

**Topology optimization for linear thermo-mechanical transient problems
Modal reduction and adjoint sensitivities**

Hooijkamp, Evert; van Keulen, Fred

DOI

[10.1002/nme.5635](https://doi.org/10.1002/nme.5635)

Publication date

2018

Document Version

Final published version

Published in

International Journal for Numerical Methods in Engineering

Citation (APA)

Hooijkamp, E., & van Keulen, F. (2018). Topology optimization for linear thermo-mechanical transient problems: Modal reduction and adjoint sensitivities. *International Journal for Numerical Methods in Engineering*, 113(8), 1230-1257. <https://doi.org/10.1002/nme.5635>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Topology optimization for linear thermo-mechanical transient problems: Modal reduction and adjoint sensitivities

E. C. Hooijkamp | F. van Keulen 

Structural Optimization and Mechanics,
Delft University of Technology, Mekelweg
2, 2628 CD Delft, The Netherlands

Correspondence

Evert C. Hooijkamp, Faculty of 3mE, Delft
University of Technology, Mekelweg 2,
2628 CD Delft, The Netherlands.
Email: e.c.hooijkamp@tudelft.nl

Funding information

Pieken in de Delta, Provincie
Noord-Brabant; Samenwerkingsverband
Regio Eindhoven

Summary

This paper focuses on topology optimization for linear transient thermo-mechanical problems. The latter are, for example, encountered for extreme precision tools and instrumentation. Due to the transient nature, a standard adjoint sensitivity analysis will result in a backward transient analysis for the adjoint variables, leading to both storage and computational inefficiencies. A method is proposed that rigorously eliminates the backward transient integration for the adjoint sensitivity analysis. At the basis is a model-order reduction technique, which relies on a reduced thermal modal basis combined with static correction. The modal amplitudes can be readily obtained semi-analytically using simple convolutions. This accurate but reduced-order model is the starting point for an adjoint sensitivity analysis. Via a tactic selection of adjoint variables, the backward transient analysis for the adjoints is completely eliminated, whereas computational efficiency and consistency are maintained. The effectiveness of the resulting adjoint sensitivities and their application in topology optimization are demonstrated on the basis of several test examples.

KEYWORDS

sensitivity analysis, adjoint sensitivities, model-order reduction (MOR), transient problems, thermal modes, topology optimization

1 | INTRODUCTION

Transient thermo-mechanical problems can be found in a wide range of applications, for example, combustion processes, energy conversion, manufacturing, and instrumentation. This paper focuses on problems with relatively small temperature fluctuations resulting in relatively small deflections and deformations, the so-called thermal errors. The latter may have a major impact on the performance. Examples are instruments operating at the nanometer scale and precision (machine) tools. Design for thermal error reduction, especially with a transient nature, is very challenging due to the interaction of multiple physics domains. Therefore, a combination of complex material layout with active thermal management is required.^{1,2} Considering this complexity, use of topology optimization for design is an obvious choice.

Thermo-mechanical design problems have been considered in several studies applying topology optimization.³ Early work was done by Rodrigues and Fernandes,⁴ and with the development of different topology optimization methods, Li et al⁵ used an evolutionary method, Shiguang⁶ used a level-set method, and Sigmund⁷ applied a density-based method to solve thermo-mechanical problems. All studies show that challenges arise because of the multiple physics domains.

An excellent exploration of these challenges for density-based methods is given by Gao and Zhang.⁸ Later, Deaton and Grandi⁹ investigated various problem formulations.

The studies mentioned above all consider steady-state thermo-mechanical problems. Topology optimization has been applied to transient problems, however. For example, Turteltaub presented a topology optimization of material properties for transient heat problems¹⁰; Li et al presented a topology optimization procedure for compliant mechanisms subjected to transient thermal and/or mechanical loading¹¹; Matzen et al presented a topology optimization scheme for transient response of photonic crystal structures¹²; and Mello et al presented a topology optimization for transient problems to design electro-thermo-mechanical micro-actuators.¹³ These papers show that *transient problems* introduce a major drawback in the sensitivity analysis for topology optimization.¹⁴ Although the adjoint variable method is computationally much more efficient than other sensitivity methods for problems with many design variables, as in topology optimization, it requires a reverse transient analysis to compute the adjoint variables. This results in storage problems for large-scale systems as the time integration has to be performed sequentially. In addition, the approach is not well suited for parallel computing. Therefore, the sensitivity analysis for transient topology optimization problems is often performed using the direct method (see, eg, Pedersen¹⁵). A different solution is the application of the equivalent static load (ESL) methods.^{16–18} In ESL, the transient topology optimization problem is transformed into a topology optimization problem with multiple static loads. These ESLs, one computed for each time step, are load cases used to solve for static problems, which produces a response equal to the transient response. When performing the sensitivity analysis, the ESL is assumed to be independent of the design variables. Thus, only a static system needs to be considered for the adjoint sensitivity analysis and, consequently, the reverse analysis is avoided. However, the contribution of the dynamic terms to the sensitivity information is neglected. In the present paper, reduced-order models are investigated to overcome the difficulties caused by the backward time integration when applying the adjoint variable method.

For efficient computation of the transient response of large-scale linear systems, model-order reduction (MOR) has been applied. Model-order reduction aims at finding a lower-dimensional system that gives a high-quality approximation and reduces the computation time and storage requirements to solve the problem (see, eg, previous studies^{19–21}). The majority of the MOR techniques are based on projection and search for an appropriate problem subspace and the associated transformation basis. Similarly, these techniques are used to decrease the computational cost and storage of the sensitivity analysis. For example, Ahmed et al²² reduced the sensitivity analysis with a transformation basis obtained by an Arnoldi algorithm, whereas Proper Orthogonal Decomposition, also known as principal component analysis, was applied by Weickum et al²³ and Ilievski.²⁴ The main drawback in computing the sensitivity of any reduced system is to determine the sensitivity of the transformation basis, which is either computationally expensive or even impossible (eg, when the transformation basis is defined based on a heuristic method or an optimization).

In the present study, MOR is applied to reduce the system and, consequently, the sensitivity analysis. We distinguish 4 strategies to approximate the response sensitivities with a reduced system as presented in Figure 1. Strategy 1, already published by Fox and Kapoor in one study,²⁵ represents the transient response as a superposition of its modal basis and uses a direct method to compute the sensitivities. Although the response computation is reduced, the computation of the eigenvector sensitivities is very expensive. The sequence of strategy 1 in Figure 1 is MOR, approximate the response, differentiate, and solve for the sensitivities. Likewise, one may opt to apply MOR on the sensitivity equations rather than the response (strategy 2 in Figure 1). When based on a reduced modal basis, this avoids the computation of the eigenvector sensitivity and is, therefore, preferred as shown in Haftka and Kamat²⁶ among others. Thus, the sequence of strategy 2 in Figure 1 is differentiate, MOR, approximate the derivative, and solve for the sensitivities. However, both approaches are essentially direct methods and, thus, inefficient for a large number of design variables (see also other studies^{27,28}). Feng, Arora, and Haug²⁹ presented an algorithm applying the adjoint variable method (strategy 3 in Figure 1). After differentiation of the governing equations, the adjoint variable is defined such that the calculation of the computationally expensive derivatives is avoided. Next, the adjoint system is reduced with the adjoint variable represented as a superposition of the eigenvectors, and the sensitivity is computed with approximated adjoint variables. Hence, the sequence of strategy 3 in Figure 1 is define adjoint, MOR, approximate the adjoint variables, differentiate, and solve for the sensitivities.

When modal reduction (MR) is allied as a MOR technique, all 3 strategies as described above are effective only if a subset of modes is included, ie, applying MR. Careful selection of modes is needed to ensure a high-quality approximation. The publications by Feng et al²⁹ and by Haftka and Kamat²⁶ apply modal truncation as reduction method following strategies 3 and 2, respectively. However, both papers do not specify a clear truncation criterion nor consider the effect of the reduction on the sensitivity analysis. Greene and Haftka^{30,31} address these 2 topics for strategy 1. They concluded that small errors due to modal truncation might be magnified in the sensitivity analysis. To select the number of modes, they used the modal participations to the pseudo-load as an indication of the error. A more thorough study on the latter was conducted

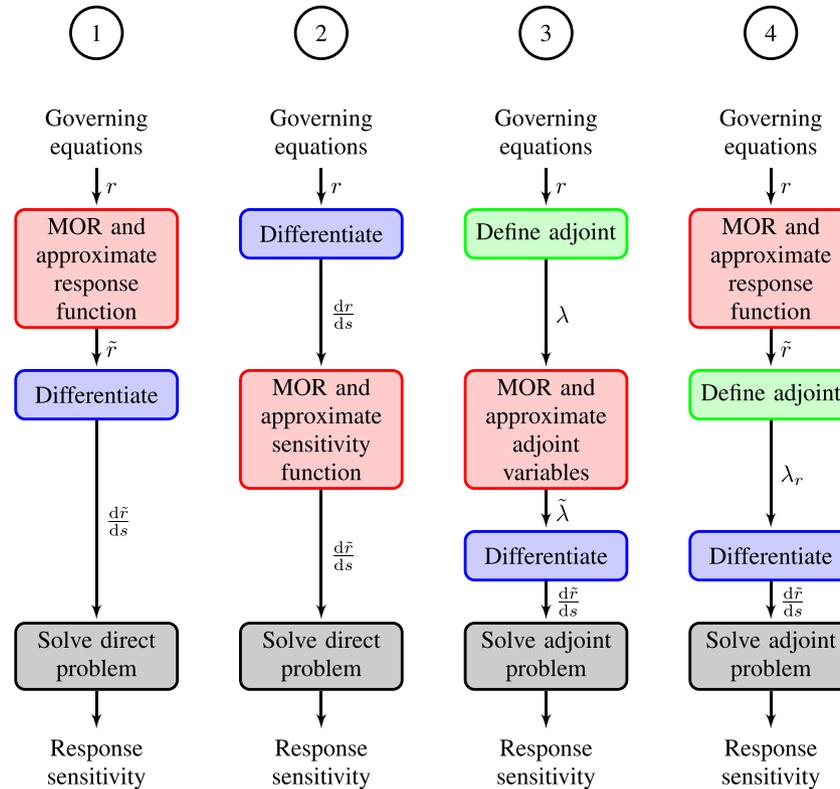


FIGURE 1 Different strategies to compute approximate response sensitivities of a transient problem based on model-order reduction. Strategy 1 approximates the governing equations to decrease the computational time. To, typically, avoid the computationally expensive eigenvector sensitivities, strategy 2 approximates the sensitivity function rather than the response. Both strategies use a direct differentiation, which is not efficient for problems introducing many design variables (eg, topology optimization) as compared to the number of response functions. Strategy 3 applies the adjoint variable method to handle a large number of design variables and uses a reduction method to approximate the adjoint variables and solve the adjoint problems. To avoid inconsistencies and computational inefficiencies of a reverse analysis in the case of transient problems, strategy 4 is presented in this paper. This strategy applies adjoint variable method in combination with the model-order reduction of the governing equations

by Paul, Dutta, and Ramakrishnan.^{32,33} They compared strategy 3, as proposed by Feng et al,²⁹ to an approach without MR but applying numerical time integration as presented by Hsieh and Arora.³⁴ They showed that the calculation of design derivatives depends highly on the number of basis functions and time steps taken into account. A proper choice of the time step and inclusion of sufficient modes is crucial for accurate computation of sensitivities. Furthermore, they point out that the forcing functions (ie, loads and pseudo-loads) are different for the response, its sensitivity, and the adjoint. Therefore, the proper basis may differ for each of them, resulting in inconsistent derivatives.

The truncation error can be decreased, while increasing computation time, by including the static contribution of the truncated part by applying the mode acceleration method (MAM) as in other studies.^{31,35} Dutta and Ramakrishnan³³ argue that design sensitivities approximated by MR are not reliable without this correction as, in their experience, the derivatives without the correction are unstable even when including a substantially large number of modes. In addition, the modal basis can be augmented with, for example, Ritz vectors (see, eg, other works^{31,35}). These augmentations to the basis are not considered in this paper.

Topology optimization in combination with MOR has been studied by several researchers. An early application was presented by Ma et al.³⁶ They performed topology optimization of a frequency response problem. Two sensitivity analysis approaches were discussed of which one was based on MR of the response and sensitivity (conform strategy 2). Ma et al concluded that the computations using the reduced sensitivities are more efficient when the sensitivities need to be calculated at many frequency points. However, they noted that the accuracy of the modal truncation is limited and improved techniques may be needed. Yoon³⁷ studied the effect of MOR techniques on density-based topology optimization. He applied the mode superposition and Ritz vector methods to approximate the response and adjoint variables, ie, strategy 3. He concluded that mode superposition methods can be inaccurate due to localized vibration modes (see, eg, Olhoff and

Du³⁸ for a discussion about localized eigenmodes). Better results were obtained with the Ritz vector methods. Recently, both Liu et al³⁹ and Zhao and Wang⁴⁰ investigated the effectiveness of modal truncation and the MAM to reduce computational time in density-based topology optimization for transient problems. They concluded that mere truncation is not suitable to approximate the adjoint variables due to low computing accuracy. Approximation with MAM, however, is very effective for problems under harmonic excitation with multiple frequencies and for problems with many time steps.

In summary, for the sensitivity analysis in topology optimization of linear transient thermo-mechanical problems, all 3 strategies in combination with MR have major disadvantages: strategies 1 and 2 are direct methods out ruling problems with large number of design variables; strategy 1 requires the computation of eigenvector sensitivities; strategy 3 needs a reverse analysis for the adjoint problem; and an additional selection of modes is needed in strategies 2 and 3. Moreover, strategies 2 and 3 lead to inconsistencies between the response evaluation and the associated design sensitivities.

Another approach to approximate the sensitivity analysis of dynamic systems is based on the modal representation of the system, shown as strategy 4 in Figure 1. Like in strategy 1, first, the governing equations are approximated based on a MOR technique, but the sensitivities are determined using the adjoint variable method to avoid the computation of the eigenvector derivatives. This approach has already been applied to problems optimizing the shape of a single or few modes for the design of resonating systems. For example, the sensitivity analysis of a distinct eigenvalue problem using the adjoint variable method is presented by Lee⁴¹ and for multiple eigenvalues.⁴² This modal-based sensitivity analysis was applied by Yu et al on, for example, an optimization of nodal lines for violin design,⁴³ and optimization of acoustic parameters of music bells.⁴⁴ The use of strategy 4 in topology optimization is presented by Tcherniak⁴⁵ for the design of resonating structures. In addition, Takezawa and Kitamura⁴⁶ used this sensitivity approach to match the eigenvectors to prescribed shapes optimizing the material design. These studies focus on the optimization of a single or a couple of mode shapes, whereas the present study aims to optimize a transient response.

In this paper, we aim for an efficient computation of sensitivities in topology optimization of *transient* thermo-mechanical problems. The proposed method follows strategy 4 and rigorously eliminates the backward time integration for the adjoint sensitivity analysis. At the basis is a MOR technique, which relies on a reduced thermal modal basis. Thus, the temperature response and induced displacements are approximated selecting a proper modal basis combined with a static correction. The resulting adjoint problem consists of a set of decoupled problems. The number of equations is equal to the number of modes in the modal basis. This decoupling makes the proposed approach very well suited for parallel computing.

In Section 2, the thermo-mechanical transient problem is defined representing the temperature and the induced displacements by mode superposition. In addition, the MOR techniques are described focusing on MR. Next, the dynamic response topology optimization problem is formulated in Section 3, including the optimization procedure and implementation. The main contribution of this paper is presented in Section 4 covering the sensitivity analysis. We show that the backward time integration for the adjoint variables can be eliminated when using mode superposition. In Section 5, the effect of MR on the sensitivities is discussed on the basis of a thermal optimization problem. We concluded that inclusion of the static correction (ie, applying the MAM) is essential for a high-quality approximation of the sensitivity information. Finally, the presented approach is applied on a thermo-mechanical topology optimization problem in Section 6. The proposed approach with sensitivities based on the combination of mode superposition and static correction is shown to result in similar optimized designs as the reference designs.

2 | THERMO-MECHANICAL MODEL

The thermo-mechanical systems considered allow to assume small temperature fluctuations with respect to a reference temperature, to assume linear material behavior, and to linearize the effect of convection and radiation. Furthermore, the mechanical systems are fast as compared to the thermal systems and mechanical loads are expected to vary relatively slow; therefore, the inertia and damping effects are neglected. Consequently, the thermo-mechanical coupling is one way, by thermal expansion only, and the thermo-mechanical system is modeled by a linear system of equations. This system is described by the uncoupled quasi-static thermo-elastic relations in combination with proper boundary and initial conditions (see, eg, other studies^{47,48}).

The structure is discretized in space by any well-known technique such as the finite element method, resulting in 2 systems of equations: one for the thermal system and one for the mechanical system. The thermal system is governed by the spatially discretized heat equation, which for N degrees of freedom can be written as^{48,49}

$$\mathbf{C}_T \dot{\boldsymbol{\theta}}[t] + \mathbf{K}_T \boldsymbol{\theta}[t] = \mathbf{q}[t], \quad (1)$$

where $\boldsymbol{\theta}$ denotes the temperature differences, with respect to a reference temperature, as a function of time and $\dot{\boldsymbol{\theta}}$ its time nodal derivative. The heat capacity matrix and heat conductivity matrix are indicated by \mathbf{C}_T and \mathbf{K}_T , respectively; both matrices are symmetric and positive definite. The linearized contributions of the heat convection and radiation are both included in the heat conductivity matrix \mathbf{K}_T . The nodal thermal load vector is denoted by \mathbf{q} and is assumed to be known and may be time dependent. In addition, the initial temperature at $t = 0$ is known as $\boldsymbol{\theta}[t = 0] = \boldsymbol{\theta}^0$.

The mechanical system needs to satisfy the spatially discretized quasi-static mechanical equilibrium equations^{48,49}:

$$\mathbf{K}_M \mathbf{u}[t] = \mathbf{f}[t] + \mathbf{f}_T[t], \quad (2)$$

where \mathbf{u} denotes the nodal kinematic degrees of freedom. The mechanical stiffness matrix is indicated by \mathbf{K}_M , which is symmetric and positive definite. \mathbf{f} and \mathbf{f}_T indicate the mechanical load vector and the equivalent load due to thermal expansion, respectively. The latter is linearly dependent on the temperature differences $\boldsymbol{\theta}$, as obtained from Equation 1, and is given by

$$\mathbf{f}_T[t] = \mathbf{A}\boldsymbol{\theta}[t], \quad (3)$$

where \mathbf{A} indicates the coupling matrix associated with thermal expansion.

Hence, the discretized thermo-mechanical system is governed by Equations 1 and 2, which can be solved sequentially. First, the temperature response is obtained from Equation 1, and next, the temperature response is used in Equation 2 to solve for the displacements. The first-order differential equation in Equation 1 can be solved using several methods.⁵⁰ Here, only time integration using mode superposition is discussed.

2.1 | Thermal mode superposition

Thermal mode superposition describes the transient behavior of the thermal system using the eigenvectors and corresponding eigenvalues (ie, the thermal modes and time constants). These are determined solving the eigenvalue problem associated with Equation 1 (with the eigenfunction: $\boldsymbol{\theta} = \boldsymbol{\phi}e^{-\lambda t}$):

$$(\mathbf{K}_T - \lambda_k \mathbf{C}_T) \boldsymbol{\phi}_k = \mathbf{0}, \quad (4)$$

where $\boldsymbol{\phi}_k$ denotes eigenvector and λ_k denotes the corresponding eigenvalue, ie, thermal mode and reciprocal of the associated thermal time constant ($\tau_k = \lambda_k^{-1}$), respectively. As the heat conductivity and heat capacity matrices are symmetric and positive definite, N independent modes can be identified. The modes are numbered such that $\tau_1 \geq \tau_2 \geq \dots \geq \tau_k \geq \dots \geq \tau_N$ and satisfy the orthonormality relations:

$$\begin{aligned} \boldsymbol{\phi}_k^T \mathbf{C}_T \boldsymbol{\phi}_l &= \delta_{kl}, \\ \boldsymbol{\phi}_k^T \mathbf{K}_T \boldsymbol{\phi}_l &= \frac{1}{\tau_k} \delta_{kl}, \end{aligned} \quad (5)$$

where δ_{kl} indicates Kronecker delta. Note that the normalization of the thermal mode shapes is taken with respect to the heat capacity matrix. Moreover, the N modes provide a complete basis.

The next step is to express the temperature as a superposition of thermal modes:

$$\boldsymbol{\theta}[t] = \sum_{k=1}^N \boldsymbol{\phi}_k \eta_k[t], \quad (6)$$

where η_k denote the time-dependent modal amplitudes. The modal amplitudes are determined by substituting Equation 6 into Equation 1 and pre-multiplying this equation by the transpose of a thermal mode ($\boldsymbol{\phi}_j^T$). Using the orthonormality relations of Equation 5, this results in N independent differential equations, ie, one per thermal mode:

$$\dot{\eta}_k[t] + \frac{1}{\tau_k} \eta_k[t] = q_k[t] = \boldsymbol{\phi}_k^T \mathbf{q}[t], \quad (7)$$

where q_k is the modal thermal load, ie, the excitation of thermal mode k by the thermal load. The general semi-analytic solution for modal amplitude η_k is

$$\eta_k[t] = \eta_k^0 e^{-\frac{t}{\tau_k}} + \int_0^t q_k[\gamma] e^{-\frac{t-\gamma}{\tau_k}} d\gamma, \quad (8)$$

here, η_k^0 is the initial modal amplitude ($\eta_k^0 = \phi_k^T \mathbf{C} \theta^0$). Note that this solution does account for zero eigenvalues (ie, for an isolated thermal system, the time constant describing the temperature rise in the entire system).

The final step is to compute the transient displacement response caused by the temperature fluctuations and mechanical loads. Because Equation 2 is linear, the displacement vector \mathbf{u} can be expressed as a linear combination of 2 responses:

$$\mathbf{u}[t] = \mathbf{u}_M[t] + \mathbf{u}_T[t], \quad (9)$$

with

$$\mathbf{u}_M[t] = \mathbf{K}_M^{-1} \mathbf{f}[t] \quad (10)$$

and

$$\mathbf{u}_T[t] = \mathbf{K}_M^{-1} \mathbf{A} \theta[t]. \quad (11)$$

The latter is the instantaneous response to the temperature fluctuations θ .

It is emphasized that so far, the modal description is complete and the only assumptions accepted are related to the linearity, the static mechanical response, and, thus, the relatively small time constants associated with the mechanics.

2.2 | Modal reduction

A major advantage of the modal representation is that the thermal model can be reduced by selecting a subset of thermal modes, ie, MR. Thus, the temperature response can be approximated by

$$\theta[t] \approx \hat{\theta}[t] = \sum_{\mathbf{m}} \phi_{\mathbf{m}} \eta_{\mathbf{m}}[t], \quad (12)$$

where \mathbf{m} indicates a subset of mode numbers, with length $R \ll N$. The key is to select a small set of modes that still gives a high-quality approximation.

A common approach is to truncate the faster thermal modes, thus selecting only the thermal modes with the higher time constants (ie, $\mathbf{m} = [1, 2, \dots, R]$):

$$\hat{\theta}[t] = \sum_{m=1}^R \phi_m \eta_m[t]. \quad (13)$$

This approach is known as mode truncation method or, in structural dynamics, mode displacement method. The selection of R can be based on the experience of the engineer or on a cutoff criteria. Several cutoff criteria were proposed in, for example, literature.^{32,51,52} However, it is noted that also, other strategies may be used (see, eg, other works⁵³).

The thermally induced displacements obtained with the reduced modal basis is the instantaneous response to the approximated temperature fluctuations and given by

$$\mathbf{u}_T[t] \approx \hat{\mathbf{u}}_T[t] = \mathbf{K}_M^{-1} \mathbf{A} \hat{\theta}[t]. \quad (14)$$

In practice, the convergence of the mode truncation method with respect to the number of modes is often slow, in particular, for thermal mode superposition. The slow convergence is due to the spatial and/or temporal nature of the thermal loads. To avoid the computation of a high number of modes, other approaches have been developed such as MR with static correction.

2.2.1 | MR with static correction

The MR method with static correction (RC) takes into account the entire basis but ignores the dynamic behavior of all the modes omitted in \mathbf{m} (see, eg, Besselink²¹ and Maddox⁵¹). It is usually presented as an extension of the mode truncation method correcting for the truncated modes, but the same idea can be used for another selection of modes.

The RC approximates the response as before including an additional correction term, ie,

$$\hat{\theta}[t] = \sum_{\mathbf{m}} \phi_{\mathbf{m}} \eta_{\mathbf{m}}[t] + \theta_{\text{cor}}[t]. \quad (15)$$

The static correction term θ_{cor} is obtained substituting the approximation into Equation 1 and using Equation 5, the expression for the uncoupled modal amplitudes. This results in

$$\theta_{\text{cor}}[t] = \left(\mathbf{K}_T^{-1} - \sum_{\mathbf{m}} \tau_{\mathbf{m}} \phi_{\mathbf{m}} \phi_{\mathbf{m}}^T \right) \mathbf{q}[t]. \quad (16)$$

Note that this correction term is time dependent and orthogonal to the reduced modal basis. If we compare Equation 15 with Equation 6, the omitted modes in Equation 6 are included in Equation 15; however, their effective thermal time constant have been set to zero. The thermally induced displacements are obtained by Equation 14 as the instantaneous response to the approximated temperature fluctuations.

The convergence of the RC with respect to the number of modes is faster than MR, but, at a cost of solving a static system, the first term in Equation 16. However, the factorization of the matrix \mathbf{K}_T is already performed to solve the eigenvalue problem and can be reused.

2.2.2 | Dominant mode selection

For thermal systems, the mode truncation method is less effective as compared to structural dynamic systems due to the delayed transfer through the system. Therefore, a smart selection of modes is needed. For example, Zhu et al used a subset of dominant thermal modes.⁵⁴ These dominant modes are identified using a modal weighting, which is defined as the magnitude of the time constant times the modal thermal load (ie, $w_{\mathbf{m}} = |\tau_{\mathbf{m}} q_{\mathbf{m}}|$). Others have applied selection approaches using so-called modal excitation, modal controllability, or modal observability to identify the most significant thermal modes of the system (see, eg, Koevoets et al⁵⁵). Currently, there is no general accepted method to select the modes that need be included in the reduction.²⁸

In this paper, only the relevant part of the response is taken into account (as in Hooijkamp et al⁵³), which is determined considering 3 indicators:

1. modal excitation by the thermal load or the initial temperature distribution;
2. modal observability on the objective; and
3. modal participation within the time frame of interest.

The relevance of a mode is determined combining these 3 indicators and the most relevant modes identified. Consequently, the response is expressed in terms of a relevance-based modal basis. This basis reduces the thermal model and can provide designers profound inside to improve their designs.

Because the indicators are determined using the modes themselves, a larger number of modes have to be computed before selection. Therefore, this approach may not have a computational advantage over the truncation method to determine the response. However, the computational advantage is obtained in the sensitivity analysis as only the sensitivities associated with the relevant modes need to be considered.

3 | DYNAMIC RESPONSE TOPOLOGY OPTIMIZATION

3.1 | Dynamic response optimization

In general, a dynamic response optimization for a spatially discretized thermo-mechanical system can be formulated as

$$\begin{aligned} \min_{\mathbf{s}} J[\mathbf{s}, \boldsymbol{\theta}, \mathbf{u}, t] & \quad \text{for } t \in [0, t_E] \\ \text{s.t. } H_a[\mathbf{s}, \boldsymbol{\theta}, \mathbf{u}, t] &= 0 \quad \text{for } a = 1, \dots, N_h \\ G_b[\mathbf{s}, \boldsymbol{\theta}, \mathbf{u}, t] &\leq 0 \quad \text{for } b = 1, \dots, N_g \\ s_i^L &\leq s_i \leq s_i^U \quad \text{for } i = 1, \dots, N_s, \end{aligned} \quad (17)$$

where J denotes the objective function, s_i denotes a design variable with a lower and upper bound indicated by s_i^L and s_i^U , respectively, and \mathbf{s} indicates the design vector containing all design variables (ie, $\mathbf{s} = [s_1, s_2, \dots, s_{N_s}]^T$). In addition, H_a

and G_b denote equality and inequality constraints, respectively. For a transient thermo-mechanical problem, the objective and constraints may dependent on the design variables \mathbf{s} , time t , the temperature differences $\theta[t]$ with respect to a reference temperature, and the mechanical degrees of freedom $\mathbf{u}[t]$. The number of design variables, equality, and inequality constraint functions is indicated by N_s , N_h , and N_g , respectively.

The objective and the constraints can be related to a certain time instant, a time frame, or a combination of both. For that reason, we consider a general response function $r[\mathbf{s}]$, which could represent the objective function or any of the constraints (see, eg, Choi and Kim²⁷ and Kang et al²⁸). This general response function is defined as

$$r[\mathbf{s}] = \int_0^{t_E} w[t] R[\mathbf{s}, \theta, \mathbf{u}, t] dt, \quad (18)$$

where w represents a weighting function in time and t_E defines the end time. The weighting function defines the importance of the response function R in time.

3.2 | Topology optimization

Topology optimization is an automated design process that places material within a prescribed design domain to obtain the optimal structural performance. After the introduction of topology optimization for mechanical design problems, several different approaches were developed⁵⁶ for a wide range of physical disciplines.³ The present paper uses density-based topology optimization to optimize a transient thermo-mechanical problem as above.

3.2.1 | Design variables and material interpolation

Developed by Bendsøe in previous study,⁵⁷ the current popular topology optimization approach is the density-based approach. In this approach, each element e is associated with a single design variable ρ_e , also known as pseudo-density, that indicates the amount of material of element e . The design variable varies continuously between 0 and 1, where $\rho_e = 0$ indicates void material and $\rho_e = 1$ indicates solid material. For $0 < \rho_e < 1$, the material is interpolated. Different interpolation schemes have been developed over the years. In this study, the rational approximation of material properties (RAMP) is applied that is recommended for thermo-mechanical problems by Gao and Zhang.⁸ Then, the global system matrices of both thermal and mechanical systems (Equations 1 and 2) are obtained by assembly of the interpolated element matrices:

$$\begin{aligned} \mathbf{C}_T &= \sum_{e=1}^{N_{el}} \frac{c_e}{c_s} [\mathbf{C}_T]_e, & \mathbf{K}_T &= \sum_{e=1}^{N_{el}} \frac{k_e}{k_s} [\mathbf{K}_T]_e, \\ \mathbf{K}_M &= \sum_{e=1}^{N_{el}} \frac{E_e}{E_s} [\mathbf{K}_M]_e, & \mathbf{A} &= \sum_{e=1}^{N_{el}} \frac{\beta_e}{\beta_s} [\mathbf{A}]_e, \end{aligned} \quad (19)$$

where $[\mathbf{M}]_e$ indicates the element matrix of system matrix \mathbf{M} associated with element e and N_{el} indicates the number of elements. The material properties—specific heat capacity, heat conductivity coefficient, Young modulus, and thermal stress coefficient—are denoted by c , k , E , and β ($\beta = E\alpha$, with thermal expansion coefficient α),⁸ respectively. These properties are interpolated according to the RAMP. For a material property p and element e , this interpolation is given by

$$p_e = \frac{\rho_e}{1 + S_p(1 - \rho_e)} (p_s - p_v) + p_v, \quad (20)$$

where p_v and p_s indicate the value of the material parameter for $\rho = 0$ (void) and $\rho = 1$ (solid), respectively, and S_p denotes the penalization parameter. Note that the penalization parameter S_p may differ between material properties.

In density-based topology optimization, the element densities are the design variables s_i . As some of the element densities are predefined, the number of densities included in the design vector \mathbf{s} may differ from the number of elements N_{el} . The elements associated with the densities in the design vector \mathbf{s} form the design domain. For convenience, we assume in this topology optimization study that design variables only consist of element densities, ie, the number of design variables N_s is equal to the number of element densities included in the design vector \mathbf{s} .

3.2.2 | Objective

The objective of the considered thermo-mechanical problem is to minimize the thermal error, ie, the undesired displacements and/or displacement differences due to temperature fluctuations. Avoiding thermal errors is of high interest for, for example, precision devices. The objective function j is defined as

$$j[\mathbf{s}] = \frac{1}{t_E} \int_0^{t_E} J[\mathbf{s}, \mathbf{u}, \boldsymbol{\theta}, t] dt, \quad (21)$$

with

$$J[\mathbf{s}, \mathbf{u}, \boldsymbol{\theta}, t] = (u_{\text{interest}}[\boldsymbol{\theta}, t] - u_{\text{observed}}[\boldsymbol{\theta}, t])^2. \quad (22)$$

The thermal error is given as the difference between the displacement of interest u_{interest} and the observed displacements u_{observed} (eg, the measurement or reference area). Both displacements are scalar functions of time, which are obtained by a linear combination of the nodal displacements. Thus, the displacement difference can be written as a vector multiplication: $u_{\text{interest}} - u_{\text{observed}} = \mathbf{p}^T \mathbf{u}$, where \mathbf{p} is a vector selecting the displacements. Then, the objective function can be rewritten as

$$J[\mathbf{s}, \mathbf{u}, \boldsymbol{\theta}, t] = \mathbf{u}^T \mathbf{p} \mathbf{p}^T \mathbf{u}. \quad (23)$$

The function J denotes the thermal error squared to avoid cancellation of positive and negative errors.

3.2.3 | Volume constraint

To limit the amount of material, the used volume is constrained. The limit is based on the volume fraction ν , where a full solid domain equals $\nu = 1$. Then, the constraint is given as

$$\frac{\nu}{\nu_{\text{lim}}} - 1 \leq 0, \quad \text{with } \nu = \frac{1}{N_{el}} \sum_{i=1}^{N_{el}} \rho_i, \quad (24)$$

where ν indicates the volume fraction of the current design and ν_{lim} denotes the volume fraction bound. Note that in this study, using one material, the mass and volume fractions are equivalent.

3.2.4 | Topology optimization problem

Combining the equations above, the dynamic response topology optimization problem can be written as

$$\begin{aligned} \min_{\mathbf{s}} \quad & \frac{1}{t_E} \int_0^{t_E} (\mathbf{u}^T \mathbf{p} \mathbf{p}^T \mathbf{u}) dt, \\ \text{s.t.} \quad & \mathbf{K}_M \mathbf{u} = \mathbf{f} + \mathbf{A} \boldsymbol{\theta} \quad \text{for } t \in [0, t_E], \\ & \mathbf{C}_T \dot{\boldsymbol{\theta}} + \mathbf{K}_T \boldsymbol{\theta} = \mathbf{q}, \\ & \frac{\nu}{\nu_{\text{lim}}} - 1 \leq 0, \\ & 0 \leq s_i \leq 1 \quad \text{for } i = 1, \dots, N_s, \end{aligned} \quad (25)$$

where s_i are the pseudo-densities ρ_e of the elements in the design domain.

To solve the optimization problem given in Equation 25, the use of a gradient-based optimizer is currently common practice. Hence, a sensitivity analysis is needed. However, the sensitivity analysis for this problem is computationally inefficient.¹⁴ On the one hand, the adjoint variable method is preferred because of the high number of design variable. On the other hand, the adjoint variable method gives rise to a backward time integration, ie, an additional transient analysis.

In this study, we propose to use MR as described in Section 2. Besides an efficient computation of the transient response for large-scale systems, MR rigorously eliminates the backward time integration for the adjoint sensitivity analysis as will be shown in Section 4. Using Equations 4, 5, 8, and 15, the considered dynamic response topology optimization problem is written using the approximated responses and is given by

$$\begin{aligned}
 \min_{\mathbf{s}} \quad & \frac{1}{t_E} \int_0^{t_E} (\hat{\mathbf{u}}^T \mathbf{p} \mathbf{p}^T \hat{\mathbf{u}}) dt, \\
 \text{s.t.} \quad & \mathbf{K}_M \hat{\mathbf{u}} = \mathbf{f} + \mathbf{A} \hat{\boldsymbol{\theta}} \quad \text{for } t \in [0, t_E], \\
 & \hat{\boldsymbol{\theta}} = \sum_m \phi_m (\eta_m - \tau_m \phi_m^T \mathbf{q}) + \mathbf{K}_T^{-1} \mathbf{q}, \\
 & \eta_m = \eta_m^0 e^{-\frac{t}{\tau_m}} + \int_0^t (q_m e^{-\frac{t-\gamma}{\tau_m}}) d\gamma \quad \forall m, \\
 & \left(\mathbf{K}_T - \frac{1}{\tau_m} \mathbf{C}_T \right) \phi_m = 0, \\
 & \phi_m^T \mathbf{C}_T \phi_m = 1, \\
 & \frac{v}{v_{\text{lim}}} - 1 \leq 0, \\
 & 0 \leq s_i \leq 1 \quad \text{for } i = 1, \dots, N_s,
 \end{aligned} \tag{26}$$

where m contains a small subset of thermal modes to generate a high-quality approximation of the temperature response $\boldsymbol{\theta}$ and the associated displacements \mathbf{u} . In addition, note that the mode truncation method with static correction (ie, MAM) is applied. When the mode truncation method (without static correction) is used, then Equation 15 is replaced by Equation 13.

3.2.5 | Filtering

The so-called checkerboard problem in topology optimization refers to patches of alternating solid and void elements favored by the optimization process due to the limited finite element modeling. To avoid the forming of checkerboard patterns and other mesh-dependent phenomena, a density filter is used (see, eg, Sigmund⁵⁸). The filtered density $\bar{\rho}_i$ of element i is given as

$$\bar{\rho}_i = \frac{\sum_{e=1}^{N_{el}} w_{ie} \rho_e}{\sum_{e=1}^{N_{el}} w_{ie}} \quad \text{for } i \in [1, N_s], \tag{27}$$

with

$$w_{ie} = \max(0, r_{\text{min}} - \text{dist}(e, i)). \tag{28}$$

where w_{ie} denotes the contribution of the original density ρ_e in the filtered density $\bar{\rho}_i$, r_{min} indicates the filter radius, and $\text{dist}(e, i)$ gives the distance between element e and i . Note that an element's filtered density is the distance-weighted average of its own density and the density of its neighboring elements.

The density filter is only used on the design variables, ie, the element densities that are part of the design domain. In addition, this filter is conducted directly after the optimization step, and thus, the filtered densities become the input variables used to obtain the material properties in Equation 20, the volume constraint in Equation 24 and other responses.

3.2.6 | Penalization of intermediate densities

Intermediate density material (material of elements with $0 < \rho_e < 1$) may perform superior over solid or void material (see also Deaton and Grandhi⁹). However, designs without intermediate densities, so-called black/white designs, are preferred because the interpretation of the design areas with intermediate densities is not evident. In the considered optimization, material interpolation alone does not result in black/white designs. Therefore, a continuation step is adopted. After convergence, first, the design is projected to a black/white design. As the projected design is not optimal, the iteration is continued with penalized intermediate densities to ensure a black/white design.

For the projection, a bound is set and all densities below are projected to void and above to solid. Hence, the design after projection $\hat{\rho}$ is given by

$$\hat{\rho}_e = \begin{cases} 1 & \text{for } \bar{\rho}_e > \rho^* \\ 0 & \text{for } \bar{\rho}_e \leq \rho^* \end{cases} \quad \forall e, \tag{29}$$

where ρ^* indicates the density bound. Next, the extra optimization is performed with penalization of intermediate densities. Therefore, a penalty term is added to the objective j (Equation 21) and the objective function j^* for the second optimization is written as⁵⁶

$$j^*[\mathbf{s}] = j[\mathbf{s}] + j_{gp}[\mathbf{s}], \quad (30)$$

with the penalty term given by

$$j_{gp} = \frac{C_{gp}}{N_s} \sum_{i=1}^{N_s} s_i (1 - s_i). \quad (31)$$

where C_{gp} denotes the weighting factor with respect to the initial objective j . Note that other researches have used this term as a constraint rather than an addition to the objective (see, eg, Sigmund⁵⁶ and Sigmund and Petersson⁵⁹). The main difficulty is to determine the weighting factor C_{gp} to ensure smooth convergence towards a black-white design without losing focus of the main goal. To avoid defocussing, the weighting factor is gradually increased during the optimization from 4 to 32 by doubling its value every 50 iterations.

In summary, to obtain black/white designs, the following strategy is adopted:

1. Optimizing for objective j (Equation 21, without gray penalization) until convergence;
2. Applying the black/white projection (Equation 29);
3. Optimizing for objective j^* (Equation 30) with a gradually increasing weighting factor.

3.2.7 | Numerical implementation

In this study, the computational algorithms are developed in Matlab to ensure full control of the sensitivity information. The finite element analysis is implemented and our structure is discretized using linear 4-noded rectangular elements. These elements are used for both thermal and mechanical domain. A trapezoidal scheme is implemented for the time integration, for example, to compute the objective function j (Equation 21). Furthermore, as an optimizer, we used Svanberg's method of moving asymptotes (MMA) (see Svanberg⁶⁰ for a detailed description).

4 | DESIGN SENSITIVITY ANALYSIS

In design optimization, sensitivities of objective and constraints with respect to the design variables are often used. Different approaches to obtain the desired design sensitivities can be found in literature, as discussed before. In this section, we present the sensitivity analysis based on MR and the adjoint variable method (following strategy 4 as shown in Figure 1). The result is an effective approach avoiding the backward time integration. The focus is on sensitivities for discretized transient thermo-mechanical models based on reduced thermal modal basis combined with static correction as described in Section 2. However, this approach is not limited to these models as will be discussed in the last part of this section.

4.1 | Design and response sensitivities

In this study, the objective and constraints are expressed by mode superposition and then differentiated with respect to the design variables. When a general response function r , as given in Equation 18, is written in terms of approximated responses, it is also approximated and given by

$$r[\mathbf{s}] \approx \hat{r}[\mathbf{s}] = \int_0^{t_E} w[t] R[\mathbf{s}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}, t] dt. \quad (32)$$

The direct method and adjoint variable method are analytical methods to determine the sensitivity. For both methods, the derivative of the general response function is first written as

$$\frac{d\hat{r}}{ds_i} = \int_0^{t_E} w \left(\frac{\partial \hat{R}}{\partial s_i} + \frac{\partial \hat{R}^T}{\partial \mathbf{u}} \left(\frac{d\mathbf{u}_M}{ds_i} + \frac{d\hat{\mathbf{u}}_T}{ds_i} \right) + \frac{\partial \hat{R}^T}{\partial \boldsymbol{\theta}} \frac{d\hat{\boldsymbol{\theta}}}{ds_i} \right) dt, \quad (33)$$

with $\hat{R} = R[\mathbf{s}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}, t]$ and where d indicates the total derivative and ∂ indicates the partial derivative. The differentiation of a scalar function, like R , with respect to a vector, like $\boldsymbol{\theta}$, denotes the differentiation of the scalar with respect to each

of the vector components, ie, $\frac{\partial R}{\partial \theta} \equiv \left[\frac{\partial R}{\partial \theta_{(1)}} \dots \frac{\partial R}{\partial \theta_{(N)}} \right]^T$. Furthermore, the weighting function w , as defined in Equation 18, is assumed to be known and independent of the design variables. In addition, the variation of the end time t_e is not included, the case where it is included can be found in Choi and Kim.²⁷

The major disadvantages of these direct derivatives are (1) all system derivatives are required and (2) all derivatives of the modes in the reduced basis are required. These disadvantages can be eliminated using an adjoint formulation.

4.2 | Adjoint variable method (strategy 4)

The derivatives of the approximate transient responses, the temperature differences $\hat{\theta}$, and the mechanical degrees of freedom $\hat{\mathbf{u}}$, with respect to the design variables \mathbf{s} , include the computational expensive derivatives of the mode shapes. To avoid the computation of these expensive derivatives and the derivatives of the system responses, the adjoint variable method is used. However, transient responses lead to backward time integration for the adjoint variables and cause both storage and computational inefficiencies. Inclusion of the semi-analytic solutions for the modal amplitudes (Equation 8) in the augmentation completely eliminates the backward transient analysis. This will be shown in the following.

To apply the adjoint variable method, the general response function is augmented by the governing equations of all the responses included in the function. Therefore, to use the adjoint variable method on the considered problem, Equation 32 is augmented with Equation 4, 5, 8, 11, and 15:

$$\begin{aligned} \hat{r}^* = \hat{r} & - \int_0^{t_E} \left\{ \boldsymbol{\beta}^T (\mathbf{K}_M \mathbf{u}[t] - \mathbf{f} - \mathbf{A}\boldsymbol{\theta}) \right. \\ & + \boldsymbol{\alpha}^T \left(\boldsymbol{\theta} - \sum_m \phi_m (\eta_m - \tau_m \phi_m^T \mathbf{q}) - \mathbf{K}_T^{-1} \mathbf{q} \right) \\ & \left. + \sum_m \omega_m \left(\eta_m - \eta_m^0 e^{-\frac{t}{\tau_m}} - \int_0^t (q_m e^{-\frac{t-\gamma}{\tau_m}}) d\gamma \right) \right\} dt \\ & - \sum_m \left(\lambda_m^T \left(\mathbf{K}_T - \frac{1}{\tau_m} \mathbf{C}_T \right) \phi_m \right) \\ & - \sum_m \left(\mu_m \left(\frac{1}{2} - \frac{1}{2} \phi_m^T \mathbf{C}_T \phi_m \right) \right), \end{aligned} \quad (34)$$

where $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, ω_m , λ_m , and μ_m denote the adjoint variables.

Next, the sensitivity of this augmented response function with respect to a design variable is determined and the adjoint variables are chosen requiring that all the terms including the derivatives of the system response, the thermal modes, and the time constants cancel out. Then, the adjoint problems are obtained as

$$\begin{aligned} \boldsymbol{\beta}[t] & = \mathbf{K}_M^{-1} \left(w[t] \frac{\partial \hat{R}}{\partial \mathbf{u}}[t] \right), \\ \boldsymbol{\alpha}[t] & = \left(w[t] \frac{\partial \hat{R}}{\partial \boldsymbol{\theta}}[t] + \mathbf{A}^T \boldsymbol{\beta}[t] \right), \\ \omega_m[t] & = \boldsymbol{\alpha}^T[t] \phi_m, \end{aligned} \quad (35)$$

and

$$\begin{bmatrix} \left(\mathbf{K}_T - \frac{1}{\tau_m} \mathbf{C}_T \right) & -\mathbf{C}_T \phi_m \\ -\phi_m^T \mathbf{C}_T & 0 \end{bmatrix} \begin{bmatrix} \lambda_m \\ \mu_m \end{bmatrix} = \begin{bmatrix} \mathbf{a}_m \\ b_m \end{bmatrix}, \quad (36)$$

with

$$\begin{aligned} \mathbf{a}_m & = \int_0^{t_E} \left\{ (\eta_m - \tau_m \phi_m^T \mathbf{q}) \boldsymbol{\alpha} - \tau_m (\boldsymbol{\alpha}^T \phi_m) \mathbf{q} + \omega_m \left(\mathbf{C}_T \boldsymbol{\theta}^0 e^{-\frac{t}{\tau_m}} + \int_{\gamma=0}^t \mathbf{q}[\gamma] e^{-\frac{t-\gamma}{\tau_m}} d\gamma \right) \right\} dt, \\ b_m & = \int_0^{t_E} \left\{ (\tau_m^2 \boldsymbol{\alpha}^T \phi_m \phi_m^T \mathbf{q}) - \omega_m \left(\eta_m^0 t e^{-\frac{t}{\tau_m}} + \int_{\gamma=0}^t q_m[\gamma] e^{-\frac{t-\gamma}{\tau_m}} (t-\gamma) d\gamma \right) \right\} dt. \end{aligned} \quad (37)$$

Note that the transient analysis needed to compute the pseudo-load (\mathbf{a}_m, b_m) is forward and that the adjoint problem, Equation 36, involves M independent problems, one per included mode.

Finally, the sensitivity of the augmented performance measure is obtained as

$$\begin{aligned} \frac{d\hat{r}^*}{ds} = \frac{d\hat{r}}{ds} = & \int_0^{t_E} \left\{ w \frac{\partial \hat{R}}{\partial s_i} \right. \\ & - \boldsymbol{\beta}^T \left(\frac{d\mathbf{K}_M}{ds_i} \mathbf{u} - \frac{\partial \mathbf{f}}{\partial s_i} - \frac{d\mathbf{A}}{ds_i} \boldsymbol{\theta} \right) \\ & + \boldsymbol{\alpha}^T \left(\mathbf{K}_T^{-1} \frac{d\mathbf{q}}{ds_i} - \mathbf{K}_T^{-1} \frac{d\mathbf{K}_T}{ds_i} \mathbf{K}_T^{-1} \mathbf{q} - \sum_m \tau_m \phi_m \phi_m^T \frac{d\mathbf{q}}{ds_i} \right) \\ & \left. - \sum_m \omega_m \left(\phi_m \frac{d\mathbf{C}_T}{ds_i} \boldsymbol{\theta}_0 e^{-\frac{t}{\tau_m}} + \int_0^t \phi_m^T \frac{d\mathbf{q}}{ds_i} e^{-\frac{t-\gamma}{\tau_m}} d\gamma \right) \right\} dt \\ & - \sum_m \left(\lambda_m^T \left(\frac{d\mathbf{K}_T}{ds_i} - \frac{1}{\tau_m} \frac{d\mathbf{C}_T}{ds_i} \right) \phi_m \right) \\ & - \sum_m \left(\mu_m \frac{1}{2} \phi_m^T \frac{d\mathbf{C}_T}{ds_i} \phi_m \right). \end{aligned} \quad (38)$$

Inspecting the resulting sensitivity, Equation 38, and the adjoint equations given in Equations 35 to 37, we observe that all terms involve forward time integrations. Most terms can be solved as soon as the response of a time instance is available. As can be seen from Equation 37, the adjoint variables λ_m and μ_m require integrations for the entire time domain and, thus, the adjoint equations in Equation 36 can be solved once a full time integration has been conducted. Hence, these equations do not call for a backward time integration or for an additional transient analysis.

4.3 | Effective time integration strategy

In the previous section, it has been shown that augmentation including the convolutions enabled to eliminate the backward transient integration commonly needed for an adjoint sensitivity analysis. The computation of the transient responses and their sensitivities need only one forward time integration, as mentioned before. The sequence of an efficient time integration is given as follows:

1. Compute and select a thermal modal basis for the response;
2. Use forward time integration, with at each time step calculation of
 - (a) general response functions,
 - (b) instantaneous adjoints: $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\omega}$,
 - (c) contributions to the pseudo-load vector to $[\mathbf{a}_m, b_m]^T$, and
 - (d) contributions to $\frac{d\hat{r}^*}{ds}$;
3. Compute adjoints: λ_m and μ_m (Equation 36);
4. Determine total sensitivity of general response functions ($\frac{d\hat{r}^*}{ds} = \frac{d\hat{r}}{ds}$).

Note that the obtained decoupling into independent computations per modes for the most computational expensive part (Equations 36 and 37) makes the proposed method well suited for parallel computing.

4.4 | Sensitivities for density-based topology optimization

The sensitivity analysis described above is generic. To compute the sensitivities needed for the density-based topology optimization as described in Section 3.2, some additional derivatives need to be taken into account. First, the sensitivities of the system matrices Equation 19 to the element densities (ie, the design variables s_i) are given by

$$\begin{aligned} \frac{d\mathbf{C}_T}{ds_i} &= \frac{1}{c_s} \frac{dc_i}{ds_i} [\mathbf{C}_T]_i, & \frac{d\mathbf{K}_T}{ds_i} &= \frac{1}{k_s} \frac{dk_i}{ds_i} [\mathbf{K}_T]_i, \\ \frac{d\mathbf{K}_M}{ds_i} &= \frac{1}{E_s} \frac{dk_i}{ds_i} [\mathbf{K}_M]_i, & \frac{d\mathbf{A}}{ds_i} &= \frac{1}{\beta_s} \frac{d\beta_i}{ds_i} [\mathbf{A}]_i, \end{aligned} \quad (39)$$

with the derivative of the interpolated material properties to the element density as

$$\frac{dp_i}{ds_i} = \frac{1 + S_p}{(1 + S_p(1 - s_i))^2} (p_s - p_v). \quad (40)$$

Note that the derivative of the interpolation according to the RAMP is nonzero in contrast to the derivative of the popular interpolation via solid isotropic material with penalization.⁸

When a density filter is used (Equation 27), the sensitivities need to be multiplied according to the chain rule of differentiation. Then, the derivative of the filtered density to the original density is needed, which is given as

$$\frac{d\bar{\rho}_i}{ds_j} = w_{ij}, \quad (41)$$

where w_{ij} denotes the contribution of the original density s_j in the filtered density $\bar{\rho}_i$ as defined in Equation 28.

For the optimization with penalization of intermediate densities, the derivative of the penalty term j_{gp} (Equation 31) is added to the sensitivity of the objective:

$$\frac{dj_{gp}}{ds_i} = \frac{C_{gp}}{N_s} (1 - 2s_i). \quad (42)$$

We like to mention that the sensitivity of the black/white projection (Equation 29) is not needed, because the black/white projection is not an optimization step but an initial design step before the second optimization.

4.5 | Method extensions

The sensitivity analysis above eliminates the reverse transient for the adjoint problem as a result of the modal representation. Thus, any problem that can be represented reasonably accurate by its modes can benefit from the proposed sensitivity approach. Although our study focuses on a thermo-mechanical system, the approach can easily be extended to, for example, electric or structural dynamic systems. Then, the augmentation in Equation 34 should include the governing equations for the reduced electric or structural dynamics problem and following the steps as described above, one will arrive at a sensitivity equations similar to Equation 38.

As was mentioned earlier, the sensitivity analysis is generic and can be applied in different optimization methods. For example, instead of the density-based topology optimization, as used in this study, the analysis can be applied in combination with evolutionary or level-set approaches for topology optimization. Then, the additions discussed in Section 4.4 need to be tailored to the appropriate topology optimization method.

Furthermore, in engineering practice, topology optimization is used to design a component while the performance is dependent of the entire system. Then, dynamic substructuring methods can provide an efficient computation of the responses.⁶¹ The presented sensitivity approach can be very well applied in combination with dynamic substructuring, especially the component mode synthesis approaches.

5 | EFFECT OF MR ON SENSITIVITIES

In the previous section, a method has been proposed which rigorously eliminates the backward transient integration for the adjoint sensitivity analysis. At the basis is a MOR technique, which relies on a reduced thermal modal basis combined with static correction. To study the effect of the model reduction, we focus on the sensitivities of the thermal system only and apply the method to a simple two-dimensional thermal transient problem.

First, the example is described. Next, the MR and selection methods presented in Section 2.2 are compared. In this section, both MR and MR with static correction (RC) are applied in combination with modal truncation (-T) or relevance-based modal basis (-R). Hence, the abbreviation “RC-T” refers to the MR in combination with modal truncation.

5.1 | Simple thermal problem

The objective is to design a cooling layout such that the temperature at the area of interest is minimized. The system is a thick slab of aluminum with a thickness of 10 mm as shown in Figure 2. The temperature at the right boundary is constant

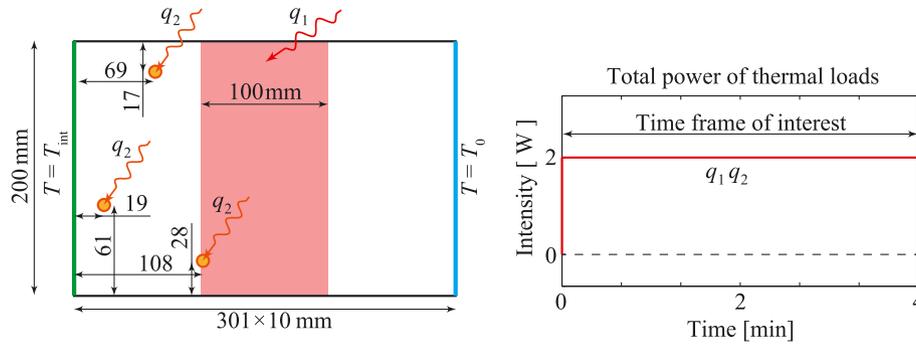


FIGURE 2 Left, the system, its boundaries and the thermal loads are shown, and on the right, the intensity of the thermal loads as function of time

and equal to the environmental temperature, and the other boundaries are thermally isolated. The area of interest is the left boundary (green boundary in Figure 2), and, thus, the aim is to minimize the temperature at the left boundary of the plate. The objective function is defined as

$$j[\mathbf{s}] = \frac{1}{t_E} \int_0^{t_E=240} J[\mathbf{s}, \boldsymbol{\theta}, t] dt \quad (43)$$

with

$$J[\mathbf{s}, \boldsymbol{\theta}, t] = \boldsymbol{\theta}^T \mathbf{p} \mathbf{p}^T \boldsymbol{\theta}, \quad (44)$$

where $\boldsymbol{\theta}$ indicates the nodal temperature vector, \mathbf{p} denotes a boolean vector to select the temperature nodes of the area of interest, and t_E is the end time of the time interval of interest ($t_E = 240$ s). Note that the function J represents the time average temperature difference squared. In addition, the weighting function w as defined in Equation 18 is chosen to be constant and equal to $\frac{1}{t_E}$; hence, the objective j represents the temporal average of the function J .

The design problem is to distribute the cooling such that the objective is minimized, whereas the total amount of cooling power is constrained. The cooling is modeled as linear convection with a constant coolant temperature equal to the environmental temperature. The design variables are the heat transfer coefficients, modeled with one variable per every element. The initial design is chosen to be a uniform cooling distribution. In practice, the cooling layout can be considered to represent the distribution of cooling channels in the system.

The dynamic response optimization problem for the described simple thermal problem is given by

$$\begin{aligned} \min_{\mathbf{s}} \quad & \frac{1}{t_E} \int_0^{t_E} (\boldsymbol{\theta}^T \mathbf{p} \mathbf{p}^T \boldsymbol{\theta}) dt, \\ \text{s.t.} \quad & \mathbf{C}_T \dot{\boldsymbol{\theta}} + (\mathbf{K}_T + \mathbf{H}) \boldsymbol{\theta} = \mathbf{q}_k \quad \text{for } t \in [0, t_E], \\ & \frac{1}{h_{\text{lim}}} \sum_{e=1}^{N_{el}} s_e - 1 \leq 0, \\ & 0 \leq s_i \leq h_{\text{lim}} \quad \text{for } i = 1, \dots, N_s, \end{aligned} \quad (45)$$

where \mathbf{q}_k indicates the nodal thermal load vector for case k and \mathbf{H} denotes the heat convection matrix that depends on the element heat transfer coefficients, ie, the design variables s_i . The total amount of cooling power is limited by h_{lim} .

Two optimization cases are considered, each with a different distributed thermal load: Case 1, a thermal load covering a large part of the system (\mathbf{q}_1), and for case 2, 3 localized thermal loads (\mathbf{q}_2) are applied, as shown on the left-hand side of Figure 2. Both thermal loads are initiated at the initial time and have a total intensity of 2 W, as shown on the right-hand side of Figure 2. The temperature response of both load cases is shown in Figure 3, where the initial temperature distribution is chosen to be uniform and equal to the environmental temperature. The objective j of the considered problem is a scalar function as given by Equation 43. In the present paper, the design is not optimized as this example is exclusively used to study the sensitivities and the effect of MR on the sensitivities. The design sensitivities of the objective, Equation 43, are computed for the full model with respect to all design variables s_i (with $i \in [1, N_s]$, where N_s denotes the

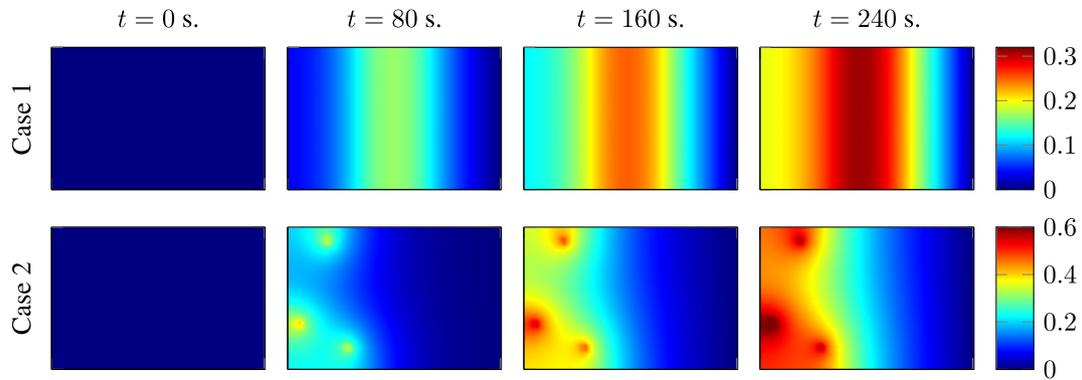


FIGURE 3 The temperature response (144×84 elements) at 4 time instants for both load cases and a uniform cooling layout. Note that the color scaling, representing the temperature distribution, is different for both cases as indicated by the color bars at the right side

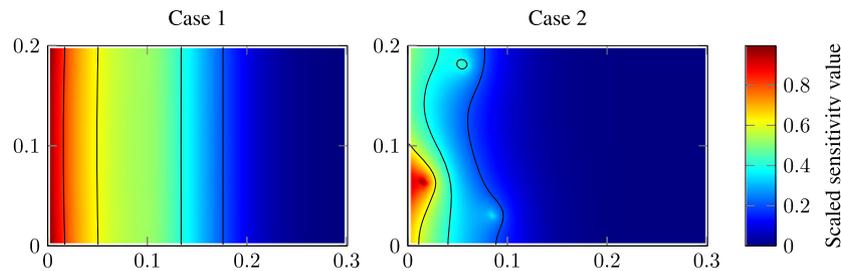


FIGURE 4 Sensitivity information of the objective for all design variables is shown for the full model. The sensitivity value represents a decrease of the objective. The sensitivity values are scaled with their maximum to show the distribution

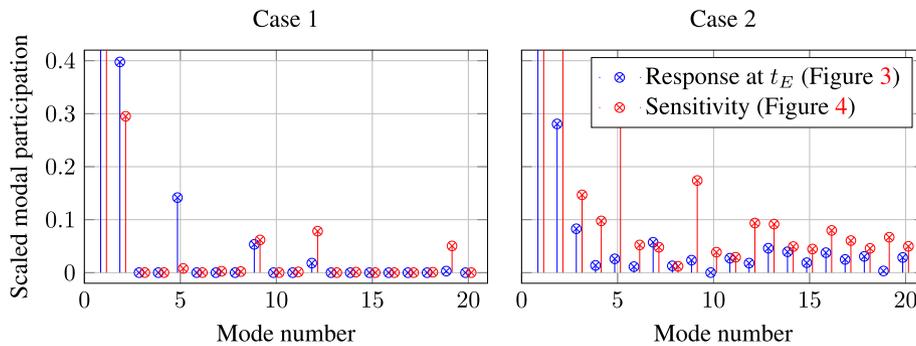


FIGURE 5 In all cases, the first mode is dominate followed by the second. This figure shows, for both cases, the participation of higher modes with respect to the first mode in the response (Figure 3) and sensitivity field (Figure 4)

number of design variables, in this case equal to the number of elements N_{el}) and shown in Figure 4. In the next section, the effect of the proposed sensitivity analysis on the sensitivity field is discussed.

5.2 | Sensitivities by MR

The results shown in Figures 3 and 4 are obtained from the full system without MR ($N = 12325$) using direct time integration ($N_t = 1001$). To reduce the computational and storage requirements, MR is applied as described in Section 2.2. The modal participations to the response, ie, modal amplitudes η , at end time t_E are shown in Figure 5. Only a few modes are relevant for the response of case 1, whereas the response of case 2 requires a combination of more modes. Nevertheless, a high-quality approximation for case 2 with a relative low number of mode is feasible when the modal acceleration method is applied. Furthermore, the modal participation to the sensitivity, as shown in Figure 4, is hardly related to the

modal participation of the response as seen in Figure 5. This is in line with, for example, Dutta and Ramakrishnan³³ in which they concluded that the reduction of the sensitivity of adjoint variable (as shown as strategies 2 and 3 in Figure 1) needs a different selection of modes for a proper approximation.

We use the proposed approach, strategy 4, on the system described above, where the response is reduced using MR and the sensitivities are determined using the adjoint variable method. To compare the results of the reduced model with the results of the full model (ie, the full system using direct time integration), an error measure is defined. Analog to the definitions in Dutta and Ramakrishnan,⁶² the error is defined as

$$\epsilon_{\theta} = \frac{\sqrt{\frac{1}{N_t} \sum_n (\|\hat{\theta}_n - \theta_n\|)^2}}{\sqrt{\frac{1}{N_t} \sum_n (\|\theta_n\|)^2}}, \quad (46)$$

where $\hat{\theta}_n$ and θ_n denote the temperature vector at t_n computed with the reduced model and the full model, respectively, and $\|\mathbf{v}\|$ denotes the Euclidean norm of the vector \mathbf{v} . Similarly, the error of the sensitivity field is expressed as

$$\epsilon_S = \frac{\sqrt{\frac{1}{N_e} \sum_i \left(\frac{d\hat{r}}{ds_i} - \frac{dr}{ds_i} \right)^2}}{\sqrt{\frac{1}{N_e} \sum_i \left(\frac{dr}{ds_i} \right)^2}}. \quad (47)$$

Figures 6 and 7 show these errors depending on the number of modes for cases 1 and 2, respectively. All figures include 4 graphs related to the MR method and mode selection method. The reduction is either by MR, Equation 12, or with static

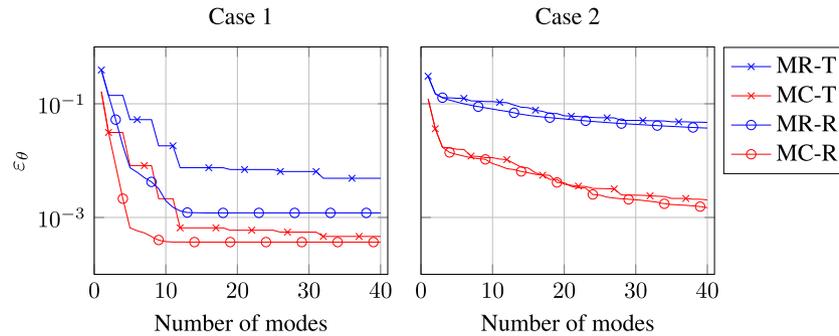


FIGURE 6 The error in the transient response, as defined in Equation 46, as a function of the number of modes. The error decreases with the number of modes, where the selection method based on relevance is only effective compared to truncation when a small number of modes participate, while a significant decrease is seen applying MC instead of MR

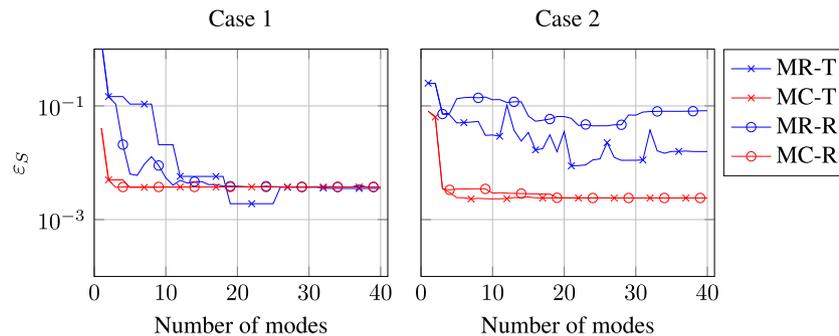


FIGURE 7 The error in the sensitivity, as defined in Equation 47, as a function of the number of modes. The main point is that a significant decrease is seen applying MC instead of MR, specifically for case 2. Figure 8 shows the sensitivity difference distributions to explain the behavior seen for case 2

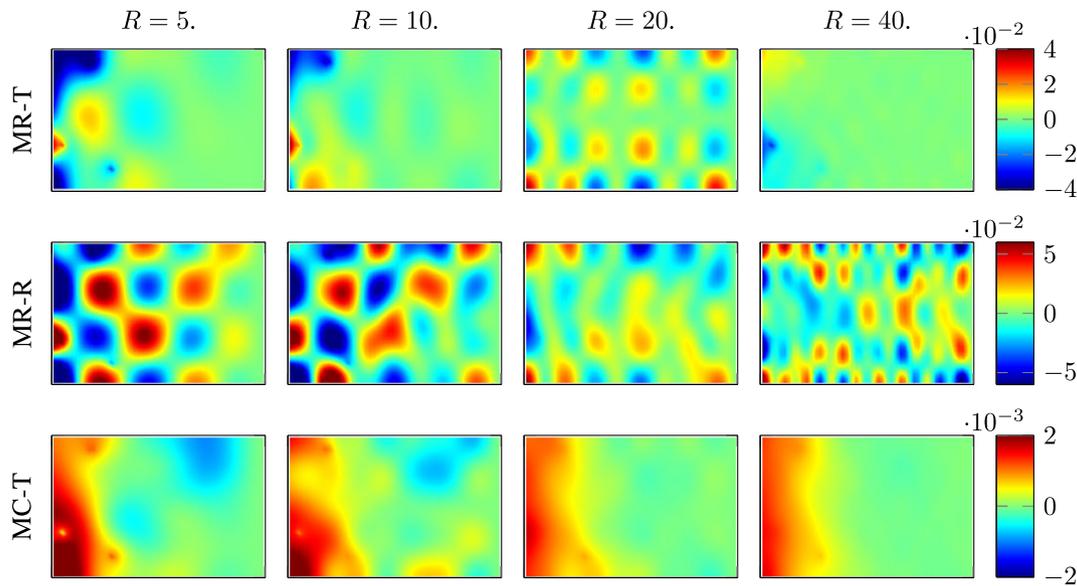


FIGURE 8 The scaled difference between the sensitivities (ΔS_i) for different modal reduction techniques and selection criteria. First, the top and middle rows are computed by the modal reduction (MR) and the bottom row is with correction (MC). Note that the color scaling is different for each row as shown by the color bars at the right side. Second, different number of modes are selected (R), which range from 5 to 40. The modes used for the top row are truncated, whereas the middle and bottom rows select the modes out of the first 100 based on their relevance. This figure shows that the large error when using MR results from high-frequency content. When applying the MC, the high-frequency content is captured and the error is significantly decreased

correction (MC), Equation 15. For the modal truncation (-T), the number of modes indicates the cutoff criteria, whereas, for the relevance-based (-R), the number of modes indicates the number relevant modes selected out of the first 100 modes.

From Figures 6 and 7, several interesting results are observed. First, in Figure 6 and more clearly in Figure 7, the error decreases for both the response and sensitivity when applying the MC instead of the MR. However, the accuracy increase is associated with an increase of computational effort to solve for the static correction and its sensitivity. Second, in none of the graphs, the error converges to zero but converges to a difference due to other errors: the truncation error in numerical time integration present in the full model and the numerical error in eigenmodes and eigenvalues computation. To decrease this difference, the number of time steps needs to be increased. Next, while the error in the responses shows a smooth convergence, the error of the sensitivity does not, especially when applying MR. To investigate this behavior, the difference of the sensitivity fields is studied, which is expressed as

$$\Delta S_i = \frac{\frac{d\hat{r}}{ds_i} - \frac{dr}{ds_i}}{\frac{dr}{ds_i}}. \tag{48}$$

Figure 8 shows that the large error when using MR results from the discarded high-frequency content. When applying the MC, this high-frequency spatial effects are captured and the error is significantly decreased. Furthermore, only slight differences are seen between the 2 selection criteria: modal truncation and relevance-based.

In conclusion, from our experience and the examples shown above, the number of modes needed for a high-quality approximation of the response is sufficient for the sensitivity. However, in line with Dutta and Ramakrishnan,³³ the use of MC is essential to avoid high-frequency spatial effects, which otherwise dominate the sensitivity information. Although a uniform distribution of design variables was used for this demonstration, similar results were obtained for nonuniform distributions.

6 | TOPOLOGY OPTIMIZATION FOR A THERMO-MECHANICAL TRANSIENT PROBLEM

The accuracy of the sensitivities have been studied in Section 5. In this section, the proposed sensitivity method is applied to a topology optimization problem. We focus on both the effect of the MOR on the topological sensitivity and obtained

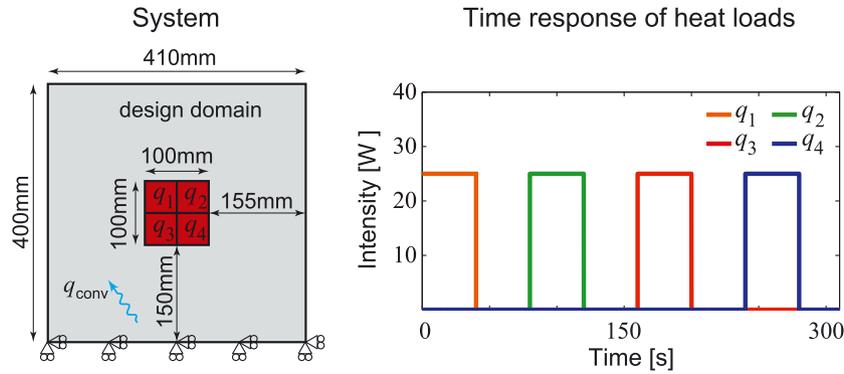


FIGURE 9 Left, the thermo-mechanical system, its boundaries and the thermal loads are shown. The boundaries are thermally isolated and heat is extracted via convective cooling over the plate. The red squares indicate the 4 quadrants subjected to a transient heat load. Their intensity over time is given on the right. The optimization aims to design a structure in the design domain (gray area, excluding the red squares), which minimizes the displacement of the center points in the red squares

designs. Furthermore, several important aspects will be studied: the effect of the mesh size, the effect of intermediate densities, and the effect of a rapidly fluctuating heat load. In this study, both MR and MR with static correction (RC) are applied in combination with modal truncation (-T) only.

6.1 | Topology optimization problem

The thermo-mechanical design problem involves an aluminum slab with a thickness of 10 mm as shown on the left in Figure 9. A thermal load is applied sequentially to 4 quadrants q_1 , q_2 , q_3 , and q_4 , as indicated by the red squares. A 25-watt heat load is applied to each quadrant for 40 seconds sequentially, with a 40-second break between each heat load. The transient responses of these thermal loads are shown on the right of Figure 9. This heat load may represent a typical measuring process within a precision tool.

All the system boundaries are thermally insulated, and the heat is extracted from the slab by cooling, which is modeled as linear heat convection with a heat transfer coefficient of $100 \text{ W}/(\text{m}^2\text{K})$. The system is mechanically supported in x, y -direction at one side of the slab, as shown in Figure 9. The initial temperature distribution is chosen to be uniform and equal to the environmental temperature, and there are no initial deformations present. As a result of the temperature fluctuations, thermal expansion will cause the areas of interest, the 4 quadrants, to displace. This may result in, for example, a misalignment during the measuring process. In this case, the so-called thermal error is defined as the absolute displacement of a quadrant's center point during the time of its loading. Hence, the position of point of interest is time dependent.

The objective of the topology optimization is to find a material layout that minimizes the thermal error of the quadrants during their individual loading. The material layout is represented by the element densities as described in Section 3.2. The area comprising the quadrants ($100 \times 100 \text{ mm}$ in the center of the slab) needs to be solid. Thus, the design variables are the element densities of all elements covering the slab apart from the quadrants (ie, the design domain is shown as the gray area in Figure 9). The initial design is a uniform material distribution over the design domain with $\rho_i = 0.5$.

Material interpolation via RAMP is applied to 5 parameters. The 4 material parameters as in Equation 19 are heat capacity, heat conduction, thermal expansion coefficient and Young modulus, and the heat transfer coefficient. In the present paper, the penalization parameters associated with the 5 interpolation functions, Equation 20, are all chosen to be equal to 1. Choosing different values for the penalization parameters has influence on the optimization result; however, this is not part of our present investigation.

The volume of the design is contained with Equation 24, where the volume fraction bound $v_{lim} = 0.5$. In addition, a density filter is implemented with a filter radius of $r_{min} = 25 \text{ mm}$ (Equation 27) to avoid mesh-dependent phenomena. The topology optimization problem for this transient thermo-mechanical system is given in Section 3.2.4.

6.2 | Performance and sensitivity of full design

Before optimizing, the original design is considered. The original design is the rectangular aluminum slab, the so-called full design (see the left of Figure 10), in which the entire design domain is filled with material ($\rho_i = 1, \forall i$). The response

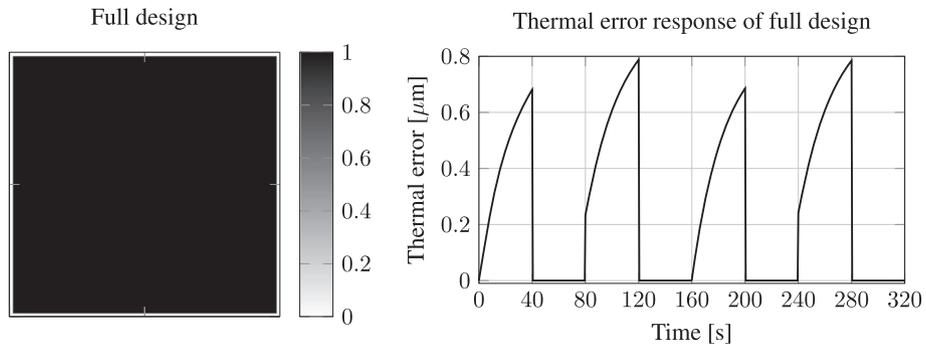


FIGURE 10 Left, the full design is given, in which the entire design domain is filled with solid material ($\rho_i = 1; \forall i$). This design is taken as a reference for the optimization. Right, the response of its thermal error, ie, the absolute displacement of the quadrant's center points, is shown. Thus, the line from 0 to 40 seconds represents the absolute displacement of the center point of the first quadrant, from 80 to 120 seconds for the second quadrant, from 160 to 200 seconds for the third quadrant, and from 240 to 280 seconds for the fourth quadrant. The objective of the topology optimization is to find a material layout, which minimizes this transient response

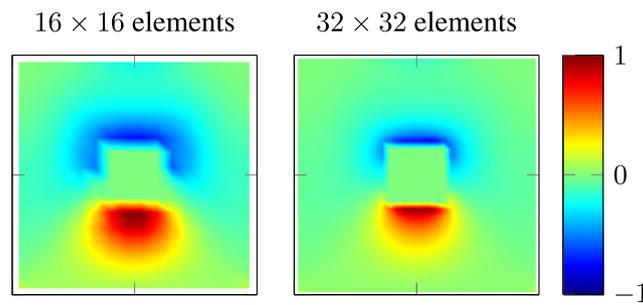


FIGURE 11 The scaled sensitivity of the objective to the design variables for two meshes computed for $\rho_i = 1$ without modal reduction. On the left, the scaled sensitivity for a mesh with 16×16 elements, and on the right, with 32×32 elements. A scaled sensitivity value > 0 indicates to remove material and a value < 0 indicates to add material

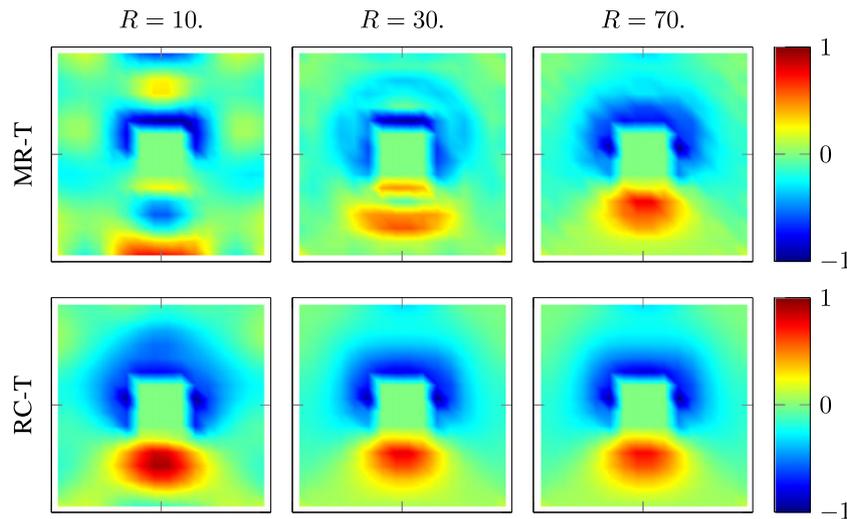


FIGURE 12 The scaled sensitivity based on modal reduction for a mesh of 16×16 elements. For both MR and RC, the sensitivity converges with the number of modes towards the sensitivity without reduction as shown on the left in Figure 11

of the full design's thermal error, ie, the absolute displacements of the quadrant's center points, is shown on the right of Figure 10. It is seen that the thermal error is nonzero only when a heat load is present. Furthermore, the absolute displacement with respect to the reference position does not have to be equal to zero as the thermal history is taken into account.

To effectively compute the response and sensitivities, the transient thermo-mechanical system is reduced by the proposed method. Hence, strategy 4, as given in Figure 1, is followed, MR is used, and the sensitivities are computed according to the time integration strategy given in Section 4.3. The sensitivities computed without reduction are shown in Figure 11 for 2 mesh sizes. Both sensitivities fields are approximated with different reduced models in Figures 12 and 13. The results are in line with the conclusions in Section 5: The sensitivity fields based on both MR methods convergence with the number of modes towards the field without reduction as shown in Figure 11. However, the MR results in large errors due to the omitted high-frequency content, which can be decreased significantly by applying RC. The latter becomes more prominent for models with larger number of thermal degrees of freedom, ie, smaller elements. Thus, the proposed sensitivity method with RC can be a effective method to compute topological sensitivities for transient problems.

6.3 | Topology optimized designs

The left of Figure 14 shows the optimized design after 200 iterations obtained with sensitivities computed without reduction. The design shows 2 main functional areas: The upper part functions for heat storage and dissipates heat via convective cooling, and the lower part functions as a compliant mechanism rotating the system such that the thermal error is minimized. The rotation is actuated by the thermal expansion resulting from the heat loads. The optimizer abuses

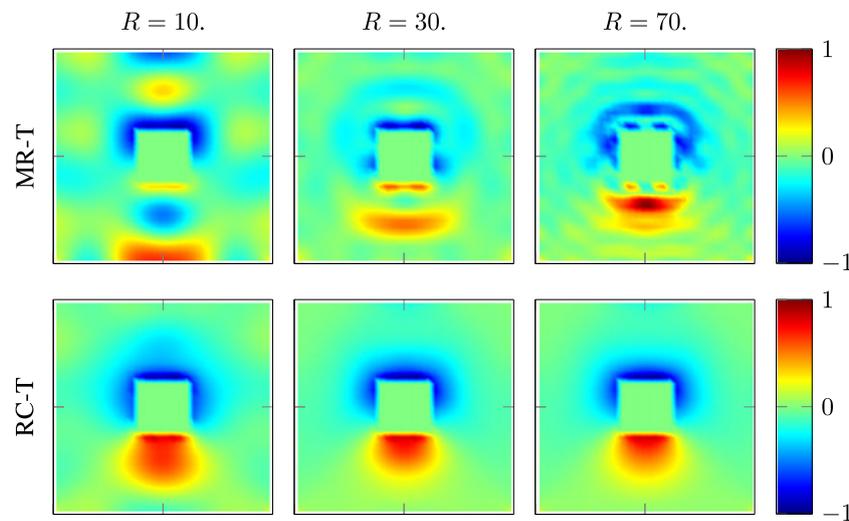


FIGURE 13 The scaled sensitivity based on modal reduction for a mesh of 32×32 elements. For both MR and RC, the sensitivity convergence with the number of modes towards the sensitivity without reduction as shown on the right in Figure 11. Similar to the results in Section 5, the sensitivity based on MR (top row) shows high-frequency spatial errors, which are avoided when using RC (bottom row)

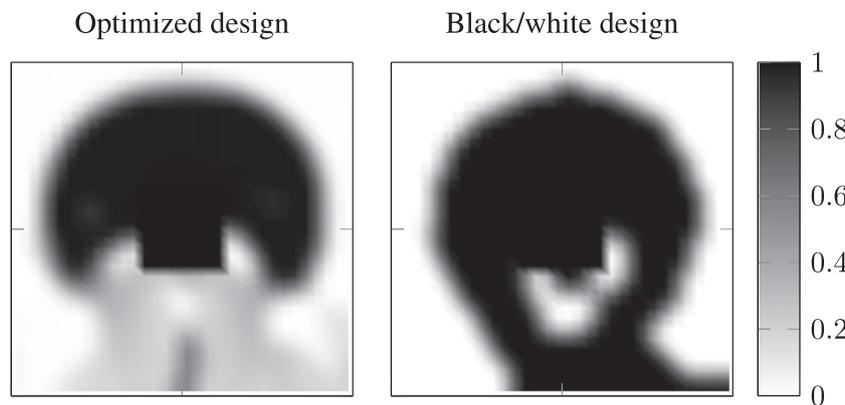


FIGURE 14 Left, the optimized design (after 200 iterations) obtained with sensitivities computed without modal reduction on a mesh with 32×32 elements. Right, the black/white design obtained after penalizing intermediate densities of the optimized design on the left following the strategy discussed in Section 3.2.6. The performance of these 2 designs compared to the full design is shown in Figure 15

the intermediate density material to obtain the best stiffness for the compliant mechanism. Because the interpretation of the design areas with intermediate densities is not evident, an additional optimization is performed in which the intermediate densities are penalized as discussed in Section 3.2.6. The resulting black/white design is shown at the right of Figure 14 in which the same 2 functional areas are seen.

The performance of the optimized and black/white design is compared to the performance of the full design. A major decrease of the thermal error is seen in Figure 15; hence, a significant performance increase is obtained after optimization. The design example demonstrates that smart material layouts can reduce the thermal error in thermo-mechanical problems. Hence, topology optimization is a good option for thermal error reduction in precision devices.

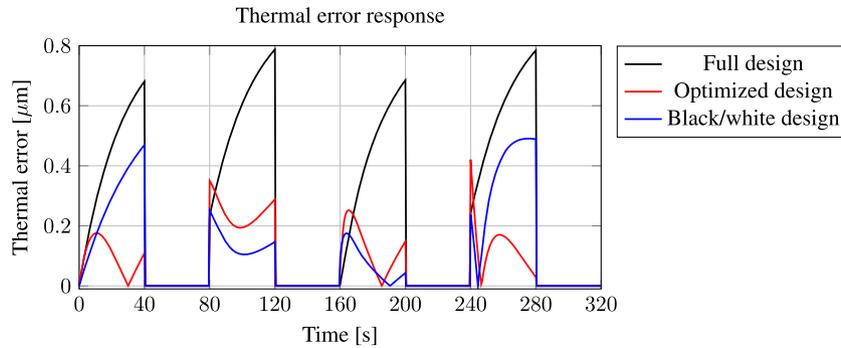


FIGURE 15 The thermal error response in time for both the optimized and black/white design compared to the full design from Figure 10. The objective shows a significant decrease after optimization; the displacement of all 4 quadrants have been decreased during their time frames of interest. Furthermore, the thermal error response of the black/white design is slightly larger than the response of the optimized design

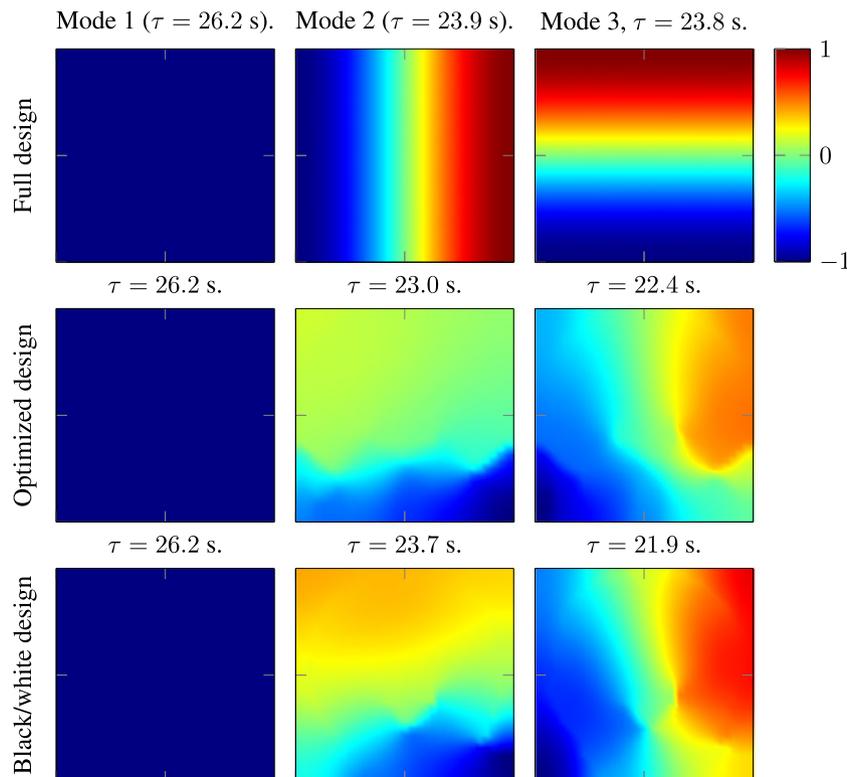


FIGURE 16 The first 3 thermal mode shapes of the 3 designs: full design (Figure 10), optimized design, and black/white design (Figure 14). The amplitude of all mode shapes is scaled such that the absolute maximum is equal to one. The first mode shape of all 3 designs is a uniform temperature distribution, which transfers heat only by heat convection. The second and third mode shapes are temperature distributions with one main heat transfer from one to another side of the system

The sensitivities based on reduced models as proposed in the present paper are also used for topology optimization. To give an idea of the thermal modes shapes, these temperature distributions for the first 3 shapes are given for the designs in Figure 16. Figure 17 show the optimized designs after 200 iterations. Similar as seen for the sensitivities in Figure 13, the designs converge with the number of modes to the design based on models without reduction (Figure 14). The spatial fluctuating errors in the sensitivities based on MR result in gray areas and material islands in the designs. These effects do not necessary have a bad influence on the performance; however, it is not preferred for interpretation reasons.

Figure 18 shows the designs after the second optimization with penalization of intermediate densities. The high-frequency spatial errors in the sensitivities lead to improper black/white designs with poor performance. Only the reduced models with static correction and sufficient modes give legitimate designs.

6.4 | Effect of intermediate density material

To study the effect of intermediate density material on the sensitivity analysis with MR, the sensitivity of the optimized design (see Figure 14) is considered. This design consists of a large area of intermediate and low density elements and is

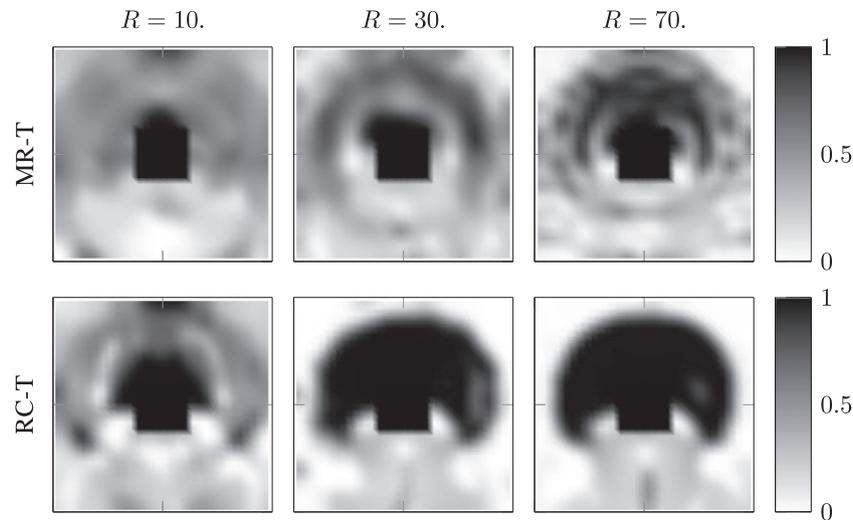


FIGURE 17 Optimized designs (after 200 iterations) obtained with modal reduction on a mesh of 32×32 elements. The errors in the sensitivities introduce designs with a lot of gray material reflecting the spatial errors in the sensitivities (see Figure 13). The design obtained with sensitivities based on RC including 70 modes is very similar to the optimized design shown in Figure 14

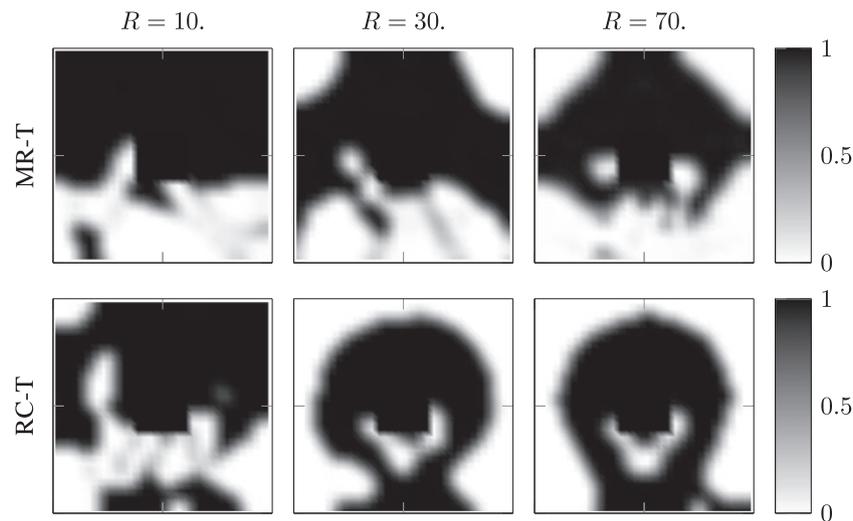


FIGURE 18 Black/white designs obtained following the strategy as described in Section 3.2.6 in which an additional optimization is performed while the intermediate densities are penalized. The errors in the sensitivities result in improper designs. Only the design obtained with sensitivities based on modal reduction with static correction and sufficient modes gives similar designs to the design shown in Figure 14

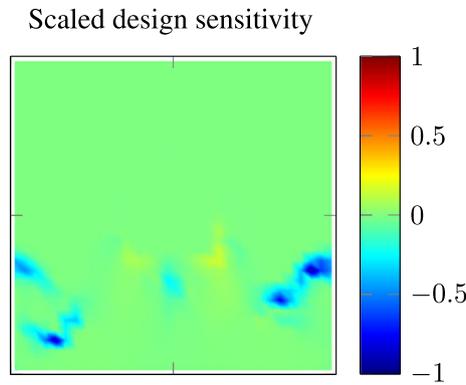


FIGURE 19 The scaled sensitivity of the objective to the design variables computed without reduction for the optimal design shown left in Figure 14. A sensitivity value > 0 indicates to remove material, and a value < 0 indicates to add material

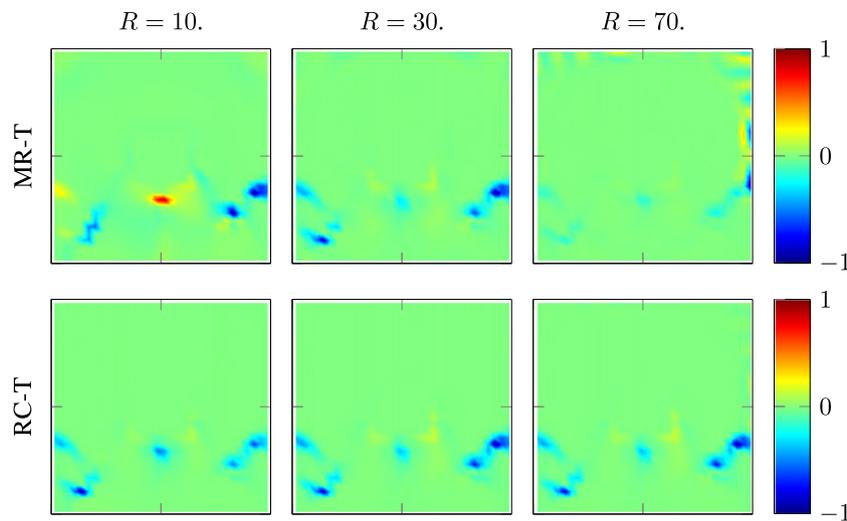


FIGURE 20 The scaled sensitivity based on modal reduction for the optimized design as shown in Figure 14 to study the effect of intermediate density material. Similar as for the full design, the MR shows high-frequency spatial errors, but the RC obtains the sensitivity very accurately

therefore expected to enable localized modes as shown in previous studies.^{38,45} However, these spurious modes are less pronounced for thermal systems and are not obtained for these designs as seen in Figure 16.

The sensitivity computed without MR is presented in Figure 19, which shows localized gradient information. Figure 20 shows the sensitivities for the optimized design based on reduced models. Although the design consists of a low density material and the essential sensitivity information is very localized, the RC is able to capture the sensitivity field very accurately, whereas the MR shows high spatial frequency errors.

6.5 | Effect of a rapidly fluctuating heat load

In this section, we like to point out a limitation of the proposed method using the RC. The RC takes into account the entire basis, but ignores the dynamic behavior of all the modes omitted, as mentioned in Section 2.2. Consequently, the RC includes valuable information that was truncated in MR. However, when the time constants of the modes are in the same order as the relevant heat load fluctuations, then the static approximation of the omitted modes is not sufficient and the RC is not able to compute the sensitivities accurately.

The described limitation is presented by using the same problem as before, but the heat loads and the time interval of interest are much shorter, as shown in Figure 21. The time interval is 32 instead of 320 seconds, and each heat load acts for 4 instead 40 seconds. The sensitivity computed without MR is presented in Figure 22 and the sensitivities based on MR in Figure 23. In this figure, it is seen that neither the MR nor the RC can capture the essential sensitivity information and

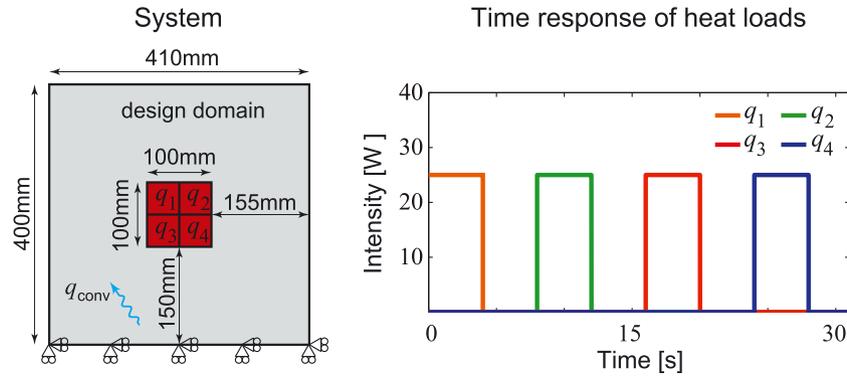


FIGURE 21 Left, the thermo-mechanical system as presented in Figure 9, and on the right, the intensity of the thermal load over time. The transient behavior of the heat loads differs from the previous example as the heat loads are applied for 4 instead of 40 seconds, and the end time t_E is 32 instead of 320 seconds

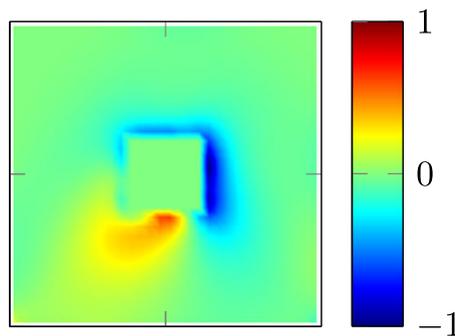


FIGURE 22 The scaled sensitivity of the objective to the design variables for the problem shown in Figure 21 (32×32 elements) computed without modal reduction. A sensitivity value > 0 indicates to remove material, and a value < 0 indicates to add material

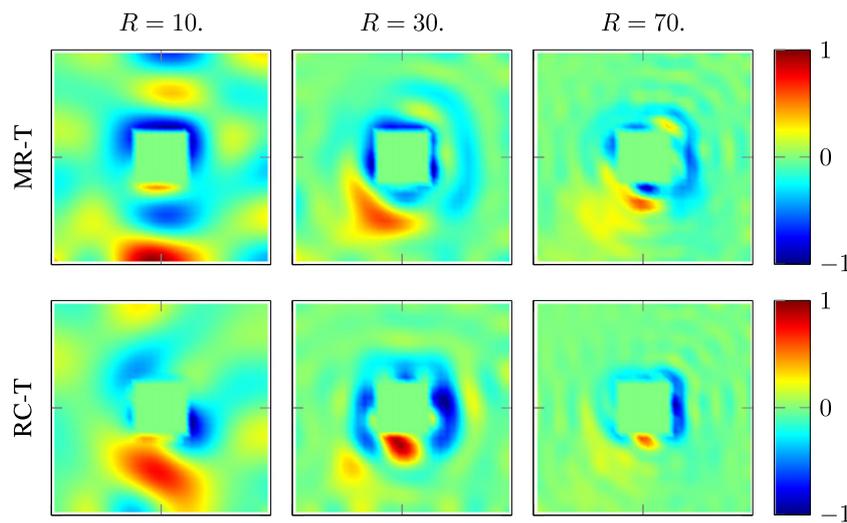


FIGURE 23 The scaled sensitivity based on modal reduction for the problem shown in Figure 21. For both MR and RC, the sensitivity convergence with the number of modes towards the sensitivity without reduction. However, both methods show high-frequency spatial errors

show high spatial fluctuating errors. The RC is not able to give an accurate result as the time constants are in the same order as the heat load fluctuations (the time constants of the 10th, 30th, and 70th modes are 13.9, 6.7, and 3.2 seconds, respectively).

This example clearly indicates that RC is capable to provide an excellent correction, unless the time constants related to the modes RC corrects for are still relevant for the overall response of the system. In case the RC correction is associated with modes that are not relevant for the overall response, then RC leads to a major improvement of the sensitivities.

7 | CONCLUSIONS

Computation of design sensitivities for *transient problems* is computationally expensive and often hinders the optimization of large-scale systems with many design variables. For many design variables and few response functions, the adjoint variable method as the preferred choice to compute sensitivities suffers from large storage requirements due to the reverse transient analysis and is not well suited for parallel computing. Common approaches to decrease the computational effort apply model-order reduction techniques. However, these methods need either an expensive sensitivity computation for the transformation basis nor an additional (new) transformation basis.

In this study, an approach is presented, which uses the adjoint variable method to a reduced model represented by modal superposition combined with static correction (ie, MAM), as shown as strategy 4, in Figure 1. It can be concluded that the method eliminates the need for a reverse transient analysis that is commonly needed for transient problems. In fact, all terms and contributions required for the adjoint sensitivity can be co-integrated. Augmentation with the static correction provides a major improvement for the accuracy of the resulting design sensitivities. In particular, rapid spatial fluctuations can be eliminated. However, including a static correction has computational disadvantages for both the response evaluation as well as the sensitivity analysis. Moreover, the method is only effective if its correction is associated with modes, which transient contribution is not relevant for the overall response.

All methods based on MR need a proper selection of the reduction basis to ensure a high-quality approximation. In Section 5, we recommend to take into account the same number of modes for the sensitivity as the reduction of the response as long as the static correction is included (ie, apply the MAM). To reduce the computational costs during optimization, one may opt to compute new eigenvectors only when the system changed significantly. Furthermore, the sensitivities determined following the proposed strategy are consistent with the reduced response.

In Section 6, the proposed method is used for topology optimization problem. The high-frequency spatial errors in the sensitivities without static correction lead to improper designs with poor performance. Sensitivities based on mode superposition with static correction and sufficient modes are able to result legitimate designs similar to the optimized designs without reduction. Furthermore, the examples demonstrate that topology optimization is a good option for thermal error reduction.

Finally, although our study concentrates on transient thermo-mechanical problems (ie, first-order linear systems), the approach can be translated nearly one-to-one to undamped mechanical system (ie, second-order linear systems).

ACKNOWLEDGEMENT

Pieken in de Delta, Provincie Noord-Brabant, and Samenwerkingsverband Regio Eindhoven are acknowledged for their support in this project.

ORCID

F. van Keulen  <http://orcid.org/0000-0003-2634-0110>

REFERENCES

1. Schellekens P, Rosielle N, Vermeulen H, Vermeulen M, Wetzels S, Pril W. Design for precision: current status and Trends. *CIRP Ann - Manuf Technol.* 1998;47(2):557-586.
2. Mayr J, Jedrzejewski J, Uhlmann E, et al. Thermal issues in machine tools. *CIRP Ann-Manuf Technol.* 2012;61(2):771-791.
3. Deaton JD, Grandhi RV. A survey of structural and multidisciplinary continuum topology optimization: post 2000. *Struct Multi Optim.* 2014;49(1):1-38.
4. Rodrigues H, Fernandes P. A material based model for topology optimization of thermoelastic structures. *Int J Numer Methods Eng.* 1995;38(12):1951-1965.
5. Li Q, Steven GP, Xie YM. Thermoelastic topology optimization for problems with varying temperature fields. *J Therm Stresses.* 2001;24(4):347-366.

6. Shiguang Deng. Augmented topological level-set for large-scale thermo-elastic topology optimization. *PhD*. University of Wisconsin, Madison; 2016.
7. Sigmund O. Design of multiphysics actuators using topology optimization—part I: one-material structures. *Comput Methods Appl Mech Eng*. 2001;190(49):6577-6604.
8. Gao T, Zhang WH. Topology optimization involving thermo-elastic stress loads. *Struct Multi Optim*. 2010;42(5):725-738.
9. Deaton JD, Grandhi RV. Stiffening of restrained thermal structures via topology optimization. *Struct Multi Optim*. 2013;48(4):731-745.
10. Turteltaub S. Optimal material properties for transient problems. *Struct Multi Optim*. 2001;22(2):157-166.
11. Li Y, Saitou K, Kikuchi N. Topology optimization of thermally actuated compliant mechanisms considering time-transient effect. *Finite Elem Anal Des*. 2004;40(11):1317-1331.
12. Matzen R, Jensen JS, Sigmund O. Topology optimization for transient response of photonic crystal structures. *J Opt Soc Am B*. 2010;27(10):2040-2050.
13. Mello LAM, Salas RA, Silva ECN. On response time reduction of electrothermomechanical MEMS using topology optimization. *Comput Methods Appl Mech Eng*. 2012;247-248:93-102.
14. Park G-J. *Analytic Methods for Design Practice*. London: Springer Science+Business Media; 2007.
15. Pedersen CBW. Crashworthiness design of transient frame structures using topology optimization. *Comput Methods Appl Mech Eng*. 2004;193(6-8):653-678.
16. Jang HH, Lee HA, Lee JY, Park GJ. Dynamic response topology optimization in the time domain using equivalent static loads. *AIAA J*. 2012;50(1):226-234.
17. Zhuang C, Xiang Z. A global heat compliance measure based topology optimization for the transient heat conduction. *Numer Heat Transfer, Part B: Fundam*. 2014;65(5):445-471.
18. Zhuang C, Xiong Z. Temperature-constrained topology optimization of transient heat conduction problems. *Numer Heat Transfer, Part B: Fundam*. 2015;68(4):366-385.
19. Antoulas AC. An overview of approximation methods for large-scale dynamical systems. *Annu Rev Control*. 2005;29(2):181-190.
20. Salleras M, Bechtold T, Fonseca L, et al. Comparison of model order reduction methodologies for thermal problems. In: Proceedings of the 6th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems. IEEE: Berlin, Germany; 2005:60-65.
21. Besselink B, Tabak U, Lutowska A, et al. A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control. *J Sound Vib*. 2013;332(2013):4403-4422.
22. Ahmed T, Gad E, Yagoub MCE. An adjoint-based approach to computing time-domain sensitivity of multiport systems described by reduced-order models. *IEEE Trans Microwave Theory Tech*. 2005;53(11):3538-3547.
23. Weickum G, Eldred MS, Maute K. A multi-point reduced-order modeling approach of transient structural dynamics with application to robust design optimization. *Struct Multi Optim*. 2009;38(6):599-611.
24. Ilievski Z. Model order reduction and sensitivity analysis. *PhD*. Eindhoven University of Technology, Eindhoven, the Netherlands; 2010.
25. Fox RL, Kapoor MP. Structural optimization in the dynamics response regime: a computational approach. *AIAA J*. 1970;8(10):1798-1804.
26. Haftka RT, Kamat MP. *Elements of Structural Optimization*. The Hague, The Netherlands: Martinus Nijhoff; 1985.
27. Choi KK, Kim NH. *Structural Sensitivity Analysis and Optimization 1*. New York, US: Springer Science+Business Media; 2005.
28. Kang B-S, Park G-J, Arora JS. A review of optimization of structures subjected to transient loads. *Struct Multi Optim*. 2006;31(2):81-95.
29. Feng T-T, Arora JS, Haug EJ. Optimal structural design under dynamic loads. *Int J Numer Methods Eng*. 1977;11(1):39-52.
30. Greene WH, Haftka RT. Computational aspects of sensitivity calculations in transient structural analysis. *Comput Struct*. 1989;32(2):433-443.
31. Greene WH, Haftka RT. Computational aspects of sensitivity calculations in linear transient structural analysis. *Struct Optim*. 1991;3(3):176-201.
32. Paul AC, Dutta A, Ramakrishnan CV. Accurate computation of design sensitivities for dynamically loaded structures with displacement constraints. *AIAA J*. 1996;34(8):1670-1677.
33. Dutta A, Ramakrishnan CV. Accurate computation of design sensitivities for structures under transient dynamic loads with constraints on stresses. *Comput Struct*. 1998;66(4):463-472.
34. Hsieh CC, Arora JS. Design sensitivity analysis and optimization of dynamic response. *Comput Methods Appl Mech Eng*. 1984;43(2):195-219.
35. Wang S, Choi KK. Continuum design sensitivity of transient responses using Ritz and mode acceleration methods. *AIAA J*. 1992;30(4):1099-1109.
36. Ma Z-D, Kikuchi N, Hagiwara I. Structural topology and shape optimization for a frequency response problem. *Comput Mech*. 1993;13(3):157-174.
37. Yoon GH. Structural topology optimization for frequency response problem using model reduction schemes. *Comput Methods Appl Mech Eng*. 2010;199(25-28):1744-1763.
38. Olhoff N, Du J. On topological design optimization of structures against vibration and noise emission. *Computational Aspects of Structural Acoustics and Vibration, CISM International Centre for Mechanical Sciences*, Vol. 505. Vienna: Springer; 2009:217-276.
39. Liu H, Zhang W, Gao T. A comparative study of dynamic analysis methods for structural topology optimization under harmonic force excitations. *Struct Multi Optim*. 2015;51(6):1321-1333.

40. Zhao J, Wang C. Dynamic response topology optimization in the time domain using model reduction method. *Struct Multi Optim*. 2016;53(1):101-114.
41. Lee TH. An adjoint variable method for structural design sensitivity analysis of a distinct eigenvalue problem. *KSME Int J*. 1999;13(6):470-476.
42. Lee TH. Adjoint method for design sensitivity analysis of multiple Eigenvalues and associated eigenvectors. *AIAA J*. 2007;45(8):1998-2004.
43. Yu Y, Jang IG, Kim IK, Kwak BM. Nodal line optimization and its application to violin top plate design. *J Sound Vib*. 2010;329(22):4785-4796.
44. Yu Y, Kwak BM. Design sensitivity analysis of acoustical damping and its application to design of musical bells. *Struct Multi Optim*. 2011;44(3):421-430.
45. Tcherniak D. Topology optimization of resonating structures using SIMP method. *Int J Numer Methods Eng*. 2002;54(11):1605-1622.
46. Takezawa A, Kitamura M. Sensitivity analysis and optimization of vibration modes in continuum systems. *J Sound Vib*. 2013;332(6):1553-1566.
47. Boley BA, Weiner JH. *Theory of Thermal Stresses*. London: John Wiley & Sons, Inc.; 1960.
48. Hetnarski RB, Eslami MR. *Thermal Stresses—Advanced Theory and Applications*, Solid Mechanics and Its Applications, vol. 158. Dordrecht, The Netherlands: Springer Science+Business Media; 2009.
49. Zienkiewicz OC, Taylor RL. *The Finite Element Method. Vol 1: Basic Formulations and Linear Problems*. 4th ed., Vol. 1. England: McGraw-Hill Book Company Europe; 1989.
50. Argyris JH, Vaz LE, William KJ. Higher order methods for transient diffusion analysis. *Comput Methods Appl Mech Eng*. 1977;12(2):243-278.
51. Maddox NR. On the number of modes necessary for accurate response and resulting forces in dynamic analysis. *Trans ASME - J Appl Mech*. 1975;42(Sur E):516-517.
52. Joo K-J, Wilson EL, Leger P. Ritz vectors and generation criteria for mode superposition analysis. *Earthquake Eng Struct Dyn*. 1989;18(2):149-167.
53. Hooijkamp EC, Eijk Jv, Keulen Fv. Thermo-mechanical design and optimization using transient modal analysis. In: Proceedings of the 10th World Congress on Structural and Multidisciplinary Optimization; 2013; Orlando, Florida, US:1-9.
54. Zhu J, Ni J, Shih AJ. Robust machine tool thermal error modeling through thermal mode concept. *J Manuf Sci Eng*. 2008;130 (6): 061006-1-9.
55. Koevoets AH, Eggink HJ, van der Sanden J, Dekkers J, Ruijl TAM. Optimal sensor configuring techniques for the compensation of thermo-elastic deformations in high-precision systems. In: Therminic 2007. EDA Publishing Association: Budapest, Hungary; 2007:17-19.
56. Sigmund O, Maute K. Topology optimization approaches—a comparative review. *Struct Multi Optim*. 2013;48(6):1031-1055.
57. Bendsoe MP. Optimal shape design as a material distribution problem. *Struct Optim*. 1989;1(4):193-202.
58. Sigmund O. Morphology-based black and white filters for topology optimization. *Struct Multi Optim*. 2007;33(4-5):401-424.
59. Sigmund O, Petersson J. Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct Optim*. 1998;16(1):68-75.
60. Svanberg K. The method of moving asymptotes—a new method for structural optimization. *Int J Numer Methods Eng*. 1987;24:359-373.
61. Klerk Dd, Rixen DJ, Voormeeren SN. General framework for dynamic substructuring: history, review, and classification of techniques. *AIAA J*. 2008;46(5):1169-1181.
62. Dutta A, Ramakrishnan CV. Error estimation in finite element transient dynamic analysis using modal superposition. *Eng Comput*. 1997;14(1):135-158.

How to cite this article: Hooijkamp EC, van Keulen F. Topology optimization for linear thermo-mechanical transient problems: Modal reduction and adjoint sensitivities, *Int J Numer Meth Engng*. 2018;113:1230–1257. <https://doi.org/10.1002/nme.5635>