AUTOMATED SHAPE OPTIMIZATION USING A MULTIGRID METHOD AND ESTIMATION OF DISTRIBUTION ALGORITHMS

S. I. Valdez*, N. Faurrieta*, J. L. Marroquín* and S. Botello*

* Center for Research in Mathematics, CIMAT, A.C. Jalisco S/N, Valenciana, Guanajuato, Gto. México, C.P. 36000
e-mail: noe.faurrieta, jlm, ivvan, botello@cimat.mx

Key words: Topological optimization, Estimation of Distribution Algorithms

Abstract. Topological shape optimization refers to the problem of finding the optimal shape of a mechanical structure by using a process for removing or inserting new holes or parts, it is to say, using a process which produces topological changes. This article introduces a method for automated topological optimization via an Estimation of Distribution Algorithm (EDA) with a suitable representation of the optimization variables. The optimum structure is such with the minimum weight which does not exceed a maximum von Mises stress and displacement. The contributions of this proposal resides in the definition of a candidate solution and the optimization method. The candidate solution representation is independent of the finer discretization used for analyzing candidate structures using the finite element method. Given a domain Ω, which corresponds to the physical space where candidate structures reside, a vector $\phi = [\phi_1, \phi_2, ..., \phi_m]$ is used to define a smooth function $\hat{\phi}(x, y)$ on $\Omega$. If $\hat{\phi}(x, y)$ is less than 0.5, such region does not have material, otherwise, it has. The smooth function $\hat{\phi}(x, y)$ provides the advantage of having continuous regions with or without material while it depends on few optimization parameters $\phi$, in addition, we can define an arbitrary number of parts or gaps as thinner or larger as needed. The EDA benefits from this representation, sampling random arbitrary structures and using probabilistic learning to determine whether a region must have material. The EDA is a global optimizer which can propose different topologies without the need of a priori knowledge neither initial solutions. In addition it uses a probabilistic model which smoothly evolve through generations. In consequence, at the beginning of the optimization process it arbitrarily proposes topologically different structures, while at the convergence phase it performs similar to a local search algorithm. The EDA uses few parameters which can be set in a straight forward manner. We report several study cases from the specialized literature, showing that our proposal outperforms reported results from up-to-date well performed algorithms.
1 INTRODUCTION

The shape optimization problem is defined as: to find the best shape of a structure [5], for given design objectives. The best shape usually is defined according to the structure stresses or rigidity [7] or weight [15].

In this article the goal is to find the structure with minimum weight which fulfills maximum von Misses and maximum displacement conditions. In contrast with other proposals we consider self weight [7], which is essential for realistic numerical simulation.

We differentiate three categories of the problem referred as shape optimization:

- **sizing optimization**. It is to find optimal dimensional parameters such as thickness, height, width, radius, etc.
- **shape optimization**. It is to find the optimal shape by modifying the inner or outer boundaries of a predefined geometrical model.
- **topology optimization**. It is to find the optimal structure by adding or removing complete structural elements or holes.

In this article we introduce a method for topological optimization based on an Estimation of distribution algorithm with a suitable representation. A topology optimization algorithm requires of several components, by instance:

- a representation of candidate solution for the optimization process,
- an optimization method,
- a procedure to evaluate a candidate solution,
- in the case that the representation used for the optimization process differs from the one used for evaluating a candidate solution, a procedure to map a candidate solution from the optimization process to the evaluation representation is needed.

In this case, the representation used for a candidate solution is a vector of continuous real variables. The optimization method is an Estimation of Distribution Algorithm, which initially simulate random vectors of continuous real variables in [0, 1], these solutions are evaluated using the Finite Element Method (FEM), the best solutions are used to learn a probability distribution, in this case, we used independent Gaussian models for each variable, then new candidate solution can be sampled.

A candidate solution in the EDA can not be directly evaluated by the finite element method, hence we use a procedure described in Section 2 for mapping a candidate solution to a structure. The size of the vector used in the EDA is around 1000 of elements, while the number of finite elements can be arbitrarily set, in our experiments it is around 20000.

Shape and topological optimization have been largely tackled, the proposals include gradient based methods as well as evolutionary algorithms [12]. The proposals based
on gradient based methods [9, 17] or level set methods [16], usually required of initial structures, where basically, initial holes are a bias for the shape is planted by the user. In contrast our proposal does not need of any initialization, gradient search methods and level set methods, usually gradually perturb the structure in certain computed directions, until no improvement is achieved, hence the structure converges to a local optimum given the initialization. An advantage of these kind of methods is that they are effective for refining initial solutions, the solution delivered by these methods usually optimal given the topology, but the method is incapable of performing abrupt changes in the topology.

On the other hand, evolutionary algorithms, such as ant colony optimization [10], particle swarm optimization[3], and genetic algorithms [4, 1], have been used for global search of the optimum topology. The advantage of these method is their capacity to propose different topologies without user bias, hence they are not restricted to a particular kind of solutions, neither required of an initial structure. The possible disadvantages are that, even no initial bias is required, the structures could be limited to certain structure, because of the representation used [14], where the possible structures are limited in the number of parts or structural elements, and their sizes. Another disadvantage, is that evolutionary algorithms permit abrupt changes any stage of the algorithm, this characteristic could be convenient to escape from local optimum but undesirable to refine solutions and obtain the maximum efficiency of a given topology.

In order to obtain the combined advantages of the mentioned algorithms hybrid proposals include global and local search, intending to explore many different topologies while obtaining the maximum efficiency of each of them, hence they can be seen as topology as well as size or shape optimizers [18, 11].

Considering the comments above, the EDA used in this approach is based on learning and sampling from a probability distribution, the distribution is computed by using the best structures found. The best structures found during the optimization process are maintained through generations, and used to reinforce the knowledge about them in the estimation of the probability distribution, consequently at the beginning the proposed topologies present a large variation, as similar structures reinforce the probabilistic learning the variation is reduced. Thus, at the beginning of the optimization process the EDA presents the very same characteristics of other evolutionary algorithms, it can explore many different topologies, and at the end it performs similar to a local search algorithm because the sampled structures are similar, and only small changes in the most uncertain regions are allowed. In addition, our representation allow to build structures with any number of gaps or structural elements, there is not a bias which forces to sample a particular kind of structure.

2 PROBLEM REPRESENTATION

The optimization problems is defined as to find the topology and shape of a mechanical structure with minimum weight which fulfills the service conditions. That is to say, the mechanical structure must support the given loads without exceeding the maximum
allowed von Misses stress and displacements in the loading points.

To approximate a solution to this problem we use an initial domain $\Omega$ in which the boundary conditions and loads are given. Figure 2a) shows the domain where the boundary conditions were set. Figure 2b) is a discretization used for optimizing. In this mesh each node represents a continuous variable $\phi_i$, hence we have as many optimization variables in the vector $\phi$ as nodes in the mesh. The $\phi_i$ vector is used to define the Poisson problem in Equation (1).

$$\lambda \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \hat{\phi}(x, y) = f(x, y)$$

In Equation (1) $f(x_i, y_i)$ has the value $\phi_i$ in the node with coordinates $(x_i, y_i)$, while $f(x, y)$ is interpolated using the finite element shape functions inside each element. The parameter $\lambda$ regulates the smoothness of the solution $\hat{\phi}(x, y)$. Figure 2c) is an example of a solution of such problem, seeing $\hat{\phi}$ as a height function.

The solution $\hat{\phi}(x, y)$ is a continuous function defined in the domain $\Omega$ and it is used to decide whether the material exist in the domain. If $\hat{\phi}(x, y) \geq 0.5$, then, there is material in such region, otherwise there is not material. In consequence, the existence of material is a property defined with arbitrary high precision. Once the regions with and without material is defined, we propose to use a second, finer, discretization for evaluating the structure feasibility (Von Mises stress and displacements) and efficiency (weight).

Figure 2a) shows the mesh used for optimizing. Figure 2b) shows a single element $r$ of this mesh, the nodes of this element have assigned four $\phi$ values from the optimization variables, it is to say, $\phi_i, \phi_j, \phi_k, \phi_l$, the curved surface in Figure 2b) is the corresponding smooth function $\hat{\phi}(x, y)$ in such element, computed as the solution of Equation (1), the flat plane is the threshold of 0.5. At the base of Figure 2b) there is a second mesh used for the Finite Element evaluation of Von Mises stresses and displacements. Notice that the function (curved surface) is continuously defined in the element, hence the discretization of the bottom mesh can be set arbitrarily. The corresponding element is shown in Figure 2b), where the domain $\Omega$ is discretized using a finer mesh in order to evaluate the Von Mises stresses and displacements.

In summary, a candidate solution is represented by a relatively low dimension vector $\phi$ of continuous variables, in our experiments we use approximately one thousand of variables, this vector is used to compute a smooth function, then the function is thresholded to determine whether the elements in a finer mesh (tens of thousands of elements) have material.

2.1 Objective function

The main goal of this approach is to find a structure with minimal weight, without exceeding a maximum Von Mises stress and nodal displacement. The stresses and displacements are measured via the finite element method (FEM), using the fine grained mesh.
The goal as well as the constraints are considered in a single function in Equation 2.
\[ \min f(x) = \begin{cases} 
(VM(x) + 1)(D(x) + 1) & \text{if } VM(x) > 0 \\
W(x) & \text{or } D(x) > 0 \\
\frac{W(x)}{W_{\text{max}}} & \text{otherwise} 
\end{cases} \] (2)

The first case in Equation (2) measures the feasibility of a candidate solution. \( VM(x) \) is the sum of the exceeding Von Mises stresses in the elements with material, computed according Equation (3), where \( \hat{\gamma}_i \) is the exceeding stress in the \( i \)-th element. The maximum
allowed Von Mises stress is a material property.

\[ VM(x) = \sum_{i=1}^{n} x_i \hat{\gamma}_i, \] (3)

In addition, \( D(x) \) is the sum of the exceeding displacements in the loaded nodes. These excesses are calculated using Equation 4, \( m \) is the number of loaded nodes, \( \delta_i \) is the exceeding displacement in the \( i-th \) node. The maximum allowed displacement is a parameter of the problem.

\[ D(x) = \sum_{i=1}^{m} |\delta_i|, \] (4)

Notice that the first case in Equation (2) always returns a value greater than 1, because \( VM(x) \) and \( D(x) \) are the exceeding stresses and displacements, hence they are lower bounded by 0. Consequently the first case applies when the candidate solution is unfeasible, and, in such a case, the returned value is always greater than 1. The second case in Equation (2) measures the relative weight of a candidate structure, where \( W(x) \) is the weight of the structure given by:

\[ W(x) = \sum_{i=1}^{n} w_i x_i, \] (5)

being \( w_i \) the weight of each present element in the fine grained mesh, \( x_i \) is a binary indicator for the presence or absence of material, \( x_i \) is computed as shown in Figure 2c), by thresholding the function \( \phi(x, y) \) at the center of the element in the fine grained mesh. \( W_{\text{max}} \) is the maximum weight of the structure, that is to say with all the elements present.

As can be observed the second case measures the weight of the structure only in the case it is feasible. Hence, the objective function is used to minimize the exceeding Von Mises stresses and displacements when the structure is unfeasible and to minimize the weight of the feasible structures. In the same sense, a feasible structure always has a lower objective function value than an unfeasible structure.

3 OPTIMIZATION ALGORITHM

The current proposal is based on Estimation of Distribution Algorithm (EDA) for global optimization. EDAs are stochastic optimization methods, which perform global search by building and sampling from probabilistic models. The EDA used for this proposal is similar to the UMDAg [8], the main differences are the initialization, the selection method, and the stopping criterion, but our proposal as well as UMDAg assumes independent Gaussian models. The EDA used in this article is described by Algorithm 1. In line 1 the iteration or generation counter is started, line 2 initializes the probability model with a uniform distribution, in line 3 we simulate the initial sample named as a population.
Each element in the population is called as an individual. An individual is a candidate solution, for this particular problem, it is a vector of $d$ values $\phi^i = [\phi^i_1, \phi^i_2, ..., \phi^i_j, ..., \phi^i_d]$, the super-index $i$ indicates that it is the $i$-th individual.

Each individual $\phi^i$ is translated into a unique binary vector $x^i$, then $x^i$ can be translated into a structure as shown in Figure 2. In the while loop, the number of individuals sampled is $n_{pop}$, but in the first population, in line 3, is of size $2n_{pop}$, with the purpose of reducing the possible initial bias in the search process.

<table>
<thead>
<tr>
<th>Algorithm 1: EDA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
</tr>
<tr>
<td>$n_{pop}$ = Population size.</td>
</tr>
<tr>
<td>$O_{min}$ = Minimum ordering parameter for the stopping criterion.</td>
</tr>
<tr>
<td>$n_{gen}$ = Maximum number of generations for the stopping criterion.</td>
</tr>
<tr>
<td><strong>Data:</strong></td>
</tr>
<tr>
<td>$n_{var}$ = Number of variables ($\phi$ size).</td>
</tr>
<tr>
<td><strong>Result:</strong></td>
</tr>
<tr>
<td>$\phi^{Best}$ = Best optimum approximation</td>
</tr>
<tr>
<td><strong>1</strong> $t = 0$;</td>
</tr>
<tr>
<td><strong>2</strong> $Q^0(\phi) =$ Initial probability model();</td>
</tr>
<tr>
<td><strong>3</strong> $\Phi^0 =$ Generate a random population*(Q$^0$);</td>
</tr>
<tr>
<td><strong>4</strong> $F^0 =$ Evaluation($\Phi^0$);</td>
</tr>
<tr>
<td><strong>5</strong> $S^0 =$ Selection*(F$^0$, $\Phi^0$);</td>
</tr>
<tr>
<td><strong>6 while</strong> stopping criterion is not met <strong>do</strong></td>
</tr>
<tr>
<td><strong>7</strong> $Q^t(\Sigma, \mu, \phi) =$ Compute probability model($S^t$);</td>
</tr>
<tr>
<td><strong>8</strong> $\Phi^t =$ Generate a random population($Q^t$);</td>
</tr>
<tr>
<td><strong>9</strong> $F^t =$ Evaluation($\Phi^t$);</td>
</tr>
<tr>
<td><strong>10</strong> $[S^t, \phi^{Best}] =$ Selection($F^t$, $\Phi^t$);</td>
</tr>
<tr>
<td><strong>11</strong> ;</td>
</tr>
<tr>
<td><strong>12</strong> Postprocess($\phi^{Best}$);</td>
</tr>
</tbody>
</table>

In line 4 the individuals are evaluated using Equation (2). In line 5 a subset of the best individuals is selected. Recall that the initial population is of size $2n_{pop}$, then the initial selected set is of size $n_{pop}$. In line 7 we compute a new probability distribution, by using the selected set in line 5, the mean an variances are computed considering independent variables in a similar fashion than the UMDAg [8]. In line 8 a new population of size $n_{pop}$ is sampled. In the selection step in line 10, the last selected set at $t - 1$ is merged with the current population, then this new set is ordered and the best $n_{pop}/2$ individuals are selected. Notice that the best $n_{pop}/2$ candidate solution are never lost, hence the best candidate solution is always in the selected set, and it is returned as the best optimum approximation when the optimization process finishes. The stopping criterion is given by the maximum number of generations.
3.1 Numerical Issues

In the numerical implementation of the proposed method, linear elements are used for the displacement analysis. In order to avoid numerical instability during the optimization process, the elemental stiffness matrix is scaled by a factor of $\epsilon = 1 \times 10^{-4}$ when an element is under the threshold of $\theta = 0.5$. Once the optimization process has finished, as a final process, a dilatation is applied to the best structure, and present objects which are unconnected with the main structure are removed. The dilatation process sets present to any element which has a present neighbor. A neighbor of one element is another element connected by a face. In the final evaluation, instead of scaling the elemental matrix of non-present elements, it is set to zero in order to ensure that the rigidity contribution is null in absent elements.

4 NUMERICAL EXPERIMENTS

To validate the proposed optimization method in this section we introduce several examples of structural optimization. For all examples the material properties are: Young modulus $E = 2.1 \times 10^{11}$ Pa, Poisson ratio $\nu = 0.3$, thickness of 1 mm and a maximum Von Mises stress of $2.8 \times 10^8$ Pa, using $\lambda = 0.2$ in Equation (1), and a maximum displacement of $\bar{\epsilon} = 1 \times 10^{-4}$. The EDA uses a population of 2000 individuals. The algorithm is stopped when it reaches 500 generations.

4.1 Short cantilever beam with load of 100N

The first study case is a short cantilever beam. In Figure 3 the design domain is shown. The beam is fixed in the upper and lower left side, and a load $P = 100$ N is applied in a downward direction in the lower right side of the beam. The optimization domain is discretized using $36 \times 36$ square cells. Additionally to show the optimal structure approximation, this case of study is used to show the approximation of the optimum structure in different generations.

Different stages of the optimization process, at generation 100, 300, 500 and final of the best individual are shown in Figure 4. The best optimum approximation is shown in Figure 4 (d). The volume in the final design structure is 7.24%. As can be observed, the solution is improved through the generations, the unnecessary objects are gradually removed, the final process which consist on removing flying objects and dilatation does not drastically change the structure found by the EDA.

4.2 Short cantilever beam with a load of 300N

For the same test case, the load is increased to $P = 300$ N. Different optimization stages (100, 300, 500 and final) of the best individual are shown in Figure 5. The best optimum approximation is shown in Figure 5 (d) and, as might be expected, the volume increases in the final design. For this study case the volume in the final structure is 25% of the initial volume. In this case, we appreciate the effects of the final process (removing
4.3 Cantilever beam with a load of 100N

The second study case is a longer cantilever beam than in the first case. In Figure 3 the design domain is shown. The initial structure is fixed at the top and bottom of the left side. A load $P = 100$ N in the middle right side of the beam is applied. The domain is discretized using $36 \times 36$ cells. In addition to show the performance of this proposal, in this study case we show the evolution of the probability distribution, which
is an indicator of the search process and convergence of the whole population to the best optimum approximation.

![Design domain](image)

**Figure 6**: Design domain, loading and boundary conditions for cantilever beam.

Figures 7(a)-(c) shows the probability of having material in the design domain at different generations: 100, 300 and 500, and the final design in (d). In Figure 7 (a) it can be seen that in early generations there is a large variability, then, as the algorithm progresses this variability decreases. In the last regions where the individuals vary, are the boundaries of the structure, as can be seen, the algorithm learns from the random samples in order to accurately finding an optimal approximation. The optimal volume approximation for this example is 10.95% with respect to the initial.

![Evolution of cantilever beam population](image)

**Figure 7**: Evolution of cantilever beam population at different generations. Color indicates probability in (a)-(c) and von Mises stress in the final design (d).

### 4.4 Cantilever beam with a load of 300N

The load is increased to $P = 300$ N. Figure 8 shows the probability of having material, in generations: 100, 300, 500, and the final design is shown in Figure 8 (d). The final design is almost symmetrical, this symmetric design is usually seen in literature for this test case. Despite our proposal does not insert any bias in order to get a symmetric design, the algorithm detects the adequate optimal topology. The volume in final design is 25% with respect to the initial.
4.5 L-beam

The last study case is a L-beam. The model is fixed at the top edge and a load of P = 100 N in the middle right side of the beam is applied, as shown in Figure 9. The domain is divided into three rectangular sections, using the dimension lines as reference. Then, each rectangular section is discretized in $22 \times 22$ cells. This last test case shows that the proposal is capable of approximating complex optimal structures with several holes, branches and curves.

4.6 Quantitative comparison

The cases of: Short Cantilever with a load of 100 N, Long Cantilever with a load of 100 N, and the L-beam are reported in a recent well-performed proposal. In [13] the desired volume is a parameter, hence it is not minimized but a feasible structure with such volume is intended to be found. Nevertheless the difference in the manner the problem is approached, the volume found can be taken as a reference of possible volume reduction. Table 1 shows the numerical comparison between the volume reported in [13] in contrast with the volume found by our approach.
Table 1: Volume comparison in optimal structures design (%).

<table>
<thead>
<tr>
<th>Example</th>
<th>reported</th>
<th>our proposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short cantilever</td>
<td>35</td>
<td>7.24</td>
</tr>
<tr>
<td>Cantilever</td>
<td>35</td>
<td>10.95</td>
</tr>
<tr>
<td>L-beam</td>
<td>45</td>
<td>36.19</td>
</tr>
</tbody>
</table>

5 CONCLUSIONS

In this proposal we introduce a novel representation for the topological optimization of mechanical structures, which is suitable to be used with evolutionary algorithms. In the best of our knowledge there are not competitive results, with respect to our proposal, reported in the up to date literature using evolutionary algorithms. We show in Table 1 that our proposal is competitive with state of the art algorithms. Very often evolutionary algorithms used for topology optimization cannot deal with high order dimensions (thousands), and suffer from premature convergence [2, 6], which consequently derives in sub-optimum structures with many holes, or with gross approximations because of a single coarse mesh is used for optimization as well as the FEM analysis.

The representation proposed in this work, considerably reduces the dimension of the optimization problem, hence the evolutionary algorithm only deals with hundreds of variables. In addition, the proposed mapping between the vector of decision variables and the candidate structure allows to define any structure, it is to say, we do not bias or constraint the search space as in similar approaches with evolutionary algorithms [14]. A impacting advantage of EDA over other evolutionary algorithms is the reduced number of parameters, in our case, we only need a parameter: the population size, and a stopping criterion. Both of them can be set straightforward. The initial population size can be set as the number of variables, and it can be increased until the performance of the algorithm
do not significantly improve. The stopping criterion can be set using the ordering parameter, when the ordering parameter is less than $1e-4$ it means that the population is not varying and we are evaluating almost the same individual, hence we are not exploring anymore and the algorithm must be stopped. Thus, we consider that 500 generations are enough to stop the exploration in all cases. As can be seen, the EDA is not difficult to configure.

Another advantage is that we can extract valuable information, as Figure 5 8 where we can appreciate the regions where material exist most probably, then even if we do not find the real optimum, we can determine a region where very probably it is, because we have sample a lot of random structures, an we can observe how the material is distributed in the best performed ones.

Our proposal considers self-weight and binary densities, which are mapped to existence or absence of material in contrast with similar approaches which consider continuous densities or Young modules, notices that in real world problems we can not manufacture materials with any density or Young module.

Finally, notice that the methodology can be applied to 3d problems without any significant change. Hence, future work will contemplate 3d problems as well as multi-grid strategies.

REFERENCES


