11th World Congress on Computational Mechanics (WCCM XI)
5th European Conference on Computational Mechanics (ECCM V)
6th European Conference on Computational Fluid Dynamics (ECFD VI)
E. Oñate, J. Oliver and A. Huerta (Eds)

TOPOLOGY OPTIMIZATION BY PREDICTING SENSITIVITIES BASED ON LOCAL STATE FEATURES

Nikola Aulig[1], Markus Olhofer[2]

Honda Research Institute Europe GmbH,
Carl-Legien-Straße 30, 63073 Offenbach/Main, Germany, www.honda-ri.de
[1]nikola.aulig@honda-ri.de [2]markus.olhofer@honda-ri.de

Key words: Topology Optimization, Finite Differencing, Sensitivity Prediction, Linear Regression, Support Vector Regression, Local State Features.

Abstract. Mathematical density-based topology optimization methods commonly require analytical sensitivity information. In this paper we propose a heuristic approach for topology optimization, targeting the optimization of objective functions for which analytical sensitivities are not available or difficult to obtain. Concretely, sensitivities are substituted by the prediction of a regression model, which is trained based on sampled sensitivity data. This information is obtained from finite differencing, combined with the assumption that local state features, associated with each design variable, are suitable for predicting the corresponding sensitivity. In order to evaluate the proposed method and in order to compare the results to a known optimal baseline solution, it is applied to the problem of optimizing a minimum compliance cantilever. In most experiments, optimized designs similar to the baseline design are obtained, while the number of finite element solver runs is reduced drastically compared to pure finite differencing gradient estimation. As solution quality and the number of required samples depend on the prediction quality, we provide recommendations for the choice of model and features based on the conducted experiments.

1 INTRODUCTION

The target of topology optimization is to find concepts for the design of mechanical structures, i.e. to find the optimal distribution of material within a design space, in order to provide the basic layout and shape of the structure. Various approaches to topology optimization exist [7]. Among these, density-based methods [4] are widely used. In density-based methods, a finite element mesh discretization of the design space is represented by a continuous density variable for each element. The problem of identifying elements which should contain material and those that should contain void can be efficiently solved with

gradient-based approaches using rigorously derived analytical sensitivities of the design variables.

However, in practical engineering optimization problems gradient information can be difficult to obtain. Examples are topology optimization of crashworthiness designs or structures consisting of complex materials like composites. In such cases non-gradient approaches or finite difference gradient estimations may be used, however these methods are usually considered computationally too expensive due to the typically high dimensionality of topology optimization problems. Therefore, in this paper we propose a method which generates a heuristic substitute for analytical sensitivities, which can be used instead.

The proposed approach is based on the observation that the sensitivity of a design variable is usually computed by using state information related to the finite element to which it is associated. This is true for the minimum compliance problem, since the sensitivity can be computed (for a constant elemental stiffness matrix) from elemental displacements and elemental density. Other examples are the problem of eigenvalue maximization, thermal conductivity maximization or compliant mechanism synthesis [4], for which the sensitivities are as well computed based on the state of the element.

Furthermore in the literature approaches exist, that utilize heuristic criteria based on which the design is optimized instead of rigorously derived sensitivities. In [3] the stress within an element is used as criterion to perform biologically inspired topology optimization. In the bi-directional evolutionary structural optimization [10] a sensitivity number is used as criterion to discretely remove or add material in elements. Another approach [11], aims for topology optimization of structures subject to crash loads. It utilizes the elemental internal energy density as criterion based on which material is distributed in order to achieve a uniform energy distribution.

Mathematically rigorous and heuristic approaches have in common that the criterion for material distribution is defined prior to the optimization, with respect to the considered objective function.

So far, there exist almost no approaches with the ability to generically devise a heuristic criterion which can be used for topology optimization instead a pre-defined one. An exception is [2] in which a neural network approximation model is optimized by an evolutionary search algorithm. The purpose of the model is to process elemental state information in order to provide an update signal for each element, which can be used as a substitute for sensitivity information. The process was shown to provide feasible structures, even for using weak features. However, the approach involves a very high computational cost and the resulting designs show a large variance.

In this paper, we propose an approach in which the sensitivity of a design variable is explicitly sampled using finite differencing and a sensitivity model is learned directly on these data. In order to perform a statistical evaluation and in order to compare the results to a well established baseline, we apply the new method in the context of the minimum compliance problem.

Section 2, is a reminder of the density-based topology optimization with respect to the

minimum compliance problem. Section 3 introduces the topology optimization based on predicting sensitivities. Experiments and obtained results are described in Section 4, and the paper is concluded in Section 5.

2 MINIMUM COMPLIANCE TOPOLOGY OPTIMIZATION

The minimum compliance problem is frequently used as a benchmark problem for topology optimization, therefore density-based topology optimization is briefly introduced in the context of the minimum compliance problem.

In density-based topology optimization [4] each element $i = 1, \ldots, N$ of the discretized design space is represented by a variable. This binary problem can be relaxed by utilizing a material interpolation scheme like the solid isotropic material with penalization (SIMP) method. In case of a minimum compliance problem the Young's modulus E_0 of the material within a mesh element is interpolated according to

$$E_i(\rho_i) = \rho_i^p E_0, \tag{1}$$

with the design variable $\rho_i \in [\rho_{\min}, 1]$, a penalization p and a minimum density $\rho_{\min}$.

The minimum compliance problem can be formulated as [4]:

$$\begin{aligned} \min_{\boldsymbol{\rho}} \, & c(\boldsymbol{\rho}) = \mathbf{u}^T \mathbf{f} \\ \text{s.t.} : \, & \mathbf{K}(\boldsymbol{\rho})\mathbf{u} = \mathbf{f}, \\ & V(\boldsymbol{\rho}) = V_f, \\ & 0 < \rho_{\min} \leq \rho_i \leq 1, \; i = 1, \ldots, N, \end{aligned} \tag{2}$$

with the compliance c, the global displacement vector $\mathbf{u}$, the forces vector $\mathbf{f}$, and the global stiffness matrix $\mathbf{K}$. A volume constraint V_f is imposed, specifying the proportion of design space which is filled with material.

The problem can be solved by a standard optimality criteria method (OC-update). A heuristic update scheme that redistributes material among the design variables and maintains the volume constraint is [4]:

$$\rho_i^{\text{new}} = \begin{cases} \max(\rho_{\texttt{min}}, \rho_i - m) \\ \quad \text{if } \rho_i B_i^\eta \leq \max(\rho_{\min}, \rho_i - m), \\ \\ \min(1, \rho_i + m) \\ \quad \text{if } \min(1, \rho_i + m) \leq \rho_i B_i^\eta, \\ \\ \rho_i B_i^\eta \text{ else}, \end{cases} \tag{3}$$

with the move-limit m, damping coefficient η, and $B_i = \frac{\partial c}{\partial \rho_i} / \Lambda$, with the assumption that every element has a unit volume. The multiplier Λ is determined by a bi-sectioning

algorithm, which tunes the design update such that the volume constraint holds. In conventional approaches to topology optimization an analytical formulation of the adjoint sensitivity $\frac{\partial c}{\partial \rho_i}$ is assumed to be known. While this is true for the minimum compliance problem this is not true in the general case. Since the compliance problem only serves as a test problem for the general case, we assume $\frac{\partial c}{\partial \rho_i}$ to be unknown. In that case the sensitivity may be estimated using a finite differencing approach as:

$$\frac{\partial c}{\partial \rho_i} \approx \frac{\Delta c}{\Delta \rho} = \frac{c(\boldsymbol{\rho} + \boldsymbol{\Delta\rho}_i) - c(\boldsymbol{\rho})}{\Delta \rho}, \tag{4}$$

with the vector elements $\Delta\rho_{ij}$ of $\boldsymbol{\Delta\rho}_i$ for $j = 1, \ldots, N$ and

$$\boldsymbol{\Delta}\rho_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ \Delta\rho & \text{if } i = j, \end{cases} \tag{5}$$

with a small $\Delta\rho > 0$.

A problem regularization can be imposed for example by filtering of sensitivities [12]:

$$\frac{\widehat{\partial c}}{\partial \rho_i} = \frac{1}{\rho_i \sum_{j \in N_i} H_{ij}} \sum_{j \in N_i} H_{ij} \rho_i \frac{\partial c}{\partial \rho_i} \tag{6}$$

where N_i is the set of elements whose center lies within the radius $r_{\min}$ around the center of element i. H_{ij} is defined as

$$H_{ij} = r_{min} - \text{dist}(i, j) \tag{7}$$

where $\text{dist}(i, j)$ is the center to center distance of elements i and j. Filtering of sensitivities imposes a minimum length scale and prevents checker-board patterns.

3 TOPOLOGY OPTIMIZATION BY PREDICTING SENSITIVITIES

The required number of finite element analysis simulations per iteration, when using gradient-estimation via finite differencing as in (4) is $N+1$. In a typical topology optimization application, however, the number of design variables easily reaches tens of thousands up to millions. In such a case the computational cost for estimating the gradient in this way is infeasibly high.

In this section we propose a heuristic alternative to computing finite differences for all design variables. We assume that the sensitivity of element i can be modeled by a predictor $S_\theta(\mathbf{s}_i)$:

$$\frac{\partial c}{\partial \rho_i} \approx S_{\boldsymbol{\theta}}(\mathbf{s}_i), \tag{8}$$

with the feature vector $\mathbf{s}_i$ and the model parameters $\boldsymbol{\theta}$. We term the approach "Topology Optimization by Predicting Sensitivities"(TOPS). It assumes that there exist features $\mathbf{s}_i$

which contain information on the sensitivity of element i. The mapping from the features to the sensitivity substitute needs to be learned by a machine learning approach. The choice and availability of features depends on the problem. In topology optimization, the governing equation needs to be solved. In the case of the minimum compliance problem this is achieved by a finite element analysis. The solution of the analysis is the state of the structure, which for a (displacement-based) finite element solver is given by the nodal displacements. Therefore the elemental displacement vector $\mathbf{u}_i$ of the nodes defining an element are naturally the basic features available for learning the model. Another basic information on the element is the design variable itself, i.e. the material density ρ_i of the element. A basic feature vector for element i can therefore be defined as $\mathbf{s_i} = [\mathbf{u}_i^T \ \rho_i]^T$. Depending on the objective function and the type of analysis, different features might be available. Naturally we would expect that the features are based on the local state of the element to which the design variable is associated, therefore we use the designation Local State Features (LSF), in accordance with [2].

In order to train a predictor, training data is required. After conducting the finite element analysis the LSF vector $\mathbf{s}_i$ is available for all elements, respectively design variables $i = 1, \ldots, N$. A training sample for the model output can be obtained by computing the sensitivity from finite differencing in (4) for a single design variable i, yielding a sensitivity sample $\left(\mathbf{s}_i, \frac{\Delta c}{\Delta \rho}\right)$. A set of training samples can be obtained by repeating this for a number of design variables. This data set can be used to train the regression model, which can predict the sensitivities based on the LSF for other design variables. Since the computational cost for training a model and predicting sensitivities for the remaining design variables is usually much smaller than that for performing finite element analysis simulations for all of the remaining variables, this is a more efficient approach than estimating the gradient by pure finite differencing.

Based on this idea we propose the TOPS algorithm. The design is initialized with an initial guess, for example a homogenous distribution of material and a finite element analysis is performed. The model parameters are initialized with zero and the database is empty. The following steps are iterated:

1. The sensitivity with respect to the objective function and the current design is estimated separately for a random subset of size N_s of the design variables by performing a finite difference step as in eq. (4). No design variable is sampled more than once for the same current design. The sampled sensitivities and the corresponding features are stored in the database.

2. The parameters of a new regression model are trained with a learning algorithm, based on the samples in the database.

3. Sensitivities are predicted for all design variables which were not sampled by finite differencing in step 1.

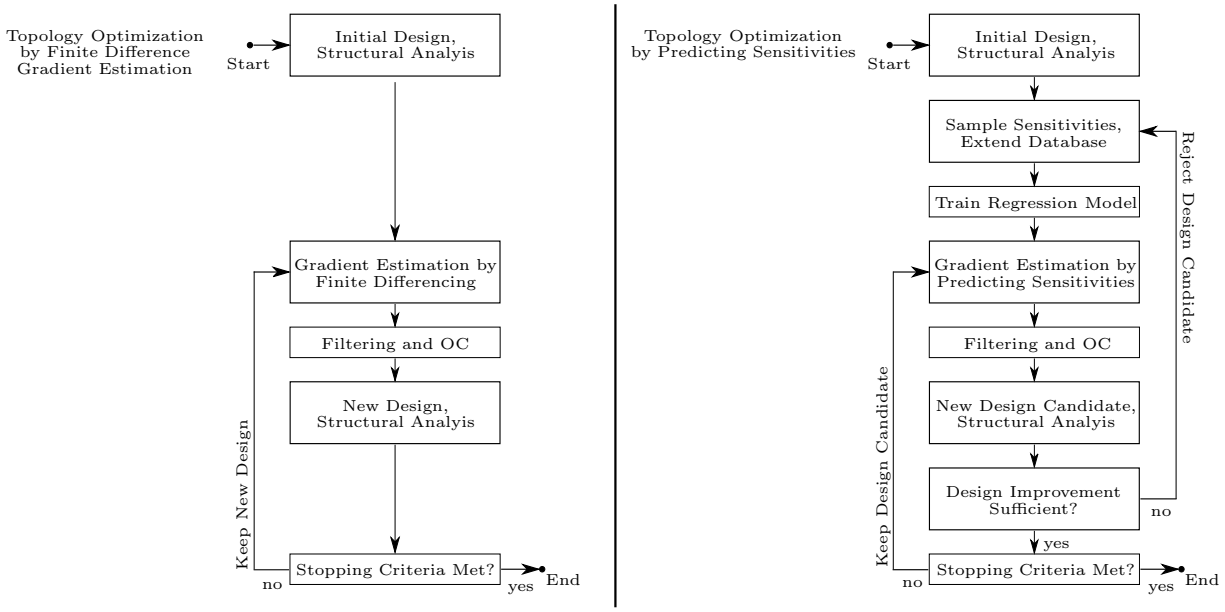

Figure 1: Comparison of the computational flow for topology optimization using pure finite difference gradient estimation (left) and topology optimization by predicting sensitivities (right).

4. The predicted sensitivities are filtered and the OC update is applied.

5. The quality of the new design candidate is assessed by a finite element analysis.

6. If the design is accepted, the design candidate becomes the current design and the algorithm continues in step 3. Otherwise the design candidate is rejected and either the design is converged and the algorithm terminates, or additional samples are taken for the current design, i.e. the algorithm continues in step 1.

In the following detailed information is provided on step 6 of the algorithm:

Check a) A design candidate is accepted and becomes the current design, if its percental improvement over the current design is equal or larger than the improvement rate f_r. Otherwise continue with check b).

Check b) If less than $N_{\mathrm{s_{req}}}$ samples have been taken for the current design, the design candidate is rejected and additional sensitivity samples are collected. Otherwise f_r is reduced according to $f_r^{\mathrm{new}} = \max(0.5 \cdot f_\mathrm{r}, f_{\mathrm{r_{min}}})$, with a minimum improvement rate $f_{\mathrm{r_{min}}}$. However, f_r is at maximum reduced once per design iteration. The design candidate is accepted and becomes the new current design if its quality is satisfying the new improvement rate, otherwise it is rejected again and it is continued with the check in c).

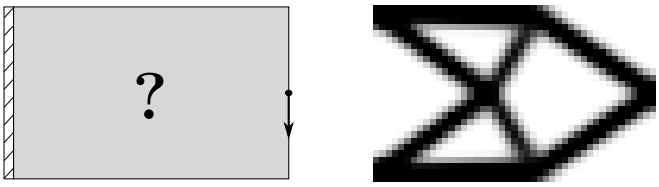

Figure 2: The design space with boundary conditions and the baseline design obtained by finite difference gradient estimation.

Check c) If the maximum number of samples $N_{\mathrm{s_{max}}}$ which can be taken per design iteration has not been reached, additional sensitivity samples are collected. Otherwise the algorithm terminates.

This slightly complicated convergence scheme is due to the fact that the optimal choice for N_s is unknown. In step 1 a fixed number of the design variables is sampled and added to the database. Iteratively, the algorithm updates the design and in case of failure the sampling step is repeated, so that only the minimum amount of samples for acceptance are required and the total number of finite element solver runs is kept low. However, when the design is not improving, this can either be caused by an actual convergence of the design or by insufficient quality of the predictions. Therefore, a limit $N_{\mathrm{s_{max}}}$ on the maximum amount of samples that can be taken is provided. Furthermore, it may happen that weak predictions still result in a marginal improvement of the design. In order to avoid this case f_{r} is specified, maintaining a higher quality of the predictions. As the optimization proceeds, the achievable improvement decreases and so f_{r} is reduced after $N_{\mathrm{s_{req}}}$ samples have been taken.

The computational flow for TOPS as well as for a baseline with pure finite differencing is depicted in Fig. 1.

4 EXPERIMENTS

4.1 Setup

In order to do a statistical evaluation of TOPS, the algorithm is applied to a 2-dimensional minimum compliance cantilever design problem. The minimum compliance problem is chosen in order to evaluate TOPS with respect to a well established baseline. In our implementation the efficient solver from [1] is used. The design space is discretized into $N = 45 \times 28 = 1260$ quadratic elements with a size of 1mm $\times$ 1mm. The following parameters are set: load $f = 1\mathrm{N}$, target volume fraction $V_f = 0.4$, Young's modulus $E_0 = 1\mathrm{N/mm^2}$, Poisson's ratio $\nu = 0.3$, penalization $p = 3$, filter radius $r_{\mathrm{min}} = 2.1\mathrm{mm}$. A baseline solution is obtained by gradient estimation via pure finite differencing, without sensitivity model. The design space with boundary conditions and the optimized baseline design are depicted in Fig. 2.

We evaluate TOPS for two different regression models [5], respectively sensitivity predictors. First, a linear regression is considered in order to test the hypothesis of linear

relations between LSF and the sensitivity. In order to avoid over fitting, a quadratic regularization term is added to the ordinary least squares problem of linear regression (termed ridge regression [8]). Secondly, a support vector regression (SVR) [13] as a state of the art nonlinear predictor is applied. The implementation is using the libsvm [6] and the recommendations from [9] are followed.

Both predictors are evaluated for several LSF sets. As explained in Sec. 3 we designate the elemental displacement vector and the density as basic LSF. This feature set represents a minimum amount of previous knowledge about the problem. In order to improve the performance of the linear regression often higher order features are constructed from the basic ones by multiplication of features with each other. Accordingly it is possible to construct second, third and fourth order features from the basic LSF. The SVR is intrinsically able to model nonlinear relations and is therefore only trained in combination with the linear versions of the basic LSF.

For the minimum compliance problem another LSF which can be used is the elemental strain energy, which is defined as:

$$W_i = x_i^p \mathbf{u}_i^{\mathrm{T}} \mathbf{k_0} \mathbf{u}_i. \tag{9}$$

This feature represents previous knowledge on the problem, as it contains direct knowledge on the material constants and the elemental stiffness matrix $\mathbf{k_0}$, as well as indirect knowledge by selecting this feature over other physical information on the element, like e.g. strains or stresses.

We chose to sample $N_s = 50$ samples in each sampling step. For the convergence parameters we chose $f_{\mathrm{r}} = 0.05$, $f_{\mathrm{r_{min}}} = 0.001$, $N_{\mathrm{s_{req}}} = 250$, $N_{\mathrm{s_{max}}} = 500$. Additionally, samples of the last 8 designs which have been sampled are remembered, samples older than that are neglected. All TOPS variants are started for 30 different random seeds in order to account for the stochastic nature of the sampling step. Two measures for evaluating TOPS are considered, the median compliance of the resulting optimized design $\overline{c}_{\mathrm{opt}}/\mathrm{mJ}$ and the computational cost measured by the median number of finite element analysis simulations, which were required N_{FEA}, compared to $N_{\mathrm{FEA,Base}}$ for the baseline. Additionally the resulting designs can be visually compared to the baseline design in Fig. 2.

4.2 Results TOPS: Linear Regression

The result of TOPS can be seen in Tab. 1. The designs resulting from TOPS with linear regression are depicted in Fig. 3. For the linear regression predictor, a trend of improving quality can be observed when using higher order variants of the basic LSF. For the basic features the linear model is not able to model the sensitivities appropriately resulting in a sub-optimal solution, which can be seen in Fig. 3. Even though a high number of samples is taken the linear mapping is not sufficient. Adding second, third and forth order terms of the basic LSF improves the compliance and reduces the required number of finite element analysis simulations. These features imply a multi-variate polynomial model in

	$\frac{\overline{c}_{\text{opt}}}{\text{mJ}}(\pm$ std)	$\frac{N_{\text{FEA}}}{1000}(\pm$std)	$\frac{N_{\text{FEA}}}{N_{\text{FEA,Base}}}(\pm$std)
TOPS Linear Regression			
Basic LSF	61.44(±1.54)	15.85(±4.04)	0.63(±0.16)
Basic LSF, 2nd Order	59.19(±1.25)	7.35(±2.41)	0.29(±0.10)
Basic LSF, 3rd Order	56.56(±1.32)	3.03(±0.80)	0.12(±0.03)
Basic LSF, 4th Order	56.50(±1.80)	3.09(±0.70)	0.12(±0.03)
Energy LSF	56.78(±1.14)	0.67(±0.38)	0.03(±0.02)
Energy and Basic LSF	55.05(±0.38)	2.62(±0.69)	0.10(±0.03)
TOPS Support Vector Regression			
Basic LSF	56.66(±21.26)	6.75(±2.29)	0.27(±0.09)
Energy LSF	56.14(±0.46)	2.06(±0.49)	0.08(±0.02)
Energy and Basic LSF	56.36(±0.90)	0.62(±0.19)	0.02(±0.01)
Baseline	55.52(±0.00)	25.22(±0.00)	1.00(±0.00)

Table 1: Results for TOPS with different predictors and sets of LSF from restarting each variant for 30 different random seeds.

the basic LSF with degree, two, three and four, which is able to much better approximate the sensitivity function. Compared to the baseline only 12% of the FEA simulations are required for third and fourth order LSF. The higher the order of the basic LSF that are used, the better the resemblance between the optimized design and the baseline.

Using the energy LSF W_i has a high impact on the results. The lowest compliance in the whole study is achieved when using the energy and basic LSF in combination with linear regression. This corresponds to the case where all considered information is available. Comparing this case to the case of only basic LSF it is obvious that the energy feature is very descriptive. Furthermore the lowest N_{FEA} for the linear regression predictor is achieved when only the energy is used as LSF. The lack of information causes the result to be less accurate, however the training of the models requires less samples, since with only one strong feature the linear relation between energy and sensitivity is easily learned. In numbers only 3% of $_{\text{FEA,Base}}$ is required. In case the energy is included as LSF, the designs are very similar to the baseline.

4.3 Results TOPS: SVR

Tab. 1 also contains the results for TOPS with SVR predictor. The corresponding optimized designs can be seen in Fig. 4.

When considering the results of TOPS with SVR in terms of compliance, all feature sets perform similar. This is also reflected in the visual resemblance of the optimized designs to each other and to the baseline design. Although the energy LSF contains much more information on the sensitivity the same information is also contained in the basic LSF. However, SVR is able to learn the required non-linear relation. This comes with

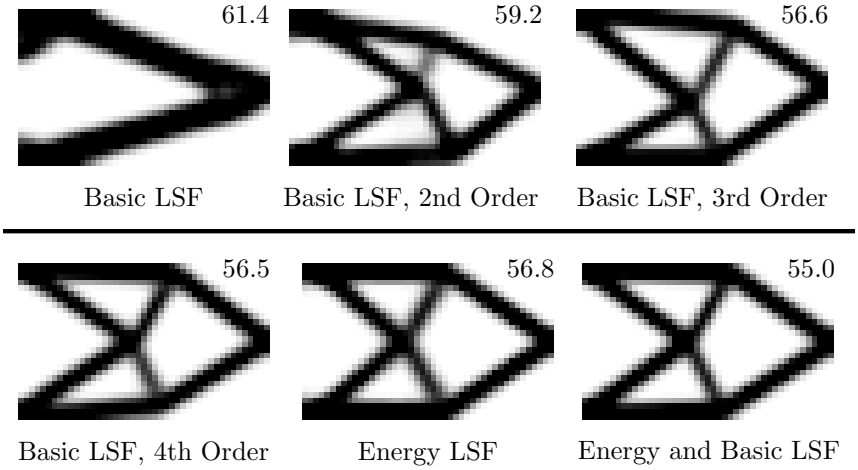

Figure 3: TOPS designs for linear regression as predictor with different LSF sets. For each set the optimized design from the run which converged closest to the median compliance is shown and its compliance /mJ is indicated next to it.

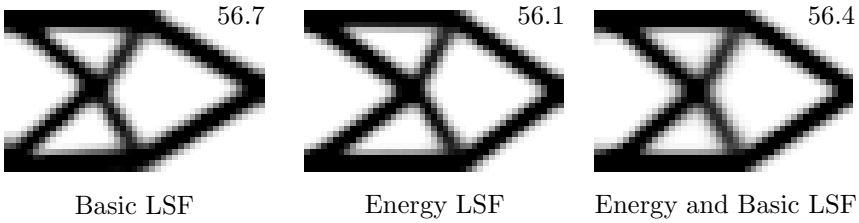

Figure 4: TOPS designs for SVR as predictor with different LSF sets. For each set the optimized design from the run which converged closest to the median compliance is shown and its compliance /mJ is indicated next to it.

a higher number of samples required for the optimization with only basic LSF, but still requires about 73% fewer FEA simulations than the baseline optimization. The overall cheapest approach of this study in terms of FEA simulations is TOPS with SVR and energy as well as basic LSF, which requires only 2.4% of the samples required by the baseline, i.e. the number of samples is reduced from 25220 to 620.

4.4 Discussion

From the results it can be seen that predictor and LSF have to be chosen carefully. The lowest compliance was achieved for linear regression with energy and basic LSF. The smallest amount of samples was required in case of the SVR for the same features.

Results can be interpreted when considering the analytical sensitivities. For our problem these are:

$$\frac{\partial c}{\partial x_i} = -p x_i^{p-1} \mathbf{u}_i^{\mathrm{T}} \mathbf{k_0} \mathbf{u}_i, \tag{10}$$

The true sensitivity is obviously strongly correlated with the strain energy in (9). This

explains why including the energy has such a high impact on the compliance for the linear regression as well as on the number of FEA simulations for the SVR. For the minimum compliance problem, using the elemental strain energy as a LSF is effectively an example for using analytical or semi-analytical gradient information.

Furthermore (10) reveals that for a constant stiffness matrix the analytical sensitivity (and $p = 3$) is a multivariate polynomial of degree 4. This explains why a linear regression with second, third or fourth order is improving the results substantially.

The energy is a strong and intuitive feature, and also higher order basic LSF are, by judging from (10), suited for modeling the minimum compliance sensitivity. For general problems no strong features, previous knowledge or intuition might be available. However, even in this case, TOPS with a SVR and only basic LSF yields a solution very similar to the baseline, with 73% less FEA simulations, since it is able to learn a suitable non-linear relation. The 6750 FEA simulation in this case seem still to be a high cost. However, since the sensitivity function is the same for all elements, it is intuitive to expect that the absolute number of samples required for learning will not rise with an increased mesh size, a hypothesis that needs to be validated by future research.

The results suggest that in general it is advisable to use a non-linear predictor like SVR and to include all LSF that might be relevant. If expert knowledge or intuition on the type of non-linearity is available this could be used to chose or construct additional LSF, like higher order versions of the basic LSF. In this case a linear regression might be more efficient than a non-linear model. Furthermore, any analytical or semi-analytical information can (and should) be easily utilized by TOPS.

5 CONCLUSIONS

The proposed algorithm is a novel heuristic for topology optimization, which substitutes analytical sensitivity information by a regression model. Sets of training data for a design are obtained by sampling finite differences for a subset of the design variables until the model quality is sufficient. The method is targeting problems for which analytical sensitivity information on the design variables is not available or difficult to obtain. However, in order to perform a statistical analysis the method was evaluated on the minimum compliance problem. Even for minimal previous problem knowledge, TOPS was demonstrated to provide designs close to the baseline, whereby best results for minimum previous knowledge were obtained when using a nonlinear SVR model. Including previous knowledge or intuition about the problem in the form of stronger LSF was demonstrated to improve the solution quality and reduce the number of solver runs.

REFERENCES

[1] E. Andreassen, A. Clausen, M. Schevenels, B. Lazarov, and O. Sigmund. Efficient topology optimization in matlab using 88 lines of code. *Structural and Multidisciplinary Optimization*, 43(1):1–16, 2011.

[2] N. Aulig and M. Olhofer. Neuro-evolutionary topology optimization of structures by utilizing local state features. To appear in *Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO'14*, Vancouver, BC, Canada, July 2014. ACM 2014.

[3] A. Baumgartner, L. Harzheim, and C. Mattheck. SKO (soft kill option): the biological way to find an optimum structure topology. *International Journal of Fatigue*, 14(6):387 – 393, 1992.

[4] M. Bendsøe and O. Sigmund. *Topology Optimization Theory, Methods and Applications*. Springer Verlag Berlin, 2nd edition, 2004.

[5] C.M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[6] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[7] J. Deaton and R. Grandhi. A survey of structural and multidisciplinary continuum topology optimization: post 2000. *Structural and Multidisciplinary Optimization*, 49(1):1–38, 2014.

[8] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67, 1970.

[9] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. A practical guide to support vector classification. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan, 2003.

[10] X. Huang and Y. Xie. *Evolutionary Topology Optimization of continuum structures*. John Wiley & Sons, 2012.

[11] N. M. Patel, B.-S. Kang, J. E. Renaud, and A. Tovar. Crashworthiness design using topology optimization. *Journal of Mechanical Design*, 131:061013, 2009.

[12] O. Sigmund. Morphology-based black and white filters for topology optimization. *Structural and Multidisciplinary Optimization*, 33(4-5):401–424, 2007.

[13] V. Vapnik, S. E. Golowich, and A. J. Smola. Support vector method for function approximation, regression estimation and signal processing. In *Advances in Neural Information Processing Systems 9 — Proceedings of the 1996 Neural Information Processing Systems Conference (NIPS 1996)*, pages 281–287, Dever, CO, USA, December 1997. MIT Press, Cambridge, MA, USA.