



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/236639/>

Version: Accepted Version

---

**Article:**

Fairclough, H.E., Bołbotowski, K., He, L. et al. (2026) Topology and geometry optimization of grid-shells incorporating design-dependent self-weight loading. *Computers & Structures*, 320. 107997. ISSN: 0045-7949

<https://doi.org/10.1016/j.compstruc.2025.107997>

---

© 2025 The Authors. Except as otherwise noted, this author-accepted version of a journal article published in *Computers & Structures* is made available via the University of Sheffield Research Publications and Copyright Policy under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Topology and geometry optimization of grid-shells incorporating design-dependent self-weight loading

Helen E. Fairclough<sup>a</sup>, Karol Bołbotowski<sup>c,d</sup>, Linwei He<sup>a</sup>, Andrew Liew<sup>b</sup>, Matthew Gilbert<sup>a</sup>

<sup>a</sup>*School of Mechanical, Aerospace and Civil Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK*

<sup>b</sup>*Unipart Construction Technologies Ltd, Advanced Manufacturing Park, Sheffield, S60 5WG, UK*

<sup>c</sup>*Department of Structural Mechanics and Computer Aided Engineering, Faculty of Civil Engineering, Warsaw University of Technology, 16 Armii Ludowej Street, Warsaw, 00-637, Poland*

<sup>d</sup>*Lagrange Mathematics and Computing Research Center, 103 rue de Grenelle, Paris, 75007, France*

---

## Abstract

This manuscript presents the first approach which is capable of finding the optimal connectivity and elevation of grid-shell structures acting in pure compression (or pure tension) for long-span scenarios; i.e. under the combined effects of a prescribed external loading and the design-dependent self-weight of the structure itself. The method derived herein involves solving a second-order cone optimization problem, thereby ensuring convexity and obtaining globally optimal results for a given discretization of the design domain. Several numerical examples are presented, illustrating hitherto unknown or unconfirmed characteristics of this class of optimal structures. For example, it is found that, as self-weight becomes more significant, both the optimal topology and the optimal elevation profile of the structure change, highlighting the importance of optimizing both topology and geometry simultaneously from the earliest stages of design. It is shown that this approach can obtain solutions with greater accuracy and several orders of magnitude more quickly than a standard 3D layout/truss topology optimization approach. Python scripts and Rhino/Grasshopper files are provided for the examples shown.

*Keywords:* Ground structure method, Form-finding, Grid-shell,

---

## 1. Introduction

Grid-shells can be used as lightweight and elegant structures, which are especially suited to long-span roofs and similar structures. The term grid-shell was originally applied to timber structures constructed by deforming an initially flat network of thin elements [1]. However, it has since come to be used for any shell structure constructed from discrete elements [2]; it is the latter definition which is adopted herein.

### 1.1. Current methods for form-finding and optimization

To ensure an axially loaded design, a wide variety of form-finding approaches have been developed. These typically correspond to problems which would, in

other structural optimization communities, be referred to as size and/or shape optimization. This means that the connectivity or topology of the design is fixed, usually by specifying an initial structure. Such methods include particle-spring approaches, dynamic relaxation and force-density methods, as well as physical modeling approaches [2].

A key issue in applying any of these approaches to long-span structures is the ability to model the weight of the structure itself. The simplest and most common approximation is to size the (straight) element based only on the axial load, and then apply a force equal to half the element's weight at each end [3, 4]. This approximation ignores the bending effects of the weight, which may be substantial, especially in longer elements. A refinement of this approach is to divide each element into multiple shorter sub-elements [5]. A more realistic approach is to distribute the loading along the element, typically resulting in parabolic forms, based on the assumption of constant cross-sections [6].

Regardless of the self-weight model used, these form-finding approaches typically do not alter the initial connectivity of the problem. Within the few studies that attempt to take this into account, genetic algorithms are typically employed [7, 8], these are very flexible in what may be considered, but there is no guarantee that even a locally optimal solution will be obtained. In [9], gradient-based optimization is used, ensuring that the results are at least locally optimal, however the problem was still non-convex and non-linear, so global optimality could not be obtained, and a only a restricted set of possible elements was employed.

Within the broader field of axially loaded structures (e.g. trusses), the optimization of topology is typically achieved through the use of the ground structure method [10], also known as truss topology optimization or layout optimization. Continuum-based topology optimization methods (e.g. based on finite element type meshes) are also available [11], however, these are less suited to structures containing slender elements such as grid-shells, and so will not be discussed here.

Layout optimization requires as input only the geometry of the permissible design domain, the location and nature of the support conditions and the forces to be supported. The design domain is then populated with nodes, and every possible pair of nodes is connected with lines representing the bars of the truss. This network of bar elements forms the ground structure. A mathematical optimization problem is then solved to find the minimum material structure contained within the ground structure that satisfies the loading and support conditions. In the basic case of a static loading regime, this becomes a linear programming problem which can be solved to global optimality rapidly, even for millions of potential elements.

There are no additional challenges involved with applying this method in 3D space, although the number of nodes required to obtain the same nodal spacing will be much larger. The computational demands (in 2D or 3D problems) can be reduced by using the member adding method [12], without impacting the resulting optimal solution. Using the member adding approach, problems containing billions of potential elements can be tackled.

One of the key challenges in applying truss topology optimization approaches in the design of grid-shells or other roof structures, is that the applied loading should move to track the structure as the topology optimization progresses. This can be implemented through the use of transmissible loads [13], where the representative point loads of the grid-shell are transmitted through or attached to vertical lines passing through the nodes. Two formulations of transmissible loads are commonly used, although if only single layer structures are permitted then they are equivalent

[4, 14]. However, to obtain accurate results with either approach, a very dense distribution of nodes must be adopted in the ground structure/domain, particularly in the vertical direction. This can result in computationally expensive problems. Jiang et al. [4] suggest an iterative approach of refining the design domain to allow results to be obtained with fewer nodes, however this was still found to be 3 to 4 orders of magnitude slower than fixed-topology approaches such as the force density method.

For roofs of particularly long spans, a large portion of the loading is generated by the self-weight of the structure itself, as other imposed loads are restricted to e.g. wind and snow loading. The classical method of addressing this within the layout optimization framework is again to add a point load at the end of each element, equal to one half of its weight [15]. As with other form-finding methods, this produces acceptable results in problems of moderate span; but at long spans this erroneously favors solutions containing long elements, maximizing the ‘free’ bending capacity. To address this, Fairclough et al. [16] developed an approach where each element is curved and variable in cross-section such that, at all points, it experiences a uniform, purely axial stress under the combined effect of the applied loads and the self-weight.

In contrast to the numerical methods described above, exact analytical results for minimum material structures (a.k.a. Michell structures) are challenging to obtain, and as such only a few are known [17]. In particular, 3D examples of exact optimal structures have been obtained only for two classes of problems. Firstly, axisymmetric problems [17, Chapter 5], most notably the torsion sphere identified in the seminal work of Michell [18]. Secondly, funicular structures, also known as Prager-structures, which are compression-only (or tension-only) structures, where the point of application of the load is to be optimized alongside the structural form.

For these Prager-structures, significant initial work focused on cases where the elements were restricted to lie along Cartesian directions [19], referred to as archgrids. The methodology involved assuming the two sets of arches could be optimized independently, and then noting that the optimal elevations of each layer were coincident, leading to a single-layer structure (the *equal elevation* condition). This work also identified that for an optimal planar arch under external loading (i.e. without self-weight), the average of the slope value squared must be 1 (the *unit mean square slope* condition).

Later, [20] extended the work on arch-grids to cases where self-weight of the structure was to be considered. It was found that the unit mean square slope condition was no longer valid and must be replaced by a much more complicated equation [20, Eqn 19], which results in a greater rise for a given problem when self-weight is included. Meanwhile, the equal elevation condition continued to hold, demonstrating that a single layer structure is still suitable when self-weight is considered.

Extending the archgrid concept to problems where the layout of elements is also to be considered was achieved by [21] for certain classes of problems. It was also noted that the vertical displacement at any point in such a structure was proportional to the elevation at that point, and that the layouts obtained achieve the minimum material usage for under either a limit on maximum stress or on minimum stiffness. It is notable that consideration of the effects of the structure’s self-weight loading was not included in [21].

Recently, there has been a revival in interest in optimal design of vault and grid-



shell structures, aligning with ambitions for efficient and sustainable construction. The most widely applicable results for the problem of combined topology and geometry optimization have been obtained by [22]. For problems where self-weight is not considered, this shows that the minimum material layout and elevation at each point can be obtained as the solution of a convex optimization problem. An analytical approach was presented, alongside a ground-structure based numerical approach which allowed the use of second-order cone programming to solve the problem. This allows for standard solvers to be used, significantly easing implementation, and as the optimization problems are convex, globally optimal results can be found. However, the self-weight of the structures was not considered, and so the applicability of the problems is limited to shorter spans and/or more lightweight materials. Structures designed for longer spans using this method would become overloaded due to the extra forces generated by their own weight and would be likely to fail.

### *1.2. Research gaps and contributions of this paper*

It has been shown that a wide range of optimization and form-finding methods already exist to support the design of axially loaded structures. However, none of the existing methods provide all of the features necessary to rigorously identify the most material-efficient form for a long-span grid-shell design problem. The requirements for such a method can be summarised as:

- To optimize both the topology and geometry of the grid-shell, allowing for a full exploration of the design space.
- To attain the globally minimum volume for the specified problem.
- To be applicable to long-span problems by accurately modeling the effects of self-weight. (When self-weight is neglected or modeled in a simplified way, the resulting designs will usually be unsafe under the true loads.)
- To be solvable with a practicable level of computational cost.

Existing form-finding and optimization methods, reviewed in Section 1.1, all fail in at least one of these criteria. In contrast, the novel methodology proposed herein satisfies all four of the above requirements, and is thus fully applicable to the problem of long-span grid-shell design.

The approach, summarised in Figure 1, is based on the use of a 2D ground structure where the elevations of each node are not explicitly included in the optimization, but are instead reconstructed from the optimal solution. This concept has never previously been applied in contexts where self-weight loadings need to be considered.

In order to correctly include self-weight within the optimization, each individual element in the ground structure should take the optimal form to resist the combined effect of the loads applied to it. This form is known as the catenary of equal stress, and is defined by the presence of uniform, axial-only stresses under the combination of external loading and self-weight. Appendix A demonstrates that the simpler lumped model, although common in the literature, cannot be used within the proposed framework.

These ideas are combined with the aim of attaining a convex optimization problem in order to allow for globally optimal solutions to be obtained. A convex

formulation is found, but only after some relaxations to the physical constraints. This means that there are feasible solutions to the optimization problem which do not correspond to a physically meaningful structure. However, proofs are also provided herein to show that for *optimal* solutions of the problem, these relaxations are not active. In other words, the globally optimal solution always corresponds to a physical structure.

By using this convex optimization approach, it now becomes possible to obtain globally optimal solutions to the problem of long-span grid-shell optimization. From these unique results, several characteristics of this class of problem are newly identified:

- The optimal topology often changes as span increases.
- Although the optimal elevation profile requires higher *maximum* rise when self-weight is more significant, the optimal elevation at each point within the structure can increase or decrease with increasing span.
- Unlike the weightless scenario, the optimal topology under self-weight may require connections which were not the locations of loads or supports in the problem.

The presented numerical approach is fully flexible in terms of the geometry of the grid-shell footprint (including non-convex domains and holes), the loading to be applied, the support conditions (including full pin and roller pin supports) and the material properties. Scenarios for which there is no possible structure correspond to infeasible optimization problems. Python scripts and Rhino/Grasshopper files are provided to encourage uptake; the approach is sufficiently computationally efficient to allow problems with hundreds of thousands of potential elements to be solved on a laptop computer in a matter of seconds.

The structure of this paper is as follows; Section 2 first discusses two simpler approaches which satisfy only some of the requirements above, but allow for key concepts and notation to be introduced. Section 3 then derives the new procedure for grid-shell optimization under self-weight loading. The new formulation is tested on a range of examples in Section 4 and concluding remarks are given in Section 5.

## 2. Current state-of-the-art

To understand the approach proposed in this paper, it is useful to first consider some simpler scenarios. These are based on existing methods in the literature, with minor additional refinements as needed. Note that the novel optimization problem derived in Section 3 begins its derivation from basic principles; it is outlined within each sub-section here why it is not possible to directly progress from these methods to an approach satisfying all the requirements set out in Section 1.2.

### 2.1. Grid-shell topology and geometry optimization without self-weight

Firstly, the weightless vault formulation of [22] is recalled. For clarity, the discretised formulation is discussed here. This method does not satisfy the requirements in Section 1.2 as it cannot consider self-weight loads and is therefore non-conservative, particularly at longer spans. Furthermore, attempting to directly modify this approach to include self-weight would result in the formulation

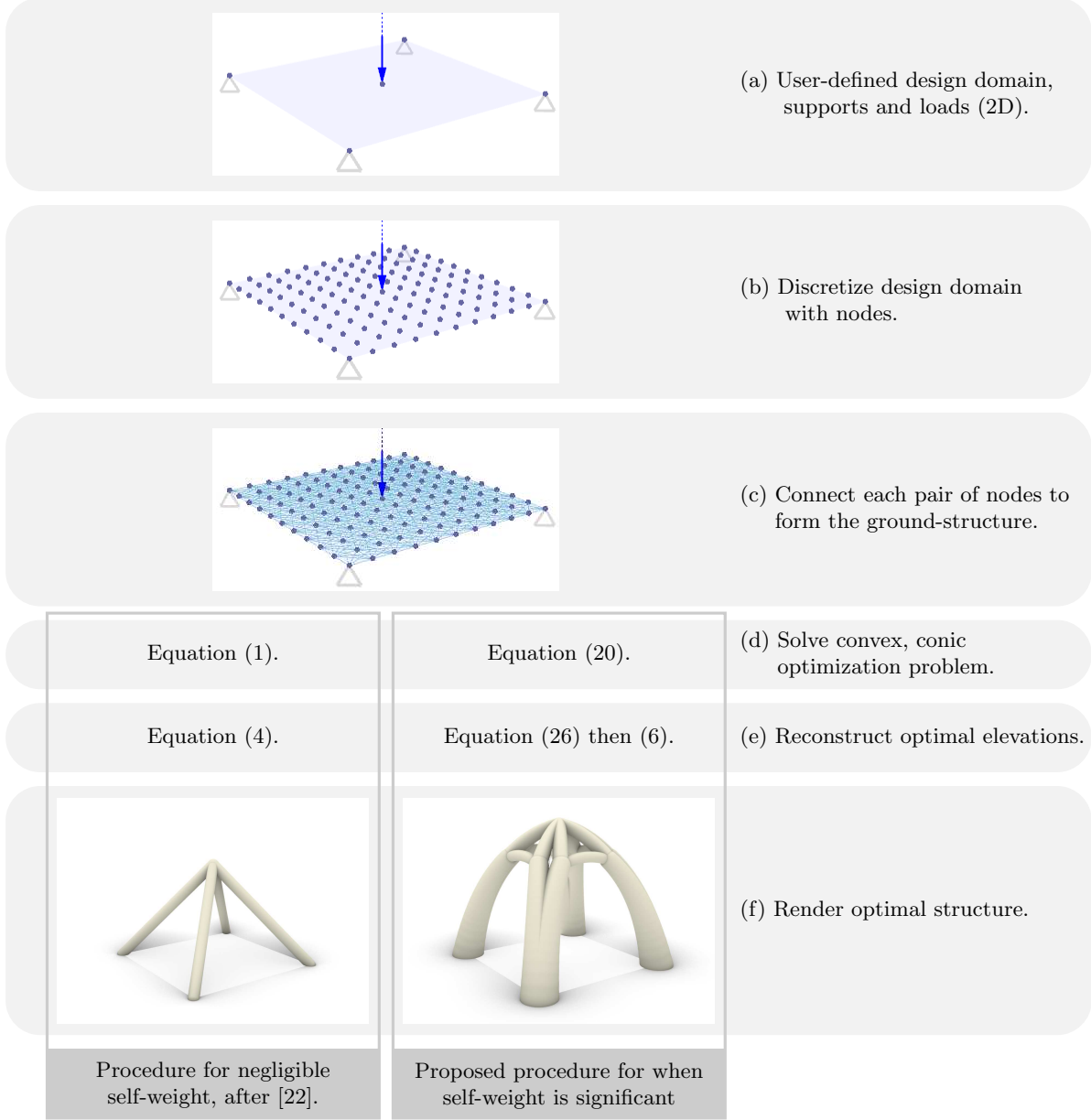


Figure 1: Summary of methodology developed herein, compared to existing approach [22] for cases where self-weight can be neglected. Note that the initial setup process is common to both cases, with the key differences being the formulation of the optimization problem and the reconstruction of the optimal elevations. The procedures for when self-weight is negligible are recapped in Section 2.1, whilst the new procedure including self-weight is derived in Section 3.

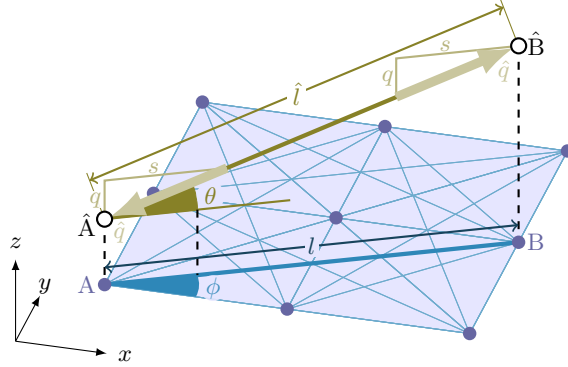


Figure 2: Notation for grid-shell topology and geometry optimization without self-weight. The blue lines show elements in the ground structure defined in the horizontal plane. The highlighted element is also shown (in brown) as it would be in the optimal solution, once the end-points are projected to their optimal elevations for the final structure. The force,  $\hat{q}$ , acting on the nodes is aligned to the final centre-line of the element, i.e. at an angle of  $\theta$  to the horizontal.

in Appendix A which is shown to produce solutions which are not physically realizable. Nonetheless, the existence of this method in the simpler, weightless context inspires confidence in the seemingly unlikely approach to obtaining convexity which will be required in Section 3.1.

In this vault layout optimization approach, the allowable design domain is defined in plan, and locations of applied loads and supports are also specified (Figure 1a). This 2D domain is discretised using a dense array of nodes (Figure 1b), and every pair of nodes is connected by a potential structural element (Figure 1c). Conceptually, nodes are allowed to move in the vertical direction to obtain the optimal shape (except where vertical support is provided at a node), and any applied forces move vertically with the associated node.

When formulating the mathematical optimization problem, for each potential element two optimization variables are required,  $s_i$  and  $q_i$ , representing the horizontal and vertical component of the axial force in element  $i$ , over the set of members  $M$ . The objective function is to minimize the total volume, calculated on a per-element basis and then summed over all elements in the ground-structure. For a single element, as shown in Figure 2 with stress  $\sigma_i$  the volume is given by  $\hat{l} \frac{\hat{q}}{\sigma_i}$ . In the original presentation, the objective function was the normalised volume (i.e. volume divided by the allowable stress), but for more practical cases it is convenient to include  $\sigma$  in the objective calculation directly. Geometrical relationships from Figure 2 can then be used to re-write this in terms of  $s_i$  and  $q_i$ .

The required optimization problem is then given by,

$$\min \sum_{i \in M} \frac{l_i}{\sigma_i} \left( s_i + \frac{q_i^2}{s_i} \right), \quad (1a)$$

$$\text{s.t. } \mathbf{B}\mathbf{s} = \mathbf{f}_{xy}, \quad (1b)$$

$$\mathbf{D}\mathbf{q} = \mathbf{f}_z, \quad (1c)$$

$$\mathbf{s} \geq 0, \quad (1d)$$

where  $l_i$  is the horizontal length of potential structural element  $i$ , and the objective function to be minimized being the summation of all element volumes. The

matrix  $\mathbf{B}$  contains in-plane direction cosines ( $\pm \cos \phi$  and  $\pm \sin \phi$ ) such that (1b) enforces in-plane equilibrium at each node subject to horizontal components of external forces  $\mathbf{f}_{xy}$ . Matrix  $\mathbf{D}$  contains entries of 0, -1 and 1 such that (1c) enforces equilibrium in the vertical direction at each node with vertical components of external forces  $\mathbf{f}_z$ . Finally, the positivity of the  $s$  variables ensures that all elements remain in compression. Conversely, the  $q$  variables may be positive or negative, corresponding to elements with an upwards or downwards slope respectively. Note that, as only a single loading condition is permitted, for practical problems  $\mathbf{f}_{xy}$  is typically zero.

Problem (1) can be transformed into a convex conic problem by assigning an additional variable  $r_i$  to each potential element, which is defined as

$$2r_i \geq \frac{q_i^2}{s_i}. \quad (2)$$

The variables  $r_i$  are then used in the objective function. Note that this transformation is a relaxation of the problem (1), in that the original problem demands equality of (2), however this would not be a convex constraint. However, as the introduction of any slack in (2) would increase the objective without affecting any other constraints, it is easy to see that this relaxation will not be in effect at the optimum solution.

Note that a feasible solution to the formulation (1) does not require that the resulting structure be single layered. However, it was proven in [22] that an *optimal* solution to the problem will permit a consistent, single-layered elevation function across the domain. A more intuitive demonstration for this was also given in [23, Fig 3]. It is possible to obtain the required heights of each node of the gridshell in the final optimal solution by progressively moving along each non-zero element in the solution, calculating its slope by the ratio between  $q_i$  and  $s_i$  [23]. However, the elevations can also be obtained directly from the dual of problem (1). The dual problem can be derived through kinematic principles (or obtained through conic duality principles [24]) and is written as,

$$\max_{\mathbf{u}, \mathbf{w}, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3} \mathbf{f}_z^T \mathbf{w} + \mathbf{f}_{xy}^T \mathbf{u} \quad (3a)$$

$$\text{s.t. } \mathbf{B}^T \mathbf{u} + \mathbf{t}_2 = \frac{1}{\sigma} \mathbf{1}, \quad (3b)$$

$$\mathbf{D}^T \mathbf{w} + \mathbf{t}_3 = \mathbf{0}, \quad (3c)$$

$$\mathbf{t}_1 = \frac{2}{\sigma} \mathbf{1}, \quad (3d)$$

$$(2t_{1,i}t_{2,i} \geq t_{3,i}^2) \forall i \in M. \quad (3e)$$

In this formulation, the variables  $\mathbf{u}$  and  $\mathbf{w}$  represent virtual displacements in the horizontal and vertical direction. The optimal elevation,  $z$  at a point is calculated from the virtual strain at the optimum by the relationship

$$z = -\frac{\sigma}{2}w. \quad (4)$$

where the inclusion of  $\sigma$  is due to the use of the real volume as an objective. Conversely, the original formulation scaled the objective function, and therefore found  $w$  directly in units of distance.

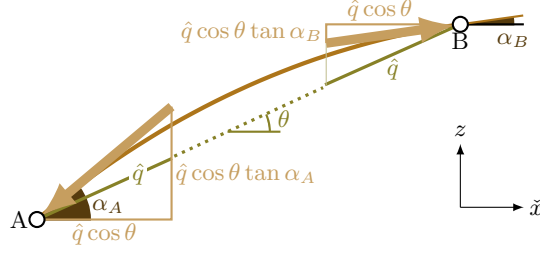


Figure 3: Notation for the problem of truss topology optimization with continuous self-weight, a single element shown in elevation in the vertical plane containing the element. The element number subscript  $i$  is omitted for clarity. The indicated forces acting on the nodes are aligned to the element centreline at the respective point.

However, a structure designed through this approach neglects the impact of the self-weight of the structure itself. This will lead to unsafe and under-designed structures, particularly as span lengths increase.

## 2.2. Truss topology optimization with continuous self-weight

As previously mentioned, it is common in truss optimization problems to model self-weight as forces applied directly onto the endpoints of an element, neglecting the effect within the element itself. Whilst this has been effective as an approximation for trusses of moderate span, it is not possible to use this approach within the framework outlined in Section 2.1. This is discussed in detail in Appendix A. Instead, a model must be applied which correctly accounts for the self-weight in a continuous manner. The catenary of equal stress is the arching/sagging shape which results in uniform axial stresses only under the combined action of self-weight and a thrust force [25]. For a given strength-to-weight ratio of material, the centre-line of this catenary is fixed. Meanwhile, the cross-sections vary along the length of the catenary, to ensure uniformity of stress; but under a change in applied load, the forces and required area at each point increase linearly. These characteristics lead to the possibility for these catenaries to be employed within the truss layout optimization framework, and will also prove to be advantageous in the formulation to be developed herein.

In this section, a catenary truss formulation is given, based on [16], but extended to 3D. To simplify the upcoming adaption to compressive grid-shell problems, it will be assumed that only compressive members are to be permitted, and that the ground structure does not contain any elements which are exactly aligned in the vertical direction (as those require special treatment). Even with these modifications, the approach in this section is not applicable to grid-shell design as the positions of external loads must be specified in advance. Transmissible loading methods e.g. [4, 14] could be incorporated to address this, but the approach then becomes hugely computationally expensive, hence why this has not been well explored in literature or practice.

The 3D case has not been outlined in previous literature, but is obtained with a few modifications. Specifically, each element can be thought of within its own local 2D coordinate system in the plane containing the element and the vertical axis; within this plane the classical 2D equations apply, as shown in Figure 3. The end-forces then have to be returned to the global coordinate system. Vertical

forces are unchanged, but the horizontal force  $\hat{q} \cos \theta$  must be resolved by using the in-plane angle  $\phi$  as defined in Fig. 2 to become  $\hat{q} \cos \theta \sin \phi$  and  $\hat{q} \cos \theta \cos \phi$  in the  $x$  and  $y$  directions respectively.

The resulting optimization problem thus becomes,

$$\min_{\hat{\mathbf{q}}} \sum_{i \in M} \frac{\hat{q}_i \cos \theta_i}{\rho g} (\tan \alpha_{B,i} - \tan \alpha_{A,i}), \quad (5a)$$

$$\text{s.t. } \bar{\mathbf{B}} \hat{\mathbf{q}} = \mathbf{f}_{xy}, \quad (5b)$$

$$\bar{\mathbf{D}} \hat{\mathbf{q}} = \mathbf{f}_z, \quad (5c)$$

$$\hat{\mathbf{q}} \geq \mathbf{0}, \quad (5d)$$

where  $\bar{\mathbf{B}}$  contains direction cosines in the  $x$  and  $y$  directions ( $\pm \cos \theta_i \cos \phi_i$  and  $\pm \cos \theta_i \sin \phi_i$  respectively). The matrix  $\bar{\mathbf{D}}$  contains the vertical coefficients,  $\cos \theta_i \tan \alpha_{A,i}$  and  $\cos \theta_i \tan \alpha_{B,i}$ . The optimization variables, here denoted by  $\hat{\mathbf{q}} = [\hat{q}_1, \hat{q}_2, \dots, \hat{q}_m]^T$ , are defined as the equivalent force directed straight between the points A and B.  $\rho g$  is the unit weight of the material.

To calculate the inclination angles  $\alpha$ , it is required to make reference to the definition of the catenary of equal stress from e.g. [16, Eqn. 2.1] or [25, Art. 453], the centre-line of the element follows a curve which can be expressed in a local 2D element coordinate system  $(\tilde{x}, z)$  as

$$\frac{\rho g}{\sigma} z = \log \left( \cos \left( -\frac{\rho g}{\sigma} \tilde{x} + C_1 \right) \right) + C_2 \quad (6)$$

where  $C_1$  and  $C_2$  are constants which define the translation of the shape in the plane. The allowable stress of the material is represented by  $\sigma$ . Within this equation, the term  $-\frac{\rho g}{\sigma} \tilde{x} + C_1$  gives the inclination angle  $\alpha$  for any point. By defining the end-points to be at  $A = (0, z_A)$  and  $B = (l, z_B)$ , explicit relationships for  $\tan \alpha_A$  and  $\tan \alpha_B$  can be obtained.

$$\tan \alpha_A = -\frac{\cos(\frac{\rho g}{\sigma} l) - \exp(\frac{\rho g}{\sigma}(z_B - z_A))}{\sin(\frac{\rho g}{\sigma} l)}, \quad (7)$$

$$\tan \alpha_B = \frac{\cos(\frac{\rho g}{\sigma} l) - \exp(-\frac{\rho g}{\sigma}(z_B - z_A))}{\sin(\frac{\rho g}{\sigma} l)}. \quad (8)$$

It should be noted that the horizontal length of an element must be less than  $\frac{\pi \sigma}{\rho g}$ , otherwise the element is excluded from the ground structure. Accordingly,  $\sin(\frac{\rho g}{\sigma} l)$  is always a positive number ranging in the open interval  $(0, 1)$ , while  $\cos(\frac{\rho g}{\sigma} l)$  can be negative as it varies in  $(-1, 1)$ .

### 3. Grid-shell topology and geometry optimization with continuously distributed self-weight

In this section, an approach suitable for obtaining the optimal grid-shell subjected to both external loading and its own self-weight is developed. Firstly, in Section 3.1 the relevant optimization problem is constructed and cast as a conic problem, which can be easily solved to global optimality. As some relaxations are required to obtain this form, Section 3.2 describes the procedure for reconstructing the physical structure from the solution of the optimization problem. Specifically, it is shown that the relaxations required do not affect the solution at the optimum points.

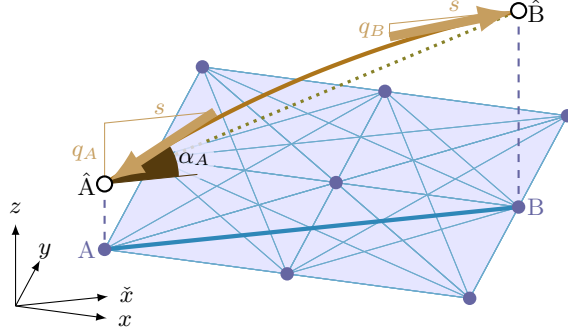


Figure 4: Notation for catenary elements in the vault formulation. The forces acting on the nodes are as shown.  $s$ ,  $q_A$  and  $q_B$  are the corresponding optimization variables.

### 3.1. Formulation of the optimization problem

To formulate the convex optimization problem, this subsection will first set out the required optimization variables (Section 3.1.1), and then derive the required objective function (Section 3.1.2) and constraints (Sections 3.1.3 and 3.1.4). The required conic optimization problem can then be written (Section 3.1.5), and strategies to speed up solving discussed (Section 3.1.6).

#### 3.1.1. Optimization variables

The optimization problem formulated in this section will model the structure's self-weight by employing elements in the shape of the catenary of equal stress. The key difference between these elements and straight elements is that the vertical forces at each end of an element are no longer equal. Thus, the required optimization variables for the long-span grid-shell problem will be:

- $\mathbf{s} = [s_1, s_2, \dots, s_m]^T$ , variables representing the horizontal component of force in each element. These variables are non-negative, with positive values representing forces acting in a compressive manner.
- $\mathbf{q}_A = [q_{A,1}, q_{A,2}, \dots, q_{A,m}]^T$ , variables representing vertical component of the force at the start of each element. Positive values represent downwards force, i.e. Fig 4 shows a positive value for  $q_{A,i}$ .
- $\mathbf{q}_B = [q_{B,1}, q_{B,2}, \dots, q_{B,m}]^T$ , variables representing vertical component of the force at the end of each element. Positive values represent downwards force, i.e. Fig 4 shows a negative value for  $q_{B,i}$ .

These variables are shown visually for a single element in Figure 4. Note that quantities defined in Figures 2 and 3 are also required for this formulation. In particular,  $l$  refers to the horizontal length of the element as in Figure 2, and  $\alpha$ ,  $\theta$  and  $\hat{q}$  are as defined in Figure 3.

Conceptually, the objective function and constraints are analogous to those of the previous formulations. Note that in problems (1) and (5) the objectives and constraints address the same physical concerns – to minimize structural volume, whilst enforcing equilibrium horizontally and vertically. This is also the overall structure of the formulation derived herein. As in (1), the nodal elevations  $z$  will



not be explicitly present in the formulation as design variables. Instead, they will be cleverly recast after solving the optimization problem to be put forth, see Section 3.2.2 below for details.

In the remainder of this section, the subscript indicating element number will be dropped for clarity.

### 3.1.2. Objective function

In the catenary self-weight formulation in Section 2.2, the horizontal component of the element force was defined based on the inline force  $\hat{q}$ . Converting this to the variables required in the grid-shell formulation necessitates the relationship,

$$s = \hat{q} \cos \theta. \quad (9)$$

From Fig. 4, the following relationships can be observed,

$$\frac{q_A}{s} = \tan \alpha_A, \quad (10)$$

$$-\frac{q_B}{s} = \tan \alpha_B. \quad (11)$$

Combining (9), (10) and (11) with (5a) gives the following expression for the volume of a single bar,  $V_i$ ,

$$\begin{aligned} V &= \frac{s}{\rho g} \left( \frac{q_B}{s} + \frac{q_A}{s} \right), \\ &= \frac{1}{\rho g} (q_B + q_A). \end{aligned} \quad (12)$$

This can also be intuitively understood as the total downwards force from the element (i.e. its weight) divided by the unit weight to give the volume.

### 3.1.3. Equilibrium

Equilibrium is enforced in a similar way to the original vault formulation, in that in-plane and out-of-plane equilibrium are considered separately. In-plane equilibrium is enforced by the constraints

$$\mathbf{B}\mathbf{s} = \mathbf{f}_{xy}, \quad (13)$$

where  $\mathbf{B}$  contains entries of  $\pm \cos \phi$  and  $\pm \sin \phi$ .

For out-of-plane equilibrium, changes are needed to allow for the presence of separate force variables in  $\mathbf{q}_A$  and  $\mathbf{q}_B$ .

$$\mathbf{D}_A \mathbf{q}_A + \mathbf{D}_B \mathbf{q}_B = \mathbf{f}_z, \quad (14)$$

where  $\mathbf{D}_A$  (resp.  $\mathbf{D}_B$ ) contains all zeros except in position  $ij$  where node  $j$  is the start (resp. end) node of element  $i$  where it contains 1.

### 3.1.4. Geometrical coupling and its relaxation to a conic constraint

The values of  $s$ ,  $q_A$  and  $q_B$  should be such that the total forces at A and B are aligned to the centreline of a catenary of equal stress connecting the points A and B whose elevations are equal to  $z_A$  and  $z_B$ , respectively, cf. Figure 4. This leads to a coupling condition for the vectors  $\mathbf{s}$ ,  $\mathbf{q}_A$ ,  $\mathbf{q}_B$ .

In the case of the grid-shells without the self-weight this coupling was simple:  $q_A = -q_B = s \frac{z_B - z_A}{l}$ . For an equally stressed catenary the relations are more involved, and they can be obtained by combining the two groups of equations for the tangents  $\tan \alpha_A$ ,  $\tan \alpha_B$ : (7), (8) and (10), (11). These lead to,

$$q_A = -s \frac{\cos(\frac{\rho g}{\sigma} l) - \exp(\frac{\rho g}{\sigma}(z_B - z_A))}{\sin(\frac{\rho g}{\sigma} l)}, \quad (15)$$

$$q_B = -s \frac{\cos(\frac{\rho g}{\sigma} l) - \exp(-\frac{\rho g}{\sigma}(z_B - z_A))}{\sin(\frac{\rho g}{\sigma} l)}. \quad (16)$$

It is crucial to observe that the elevation differences  $z_B - z_A = z_{B,i} - z_{A,i}$  cannot vary arbitrarily from element to element. In fact, if  $\mathbf{z}$  is the vector of nodal elevations, then  $z_{A,i}$ ,  $z_{B,i}$  are the  $i$ -th entries of, respectively,  $\mathbf{D}_A^T \mathbf{z}$ ,  $\mathbf{D}_B^T \mathbf{z}$ . Accordingly, (15), (16) not only entail local couplings between  $s_i$ ,  $q_{A,i}$ ,  $q_{B,i}$  for each element  $i$ , but also a global compatibility condition for the vectors  $\mathbf{s}$ ,  $\mathbf{q}_A$ ,  $\mathbf{q}_B$ .

A simple and natural fix would be to add the nodal elevations  $\mathbf{z}$  as variables and explicitly enforce (15), (16). This, however, would lead to a non-convex optimization problem and, as a result, to compromising the efficiency of the method. Instead, following the idea of [22] for grid-shells without self-weight, conditions (15), (16) will undergo a suitable relaxation.

The relaxation will be performed in two steps. The first consists in eliminating the slopes  $z_B - z_A$ . A slight rewriting of (15), (16) and putting  $\bar{l} = \frac{\rho g}{\sigma} l$  for brevity yields

$$\sin \bar{l} q_A + \cos \bar{l} s = s \exp\left(\frac{\rho g}{\sigma}(z_B - z_A)\right), \quad (17a)$$

$$\sin \bar{l} q_B + \cos \bar{l} s = s \exp\left(-\frac{\rho g}{\sigma}(z_B - z_A)\right). \quad (17b)$$

Side-wise multiplication of the two equalities eliminates the  $z$  values and leads to

$$(\sin \bar{l} q_A + \cos \bar{l} s)(\sin \bar{l} q_B + \cos \bar{l} s) = s^2. \quad (18)$$

The above is a local coupling; when  $s > 0$ , it gives the condition for the triple  $s, q_A, q_B$ , which corresponds to the shape of an equally stressed catenary *for some* slope  $z_B - z_A$ , see Figure 4. The global compatibility with an elevation vector  $\mathbf{z}$  is now forgotten (although it will be shown in Section 3.2 that this global compatibility can be later recovered for optimal solutions).

Adding the equation (18) would still deprive the formulation of convexity. It must be further relaxed to an inequality,

$$(\sin \bar{l} q_A + \cos \bar{l} s)(\sin \bar{l} q_B + \cos \bar{l} s) \geq s^2. \quad (19)$$

Indeed, the latter can be written as the standard rotated conic quadratic constraint:  $2t_1 t_2 \geq t_3^2$ , where  $t_1, t_2$  are the bracketed terms on the left-hand side and  $t_3 = \sqrt{2}s$ . It is a convex constraint that can be efficiently tackled by modern convex optimization software. It is worth emphasizing that imposing the conic constraint  $2t_1 t_2 \geq t_3^2$  canonically comes with the inequalities  $t_1 \geq 0$ ,  $t_2 \geq 0$ . These are valid here: the two factors on the left-hand side of (19) are always non-negative, which can be seen from (17).

### 3.1.5. Complete formulation

Combining the equations found in Sections 3.1.2, 3.1.3 and 3.1.4 gives the following conic optimization problem,

$$\min_{\mathbf{s}, \mathbf{q}_A, \mathbf{q}_B} \frac{1}{\rho g} \mathbf{1}^T (\mathbf{q}_A + \mathbf{q}_B), \quad (20a)$$

$$\text{s.t.: } \mathbf{B}\mathbf{s} = \mathbf{f}_{xy}, \quad (20b)$$

$$\mathbf{D}_A \mathbf{q}_A + \mathbf{D}_B \mathbf{q}_B = \mathbf{f}_z, \quad (20c)$$

$$(\sin \bar{l} q_A + \cos \bar{l} s)(\sin \bar{l} q_B + \cos \bar{l} s) \geq s^2, \quad (20d)$$

$$\mathbf{s} \geq 0, \quad (20e)$$

where the conic constraint is repeated for each potential element, while  $\bar{l} = \frac{\rho g}{\sigma} l_i$ . Recall that the canonical form of the conic constraint implies the two extra inequalities:  $\sin \bar{l} q_A + \cos \bar{l} s \geq 0$  and  $\sin \bar{l} q_B + \cos \bar{l} s \geq 0$ .

The dual of this problem can be derived using the classical systematic methods, see [24],

$$\max_{\mathbf{u}, \mathbf{w}, \mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3} \mathbf{f}_z^T \mathbf{w} + \mathbf{f}_{xy}^T \mathbf{u}, \quad (21a)$$

$$\text{s.t.: } \mathbf{B}^T \mathbf{u} + \mathbf{C}\mathbf{g}_1 + \mathbf{C}\mathbf{g}_2 + \sqrt{2}\mathbf{g}_3 \leq \mathbf{0}, \quad (21b)$$

$$\mathbf{D}_A^T \mathbf{w} + \mathbf{S}\mathbf{g}_1 = \frac{1}{\rho g} \mathbf{1}, \quad (21c)$$

$$\mathbf{D}_B^T \mathbf{w} + \mathbf{S}\mathbf{g}_2 = \frac{1}{\rho g} \mathbf{1}, \quad (21d)$$

$$2 \mathbf{g}_1 \circ \mathbf{g}_2 \geq \mathbf{g}_3^{\circ 2}, \quad (21e)$$

where  $\mathbf{x} \circ \mathbf{y}$  represents element-wise product of vectors, i.e. its  $i$ -th element is  $x_i y_i$ , and  $\mathbf{x}^{\circ n}$  represents element-wise power, i.e. each element of vector  $\mathbf{x}$  is raised to the power  $n$ . Meanwhile  $\mathbf{S} = \text{diag}([\sin \bar{l}_1, \sin \bar{l}_2, \dots, \sin \bar{l}_m])$ , and matrix  $\mathbf{C} = \text{diag}([\cos \bar{l}_1, \cos \bar{l}_2, \dots, \cos \bar{l}_m])$ . The optimization variables  $\mathbf{w} = [w_1, w_2, \dots, w_n]^T$  and  $\mathbf{u} = [u_1^x, u_1^y, u_2^x, \dots, u_n^y]^T$  contain virtual displacements in the vertical and horizontal directions respectively. The remaining optimization variables  $\mathbf{g}_1$  to  $\mathbf{g}_3 \in \mathbb{R}^m$  are auxiliary variables with no specific physical meaning, which are dual to each of the terms in (20d).

Note that, in practice, it is more convenient to set-up and solve the primal problem (20). However, consideration of various aspects of the dual will be necessary below. When using modern conic solvers, it is typically easy to obtain the solutions for both problems after solving either one.

### 3.1.6. Member adding

Member adding [12] is a specialization of column generation which can be used to significantly improve the speed and memory usage of a ground structure optimization problem, whilst guaranteeing that the same globally optimal volume will be obtained.

The procedure begins with an active ground structure consisting of a subset of all possible connections (adjacent connectivity is usually employed). Once the initial problem is solved, the constraints in the dual problem (21) can be checked for inactive elements, i.e. those which were not included in the initial ground

structure. Formally, for each element it is necessary to obtain a  $g_1, g_2, g_3$  satisfying the relevant row of each constraint (21b-e). Candidate values for  $g_1, g_2, g_3$  are obtained by solving the first three constraints, taking (21b) as an equality to give a system which can be solved by simple rearrangement. These candidate values are then checked within the conic constraint (21e), and the element is thus categorized as violating or not violating.

If no elements were found to be violated, i.e. (21e) was satisfied for each set of  $g$ , then the obtained values of  $g$ , combined with the incumbent solution provide a feasible solution to (21) for the full ground structure. A feasible primal solution is easily obtained by inserting zero values for  $q_A, q_B$  and  $s$  for inactive elements. These now form a pair of primal-feasible and dual-feasible solutions, with the same objective. Thus the strong duality theorem [24] proves that the obtained solution is optimal for the fully connected ground structure.

This situation rarely occurs after the solution of the first sub-problem. Several iterations are usually required, adding some violated elements to the ground structure at each iteration. When one or more sets of candidate values for  $g$  do not satisfy (21e), then the corresponding element is considered for addition to the ground structure in the next iteration.

For example, the results presented later in Figure 16 involve a ground structure of over 250,000 potential elements. Directly solving the problem (20) for the full ground structure takes around between 1 and 2.5 minutes. If member adding is used, the total time required to solve the problem is 20-40 seconds, requiring between 5 and 7 iterations. Typically, the speed-up from member adding increases for larger problem sizes. An additional benefit for larger problems is that much less memory is required when member adding is used (reduced by around 90% in the cases mentioned here), which means even larger problems can be solved. It should be emphasized that this process is guaranteed to obtain the same optimal volume as directly solving the full problem.

### 3.2. From numerical solution to physical structure

This section discusses how a solution to the pair of problems (20) and (21) is used to construct the corresponding structural form. Reconstruction of the optimal grid-shell is possible under two conditions stated and discussed in Section 3.2.1. Under those conditions, primal and dual optimal solutions meet the optimality criteria listed in Section 3.2.2. These, in turn, will pave a way to a rigorously justified algorithm for recasting the nodal elevations  $\mathbf{z}$ . The global compatibility conditions (15), (16) will be evidenced, rendering the pair (20), (21) a convex reformulation of the original, non-convex, optimal form-finding problem<sup>1</sup> rather than merely its relaxation.

#### 3.2.1. Criteria for optimal fully-stressed grid-shell structures

In order that the grid-shell form is possible to construct based on the solutions of the problems (20), (21), two criteria must be met:

- the primal problem (20) is feasible,

---

<sup>1</sup>The original, non-convex, form-finding problem is the primal problem (20), but where the relaxed geometry constraint (20d) is replaced by equations (17) for each element, and  $z$  elevation for each node is explicitly included as an optimization variable.

- the solution  $(\mathbf{u}, \mathbf{w})$  of the dual problem (21) satisfies

$$\mathbf{w} < \frac{1}{\rho g} \quad (22)$$

(note that, *a priori*,  $\mathbf{w} \leq \frac{1}{\rho g}$  by the dual conic constraints).

These criteria are of *a posteriori* nature since, in practice, they can be verified only after solving the pair of problems (20), (21). Whilst the first condition is trivially necessary for the correctness of the method, the second one is rather abstract. Nevertheless, both conditions can be translated to natural structural requirements, as explained below.

Firstly, it is worthwhile to mention that the analysis of the two criteria simplifies greatly when there are no upward loads, i.e.

$$\mathbf{f}_z \leq 0. \quad (23)$$

Note that this loading scenario is the most frequent in practice. It turns out that, in this case, the second criteria, (22), is automatically satisfied whenever the problem (20) is feasible. This can be made rigorous with a rather long and technical proof, available in the supplementary materials. But intuitively, the term  $\mathbf{f}_z^T \mathbf{w}$  in the dual objective favours negative virtual deflections  $\mathbf{w}$  for downward loads, and hence for these cases it is found that  $\mathbf{w} \leq 0 < \frac{1}{\rho g}$ .

Regardless of the loading applied, if the first criteria does not hold, i.e. the primal problem (20) is infeasible, then there is no possible structure capable of satisfying the specified scenario. This is because the relaxations introduced in Section 3.1.4 only increase the possible feasible solutions, they do not exclude any feasible structures.

To assure the feasibility of (20), it is naturally necessary to guarantee a proper connectivity of the ground structure with respect to the loads and supports, minding that only compression is allowed. This requirement is no different than the one for the weightless variant of the problem in (1). However, when the self-weight is present, infeasibility may also be caused by excessive weight-to-stress ratio  $\frac{\rho g}{\sigma}$  for the spans involved. Indeed, elements  $i$  whose horizontal length  $l_i$  is  $\frac{\pi \sigma}{\rho g}$  or more are automatically excluded from the ground structure. Nonetheless, guaranteeing that  $l_i < \frac{\pi \sigma}{\rho g}$  for the whole ground structure is not a sufficient condition for feasibility, the specific geometry and connectivity of the problem must also be considered.

It turns out that the condition (22) is also related to keeping the ratio  $\frac{\rho g}{\sigma}$  below a certain, problem-specific, limit. Intuitively, observe that as  $\rho g \rightarrow 0$ ,  $\frac{\sigma}{\rho g} \rightarrow \infty$ , and so (22) does not constrain the value of  $\mathbf{w}$ .

Thus, both criteria correspond to requiring feasibility on the equivalent weightless problem, plus a requirement that the material weight-to-stress ratio is ‘sufficiently small’ for the spans and geometry concerned. Finding the critical value for what is ‘sufficiently small’ is difficult as this depends on all of the problem data including supports, domain geometry and loading (including whether these are upwards or downwards). However as a first approximation and intuition, this often relates to when some distance in the domain (e.g. between adjacent supports) reaches the maximum span of the catenary  $\frac{\pi \sigma}{\rho g}$ .

If upward loads are present, then one may obtain numerical solutions for which the condition (22) fails, without causing infeasibility. Such a numerical solution does not allow for constructing a corresponding structure that is a ‘pure’ grid-shell,

as the original non-convex form finding problem is ill-posed. However, a physically meaningful structure can still be proposed if lumped masses are introduced at nodes for which  $w = \frac{1}{\rho g}$ . Although this structure is less practical to construct than the pure grid-shell structures, it provides a valid bound on the minimum material usage. Section 3.2.3 below outlines this case.

### 3.2.2. Reconstruction of the optimal grid shell from an optimal primal and dual solution

To show that the relaxations introduced in Section 3.1.4 are never active in optimal solutions of (20), (21), one must appeal to additional properties that are exhibited at optimality. Under the *a posteriori* criteria in Section 3.2.1, the optimality conditions will pave the way to a globally-optimal, fully-stressed grid-shell. Below, the essential features of the optimal primal and dual variables  $\mathbf{s}, \mathbf{q}_A, \mathbf{q}_B$  and  $\mathbf{u}, \mathbf{w}$  are listed; rigorous derivation of these from duality principles can be found in Appendix B.

- (I) The primal conic constraint (20d) is always satisfied as an equality,

$$(\sin \bar{l} q_A + \cos \bar{l} s)(\sin \bar{l} q_B + \cos \bar{l} s) = s^2. \quad (24)$$

- (II) Whenever  $s \neq 0$ , the complementary slackness conditions give:

$$\begin{aligned} \left(\frac{1}{\rho g} - w_A\right)(\sin \bar{l} q_A + \cos \bar{l} s) \\ = \left(\frac{1}{\rho g} - w_B\right)(\sin \bar{l} q_B + \cos \bar{l} s). \end{aligned} \quad (25)$$

- (III) For every element with  $s = 0$  there is also  $q_A = q_B = 0$ .

A few comments about the foregoing properties are in order. The property (I) shows that for each non-zero element ( $s > 0$ ) the forces  $s, q_A, q_B$  are aligned with the shape of an equally stressed catenary as in Figure 4. However, in the case when  $s = 0$ , the equality (24) only implies that  $q_A q_B = 0$ , which allows that, e.g.,  $q_A = 0, q_B > 0$ . This does not adhere to the conditions (15), (16). Physically, it corresponds to the scenario where  $q_B$  does not come from a bar but represents a lumped mass at the end  $B$ , cf. Section 3.2.3 below. The property (III) (proved separately from (I)) rules out such a possibility.

Finally, the property (II), combined with (I), will be used below to reconstruct the nodal elevations  $\mathbf{z}$  that validate the global compatibility conditions (15), (16).

In the original vault formulation without self-weight (Sec. 2.1), the required elevation at each node may be obtained from the virtual deflections in the vertical direction, as found in the optimal dual variable  $\mathbf{w}$ , through a simple formula:  $z = -\frac{\sigma}{2}w$ . For the case with self-weight the required relationship turns out to be:

$$z = \frac{\sigma \log(1 - \rho g w)}{2\rho g}. \quad (26)$$

This formula provides an extra intuition behind the second criteria in Section 3.2.1, i.e. the inequality (22): it is necessary to guarantee that the argument of the logarithm above is positive. The very same formula was used for optimal archgrids with self-weight by Rozvany et al. [20, Eqn. 16]. To justify it in the broader framework developed in this paper, for each non-zero element ( $s > 0$ ) the compatibility relations (15), (16) must be evidenced. To that aim, the properties (I), (II) of the optimal solutions will be used.

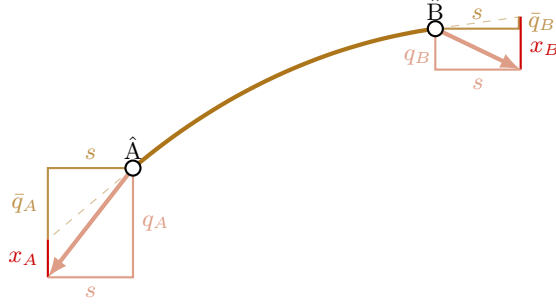


Figure 5: A single catenary element AB in the catenary vault formulation, including representation of the required ‘lumped mass’ relaxation. Note that, as in Figure 4,  $\bar{q}_A$  is shown with a positive value, whilst  $\bar{q}_B$  is shown with a negative value, meanwhile  $x$  and  $s$  are always restricted to non-negativity.

Using the formula (26) leads to:

$$\begin{aligned} z_B - z_A &= \frac{\sigma}{2\rho g} \left( \log(1 - \rho g w_B) - \log(1 - \rho g w_A) \right) \\ &= \frac{\sigma}{2\rho g} \log \left( \frac{1 - \rho g w_B}{1 - \rho g w_A} \right) \\ &= \frac{\sigma}{2\rho g} \log \left( \frac{\sin \bar{l} q_A + \cos \bar{l} s}{\sin \bar{l} q_B + \cos \bar{l} s} \right), \end{aligned}$$

where to pass to the last line the complementary slackness condition (25) was employed. Next, the equality (24) furnishes  $\sin \bar{l} q_B + \cos \bar{l} s = s^2 / (\sin \bar{l} q_A + \cos \bar{l} s)$ , thus allowing to continue,

$$\begin{aligned} \frac{\rho g}{\sigma} (z_B - z_A) &= \frac{1}{2} \log \left( \left( \frac{\sin \bar{l} q_A + \cos \bar{l} s}{s} \right)^2 \right) \\ &= \log \left( \frac{\sin \bar{l} q_A + \cos \bar{l} s}{s} \right). \end{aligned}$$

Inverting the logarithm leads to the condition (17a). The condition (17b) can be checked similarly, and, together, they are equivalent to (15), (16). Readily, the solutions, and the method itself by extension, are rigorously validated for data that satisfy the criteria in Section 3.2.1.

Once the optimal elevations of the nodes are known from (26), centerlines for each element can be obtained using the centerline equation (6). This allows for graphical representation of the structure. The cross-section area at each point along the element is found based on the total force at that point,

$$\text{cross-section area} = \frac{s}{\sigma \cos \alpha} = \frac{s}{\sigma \cos \left( -\frac{\rho g}{\sigma} \tilde{x} + C_1 \right)}, \quad (27)$$

(refer to Section 2.2 for meanings of symbols).

### 3.2.3. A lumped mass interpretation for cases where the optimal structure is not fully stressed

This section will discuss the case where an optimal solution is obtained to (20), (21) which does not comply with the a posteriori criteria (22), i.e.  $w = \frac{1}{\rho g}$  for at

least some nodes. Note that, as discussed in Section 3.2.1, this case occurs only when there are upwards loads present and the self weight is relatively large (or, equivalently, spans are long).

As  $w \rightarrow \frac{1}{\rho g}$ , equation (26) implies that  $z \rightarrow -\infty$ . So the original, non-convex problem attains the optimum only at the limit. Nonetheless, the convex problem (20) attains the same optimum value with no issue as  $z$  is not explicitly present, however the reconstruction process in Section 3.2.2 no longer applies.

The original non-convex form-finding problem inherently assumes that all material must be fully stressed. The relaxation of moving from the equality (18) to the inequality (19) can be understood as removing this assumption, and in so doing unlocks a more interpretable physical structure for these cases.

Assume that, for some element, (19) is a strict inequality. This means that the vertical forces  $q_A$  and  $q_B$  are not aligned to the end-points of a catenary of equal stress with horizontal length  $l$  and thrust  $s$ , see Figure 4. From the direction of the inequality, the values of  $q_A$  and/or  $q_B$  can only be larger (in the downwards direction) than expected from the definition of the equally stressed catenary. Physically, such surplus can be understood as putting lumped masses  $x_A, x_B \geq 0$  at the respective endpoints (which always act downwards, i.e. are positive). This is illustrated graphically in Figure 5, from which the following relations may be observed,

$$q_A = \bar{q}_A + x_A, \quad q_B = \bar{q}_B + x_B, \quad (28)$$

where  $\bar{q}_A, \bar{q}_B$  are the aligned forces, i.e. they satisfy  $(\sin \bar{l} \bar{q}_A + \cos \bar{l} s)(\sin \bar{l} \bar{q}_B + \cos \bar{l} s) = s^2$ .

Note that the objective function is calculated based on the total vertical forces  $q$ , not the aligned force  $\bar{q}$ . This means that any lumped masses which occur in a solution are correctly counted as part of the structural volume.

There are now two possible ways in which upwards loads may be resisted by the structure. These correspond to (22) being true or false respectively:

- *Grid-shell regions*, as described above, where fully-stressed frameworks transmit the load back to some supports, and  $w < \frac{1}{\rho g}$ .
- *Counterweight regions* where the weight of material acting at each node is exactly equal to the upward force applied at that node, permitting vertical equilibrium to be attained without recourse to supports, and  $w = \frac{1}{\rho g}$ .

Note that these definitions operate node-wise, i.e. it is possible for the minimum material solution to contain some grid-shell regions and some counterweight regions.

Supplementary material section S2 proves these further features for these optimal solutions:

- (SI) No real element (i.e. with  $s > 0$ ) can connect between a node in a counterweight region and a node in a grid-shell region.
- (SII) Within counterweight regions, in the usual case where  $\mathbf{f}_{xy} = \mathbf{0}$ , there is an optimal solution consisting only of lumped masses, which allows elevations of each node to be chosen arbitrarily. Otherwise, the procedure to obtain valid elevations is shown in the supplementary material.



So for cases containing both region types, the optimal structure would be disconnected between the grid-shell and counterweights. Disconnected ‘floating’ regions of material are likely to be impractical, and are clearly in unstable equilibrium. Nonetheless, this solution can still be used as a lower bound on the material required to support the given scenario. Extra elements added to ensure connectivity would require only nominal cross-sections. Furthermore, the solution in the grid-shell regions acts independently and can still be utilized even if the counterweight areas are deemed impractical.

#### 4. Examples

A series of numerical examples are now presented to demonstrate the efficacy of the method and to demonstrate characteristics of optimal vaults under significant self-weight.

To allow illustration and comparisons between structures with negligible impact from self-weight and those where self-weight is dominating, each example is considered across a varying range of values for unit weight ( $\rho g$ ) of the material used. Note that, throughout the derivation section, it is the strength-to-weight ratio of the material which is used in calculations. Thus, a proportional increase in both unit weight and allowable stress implies only a uniform scaling of the thickness of each element. In other words, increasing the unit weight could be equivalently thought of as increasing the span (for fixed material properties) or reducing the allowable stress (for a constant span and material weight). Presentation in terms of unit weight is chosen here for ease of rendering and visually comparing the resulting forms – in this way, changes in structure elevations and element thicknesses in the images are solely due to the impact of the self-weight.

To highlight this, results are presented in normalised form. For context, values given at  $\rho g = 1 \frac{\sigma}{L}$  correspond to  $L = 4.4m$  for the material values observed in the Armadillo vault [26, Stress = 0.1 MPa, density 2320 kg/m<sup>3</sup>]. The impact of self weight can be significantly reduced by increasing the allowable design stress, however this is likely to be less desirable in practical terms as it may increase the impact of second-order effects such as buckling.

Problems have been solved using MOSEK [27], interfaced via the CVXPY programming interface [28]. The CVXPY interface has a substantial processing overhead when adding many rotated quadratic cones. Therefore, the rotated quadratic cone (20d) is re-written as the equivalent standard quadratic form

$$(q_A + q_B) \sin \bar{l} + 2s \cos \bar{l} \geq \sqrt{(\sin \bar{l}(q_A - q_B))^2 + (2s)^2}. \quad (29)$$

Example python scripts and Rhino/Grasshopper files are provided [29], and the use of CVXPY means that alternative open-source solvers can easily be used if MOSEK license is not available.

##### 4.1. One-way spanning barrel vault

To validate the formulation proposed here, a very simple example will be considered. The example is as shown in Fig. 6, and consists of a ground structure of 12 nodes, with point loads applied at the mid-span and pinned supports along 2 edges.

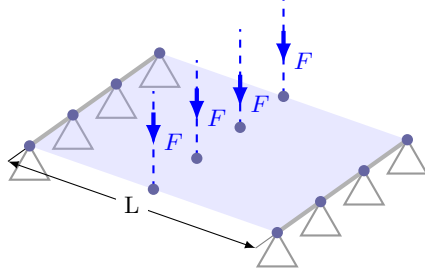


Figure 6: Barrel vault: Problem setup, consisting of nodal supports at two edges of a rectangular domain with span  $L$ , with point loads  $F$  applied to the centre/apex line.

As this is essentially a 1-dimensional problem, the expected solution will be a series of four identical parallel pointed arches. Because of this, it is possible to enumerate all possible solutions to this problem through the changing of a single height variable. This will give a suitable comparison by which to validate the numerical solutions. Here, the enumerated solutions have been obtained through the 3D ground structure method using Peregrine [30], which implements the catenary self-weight model from [16]. This has been solved for pre-determined loading points at various heights as shown in Figure 7.

This procedure has been repeated for materials ranging from very lightweight to very heavy. The black crosses in Figure 7 show the optimal height and volume obtained by the vault formulation proposed here. This allows more precise identification of the optimal height compared to the finite step size used in the exhaustive approach.

From Figure 7 it can be seen that the problem is relatively in-sensitive to changes in elevation when self-weight is negligible, particularly with respect to structures that are taller than the optimal height. However, when self-weight is more significant, the importance of selecting the optimal height is increased, with larger volume increases observed for structures which are either too tall or too short. Furthermore, there are significant differences in the results between the case with negligible self-weight and when self-weight is significant. Adding self-weight requires significantly more material to be used and the curvature of the elements to be increased, demonstrating the importance of considering self-weight in the form-finding stage.

Figure 8 shows the same example but with additional nodes in the ground structure. From this it can be seen that the shape of the catenary of equal stress is successfully obtained as a series of curved segments, attaining the same material usage, as expected. This is in contrast to the lumped mass formulation (Appendix A) where the optimal volume changes with nodal resolution.

#### 4.2. Square domain with distributed load

The next problem to be considered is a square domain with pin supports at all edge nodes, with the domain discretized by an  $11 \times 11$  grid of nodes (121 nodes total). Point loads of equal magnitude are added at all unsupported internal nodes. The problem has again been solved with materials ranging from very light to very heavy and the total volumes are shown in Figure 9. The problem has been solved with the load applied in both an upward or downward direction. It can be

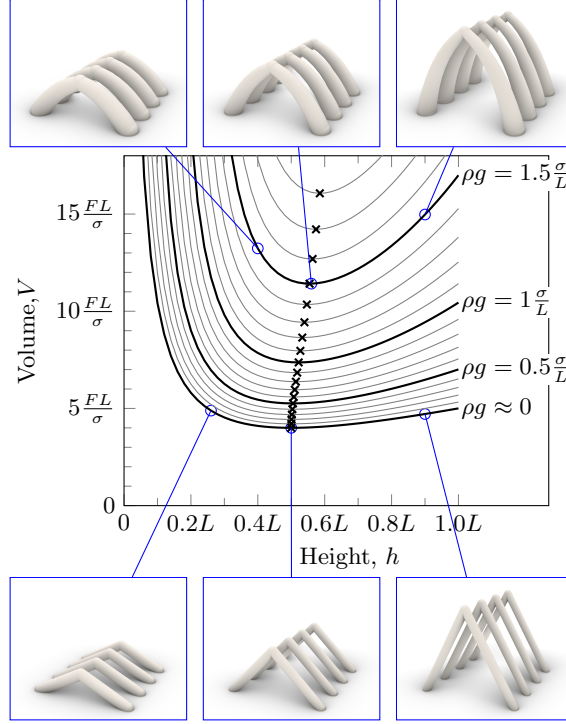


Figure 7: Barrel vault: Enumerated solutions for problem heights at  $0.02L$  spacing, and material unit weights in  $0.1 \frac{\sigma}{L}$  increments. Selected rendered forms shown for  $\rho g = 0.001 \frac{\sigma}{L} \approx 0$  and  $\rho g = 1.5 \frac{\sigma}{L}$ . Black crosses mark the obtained solutions using formulation (20d) for the same values of unit weight.

seen that when self-weight is less significant, the upward and downward loadings produce structures with similar volumes, and where the forms are close to mirror images of one another (Figure 9b and d). However, as self-weight becomes more significant, the forms diverge (Figure 9c and e). In both cases, the arching/sagging of the elements becomes more pronounced, with the downward load this causes an increase in height and with upward load the height reduces. For this example, the structure for upward loading undergoes only minor changes until  $\rho g \approx 15 \frac{\sigma}{L}$ , when it abruptly switches to a counterweight solution across the entire domain near-simultaneously, which continues to be the optimal solution for  $\rho g \rightarrow \infty$ . For an example with upward loading which demonstrates the separation of the solution into regions, see Figure S1 in the supplementary materials.

For downward loads, the volume increases as self-weight becomes more significant. The volume tends to infinity, with an asymptote at  $\rho g \approx 4 \frac{\sigma}{L}$ , beyond which the problem is infeasible. Rozvany et al. [20] considered a problem similar to this, but with a fixed topology of elements parallel to the domain edges – an archgrid. This can be easily modeled using the proposed approach by simply restricting the elements available in the ground structure, and the results of this are also shown in Figure 9, where it can be seen that optimizing the topology permits a significant material saving of almost 25% in cases with significant self-weight ( $\rho g = 2 \frac{\sigma}{L}$ ).

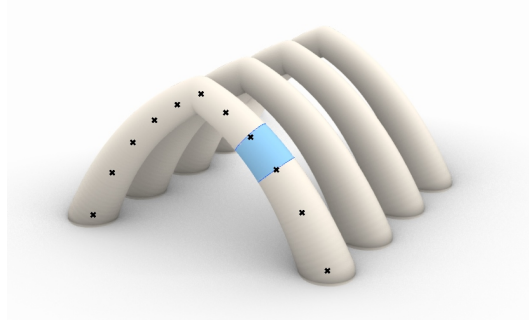
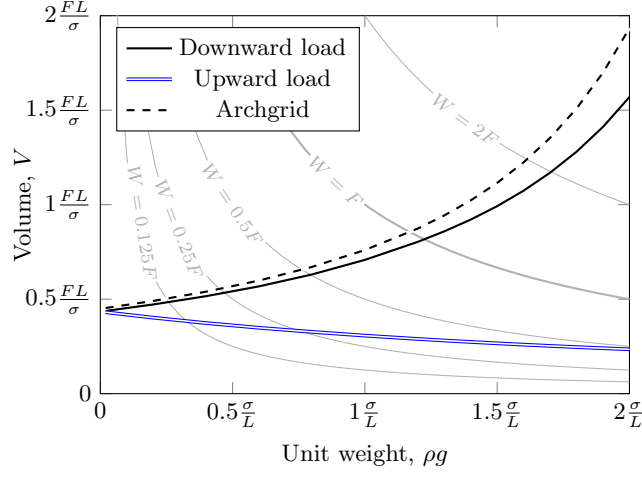


Figure 8: Barrel Vault: Output from catenary vault formulation at higher resolution showing how an individual element (blue) forms part of the arch and the correctly determined heights of intermediate nodes (black).

To further explore the influence of different topologies, Figure 10 explores results with the shown selected restricted topologies. It can be seen that the decision of which topology is preferable varies significantly for different values of self-weight, indicating that some forms are better suited to carrying self-weight loads and others to carrying uniform external loading. For example, the Opt-heavy form provides the minimum material result for larger values of self-weight, but when used to design a structure with a lightweight material it is close to the worst performing topology tested. Meanwhile the converse is true for the Opt-light form. Even between the less efficient forms of the Archgrid and Diagonals topologies, there is a switch in preferable topology as self-weight increases.

Generally, the impact of selecting an incorrect topology appears to increase when self-weight is more significant, with only 3.5% between the options tested when self-weight is negligible, and 30% variations when  $\rho g = 2 \frac{\sigma}{L}$ . Furthermore, Figure 10 shows the importance of optimizing topology and elevations simultaneously, since the optimal elevation profiles vary between the different topologies tested.

To show the benefit of the proposed approach against existing methodologies, results will be compared for accuracy and speed against the 3D truss layout optimization method with transmissible loads [14]. The comparison is carried out with a horizontal discretization of  $0.1L$  as previously, and with an unrestricted topology (fully connected ground structure). To use the 3D truss layout optimization approach, the domain must also be defined and discretized in the vertical direction, and the choice of this discretization can greatly influence the speed and accuracy of the solution. Here a design domain with height  $0.8L$  has been used (note that the corresponding result in Figure 9c has maximum height  $0.604L$ ), and resulting volumes and solution time with a range of vertical spacings are shown in Figure 11. The results show good convergence of the volume towards the vault formulation result, providing further validation of the approach proposed here. However, the transmissible load approach generally produces a non-single-surface structure, leading to additional difficulty in interpreting the results, as shown in Figure 11c and d, where multiple elements are seen on each vertical grid-line. Furthermore, the time required to solve the problem using the vault formulation presented here



(a)

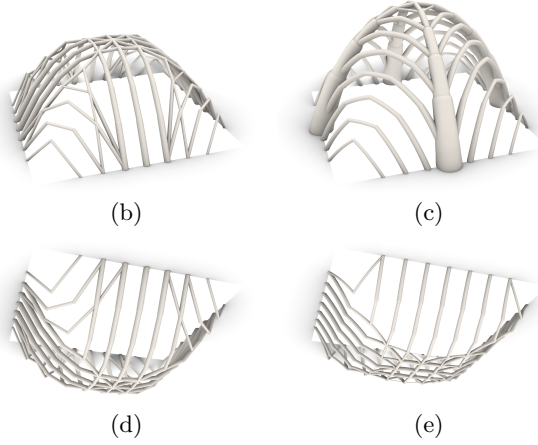


Figure 9: Square domain example with distributed loads and pinned edges: (a) Optimal volumes for materials of various unit weights. For context, relationships between the structure weight  $W = \rho g V$  and the total external loading  $F$  are given. (b)-(e) optimal solutions; (b), (c) for downward load and (d), (e) for upward load; (b), (d) for  $\rho g = 0.2 \frac{\sigma}{L}$  and (c), (e) for  $\rho g = 2 \frac{\sigma}{L}$ .

was less than 0.6 seconds, whereas the transmissible load approach required several orders of magnitude more computation time (up to 6 hours at the discretizations tested here) to produce a less accurate result <sup>2</sup>.

#### 4.3. Square domain with point load

The next problem again concerns a square domain with sides of length  $L$ . However, for this example, the supports are available only at the four corners, and

<sup>2</sup>Note that member adding has been used for the transmissible loading approach as well as the vault formulation, otherwise the required times would likely be 1-2 orders of magnitude larger, and the memory requirements would preclude solving on a standard laptop.

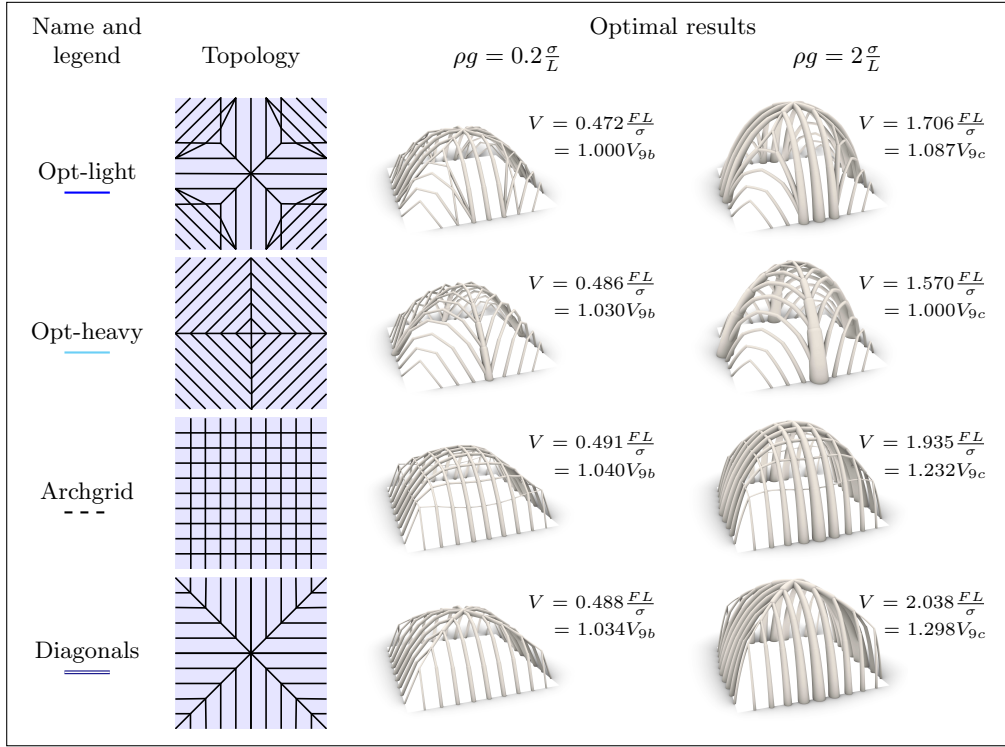
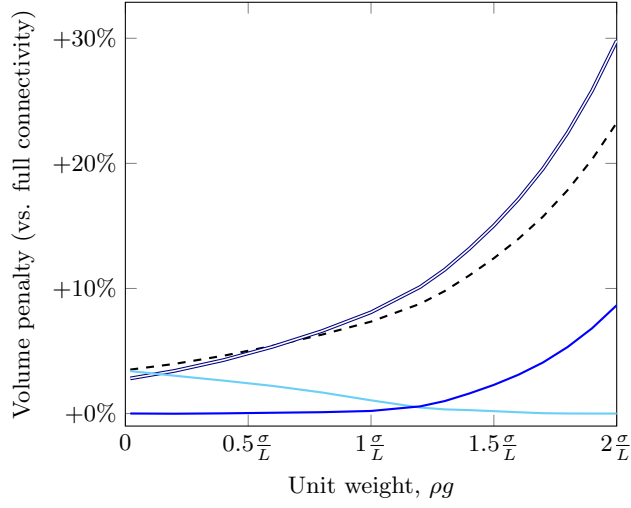


Figure 10: Square based example: Results with restricted topologies. ‘Topology’ images show the available elements in the restricted ground structures. Opt-light is based on the optimized structure at low self-weight values (e.g. Figure 9b), manually simplified to remove equally-optimal elements around the center of the domain. Opt-heavy is the topology identified as optimal at high self-weight values (e.g. Figure 9c). Archgrid allows only elements aligned to the  $x$  or  $y$  axes. Diagonals allows the two main diagonals, plus elements perpendicular to the closest edge of the domain, based on extending the pattern observed at the center of the Opt-light topology. Note that some ground-structure elements are not used in the optimized results, e.g. the elements connecting to the corner points in the Diagonals result. Structure volumes,  $V$  are shown as raw values, and also in the normalised form plotted above, i.e. normalised by the volume of the full connectivity result from Figure 9b ( $V_{9b}$ ) or Figure 9c ( $V_{9c}$ ) as appropriate.

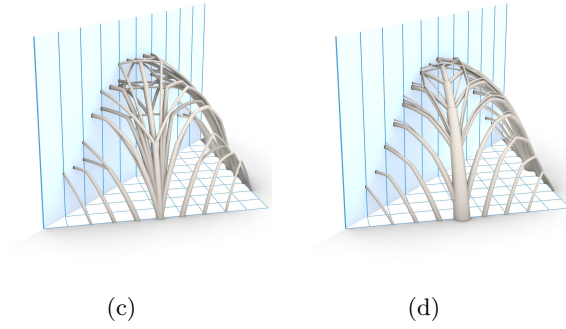
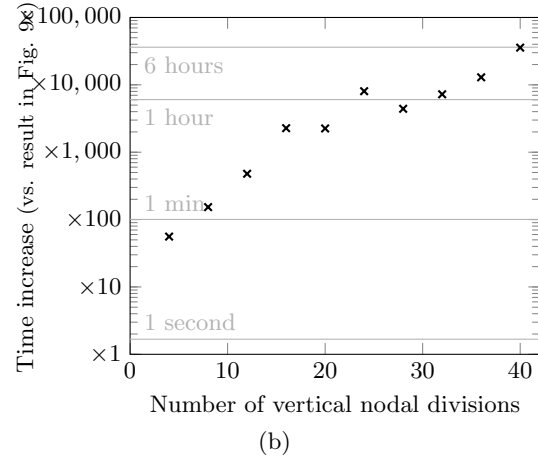
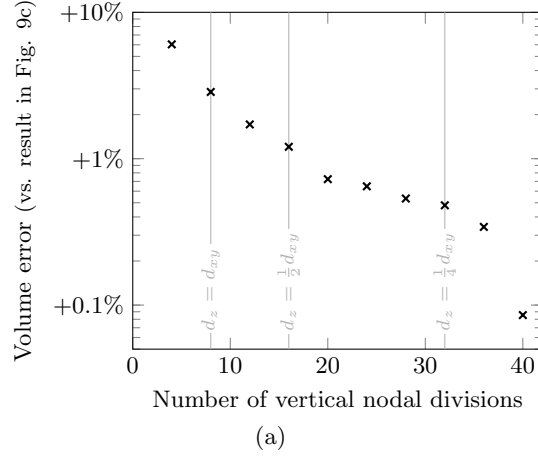


Figure 11: Square based example: Comparison of vault formulation with transmissible load approach for  $\rho g = 2 \frac{\sigma}{L}$ , and maximum permitted height  $0.8L$ . (a) Volume, with indication of nodal spacings in the horizontal directions,  $d_{xy}$  compared to the vertical nodal spacing  $d_z$  (b) Time required and (c)-(d) forms obtained using the transmissible load approach. The blue plane cuts the domain corner-to-corner and the lines of the nodal grid are shown to highlight the non-single-surface nature of the results. (c)  $d_z = d_{xy}$  i.e. 8 nodal divisions, (d)  $d_z = \frac{1}{5}d_{xy}$  i.e. 40 nodal divisions.

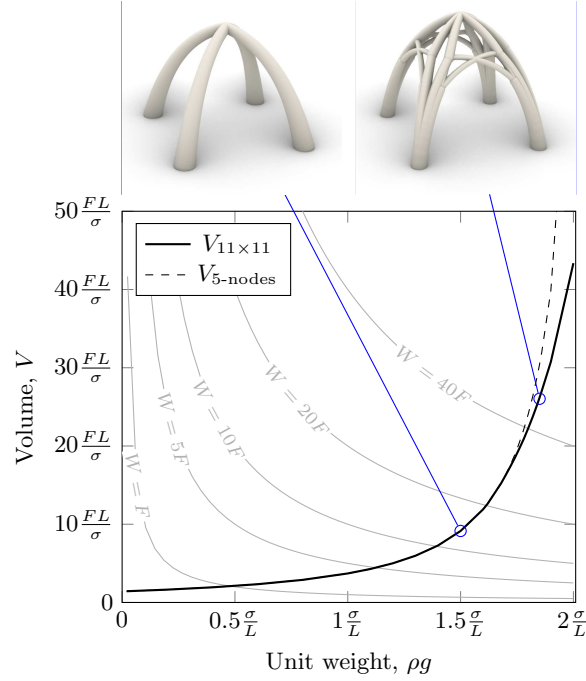


Figure 12: Square example with point load: Optimal volumes for fully connected  $11 \times 11$  grid,  $V_{11 \times 11}$  and for minimal 5 node (load and supports only) problem,  $V_{5\text{-nodes}}$ . Example results shown for  $\rho g = 0.75 \frac{\sigma}{L}$  (identical for 5-nodes and  $11 \times 11$  result) and  $\rho g = 1.85 \frac{\sigma}{L}$  (for  $11 \times 11$  grid). For context, relationships between the structure weight  $W = \rho g V$  and the total external loading  $F$  are given.

the load is applied as a point load at the center of the domain acting downward. The resulting volumes based on an  $11 \times 11$  grid of nodes with a fully connected ground-structure are plotted in Figure 12 as  $V_{11 \times 11}$ . It can be seen that the self-weight dominates at shorter spans than in Section 4.2, with a much more rapid increase in volume observed; this is to be expected as there are fewer supports and there is no longer a proportion of the loading applied very close to the supports. When self-weight is significant, these solutions make use of nodes other than the loads or supports, see example in Figure 12.

When self-weight is not considered, it has been supposed that an optimal topology must exist that involves only joints at points that are loaded or supported (Conjecture 9.1 in [22]) although there may be equally optimal solutions with extra joints. In this case, this would imply an equally optimal solution containing 5 nodes – the loaded point plus the four supports – and a simple  $\times$ -shaped structure. To verify that the more complex solutions are beneficial, the problem was re-analyzed using a restricted ground structure containing just those 5 nodes, and these results are also shown in Figure 12 and Table 1 as  $V_{5\text{-nodes}}$ . It can be seen that the simpler topology uses significantly more material than the optimized topology, with the optimal topology requiring around half (53.7%) of the material needed for the simpler design when  $\rho g = 2 \frac{\sigma}{L}$ . However, for lower values of self-weight ( $\rho g < 1.65 \frac{\sigma}{L}$ ) both problems give the same solution, again implying a change in



Table 1: Square example with point load: Optimal volumes for selected values of self-weight,  $\rho g$  as highlighted in Figures 12, 13 and/or 14. Results given for  $V_{5\text{-nodes}}$  allowing nodes at supports and loaded points only;  $V_{11\times 11}$  using a fully connected  $11 \times 11$  grid at fixed positions; and  $V_{13\text{-nodes}}$  with nodes at supports, loaded point and one optimally selected additional node per segment as shown in 13.

$\rho g \left( \frac{\sigma}{L} \right)$	$V_{5\text{-nodes}} \left( \frac{FL}{\sigma} \right)$	$V_{11\times 11} \left( \frac{FL}{\sigma} \right)$	$V_{13\text{-nodes}} \left( \frac{FL}{\sigma} \right)$
1.65	13.8394	13.8394	13.8394
1.68	15.2528	15.2516	15.2415
1.72	17.5301	17.3435	17.3056
1.76	20.4014	19.6510	19.6100
1.80	24.0981	22.2817	22.2180
1.85	30.4425	26.1884	26.0334
2.00	80.7391	43.3682	43.3575

optimal topology when self-weight is more significant.

The results given for  $V_{11\times 11}$  are the globally optimal solutions for that discretization of the domain. However, moving the (non-loaded) nodes may further improve the solution, as in geometry optimization which is commonly employed for truss problems. Here a simple, semi-manual approach is used for demonstration purposes. From observation of the optimal  $11\times 11$  forms, it is assumed that the optimal topology involves the 5 loaded/supported points plus an additional point in each of the 8 symmetrical segments of the domain defined by the horizontal, vertical and diagonal symmetry of the problem; one of these segments is highlighted in Figure 13 (duplicate nodes were removed for degenerate cases where the extra node is located on a symmetry plane). The position of all the extra nodes is thus controlled by just two variables, the horizontal and vertical position.

Figure 13a shows the resulting volumes – for a given unit weight – for all possible locations of the additional point. This shows the simple behavior of the problem which exhibits a single optimum location for the point, denoted as  $V_{13\text{-nodes}}$ . Thus, using the initial  $11\times 11$  node solution as a guide, an optimal node location can be found easily by manually exploring different locations. With each problem taking less than 0.1 seconds to solve, this is easily achieved using parametric modeling software such as Rhino/Grasshopper [31]. In all cases, this proved to make only a minor improvement to the total material usage (generally around 0.2%, with some cases up to 0.6% where the optimal node position fell directly between points in the original ground structure). The resulting volumes,  $V_{13\text{-nodes}}$ , are also shown in Table 1. The forms of these solutions are shown in Figure 14, while Figure 13b also shows the location of one of the additional nodes for various levels of self-weight.

From Figures 12 and 14 it can be seen that when self-weight is low ( $\rho g \leq 1.65$ ), the optimal structure employs the most direct load-path from the point load to the support, aligning the structure to the  $\sqrt{2}L$  length span between diagonally opposing supports. However, when self-weight is more dominant (e.g.  $\rho g \geq 1.85$ ) it is preferable to locate as much material as possible close to the shorter spans of length  $L$  between adjacent supports, with a secondary set of elements spanning from the point load to these outer elements. For intermediate values of self-weight, the transition occurs by a splitting of the original, diagonal elements.

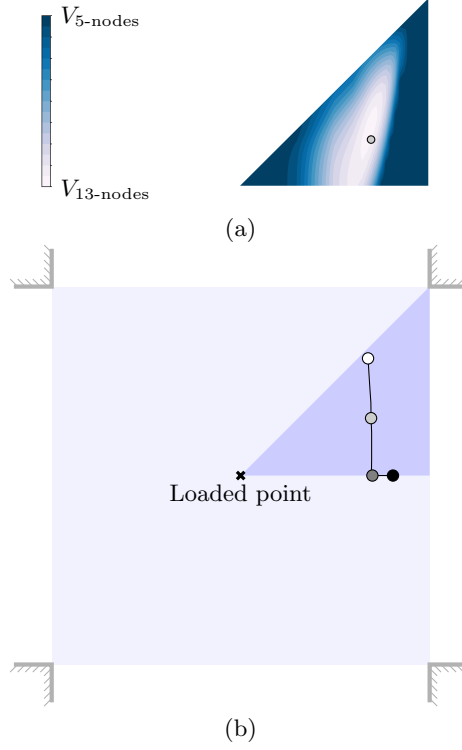


Figure 13: Square example with point load: Optimization of 13-node structures. (a) Contour plot of volumes for all possible locations of the additional node for  $\rho g = 1.76 \frac{g}{L}$ , with optimal location indicated. (b) Optimal location of additional node for  $\rho g = 1.68 \frac{g}{L}$  (white),  $\rho g = 1.76 \frac{g}{L}$ ,  $\rho g = 1.85 \frac{g}{L}$  and  $\rho g = 2 \frac{g}{L}$  (black).

#### 4.4. Example with two holes

The example in Figure 15 will now be considered. To generate a pure compression structure, it is essential that supports on the convex edges<sup>3</sup> of the domain support the structure both vertically and horizontally, allowing the required thrust to be generated. However, for supports on the interior of the domain, or on concave edges, it is possible to model the existence of vertical-only (roller pin) supports. This example has support in all three directions (full pin) at the outer boundary (solid lines in Figure 15a), and vertical-only support at one of the inner hole boundaries (dashed lines in Figure 15a). The boundary of the other hole is not supported. A regular grid of nodes is used across the rectangle with  $\frac{L}{40}$  spacing, nodes within  $0.15L$  of the centers of the holes are removed, and 32 equally spaced nodes are added around each hole. Full connectivity is permitted, excluding elements that would overlap the holes; adjacent nodes around the edge of each hole

<sup>3</sup>The term convex edge is used here to denote locations on the domain boundary where the tangent is (locally) on the outside of the domain, or the point is a corner with internal angle less than  $\pi$  radians. Conversely, concave edges are those portions of the boundary where the tangent lies (locally) inside the domain or at corners with internal angles greater than  $\pi$  radians. Straight edges may also be unsupported, although equilibrium enforces that an unsupported straight edge will have only a single arch along it, not connected to the interior of the domain.

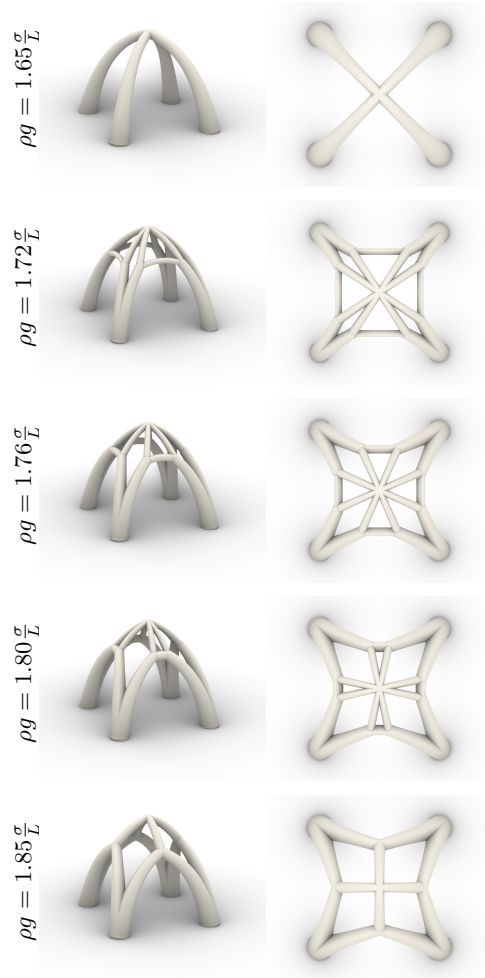


Figure 14: Square example with point load: Optimal '13-node' solutions (plan and perspective view) for various values of self-weight

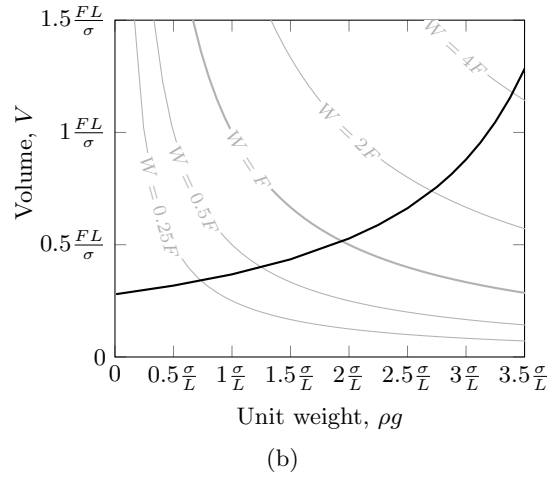
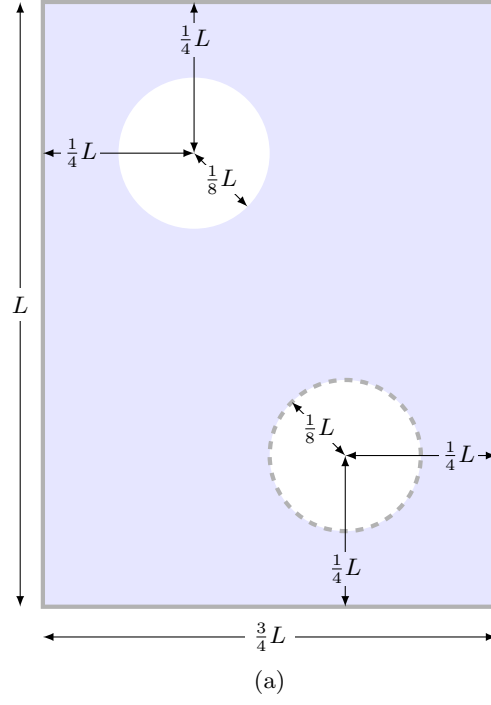


Figure 15: Example with two holes: (a) Problem specification, fully pinned ( $x$ ,  $y$  and  $z$  direction) supports defined around outer edges indicated with solid line, and vertical-only supports defined around one of the holes indicated with dashed line. Loading is uniform across the domain. (b) Optimal structure volume for various levels of self-weight. For context, relationships between the structure weight  $W = \rho g V$  and the total external loading  $F$  are given.

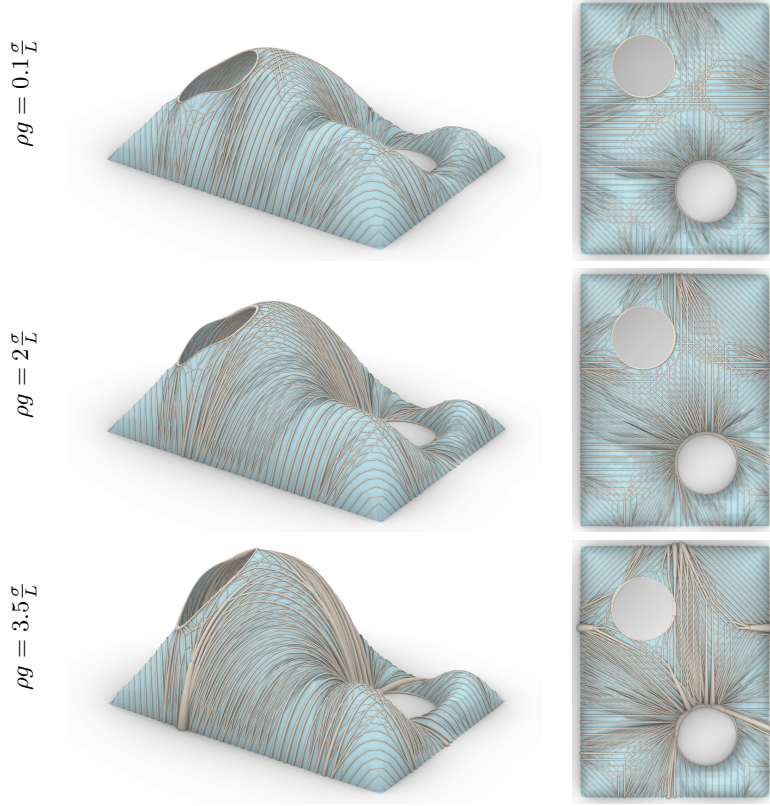


Figure 16: Example with two holes: Optimal results for various values of self-weight, plan and perspective views. These structures are identified from a ground structure of over 250,000 potential elements, requiring between 20-40 seconds to solve the optimization problem (20).

are connected with potential elements, even though these pass slightly within the circles. Co-linear elements within the grid are removed, and the problem therefore contains 1117 nodes and 258,856 potential elements. Around 20-40 seconds is required to solve the problem (20), while around 15 seconds is then needed to construct and render the structure (depending on the detail of rendering the curves); when using Grasshopper to construct the problem a further 30 seconds is needed to construct the geometry, nodal grid and ground-structure (largely attributable to calculating the contributing area for the applied force at each node).

This problem has been solved with various levels of self-weight, and the resulting volumes are shown in Figure 15b. The optimal forms are shown in Figure 16 for self-weight values of  $\rho g = 0.1 \frac{\sigma}{L}$  (self-weight negligible),  $\rho g = 2 \frac{\sigma}{L}$  (total structure weight  $\approx$  total external load, see Figure 15b) and  $\rho g = 3.5 \frac{\sigma}{L}$  (self-weight more than four times applied load, see Figure 15b). The optimal topologies observed are reminiscent of those seen in the simpler square example, see Figure 10. When self-weight is low, the optimal topology contains many elements perpendicular to the domain edges, each with similar size; when self-weight is more significant, certain support points appear preferable, with larger elements and more elements fanning

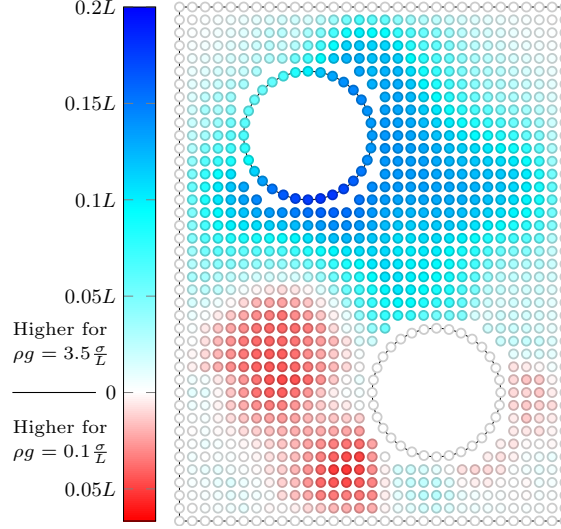


Figure 17: Example with two holes: Elevation difference between optimal structures for  $\rho g = 0.1 \frac{\sigma}{L}$  and  $\rho g = 3.5 \frac{\sigma}{L}$ , shown at nodal points. The top hole is the unsupported boundary. For context, the  $\rho g = 0.1 \frac{\sigma}{L}$  result has a maximum height of approximately  $0.32L$

out from a single point. In all the results, the vertical-only internal supports result in relatively low rises on the shorter spans, efficiently allowing a larger thrust to be generated to ensure horizontal equilibrium at those points.

In the previous examples, it has been seen that the optimal elevation increased across the whole domain when the self-weight increased. However, whilst the maximum elevation increases with self-weight in Figure 16, this is not the case across the whole domain. Figure 17 shows the difference in elevation for each node in the problem. Whilst the largest span regions (around the unsupported hole) show significant increases in elevation, notable reductions in elevations are also seen, particularly around the supported hole. This may be due to the need to generate larger horizontal thrusts to ensure horizontal equilibrium of the supported circle.

#### 4.5. Self-intersecting example

It is interesting to note that, although the problem is specified in 2D, it is not essential that the domain's outer boundary is a simple closed curve. Specifically, it is possible to consider self-intersecting domains, whilst defining that the intersecting parts are disconnected. To demonstrate this, the problem shown in Figure 18a will be considered. The boundary is a planar self-intersecting curve, formed from a series of arcs. Supports are available along the boundary sections shown with grey lines in Figure 18a, and the maximum straight line distance between two points in the domain is denoted as  $L$ .

For such a case, no special changes need be made to the method described. However, great care must be taken in constructing the nodal grid and connectivity of the ground structure, as standard algorithms will not be able to distinguish between the overlapping regions. Visual modelling environments such as

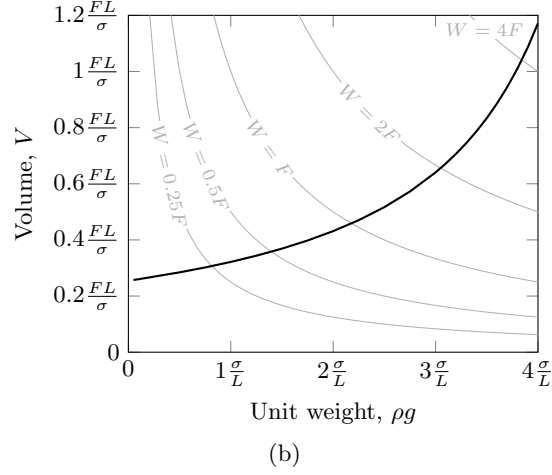
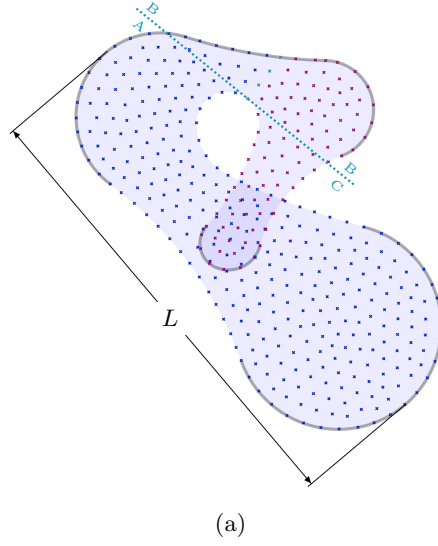


Figure 18: Self-intersecting example: (a) Problem specification. Grey edges represent locations of supports, note that the self intersecting regions (i.e blue and purple nodal points) do not interact, and the support shown there belongs only to the portion of the domain containing purple node markers. The dashed tangent line and labels A, B, C refer to the regions used to calculate the permissible ground structure. (b) Optimal volumes for different values of self-weight. For context, relationships between the structure weight  $W = \rho g V$  and the total external loading  $F$  are given.

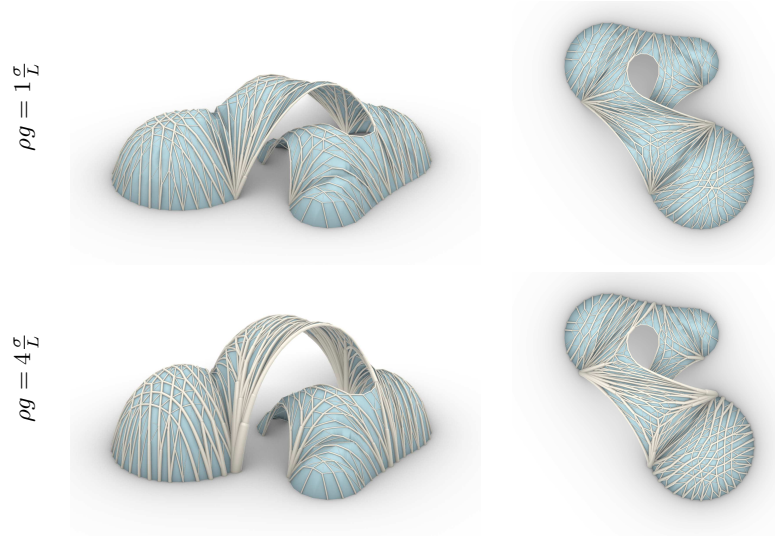


Figure 19: Self-intersecting example: Results, plan and perspective view. These structures were identified from a ground structure of over 51,000 potential elements, requiring between 2 and 5 seconds to solve the problem (20).

Rhino/Grasshopper [31] can be invaluable for this to evaluate the problem-specific logic at each stage.

Here, the approach used was to select a point on the interior hole, and from there take a normal and tangent line to split up the domain. Nodal points were defined separately in the regions each side of the normal line (purple/blue points in Figure 18), with evenly spaced points pre-defined along the normal line. The Grasshopper Triremesh component was used to generate approximately equally spaced points across the irregular shaped regions, and a Voronoi diagram was used to assign forces to each point representing a uniformly distributed imposed load. The potential element list was then generated by considering three regions (A, B, C) of the domain defined by the tangent line (shown as a dotted line in Figure 18). Note that any element connecting section A to section C would not be permitted, regardless of the proximity of the points in actual 2D space. By separately considering the union of A and B and the union of B and C, potential elements can be checked using standard approaches, and finally the ground structure is combined for solving. The complete list of nodal coordinates and potential elements can be found in the supplementary information.

The results shown in Figure 18b required between 2 and 5 seconds to solve the optimization problem for each value of self-weight. Selected optimal forms are shown in Figure 19. From Figure 18b, it can be seen that the two cases shown in Figure 19 correspond to the cases where the self-weight is just over one quarter of the external load, and where self-weight is just over four times the external load, respectively.

For self-intersecting examples such as this one, it is not possible to specify which of the overlapping regions should occur on top, or indeed to prevent the two areas from crossing over. Nonetheless, when the overlapping regions have very



Table 2: A qualitative comparison between the approach proposed herein and existing methods. Green ticks indicate desirable characteristics, red crosses indicate limitations, and orange circles indicate intermediate values or partial ability.

	Self-weight included (i.e. applicable to long spans)	Optimal topol- ogy found	Material usage	Computational difficulty
This contribution	✓Yes	✓Yes	✓Global minimum	✓Low
Ground structure method for weightless vaults [22]	✗No	✓Yes	✓Global minimum	✓Low
3D layout optimization with transmissible loads, e.g. Fig. 11 herein, [14] or [4]	✓Yes	✓Yes	○Tends to global min- imum for dense verti- cal discretization	✗Very high
Force Density Method	○Approximate methods only	✗No	✗Not minimised	✓Low
Thrust Network Analysis	○Approximate methods only	✗No	✗Not minimised	✓Low-Moderate
Dynamic Relaxation	○Approximate methods only	✗No	✗Not minimised	○Moderate
Particle-Springs	○Approximate methods only	✗No	✗Not minimised	○Moderate
Continuum topology opti- mization with transmissible loads (e.g. [14, Fig 5c/7c]) in 3D	○Good approxi- mations possible, but not tested with transmissible loads	✓Yes	○Minimum found for high-resolution greyscale results only, black-and-white re- sults prone to local optima.	✗Very high

different spans, it will generally be the case that the optimal elevations will differ sufficiently for useful results to be obtained.

## 5. Concluding remarks

This paper provides two key contributions, a novel methodology for the topology and geometry optimization of long-span grid-shells, and the identification of characteristics of such optimal structures. The concluding remarks have been divided into those pertaining to the novel methodology itself (Section 5.1), and those relating to the characteristics of the optimal structures which can now be obtained (Section 5.2).

### 5.1. The proposed method: Summary, advantages and limitations

A methodology has been presented to obtain the globally-optimal topology and elevation profile for a pure compression grid-shell structure carrying a predefined external loading plus the self-weight of the structure itself. The inclusion of self-weight is a significant practical advantage compared to previous approaches, as grid-shells are very often used for long-span structures where self-weight is a significant proportion of the loading. By solving a convex conic programming problem, globally optimal solutions (for the given ground structure) can be obtained rapidly. The method provides speed-up of several orders of magnitude compared to a 3D truss optimization approach, whilst simultaneously increasing accuracy and increasing the clarity of the solution by ensuring a single-layer structure is returned.

A direct comparison with traditional form-finding approaches, such as the force density method or dynamic relaxation, is not possible as comparable results with accurately modeled self-weight are not available for these approaches. Nonetheless, as form-finding approaches do not aim to minimise the material usage, they generally result in less materially-efficient solutions; for the non-self-weight case, comparison between the layout optimization form-finding approach and force density methods can be found within [23] where it is found that – even for the same topology – the ground-structure based method demands up to 20% less material than designs obtained by the force density method, with further savings available if the topology is also optimized. Due to the compounding effects of self-weight loads, even greater level of savings may be obtained for long span scenarios. A summary of the advantages of the proposed approach compared to previously available methods can be found in Table 2.

The most significant limitation of the proposed approach is one which is shared with the vast majority of available form-finding approaches; the structure is designed to resist only a single load-case. In general, it is not possible to define a structure which is funicular under multiple load-cases, and so it is necessary to invoke, e.g., bending, large deformations, or move away from single layer structures. Of the approaches in Table 2, 3D layout optimization can consider multiple load cases, but it produces a structure which is no longer single-layered, and therefore no longer meets many definitions of a grid-shell.

A more unique restriction on the applicability of this method particularly (in common with [22]), is that the supports are defined as having an elevation of zero. In other words, all the support points are at the same height. The weightless formulation has been extended to cases where the supports lie on an inclined plane; it is an open question whether this approach can be extended to supports of different heights.

The approach presented here does not explicitly consider elastic compatibility or a constitutive relationship. However, this is not necessary; there will always be an optimal solution which is statically determinate, and therefore automatically satisfies elastic compatibility. This was shown for the weightless truss case by Rozvany [32, Appendix], via an argument based on the structure of the linear programming problem (statically determinate structures correspond to basic solutions of the optimization problem). It can be observed that the truss problem (5) displays the same structure, and so the argument can be directly transferred. Furthermore, once the method herein has been performed, the obtained nodal positions can easily be used to construct the problem (5) for that case, which will

attain the same optimal volume as obtained from (20) and almost always<sup>4</sup> the same optimal form.

The approach permits a limitless variety of scenarios to be investigated. Support conditions may include both pins and roller-pins, and loadings may be chosen freely (albeit that when considering only a single load case, horizontal forces may be of less practical interest). Footprint geometries can include holes and can even self-intersect, while customised ground structures may be used to investigate the effects of limiting the available topologies. If a scenario is specified for which no solution is possible, then this is easily identified, as the corresponding primal problem (20) will be infeasible. The method could equivalently be used to design pure tension structures by following the logic herein with signs reversed.

### 5.2. Characteristics of optimal grid-shell structures

Through the use of this approach, many features of optimal grid-shell structures under self-weight loading have been discovered. It is observed that as self-weight becomes more significant, the optimal peak height of the structure typically increases. For the weightless case, the maximum elevation cannot be more than the radius of the domain's circumcircle, a limit which the solutions with self-weight frequently violate. This means that these optimal solutions could not be obtained by approximating the self-weight as an external load, even if an iterative approach was used to re-distribute the load in the most appropriate proportion.

Whilst the overall increase in optimal height with increasing influence of self-weight has been previously noted for simpler structures, the numerical method presented here allows more complex scenarios which show that it is not necessarily true that elevation should increase at *all* points in a structure. Particularly when vertical-only supports are present, reducing the rise in some areas can be beneficial in generating required horizontal thrusts.

Furthermore, it has been shown that the optimal topology can change markedly when self-weight becomes significant. The use of an incorrect topology in either direction (i.e. the optimal topology found for the weightless case being used for a problem where self-weight is significant, or vice versa) is shown to increase the material demands of the structure. Nonetheless, it is generally observed that the penalty for an inefficient topology increases as self-weight becomes more significant, echoing previous findings in optimization of other structure types.

Finally, it has been shown that when self-weight loading is considered, it is possible for the optimal structure to require joints at points which are not the locations of external loads or supports. This increases the need to optimize the topology in these scenarios, in conjunction with the elevation function. Furthermore, it can be seen from the results that the chosen topology and the vertical shape are closely linked, highlighting the importance of the coupled optimization method presented here.

---

<sup>4</sup>The exception to this is a case where there are two or more equally optimal solutions, (e.g. in the top row of Fig. 14 either of the two crossing arches is a valid optimal, basic, statically determinate solution). In this case, any convex combination of the solutions is a valid plastic design, but as it is statically indeterminate it *may* not be elastically valid. The interior point method used to solve the conic problem (20) will generally return a mixed, non-basic solution, whereas if the simplex method is used to solve (5) then one of the basic solutions will be returned. Equally-optimal solutions are observed most frequently in cases with a single point load and an axisymmetric domain (e.g. Fig 14), wherein the symmetry implies that the statically indeterminate structures also satisfy compatibility. Nonetheless, there may be cases (albeit not yet encountered by the authors) where obtaining the basic, statically determinate structure becomes a necessity.

Overall, the method presented here provides a powerful tool for investigating this class of material-efficient structures. This study has shown that these structures have many interesting properties, many of which demonstrate significant challenges to purely intuitive design approaches, or to existing fixed-topology methods. By overcoming these issues, this method can facilitate the adoption of such structural forms, and Python scripts and Grasshopper files implementing examples from this paper have been made available [29] to encourage this.

#### *Declaration of Competing Interest*

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### *CRedit authorship contribution statement*

**Helen Fairclough:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft. **Karol Bołbotowski:** Formal Analysis, Methodology, Writing - original draft. **Linwei He:** Software, Writing - review and editing. **Andrew Liew:** Writing - review and editing. **Matthew Gilbert:** Conceptualization, Writing - review and editing.

## References

- [1] J. Chilton, G. Tang, Timber gridshells: architecture, structure and craft, Routledge, 2016.
- [2] S. Adriaenssens, P. Block, D. Veenendaal, C. Williams, Shell structures for architecture: form finding and optimization, Routledge, 2014.
- [3] K. Linkwitz, Force density method: design of a timber shell, in: Shell Structures for Architecture, Routledge, 2014, pp. 59–70.
- [4] Y. Jiang, T. Zegard, W. F. Baker, G. H. Paulino, Form-finding of gridshells using the ground structure and potential energy methods: a comparative study and assessment, Structural and Multidisciplinary Optimization 57 (2018) 1187–1211.
- [5] M. Hüttner, J. Maca, P. Fajman, The efficiency of dynamic relaxation methods in static analysis of cable structures, Advances in Engineering Software 89 (2015) 28–35.
- [6] H. Deng, Q. Jiang, A. Kwan, Shape finding of incomplete cable-strut assemblies containing slack and prestressed elements, Computers & structures 83 (2005) 1767–1779.
- [7] J. N. Richardson, S. Adriaenssens, R. F. Coelho, P. Bouillard, Coupled form-finding and grid optimization approach for single layer grid shells, Engineering structures 52 (2013) 230–239.
- [8] J. N. Richardson, S. Adriaenssens, R. F. Coelho, P. Bouillard, Discrete topology optimization: connectivity for gridshells, in: Shell Structures for Architecture, Routledge, 2014, pp. 171–180.

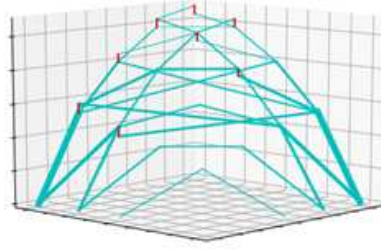
- [9] W. Gythiel, M. Schevenels, Gradient-based size, shape, and topology optimization of single-layer reticulated shells subject to distributed loads, *Structural and Multidisciplinary Optimization* 65 (2022) 144.
- [10] W. S. Dorn, R. E. Gomory, H. J. Greenberg, Automatic design of optimal structures, *J Mecanique* 3 (1964) 25–52.
- [11] M. P. Bendsoe, O. Sigmund, *Topology optimization: theory, methods, and applications*, Springer Science & Business Media, 2003.
- [12] M. Gilbert, A. Tyas, Layout optimization of large-scale pin-jointed frames, *Engineering computations* 20 (2003) 1044–1064.
- [13] M. Fuchs, E. Moses, Optimal structural topologies with transmissible loads, *Structural and Multidisciplinary Optimization* 19 (2000) 263–273.
- [14] H. Lu, A. Tyas, M. Gilbert, A. V. Pichugin, On transmissible load formulations in topology optimization, *Structural and Multidisciplinary Optimization* 64 (2021) 23–37.
- [15] M. P. Bendsøe, A. Ben-Tal, J. Zowe, Optimization methods for truss geometry and topology design, *Structural optimization* 7 (1994) 141–159.
- [16] H. E. Fairclough, M. Gilbert, A. V. Pichugin, A. Tyas, I. Firth, Theoretically optimal forms for very long-span bridges under gravity loading, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 474 (2018) 20170726.
- [17] T. Lewiński, T. Sokół, C. Graczykowski, *Michell structures*, Springer, 2018.
- [18] A. G. M. Michell, The limits of economy of material in frame-structures, *Phil Mag* 8 (1904) 589–597.
- [19] G. I. N. Rozvany, W. Prager, A new class of structural optimization problems: optimal archgrids, *Computer methods in applied mechanics and engineering* 19 (1979) 127–150.
- [20] G. I. N. Rozvany, H. Nakamura, B. T. Kuhnell, Optimal archgrids: allowance for selfweight, *Computer Methods in Applied Mechanics and Engineering* 24 (1980) 287–304.
- [21] G. I. N. Rozvany, C. M. Wang, C.-M. Dow, Prager-structures: archgrids and cable networks of optimal layout, *Computer Methods in Applied Mechanics and Engineering* 31 (1982) 91–113.
- [22] K. Bołbotowski, Optimal vault problem—form finding through 2d convex program, *Computers & Mathematics with Applications* 109 (2022) 280–324.
- [23] L. He, H. E. Fairclough, M. Gilbert, A. Liew, K. Bołbotowski, Minimum material vault designs generated via adaptive layout optimization, *Engineering Structures* 334 (2025) 120213.
- [24] S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge UP, 2004.
- [25] E. J. Routh, *A treatise on analytical statics*, volume 1, University Press, 1896.

- [26] T. V. Mele, A. Mehrotra, M. T. Echenagucia, U. Frick, J. Ochsendorf, M. Dejong, P. Block, Form finding and structural analysis of a freeform stone vault, in: Proceedings of IASS Annual Symposia, 8, International Association for Shell and Spatial Structures (IASS), 2016, pp. 1–10.
- [27] MOSEK ApS, MOSEK Optimizer API for Python 10.2, 2024. URL: <https://docs.mosek.com/10.2/pythonapi/index.html>.
- [28] S. Diamond, S. Boyd, CVXPY: A Python-embedded modeling language for convex optimization, *Journal of Machine Learning Research* 17 (2016) 1–5.
- [29] H. E. Fairclough, K. Bolbotowski, L. He, A. Liew, M. Gilbert, Example scripts for topology and geometry optimization of gridshells under self-weight loads, <https://doi.org/10.15131/shef.data.29046755>, 2025.
- [30] L. He, T. Pritchard, J. Maggs, M. Gilbert, H. Lu, Peregrine User Manual v7.1, 2025. URL: <https://www.limitstate.com/peregrine>.
- [31] Robert McNeel & Associates, Rhino 8 and Grasshopper, 2025. URL: <https://www.rhino3d.com/>.
- [32] G. I. Rozvany, On symmetry and non-uniqueness in exact topology optimization, *Structural and Multidisciplinary Optimization* 43 (2011) 297–317.
- [33] A. Ben-Tal, A. Nemirovski, Lectures on modern convex optimization: analysis, algorithms, and engineering applications, SIAM, 2001.

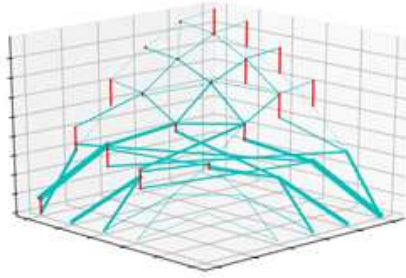
## Appendix A. Lumped mass approach

Within truss optimization, it is common to consider the effects of self-weight using a ‘lumped mass’ approach, where it is assumed that the weight of each element can be modeled as two equal point loads acting directly on the end-points of the element. This approach neglects the bending within a straight element which would be required to transmit the self-weight load to the nodes. However the simplicity of the model means it remains popular. This section will discuss the seemingly promising approach of combining this self-weight model with the vault design formulation. However, it will be shown that this does not produce valid solutions.

From the weightless vault formulation (1), it can be seen that the volume of an element  $i$  is given by  $\frac{l_i}{\sigma} \left( s_i + \frac{q_i^2}{s_i} \right)$ . Using the material’s unit weight  $\rho g$  and the definition of the variable  $r_i$  given in (2), then the lumped mass model in this case would involve the imposing of self-weight forces with magnitude  $0.5 \frac{\rho g}{\sigma} l_i (s_i + r_i)$  at each end of element  $i$ . Applying these forces involves adding these terms to the



(a)



(b)

Figure A.20: Lumped mass formulation: Results showing inconsistency in nodal elevation calculations. Vertical red lines link elements ends which should connect at a single point, highlighting the inconsistency

vertical equilibrium constraints, to give the optimization problem:

$$\min \sum_{i \in M} \frac{l_i}{\sigma_i} (s_i + r_i), \quad (\text{A.1a})$$

$$\text{s.t. } \mathbf{B}\mathbf{s} = \mathbf{f}_{xy}, \quad (\text{A.1b})$$

$$\mathbf{D}\mathbf{q} + \mathbf{Z}\mathbf{s} + \mathbf{Z}\mathbf{r} = \mathbf{f}_z, \quad (\text{A.1c})$$

$$(r_i \geq \frac{q_i^2}{s_i})_{\forall i} \quad (\text{A.1d})$$

$$\mathbf{s} \geq 0, \quad (\text{A.1e})$$

where the new  $\mathbf{Z}$  matrix contains zeros everywhere except at entry  $i, j$  when element  $i$  is connected to node  $j$ , at which point it contains the value  $Z_i = -0.5l_i \frac{\rho g}{\sigma}$ .

Problem (A.1) can be easily set up and solved for a given scenario. However, the issue arises in reconstructing a physical structure corresponding to the solution. This is not possible, as there is not a set of nodal elevations which are compatible with the forces,  $q$  and  $s$ , in all of the bars simultaneously. The variables  $q$  and  $s$  describe the vertical and horizontal components of the force in an element, and thus define the inclination of its centre-line<sup>5</sup>. From this, the elevation difference between

<sup>5</sup>Note that here we cannot assume  $q$  to include an additional mass contribution as used in Figure 5. In this formulation  $q$  defines two equal and opposite forces, so any variation in  $q$  would be imposed downward at one end of the element and upward at the other.

the two ends can be calculated. It is then possible to work through the structure, starting from the supported points with 0 elevation, and thereby establish heights for all of the nodes. (Note: this procedure was suggested and successfully used in the weightless case by [23] as a conceptually simpler alternative to the use of the dual variables and also as a way to validate the solutions.)

To demonstrate the issue, (A.1) has been solved for problems representing a quarter of the square-based domain problem, with distributed loads. When the procedure is followed, it is found that different routes through the structure result in different elevations of the same node. These errors are represented in Figure A.20 by the vertical red lines. This therefore shows that the solutions of (A.1) do not represent physically meaningful solutions.

Note that within the catenary formulation presented in this paper, the equivalent process of validation via primal variables is the calculation of (non-negative) lumped masses from the relaxed geometrical coupling condition described in 3.1.4. This has been successfully carried out for all examples herein, validating the analytical proof of correctness of the catenary approach.

## Appendix B. Analysis of the optimal primal and dual solutions

Below the properties (I)-(III) of the optimal primal and dual solutions  $\mathbf{s}, \mathbf{q}_A, \mathbf{q}_B$  and  $\mathbf{u}, \mathbf{w}, \mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$  are derived, see Section 3.2.2. It must be emphasized that the criteria from Section 3.2.1 are assumed to hold throughout. In particular, the strict inequality below holds for every node,

$$\mathbf{w} < \frac{1}{\rho g} \quad (\text{B.1})$$

*Proof of properties (I), (II):* These properties will be shown for every element for which  $s > 0$  at optimality.

First, from the complementary slackness for the primal constraint  $s \geq 0$ , it immediately follows that the dual linear constraint (21b) is an equality for such an element. This allows to explicitly express all  $g_j$  that enter the dual conic constraint  $2g_1g_2 \geq g_3^2$ ,

$$g_1 = \frac{1}{\sin \bar{l}} \left( \frac{1}{\rho g} - w_A \right), \quad g_2 = \frac{1}{\sin \bar{l}} \left( \frac{1}{\rho g} - w_B \right), \quad (\text{B.2})$$

$$g_3 = -\sqrt{2} \left( \frac{1}{2} \Delta u + \frac{\cos \bar{l}}{\sin \bar{l}} \left( \frac{1}{\rho g} - \frac{w_A + w_B}{2} \right) \right), \quad (\text{B.3})$$

where  $\Delta u$  represents the horizontal extension of the element, calculated as the relevant entry of  $\mathbf{B}^T \mathbf{u}$ . On the other hand, the primal conic constraint (20d) can be expressed as  $2t_1t_2 \geq t_3^2$ , where

$$t_1 = \sin \bar{l} q_A + \cos \bar{l} s, \quad (\text{B.4})$$

$$t_2 = \sin \bar{l} q_B + \cos \bar{l} s, \quad t_3 = \sqrt{2}s. \quad (\text{B.5})$$

It is well established that at optimality there holds a complementary slackness condition associated to the conic constraints, see Section 2.5 in [33],

$$t_1g_1 + t_2g_2 + t_3g_3 = 0. \quad (\text{B.6})$$

Since  $s > 0$ , there is  $t_3 > 0$  and thus also  $t_1, t_2 > 0$  due to the conic constraint. In addition,  $g_1, g_2 > 0$ , due to the assumption (B.1). As result, the condition (B.6)



implies that  $g_3 < 0$ . Starting from the simple inequality  $a^2 + b^2 \geq 2ab$ , the chain below follows,

$$\begin{aligned} t_1 g_1 + t_2 g_2 &\geq 2\sqrt{t_1 g_1} \sqrt{t_2 g_2} \\ &= \sqrt{2t_1 t_2} \sqrt{2g_1 g_2} \geq |t_3| |g_3| = -t_3 g_3, \end{aligned}$$

where the second inequality simply combines the two conic constraints. By virtue of the condition (B.6), this chain must be a chain of equalities only. The first equality implies that  $t_1 g_1 = t_2 g_2$ . Indeed, for positive numbers,  $a^2 + b^2 = 2ab$  ensures that  $a = b$ . Up to multiplying by  $\sin \bar{l}$  on each side, this is exactly the statement (III). Next, the equality previous to last ensures that both conic constraints are satisfied as equalities. In particular,  $2t_1 t_2 = t_3^2$ , which furnishes the statement (II) in the case  $s > 0$ . For  $s = 0$  it will follow from the assertion (III), proved below.

Note that (B.1) was fundamental, since, otherwise, the complementary slackness condition (B.6) could have been trivially satisfied for  $g_1 = g_2 = g_3 = 0$ . It was thus necessary to make sure that  $\mathbf{w} < \frac{1}{\rho g}$ .

*Proof of property (III):* For an element with  $s = 0$ , there must also be  $t_3 = 0$  and so the complementary slackness condition becomes

$$t_1 g_1 + t_2 g_2 + 0 = 0. \tag{B.7}$$

The definition of the primal cone requires  $t_1, t_2 \geq 0$ , while property (I) furnishes  $g_1, g_2 > 0$ . Hence both remaining terms are  $\geq 0$ , and so both must be equal to 0. This implies  $t_1 = t_2 = 0$ , which when combined with the initial definition  $s = 0$  proves that  $q_A = q_B = 0$ .