**Structural Multidisciplinary Optimization manuscript No.**
(will be inserted by the editor)

# A bi-level scalable methodology for mixed categorical-continuous structural optimization problems

**Pierre-Jean Barjhoux · Youssef Diouane · Stéphane Grihon · Dimitri Bettebghor · Joseph Morlier**

Received: date / Accepted: date

**Abstract** In this paper, large scale structural optimization problems involving both non-ordinal categorical and continuous design variables are investigated. The aim is to minimize the weight of a truss structure with respect to the cross-section areas, with optimal materials and cross-section type selection. The targeted structure counts more than one hundred elements. The proposed methodology consists of using a bi-level decomposition involving two problems, named master and slave. For given categorical choices, the slave addresses the continuous variables of our optimization problem. The master consists of minimizing a first order like approximation of the slave problem with respect to the categorical design variables. Such approximation helps to drastically reduce the combinatorial raised by the categorical variables. The proposed heuristic algorithm is tested on three different structural optimization test cases. The comparison to state-of-the-art algorithms emphasizes the efficiency of the proposed algorithm in terms of the optimum quality, computation cost, as well as its scalability with respect to the problem dimension. A result of a 120-bar truss mixed-categorical optimization problem instance solved by the proposed heuristics is discussed.

Pierre-Jean Barjhoux
Institute of Technology IRT Saint Exupéry,
Institut Clément Ader (ICA)
Toulouse, France
E-mail: pierre-jean.barjhoux@irt-saintexupery.com

Youssef Diouane
ISAE SUPAERO, Université de Toulouse
Toulouse, France
E-mail: youssef.diouane@isae-supaero.fr

Stéphane Grihon
Airbus Operations SAS
Toulouse, France
E-mail: stephane.grihon@airbus.com

Dimitri Bettebghor
ONERA – The French Aerospace Lab
Châtillon, France
E-mail: dimitri.bettebghor@onera.fr

Joseph Morlier
Université de Toulouse,
Institut Clément Ader (ICA),
CNRS-ISAE SUPAERO-INSA-Mines Albi-UPS
Toulouse, France
E-mail: joseph.morlier@isae-supaero.fr

## 1 Introduction

In the field of structural design, weight minimization of the structure is a major concern for engineers. In the aircraft industry for example, structural optimization problems can combine changes in choices of materials, cross-section types, or sizes of elements based on manufacturer catalogs (Grihon, 2012, 2018). As a consequence, the number of design variables grows significantly and prevents practical resolution of the associated optimization problems.

In this article, we aim to solve large scale structural weight minimization problems with both categorical and continuous variables, subject to stress and displacements constraints. The topology of the structure is fixed. The categorical variables take values belonging to an unordered set. Typically, in the context of structural optimization, the choices of materials or cross-section types are depicted by categorical variables. Most existing algorithms used to handle such classes of problems (Fister et al., 2013) are known to scale badly as the number of the categorical design variables increases. To illustrate the curse of dimensionality encountered for such problems, an order of magnitude of the targeted industrial problem's dimensions can be as follows. Consider an engine pylon structural model as addressed in (Gazaix et al., 2019) with 100 elements, each one having a hundred possible choices of material and stiffening principles. In this case, the categorical design space counts $100^{100}$ possible configurations to describe the whole structure. Thus, the high combinatorial dimension of the categorical design space enforces the need for a methodology to solve such problems efficiently.

In general, to handle mixed categorical-continuous optimization problems, many optimization algorithms can be used, e.g., metaphor-based metaheuristics and swarm intelligence algorithms (Liao et al., 2014; Goldberg, 1989; Nouaouria and Boukadoum, 2011). However, this type of methods is not designed to solve efficiently large scale optimization problems (Sigmund, 2011; Stolpe, 2011). Pattern search strategies have also been proposed to solve mixed variable optimization problems with categorical variables (Audet and Dennis, 2001; Abramson et al., 2009; Audet et al., 2018). In these approaches, mixed variables programming (MVP) is combined with mesh adaptive direct search (MADS) and a surrogate-assisted strategy (Audet et al., 2018). The drawback of such approaches is mainly related to the definition of a suitable neighborhood to be able to handle the categorical choices. Other approaches based on the discrete global descent method have been proposed to solve mixed optimization problems (Lindroth and Patriksson, 2011).

In the context of structural optimization problems, various surrogate-based optimization strategies have been extended to categorical variables (Filomeno Coelho, 2014; Müller et al., 2013; Herrera et al., 2014; Roy et al., 2017, 2019; Garrido-Merchán and Hernández-Lobato, 2018; Pelamatti et al., 2019). One of the main challenges of such approaches is related to their efficiency when handling large dimension categorical design space. Furthermore, a definition of a neighborhood is often required during the construction of the surrogate model. As an example, in (Pelamatti et al., 2019), the neighborhood is defined through an appropriate kernel definition. In these approaches, once the surrogate model is built, the optimizer still faces a large scale discrete optimization problem. A recent work (Gao et al., 2018) also suggests reducing the dimension of a structural optimization problem by finding implicit correlation between the design variables. Existing works propose to solve structural optimization problems (with multi-material and multi-cross-section design variables) using a continuous formulation of the design space which is provided by means of interpolation schemes (Stegmann and Lund, 2005; Krogh et al., 2017). Although such approaches allow to leverage the efficiency of gradient-based optimization algorithms, there is no guarantee that the optimization will retrieve integer values corresponding to a given material, for instance.

Other existing approaches rely on the structure of the mathematical mixed variable problem to decompose the intial problem into several more tractable subproblems. For instance, by the use of Benders decomposition (Benders, 1962; Geoffrion, 1972) or by outer approximation schemes (Duran and Grossmann, 1986; Hijazi et al., 2014). For structural optimization, decomposition schemes have been mostly applied to continuous optimization problems, e.g., StiffOpt (Samuelides et al., 2009), Quasi Separable Decomposition (QSD) (Haftka et al., 2006; Schutte et al., 2004). The QSD has then been applied to structural optimization of large scale composite structures (Bettebghor et al., 2011, 2018). In this context, the composite stacking sequences were formulated as continuous variables by using lamination parameters. In (Allaire and Delgado, 2015), both the composite fiber, lay-up sequence and the ply topology are optimized into a bi-level scheme. The main difficulty of existing decomposition schemes is related to the fact that they are not able to handle large scale mixed optimization problems with categorical variables. In the industry, methodologies have emerged to tackle the curse of dimensionality when dealing with categorical variables in structural optimization. For instance, (Grihon, 2018) uses a bi-step strategy involving massively parallel element-wise optimizations. This approach industrially used at Airbus simplifies the impact of each categorical choice on the overall optimal internal loads distribution by deporting the optimization at element (subsystem) level. Although the proposed approach is highly scalable, it can not handle system-level behavior (optimum internal load distribution) nor system-level constraints (e.g., flutter, modal or displacement constraints). The absence of such constraints in the problem formulation is not representative of aircraft structure design problems, in a multidisciplinary context for instance.

In this article, a bi-level methodology is proposed to solve mixed categorical-continuous structural optimization problems while trying to cover all the aforementioned gaps. To the best of our knowledge, in the context of weight minimization subject to stress and displacement constraints, problems combining simultaneously both materials, cross-section types, and continuous design variables are not tackled in the literature. The problem is formulated using a bi-level decomposition involving master and slave problems. The continuous design variables are handled by the slave problem, where the categorical variables are driven by the master. The latter consists of solving a first order-like approximation of the slave problem with respect to the categorical design variables. This helps to drastically reduce the combinatorial explosion raised by the categorical variables. In a previous work (Barjhoux et al., 2017), the master problem was solved without approximation by a branch and bound algorithm. However, the computational cost scaling with the number of elements prevents the use of the algorithm for large scale industrial applications. For this reason, we choose here to solve the master problem by using a first order like approximation.

This paper is organized as follows. In Section 2, we describe the problem formulation, the physical model involved, and the links with the design variables. In Section 3, the bi-level decomposition and the approximation at the upper level are presented. Finally, the accuracy of the optimum and the scalability of the proposed approach are compared with state-of-the-art algorithms in Section 4. Possible future extensions and concluding remarks will be given in the last Section.

## 2 Problem Statement

### 2.1 Design variables

In this article, our goal is to minimize the weight of a structure, at fixed topology, by exploring the internal geometry as well as material description of all the structural elements of the problem. Two kinds of design variables are thus involved when handling these kind of problems.

First, the areas of stiffeners cross-section, also named sizing variables, are continuous design variables. Formally, the areas can be represented as a vector $\mathbf{a} \in \mathbb{R}^n$ where the number of components $n$ corresponds to the number of structural elements. For a given choice of cross-section type, the areas scale the internal shape of the structural elements (Grihon, 2018; Barjhoux et al., 2017). Fig. 1 shows how internal parameters, and thus area moments of inertia, can be scaled using the area of the cross-section. Several examples of cross-section types are given in Fig. 2. This means that the parameters that describe the internal cross-section are not directly driven by the optimizer, they are latent variables. One advantage of such parameterization is that it keeps the number of sizing variables independent from the number of detailed geometrical parameters of the stiffeners. Furthermore, this removes from the design space impractical or non-physical configurations (Gao et al., 2018).

The second type of design variables that are involved during the optimization process are categorical choices. Indeed, in this work, the possible choices of material and stiffener for each element will also be regarded as a part of the design variables and that have to be explored as well. The categorical choices will be represented by a vector $\mathbf{c}$ that has $n$ components where $n$ is the number of elements in the structure. In this context, we assign to each element a choice of material and stiffener, all described by one categorical design variable. Where $\Gamma$ is the enumerated set that contains all possible choices of materials and stiffeners for each element of the structure. Assuming that there are $p$ possible choices per element, the categorical set is given by $\Gamma = \{1, \ldots, p\}$. Each value of this set is called a catalog. The vector of categorical variables $\mathbf{c}$ belongs to $\Gamma^n$, meaning that each component of the categorical variable can take a value among the same set of catalogs $\Gamma$. The size of the categorical and continuous design space remains fixed during the optimization.

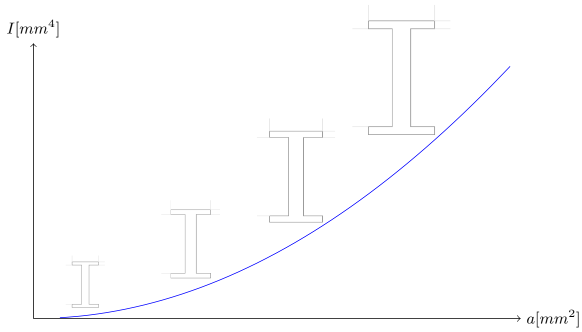


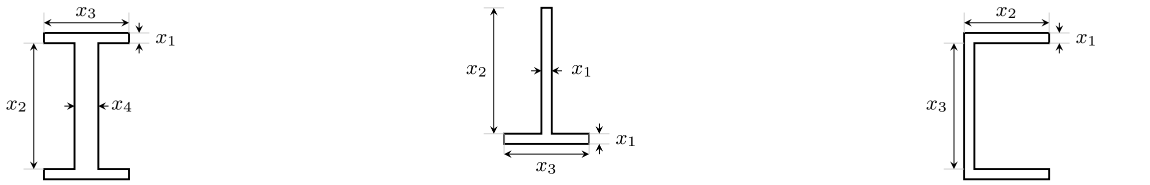

Fig. 1: Trend of area moment of inertia of a "I"-stiffener in Fig. 2a with respect to areas.

### 2.2 Problem definition

In this paper, we consider the following mixed categorical continuous optimization problem:

$$\begin{aligned} \underset{\mathbf{c}\in\Gamma^n, \mathbf{a}\in\mathbb{R}^n}{\text{minimize}} \quad & w(\mathbf{a},\mathbf{c}) \\ \text{subject to} \quad & \boldsymbol{s}(\mathbf{a},\mathbf{c}) \leq \mathbf{0}_{n,m} \\ & \boldsymbol{\delta}(\mathbf{a},\mathbf{c}) \leq \mathbf{0}_d \\ & \underline{\mathbf{a}} \leq \mathbf{a} \leq \bar{\mathbf{a}} \end{aligned} \tag{P}$$

where $\underline{\mathbf{a}} \in \mathbb{R}^n$ and $\bar{\mathbf{a}} \in \mathbb{R}^n$ are the lower and upper bounds on areas, respectively. The constraints $\boldsymbol{\delta}$ on displacements $\mathbf{u}$ ensures that on $d$ given nodes of the truss the displacements will not exceed predefined upper bounds $\bar{\mathbf{u}} \in \mathbb{R}^d$. With $\boldsymbol{P}$ a projector that selects the elements on which the displacement constraint will




(a) Example of an "I"-stiffener, described by 4 geometrical variables.

(b) Example of a "T"-stiffener, described by 3 geometrical variables.

(c) Example of a "C" profile, described by 2 geometrical variables.

Fig. 2: Examples of commonly used stiffeners in aircraft structural design. The internal geometrical variables are latent variables, scaled by the area of the cross-section.

apply, the definition of $\boldsymbol{\delta}$ function is given as follows:

$$\begin{aligned}\boldsymbol{\delta}\colon \mathbb{R}^n \times \Gamma^n &\to \mathbb{R}^d \\ (\mathbf{a}, \mathbf{c}) &\mapsto \boldsymbol{P}\boldsymbol{u}(\mathbf{a}, \mathbf{c}) - \bar{\mathbf{u}}.\end{aligned}$$

The structural constraints function $\boldsymbol{s}$

$$\boldsymbol{s}\colon \mathbb{R}^n \times \Gamma^n \to \mathcal{M}(\mathbb{R}^{n,m})$$

is of the form

$$\begin{array}{c} \\ elt_1 \\ elt_2 \\ \vdots \\ elt_n \end{array}
\begin{array}{c} \text{Constraint type 1} \quad \dots \quad \text{Constraint type } m \\
\begin{pmatrix} \boldsymbol{s}_{11}(\mathbf{a}_1, \mathbf{c}_1, \boldsymbol{\Phi}_1(\mathbf{a}, \mathbf{c})) & \dots & \boldsymbol{s}_{1m}(\mathbf{a}_1, \mathbf{c}_1, \boldsymbol{\Phi}_1(\mathbf{a}, \mathbf{c})) \\ \boldsymbol{s}_{21}(\mathbf{a}_2, \mathbf{c}_2, \boldsymbol{\Phi}_2(\mathbf{a}, \mathbf{c})) & \dots & \boldsymbol{s}_{2m}(\mathbf{a}_2, \mathbf{c}_2, \boldsymbol{\Phi}_2(\mathbf{a}, \mathbf{c})) \\ \vdots & \vdots & \ddots \\ \boldsymbol{s}_{n1}(\mathbf{a}_n, \mathbf{c}_n, \boldsymbol{\Phi}_n(\mathbf{a}, \mathbf{c})) & \dots & \boldsymbol{s}_{nm}(\mathbf{a}_n, \mathbf{c}_n, \boldsymbol{\Phi}_n(\mathbf{a}, \mathbf{c})) \end{pmatrix}\end{array}$$

and ensures, element per element, that the structural stress does not exceed a limit stress value.

Note that the continuous and categorical variables (i.e., the areas $\mathbf{a}$ and the categorical variables $\mathbf{c}$ hiding material or stiffeners choices) have a significant role in this problem. In fact, the categorical variables affect the weight $w$, the internal forces $\boldsymbol{\Phi} \in \mathbb{R}^n$, the constraints $\boldsymbol{\delta}$ on displacements and the stress constraints $\boldsymbol{s}$. On the other hand, the continuous variables affect the weight, the internal forces, the stress constraints and the constraints on displacements. It is worth noting that a change in a categorical variable or area will modify the loads distribution $\boldsymbol{\Phi}$ along the whole structure. Since the stresses $\boldsymbol{s}$ require the value of $\boldsymbol{\Phi}$, each component of $\boldsymbol{s}$ vector depends on the whole structure description.

In the context of this work, there is no change in the topology of the structure. Internal forces $\boldsymbol{\Phi}$ and displacements $\mathbf{u}$ will be computed using the direct stiffness method, introduced in (Turner, 1959; Turner et al., 1964). Structural elements are considered as truss elements with pin-jointed connections. This means that the bars will only carry axial forces. At each node, displacements are allowed along the global axes. Each element $i$ is defined by the elementary stiffness matrix $\boldsymbol{K}_i^e(\mathbf{a}_i, \mathbf{c}_i) \in \mathbb{R}^{q,q}$, with $q$ the number of free nodes multiplied by the number of physical space dimensions. The global stiffness of the whole truss is given by the matrix $\boldsymbol{K}(\mathbf{a}, \mathbf{c}) \in \mathbb{R}^{q,q}$ in global coordinates. Such matrix can be computed as the sum of each elementary stiffness matrix expressed after its transformation with the $i^{th}$ element rotation matrix $\mathbf{T}_i$, i.e., (Turner, 1959; Turner et al., 1964):

$$\boldsymbol{K}(\mathbf{a}, \mathbf{c}) = \sum_{i=1}^{n} [\mathbf{T}_i^t \boldsymbol{K}_i^e(\mathbf{a}_i, \mathbf{c}_i) \mathbf{T}_i].$$

Given a vector $\boldsymbol{f} \in \mathbb{R}^q$ of external loads applied on each of the free nodes in the global coordinates, the vector of displacements $\boldsymbol{u} \in \mathbb{R}^q$ can be obtained by solving the following equation:

$$\boldsymbol{K}(\mathbf{a}, \mathbf{c})\boldsymbol{u}(\mathbf{a}, \mathbf{c}) = \boldsymbol{f}.$$

The vector of internal forces $\boldsymbol{\Phi} \in \mathbb{R}^n$ is then given by: $\forall i \in \{1, \dots, n\}$,

$$\boldsymbol{\Phi}_i(\mathbf{a}, \mathbf{c}) = \boldsymbol{K}_i^e(\mathbf{a}_i, \mathbf{c}_i) \mathbf{T}_i \boldsymbol{u}_i(\mathbf{a}, \mathbf{c}),$$

where $\boldsymbol{\Phi}_i$ is the axial force through element $i$ and $\boldsymbol{u}_i$ its displacement vector.

## 3 Methodology

### 3.1 Decomposition

For a given $\mathbf{c}$, let $\Omega(c)$ be the set of feasible constraints given by

$$\begin{aligned}\Omega(\mathbf{c}) := \{\ \mathbf{a} \in \mathbb{R}^n ; \\ \boldsymbol{s}(\mathbf{a}, \mathbf{c}) \le \mathbf{0}_{n,m}\ ; \\ \boldsymbol{\delta}(\mathbf{a}, \mathbf{c}) \le \mathbf{0}_d\ ; \\ \underline{\mathbf{a}} \le \mathbf{a} \le \bar{\mathbf{a}}\ \}.\end{aligned}$$

An efficient way to solve pure continuous optimization problems is by taking advantage of gradient-based algorithms. In the problem introduced in Section 2, it can be seen that by fixing (temporarily) the categorical variables in (P), the problem becomes a continuous optimization problem, parameterized with $\mathbf{c}$. This means that given $\mathbf{c}$, the weight $w$ can be minimized with respect to the continuous design variables, that are the areas $\mathbf{a}$ subject to $\Omega(\mathbf{c})$. This leads to the following slave problem, that reduces to a structural sizing optimization problem (sP):

$$\Psi(\mathbf{c}) := \min_{\mathbf{a}\in\Omega(\mathbf{c})} w(\mathbf{a},\mathbf{c}). \tag{sP}$$

The structure of the problem is such that this remaining optimization problem becomes more tractable. In fact, the decomposition leverages the use of the gradients (with respect to $\mathbf{a}$) of the objective and constraints to solve (sP). This is the main motivation in handling the continuous variables separately from the categorical ones. In our approach, the categorical variables will be handled by a master problem (mP) of the form

$$\min_{\mathbf{c}\in\Gamma^n} \Psi(\mathbf{c}), \tag{mP}$$

where $\Psi(\mathbf{c})$ is the result of the slave Problem (sP).

The slave Problem (sP) takes these complicating variables as parameters while optimizing with respect to continuous design variables. This follows the generalized Benders decomposition in (Geoffrion, 1972), initially designed to handle linear optimization problems in (Benders, 1962). For given choices of materials and cross-sectional types for all elements, the continuous optimization will be performed using a gradient-based method. The result of this optimization can be seen as a function $\Psi(\mathbf{c})$ which is parameterized by the categorical choices. Namely, $\Psi(\mathbf{c})$ corresponds to the optimal weight of the slave problem knowing the categorical variables $\mathbf{c}$. This function is then taken as the objective of the master optimization Problem (mP). Although the slave problem can be easy to handle using gradient-based algorithm, the difficult part remains in the master problem. In fact, the (mP) problem is a large-scale categorical optimization problem, that usual metaheuristic algorithms fail to solve efficiently. In this work, we propose to consider, at the master level, the minimization of an approximated model $\hat{\Psi}(\mathbf{c})$ instead of $\Psi(\mathbf{c})$, so that the combinations can be reduced drastically, as explained later.

For that, we use the following iterative scheme: given an iteration $(k)$, the master problem (mP) of the bi-level formulation reduces to the following :

$$\mathbf{c}^{(k+1)} := \operatorname*{argmin}_{\mathbf{c}\in\Gamma^n} \hat{\Psi}_k(\mathbf{c}) \tag{amP}$$

where $\hat{\Psi}_k$ is a given approximation function of $\Psi$ at the iteration $(k)$ that depends locally on the previous iteration. Such a problem will be called the approximation master mixed Problem (amP) at iteration $(k)$. At each iteration $(k)$ of the algorithm, instead of $\Psi$, an approximation function $\hat{\Psi}_k$ is minimized with respect to the global variable $\mathbf{c}$.

Let's call $\mathbf{a}^{(k)}$ the optimal areas obtained by solving Problem (sP) for given choices $\mathbf{c}^{(k)}$ :

$$\mathbf{a}^{(k)} := \operatorname*{argmin}_{\mathbf{a}\in\Omega(\mathbf{c}^{(k)})} w(\mathbf{a},\mathbf{c}^{(k)}).$$

Let also $w^{(k)}$ be the optimal weight returned by the evaluation of the weight function taken at $\mathbf{a}^{(k)},\mathbf{c}^{(k)}$, i.e.,

$$w^{(k)} := w(\mathbf{a}^{(k)},\mathbf{c}^{(k)}).$$

In this article, the termination criterion is based on the stationarity of the optimal weights, i.e., $|w^{(k+1)} - w^{(k)}| \le \epsilon$ for a given small $\epsilon > 0$. However, there will be no guarantee that a weight decreases lower than $\epsilon$ during the optimization process. A possible way to justify such decrease would be to prove that the proposed algorithm converges to a fixed point $w^*$ independently of the starting point. Ensuring this in the general case would require imposing a decrease on the weight sequence $\{w^{(k)}\}$ over the iterations. In Section 3.4, a heuristic is proposed to handle that by further exploration of the design space.

The generic process of the proposed iterative scheme is given in Algorithm 1.

**Algorithm 1** A generic Bi-level framework

1: **initialize** $\mathbf{c}^{(0)}$and set $k=0$
2: **while** a termination criteria is not reached **do**
3: `Step 1` $\mathbf{c}^{(k+1)} \leftarrow \operatorname{argmin}\ \hat{\Psi}_k(\mathbf{c})$ s.t $\mathbf{c}\in\Gamma^n$
4: `Step 2` $\mathbf{a}^{(k+1)} \leftarrow \operatorname{argmin} w(\mathbf{a},\mathbf{c}^{(k+1)})$ s.t. $\mathbf{a} \in \Omega(\mathbf{c}^{(k+1)})$
5: Increment $k$
6: **end while**
7: **return** $\mathbf{a}^{(k+1)}$, $\mathbf{c}^{(k+1)}$, and $w^{(k+1)} \leftarrow w(\mathbf{a}^{(k+1)},\mathbf{c}^{(k+1)})$.

### 3.2 On the approximation $\hat{\Psi}_k$

The problem (mP) involves a number of categorical combinations that increases exponentially $(p^n)$ with the number of catalogs and structural elements. In this Section, we aim to propose an approximation $\hat{\Psi}_k$ to the function $\Psi$ so that one can reduce the resulting problem complexity. In fact, we propose solving an approximation $\hat{\Psi}_k$ by using a first order approximation of $\Psi$. The expression of the approximation $\hat{\Psi}_k$ around the categorical variable $\mathbf{c}^{(k)}$ is given by:

$$\hat{\Psi}_k(\mathbf{c}) = \Psi(\mathbf{c}^{(k)}) + \sum_{i=1}^{n} \left.\frac{\Delta\Psi}{\Delta\mathbf{c}_i}\right|_{\mathbf{c}^{(k)}} (\mathbf{c}_i - \mathbf{c}_i^{(k)}), \tag{1}$$

where the scalar value $\Delta \mathbf{c}_i$ denotes the perturbation of the $i^{th}$ component of $\mathbf{c}$ starting from $\mathbf{c}^{(k)}$, $\left.\frac{\Delta \Psi}{\Delta \mathbf{c}_i}\right|_{\mathbf{c}^{(k)}} \in \mathbb{R}$ model the rate of the $\Psi$ function taken at $\mathbf{c}^{(k)}$ after a perturbation $\Delta \mathbf{c}_i$. The term $\left.\frac{\Delta \Psi}{\Delta \mathbf{c}_i}\right|_{\mathbf{c}^{(k)}}$ is computed as follows :

$$\left.\frac{\Delta \Psi}{\Delta \mathbf{c}_i}\right|_{\mathbf{c}^{(k)}} = \frac{\Psi(\mathbf{c}^{(k)} + \Delta \mathbf{c}_i \mathbf{e}_i) - \Psi(\mathbf{c}^{(k)})}{\Delta \mathbf{c}_i}, \qquad (2)$$

with $\mathbf{e}_i$ a vector of size $n$ where the $i^{th}$ component is equal to 1 and 0 everywhere else.

We note that the term $\Delta \mathbf{c}_i$ has no physical meaning and, due to the categorical nature of the set $\Gamma$, there is no straightforward neighborhood definition. This prevents from choosing $\mathbf{c}_i$ so that $\mathbf{c}_i - \mathbf{c}^{(k)}$ is close to the perturbation $\Delta \mathbf{c}_i$ used to compute the rate $\left.\frac{\Delta \Psi}{\Delta \mathbf{c}_i}\right|_{\mathbf{c}^{(k)}}$. To overcome this issue, we propose setting the perturbation $\Delta \mathbf{c}_i$ equal to $\mathbf{c}_i - \mathbf{c}_i^{(k)}$. In this case, combining (1) and (2), one results in

$$\hat{\Psi}_k(\mathbf{c}) = \Psi(\mathbf{c}^{(k)}) + \sum_{i=1}^{n} \left( \Psi(\mathbf{c}^{(k)} + \Delta \mathbf{c}_i \mathbf{e}_i) - \Psi(\mathbf{c}^{(k)}) \right). (3)$$

Equation (3) verifies, in the trivial case where the structure is composed of one element ($n = 1$), that the approximation $\hat{\Psi}_k(\mathbf{c})$ is equal to $\Psi(\mathbf{c})$, for every $\mathbf{c}^{(k)} + \Delta \mathbf{c}_i \mathbf{e}_i$ in $\Gamma$. Knowing $\Delta \mathbf{c}_i = \mathbf{c}_i - \mathbf{c}_i^{(k)}$, the term $\mathbf{c}^{(k)} + \Delta \mathbf{c}_i \mathbf{e}_i$ is equal to $\mathbf{c}^{(k)}$ except the $i^{th}$ component which is equal to $\mathbf{c}_i$.

Physically, the approximation (1) relies on the hypothesis that the effects of the couplings between the categorical variables on the optimized weight solutions of (sP) can be neglected. The block-diagonal dominance property of the stress constraints jacobian (Haftka and Gürdal, 1992), in our case with respect to both material properties and quadratic moments, makes the problem quasi-separable. This property has been largely used in the literature in the context of structural sizing problems, as in (Haftka and Watson, 2005; Haftka et al., 2006), and in the context of simultaneous sizing and material optimal selection (Bettebghor et al., 2018). However, a breakdown of the categorical variables as proposed in the Quasi Separable Decomposition scheme (Haftka and Watson, 2005) in the context of sizing variables, is not investigated here. The categorical variables are taken as design variables in the master problem (mP) only. The same remark applies for the areas, that are optimized in the slave problem only. In the proposed approach, the quasi-separable property is leveraged through the result of the areas optimizations in (sP). It is worth noting that the couplings between the elements are still partially impacting the optimizations thanks to the optimized areas, solutions of (sP). Elements are not optimized independently, since at each change of categorical variable, a sizing of the whole structure is performed. We note that the first order-like approximation does not introduce any loss of symmetry (when symmetric results are expected).

### 3.3 On the minimization of $\hat{\Psi}_k$

In the previous Section, it was proposed to replace the expression of $\Psi$ with the approximation $\hat{\Psi}_k$ when solving the master level problem (amP). Since $\Psi(\mathbf{c}^{(k)})$ is constant, the minimization of (3) according to `Step 1` of the Algorithm 1 is equivalent to :

$$\forall i \in \{1, ..., n\}, \min_{\mathbf{c}_i \in \Gamma} \Psi\left(\mathbf{c}^{(k)} + (\mathbf{c}_i - \mathbf{c}_i^{(k)})\mathbf{e}_i\right).$$

Indeed, the approximated master problem (amP) can be written as a number of $n$ independent categorical optimizations (sP). This reduces drastically the combinatorial explosion, i.e., instead of minimizing over $\Gamma^n$ we get $n$ minimizations but only over the space $\Gamma$. This is a crucial point of the methodology proposed in this article. In fact, at each iteration $(k)$ of the algorithm, the first order like assumption of the model $\hat{\Psi}_k$ makes the combinatorial of problem (mP) drop from $p^n$ to $k \times n \times (p-1)$ combinations of choices with $k$ the number of overall iterations.

Namely, in `Step 1` of the Algorithm 1, the current categorical choices $\mathbf{c}^{(k+1)} := \left[\mathbf{c}_1^{(k+1)}, \ldots, \mathbf{c}_n^{(k+1)}\right]$ is given by solving the following approximated master problem:

$$\mathbf{c}_1^{(k+1)} := \operatorname*{argmin}_{\mathbf{c}_1 \in \Gamma} \Psi\left(\left[\mathbf{c}_1,\ \mathbf{c}_2^{(k)},\ \ldots,\ \mathbf{c}_n^{(k)}\right]\right),$$

for $i \in \{2, \ldots, n-1\}$:

$$\mathbf{c}_i^{(k+1)} := \operatorname*{argmin}_{\mathbf{c}_i \in \Gamma} \Psi\left(\left[\ldots,\ \mathbf{c}_{i-1}^{(k)},\ \mathbf{c}_i,\ \mathbf{c}_{i+1}^{(k)},\ \ldots\right]\right), \quad \text{(amP2)}$$

and

$$\mathbf{c}_n^{(k+1)} := \operatorname*{argmin}_{\mathbf{c}_n \in \Gamma} \Psi\left(\left[\mathbf{c}_1^{(k)},\ \ldots,\ \mathbf{c}_{n-1}^{(k)},\ \mathbf{c}_n\right]\right).$$

Each of these $n$ optimizations is solved by enumeration of all the remaining values that can take $\mathbf{c}_i$ over $\Gamma$. This means that at each iteration $(k)$ of the algorithm, $\Psi$ is evaluated $n \times (p-1)$ times to build a new solution $\mathbf{c}^{(k+1)}$. All these evaluations can be performed in parallel. The results of all the optimal weights computed during this enumeration process will be stored in a matrix $W^{(k)}$, i.e.,

$$(\forall i \in \{1, \ldots, n\}) \quad \mathbf{c}_i^{(k+1)} = \operatorname*{argmin}_{j \in \Gamma} W_{ij}^{(k)}. \quad \text{(amP3)}$$

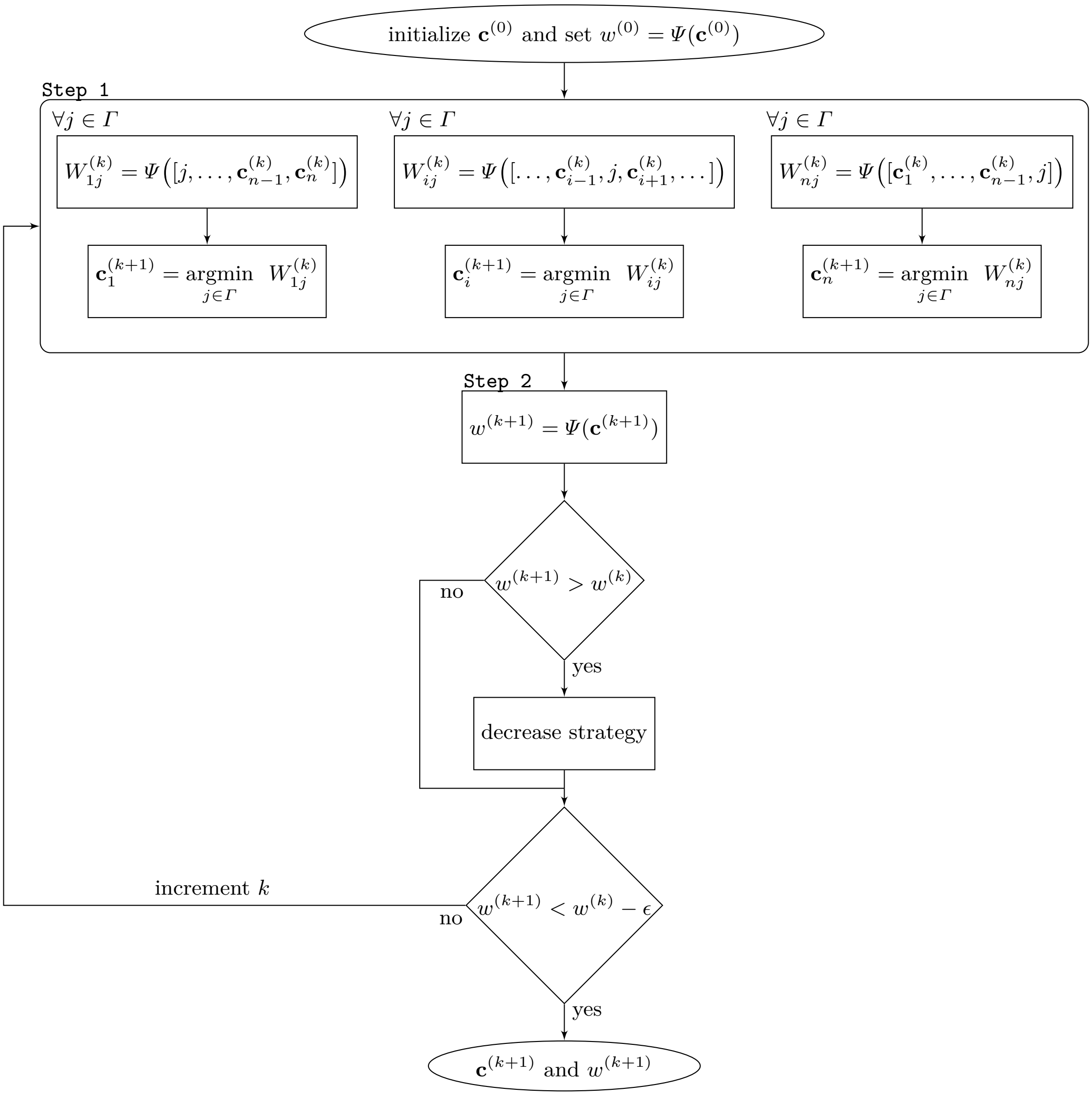



Fig. 3: Illustration of the proposed methodology.

A detailed view of the resulting Bi-level process is given in Fig. 3. As described in Algorithm 1, each iteration $(k)$ counts two main steps. `Step 1` consists of building a new solution $\mathbf{c}^{(k+1)}$. Physically, for every change of material and cross-section type of one element in the structure, sizing optimization is performed. The new categorical choice of the current element is made so that the corresponding optimal weight (with respect to the continuous variables) is the lowest. This process is repeated for each of the $n$ elements, in parallel. In `Step 2`, the materials and cross-section types of every elements are updated with the new categorical variables. A sizing optimization is then performed, leading to a new optimal weight.

However, in practice the first order approximation (3) can be responsible for convergence issues during the optimization : an increase in the optimized weight could be observed from one iteration to the next. For cases where this situation occurs, it is proposed applying a strategy that iteratively builds a new solution based on information stored in $W^{(k)}$. The proposed strategy is detailed in the next Section.

### 3.4 A strategy to ensure weight decrease

As only a first order like approximation of $\Psi$ is used, the coupling between the structural elements through the categorical variables are neglected. This may be responsible for convergence issues as the resulting optimal weight at the current iteration $w^{(k+1)}$ may be greater than the previous optimal weight $w^{(k)}$. In the context of continuous optimization, adaptive strategies based on step-size control are typically used (e.g., line search, trust-region methods) (Ivanov and Zadiraka, 1975). However, in the context of categorical optimization, the use of such a strategy is not straightforward anymore due to the lack of a neighborhood definition.

In order to fix this convergence issue observed on some cases, the following iterative strategy is proposed. At a given iteration $(k)$, the proposed process is triggered if the new optimal weight $w^{(k+1)}$ candidate is higher than the previous one $w^{(k)}$. It can be seen as an additional third step that would be introduced in Algorithm 1.

For each iteration $(k)$ where the $w^{(k+1)}$ candidate is higher than the previous candidate $w^{(k)}$, we activate the following iterative strategy starting from $\mathbf{c}^{(k+1)}$. Where $(t)$ denotes the outer iteration number of the proposed strategy related to the $k^{th}$ iteration of Algorithm 1. At each iteration $(t)$ of the strategy, a new categorical solution $\mathbf{c}^{(t+1)}$ will be built. The first step of the process consists of retrieving the element number $i$ and corresponding choice $j$ of the $(t+1)^{th}$ best weight in $W^{(k)}$, i.e.,

$$i, j := \underset{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \setminus \mathcal{F}_t}{\operatorname{argmin}} W^{(k)},$$

where $\mathcal{F}_t$ denotes the set of indices $i, j$ of all sub-iterations anterior to $t+1$. The new candidate solution $\mathbf{c}^{(t+1)}$ is given by

$$\mathbf{c}^{(t+1)} := [\mathbf{c}_1^{(t)}, \ \ldots, \ \mathbf{c}_{i-1}^{(t)}, \ j, \ \mathbf{c}_{i+1}^{(t)}, \ \ldots, \ \mathbf{c}_n^{(t)}].$$

Once $\Psi(\mathbf{c}^{(t+1)})$ is evaluated, the corresponding optimal weight $w$ is compared to $w^{(k)}$ in the following way: if the new optimal weight is found lower than $w^{(k)}$, then the candidate solution $\mathbf{c}^{(t+1)}$ becomes the new $\mathbf{c}^{(k+1)}$. The process stops and goes back to the main algorithm with the new optimal solution of iteration $(k+1)$. Algorithm 2 gives a detailed description of the proposed strategy.

Note that another possible decrease strategy would consist of repeating `Step 1` by changing only one structural element. In other words, from an iteration $(k)$ to $(k+1)$, one would change only one categorical variable. In this case, equation (3) would lead to $\Psi(\mathbf{c}) = \hat{\Psi}(\mathbf{c})$, meaning that the weight decrease would be ensured by minimization of $\hat{\Psi}$. However, in practice the weight decrease has shown to be lower compared to the weight decrease obtained by Algorithm 2.

**Algorithm 2** A proposed weight decrease strategy

**function** DECREASE_STRATEGY($w^{(k)}, \mathbf{c}^{(k+1)}, W^{(k)}$)
  $\mathbf{c}^{(0)} \leftarrow \mathbf{c}^{(k+1)}$, $\mathcal{F}_0 \leftarrow \{\emptyset\}$ and set $t = 0$
  **while** $\mathcal{F}_t \neq [\![1,n]\!] \times [\![1,p]\!]$ **do**
    $i, j \leftarrow \underset{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \setminus \mathcal{F}_t}{\operatorname{argmin}} W^{(k)}$
    $\mathbf{c}^{(t+1)} \leftarrow [\mathbf{c}_1^{(t)}, \ \ldots, \ \mathbf{c}_{i-1}^{(t)}, \ j, \ \mathbf{c}_{i+1}^{(t)}, \ \ldots, \ \mathbf{c}_n^{(t)}]$
    $\mathbf{a}^{(t+1)} \leftarrow \operatorname{argmin} w(\mathbf{a}, \mathbf{c}^{(t+1)})$ s.t. $\mathbf{a} \in \Omega(\mathbf{c}^{(t+1)})$
    $w^{(t+1)} \leftarrow w(\mathbf{a}^{(t+1)}, \mathbf{c}^{(t+1)})$
    $\mathcal{F}_{t+1} \leftarrow \mathcal{F}_t \cup \{(i,j)\}$
    **if** $w^{(t+1)} < w^{(k)}$ **then**
      break
    **end if**
    increment $t$
  **end while**
  $\mathbf{c}^{(k+1)} \leftarrow \mathbf{c}^{(t)}$
  $\mathbf{a}^{(k+1)} \leftarrow \mathbf{a}^{(t)}$
  $w^{(k+1)} \leftarrow w^{(t)}$
**return** $w^{(k+1)}, \mathbf{a}^{(k+1)}, \mathbf{c}^{(k+1)}$
**end function**

## 4 Numerical results

In the present section, our proposed methodology will be applied to three different test cases: (i) the well-known 10-bar truss structure (Haftka and Gürdal, 1992) adapted in (Merval, 2008), (ii) a 2D cantilever structure (Shahabsafa et al., 2018), and (iii) a 120-bar dome truss structure (Saka and Ulker, 1992). In order to evaluate the scalability of the methodology with respect to large number of structural elements, the 2D cantilever structure is made scalable by varying the number of blocks. The third test case is included to evaluate the capability of handling more complex structures.

We note that for all the test cases, we will consider four different structural constraints per element, i.e., $m = 4$ in the generic structural constraints expressions given in (2.2). For that, first, one has two constraints in tension and compression, respectively, given by

$$s_{i1}(\mathbf{a}, \mathbf{c}) := \frac{\mathbf{\Phi}_i(\mathbf{a}, \mathbf{c})}{\mathbf{a}_i} - \sigma^t(\mathbf{c}_i)$$
$$s_{i2}(\mathbf{a}, \mathbf{c}) := \frac{\mathbf{\Phi}_i(\mathbf{a}, \mathbf{c})}{\mathbf{a}_i} - \sigma^c(\mathbf{c}_i)$$

with $\sigma^t(\mathbf{c}_i) \in \mathbb{R}$ the stress limit in tension and $\sigma^c(\mathbf{c}_i) \in \mathbb{R}$ the stress limit in compression. The two other constraints are the Euler and local buckling constraints, respectively, given by

$$s_{i3}(\mathbf{a}, \mathbf{c}) \ := \ \frac{\mathbf{\Phi}(\mathbf{a}, \mathbf{c})}{\mathbf{a}_i} - \frac{\pi^2 E(\mathbf{c}_i) I(\mathbf{a}_i, \mathbf{c}_i)}{\mathbf{a}_i \mathbf{L}_i^2}$$
$$s_{i4}(\mathbf{a}, \mathbf{c}) \ := \ \frac{\mathbf{\Phi}(\mathbf{a}, \mathbf{c})}{\mathbf{a}_i} - \frac{4\pi^2 E(\mathbf{c}_i) \mathcal{K}(\mathbf{c}_i)}{12(1 - \nu^2(\mathbf{c}_i))}$$

with $E(\mathbf{c}_i)$, $I(\mathbf{a}_i, \mathbf{c}_i)$, $\mathbf{L}_i$, $\nu(\mathbf{c}_i)$ respectively the Young modulus, the area moment of inertia, the length and the poisson coefficient of element $i$. The ratio between cross-section internal sizes, depending on the stiffener profile, is given by $\mathcal{K}(\mathbf{c}_i)$.

Finally, the objective function (i.e., the global weight of the structure) is defined as follows :

$$w(\mathbf{a},\mathbf{c}) = \sum_{i=1}^{n} \mathbf{a}_i \boldsymbol{\rho}_i(\mathbf{c}_i)\mathbf{L}_i,$$

where $\rho(\mathbf{c}_i)$ corresponds to the material density of element $i$ driven by the choice $\mathbf{c}_i$.

### 4.1 Implementation details and comparison solvers

Algorithm 1 was implemented using the Generic Engine for MDO Scenarios (GEMS) (Gallard et al., 2018) in Python. The tool offers an efficient way to test multi-level formulations, with built-in classes that facilitate optimization problems manipulations (Gallard et al., 2019). The continuous optimization problems (i.e., evaluations of $\Psi$) are solved with the Method of Moving Asymptotes (MMA) (Svanberg, 2002) as implemented in the nonlinear-optimization (NLOPT) package (Johnson, 2008). All the default parameters are kept unchanged except the tolerance on the objective function which is set to $10^{-6}$ $kg$. In what comes next, the resulting implementation of Algorithm 1 will be called Bi-level.

Three solvers will be compared to Bi-level. First, a baseline solver where we proceed with an exhaustive enumeration of continuous optimizations w.r.t. $\mathbf{a}$ (Problem sP) taken at every available choice in $\Gamma^n$, the solution resulting with this solver will be denoted as Baseline. Second, a hybrid branch and bound (Barjhoux et al., 2017) (noted h-B&B) where the proposed methodology is also based on a categorical-continuous problem decomposition where, instead of approximating the master problem, one uses a relaxation procedure combined with a branch and bound algorithm. Similarly to Bi-level, h-B&B uses the MMA method from NLOPT to solve the slave problem. Under the assumption that these problems are convex with respect to the sizing variables $\mathbf{a}$, Baseline and h-B&B return the global optimum of the overall problem. The third solver used in the comparison is a Genetic algorithm (Deb and Goyal, 1998) using the implementation given by Distributed Evolutionary Algorithms in Python (DEAP) (Fortin et al., 2012). The latter solver will be referred by Genetic in our comparison tests. Due to the stochastic nature of Genetic, the obtained results (for this solver) will be displayed as the average of ten runs.

In all what comes next, the computation effort of a given solver will be measured by counting the number of structural analyses (noted #FEM) including those required by the computation of the gradients (when needed). The obtained optimal weights (by each solver) will be noted $w^*$, the latter will allow us to evaluate the quality of the optima found by each solver. We note also that in our setting, the Baseline solution can be seen as the best known categorical choices for the corresponding problem instance. Thus, in this context, it is important to evaluate how far the categorical choices (obtained by the tested solvers) are from the Baseline optimal choices. This information will be given using the Hamming distance (noted $d_h$) where we will count the number of structural elements that have an optimal choice different to the Baseline categorical choices.

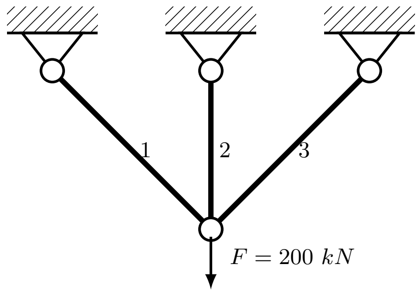



Fig. 4: A 3-bar truss structure where a downward load $F = 200$ $kN$ is applied on the free node.

### 4.2 A worked example: a 3-bar truss structure

To illustrate how the Bi-level method works, we will now describe in detail its application to a simple 3-bar truss structure. For this problem, each element can take a value among three possible choices that respectively point to materials AL2139, AL2024, TA6V and the same "I"-profile (see Fig. 2). The materials properties are listed in Appendix A. For this simple case, one has $n = 3$, $p = 3$, and $\Gamma = \{1, 2, 3\}$. For all elements, the lower and upper bounds on areas are respectively fixed to 100 $mm^2$ and 2000 $mm^2$. A maximum downward displacement equal to 1 $mm$ is allowed on the only free node of the structure.

During application of the Bi-level method, the initialization of $\mathbf{c}$ is such that :

$$\mathbf{c}^{(0)} = [1, 2, 3], \quad w^{(0)} = 13.82\ kg.$$

The first iteration (i.e. $k = 1$) of the Bi-level method can be described as follows: first, the optimization problems (given by (amP2)) are solved by enumeration of the evaluation of $\Psi$ for all values in $\Gamma$. Both categorical choices and corresponding optimal weights are given in Table 1.

| | $j=1$ | $j=2$ | $j=3$ |
|---|---|---|---|
| $i=1$ | $\mathbf{c}^{(0)}$<br>$W_{11}^{(0)} = w^{(0)}$ | $[②, 2, 3]$<br>$W_{12}^{(0)} = 13.62$ | $[③, 2, 3]$<br>$W_{13}^{(0)} = 13.92$ |
| $i=2$ | $[1, ①, 3]$<br>$W_{21}^{(0)} = 14.85$ | $\mathbf{c}^{(0)}$<br>$W_{22}^{(0)} = w^{(0)}$ | $[1, ③, 3]$<br>$W_{23}^{(0)} = 8.83$ |
| $i=3$ | $[1, 2, ①]$<br>$W_{31}^{(0)} = 13.74$ | $[1, 2, ②]$<br>$W_{32}^{(0)} = 13.53$ | $\mathbf{c}^{(0)}$<br>$W_{33}^{(0)} = w^{(0)}$ |

Table 1: Enumeration of the $n \times (p-1) = 6$ perturbed categorical variables (circled components) at first iteration, with the corresponding optimal weight in $(kg)$.

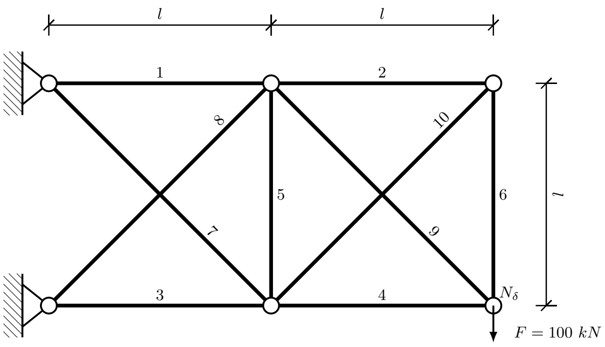



Fig. 5: 10-bar truss, seen as a scalable 2D cantilever problem with 2 blocks.

As described in problems (amP2), the vector of categorical variables at the first iteration is composed of the values in $\Gamma$ corresponding to the best weights given in Table 1, element per element. For example, $\mathbf{c}_1^{(1)} = \underset{j \in \Gamma}{\text{argmin}}\ W_{1j}^{(0)} = 2$. The new optimal vector of categorical variables $\mathbf{c}$ at iteration $k=1$ is thus $\mathbf{c}^{(1)} = [2, 3, 2]$, leading to an optimal weight $w^{(1)} = 8.63\ kg$. After this first iteration, the optimal weight drops from 13.82 $kg$ to 8.63 $kg$.

According to Algorithm 1, the same steps (that are not detailed) are executed in the next iteration $k=2$. The same categorical vector solution is found, leading to the same optimal weight. Since a stationarity of the optimal weight is reached, the method stops. Using the Baseline method (by enumerating evaluations of (sP) over the space $\Gamma^3$), we find that the best solution is indeed equal to $\mathbf{c}^{(1)}$.

### 4.3 A 10-bar truss structure

This well-known low-dimensional 10-bar truss problem (Haftka and Gürdal, 1992) is used to solve the mixed categorical-continuous optimization problem by enumeration or hybrid branch and bound (h-B&B) (Barjhoux et al., 2017). As explained in subsection 4.1, these approaches provide global solutions, that are taken as reference solutions to evaluate quality solutions of the Bi-level algorithm.

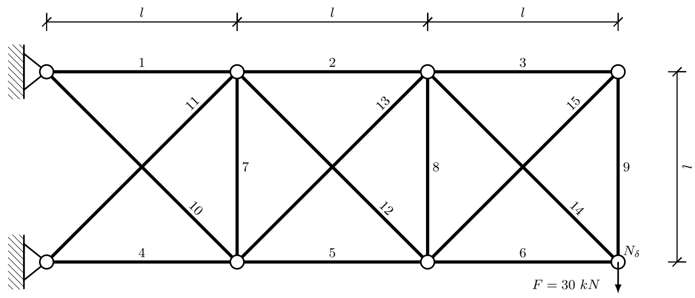



Fig. 6: An example of 2D cantilever problem with 3 blocks.

The 10-bar truss problem is illustrated Fig. 5. A downward load $F = 100\ kN$ is applied vertically on node $N_\delta$. A constraint on displacements is applied on the same node. Five cases with different bounds values $\bar{\mathbf{u}}$ on displacements are considered. For each of these cases, the displacements constraint is applied on node $N_\delta$. Catalogs 1 and 2 point to materials AL2139 and TA6V, respectively. Materials properties are listed in Appendix A. In this case, $n = 10$ and $p = 2$, $\Gamma = \{1, 2\}$.

The results of the proposed methodology (Bi-level) are thus compared to the global optima, as shown in Table 2. In all these cases, the optima obtained with `Baseline` (obtained by enumeration), h-B&B the Bi-level approaches are identical. This means that the Bi-level, in these cases, provides the global solution. On the other hand, the weights returned by the Genetic algorithm are greater than the optimal weight found by the Bi-level approach. Finally, it is shown that when the displacement constraint becomes more stringent, the material choice goes to the stiffest one despite of its high density. The optimal solutions of cases with displacements lower than $18mm$ and $17mm$ contain indeed only TA6V.

### 4.4 A scalable 2D cantilever problem

The objective of this test case is to describe the evolution of the computational cost with respect to the number of structural elements. This case can be seen as a generalization of the well-known 10-bar truss structure (Haftka and Gürdal, 1992). It has been used in the literature to demonstrate the scalability of algorithms, for example in (Shahabsafa et al., 2018). The structure is made scalable by varying the number of blocks. Each block is composed of 4 nodes that are linked by 5 bars.

Table 2: Results of 10-bar truss mixed optimization with 5 different values of constraint on displacements. Comparison between the Bi-level, the Baseline solutions obtained by enumeration of the $2^{10}$ continuous optimizations, h-B&B, and the Genetic algorithm. The catalog 1 corresponds to material AL2139 and catalog 2 to TA6V.

| $\bar{\mathbf{u}}$ $(mm)$ | Baseline | | h-B&B | | Genetic | | Bi-level | |
|---|---|---|---|---|---|---|---|---|
| | $c^*$ | $w^*(kg)$ | $d_h$ | $w^*(kg)$ | $d_h$ | $w^*(kg)$ | $d_h$ | $w^*(kg)$ |
| -22 | [2,2,1,1,1,2,2,1,2,1] | 12.988 | 0 | 12.988 | 0 | 13.283 | 0 | 12.988 |
| -20 | [2,1,1,1,1,1,2,1,1,1] | 13.996 | 0 | 13.996 | 0 | 14.423 | 0 | 13.996 |
| -19 | [2,1,1,1,1,1,2,1,1,1] | 14.570 | 0 | 14.570 | 0 | 14.802 | 0 | 14.570 |
| -18 | [1,1,1,1,1,1,1,1,1,1] | 15.175 | 0 | 15.175 | 2 | 15.642 | 0 | 15.174 |
| -17 | [1,1,1,1,1,1,1,1,1,1] | 15.912 | 0 | 15.912 | 3 | 16.258 | 0 | 15.912 |

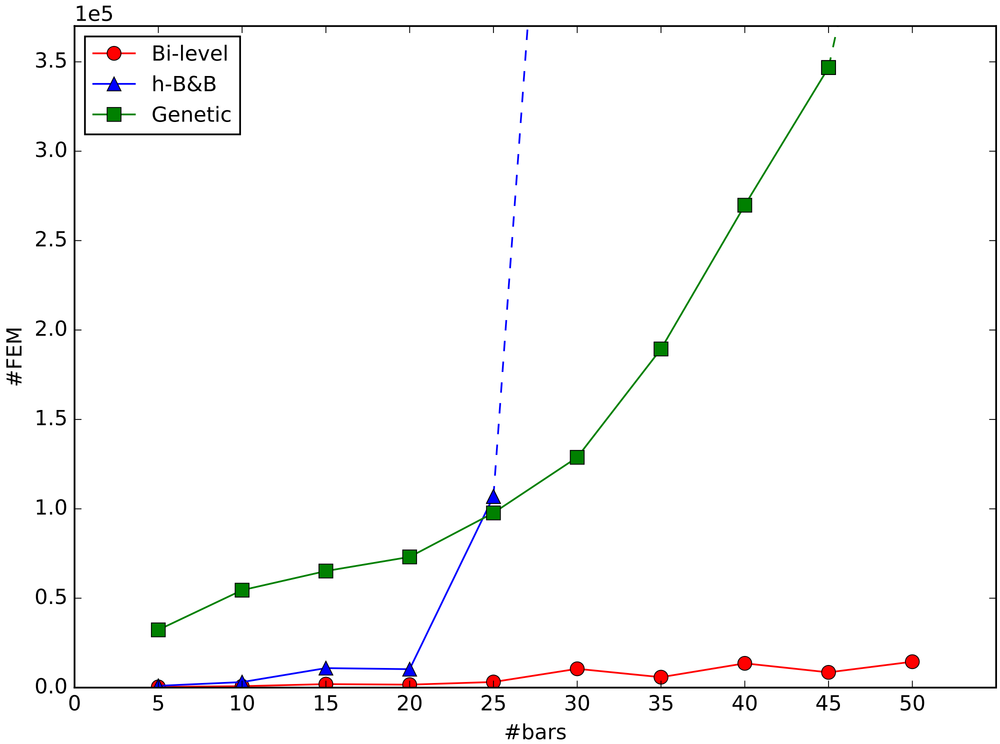



Fig. 7: Scalability of the Bi-level w.r.t. the number of elements. The computational cost's scaling of Bi-level with respect to the number of bars is quasi-linear, compared to the exponential computational cost of the h-B&B and Genetic solvers. The computational cost's scaling of the h-B&B prevents from obtaining a solution for cases greater than 25 elements.

An example of a scalable 2D cantilever structure with 3 blocks is given in Fig. 6. In Table 3 are presented the results obtained with structures composed of 1 to 10 blocks. In all cases, a downward load $F = 30\ kN$ is applied on the node $N_\delta$.

For each of the 10 cases, the results obtained by the Bi-level are compared to those obtained with reference solutions (Baseline & h-B&B) when available. First, for cases with 5 to 30 elements where a reference solution is available, it can be observed the global solution is found by the Bi-level. For cases with more than 30 elements, the optima found by the Bi-level are slightly better than those obtained by the Genetic algorithm. The h-B&B solutions are noted with (*) since they are intermediate solutions : the solver was stopped after 24 hours. The Bilevel solutions are very close (difference of $10^{-2}$ kg) to those obtained by the h-B&B. Furthermore, the number of analyses required by Bi-level is always lower than

| #bars | Baseline | h-B&B | | | | Genetic | | | | Bi-level | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $w^*(kg)$ | $d_h$ | $w^*(kg)$ | #iter | #FEM | $d_h$ | $w^*(kg)$ | #iter | #FEM | $d_h$ | $w^*(kg)$ | #iter | #FEM |
| 5 | 2.56 | 0 | 2.56 | 10 | 1004 | 0 | 2.57 | 32 | 32300 | 0 | 2.56 | 2 | 400 |
| 10 | 6.06 | 0 | 6.06 | 26 | 3097 | 1 | 6.14 | 54 | 54500 | 0 | 6.06 | 2 | 792 |
| 15 | 10.23 | 0 | 10.23 | 95 | 10907 | 2 | 10.27 | 65 | 65200 | 0 | 10.23 | 4 | 1955 |
| 20 | * | * | 15.33 | 135 | 10315 | * | 15.59 | 73 | 73100 | * | 15.33 | 2 | 1659 |
| 25 | * | * | 21.36 | 1199 | 610347 | * | 22.06 | 98 | 97700 | * | 21.36 | 3 | 3142 |
| 30 | * | * | 28,30 | 4432 | 723388 | * | 28.84 | 129 | 128800 | * | 28.30 | 8 | 10522 |
| 35 | * | * | $36,17^{(*)}$ | $5793^{(*)}$ | $1096968^{(*)}$ | * | 37.00 | 189 | 189400 | * | 36.19 | 3 | 5830 |
| 40 | * | * | $44,97^{(*)}$ | $5570^{(*)}$ | $939726^{(*)}$ | * | 45.64 | 270 | 269800 | * | 44.97 | 7 | 13577 |
| 45 | * | * | $54,70^{(*)}$ | $4181^{(*)}$ | $818455^{(*)}$ | * | 55.98 | 347 | 346800 | * | 54.71 | 4 | 8531 |
| 50 | * | * | $65,35^{(*)}$ | $4316^{(*)}$ | $717627^{(*)}$ | * | 67.48 | 561 | 561200 | * | 65.34 | 6 | 14487 |

Table 3: A comparison of the obtained solutions for 10 instances of the scalable 2D cantilever problem are compared, with a varying number of bars (from 5 to 50 bars). We note that when optimizations last more than 24 hours, the solver (Baseline, h-B&B) is stopped and the current solution (if exists) is marked by $(*)$.

the number needed by the compared approaches. The trends in terms of computational cost with respect to the number of elements are graphically represented in Fig. 7. The cost of the Genetic algorithm dominates the cost of h-B&B and Bi-level. The scaling of the Bi-level approach is nearly linear when compared to the h-B&B and Genetic approach. The observed efficiency makes the proposed approach relevant for higher dimensional problems.

### 4.5 120-bar truss

In this example, the structure of a 120-bar dome truss (Saka and Ulker, 1992) is considered and described in Fig. 8. In this case, $n = 120$ and $p = 4$, $\Gamma = \{1, 2, 3, 4\}$ There is no grouping of elements, meaning that the design space counts 120 categorical design variables and 120 continuous design variables. For each element, the categorical variable can take a value among 4 catalogs, that point to combinations of I and C-profiles with materials AL2139 and AL2024. Materials properties are listed in Appendix A. The structure is subjected to an active constraint on displacements : a maximum downward displacement of 10 $mm$ is allowed. A downward load of 60 $kN$ is applied on node 1, while a downward load of 30 $kN$ is applied on nodes 2 to 13 and 10 $kN$ on nodes 14 to 37.

The graphical solution obtained by the Bi-level algorithm is displayed on Fig. 9. To each categorical choice is associated a color on the structure. The continuous variables are qualitatively illustrated by the size of each truss element. The solution found by the algo-

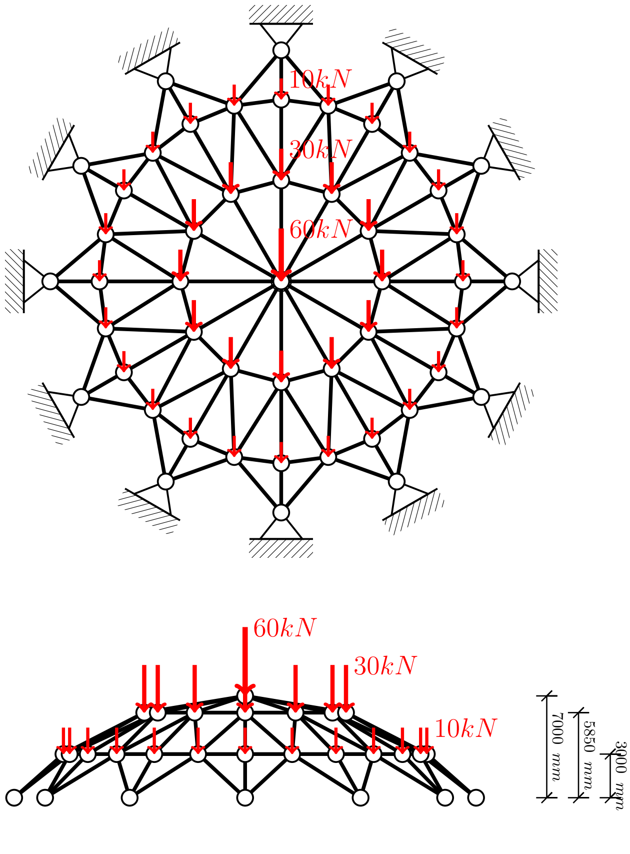



Fig. 8: Top and side view a 120-bar truss structure. Downward loads with three different magnitudes are applied.

rithm shows that a number of two catalogs in $\Gamma$ have been selected. A list of optimal choices and areas for each element is given in Appendix B. On elements 1 to 12 and 25 to 48, the catalog 4 is the optimal one : AL2024 with C-profile. For the other elements, the optimal choice is catalog 3 : AL2024 with I-profile. Thus for the entire truss, the stiffest and lightest material has been selected. Euler buckling constraints applied to elements 49 to 96 are active. The optimal choice for these elements is thus confirmed by the fact that the area moment of inertia of the I-profile is higher than the C-profile. The material with the highest Young modulus has been selected for the entire truss. This is in accordance with the fact that the other active constraint is the global constraint on displacements. The Genetic algorithm has been applied on this case with settings adapted to the problem scale. However, it was not able to find a feasible region.

The obtained results on the 120-bar truss example showed the ability of the proposed method in handling large scale instances of the problem (P). Through this 120-bar truss example and also the scalable test-cases from Section 4.4, we present a methodology that offers an interesting compromise between the computation cost and the quality of the targeted results. However, the impact of neglecting the information on the coupling of the categorical choices while using the first order-like approximation has not been formally proven yet. The consequences of such approximation need further investigation. The decrease strategy (see Section 3.4) is included to ensure a weight decrease when the first order like approximation is not capable of that. We note that, for instance, during the optimization process of the 120-bar truss, the decrease strategy was not activated which suggest the efficiency of the proposed first order-like approximation for the master problem.

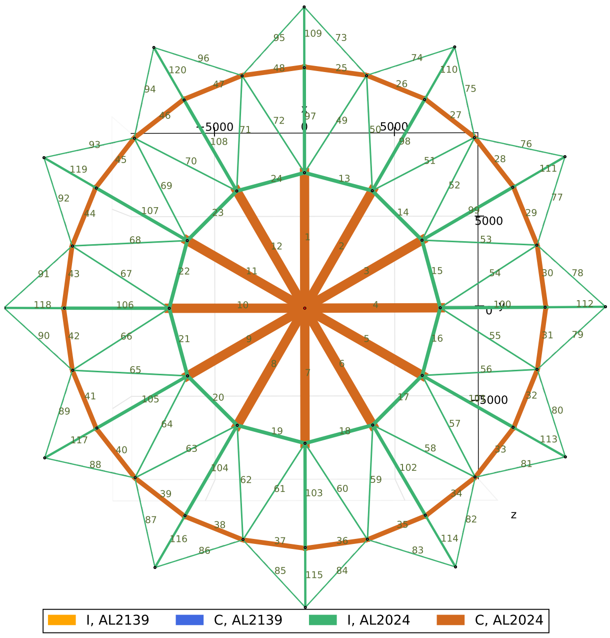



Fig. 9: Top view of the 120-bar truss mixed categorical-continuous optimization result.

## 5 Conclusions

This paper proposed an efficient heuristic algorithm to handle large scale categorical-continuous structural weight minimization problems subject to displacements constraints. The proposed methodology consisted of using a bi-level decomposition involving two problems: master and slave. The master problem was driven by a first order like-approximation, this made it possible to reduce drastically the combinatorial exploration cost raised by the categorical design space. Once the categorical decisions are driven by the master problem, the continuous variables are handled by the slave problem using a gradient-based approach. Using the proposed implementation, one was able to find the exact solutions on low-dimensional cases. Furthermore, on larger test cases, the scaling of our method revealed to be quasi-linear with respect to the number of structural elements. Particularly, our approach allowed us to solve problems that are very hard to solve with standard algorithms. It enables to perform an optimal catalog selection with optimal internal load distribution and subject to stiffness constraints (here illustrated by displacements constraints). This was not the case for the simplified approach mentioned in the introduction (Grihon, 2018). The scaling with respect to the number of catalogs still needs to be investigated. The proposed methodology offers thus an interesting compromise between the quality of the results, the computational effort and the ease of implementation. It is worth to note that, unlike methods using continuous formulation (Stegmann and Lund, 2005; Krogh et al., 2017), the proposed methodology guarantees to retrieve specific catalog selections. Further work will consist of formulating the problem as a multi-objective one minimizing both weight and cost of composite structures.

## 6 Replication of results

This section is intended to help readers to replicate the results provided in this paper. A supplementary material allows to replicate the 3-bar truss example detailed in section 4.2. In order to replicate the scalable

2D cantilever (section 4.4) and 120-bar truss (section 4.5), the reader needs to adapt the 3-bar truss physical model consequently. The geometries are depicted on Fig. 6 and Fig. 8, material data is provided in Table 4, and the solution of the 120-bar truss is given in Table 5.

## A Materials definition

| | AL2139 | AL2024 | TA6V |
|---|---|---|---|
| Density ($kg/mm^3$) | $2.8\ 10^{-6}$ | $2.77\ 10^{-6}$ | $4.43\ 10^{-6}$ |
| Young modulus ($MPa$) | $7.1\ 10^4$ | $7.4\ 10^4$ | $11.0\ 10^4$ |
| Poisson coefficient ($-$) | 0.3 | 0.33 | 0.33 |
| Tension allow. ($MPa$) | $1.5\ 10^2$ | $1.6\ 10^2$ | $11.0\ 10^2$ |
| Compression allow. ($MPa$) | $2.0\ 10^2$ | $2.1\ 10^2$ | $8.6\ 10^2$ |

Table 4: Numerical details on materials attributes.

## B 120-bar solution

| bars | **a** $[mm^2]$ | **c** $[-]$ |
|---|---|---|
| 1-12 | 4553,6 | 4 |
| 13-24 | 1800,1 | 3 |
| 25-48 | 2313,0 | 4 |
| 49-72 | 745,6 | 3 |
| 73-96 | 589,9 | 3 |
| 97-108 | 1609,0 | 3 |
| 109-120 | 1109,2 | 3 |

Table 5: Solution of 120-bar truss mixed categorical-continuous optimization.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## Acknowledgments

This work is part of the MDA-MDO project of the French Institute of Technology IRT Saint Exupery. We wish to acknowledge the PIA framework (CGI, ANR) and the project industrial members for their support, financial backing and/or own knowledge: Airbus, Altran Technologies, Capgemini DEMS France and CERFACS. The support of François Gallard (IRT Saint Exupery) for his scientific and technical assistance on multi-level schemes implementation is gratefully acknowledged.

## References

Abramson MA, Audet C, Chrissis JW, Walston JG (2009). Mesh adaptive direct search algorithms for mixed variable optimization. Optimization Letters 3(1):35–47

Allaire G, Delgado G (2015). Stacking sequence and shape optimization of laminated composite plates via a level-set method. arXiv :1–33

Audet C, Dennis JE (2001). Pattern Search Algorithms for Mixed Variable Programming. SIAM Journal on Optimization 11(3):573–594

Audet C, Kokkolaras M, Le Digabel S, Talgorn B (2018). Order-based error for managing ensembles of surrogates in mesh adaptive direct search. Journal of Global Optimization 70(3):645–675

Barjhoux PJ, Diouane Y, Grihon S, Bettebghor D, Morlier J (2017). Mixed Variable Structural Optimization: toward an Efficient Hybrid Algorithm. In: WCSMO12

Benders JF (1962). Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik 4(1):238–252

Bettebghor D, Bartoli N, Grihon S, Morlier J, Samuelides M (2011). Approche en paramètres de stratification pour l'optimisation biniveau de structures composites. In: 10e colloque national en calcul des structures. Giens, France

Bettebghor D, Grihon S, Morlier J (2018). Bilevel optimization of large composite structures based on lamination parameters and post-optimal sensitivities. part 1: Theoretical aspects

Deb K, Goyal M (1998). A Flexible Optimization Procedure for Mechanical Component Design Based on Genetic Adaptive Search. Journal of Mechanical Design 120(2):162

Duran MA, Grossmann IE (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical Programming 36(3):307–339

Filomeno Coelho R (2014). Metamodels for mixed variables based on moving least squares: Application to the structural analysis of a rigid frame. Optimization and Engineering 15(2):311–329

Fister I, Yang XS, Brest J, Fister D (2013). A brief review of nature-inspired algorithms for optimization. Elektrotehniski Vestnik/Electrotechnical Review 80(3):116–122

Fortin FA, De Rainville FM, Gardner M, Parizeau M, Gagné C (2012). DEAP : Evolutionary Algorithms Made Easy. Journal of Machine Learning Research :2171–2175

Gallard F, Barjhoux PJ, Olivanti R, Gazaix A (2019). GEMS, an Open-Source Generic Engine for MDO Scenarios : Key Features In Application. In: AIAA Aviation 2019

Gallard F, Vanaret C, Guenot D, Gachelin V, Lafage R, Pauwels B, Barjhoux PJ, Gazaix A (2018). GEMS: A Python Library for Automation of Multidisciplinary Design Optimization Process Generation. In: 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. American Institute of Aeronautics and Astronautics, Reston, Virginia

Gao H, Breitkopf P, Coelho RF, Xiao M (2018). Categorical structural optimization using discrete manifold learning approach and custom-built evolutionary operators. Structural and Multidisciplinary Optimization 58(1):215–228

Garrido-Merchán EC, Hernández-Lobato D (2018). Dealing with Categorical and Integer-valued Variables in Bayesian Optimization with Gaussian Processes :1–18

Gazaix A, Gallard F, Ambert V, Guénot D, Hamadi M, Sarouille P, Grihon S, Druot T, Brezillon J, Thierry L, Bartoli N, Lafage R, Gachelin V, Plakoo J, Desfachelles N, Gurol S, Pauwels B, Vanaret C (2019). Application of an advanced bi-level mdo formulation to an aircraft engine pylon optimization. American Institute of Aeronautics and Astronautics, Dallas, Texas

Geoffrion AM (1972). Generalized Benders decomposition. Journal of Optimization Theory and Applications 10(4):237–260

Goldberg DE (1989). Genetic Algorithms in Search, Optimization, and Machine Learning, volume Addison-We. Addison-Wesley Longman Publishing Co., Inc.

Grihon S (2012). PRESTO: A rapid sizing tool for airframe conceptual design studies. In: LMS European Aeronautical Conference. Toulouse

Grihon S (2018). Structure sizing optimization capabilities at airbus. In: A Schumacher, T Vietor, S Fiebig, KU Bletzinger, K Maute, editors, Advances in Structural and Multidisciplinary Optimization. Springer International Publishing, Cham, 719–737

Haftka RT, Gürdal Z (1992). Elements of Structural Optimization, volume 11 of Solid Mechanics And Its Applications. Springer Netherlands, Dordrecht

Haftka RT, Watson LT (2005). Multidisciplinary Design Optimization with Quasiseparable Subsystems. Optimization and Engineering 6:9–20

Haftka RT, Watson LT, Haftka RT, Watson LT (2006). Decomposition theory for multidisciplinary design optimization problems with mixed integer quasiseparable subsystems. Optim Eng 7(7):135–149

Herrera M, Guglielmetti A, Xiao M, Filomeno Coelho R (2014). Metamodel-assisted optimization based on multiple kernel regression for mixed variables. Structural and Multidisciplinary Optimization 49(6):979–991

Hijazi H, Bonami P, Ouorou A (2014). An Outer-Inner Approximation for Separable Mixed-Integer Nonlinear Programs. INFORMS Journal on Computing 26(1):31–44

Ivanov VV, Zadiraka VK (1975). Numerical optimization. Cybernetics 9(4):714–715

Johnson SG (2008). The nlopt nonlinear-optimization package

Krogh C, Jungersen MH, Lund E, Lindgaard E (2017). Gradient-based selection of cross sections: a novel approach for optimal frame structure design. Structural and Multidisciplinary Optimization 56(5):959–972

Liao T, Socha K, Montes de Oca MA, Stutzle T, Dorigo M (2014). Ant Colony Optimization for Mixed-Variable Optimization Problems. IEEE Transactions on Evolutionary Computation 18(4):503–518

Lindroth P, Patriksson M (2011). Pure Categorical Optimization: a Global Descent Approach. Technical report, Chalmers University of Technology

Merval A (2008). L'Optimisation Multiniveaux D'une Structure. PhD Thesis

Müller J, Shoemaker CA, Piché R (2013). SO-MI: A surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems. Computers & Operations Research 40(5):1383–1400

Nouaouria N, Boukadoum M (2011). A Particle Swarm Optimization approach to mixed attribute data-set classification. In: 2011 IEEE Symposium on Swarm Intelligence. IEEE, 1–8

Pelamatti J, Brevault L, Balesdent M, Talbi EG, Guerin Y (2019). Efficient global optimization of constrained mixed variable problems. Journal of Global Optimization 73(3):583–613

Roy S, Crossley WA, Stanford B, Moore KT, Gray JS (2019). A mixed integer efficient global optimization algorithm with multiple infill strategy - applied to a wing topology optimization problem. In: AIAA Scitech 2019 Forum

Roy S, Moorey K, Hwang J, Gray J, Crossley W, Martins J (2017). A mixed integer efficient global optimization algorithm for the simultaneous aircraft allocation-mission-design problem. 58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2017 :1–18

Saka M, Ulker M (1992). Optimum design of geometrically nonlinear space trusses. Computers & Structures 42(3):289 – 299

Samuelides M, Bettebghor D, Grihon S, Merval A, Morlier J (2009). Modèles réduits en optimisation multiniveau de structures aéronautiques. In: 9e Colloque national en calcul des structures. CSMA, Giens, France

Schutte JF, Haftka RT, Watson LT (2004). Decomposition and Two-level Optimization of Structures with Discrete Sizing Variables. Constraints 0(1):1–5

Shahabsafa M, Mohammad-Nezhad A, Terlaky T, Zuluaga L, He S, Hwang JT, Martins JRRA (2018). A novel approach to discrete truss design problems using mixed integer neighborhood search. Structural and Multidisciplinary Optimization 58(6):2411–2429

Sigmund O (2011). On the usefulness of non-gradient approaches in topology optimization. Structural and Multidisciplinary Optimization 43(5):589–596

Stegmann J, Lund E (2005). Discrete material optimization of general composite shell structures. International Journal for Numerical Methods in Engineering 62(14):2009–2027

Stolpe M (2011). To bee or not to bee—comments on “Discrete optimum design of truss structures using artificial bee colony algorithm”. Structural and Multidisciplinary Optimization 44(5):707–711

Svanberg K (2002). A class of globally convergent optimization methods based on conservative convex separable approximations. SIAM Journal on Optimization 12(2):555–573

Turner M (1959). The Direct Stiffness Method of Structural Analysis. In: Structural and Materials Panel Paper, AGARD Meeting. Aachen

Turner MJ, Martin HC, Leible RC (1964). Further Development and Applications of Stiffness Methods. In: Matrix Methods of Structural Analysis, Macmillian, New York, volume 1. 1st edition, 203–266