Department of Precision and Microsystems Engineering

Topology Optimization of Opto-Mechanical Structures Leveraging Differential Ray Tracing Based STOP Analysis

R.C. Looman

Report no : 2025.040
Coach : Sanne van den Boom
Professor : Nandini Bhattacharya
Specialisation : Optomechatronics
Type of report : Master Thesis
Date : 14 August, 2025

Topology Optimization of Opto-mechanical Structures Leveraging Differential Ray Tracing Based STOP Analysis

Robin C. Looman[1,2], Sanne van den Boom[1,2], Nandini Bhattacharya[2]

[1]Department of Optomechatronics, Netherlands Organisation for Applied Scientific Research (TNO), Stieltjesweg 1, Delft, 2628 CK, Netherlands.
[2]Faculty of Mechanical, Maritime & Materials Engineering(ME), Department of Precision & Microsystems Engineering (PME), Delft University of Technology (Tu Delft), Mekelweg 2, Delft, 2628 CD, Netherlands.

Contributing authors: r.c.looman@student.tudelft.nl; sanne.vandenboom@tno.nl; n.bhattacharya@tudelft.nl;

Abstract

Due to the trend towards minimization of optical space instruments, in combination with ever increasing performance requirements, their optical and mechanical design are becoming more and more intertwined. As a result, it is invaluable to have insights into how mechanical and thermomechanical disturbances influence the system's optical performance already in early stages of the design process. In this work, a differential ray tracer is implemented into a topology optimization framework, allowing for direct optical performance analysis of optical instruments under disturbance loads. The gradient information of the optical performance is provided by semi-automatic differentiation. A wide range of optics can be used including freeform optics, that have become frequently used in state-of-the-art instruments. The method has been verified and shows a strong agreement with ZEMAX. To demonstrate the advantages of the STOP-based computational design workflow, a fully coupled opto-thermo-mechanical topology optimization is performed on a case inspired by CHAPS-D, a hyperspectral air pollution sensor. The results show the potential of direct optical performance analysis in topology optimization, creating a structure that maintains optical performance at or above nominal performance, while highlighting challenges posed by the optimization process.

Keywords: Topology optimization, Ray tracing, STOP analysis, Automatic differentiation

1 Introduction

Space-based opto-mechanical instruments have to meet strict performance requirements. The main challenge is to maintain the position and shape of optical components in all operating conditions. These conditions can include gravity release, vibrations and thermal loads. For instruments in orbit, thermal disturbances are especially common due to heat produced by the sun or active components, combined with the lack of convective heat dissipation. Designing a structure that performs well in all conditions is not trivial. A perfectly thermally isolated system is one that is not connected to its surroundings. However, the structure should also provide enough stiffness for the optical components to stay in nominal position, survive transport or launch and keep the fundamental

eigenfrequency above the critical minimum. Moreover, the coupled interactions between components are further intertwined by miniaturization of the instruments, enabled by freeform optics. These optics allow for more compact designs over their conventional optics counterparts [1]. Some notable examples for space-based spectrometers include Sentinal-5, CHAPS-D and TANGO [2].

In the current workflow of opto-mechanical instruments, optical designers create an optical design with a corresponding error budget on displacements and deformations for all individual components. The structural designer then has to create a design to keep all the components under their allocated budget for a variety of above mentioned load cases. Optical performance is not directly considered, limiting the design space and prevents realizing full performance potential.

A method to directly analyze an optical instrument's performance under thermal load is the structural thermal optical (STOP) analysis. It works in two steps first, the displacement and deformation of the structure due to the thermal load is determined. Then, the displacements and deformations of the structure and components are translated into optical performance [3]. This performance can be expressed by image quality, optical resolution and image position accuracy. Image quality and optical resolution can be quantified by the RMS spot radius and image position by the Line Of Sight (LOS) error. Several methods exist to map displacements and deformations to optical performance such as linear optical sensitivities, Zernike fitting, or ray tracing with the latter being the most accurate [4]. Ray tracing is a computational method that determines the propagation of light through an optical system. Both the RMS spot radius and the line of sight error can be determined using ray tracing.

In literature a host of methods exist to trace rays through a system. A well known method is using the ray transfer matrix. It simplifies the optical physics by using the paraxial approximation, which assumes light only hits the optical surface under low angles. It allows a ray to be described as a vector and optical components to be represented by a matrix. Multiplying the ray vector with the optical component matrix together gives the ray that results from their interaction. Due to the paraxial approximation, it is only possible to accurately describe rays moving through a system if all components lay on the optical axis. To alleviate this problem Corcovilos [5] developed a method using a homogeneous coordinate description for a ray. The ray can be transformed to the local reference frame of an optical component, after which the resulting ray is transformed back to the global reference frame. This allows off axis systems to be analyzed without violating the conditions for the paraxial approximation. Expanding on this framework is Dorst [6], using algebraic geometry to extend the method to 3D systems and being able to not only image rays and points but to also image lines, planes and volumes. Yuan *et al.* [7] use the three dimensional version of the ray transfer matrix and added an extra column to describe error terms due to mirror deformation and displacement. A more elaborate method, not limited by the paraxial approximation, is proposed by Caron and Baumer [8] [9]. In this work, equations relating freeform shape coefficients to third order aberrations are composed. It allows ray tracing through a system with off axis freeform elements. The proposed use case is to identify special solutions to reduce low order aberrations and give insight in the number of degrees of freedom that are required to cancel certain aberrations. A few essential steps of the method can be recognized. The intersection of an incoming ray with a surface is first calculated. This intersection point is used to calculate the surface normal vector which is used to determine the reflecting direction. Combining this procedure with the aforementioned homogeneous coordinates to transform rays between reference frames should allow for a robust and general ray tracing algorithm.

The trade-offs posed in opto-mechanical design lend themself well to optimization techniques. Several authors have already used optical performance as an objective for the optimization of structures and show promising results. Liang *et al.* optimize the position of several mirror supporting pins to reduce the minimal wave front error under the mirrors' own weight [10]. Wang *et al.* optimize the thickness of beams on a frame of a spectrometer under gravity release and thermal load, using ZEMAX for the optical performance evaluation [11]. Song *et al.* propose a method to model the

impact of the dynamic behavior of a photoelectric detection system on optical performance using homogeneous coordinate transformation and ray tracing [12].

Another structural optimization technique is topology optimization. Given loads, boundary conditions and constraints, it distributes material within the design domain to maximize a specified performance measure [13]. Regularly used performance measures include stiffness, mass and eigenfrequency. Thermal loads have been investigated using topology optimization, where minimum compliance was required under prescribed temperatures [14] [15]. Topology optimization has been used to design mirror mounts for optical performance. For example, Sahu *et al.* minimize mirror deformation under static loads [16] and Van der Kolk *et al.* use the placement of viscoelastic material to attenuate the resonant modes present in instruments [17]. Koppen *et al.* performed a system design optimization combining topology optimization and a full STOP analysis to optimize a two mirror system under a thermomechanical load for optical performance [18]. The results show that directly including optical performance, in the form of the RMS spot radius, allows the two mirrors to compensate for each others optical errors. However, this analysis is only performed in 2D and uses the ray transfer description of Yuan *et al.* [7], which is limited by the paraxial approximation.

In optical system optimization, differential ray tracing is widely used to optimize individual surface shapes and component positions for better optical performance of the system. It is especially useful when considering freeform optics as they can poses a large amount of surface coefficients. For example Volatier *et al.* use differential ray tracing and automatic differentiation to optimize surface shape and position with the optical surfaces being described as Non-uniform rational B-splines (NURBS) [19]. Furthermore Pflaum uses differential ray tracing to successfully design a freeform lens to project a specified irradiance distribution [20]. Using differential ray tracing for structural optimization would bridge the gap between structural displacements and optical performance and can provide the necessary gradient information.

In this work, differential ray tracing is implemented in a fully coupled structural thermal optical analysis to drive the topology optimization of opto-mechanical structures. This alleviates the need for optical designers to assign an error budget per component, and create systems with improved optical performance under equivalent loads. By using the general description of ray tracing it will allow 3D analysis on a wide range of optical components including more complex optics such as freeform optics, while not being limited by the paraxial approximation. In Chapter 2, the working principles of differential ray tracing is explained. Chapter 3 introduces two optimization problems: a simple test case of a single mirror on the end of a cantilever beam, and a more complex spectrometer consisting of 12 optical components. The results are discussed and recommendations given in Chapter 4.

2 Differential ray tracing

In this section the implementation of differential ray tracing is described. Starting with the homogeneous coordinates description of points and vectors making transformation and rotations possible by matrix multiplication. Then, a description is given on how ray intersections, reflections and refraction are calculated. Lastly, the calculation of the derivatives is explained.

2.1 Homogeneous coordinates

Homogeneous coordinates are used to simplify calculations that would be complex in Cartesian coordinates. An example is affine and projective transformations, which in homogeneous coordinates can simply be calculated by a matrix multiplication. Given a point $[x, y, z]$ in Euclidean space, for a non-zero real number w, $[x, y, z, w]$ is the homogeneous coordinates set of the point. The Cartesian coordinate represented by a homogeneous coordinate can be determined by dividing x, y, z by w. With this property homogeneous coordinates $[1, 2, 4, 1]$ and $[2, 4, 8, 2]$ represent the same point in Cartesian coordinates. A point in homogeneous coordinates with a w value of zero is a point at infinity.

A ray can be described in Cartesian as a point and a direction. In homogeneous coordinates this translates well to a point $\mathbf{p}$, using a value of 1

for w, giving x,y and z the exact same values in both coordinate systems, and a point at infinity as a direction, $\mathbf{d}$. This results in the following ray description:

$$\mathbf{r} = \begin{bmatrix}\mathbf{p} & \mathbf{d}\end{bmatrix}, \mathbf{p} = \begin{bmatrix}x_0 \\ y_0 \\ z_0 \\ 1\end{bmatrix}, \mathbf{d} = \begin{bmatrix}dx \\ dy \\ dz \\ 0\end{bmatrix} \tag{1}$$

This formulation allows for easy transformations between global and local reference frames of the optical components. A ray can be transformed using the translation matrix $\mathbf{T}$ and rotation matrix $\mathbf{R}$, where u,v,w,ϕ,θ and ψ describe position and rotation in the global reference frame, as:

$$\mathbf{T} = \begin{bmatrix}1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & w \\ 0 & 0 & 0 & 1\end{bmatrix} \tag{2}$$

$$\mathbf{R} = \begin{bmatrix}[\mathbf{R}_x(\phi)] & \mathbf{0} \\ \mathbf{0} & 1\end{bmatrix} \cdot \begin{bmatrix}[\mathbf{R}_y(\theta)] & \mathbf{0} \\ \mathbf{0} & 1\end{bmatrix} \cdot \begin{bmatrix}[\mathbf{R}_z(\psi)] & \mathbf{0} \\ \mathbf{0} & 1\end{bmatrix} \tag{3}$$

The surface shape of an optical component is described in its local reference frame. To calculate where the globally defined ray intersects the locally defined component it needs to be transformed as described in Equation (4):

$$\tilde{\mathbf{r}} = \mathbf{R}^{-1}\mathbf{T}^{-1}\mathbf{r}, \tag{4}$$

where $\tilde{\mathbf{r}}$ denotes the ray in local coordinates. The resulting ray is then transformed back to the global reference frame as:

$$\mathbf{r} = \mathbf{T}\mathbf{R}\tilde{\mathbf{r}}. \tag{5}$$

2.2 Surface intersection, reflection, refraction and diffraction

An optical surface is described in its local reference frame as $z = \text{sag}(x, y)$. To calculate the intersection between a local ray $\tilde{\mathbf{r}}$, and a surface, the local ray's description is converted to a parametric line equation and placed in the optical surface equation, as:

$$f(t) = \text{sag}(x_0 + tdx, y_0 + tdy) - (z_0 + tdz). \tag{6}$$

$$= c_0 + c_1 t + \cdots + c_{n-1}t^{n-1} + t^n$$

The intersection $\mathbf{p}_{intersect}$, is calculated by finding $t = t_0$ for which $f(t_0) = 0$. In practice this is done by composing a companion matrix using the coefficients of the polynomial in t and calculating its eigenvalue:

$$\mathbf{C} = \begin{bmatrix}0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1}.\end{bmatrix} \tag{7}$$

The eigenvalue of $\mathbf{C}$ can be plugged back into the parametric line description of the original ray to find the intersection point:

$$\mathbf{p}_{\text{intersect}} = \begin{bmatrix}x_0 + t_0 dx \\ y_0 + t_0 dy \\ \text{sag}(x_0 + t_0 dx, y_0 + t_0 dy) \\ 1\end{bmatrix}. \tag{8}$$

This intersection point is then used as the origin point for the next ray. For reflective surfaces, the new direction $\mathbf{d}_{\text{reflect}}$, is calculated using the surface normal vector $\mathbf{n}$, as stated in Equation (9) and depicted in Figure 1 [21]:

$$\mathbf{d}_{\text{reflect}} = \mathbf{d} - 2(\mathbf{d} \cdot \mathbf{n})\mathbf{n}. \tag{9}$$

The refracted ray direction for a transmissive surface is calculated using Equation (10), where n is the refractive index of the medium surrounding the transmissive component and n_t the refractive index of the material of the component:

$$\mathbf{d}_{\text{refract}} = \frac{n(\mathbf{d} - \mathbf{n}(\mathbf{d} \cdot \mathbf{n}))}{n_t} - \mathbf{n}\sqrt{1 - \frac{n^2\left(1 - (\mathbf{d} \cdot \mathbf{n})^2\right)}{n_t^2}}. \tag{10}$$

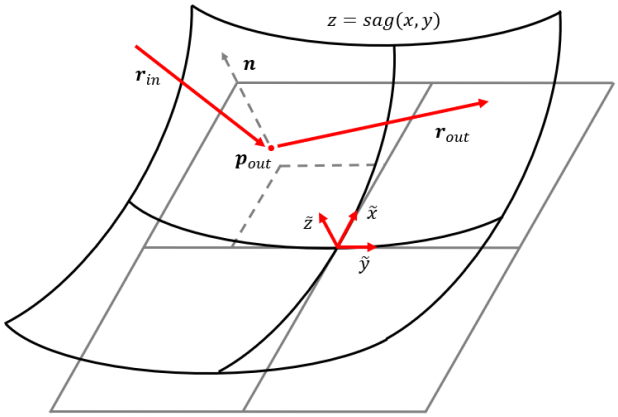


Fig. 1: Reflection on a locally defined mirror surface, using the surface normal vector

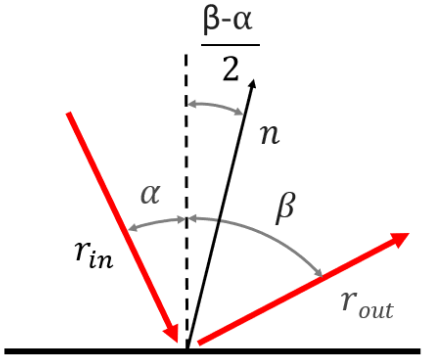


Fig. 2: Diffraction by rotation of the normal vector

Besides reflection and refraction, diffraction is also commonly used in optical systems, specifically spectrometers. For a reflective grating, the normal vector is artificially rotated in the direction of diffraction to simulate a mirror reflection (Figure 2). The incoming angle α, of a ray is first determined using Equation (11). Depending on the diffraction direction either ray component dx or dy is used.

$$\alpha = \tan^{-1}(\frac{dy}{dz}) \tag{11}$$

$$\beta = \sin^{-1}(\frac{k\lambda}{d} + \sin(\alpha)) \tag{12}$$

The angle that the reflected ray should have β, determined with k, the diffraction order d, the line spacing and λ, the wavelength. The normal vector can then be rotated by half of the difference between α and β *i.e.* $\frac{\beta-\alpha}{2}$ to make the ray reflect with the correct angle.

2.3 Ray differentiation

Several methods can be considered for calculating the derivatives for ray tracing. The derivative can be implemented manually, allowing for it to be calculated relatively fast. However, the operation has to be described analytically. Moreover, its not a flexible method as each new derivative has to be determined manually again making it a tedious process. The finite difference method can also be used, it does not require the operation to be analytically describable. Finite differencing performs two function evaluations separated by a small value of h, where the slope between the function evaluations is the derivative. Finite differencing relies on h being a small enough number that the section in between the two evaluations can be linearly approximated. The difficulty then lies in choosing the right value for h, as to not make it too large that the function can become nonlinear and not too small that numerical noise starts to take effect. Even still, finite differencing only gives estimating values and it requires two function evaluations, which can get costly in the case of many variables or long calculation times per evaluation. Another approach is symbolic differentiation, which takes an input function and recursively solves for the derivative to get its symbolic expression. The advantages are that the derivative can be given as a function, allowing it to be interpreted and it generates exact results. However, when the expression gets large and the recursion depth increases, symbolic differentiation will become slow to use. An additional differentiation method is automatic differentiation, its working principle is to convert an expression into a sequence of elementary operations. For each elementary operation the computer knows what the derivative is. Using the chain rule the computer can then accumulate derivatives together to get the derivative for the full expression. This accumulation can be done in a forward mode and a reverse mode. Forward mode traverses the chain rule from the inside starting with an input and calculating the derivatives to all outputs. Reverse

mode starts from an output and directly calculates the derivatives to all inputs. The benefit of automatic differentiation is that it is fast, especially when there are either many inputs or many outputs, and that it gives exact derivatives [22].

An example of a sequence is shown in Figure 3, where the function $y = (x_1 x_2 x_3)^2 + sin(x_1)$ is split into elementary operations. Table 1 shows how the forward evaluation is divided and how the derivatives are calculated in reverse mode. More generally, if w_j represent all successive operations of w_i then Equation (13) describes how the derivative of w_i is calculated. In reverse mode, the derivative of an output is seeded with a value of one as the derivative towards itself is one.

$$\bar{w}_i = \sum_j \bar{w}_j \frac{\partial w_j}{\partial w_i} \tag{13}$$

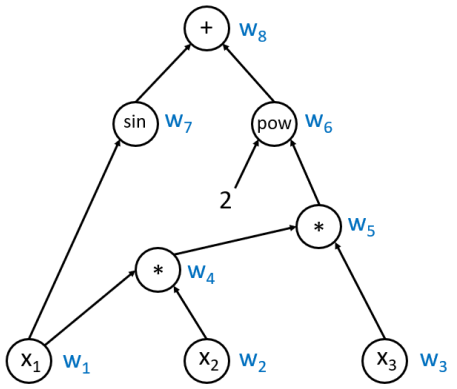


Fig. 3: Sequence of elementary operations

Table 1: AD evaluation and reverse derivative calculations

Forward evaluation	Reverse adjoint
w_1	$\bar{w_8} = 1$
w_2	$\bar{w_7} = \bar{w_8} \frac{\partial w_8}{\partial w_7}$
w_3	$\bar{w_6} = \bar{w_8} \frac{\partial w_8}{\partial w_6}$
$w_4 = w_1$	$\bar{w_5} = \bar{w_6} \frac{\partial w_6}{\partial w_5}$
$w_5 = w_4$	$\bar{w_4} = \bar{w_5} \frac{\partial w_5}{\partial w_4}$
$w_6 = w_5^2$	$\bar{w_3} = \bar{w_5} \frac{\partial w_5}{\partial w_3}$
$w_7 = sin(w_1)$	$\bar{w_2} = \bar{w_4} \frac{\partial w_4}{\partial w_2}$
$w_8 = w_6 + w_7$	$\bar{w_1} = \bar{w_4} \frac{\partial w_4}{\partial w_1} + \bar{w_7} \frac{\partial w_7}{\partial w_1}$

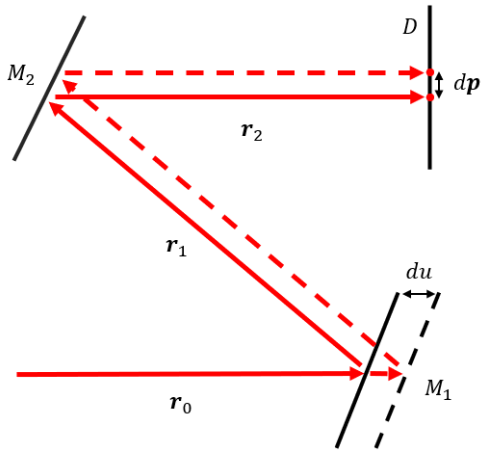


Fig. 4: Effect on the ray path due to displacements in a mirror

To differentiate the ray tracing, semi-automatic differentiation is used. It differs from automatic differentiation in that the expression is not converted into a sequence of elementary operations but a sequence of manually selected operations. An optical system can already be viewed as a sequences of operations that change the properties of light, making it a good fit. As an example, the optical system in Figure 4 is decomposed as shown in Figure 5. Each node represents an optical element in the system, with the last node f representing the performance metric. The optical element nodes have two inputs a ray $\mathbf{r}_i$ and position $\mathbf{x} = [u, v, w, \phi, \theta, \psi]$.

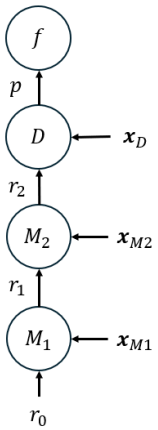


Fig. 5: Sequence of operations for Figure 4

The derivative of objective f towards the position of the first mirror can be written as Equation (14). The semi automatic differentiation calculates the total derivative by multiplying the individual derivatives together, while the individual derivatives are analytically implemented.

$$\frac{\partial f}{\partial \mathbf{x}_{M1}} = \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial \mathbf{p}}{\partial \mathbf{r}_2} \cdot \frac{\partial \mathbf{r}_2}{\partial \mathbf{r}_1} \cdot \frac{\partial \mathbf{r}_1}{\partial \mathbf{x}_{M1}} \tag{14}$$

For ease of implementation and insight into the process, the derivatives are divided further using the chain rule. As an example, the most comprehensive derivative $\frac{\partial \mathbf{r}_2}{\partial \mathbf{r}_1}$ is broken down into its fundamental parts. As previously mentioned the ray is composed of a point and a direction:

$$\frac{\partial \mathbf{r}_2}{\partial \mathbf{r}_1} = \begin{bmatrix} \frac{\partial \mathbf{p}_2}{\partial \mathbf{r}_1} & \frac{\partial \mathbf{d}_2}{\partial \mathbf{r}_1} \end{bmatrix}. \tag{15}$$

Starting with the point, its derivative can be broken down, separating the derivatives that are created transforming the ray in and out of the mirrors local reference frame. These derivatives simply reduce to the matrices applied for transformation:

$$\frac{\partial \mathbf{p}_2}{\partial \mathbf{r}_1} = \frac{\partial \mathbf{p}_2}{\partial \tilde{\mathbf{p}}_2} \frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1} \frac{\partial \tilde{\mathbf{r}}_1}{\partial \mathbf{r}_1} = TR \frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1} R^{-1} T^{-1}. \tag{16}$$

Calculating $\frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1}$ requires the derivative of the root of Equation (6). This can be obtained by calculating the derivative to the eigenvalue of the companion matrix in (7) using its left and right eigenvectors as stated:

$$\frac{\partial t_0}{\partial \tilde{\mathbf{r}}_1} = \mathbf{w}^T \frac{\partial \mathbf{C}}{\partial \tilde{\mathbf{r}}_1} \mathbf{v}. \tag{17}$$

The derivative of the direction is similarly obtained by first separating it from its transformation to the global reference frame:

$$\frac{\partial \mathbf{d}_2}{\partial \mathbf{r}_1} = \frac{\partial \mathbf{d}_2}{\partial \tilde{\mathbf{d}}_2} \frac{\partial \tilde{\mathbf{d}}_2}{\partial \mathbf{r}_1} = TR \frac{\partial \tilde{\mathbf{d}}_2}{\partial \mathbf{r}_1} \tag{18}$$

Then $\frac{\partial \tilde{\mathbf{d}}_2}{\partial \mathbf{r}_1}$ can be determined as the derivative of the reflection Equation (9):

$$\frac{\partial \tilde{\mathbf{d}}_2}{\partial \mathbf{r}_1} = \left(\frac{\partial \tilde{\mathbf{d}}_1}{\partial \mathbf{r}_1} - 2\left(\left(\frac{\partial \tilde{\mathbf{d}}_1}{\partial \mathbf{r}_1}\mathbf{n} + \tilde{\mathbf{d}}_1 \frac{\partial \mathbf{n}}{\partial \mathbf{r}_1}\right)\mathbf{n} + (\tilde{\mathbf{d}}_1 \cdot \mathbf{n}) \frac{\partial \mathbf{n}}{\partial \mathbf{r}_1}\right)\right) \tag{19}$$

$\frac{\partial \mathbf{n}}{\partial \mathbf{r}_1}$ can be separated into three parts with $\frac{\partial \mathbf{n}}{\partial \tilde{\mathbf{p}}_2}$ being the second derivative of the mirrors surface and $\frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1}$ is already determined in calculating the derivative towards the point. The derivative can therefore be written down as:

$$\frac{\partial \mathbf{n}}{\partial \mathbf{r}_1} = \frac{\partial \mathbf{n}}{\partial \tilde{\mathbf{p}}_2} \frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1} \frac{\partial \tilde{\mathbf{r}}_1}{\partial \mathbf{r}_1} = \frac{\partial \mathbf{n}}{\partial \tilde{\mathbf{p}}_2} \frac{\partial \tilde{\mathbf{p}}_2}{\partial \tilde{\mathbf{r}}_1} R^{-1} T^{-1}. \tag{20}$$

Using this method makes it easy to quickly evaluate different optical systems. It relies on manually implementing the individual derivatives of optical components, but once implemented these derivatives can be reused for any similar component with a different surface shape and easily chained together by the semi automatic differentiation. To perform a topology optimization the performance sensitivity needs to be extended to the design variables. Using the adjoint method this will take a single solve per objective h to get the adjoint vector λ, with $\mathbf{K}$ and $\mathbf{u}$ being the stiffness matrix and displacement vector and x_i the design densities.

$$\lambda = \mathbf{K}^{-T} \frac{dh}{d\mathbf{u}} \tag{21}$$

$$\frac{dh}{dx_i} = -\lambda^T \frac{\partial \mathbf{K}}{\partial x_i} \mathbf{u} \tag{22}$$

In this case, objective h is the rigid body translations and rotations of all the optical components.

3 Numerical examples

In this Chapter four test cases are introduced and their results shown. In Table 2, a short description of their objectives is given.

3.1 Ray tracing verification

To verify that the proposed ray tracing method is accurate, it is compared to ZEMAX. To compare results, the telescope section of CHAPS is used (Figure 6) and its TM1 mirror translated

Test	Test objective
1	Verifying the ray tracing response by comparing it to ZEMAX
2	Verifying the accurate modeling of homogeneous expansion
3	Comparing results and behavior of a simple topology optimization problem using ray tracing
4	Comparing results and behavior of a complex topology optimization problem using ray tracing

Table 2: Overview of test cases and their objective

and rotated in all degrees of freedom. The resulting RMS spot radii are shown in Figure 8 and overlap perfectly with the response from ZEMAX. This confirms that the ray tracing method can accurately model the behavior of multi component systems.

Fig. 6: CHAPS telescope section

3.2 Homogeneous expansion

A second test to validate the ray tracing method is to check the behavior of optical performance while being subjected to homogeneous expansion. This is because optical instruments retain their exact optical properties when homogeneously expanded. For this test a telescope is used as shown in Figure 7. It uses four off-axis parabolic to reduce the incoming beam diameter by ten times. Hexapolar sampling is used to sample 37 rays over the aperture creating a RMS spot radius of 4.011 mm. The position and surface coefficients are scaled with a factor s as described in Equation (23), using a expansion coefficient of 22e-6 1/K and temperature difference of 1000 K. The temperature difference is an exaggerated large value, which should demonstrate how accurate, homogeneous expansion can be modeled. The scaling simulates the homogeneous expansion of the system. With scaling applied, the RMS spot radius remains 4.011 mm and is exactly the same up to the 12th decimal, proving the ray tracing method

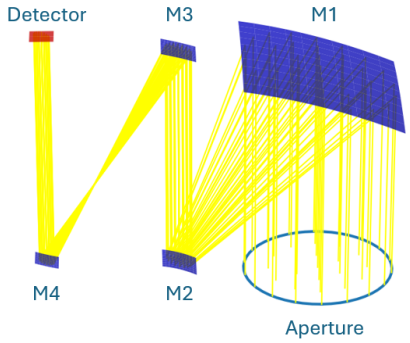


Fig. 7: Telescope using four off-axis parabolic mirrors

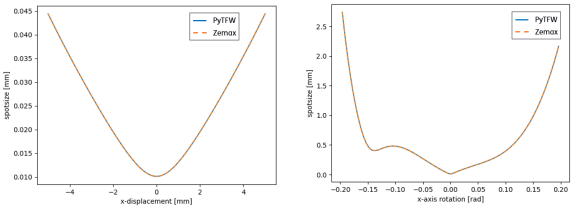
can accurately model optical systems undergoing homogeneous expansion.

$$s = \alpha \cdot dT \tag{23}$$

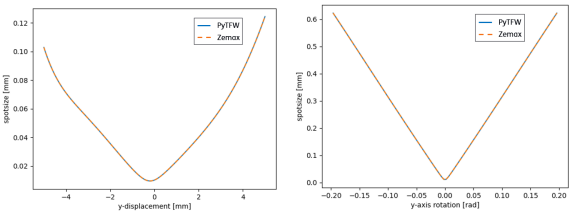
3.3 Cantilever beam

A simple case is set up to test the behavior of the differential ray tracing being used in topology optimization. It consist of a cantilever beam fully constrained on the left side, with a length of 250 mm and a cross section of 50x50 mm and made out of titanium alloy Ti6Al4V. A parabolic mirror is modeled at the other end of the beam. As Figure 9 shows, the mirror perfectly focuses a collimated bundle of light on the detector placed above. Below the mirror a force of 250 kN is exerted on the structure to bring the system out of focus. This is an exaggerated force to impose a large deformation on the structure to show the difference between optimization methods. Three different optimizations are performed, a compliance minimization, a displacement minimization and a RMS spot radius minimization. Function g_0 is the objective of the problem and is varied for the three cases, this makes the general problem statement:

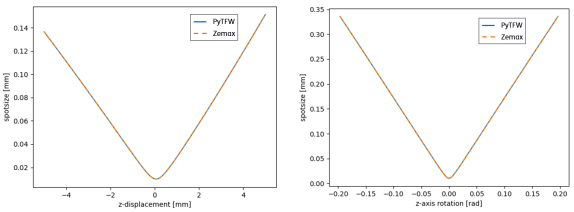
$$\begin{aligned} \min_{\rho} \quad & g_0 \\ \text{s.t.} \quad & K(\rho)\mathbf{u} = \mathbf{f} \\ & m = \sum_{e \in \Omega^d} \rho_e v_e \leq m_{max} \\ & 0 \leq \rho \leq 1 \end{aligned}$$



(a) Translation x-axis (b) Rotation x-axis



(c) Translation y-axis (d) Rotation y-axis



(e) Translation z-axis (f) Rotation z-axis

Fig. 8: Ray tracing method compared to ZEMAX

where Ω^d, defines the discritized design domain made up of elements, e. For the first optimization problem objective g_0, is compliance:

$$g_0 = C(\rho) = \mathbf{f}^T \mathbf{u}(\rho). \tag{24}$$

For the second optimization problem, all translations and rotations of the parabolic mirror d_i, minus the average displacement of the whole structure d_{avg}, are aggregated together using the P-Norm:

$$g_0 = d_{PN} = \left(\sum_{i=1}^{N} \left(\frac{d_i - d_{avg}}{d_{max}} + 1 \right)^P \right)^{(1/P)}. \tag{25}$$

The average displacement is subducted from the total as only relative displacement is of impact on the optical performance. In the third optimization

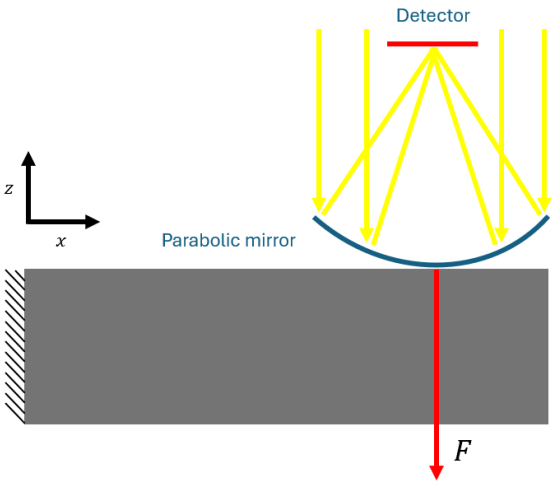


Fig. 9: A cantilever beam with a parabolic mirror, focusing light on a detector

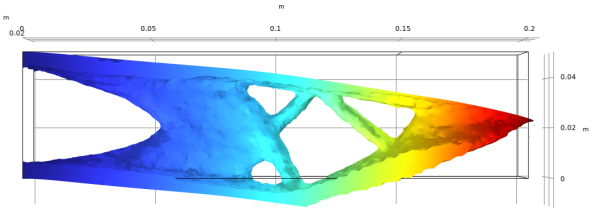
problem, r is minimized and is defined as the RMS spot radius:

$$g_0 = r = \sqrt{\frac{1}{N} \sum_{i=0}^{i<N} (x_i - x_c)^2 + \frac{1}{N} \sum_{i=0}^{i<N} (y_i - y_c)^2} \tag{26}$$

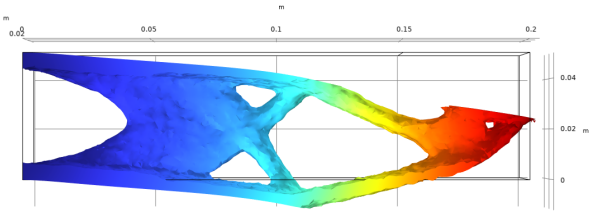
where, x_i and y_i is the location of a ray on the detector with N the total amount of rays and x_c and y_c the location of the chief ray on the detector.

The resulting three structures are shown in Figure 10. They show that optimizing directly for RMS spot radius gives about a 7% improvement over displacement minimization and about a 61% improvement over compliance minimization. Even more interesting is the way it reaches this performance, Figure 10c and Table 3 shows that it deforms more than the other two results. However it deforms in a different way, prioritizing a smaller rotation over the y-axis over displacement in the z-direction. Table 4 show that it moves about 3 mm less in the z-direction but has a reduced y-rotation. This behavior could be explained when considering the aberrations each movement makes. A displacement in the z-direction creates defocus while rotation over the y-axis creates tilt. Tilt creates a considerably larger RMS spot radius as it elongates the spot in the tilt direction, while defocus homogeneously increases the radius of the

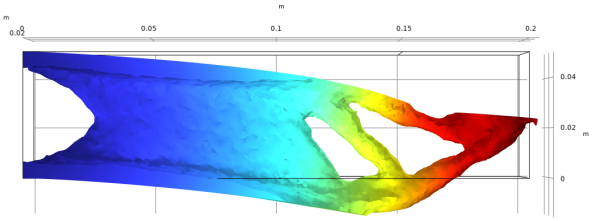
spot (Figure 11). Because the optimizer has access to the ray tracing it is able to indirectly discern which aberration is most influential on the RMS spot radius and decrease the movement that creates that aberration. Looking at Figure 10c, the structure and the end of the beam is similar to a compliant coupling which rotates inward due to the force applied. A comparable structure can be seen in Figure 10b, however it creates less inward rotation.



(a) Compliance minimization



(b) Deformation minimization



(c) RMS spot radius minimization

Fig. 10: Optimized cantilever beam for all three optimization problems. The color gradient depicts the magnitude of displacement.

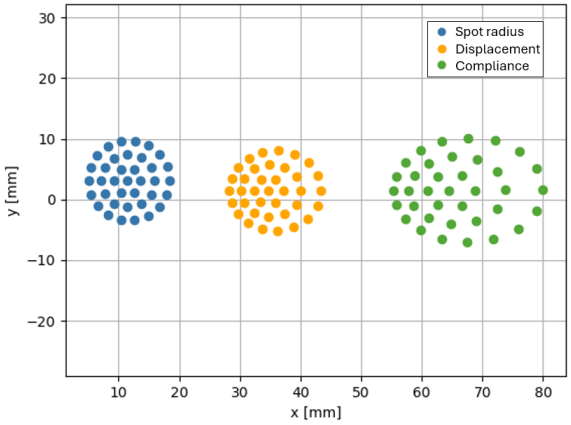


Fig. 11: Spots of the three different cantilevers structures on the detector

Optimization	RMS spot radius [mm]	Strain energy [J]
1	8.60	2730
2	5.72	2835
3	5.34	3216

Table 3: RMS spot radius and strain energy for each optimization result

Optimization	z-displacement [mm]	y-rotation [rad]
1	-21.8	0.224
2	-22.7	0.120
3	-25.7	0.027

Table 4: Cantilever displacement and rotation for each optimization result

3.4 CHAPS spectrometer

An additional test case will focus on a more elaborate optical system, inspired by a the Compact Hyperspectral Air Pollution Sensor (CHAPS) [23]. It is a spectrometer designed to detect and map trace gasses in the atmosphere. The optical path of CHAPS is shown in Figure 12, where TM1, TM2, Collimator, IM1, IM2 and IM3 are freeform mirrors. An optimization using CHAPS will be a useful test case to see how the optical optimization behaves for a state of the art optical instrument with multiple optical components. First, an optimization just constraining the optical component displacements is performed and this is compared to the second optimization using the ray tracer to directly evaluate optical performance.

The design domain contains the optical components as well as a light tight shell. Due to manufacturing limitations the system is split up into two parts along the dotted line in Figure 13a. Six interface locations have been predetermined to fully constrain the system.

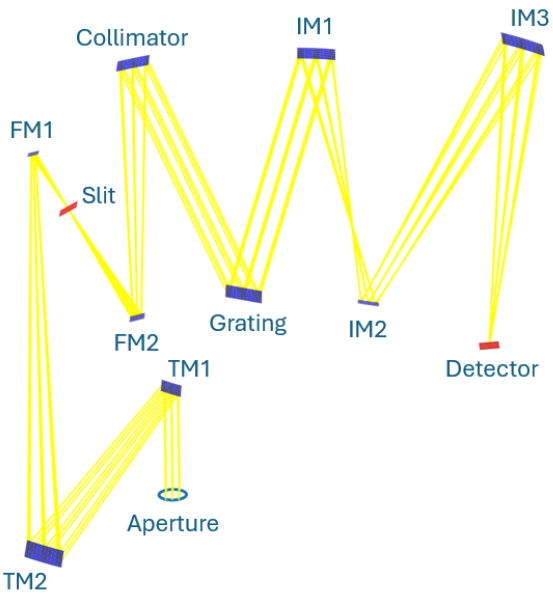
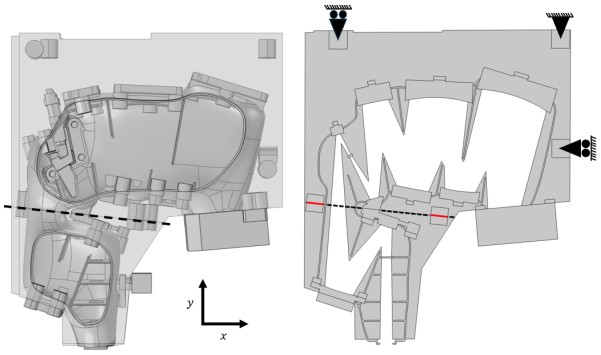


Fig. 12: Optical path and components in CHAPS

The analysis is performed in 2D to limit the calculation time required for each optimization. To go from a 3D design domain to a 2D domain, a slice of the middle section of the 3D domain is taken. Important components that are not positioned in the middle such as the interface location are projected onto the 2D domain. The interfaces are constrained as shown in Figure 13b, with the top right interface fixed in x and y direction and the other interfaces fixed in x and y, respectively. This allows the structure to freely expand under thermal load. The design domain is split along the dotted line, with two interfaces on either side indicated in red, connecting the two halves together.



(a) 3D design domain (b) 2D design domain

Fig. 13: 2D design domain compared to 3D design domain

The design is analyzed for six load cases, a gravity release load in x and y directions, a thermal gradient in x and y directions and a quasi-static load simulating launch conditions at 200 g acceleration in x and y direction. For the gravity release and thermal gradient, optical performance is analyzed, while for the launch condition the stress is analyzed.

In the first case, optical performance is constrained by means of limiting the displacement of the optical components. In the second case, the ray tracer is used to directly assess and constrain the impact on the optical performance. The performance measures used are the RMS spot radius and the line of sight (LOS) error. The spectrometer has a wavelength range of 300 - 500 nm, however, likely due to the shape of the spot, using a wavelength of 398.4 nm makes the chief ray hit the middle of the detector and is therefore used in the analysis. A perpendicular beam of 49 collimated rays is traced to get a good balance between accuracy of the spot radius and calculation time. The rays are sampled in a square pattern on the aperture, giving a nominal RMS spot radius of 27.4646 μm.

The stress constraint Ψ, is formulated by Verbart *et al.* [24]. Then for k amount of load cases, where $k = 1, 2$ are the launch accelerations and $k = 3, 4, 5, 6$ are gravity and thermal loads, the first optimization problem becomes:

$$
\begin{aligned}
\min_{\rho} \quad & m = \sum_{e \in \Omega^d} \rho_e v_e \\
\text{s.t.} \quad & \Psi - 1 \leq 0, \quad k = 1, 2 \\
& d_{PN} - 1 \leq 0, \quad k = 3, 4, 5, 6 \\
& 0 \leq \rho \leq 1
\end{aligned}
$$

Similarly, the second optimization problem using ray tracing is defined as:

$$
\begin{aligned}
\min_{\rho} \quad & m = \sum_{e \in \Omega^d} \rho_e v_e \\
\text{s.t.} \quad & \Psi - 1 \leq 0, \quad k = 1, 2 \\
& r/r_{max} - 1 \leq 0, \quad k = 3, 4, 5, 6 \\
& p/p_{max} - 1 \leq 0, \quad k = 3, 4, 5, 6 \\
& 0 \leq \rho \leq 1
\end{aligned}
$$

Here, r is defined as the RMS spot radius with a maximum allowable value of r_{max} and p is the LOS error, with a maximum value of p_{max}. The maximum allowable RMS spot radius is set to the spot radius in nominal performance *i.e.* 27.4646 μm and the maximum LOS error is set to 1 μm. The RMS spot radius is defined in Equation (26) and the LOS error is defined as the absolute distance of the chief ray to the center of the detector:

$$
p = \sqrt{x_c^2 + y_c^2}. \tag{27}
$$

The resulting structure of the first optimization problem is shown in Figure 14a, and the results of the second optimization problem are shown in Figure 14b. The result of the first optimization problem shows well defined thin beams connecting the components to the interface locations, with a larger block of material on the top right. Interestingly the connection to the interfaces are relatively thin compared to the connection to the bottom right interface. These connections seem to minimize the heat transfer from the interfaces to the components. The structure is able to keep the RMS spot radius less or equal to the nominal RMS spot radius for the gravity load in the x direction and thermal load in the y direction as can be seen in Table 5. The LOS error is greater for the thermal loads compared to the gravity case, as expected.

The optimization with the ray tracer is able to get the RMS spot radius at of below the nominal RMS spot radius for all load cases and it is able to keep the LOS error to 1 μm as can be seen in Table 6. This is a promising improvement over the displacement constrained optimization. However the result still contains a lot of intermediate (grey) design densities which do not translate well into a real design.

	RMS radius [μm]	LOS error [μm]	Stress [-]
Gravity x	27.4556	6.247	1.53
Gravity y	27.4647	1.685	1.61
Thermal x	27.6597	22.11	NA
Thermal y	27.4582	15.50	NA

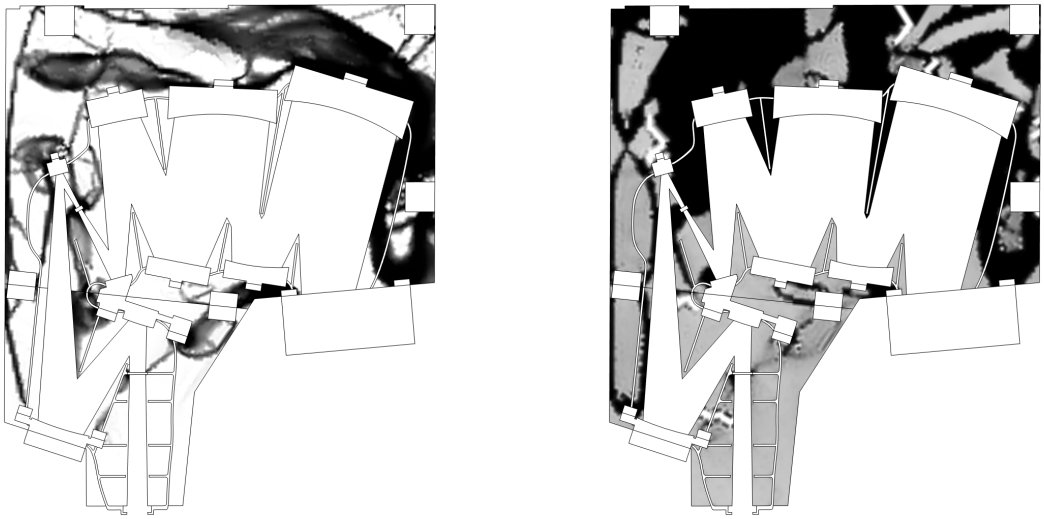
Table 5: Displacement constraint performance

	RMS radius [μm]	LOS error [μm]	Stress [-]
Gravity x	27.4532	0.112	1.2177
Gravity y	27.4646	0.994	1.2131
Thermal x	27.4643	1.051	NA
Thermal y	27.4646	0.841	NA

Table 6: RMS spot radius and LOS error constraint performance

A threshold is placed on the design densities in postprocessing to eliminate the intermediate values and get a clear separation of material and non-material (Figure 15). For the first optimization result the threshold is placed on density value of 0.16 and for the second optimization result the threshold is placed on 0.4. For the best comparison the thresholds should be equal, however the results are too different to allow it. The first optimization needs a low enough threshold value for all components to stay connected while the second optimization needs a threshold value high enough to remove the intermediate values. In doing so, the mass of the two design become more similar which will make a better comparison.

The structures are evaluated again for optical performance under the same loads and presented in Table 7 and 8. While the displacement constrained result reduces LOS error under the gravity loads, it increases the RMS spot radius and LOS error



(a) Displacement constraint, mass = 40.4 kg/m (b) Optical performance constraint, mass = 54.8 kg/m

Fig. 14: Optimized design domains

under thermal load. The optical performance constrained result generally reduces the RMS spot radius and increases the LOS error.

	RMS spot radius [μm]	LOS error [μm]
Gravity x	27.4594	2.52
Gravity y	27.4683	1.32
Thermal x	27.7270	27.01
Thermal y	27.4533	25.06

Table 7: Performance of the projected design domain, constrained by displacement

	RMS spot radius [μm]	LOS error [μm]
Gravity x	27.4589	1.735
Gravity y	27.4660	0.171
Thermal x	27.4093	6.956
Thermal y	27.4021	7.796

Table 8: Performance of the projected design domain, constrained by optical performance

While both structures support the stress constraint during optimization, this is not the case anymore when putting a threshold on the design densities. Intermediate gray values that would otherwise deform under load now take up this force leading to greater stresses. In Figure 16 several stress concentrations appear in the long beams connecting to the top left interface with a magnitude of 9.25. In Figure 17 the stress is better distributed over the structure, not having as high stress concentrations, with a peak magnitude of 2.38.

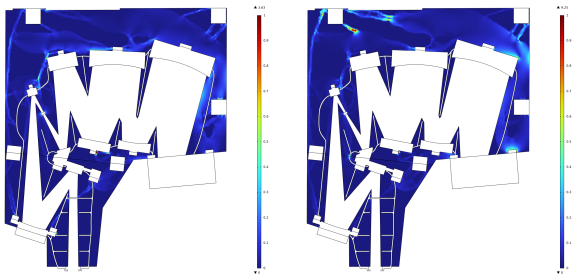
4 Discussion and recommendations

As can be seen in the case of optimizing the cantilever beam, optimizing directly for optical performance allows the optimizer to create better solutions compared to optimizing for displacements or stiffness. The same can be said in the optimization inspired by CHAPS, yet its design suffers from a large amount of intermediate density values between [0.25-0.35]. These values could be due to the thermal load as intermediate density values allow less heat transfer to occur.

During optimization, a recurring problem was the mass of the design going to zero. The parameter that seemed to have the most influence on this behavior was the value of the stress constraint. A too loose constraint would make the mass go to zero and a too strict constraint would prevent the optimizer from reaching the low stress value. The reason this happens might have to do with the fact that the other two load cases, namely the gravity

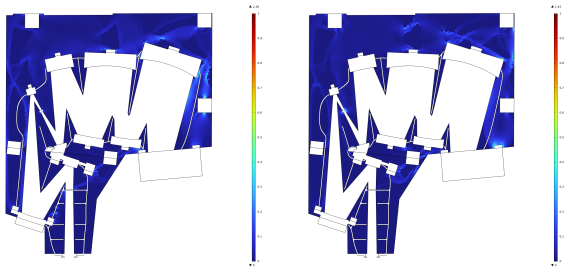
(a) Displacement constraint, mass = 48.6 kg/m (b) Optical performance constraint, mass = 52.0 kg/m

Fig. 15: Projected design domains



(a) Acceleration in x direction (b) Acceleration in y direction

Fig. 16: Stress due to 200 g acceleration in projected design domain



(a) Acceleration in x direction (b) Acceleration in y direction

Fig. 17: Stress due to 200 g acceleration in projected design domain

release and thermal gradient do not give a strong incentive to connect the structure to its interfaces. The difference between constraining RMS spot radius and LOS error compared to displacement is that relative motion between optical components now become important instead of absolute motion. This change shows itself during optimization by not connecting any of the optical components. For example, under a gravity load an unconnected system would be free floating through space but it would retain the relative position between optical components, and thus, its optical performance. Similar behavior is observed for a thermal load, as the best thermally isolated design is one that is not connected to its surroundings. It is likely that a too strict stress constraint was necessary for the design to stay connected. This could also be an explanation for the intermediate density values. Because removing the intermediate values with a threshold showed an doubling of peak stress in Figure 17, mainly around the thin flexure like structures that would otherwise be supported by intermediate densities.

Another problem is the discrepancy in performance between the output of the optimizer and its threshold design. A large part can be attributed to the presence of intermediate density values. A solution for this could be the robust formulation

creating a design that can vary slightly in size and still perform as expected and could also help in reducing the numerical instability [25].

The rays traced during optimization are from a single collimated beam of light, therefore all the rays hit the same location on the detector. During real operation, optical instruments take in a range of angles all being projected on a different part of the detector. To guarantee real world performance for all incoming angles, a range of angles need to be sampled during optimization and the same goes for a range of wavelengths. Depending on the calculation time for the FEM solve, tracing many spots on the detector will significantly increase total calculation time, so a small but representative selection needs to be made. The square sampling pattern can also create artifacts in spot diagram and skew the value of the RMS spot radius. A stochastic method like Poisson disk sampling could be used to remove artifacts from the spot diagram [26].

In the optimization of the spectrometer, the slit is not yet taken into account due to its low perceived impact in ray optics. However, to fully capture the behavior of the spectrometer, an accurate representation of the slit should be added.

To continue the work presented, the deformations of the mirrors could be taken into account. In the existing framework mirrors are assumed to be rigid, whereas in reality, mirrors deform due to gravity release and thermal gradients. This addition is needed to, for example, correctly determine the RMS spot radius during homogeneous expansion, where the curvature of the optical components slightly expand along with the rest of the structure resulting in nominal performance. Correctly modeling homogeneous expansion is now only possible by scaling the surface coefficients beforehand using the known temperature increase.

Another current limitation of this study is the impact on the computation time in going from 2D to 3D structures. In 3D, the number of degrees of freedom per optical element is doubled from 3 to 6, and, as in every topology optimization, the number of design variables is increased drastically. In the current implementation, which uses COMSOL for the finite element analysis, this requires a system solve per degree of freedom, which is a limiting factor in the computation time. It should be noted that this is only a limitation of the current implementation, and not of the method itself.

Eventually, it should be possible to optimize the structural design together with the optical design. Position, orientation and surface coefficients of the optical elements can be added as design variables giving the optimizer more feasible domain and freedom to create better performing systems.

5 Conclusion

A differential ray tracing approach has been presented that combined with topology optimization can generate opto-mechanical structures with higher optical performance subjected to several disturbance loads. It eliminates the need for optical designers to create an error budget for individual optical components allows the them to compensate each others errors. The optical response of the ray tracing method was tested against ZEMAX and show strong agreement. Two optimization cases were put forward to show the effectiveness of the approach, a cantilever beam subjected to a large displacing force and an optimization case inspired by a real state-of-the-art spectrometer, subjected to gravity, thermo-mechanical and launch acceleration loads. In the case of the cantilever beam the ray tracing approach showed a 61% improvement in RMS spot radius over a stiffness optimized beam and a 7% improvement over a displacement optimized beam. Moreover, in the spectrometer optimization case, it was able to keep nominal optical performance and retain more performance when densities were projected.

Acknowledgments The author thanks L.Dorst for his help in understanding the work he has done in the field of Algebraic geometry and T.J.Blok for his helping hand in formulating the ray tracing method.

References

[1] Rolland, J.P., Davies, M.A., Suleski, T.J., Evans, C., Bauer, A., Lambropoulos, J.C., Falaggis, K.: Freeform optics for imaging. Optica **8**(2), 161 (2021) https://doi.org/10.1364/OPTICA.413762

[2] Caron, J., Bäumer, S.: Progress in freeform mirror design for space applications. In: International Conference on Space Optics — ICSO 2020, vol. 11852, pp. 747–767. SPIE, ??? (2021). https://doi.org/10.1117/12.2599322

[3] Holzlöhner, R., Kellerer, A., Lampater, U., Lewis, S., Zanoni, C.: Structural, thermal, and optical performance analysis applied to subsystems of the European Extremely Large Telescope. Journal of Astronomical Telescopes, Instruments, and Systems **8**(2), 021504 (2022) https://doi.org/10.1117/1.JATIS.8.2.021504 . Publisher: SPIE

[4] Stoeckel, G., Crompton, D., Perron, G.: Advancements in integrated structural/thermal/optical (STOP) analysis of optical systems. In: Optical Modeling and Performance Predictions III, vol. 6675, pp. 140–153. SPIE, Bellingham, WA, USA (2007). https://doi.org/10.1117/12.732514

[5] Corcovilos, T.A.: Beyond the ABCDs: A better matrix method for geometric optics by using homogeneous coordinates. American Journal of Physics **91**(6), 449–457 (2023) https://doi.org/10.1119/5.0083069

[6] Dorst, L.: Paraxial Geometric Optics in 3D Through Point-Based Geometric Algebra, pp. 340–354 (2023). https://doi.org/10.1007/978-3-031-50078-7_27

[7] Yuan, J., Long, X., Chen, M.: Generalized ray matrix for spherical mirror reflection and its application in square ring resonators and monolithic triaxial ring resonators. Optics Express **19**(7), 6762 (2011) https://doi.org/10.1364/oe.19.006762

[8] Caron, J., Bäumer, S.: Aberrations of plane-symmetrical mirror systems with freeform surfaces. Part I: generalized ray-tracing equations. Journal of the Optical Society of America. A, Optics, Image Science, and Vision **38**(1), 80–89 (2021) https://doi.org/10.1364/JOSAA.411317

[9] Caron, J., Bäumer, S.: Aberrations of plane-symmetrical mirror systems with freeform surfaces. Part II: closed-form aberration formulas. Journal of the Optical Society of America. A, Optics, Image Science, and Vision **38**(1), 90–98 (2021) https://doi.org/10.1364/JOSAA.411323

[10] Liang, P., Song, X., Xiang, Y., Li, Z., Zhang, T., Zhang, J.: Mathematical modelling for large-aperture mirror systems: enabling optical-mechanical performance analysis and support structure optimization. Applied Mathematical Modelling **141**, 115938 (2025) https://doi.org/10.1016/j.apm.2025.115938

[11] Wang, Z., Ding, L., Cheng, L., Li, S., Lei, S.: Optomechanical integrated performance analysis and optimization of a greenhouse gas absorption spectrometer. Optik **295**, 171489 (2023) https://doi.org/10.1016/j.ijleo.2023.171489

[12] Song, X., Liang, P., Zhang, S., Pang, Y., Gong, Z., Yang, K., Zhang, J., Yuan, Z.: A modeling method for the opto-mechanical coupling problems of photoelectric detection and tracking systems in dynamics process. Structural and Multidisciplinary Optimization **67**(9), 1–18 (2024) https://doi.org/10.1007/s00158-024-03889-6

[13] Bendsoe, M.P., Sigmund, O.: Topology Optimization: Theory, Methods, and Applications. Springer, Berlin, Heidelberg, Germany (2013)

[14] Deaton, J.D., Grandhi, R.V.: Stiffening of restrained thermal structures via topology optimization. Structural and Multidisciplinary Optimization **48**(4), 731–745 (2013) https://doi.org/10.1007/s00158-013-0934-5

[15] Gao, T., Zhang, W.: Topology optimization involving thermo-elastic stress loads. Structural and Multidisciplinary Optimization **42**(5), 725–738 (2010) https://doi.org/10.1007/s00158-010-0527-5

[16] Sahu, R., Patel, V., Singh, S.K., Munjal, B.S.: Structural optimization of a space mirror to selectively constrain optical aberrations. Structural and Multidisciplinary Optimization **55**(6), 2353–2363 (2017) https://doi.

org/10.1007/s00158-016-1616-x

[17] Kolk, M., Veen, G.J., Vreugd, J., Langelaar, M.: Multi-material topology optimization of viscoelastically damped structures using a parametric level set method. Journal of Vibration and Control **23**(15), 2430–2443 (2017) https://doi.org/10.1177/1077546315617333

[18] Koppen, S., Kolk, M., Kempen, F.C.M., Vreugd, J., Langelaar, M.: Topology optimization of multicomponent optomechanical systems for improved optical performance. Structural and Multidisciplinary Optimization **58**(3), 885–901 (2018) https://doi.org/10.1007/s00158-018-1932-4

[19] Volatier, J.-B., Beaussier, S.J., Druart, G., Jougla, P., Keller, F.: Implementation of FORMIDABLE: A generalized differential optical design library with NURBS capabilities. Journal of the European Optical Society-Rapid Publications **20**(1), 2 (2024) https://doi.org/10.1051/jeos/2023043

[20] Pflaum, M.: Optimization with Algorithmic Differentiable Ray Tracing. Master's thesis, NTNU (2025)

[21] Marschner, S., Shirley, P.: Fundamentals of Computer Graphics. CRC Press, Boca Raton, FL, USA (2018)

[22] Margossian, C.C.: A Review of automatic differentiation and its efficient implementation. arXiv (2019). https://doi.org/10.48550/arXiv.1811.05031

[23] Stewart, B., Swartz, W.H., Morgan, F., Zimbeck, W., Palmer, T., Linden, J., Newport, R., Strang, J., Otter, G., Van Kempen, F., Van De Boom, S., Ferrario, I.: CHAPS-D: The Compact Hyperspectral Air Pollution Sensor–Demonstrator. In: 2024 IEEE Aerospace Conference, pp. 1–9 (2024). https://doi.org/10.1109/AERO58975.2024.10521395 . ISSN: 1095-323X

[24] Verbart, A., Langelaar, M., Keulen, F.v.: A unified aggregation and relaxation approach for stress-constrained topology optimization. Structural and Multidisciplinary Optimization **55**(2), 663–679 (2017) https://doi.org/10.1007/s00158-016-1524-0

[25] Wang, F., Lazarov, B.S., Sigmund, O.: On projection methods, convergence and robust formulations in topology optimization. Structural and Multidisciplinary Optimization **43**(6), 767–784 (2011) https://doi.org/10.1007/s00158-010-0602-y

[26] Bridson, R.: Fast Poisson disk sampling in arbitrary dimensions. In: ACM SIGGRAPH 2007 Sketches, p. 22. ACM, San Diego California (2007). https://doi.org/10.1145/1278780.1278807