979</td>
<td>564</td>
<td>3009</td>
<td>16.0%</td>
<td>14.1%</td>
<td>970</td>
<td>906</td>
</tr>
</tbody>
</table>

Table 5: Comparison of optimization results when decreasing the value of the stress limit for optimization. Consult Table 1 for definitions of the notation used here.
4.3.1 Optimization result

We solve the optimization problem (17) using 10 aggregation regions and an aggregation multiplier of \( p = 15 \) and the ROM using a 99.9\% total energy criterion. The convergence tolerance for Ipopt is set to \( 10^{-6} \). For this optimization, no reduction of the upper stress bound \( \hat{\sigma}_{\text{max}} \) was necessary to obtain a conservative design. The optimization algorithm converges in 1,944 iterations with 2,576 constraint evaluations. At the local optimum, the mass fraction relative to the ground structure is 19.2\%, with a maximum relaxed stress of \( \sigma_r = 849 \) MPa. After postprocessing, the design has a mass fraction of 15.1\% and maximum stress of 784 MPa. The result is postprocessed using dropping tolerance \( \rho_{\text{min}} = 0.25 \).

Fig. 9 shows the optimized density field and the result of postprocessing. As in the previous optimization results, many components have intermediate densities, but the chosen dropping tolerance clearly separates material and void regions. The postprocessed design resembles well-known results for compliance minimization problems and also has characteristics of functionally-graded structures – note the use of small repeating cells on the upper and lower boundaries, and a structure with a much larger void fraction in the interior.

4.3.2 Comparison to an optimization using different random aggregation regions

To justify the random assignment of aggregation regions, we compare optimization results for different choices of aggregation regions to the result shown previously. The postprocessed designs are compared to the previous one in Figure 10. The third result is postprocessed using a dropping tolerance of \( \rho_{\text{min}} = 0.2 \) (vs. \( \rho_{\text{min}} = 0.25 \) for the first two designs) because of component densities at the optimum that fall between these values but correspond to structurally important components.

Changing aggregation regions does make a significant difference in optimization results. Whereas the original optimization required 1,944 iterations to converge to tolerance, the optimizations with redistributed regions required only 793 and 1,428 respectively. We believe this is more indicative of the difficulty and sensitivity of the optimization problem than it is of a deficiency in the random strategy for region assignment. The local optima with new aggregation regions have mass fractions of 20.3\% and 19\%, vs. 19.2\% in the original; after postprocessing, the mass fractions are 17.4\% and 14.9\% vs. the original 15.1\%. Maximum relaxed stresses at the local optimum are 849 MPa and 887 MPa, vs. the 860 MPa of the original design, and after postprocessing the new designs have maximum stresses of 806 MPa and 1080 MPa. Clearly the assignment of aggregation regions influences the quality of optimization results. We note, however, that the third design’s violation of the maximum stress constraint is due to postprocessing, not the assignment of aggregation regions; postprocessing creates a stress concentration that does not exist at the local optimum. Otherwise, random region assignments do result in fairly consistent designs.

4.3.3 Comparison to a design using a simpler ground structure

Our final example illustrates the impact of the choice of reference components and ground structure on the final design. We optimize a cantilever beam with the same dimensions as in the previous sections, but with a ground
Figure 9: Results of the cantilever beam optimization. Top: optimized component-wise density field. Bottom: the postprocessed structure.

structure created using only the four-port connector and vertical and horizontal struts from Fig. 3. Loading and boundary conditions are the same as in the previous sections, and optimization settings are the same as in Sec. 4.3.1. The result is postprocessed with the same dropping tolerance of $\rho_{\text{min}} = 0.25$. This ground structure and the optimized design are pictured in Fig. 11.

Because the ground structure for this example contains fewer components than the more complex ground structure used previously (5001 components compared to 6601), there are fewer degrees of freedom in the model as well as fewer optimization parameters, and optimization is completed faster. For the simpler ground structure optimization took about 5,500 seconds in total, compared to 12,500 for the original ground structure. The performance of the final design suffers, however. The postprocessed design respects the stress constraint, with a maximum stress of 856 MPa, but its mass is 18% greater than the postprocessed design using the full component library. The four-port design also has a compliance 61% greater than when using a richer ground structure.

These results illustrate the importance of the choice of reference components and ground structure. The component-wise methodology trades flexibility in the design space for computational performance, but limiting the design space too much will significantly limit the performance that can be achieved. Additionally, while not illustrated here, some choices of ground structure could not even permit a feasible solution; for example, if the ground structure contains sharp, concave corners that cannot be eliminated, stress singularities might be unavoidable. Reference components must be chosen to avoid such deficiencies in the reachable design space.

5 Conclusions

We have presented a novel application of component-wise ROMs to stress constrained TO. The methodology succeeds in finding solutions with a large mass reduction relative to the ground structure while respecting stress constraints. While the presence of many local optima as well as the approximation of the maximum stress mean that it is impossible to guarantee a conservative solution, adjusting the aggregation multiplier and upper bound on stress proved effective for stress control. The reduced order model provides an error in the maximum stress of less than 10% in most cases and a relative error of about 3% in the $L^2$ norm, while providing a 150x speedup relative to a full order static condensation model. If greater accuracy is required, a ROM using 99.99% of total energy still provides a 60x speedup while reducing the error in the maximum stress by approximately an order
Figure 10: Comparison of the resulting postprocessed designs from three different random assignments of aggregation regions. Top: the design from Section 4.3.1 with a mass fraction of 15.1% and maximum stress of 784 MPa. Middle: the first design with redistributed aggregation regions, with mass fraction of 17.4% and a maximum stress of 806 MPa. Bottom: A third design with different aggregation regions, with mass fraction of 14.9% but a maximum stress of 1080 MPa.
Figure 11: Top: ground structure for the cantilever beam using only four-port connectors. Bottom: the optimized structure after postprocessing.

of magnitude. Due to the acceleration provided by the component-wise ROM, we are able to solve the TO problem efficiently on a ground structure containing millions of second-order finite elements. Component ROMs provide a high accuracy surrogate model that may be reused to study a multitude of both homogeneous and heterogeneous lattice-like structures. As an additional benefit, we avoid sharp corners in the design altogether by a judicious choice of reference components, obviating difficulties associated with handling of stress singularities. When manufactured, however, the smooth boundary geometry may not be realized; to account for these cases, future work is needed to quantify the effect of manufacturing uncertainties.

There are several other potential areas for future work. First, local stress constraint handling using augmented Lagrangian techniques could directly handle stress constraints and reduce the number of adjoint solves required to one per iteration, potentially providing superior performance. Second, the cost of model evaluations could be reduced using an adaptive refinement strategy, made simple by the component-wise discretization; it is easy to substitute a few components with higher ROM accuracy only where needed. Third, the performance of our method could be increased by using hyperreduction techniques to reduce the cost of computing aggregated constraints. Finally, the current formulation does not address buckling constraints, which will be necessary for practical use. There are multiple approaches that could be taken to solve the buckling problem. First, a buckling constraint could be imposed directly by solving the associated eigenvalue problem; The component-wise theory for solving symmetric eigenproblems has been developed in [35]. Extensive work on implementation and tuning would be necessary to address the difficulties of buckling-constrained TO documented in [13, 12] and others, however. It might instead be preferred to address buckling by heuristic methods. Since the choice of yield criterion is not limited to the von Mises criterion, we can instead choose a criterion that distinguishes between tension and compression, such as the Drucker-Prager criterion. Intuitively, reducing the allowable stress in compression should result in designs more resistant to buckling. Further work is necessary to investigate the merits of these different approaches to the buckling problem.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Acknowledgements

Funding: This work was supported in part by the AEOLUS center under US Department of Energy Applied Mathematics MMICC award DE-SC0019303. The second author was funded by LDRD (21-FS-042) at Lawrence Livermore National Laboratory. Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344 and LLNL-JRNL-834458.

References


25


