



Universiteit
Leiden
The Netherlands

MECHBench: a set of black-box optimization benchmarks originated from structural mechanics

Olarte Rodríguez, I.; Santoni, M.L.; Duddeck, F.; Doerr, C.; Bäck, T.H.W.; Raponi, E.

Citation

Olarte Rodríguez, I., Santoni, M. L., Duddeck, F., Doerr, C., Bäck, T. H. W., & Raponi, E. (2025). MECHBench: a set of black-box optimization benchmarks originated from structural mechanics. *Arxiv E-Prints*. doi:10.48550/arXiv.2511.10821

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from:

Note: To cite this publication please use the final published version (if applicable).

MECHBench: A Set of Black-Box Optimization Benchmarks originated from Structural Mechanics

IVÁN OLARTE RODRÍGUEZ, Leiden University, The Netherlands

MARIA LAURA SANTONI, Sorbonne Université, CNRS, LIP6, France

FABIAN DUDDECK, Technical University of Munich, Germany

CAROLA DOERR, Sorbonne Université, CNRS, LIP6, France

THOMAS BÄCK, Leiden University, The Netherlands

ELENA RAPONI, Leiden University, The Netherlands

Benchmarking is essential for developing and evaluating black-box optimization algorithms, providing a structured means to analyze their search behavior. Its effectiveness relies on carefully selected problem sets used for evaluation. To date, most established benchmark suites for black-box optimization consist of abstract or synthetic problems that only partially capture the complexities of real-world engineering applications, thereby severely limiting the insights that can be gained for application-oriented optimization scenarios and reducing their practical impact. To close this gap, we propose a new benchmarking suite that addresses it by presenting a curated set of optimization benchmarks rooted in structural mechanics. The current implemented benchmarks are derived from vehicle crashworthiness scenarios, which inherently require the use of gradient-free algorithms due to the non-smooth, highly non-linear nature of the underlying models. Within this paper, the reader will find descriptions of the physical context of each case, the corresponding optimization problem formulations, and clear guidelines on how to employ the suite.

List of Symbols

δ	Intrusion	mm
$\mathbf{x}$	Input design/parameter vector	
ν	Poisson's ratio	
ρ	Mass density	kg/m ³
σ_y	Yield Strength	MPa
SEA	Specific Energy Absorbed	J/kg
d	Problem Dimension	
E	Young's Modulus	GPa
E_{abs}	Absorbed Energy by structure	J
F_{mean}	Mean Force over the loading duration	kN
F_{peak}	Peak Force during impact	kN
m_s	Mass of the structure	kg
m_I	Mass of the impactor	kg
z	z position in $\mathbb{R}^3$	mm
LU	Load Uniformity	kN/kN

1 Introduction

Optimization in structural mechanics plays a critical role in the design of lightweight, cost-efficient, and crash-resistant structures. In specific contexts such as vehicle crashworthiness, where occupant safety must be balanced with structural performance and weight constraints, optimization techniques are essential for developing effective energy-absorbing designs [7].

Many applications in structural mechanics are modeled as non-linear dynamical systems due to phenomena such as fracture, damage, and material plasticity, which inherently exhibit non-linear behavior. This hinders

Authors' Contact Information: Iván Olarte Rodríguez, Leiden University, LIACS, Leiden, The Netherlands, i.olarte.rodriguez@liacs.leidenuniv.nl; Maria Laura Santoni, Sorbonne Université, CNRS, LIP6, Paris, France, maria-laura.santoni@lip6.fr; Fabian Duddeck, Technical University of Munich, TUM School of Engineering & Design, Munich, Germany, duddeck@tum.de; Carola Doerr, Sorbonne Université, CNRS, LIP6, Paris, France, carola.doerr@lip6.fr; Thomas Bäck, Leiden University, LIACS, Leiden, The Netherlands, t.h.w.baeck@liacs.leidenuniv.nl; Elena Raponi, Leiden University, LIACS, Leiden, The Netherlands, e.raponi@liacs.leidenuniv.nl.

the use of gradient-based optimizers since this implies that computation of sensitivities of the objectives with respect to each design variable is rather inefficient as small step sizes must be used [9]. Furthermore, works like [29] showed that in cases of highly nonlinear systems, the estimation of sensitivities via Finite Differences are uninformative due to noise.

Due to the limitations of gradient-based methods in the aforementioned applications, gradient-free optimization algorithms have gained significant attention. Notably, Duddeck [8] demonstrated that population-based gradient-free optimizers, when coupled with simulations of frontal crash scenarios using the Explicit Finite Element Method (EFEM), can achieve promising results. Nevertheless, the adoption of these methods remains a topic of ongoing debate.

On one hand, gradient-free algorithms typically require a large number of objective function evaluations. Despite advances in computational resources, conducting numerous high-fidelity crash simulations remains computationally expensive and time-consuming. Additionally, many sampled designs may be unfeasible or exhibit poor performance [9].

On the other hand, standard gradient-free methods offer no theoretical guarantee of finding globally optimal designs, and their effectiveness depends on the problem at hand [2, 25]. This observation is consistent with the No Free Lunch (NFL) theorems for optimization, which state that no algorithm can universally outperform others across all problem classes [28]. Consequently, the selection of an appropriate black-box optimization strategy becomes a critical challenge on top of a simulation or experimental pipeline.

The performance of black-box optimizers is typically assessed through experiments conducted on standardized benchmarking suites [22]. Among the most widely used are COmparing COntinuous Optimizers (COCO) [13] and CEC Benchmarking Suites [23], which offer collections of synthetic test functions specifically designed for scalability. These benchmarks provide multiple instances by systematically shifting the location of the global optimum and modifying the optimization landscape using transformations such as rotations. A key feature of these benchmarks is their computational efficiency: objective evaluations are fast and deterministic, allowing extensive experimentation across multiple runs. This makes them a convenient and reproducible testbed for comparing black-box optimization algorithms.

However, an ongoing debate concerns whether strong performance on such synthetic benchmarks reliably translates to success on real-world expensive optimization problems. As highlighted by Bliet et al. [3], real-world problems often exhibit significantly more complex and irregular landscapes that are not well captured by simple mathematical functions, making the generalization of benchmark-based performance uncertain. In the domain of multi-objective optimization, a best-of-both-worlds approach has been pursued, where simplified and empirical models are used to design physical systems, as in the suites developed by Jain and Deb [18] and Tanabe and Ishibuchi [26]. Nevertheless, this comes at the cost of reduced precision in evaluating the objectives. Moreover, simplified models are typically valid only within a limited range of the design variables, which also hinders global applicability.

In turn, several benchmarking suites such as EXPOBench [3], SUMO [12], DAfoam [15, 16], and the CFD Optimization Suite [4] feature expensive, real-world optimization problems involving simulation tools like OpenFOAM. However, these frameworks do not include problems derived from structural mechanics.

In contrast, the MOPTA08 benchmark [19], which is directly inspired by vehicle crashworthiness optimization, is based on a surrogate model trained on a large set of high-fidelity finite element simulations. While this approach eliminates the cost of online simulations, it suffers from limited scalability, as its formulation is fixed to a 124-dimensional design space. A similar limitation exists in the Mazda CdMOBP benchmark [21]; however, in this case, all design variables are discrete. Furthermore, an issue with these aforementioned benchmarks revolves around the complexity of setting up the cases and the simulation environments. In our case, we drew inspiration from the IOHExperimenter [5] interface, ensuring that our problems can be called in a straightforward and

standardized way. This design choice not only reduces setup complexity but also makes it easy to embed the benchmark problems into other codebases by following the structure from Doerr et al. [6] for function calling.

To address the gap between structural designers seeking effective optimization tools and researchers developing black-box optimization algorithms in need of realistic test environments, we introduce MECHBench. The benchmark focuses on three representative problems designed to enable consistent and reproducible evaluation of algorithms in terms of objective quality, constraint satisfaction, convergence behavior, and computational efficiency. Each of those can generate a broad family of related optimization tasks. Every simulation produces multiple physical outputs that can be flexibly combined to define diverse objectives and constraints.

In summary, our **main contributions** are:

- We introduce a suite of optimization problems motivated by structural mechanics, designed to capture key challenges such as high dimensionality, nonlinearity, and symmetries.
- To support rigorous and comparable evaluation, we establish standardized baselines and performance metrics
- All code and datasets are released openly to foster transparency and reproducibility within the optimization, machine learning and the engineering design communities. The benchmarks are publicly available at our GitHub Repository (<https://github.com/BayesOptApp/MECHBench>).

2 Implementation

To enable consistent evaluation of optimization algorithms in structural design, we introduce a set of mechanical benchmarks based on simplified crash simulations. Each benchmark, represented by a problem instance within a standardized pipeline, is designed to decouple the simulation process from the optimization logic. This modular structure allows integration with a wide range of optimizers, regardless of the programming interface, including non-Python environments. This approach contrasts with earlier works by Hunkeler et al. [17] and Volz et al. [27], in which the optimization loop was tightly integrated into a commercial software, thereby limiting flexibility and external interfacing.

2.1 Benchmark Workflow

The benchmarking process is illustrated in Figure 1 and consists of the following key stages:

- **DESIGN VARIABLE UPDATE:** The optimization algorithm proposes a new configuration by modifying a set of d design variables (DVs). All the cases analyzed in these benchmarks correspond to shape optimization; thus, the design variables only affect the geometry of the underlying structure.
- **MODEL ASSEMBLY:** The updated input data is used to assemble a new finite element (FE) model. This includes integrating the DV changes into the model structure and preparing it for simulation.
- **SIMULATION EXECUTION:** The assembled FE model is simulated using OpenRadioss [1], which is an open-source simulation solver based on EFEM formulation. This step generates detailed time-history and summary results characterizing the crash response.
- **POST-PROCESSING AND OBJECTIVE EVALUATION:** Simulation output files, particularly the time histories in *.csv format, are automatically parsed to extract relevant variables to compute the objectives. Furthermore, for visual inspection and further post-processing, the user can set an option to generate *.vtk files for different simulation times.

2.2 System Design Principles

The benchmark system is designed with the following principles:

- **MODULARITY:** Each stage of the workflow is modularized to allow for easy substitution or extension (e.g., changing the FE solver or the objective functions).

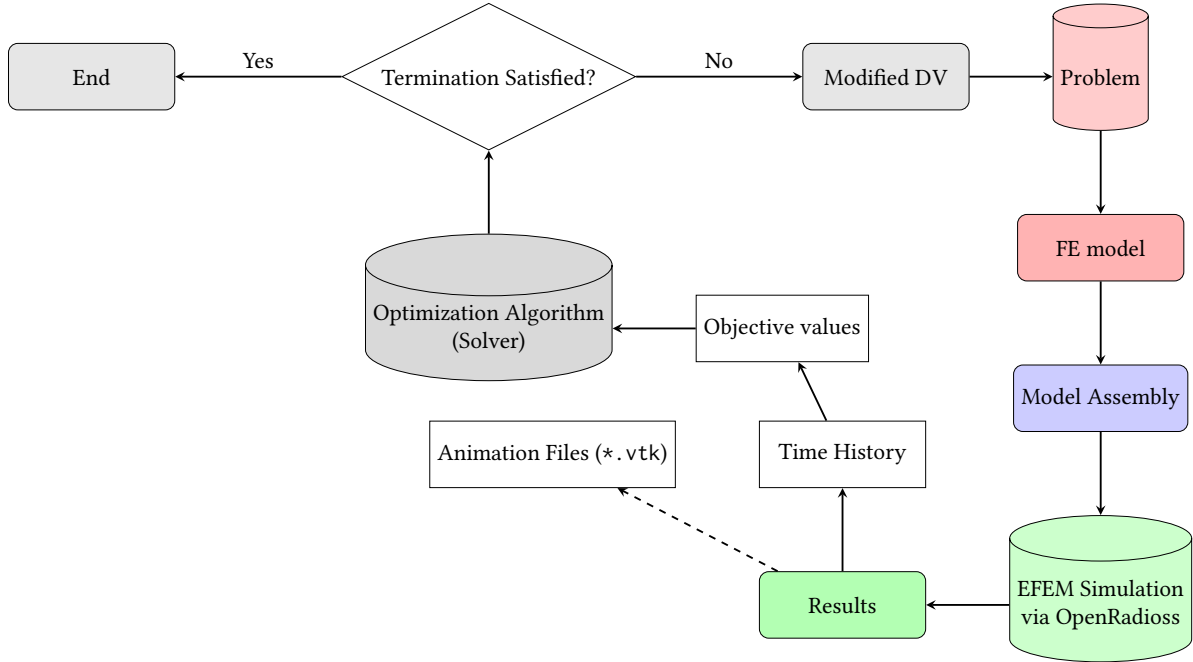


Fig. 1. Proposed optimization loop. To initiate the loop, the optimization algorithm, objective values, and problem number must be defined *a priori*. The blocks connected with solid arrows represent the necessary steps of the loop, whereas the dashed arrows indicate optional steps or processes that do not interrupt the loop. Adapted from [17, 27].

- **SEPARATION OF CONCERNS:** The optimization algorithm and the crash simulation pipeline operate independently and communicate through a clear interface. This facilitates the integration of new optimizers without modifying the core benchmark logic.
- **OPEN-SOURCE:** All steps in the pipeline, except for the optimizer, rely on open-source packages (such as NumPy [14] and Gmsh [11]) and produce outputs in open file formats, eliminating the need for proprietary software.
- **EXECUTION SCALABILITY:** The framework supports execution on local machines or high-performance computing (HPC) clusters.

2.3 Usage

Although Section 3 presents the physical bounds of the design variables, these bounds are not directly imposed in the optimizer. Instead, each design variable is mapped to the normalized domain $[-5, 5]^d$, consistent with the COCO benchmarks.

The interface is mainly built in Python and it works in a sequential fashion, as each problem object processes design variables vectors one at a time. This is because of the arrangement of files that OpenRadioss generates such that these shall be stored in a distinctive directory. Detailed instructions are available in our GitHub repository. For completeness, the usage and setup are outlined below in three steps:

- (1) **CLONE THE REPOSITORY:** The structure of the repository must follow the one shown in Figure 2.
- (2) **DOWNLOAD OR BUILD THE OPENRADIOSS BINARIES:** Precompiled OpenRadioss binaries are available for both Windows and Linux systems from the official website. Alternatively, you can build the binaries

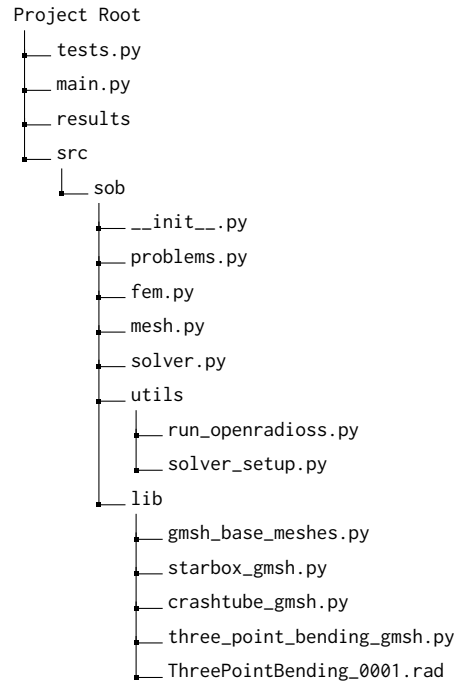


Fig. 2. MECHBench directory structure. Currently, the project contains 3 levels; Level 1 is where the `main.py` should go. On the same level, the repository has the folder `results` which contain sample simulation-generated files. The main source code is located under `src/`, where the `sob/` subpackage holds the core functionality. This includes modules for problem definitions (`problems.py`), finite element routines (`fem.py`), mesh handling (`mesh.py`), and solvers (`solver.py`). Within the second level or `sob/`, there are two utility subdirectories. The `utils/` folder provides helper scripts such as `run_openradioss.py` for starting the subprocesses and `solver_setup.py` for preparing solver configurations. Finally, the third level or `lib/` folder contains geometry and mesh generation scripts, including `gmsh_base_meshes.py` as a foundation, and specialized generators like `starbox_gmsh.py`, `crashtube_gmsh.py`, and `three_point_bending_gmsh.py`.

from the source repository, though the resulting executables are functionally equivalent to the prebuilt versions. The repository is actively maintained by Altair, the company behind OpenRadioss, along with a dedicated community of contributors.

- (3) **SET THE CASE:** The problem object from Figure 1 needs four ingredients, namely: i) The path to the OpenRadioss binaries, ii) The problem number (identified from 1 to 3), iii) the problem dimension to evaluate, and iv) the list of objectives to extract. After this setting, the user might start testing different design variable vectors.

To maintain modularity in the benchmarks, users can either call problem-specific objective functions (based on those presented in Section 3) or define tailor-made problems by extracting energy-based crashworthiness metrics, such as those proposed by Fang et al. [9].

3 Benchmark Problems

The numerical benchmark problems introduced in this section are constructed to rigorously evaluate optimization algorithms in the context of nonlinear crashworthiness analysis. Each benchmark comprises a structural component subjected to dynamic impact loading, where the design task involves adjusting geometric parameters such as thickness distributions, trigger configurations or rib dimensions, with the aim of enhancing energy absorption capacity, reducing peak force levels, or minimizing structural mass while respecting prescribed deformation limits.

To provide practitioners with practical guidance regarding computational requirements, we evaluated the wall-clock time associated with a single OpenRadioss simulation for each benchmark problem under varying levels of parallelization. The runtime of an individual simulation is not influenced by the dimensionality of the optimization problem, as the finite element model and mesh are fixed once defined. However, it does depend on the specific design variable configuration, since the elemental time step enforced by OpenRadioss is sensitive to geometric features and local stiffness variations dictated by the chosen parameters.

Table 2 summarizes the average runtime per simulation, together with the corresponding number of CPU cores utilized (both before and after the OpenRadioss call), for all three benchmark cases. All computations were carried out on the Academic Leiden Interdisciplinary Cluster Environment (ALICE) at Leiden University, exclusively employing Intel Xeon 6126 processors operating at 2.60 GHz. Each compute node is equipped with 24 cores and 384 GB of RAM.

Table 2. Average wall-clock time (in seconds) per simulation, across 45 evaluations, for each of the three problems.

		Number of Cores			
		1	2	4	8
Problem	<i>Star Shaped Crash-Box</i>	842.5 $\pm$ 19.1	757.9 $\pm$ 16.8	481.3 $\pm$ 12.4	328.5 $\pm$ 8.1
	<i>Three Point Bending of Layered Beam</i>	234.3 $\pm$ 1.6	203.6 $\pm$ 1.3	117.4 $\pm$ 1.0	71.7 $\pm$ 0.5
	<i>Long Crash Tube</i>	1918.5 $\pm$ 16.0	1532.4 $\pm$ 12.6	858.6 $\pm$ 9.6	489.2 $\pm$ 5.3

3.1 Star Shaped Crash-Box

3.1.1 Problem Description. The objective in this case is to optimize both the thickness profile and the transverse profile of the crash box shown in Figure 3. In this configuration, the structure is clamped to the ground at its base, preventing any rotation or displacement, while an impactor moves toward the crash box from above. The goal is to maximize the specific energy absorption (SEA), subject to an intrusion constraint of 60 mm, corresponding to half the extrusion length of the crash box. This formulation is inspired by the work of Hunkeler et al. [17], with the key difference that the constraint here is defined in terms of maximum intrusion rather than peak crash force.

Put simply, the design must absorb as much crash energy as possible while preventing the structure from being compressed by more than 60 mm. The goal is to ensure efficient energy dissipation during impact while maintaining a deformation limit that preserves the required structural clearance.

3.1.2 Variable Definition. For a lower-dimensional setting, we directly retrieve the cases shown in Hunkeler et al. [17], for $d \in \{2, 3, 4, 5\}$ thus we added a $d = 1$ case by just defining a square. These settings are represented in Figure 4. For higher-dimensional settings ($6 \leq d \leq 34$), the definitions of (x_1, x_2, x_3, x_4) remain identical to those in the four-dimensional case shown in Figure 4. For indices i ranging from 6 to 34, the variables define the control points of a piecewise-linear wall thickness profile of the crash box. An illustrative example for $d = 7$, corresponding to three control points, is provided in Figure 5. For $d < 34$, the thickness of each of the 30 rows of elements is obtained by interpolating the distributions at the z -coordinate of the element barycenters. At $d = 34$,

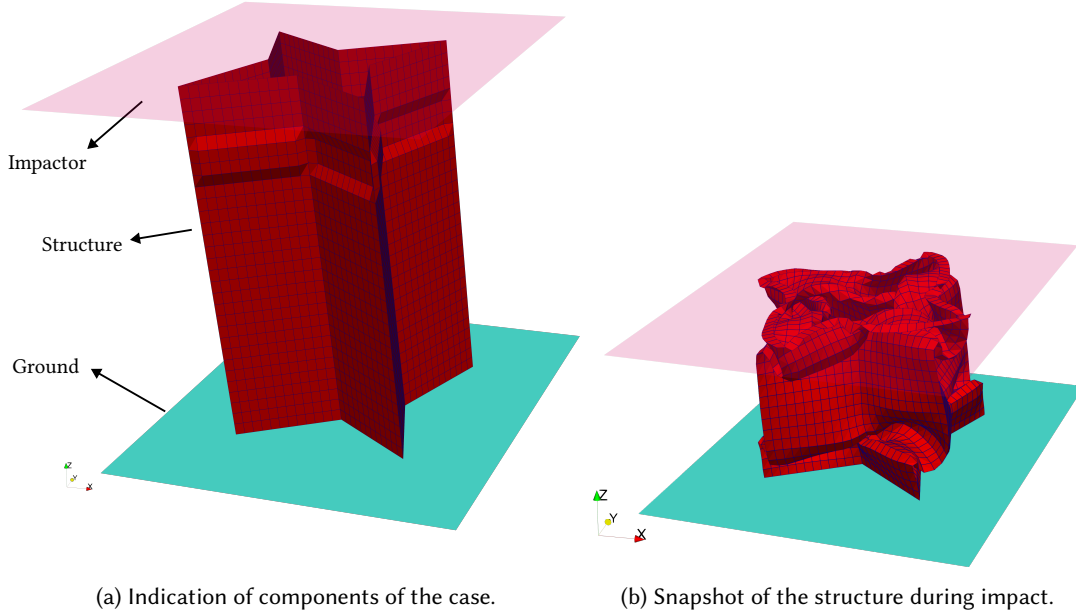


Fig. 3. Star-shaped crash-box case. The structure is subdivided into smaller connected parts, forming a finite element mesh shown by the blue lines. In the z -direction, the mesh is arranged in 30 rows of elements, wherein each row is assigned the same thickness value t_h .

each row of the finite element mesh is assigned an independent thickness value, without interpolation. Consistent with the bounds in Figure 4, all control points satisfy $0.7 \text{ mm} \leq x_i \leq 3 \text{ mm}$ for every index i between 6 and 34.

3.1.3 Standard Mathematical Formulation.

$$\begin{aligned} \max_{\mathbf{x}} \quad & \text{SEA}(\mathbf{x}) = \frac{E_{\text{abs}}(\mathbf{x})}{m_s(\mathbf{x})} \\ \text{s.t.} \quad & \delta(\mathbf{x}) \leq 60 \text{ mm.} \end{aligned} \quad (1)$$

With this objective, the structure is modified to maximize the specific energy absorption (SEA), defined as the ratio of the absorbed impact energy (E_{abs}), which is dissipated primarily through plastic deformation, to the total mass of the structure (m_s). This metric promotes designs that achieve high energy absorption while maintaining low structural weight.

3.1.4 Reformulation. Given that the problem is originally constrained, we formulated a new objective function which can be used with unconstrained optimizers. Contrary to Raponi et al. [24], who constructed a penalized objective function, we define a piecewise function instead as:

$$\min_{\mathbf{x}} \text{ Penalized SEA}(\mathbf{x}) = \begin{cases} -\text{SEA}(\mathbf{x}), & \text{if } \delta(\mathbf{x}) \leq 60 \text{ mm,} \\ 100 (\delta(\mathbf{x}) - 60), & \text{if } \delta(\mathbf{x}) > 60 \text{ mm.} \end{cases} \quad (2)$$

With this modification, we specifically aim to avoid the need for assigning a de facto penalty weight through multiple preliminary runs. In the aforementioned case, such weight may still be insufficient to ensure that the

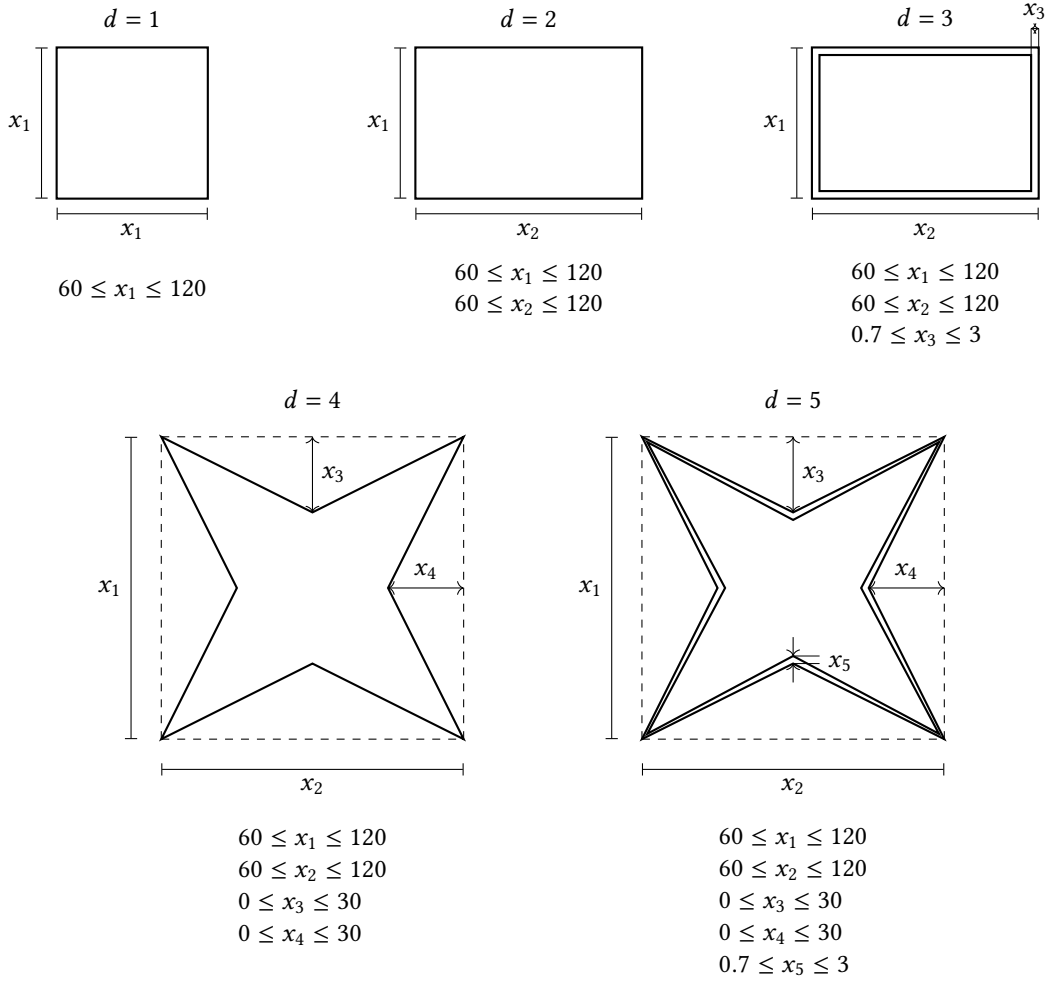


Fig. 4. Optimization problem parameterization for $d \leq 5$. The ranges of each variable are expressed in mm. The drawings are not in scale. For the cases where thickness is not controlled such as the cases where $d = \{1, 2, 4\}$, the wall thickness t_h is set constant to 2.1 mm.

maximum objective value within the feasible region exceeds that of the infeasible region, particularly when the latter corresponds to the lower bounds of wall thickness.

3.2 Three Point Bending of Layered Beam

3.2.1 Problem Description. In this case, the beam consists of five ribs, as illustrated in Figure 6a. The objective is to reduce the overall weight of the structure by adjusting the thickness of each rib. During testing, the beam is fixed firmly at both ends (clamped supports), as shown in Figure 6b, so that it cannot move or rotate at the boundaries. This setup allows the ribs to deform under load in a controlled way, while the optimization seeks the lightest design that can still withstand the applied forces.

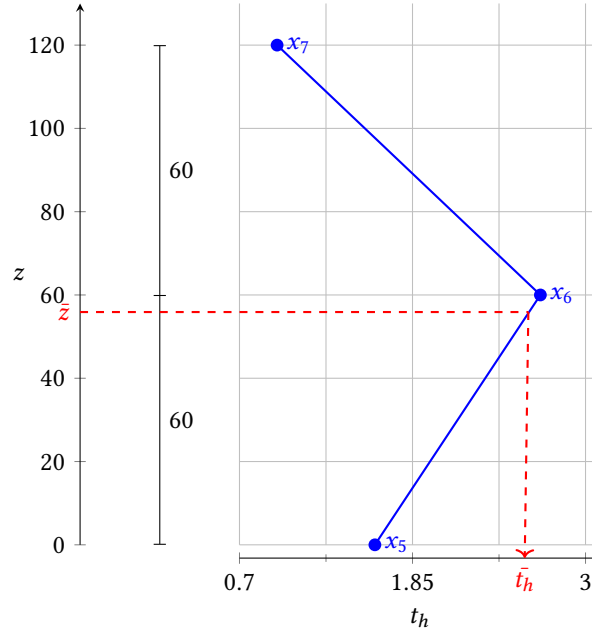


Fig. 5. Thickness distribution example for $d = 7$, wherein 3 control points are set at uniformly distributed heights between 0 and 120 mm and each point has an assigned $t_h(z)$. The wall thickness t_h at any intermediate point $\bar{z}$ is computed via linear interpolation of assigned t_h of the neighboring control points. This procedure is used to compute the wall thickness t_h of a row of finite elements by getting the interpolated $\bar{t}_h$ at the element barycenter.

A helpful analogy is to think of the ribs like the wooden slats of a bookshelf: making them thinner reduces the weight of the shelf, but if they are too thin, the shelf will bend or even break under load. The goal is therefore to find the right balance between lightness and strength.

We drew inspiration from the case presented by Kaps et al. [20], adopting the same mesh sizing and material properties described therein. Notably, Kaps et al. [20] report conducting their experiments with two mesh resolutions: a coarse mesh used for the *low-fidelity model* and a fine mesh used for the *high-fidelity model*. The reference mesh size defined in this work corresponds to the coarse mesh. This choice is motivated by the need to keep the function evaluation time as short as possible.

3.2.2 Variable Definition. Like in the previous case, we make a distinction between the *lower-dimensional* and the *higher-dimensional* setting because for the *lower-dimensional* case, the thickness variations are coupled for several ribs and others are kept to a fixed thickness. The aforementioned description is represented in Figure 7 in the cases where $d \leq 5$, the design variables account for a uniform wall thickness of the ribs.

For $d > 5$, wall thickness profiles are generated for each rib by incrementally by adding control points. An example is provided in Figure 8 for $6 \leq d \leq 10$. The procedure is extended to $d = 40$, such that each row of elements in the z -direction of a rib is assigned an independent t_h value. As in the star shaped crash-box case, thickness values are interpolated with respect to the z -coordinate of the element barycenters, with eight rows of elements defined along this direction.

Each x_i variable of this case has the same bounds, namely: $0.5 \text{ mm} \leq x_i \leq 3.0 \text{ mm}$.

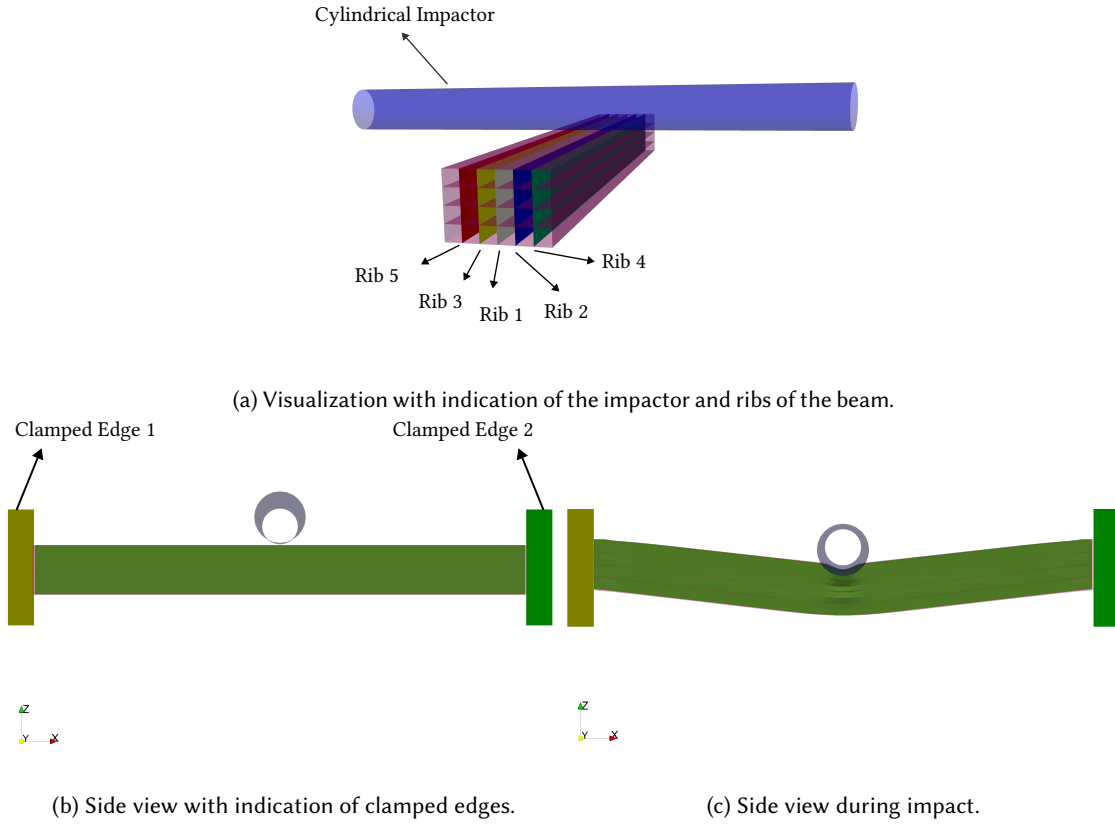


Fig. 6. Representation of the Three Point Bending of Layered Beam case.

3.2.3 Standard Mathematical Formulation. The problem is originally formulated as a constrained optimization task:

$$\begin{aligned} \min_{\mathbf{x}} \quad & m_s(\mathbf{x}) \\ \text{s.t.} \quad & \delta(\mathbf{x}) \leq 50 \text{ mm.} \end{aligned} \quad (3)$$

As in the previous case, the constraint $\delta(\mathbf{x}) \leq 50$ mm imposes an intrusion limit at the point of impact. This requirement promotes an efficient mass distribution through the rib thicknesses to withstand crash loading. In crashworthiness terms, the intrusion constraint is defined to ensure that the beam neither exceeds permissible deformation levels nor fractures under the specified impact conditions, thereby satisfying both structural integrity and occupant safety requirements.

3.2.4 Reformulation. Similar to the star shaped crash-box benchmark, we do not employ a penalized objective function as in Kaps et al. [20]. Instead, the objective was defined as a piecewise function based on feasibility, determined by the measured maximum intrusion after impact. This is expressed as:

$$\min_{\mathbf{x}} \text{Penalized } m_s(\mathbf{x}) = \begin{cases} m_s(\mathbf{x}), & \text{if } \delta(\mathbf{x}) \leq 50 \text{ mm,} \\ 4.25952 + 10 \left(\frac{\delta}{50} - 1 \right), & \text{if } \delta(\mathbf{x}) > 50 \text{ mm.} \end{cases} \quad (4)$$

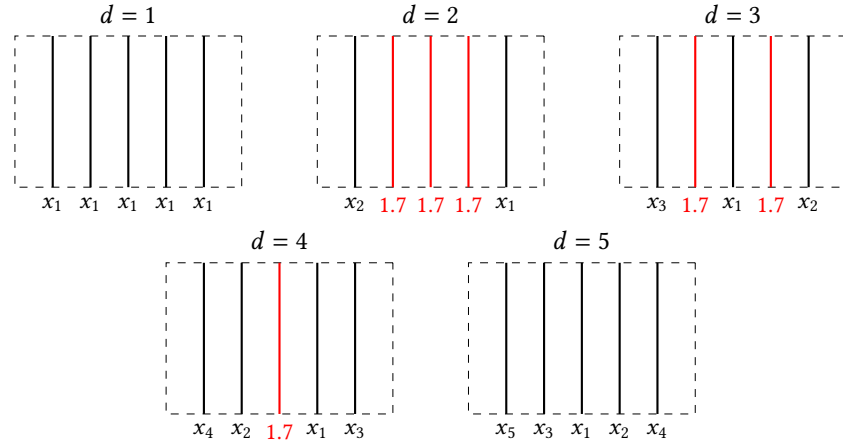


Fig. 7. Optimization problem parameterization for $d \leq 5$. Each variable is expressed in mm. In **red** are the ribs which have a predefined thickness of 1.7 mm. Each line is annotated by the design variable denoting the wall thickness of the respective rib.

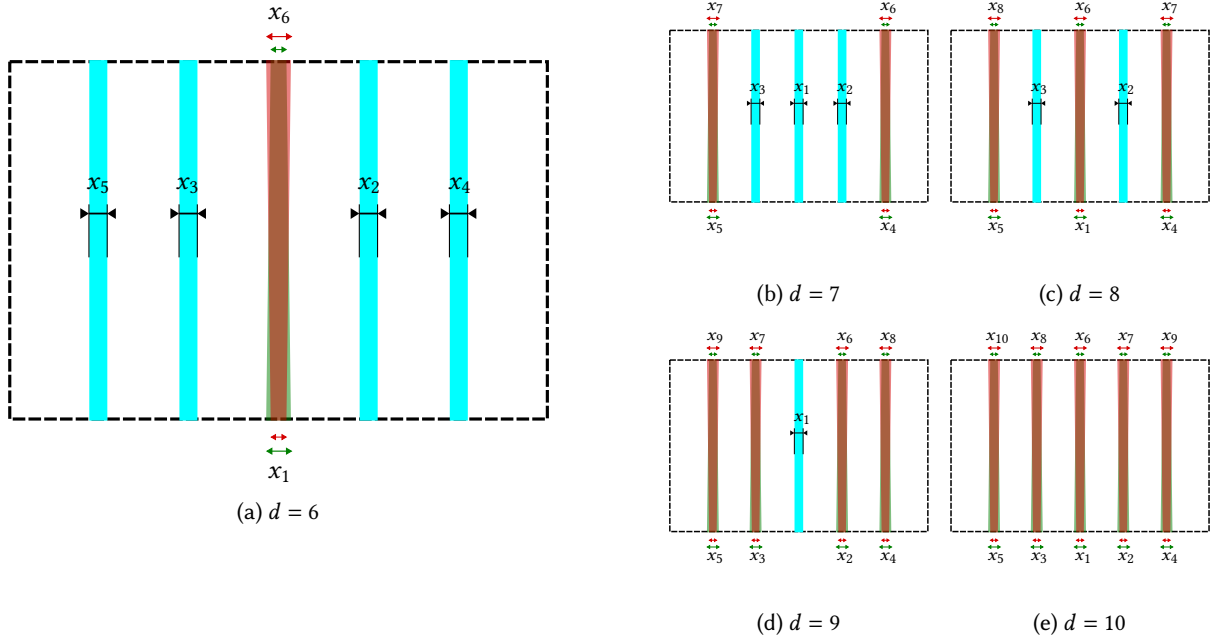


Fig. 8. Description of point addition used to develop wall thickness profiles for each of the five ribs of the beam for $d = 6, 7, 8, 9, 10$. For the dimensions shown, the straight rectangles colored in **cyan** indicate a constant thickness along the corresponding rib. In contrast, the two sets of trapezoids, colored in **semi-transparent red** and **semi-transparent green**, represent linearly increasing and decreasing thickness profiles, respectively. The superposition of these trapezoids is used to emphasize that the thickness profile of a rib may be expressed as either increasing or decreasing, depending on the variable values. For $10 < d \leq 40$, control points are added in a similar fashion as the shown sequences herein.

3.3 Long Crash Tube

3.3.1 Problem Description. The primary objective in this study is to mitigate peak force amplification by optimizing the shape and placement of triggers in a long crash box. Modifying the trigger characteristics alters the deformation behavior of the structure, enabling better energy absorption and more controlled plastic collapse modes.

We took inspiration from the case analyzed in Kaps et al. [20], but we add additional modifications to enable scalability. Yet the definition of the case is mostly the same. In Figure 9, there's a representation of the objects of the case, where the impactor is represented as a solid moving towards the structure, which is clamped to the ground.

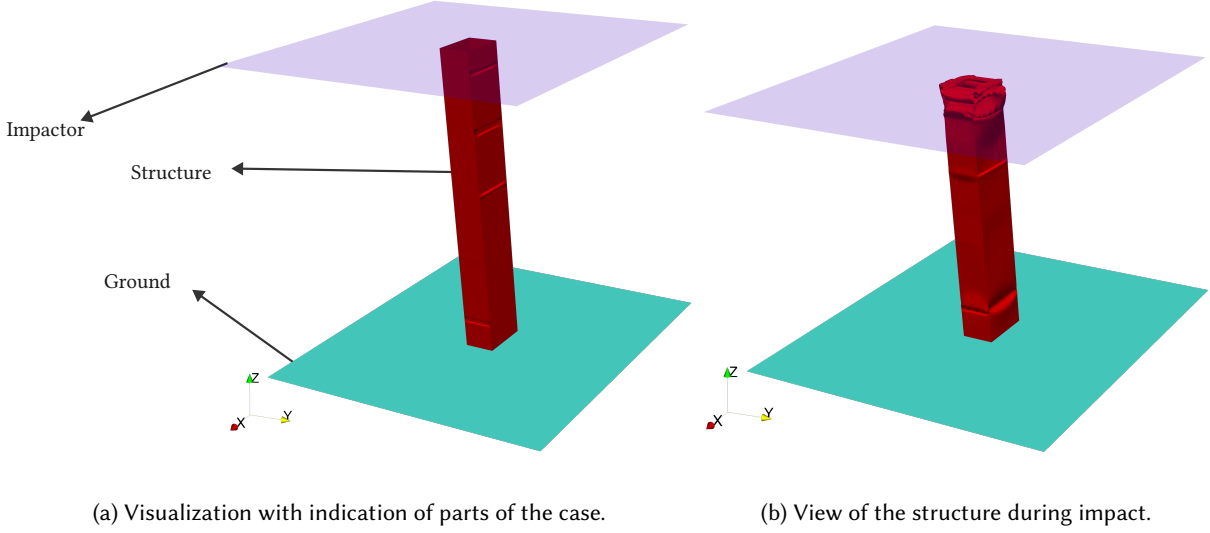


Fig. 9. Representation of the Long Crash Tube Case.

3.3.2 Variable Definition. To provide a clearer understanding of the design (optimization) variables involved in this study, we refer to the labeling scheme illustrated in Figure 10. The cross-sectional geometry of the structure is rectangular, consisting of two pairs of opposing faces. We define Face A as the pair corresponding to the longer edges of the rectangle, and Face B as the pair corresponding to the shorter edges. A total of ten triggers are distributed along these faces: five on Face A, indexed from 1 to 5, and five on Face B, indexed from 6 to 10.

Each trigger is characterized by a triplet (z_i, ϵ_i, h_i) , where z_i denotes the axial (along the z-axis) position of the trigger's centroid relative to a fixed reference point, ϵ_i represents the trigger's protrusion (depth) from the face surface, and h_i corresponds to its vertical extent (height). The representation of the variables on the structure are shown in Figures 11 and 12. These parameters serve as the primary design variables in the optimization process. The bounds of the triplet are the following:

$$\begin{aligned} -40 \text{ mm} &\leq z_i \leq 40 \text{ mm}, \\ -4 \text{ mm} &\leq \epsilon_i \leq 4 \text{ mm}, \\ 0 \text{ mm} &\leq h_i \leq 16 \text{ mm}. \end{aligned}$$

Finally, we developed a table of equivalences, wherein it is stipulated which design variable x_i controls the

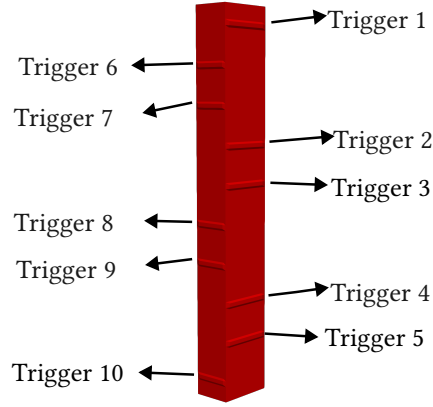


Fig. 10. Labeling of triggers of the Long Crash Tube Case.

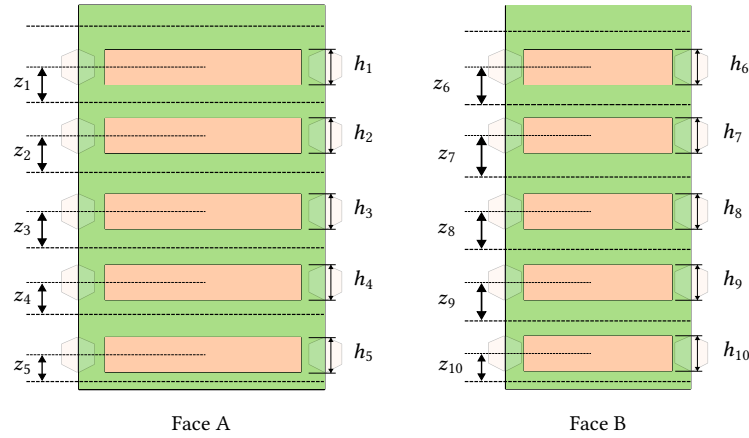


Fig. 11. Representation of the pairs (z_i, h_i) , where z_i denotes the position along the z -axis of the barycenter of a trigger, and h_i is the corresponding height. In this representation, the triggers are shown in *light peach*, while the rest of the structure is shown in *light green*. The z -position is measured from a reference point. The thick dashed lines indicate the possible limits of each z_i , whereas the thin dashed lines mark the z -position of the barycenter of a trigger. The drawing is not in scale. The triggers located at complementary faces are shown transparent.

triplets $(z_i, \varepsilon_i, h_i)$ for different dimensions. We split such a table of equivalence into two parts, which are shown in Tables 6 and 7. This splitting is worthy since for $d \leq 15$, the design variables control the triggers of the two faces in a mirrored way, whereas for $15 < d \leq 30$, the triggers of both faces are controlled independently.

3.3.3 Standard Mathematical Formulation. This problem is treated as unconstrained and is formulated as:

$$\min_{\mathbf{x}} \text{LU}(\mathbf{x}) = \left| \frac{F_{\text{peak}}}{F_{\text{mean}}} \right|, \quad (5)$$

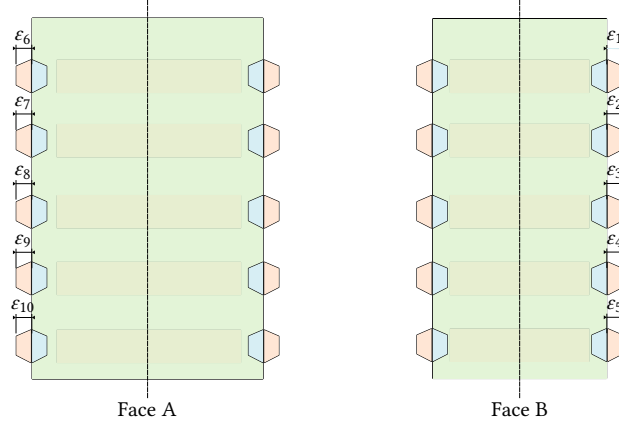


Fig. 12. Representation of each ε_i on the structure. The reference point for each ε_i corresponds to a plane attached to either Face A or Face B. In this figure, each ε_i is measured is visualized as an outward extrusion, shown in *pastel peach*. When ε_i is negative, the measure is instead represented as an inward extrusion, shown in *sky blue*. The dashed line drawn at the middle of the structure corresponds to a symmetry axis. This means that the same ε_i value is imposed to the complementary face. The drawing is not in scale. Similarly, the triggers of frontal face are shown transparent to point that from such views the triggers are present, yet not related to show the meaning of each ε_i .

where LU denotes the load uniformity, F_{peak} is the maximum force recorded during impact, and F_{mean} is the mean force over the loading duration.

Minimizing load uniformity in this context reduces the amplification of peak forces relative to the average crash load. High peak forces are undesirable, as they can cause sudden stress concentrations and localized failures, whereas smoother force distributions enable the structure to dissipate energy more steadily throughout the deformation process.

4 Conclusions and future work

We presented a benchmarking suite for mechanical design and crashworthiness optimization, designed to foster reproducibility, comparability, and scalability in black-box optimization applied to engineering simulations. The benchmarks are modular and moderately lightweight, supporting seamless integration with external optimizers through standardized interfaces, thereby minimizing setup overhead.

At present, the suite focuses on single-objective problems involving shell-based finite element models. These cover a relevant subset of shape optimization tasks but do not yet capture the full complexity of many practical design scenarios. To move towards a more general and representative evaluation platform, we plan to extend the framework along two axes.

First, we will implement problem instances by applying transformations in a manner similar to the noiseless black-box optimization benchmarks (BBOB) described in [10]. We believe this addition will make the suite particularly valuable to the community, as other real-world-inspired benchmarking suites rarely account for such variations. This will enable a more robust and fair evaluation of algorithms, since no universal shortcut will exist for tailoring an optimizer to a single fixed problem.

Second, we acknowledge that a major limitation of the current setup is the high simulation time displayed in Table 2, which can restrict the scale of experimental evaluations. To address this, we plan to incorporate surrogate models and aliasing strategies into the benchmarking repository, enabling faster and more extensive testing. In

addition, we will embed these surrogates within multi-fidelity evaluations to reflect different levels of model precision and better approximate real engineering design processes.

Finally, beyond the optimization interface itself, we will broaden the benchmark set by integrating more representative components and assemblies, and by incorporating datasets from public initiatives and industrial collaborations.

Acknowledgments

We thank the Chair of Computational Solid Mechanics (CSM) at the Technical University of Munich, and in particular Arne Kaps and Paolo Ascia, for their many insightful discussions. We also thank Fei Fan Li for her contributions to the initial development of the benchmarks.

References

- [1] Altair and Contributors. 2025. *OpenRadioss*. <https://github.com/OpenRadioss/OpenRadioss> GitHub commit 6c16e4b.
- [2] Charles Audet and Warren Hare. 2017. *Derivative-Free and Blackbox Optimization*. Springer International Publishing, Cham. doi:10.1007/978-3-319-68913-5
- [3] Laurens Blik, Arthur Guijt, Rickard Karlsson, Sicco Verwer, and Mathijs De Weerd. 2023. Benchmarking surrogate-based optimisation algorithms on expensive black-box functions. *Applied Soft Computing* 147 (Nov. 2023), 110744. doi:10.1016/j.asoc.2023.110744
- [4] Steven J. Daniels, Alma A. M. Rahat, Richard M. Everson, Gavin R. Tabor, and Jonathan E. Fieldsend. 2018. A Suite of Computationally Expensive Shape Optimisation Problems Using Computational Fluid Dynamics. In *Parallel Problem Solving from Nature – PPSN XV*, Anne Auger, Carlos M. Fonseca, Nuno Lourenço, Penousal Machado, Luís Paquete, and Darrell Whitley (Eds.). Vol. 11102. Springer International Publishing, Cham, 296–307. doi:10.1007/978-3-319-99259-4_24 Series Title: Lecture Notes in Computer Science.
- [5] Jacob de Nobel, Furong Ye, Diederick Vermetten, Hao Wang, Carola Doerr, and Thomas Bäck. 2021. IOHexperimenter: Benchmarking Platform for Iterative Optimization Heuristics. *arXiv e-prints:2111.04077* (Nov. 2021). arXiv:2111.04077 <https://arxiv.org/abs/2111.04077>
- [6] Carola Doerr, Hao Wang, Furong Ye, Sander van Rijn, and Thomas Bäck. 2018. IOHprofiler: A Benchmarking and Profiling Tool for Iterative Optimization Heuristics. doi:10.48550/arXiv.1810.05281 arXiv:1810.05281 [cs].
- [7] Paul DuBois, Clifford Chou, Bahig Fileta, Tawfik Khalil, Albert King, Hikmat Mahmood, Harold Mertz, Jac Wismans, Priya Prasad, and Jamel Belwafa. 2004. *Vehicle Crashworthiness and Occupant Protection*. American Iron and Steel Institute, Southfield, Michigan, United States of America.
- [8] Fabian Duddeck. 2008. Multidisciplinary optimization of car bodies. *Structural and Multidisciplinary Optimization* 35, 4 (April 2008), 375–389. doi:10.1007/s00158-007-0130-6
- [9] Jianguang Fang, Guangyong Sun, Na Qiu, Nam H. Kim, and Qing Li. 2017. On design optimization for structural crashworthiness and its state of the art. *Structural and Multidisciplinary Optimization* 55, 3 (March 2017), 1091–1119. doi:10.1007/s00158-016-1579-y
- [10] Steffen Finck, Nikolaus Hansen, Raymond Ros, and Anne Auger. 2009. *Real-Parameter Black-Box Optimization Benchmarking 2009: Noiseless Functions Definitions*. Technical Report RR-6829. INRIA. <https://inria.hal.science/inria-00362633v2/document> Updated version as of February 2019.
- [11] Christophe Geuzaine and Jean-François Remacle. 2009. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *Internat. J. Numer. Methods Engrg.* 79, 11 (Sept. 2009), 1309–1331. doi:10.1002/nme.2579
- [12] Dirk Gorissen, Ivo Couckuyt, Piet Demeester, Tom Dhaene, and Karel Crombecq. 2010. A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design. *J. Mach. Learn. Res.* 11 (2010), 2051–2055. ISBN: 1532-4435 Publisher: JMLR.org.
- [13] Nikolaus Hansen, Anne Auger, Raymond Ros, Olaf Mersmann, Tea Tušar, and Dimo Brockhoff. 2021. COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting. *Optimization Methods and Software* 36, 1 (Jan. 2021), 114–144. doi:10.1080/10556788.2020.1808977 arXiv:1603.08785 [cs].
- [14] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. Array programming with NumPy. *Nature* 585, 7825 (Sept. 2020), 357–362. doi:10.1038/s41586-020-2649-2
- [15] Ping He, Charles A. Mader, Joaquim R.R.A. Martins, and Kevin J. Maki. 2018. An aerodynamic design optimization framework using a discrete adjoint approach with OpenFOAM. *Computers & Fluids* 168 (May 2018), 285–303. doi:10.1016/j.compfluid.2018.04.012
- [16] Ping He, Charles A. Mader, Joaquim R. R. A. Martins, and Kevin J. Maki. 2020. DAfoam: An Open-Source Adjoint Framework for Multidisciplinary Design Optimization with OpenFOAM. *AIAA Journal* 58, 3 (March 2020), 1304–1319. doi:10.2514/1.J058853

- [17] Stephan Hunkeler, Fabian Duddeck, Milan Rayamajhi, and Hans Zimmer. 2013. Shape optimisation for crashworthiness followed by a robustness analysis with respect to shape variables: Example of a front rail. *Structural and Multidisciplinary Optimization* 48, 2 (Aug. 2013), 367–378. doi:10.1007/s00158-013-0903-z
- [18] Himanshu Jain and Kalyanmoy Deb. 2014. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point Based Nondominated Sorting Approach, Part II: Handling Constraints and Extending to an Adaptive Approach. *IEEE Transactions on Evolutionary Computation* 18, 4 (Aug. 2014), 602–622. doi:10.1109/TEVC.2013.2281534
- [19] Donald R. Jones. 2008. *Large-Scale Multi-Disciplinary Mass Optimization in the Auto Industry*. Technical Report. MOPTA 2008, Detroit, MI, USA.
- [20] Arne Kaps, Catharina Czech, and Fabian Duddeck. 2022. A hierarchical kriging approach for multi-fidelity optimization of automotive crashworthiness problems. *Structural and Multidisciplinary Optimization* 65, 4 (April 2022), 114. doi:10.1007/s00158-022-03211-2
- [21] Takehisa Kohira, Hiromasa Kemmotsu, Oyama Akira, and Tomoaki Tatsukawa. 2018. Proposal of benchmark problem based on real-world car structure design optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, Kyoto Japan, 183–184. doi:10.1145/3205651.3205702
- [22] Jakub Kudela. 2022. A critical problem in benchmarking and analysis of evolutionary computation methods. *Nature Machine Intelligence* 4, 12 (Dec. 2022), 1238–1245. doi:10.1038/s42256-022-00579-0
- [23] Z. Michalewicz, K. Deb, M. Schmidt, and T. Stidsen. 2000. Test-case generator for nonlinear continuous parameter optimization techniques. *Trans. Evol. Comp* 4, 3 (Sept. 2000), 197–215. doi:10.1109/4235.873232
- [24] Elena Raponi, Ivan Olarte Rodriguez, and Niki van Stein. 2025. Global Sensitivity Analysis Is Not Always Beneficial for Evolutionary Computation: A Study in Engineering Design. In *Explainable AI for Evolutionary Computation*, Niki van Stein and Anna V. Kononova (Eds.). Springer Nature Singapore, 13–40. doi:10.1007/978-981-96-2540-6_2
- [25] Luis Miguel Rios and Nikolaos V. Sahinidis. 2013. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization* 56, 3 (July 2013), 1247–1293. doi:10.1007/s10898-012-9951-y
- [26] Ryoji Tanabe and Hisao Ishibuchi. 2020. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing* 89 (April 2020), 106078. doi:10.1016/j.asoc.2020.106078
- [27] Karl-Heinz Volz, Bernhard Frodl, Ferdinand Dirschmid, Rafael Stryczek, and Hans Zimmer. 2007. Optimizing topology and shape for crashworthiness in vehicle product development. In *International Automotive Body Congress (IABC)*. Berlin, Germany.
- [28] D.H. Wolpert and W.G. Macready. 1997. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1, 1 (April 1997), 67–82. doi:10.1109/4235.585893
- [29] Nicholas Zabarar, Shankar Ganapathysubramanian, and Qing Li. 2003. A continuum sensitivity method for the design of multi-stage metal forming processes. *International Journal of Mechanical Sciences* 45, 2 (Feb. 2003), 325–358. doi:10.1016/S0020-7403(03)00048-1

A Material Properties

In Table 3 and 4, the material properties for each of the problems is shown. We tried to use the same formulations from the respective sources of the problems and adapted the original formulations into OpenRadioss framework.

Table 3. Material properties and strain rate parameters for Problems 1 and 3 (Star Box and Long Crash Tube). The same parameters are listed in Kaps et al. [20].

Parameter	Symbol	Value
Young's modulus	E	200 GPa
Poisson's ratio	ν	0.3
Mass density	ρ	7830 $\frac{\text{kg}}{\text{m}^3}$
Yield strength	σ_y	360 MPa
Strainrate model		Cowper-Symmonds
Strainrate parameter C	C	40
Strainrate parameter p	p	5
Plasticity		Linear Piecewise (see Figure 13a)

Table 4. Material properties and strain rate parameters for problem 2 (Three Point Bending of Layered Beam). The same parameters are listed in Kaps et al. [20].

Parameter	Symbol	Value
Young's modulus	E	70 GPa
Poisson's ratio	ν	0.33
Mass density	ρ	2700 $\frac{\text{kg}}{\text{m}^3}$
Yield strength	σ_y	180 MPa
Plasticity		Linear Piecewise (see Figure 13b)

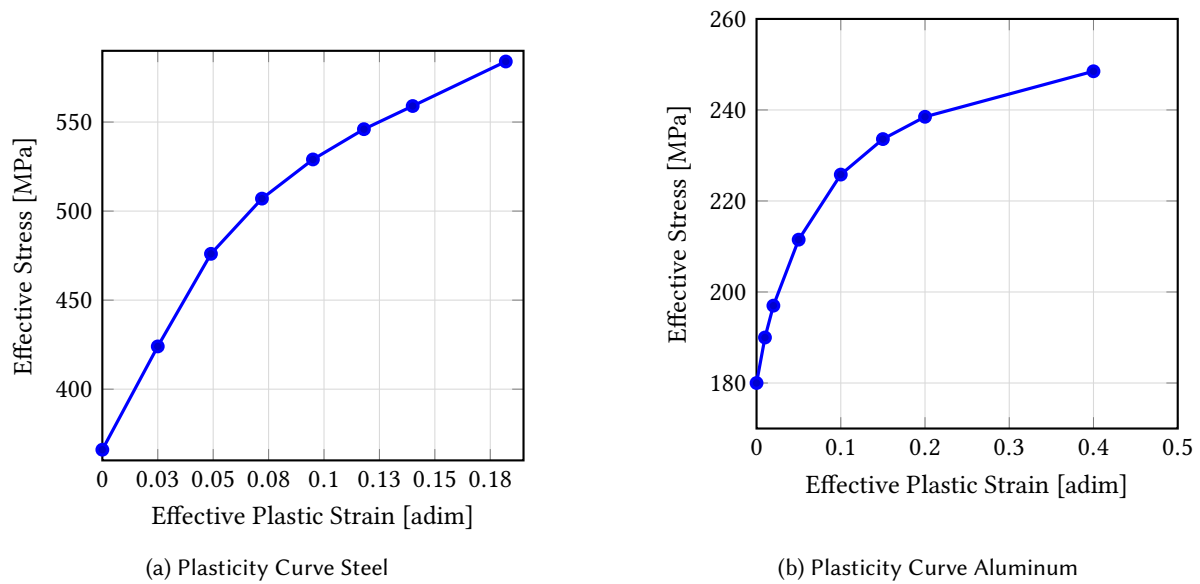


Fig. 13. Plasticity Curves used for the Material Models. Adapted from Kaps et al. [20].

B Simulation Setup

In Table 5 we show the settings of the simulations for the impactor, the volume of the structure as well as the simulation time. As shown there, we kept the same shell element formulation and the same contact formulation (the surface to surface contact) for all the problems. In problem 1, since the volume of the structure changes, we show a range instead.

C Equivalence Matrices to define variables for the Long Crash Tube case

The equivalences are shown in the tables 6 and 7.

[illegible]

[illegible]

Table 6. Equivalence Matrix of trigger characterization with each design variable for $1 \leq d \leq 15$. Units in mm.

[illegible]

z_5	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}	x_{13}
ε_5	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}	x_{14}
h_5	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}	x_{15}
z_6	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}	x_{16}
ε_6	0	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}	x_{17}
h_6	8	8	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}	x_{18}
z_7	0	0	0	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}	x_{19}
ε_7	0	0	0	0	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}	x_{20}
h_7	8	8	8	8	8	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}	x_{21}
z_8	0	0	0	0	0	0	x_{22}	x_{22}	x_{22}	x_{22}	x_{22}	x_{22}	x_{22}	x_{22}	x_{22}
ε_8	0	0	0	0	0	0	0	x_{23}	x_{23}	x_{23}	x_{23}	x_{23}	x_{23}	x_{23}	x_{23}
h_8	8	8	8	8	8	8	8	8	x_{24}	x_{24}	x_{24}	x_{24}	x_{24}	x_{24}	x_{24}
z_9	0	0	0	0	0	0	0	0	0	x_{25}	x_{25}	x_{25}	x_{25}	x_{25}	x_{25}
ε_9	0	0	0	0	0	0	0	0	0	0	x_{26}	x_{26}	x_{26}	x_{26}	x_{26}
h_9	8	8	8	8	8	8	8	8	8	8	8	x_{27}	x_{27}	x_{27}	x_{27}
z_{10}	0	0	0	0	0	0	0	0	0	0	0	0	x_{28}	x_{28}	x_{28}
ε_{10}	0	0	0	0	0	0	0	0	0	0	0	0	0	x_{29}	x_{29}
h_{10}	8	8	8	8	8	8	8	8	8	8	8	8	8	8	x_{30}

Table 7. Equivalence Matrix of trigger characterization with each design variable for $16 \leq d \leq 30$. Units in mm.

Table 5. Settings for the simulation for each of the problems.

Parameter	Problem 1	Problem 2	Problem 3
Element Formulation	Belytschko-Lin-Tsay		
Contact Formulation	/INTER/TYPE24 from Radioss		
Height [mm]	120	120	800
Width [mm]	60-120	800	120
Depth [mm]	60-120	80	80
$v_{I,0}$ [km/hr]	25.2	36	30
m_I [kg]	250	86	300
Impactor Radius [mm]		36	
Simulation Time [ms]	45	40	45