Topology optimization of compliant mechanisms with multiple degrees of freedom

Koppen, S.

DOI
10.4233/uuid:21994a92-e365-4679-b6ac-11a2b70572b7

Publication date
2022

Document Version
Final published version

Citation (APA)
Koppen, S. (2022). *Topology optimization of compliant mechanisms with multiple degrees of freedom*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:21994a92-e365-4679-b6ac-11a2b70572b7

Topology optimization of compliant mechanisms with multiple degrees of freedom

Topology optimization of compliant mechanisms with multiple degrees of freedom

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 28 November 2022 om 12:30 uur

door

Stijn KOPPEN

werktuigkundig ingenieur,
Technische Universiteit Delft, Nederland,
geboren te Haarlem, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. A. van Keulen,	Technische Universiteit Delft, promotor
Dr. ir. M. Langelaar,	Technische Universiteit Delft, promotor
Onafhankelijke leden:	
Prof. dr. M. Wallin,	Lund University
Prof. dr. O. Sigmund,	Technical University of Denmark
Prof. dr. S.D. Guest,	Cambridge University
Prof. dr. ir. P. Breedveld,	Technische Universiteit Delft
Prof. dr. ir. J. Herder,	Technische Universiteit Delft

Keywords: Topology optimization, Compliant mechanisms

Printed by: Gildeprint

Cover: Stijn Koppen

ISBN 978-94-6366-627-5

An electronic version of this dissertation is available at http://repository.tudelft.nl/.

Contents

Summary

High-tech equipment critically relies on the precise and reliable fine alignment of components such as mirrors and lenses for calibration and adaptation of instrumentation. To meet the ever-increasing requirements on precision, engineers typically resort to monolithic compliant mechanisms. These mechanisms gain mobility by deformation of the material, eliminating any friction and backlash. The design of compliant mechanisms with multiple degrees of freedom, so-called multi-DOF compliant mechanisms, is complex, and the resulting designs are sensitive to exhibit crosstalk between the actuation modes. The manual manipulation of coupled mechanisms is unintuitive and time-consuming, and automated actuation requires complex control scenarios. Computational approaches can greatly improve designing multi-DOF compliant mechanisms without such undesired characteristics.

Topology optimisation methods take a mathematical approach to designing a structure. Such methods optimize the material layout in a design domain for a given performance measure, considering a provided set of boundary conditions, loads and design constraints. Topology optimization methods have demonstrated capable as synthesis tools for designing single-DOF compliant mechanisms. The development of topology optimisation approaches for solving multi-DOF compliant mechanism design problems is relatively undeveloped and comes with severe challenges. These design problems typically involve many different loading conditions and stringent design requirements, increasing the complexity of the optimisation problem and required computational effort. Available formulations only partly address these issues and tend to be complex to understand, implement, and use or have limited applicability.

This dissertation focuses on developing topology optimisation approaches for synthesising multi-DOF compliant mechanisms with relatively short strokes, which justifies the use of linear elasticity theory. The objective is the development of a topology optimisation problem formulation that is simple to understand, implement and use, applicable to a wide range of problems and relatively computationally efficient.

When parts of the structure are forced into a prescribed motion, the energy contained in a compliant system is an indirect measure of the resistance to this motion. One can thus capture the characteristic stiffness of arbitrarily complex kinematics using a single energy measure. The main discovery of this study is that topology optimisation problem formulations based on specific combinations of such energy measures provide a unique combination of simplicity, versatility and computationally efficiency. While similar to the classic compliance minimisation problem, the proposed generalisation for compliant mechanism problems holds similar advantageous optimisation properties. It minimises the number and strictness of design constraints simplifying the optimisation problem. Despite the advantages, such integrated measures come with the loss of exact control over individual displacements and stiffnesses. This dissertation demonstrates the broad

applicability of this formulation to the design of high-resolution decoupled multi-DOF compliant mechanisms, as well as flexures and shape-morphing structures.

Furthermore, this dissertation studies the impact of design for additive manufacturing constraints on the optimization of compliant mechanisms. A critical observation to designing practically relevant compliant mechanisms is that design for additive manufacturing considerations predominantly impacts thin flexural elements. One may exploit the observation of local impact to reduce the typically negative impact of design for additive manufacturing constraints on the performance of the optimised compliant system. This dissertation introduces a computationally efficient approach to redesign the most critical regions of compliant mechanisms considering design for additive manufacturing constraints while minimizing the negatively influence on the mechanism performance. This redesign approach allows for high-resolution design and accurate modelling of sensitive flexures, providing solutions that are superior to imposing the same restrictions on the entire design domain without substantial additional computational cost.

This dissertation also addresses the aspect of computation effort. The relationship between input and output ports defines the working principle of a compliant mechanism. As a result, the response functions standard in multi-DOF design problems are typically a function of the motion at those ports, and the loads often apply to the same ports. This property provides the possibility to reduce computational costs.

Such optimisation problems are typically characterised by multiple combinations of boundary and loading conditions and many constraint functions, substantially increasing the computational cost of calculating the response functions and accompanying sensitivity analysis. By exploiting the characteristics of the multi-DOF compliant mechanism design problem and using static condensation, we demonstrate increased computational efficiency in solving problems with different boundary conditions. Although this is a well-known technique, the use of static condensation and corresponding advantages have not been studied in-depth in this context. The sensitivities of the procedure can be calculated without solving other systems of equations of high dimensionality, making this approach very suitable for use in gradient-based optimisation methods.

In addition to problems with varying boundary conditions, there is a significant potential for reducing the computational cost for problems involving similar boundary conditions, common in multi-DOF compliant mechanism design problems. Although not commonly detected, such problems contain linear dependencies between the encountered applied loads and adjoint loads. Manually keeping track of such dependencies becomes tedious for real-world design problems that become increasingly involved. This dissertation introduces a linear-dependency-aware-solver that can efficiently detect such linear dependencies between all loads to automatically avoid solving unnecessary equations.

In summary, insights and tools are provided to efficiently and effectively (re)design practically relevant high-resolution three-dimensional multi-DOF compliant mechanisms. Energy-based measures under prescribed motion scenarios offer a versatile and straightforward basis for optimising problem formulations, allowing quantitative control over mechanism stiffness and motion transmission. We envision that such problem formulations will find widespread use in industry to design complex compliant systems such as implants, optical mounts and manipulation stages.

Samenvatting

High-tech apparatuur is kritisch afhankelijk van de kalibratie en afstelling van instrumentatie. Dit wordt gedaan middles zeer nauwkeurige en betrouwbare uitlijning van componenten zoals spiegels en lenzen. Om te kunnen voldoen aan de steeds hogere nauwkeurigheidseisen gebruiken ingenieurs vaak monolitische compliante mechanismen. Zulke mechanismen verkrijgen hun mobiliteit door middel van vervorming van het materiaal. Hierdoor worden zowel wrijving als speling gelimineerd. Het ontwerpen van compliante mechanismen met meerdere vrijheidsgraden, zogenoemde multi-DOF compliante mechanismen is complex. De resulterende ontwerpen zijn normaliter gevoelig voor parasitaire verplaatsingen. Handmatige actuatie van zulke gekoppelde systemen is niet intuitief en tijdrovend. Geautomatiseerde bendiening vereist complexe besturingsscenarios. Computer methoden kunnen van grote waarde zijn voor het ontwerpen van multi-DOF compliant mechanismen zonder dergelijke ongewenste kenmerken.

Topologie optimalisatie methoden benaderen het ontwerp van een structuur op een wiskundige wijze. Dergelijke methoden optimaliseren de verdeling van materiaal binnen een gegeven domein voor een gegeven doelfunctie, rekening houdend met randvoorwaarden, belastingen en ontwerpbeperkingen. Het is reeds aangetoond dat zulke methoden geschikt zijn voor het ontwerpen van single-DOF compliante mechanismen.

Dit proefschrift richt zich op het ontwikkelen van topologie optimalisatie methoden voor het ontwerpen van multi-DOF compliante mechanismen met relatieve kleine verplaatsingen en rotaties, wat het gebruik van lineare elasticiteits-theorie rechtvaardigt. Het doel is het ontwikkelen van een topologie optimalisatie probleem-formulering die eenvoudig te begrijpen, implementeren en gebruiken is, toepasbaar is op een breed scala aan problemen en rekenkundig relatief efficient is.

Wanneer delen van een structuur in een voorgeschreven beweging worden gedwongen, is de energie in een compliant systeem een indirecte maat voor de weerstand tegen deze beweging. Men kan dus de karakteristieke stijfheid van willekeurig complexe kinematica bepalen met behulp van een enkele energie term. De voornaamste ontdekking van dit proefschrift is dat topologie optimalisatie probleemformuleringen op basis van specifieke combinaties van dergelijke energie termen een unieke combinatie van eenvoud, veelzijdigheid en rekenefficiëntie bieden. De probleem-formulering minimaliseert het aantal en de striktheid van ontwerprestricties en vereenvoudigt het optimalisatie-probleem. Ondanks de voordelen gaat het gebruik van de voorgestelde formuleringen gepaard met het verlies van de exacte controle over individuele verplaatsingen en stijfheden. Dit proefschrift demonstreert de brede toepasbaarheid van deze formulering op het ontwerp van ontkoppelde multi-DOF compliante mechanismen met hoge resolutie, evenals flexures en vormveranderende structuren.

Deze proefschrift bestudeert tevens de impact van het ontwerpen voor additieve fabricagebeperkingen op de optimalisatie van compliante mechanismen. Een belangrijke observatie bij het ontwerpen van praktisch relevante compliante mechanismen is dat on-

twerpaanpassingen voor additieve fabricageoverwegingen voornamelijk invloed heeft op flexures. Dit proefschrift introduceert een rekenkundig efficiënte aanpak om de meest kritieke gebieden van compliante mechanismen opnieuw te ontwerpen, rekening houdend met additieve fabricagebeperkingen terwijl de negatieve invloed hiervan op de prestaties van het mechanisme wordt geminimaliseerd. Deze herontwerpbenadering maakt ontwerp met hoge resolutie en nauwkeurige modellering van gevoelige flexures mogelijk, waardoor ontwerpen worden verkregen die superieur zijn aan mechanismes waarbij het opleggen van dezelfde beperkingen aan het hele ontwerpdomein worden opgelegd, zonder substantiële extra rekenkosten.

Dit proefschrift beschouwd ook het aspect van de benodigde rekeninspanning. De kracht-verplaatsing relaties tussen de in- en uitvoerpoorten definieert het werkingsprincipe van een compliant mechanisme. Als gevolg hiervan zijn de standaard doelfuncties in multi-DOF-ontwerpproblemen typisch een functie van de beweging bij die poorten, en de belastingen zijn vaak opgelegd op dezelfde poorten. Deze eigenschap biedt mogelijkheden om de rekenkosten te verlagen.

Multi-DOF optimalisatieproblemen worden doorgaans gekenmerkt door toepassing van meerdere combinaties van randvoorwaarden en belastingscondities evenals een groot aantal beperkingen. Als gevolg hiervan nemen de kosten van het berekenen van de doelfuncties en de bijbehorende gevoeligheids-analyses substantieel toe. Door gebruik te maken van de kenmerken van het multi-DOF ontwerpprobleem en met behulp van statische condensatie kan een verhoogde rekenefficiëntie bij het oplossen van problemen met verschillende randvoorwaarden worden aangetoond. Hoewel statische condensatie een bekende techniek is, is het gebruik ervan in deze context nog niet eerder grondig bestudeerd. De ontwerpgevoeligheden van de procedure kunnen worden berekend zonder andere stelsels van vergelijkingen met een hoge dimensionaliteit op te lossen, waardoor deze benadering zeer geschikt is voor het gebruik in gradient gebaseerde optimalisatiemethoden.

Naast problemen met wisselende randvoorwaarden, is er een aanzienlijk potentieel voor het verminderen van de rekenkosten voor problemen met vergelijkbare randvoorwaarden, zoals ook vaak voorkomt bij multi-DOF ontwerpproblemen. Hoewel niet vaak opgemerkt, bevatten dergelijke problemen lineaire afhankelijkheden tussen de opgelegde krachten en de zogenoemde adjoint krachten. Bij steeds complexer wordende ontwerpproblemen wordt het handmatig bijhouden van dergelijke afhankelijkheden erg gecompliceerd. Dit proefschrift introduceert een lineare-afhankelijkheidsbewuste oplosmethode die dergelijke lineaire-afhankelijkheids tussen alle krachten efficient kan detecteren ter voorkoming van het onnodig oplossen van stelsels van vergelijkingen.

Samenvattend, inzichten en hulpmiddelen worden geboden om op efficiënte en effectieve wijze praktisch relevante driedimensionale multi-DOF compliante mechanismen met hoge resolutie te (her)ontwerpen. Op energie gebaseerde ontwerpcriteria onder voorgeschreven bewegingsscenario's bieden een veelzijdige en relatief simpele basis voor het ontwerp van probleemformuleringen, waardoor kwantitatieve controle over zowel stijfheid van het mechanisme als de verplaatsingsoverdracht mogelijk is. De voorgestelde probleemformuleringen hebben wijdverbreid gebruikspotentieel in de industrie om complexe, compliante systemen te ontwerpen, zoals implantaten, optische houders en manipulatie apparatuur.

1 Introduction

1.1 Background

From exploring the galaxy down to measurement and manipulation at the nanometer scale, the fields of space, medical and semiconductor instrumentation push the limits of existing technology. Such technology paves the way for innovations that sustain the society and economy. This high-tech equipment requires unprecedented and continuously increasing precision and reliability in the alignment of critical components such as mirrors and lenses. The alignment allows for instrument customization, pre-use calibration and in-use adaptation to fulfill next-generation system functionalities and performance.

To meet the unique requirements imposed on the design of such alignment mechanisms, engineers use short-stroke *compliant mechanisms* to eliminate the effects of friction and backlash inevitably present in traditional rigid-body mechanisms (Howell 2001), see for example the the prototyped spinal implant shown in Figure 1.1. Their monolithic nature makes compliant mechanisms the undisputed option for small-scale or harsh-environment applications (Kota *et al.* 2001). Compliant mechanisms gain their mobility from structural deformation. The lack of backlash and friction allows for precise manipulation. Unfortunately, plastic deformation limits the range of motion. The energy involved in the actuation of a compliant mechanism is wholly or partly stored in the structure as strain energy. This property opens up exciting possibilities to dissipate, store and transfer energy (Howell 2013). However, it makes the motion of compliant mechanisms dependent on the applied loading, which eliminates strict kinematics and drastically complicates systematic design (Herder 2017). Considering the limited desired range of motion short-stroke compliant mechanisms (relative to the build volume), both deflection and rotations are so small that the use of linear elasticity is assumed to be sufficiently accurate. The use of monolithic elastic structures makes the design of parallel mechanisms with multiple controllable degrees of freedom particularly prone to parasitic motions such as crosstalk. These coupling effects lead to unintuitive, complex and time-consuming manual manipulation. Reducing these effects is a crucial challenge, which gets more demanding with an increased number of degrees of freedom.

A fundamental building block of many compliant mechanisms is flexures. Geometrically simple elementary flexures such as beams and plates have been extensively studied,

1

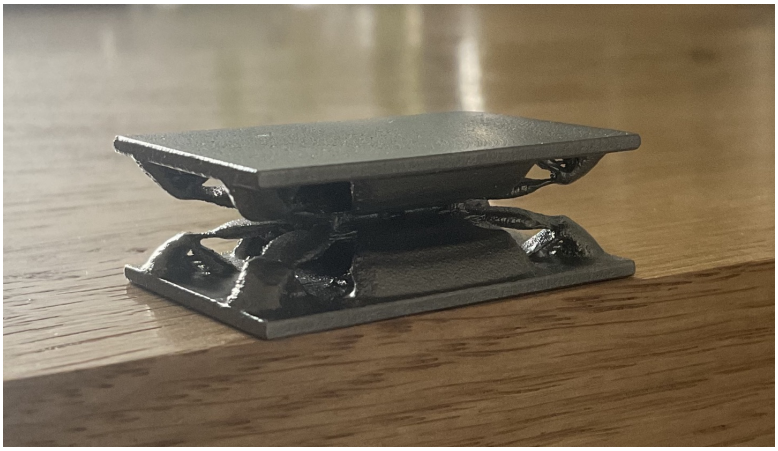


Figure 1.1: Prototyped conceptual design of a patient-specific compliant non-fusion spinal implant. This multi-DOF compliant mechanism has similar directional stiffness properties as intervertebral discs. The monolithic compliant device is used to correct the deformity in the spine as well as stabilize and strengthen the spine while retaining range of motion. The conceptual design has been performed using the methods that will be presented in this dissertation and in close collaboration with BAAT Medical Products (patent pending). The prototype is additively manufactured by Materialise via direct metal laser sintering of Titanium.

and their behaviour is well understood (Linss *et al.* 2019). The synthesis of compliant mechanisms with multiple degrees of freedom traditionally follows approaches that replace the joints of a rigid-body mechanism counterpart with such flexures, thereby introducing parasitic motions (Gallego *et al.* 2009). Such parasitic motions may be circumvented by approaches that combine elementary flexural elements, building complex, often symmetric, yet voluminous architectures. As a result, current manual or actuated adjustable mounts and mechanisms typically are voluminous and lack crosstalk-free control of individual degrees of freedom.

Gradient-based numerical optimization, and in particular topology optimization (Bendsøe 1989), can generate suitable topology, shape and dimensions for a multitude of single-DOF short-stroke compliant mechanism design problems (Frecker *et al.* 1997; Larsen *et al.* 1997; Sigmund 2001). Given an initial design, the topology optimization method optimizes the material layout for a combination of arbitrary design requirements. The optimization process maximally utilizes the available material and solves complex problems in a confined design volume. Topology optimization shows excellent potential to form the basis of a more systematic and general synthesis tool for designing short-stroke crosstalk-free multi-degree-of-freedom compliant mechanisms.

1.2 Focus and scope

This work focuses on developing topology optimization problem formulations to design compliant mechanisms involving multiple, decoupled and arbitrarily complicated degrees of freedom, such as decoupled adjustable mounts, flexures and shape-morphing structures. Aware of the numerous relevant design criteria, such as stress levels and manufacturability,

we primarily focus on the motion transmission and stiffness requirements of compliant mechanism synthesis.

We opt for the most basic and widely used design parameterization, analysis, and optimization algorithms to focus on the optimization problem formulation. We aim to minimize the need for other methodologies, such as variable projection and parameter continuation schemes. In line with the intended applications, we focus on designing compliant mechanisms involving relatively small displacements and deformations. A continuum mechanics model employing linear finite-elements for the elastic analysis is considered quantitatively sufficiently accurate. This linearization allows for a robust and efficient finite element analysis. However, it critically limits the optimization requirements to functions that depend on the behaviour in the initial configuration.

1.3 Challenges and aim

We will briefly introduce the trends in the topology optimization of compliant mechanisms with multiple degrees of freedom. A detailed review can be found in Chapter 3. The development of topology optimization problem formulations (*i.e.* choice of objective and constraint functions) for the synthesis of multi-degree-of-freedom compliant mechanisms is initiated around the twenty-first century by the works of Larsen *et al.* (1997) and Sigmund (2001). The proposed methods, as well as later variations (Alonso *et al.* 2014; Rong *et al.* 2021; Zhu *et al.* 2018), are straightforward extensions of the single-degree-of-freedom optimization problem formulations available and take on a qualitative approach, generally maximizing the motion or energy at the intended output(s). As such, the kinematics (desired motion transmission) is more or less an outcome of the optimization process, as opposed to input from the user. Motion decoupling is typically only partly considered by the addition of numerous explicit constraints.

A different branch of studies on this topic originates from a mechanism design philosophy, initiated by the works of Hasse *et al.* (2008) and Wang (2009a). Those proposed methods, and later extensions to multi-degree-of-freedom formulations (Jin *et al.* 2018; Kirmse *et al.* 2021c), formulate optimization problems based on a reduced model that describes the mechanical behaviour of the mechanism. Such reduced models appear a natural bridge between mechanism design and topology optimization, allowing the problem formulations to be expressed in terms of functions of the mechanisms' intrinsic properties independent of loading conditions, as typically used in traditional compliant mechanism design. However, existing problem formulations of this nature tend to use highly nonlinear response functions, resulting in a relatively complex understanding and use.

Despite more than two decades of research, and in contrast to 'design for stiffness', existing engineering software does typically not contain the tools to design compliant mechanisms via topology optimization, let alone the design of complex multi-degree-of-freedom variations. A straightforward translation of design requirements to a topology optimization problem formulation generally leads to an unnecessary complex optimization problem. Especially for the design of compliant mechanisms, the number of loading conditions and response functions grows fast with the number of degrees of freedom of the mechanism. Combined with the need for high-resolution three-dimensional designs, this puts a significant burden on optimization and computational costs. Existing formulations tend to be complex to understand, implement or use, applicable to a limited range of

problems, and do not provide the user with the most relevant design parameters such as motion transmission and mechanism stiffness. Thus, current topology optimization problem formulations for synthesizing multi-degree-of-freedom compliant mechanisms face challenges in simplicity, versatility and computational efficiency.

The goal of this dissertation is the development of a topology optimization problem formulation for the design of multi-DOF compliant mechanisms that is versatile, while remaining simple to understand, implement, and use. To achieve this goal, we first aim to research the optimization problem properties and ensure the proposed problem formulation is easy to solve and naturally converges to a meaningful solution. Next we provide a formulation that remains competitive in terms of required computational effort for an increased number of design variables and mechanism degrees of freedom. We seek to define the minimum set of parameters that define the proper functioning of various compliant mechanism problems, considering the essential design requirements per standard industry practice.

1.4 Outline

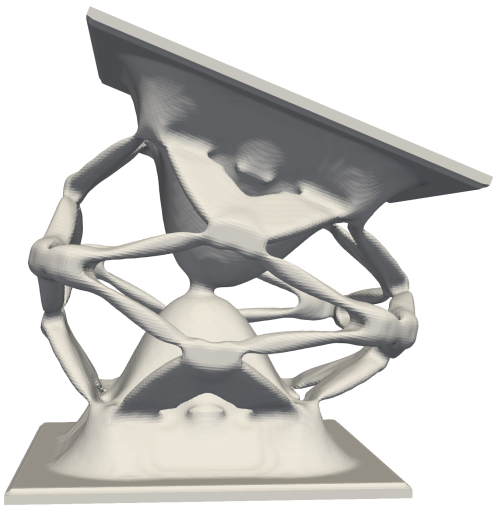
This dissertation is based on peer-reviewed and submitted papers. The upcoming chapters are slight modifications of the papers as published. Considering each paper is stand-alone, repetition of introductory content is inevitable, and we ask for the reader's tolerance for this inconvenience. Definitions and symbols are primarily consistent throughout the chapters, though, for clarity, these are (re)defined throughout. There are no strict dependencies between chapters or preferences for reading sequences.

The remainder of this dissertation is organized based on two related but independent parts. The first part (Chapters 2 to 4) focuses on the development of topology optimization problem formulations for the (re)design of (multi-degree of freedom) flexures and compliant mechanisms. Chapters 2 and 3 propose novel topology optimization problem formulations for synthesizing short-stroke flexures and multi-degree of freedom compliant mechanisms. Both proposed formulations come with associated source code to encourage the formulation to be explored and applied in academia and industry. Chapter 4 proposes a method to locally redesign selected refined local domains of a topology optimized compliant mechanism (*e.g.* as obtained via the formulations proposed in Chapters 2 and 3) whilst considering the global mechanism performance.

The second part (Chapters 5 and 6) focuses on developing techniques to reduce the high computational effort typically associated with the analysis and computational design optimization of structures subjected to multiple loading conditions and response functions, as typical in topology optimization multi-degree of freedom compliant mechanisms, among others. Chapter 5 specifically aims to reduce the computational cost of calculating the responses and accompanying design sensitivities associated with problems involving multiple partitions of the same discretization, typically corresponding to different loading scenarios. In contrast, Chapter 6 aims to increase computational efficiency by exploiting possible linear dependencies between the adjoint loads (loading terms for obtaining design sensitivities) and applied loads in the design of structures subject to multiple loads and multiple constraints but single system matrix partition. In Chapter 8 we conclude this work by discussing a selected set of overarching conclusions and most relevant recommendations for future work.

2

Topology optimization of flexures



This chapter proposes a novel topology optimization formulation for the synthesis of short-stroke flexures uniquely based on strain energy measures under prescribed displacement scenarios.

This chapter is based on peer-reviewed journal paper:
Koppen, S., Langelaar, M. and van Keulen, F. (2022). *A simple and versatile topology optimization formulation for flexure synthesis.* Mechanism and Machine Theory, 172.

2

A simple and versatile topology optimization formulation for flexure synthesis

Abstract *High-tech equipment critically relies on flexures for precise manipulation and measurement. Through elastic deformation, flexures offer extreme position repeatability within a limited range of motion in their degrees of freedom, while constraining motion in the degrees of constraint. Topology optimization proves a prospective tool for the design of short-stroke flexures, providing maximum design freedom and allowing for application-specific requirements. State-of-the-art topology optimization formulations for flexure synthesis are subject to challenges like ease of use, versatility, implementation complexity, and computational cost, leaving a generally accepted formulation absent. This study proposes a novel topology optimization formulation for the synthesis of short-stroke flexures uniquely based on strain energy measures under prescribed displacement scenarios. The resulting self-adjoint optimization problem resembles great similarity to 'classic' compliance minimization and inherits similar implementation simplicity, computational efficiency, and convergence properties. Numerical examples demonstrate the versatility in flexure types and the extendability of additional design requirements. The provided source code encourages the formulation to be explored and applied in academia and industry.*

2.1 Introduction

A flexure is a monolithic compliant element that connects two or more (assumed) rigid links, allowing for selectively chosen movements. In contrast to conventional hinges, flexures achieve their range of motion through elastic deformation. The finite dimension and operation below a critical stress limit the attainable range of motion. Due to the monolithic nature, flexures hardly require maintenance and have a long lifetime if used within the intended range of motion. Due to the lack of friction and backlash, flexures have high repeatability in use. Given these advantages, flexures are commonly applied in precision applications such as positioning stages and optical mounts (Hu *et al.* 2017). The present work is focused on so-called short-stroke flexures, for which—in contrast to large-stroke flexures—the assumptions of a linear stress-strain *and* a linear strain-displacement relationship suffice.

Flexures are engineered to have desired characteristic stiffness for specific relative rigid link movements. These movement are hereafter denoted by *Motion Patterns* (MPs). The mechanism Degrees of Freedom (DOFs) are the number of MPs with relatively low characteristic stiffness; that are the *free motion patterns.* In contrast, the mechanism Degrees

Figure 2.1: Computational design of a multi-axis flexure and its prototyped counterpart. The flexure is compliant in the rotations about the x and y-axis while stiff in the x, y and z translations as well as rotation about the z-axis.

of Constraint (DOCs) are the number of MPs with a much higher characteristic stiffness. An example of a complex three-dimensional flexure with two mechanism DOFs and four mechanism DOCs is shown in Figure 2.1.[1]

Common single-axis flexures are (i) compliant revolute joints[2], such as notch hinges that allow relative rotation *about* a single axis (Machekposhti *et al.* 2015), and (ii) compliant prismatic joints[3], such as a pair of parallel leaf-springs, that allow relative motion *along* a single axis. Common multi-axis flexures are, *e.g.*, compliant cylindrical, universal, spherical and planar joints (Machekposhti *et al.* 2015). Complex flexures typically combine multiple primitive flexures as building blocks, thus enabling more complex kinematics (Gallego *et al.* 2009). Also, flexures can be classified by the degree of localization of the deformation, ranging between lumped (*i.e.* highly localized) and distributed compliance (Yin *et al.* 2003).

The primary design requirement of a short-stroke flexure is the relative stiffness between the mechanism DOFs and DOCs. Secondary considerations are range of motion, axis drift, deformation and stress, fatigue, volume and mass, as well as the sensitivity of those aspects to, *e.g.*, manufacturing errors. The synthesis methods often used for rigid-body mechanisms, cannot straightforwardly be applied to compliant mechanisms. There is always mechanical stress involved in any motion, and the behaviour is dependent on the loading condition. This implies that kinematics (motion) and kinetics (load case) must be treated simultaneously. As a result, the concept of mechanism DOFs fades in compliant mechanisms, because they behave differently for any loading conditions (Herder 2017). Furthermore, the complex deformation and motion behaviour of compliant mechanisms complicates both their accurate analytical modeling as well as purposeful design. Hence, the synthesis process is iterative, and often time-consuming (Linss *et al.* 2019).

Systematic flexure synthesis methods rely on kinematic or building block approaches,

[1] A movie of the prototypes in motion can be found on Youtube (Koppen 2020).
[2] Also called (flexural) hinge, flexure bearing or flexure pivot.
[3] Also called translational (flexure) hinges.

such as rigid-body replacement techniques or the 'freedom and constraint topology' method (Gallego *et al.* 2009). However, these approaches do not exploit the full range of design possibilities. The use of gradient-based structural optimization techniques to design flexures has gained increasing interest because of the possibility to design optimized flexures, satisfying application-specific requirements (Zhang *et al.* 2018). Topology Optimization (TO) in particular, allows for maximum design freedom, while requiring minimal designer input regarding the flexure concept (Bendsøe *et al.* 2004).

Owing to the potential benefits of TO, academics, engineers and designers could benefit from a versatile, simple, easy to implement and use as well as computationally efficient TO method for short-stroke flexure design. Multiple different TO problem formulations are previously proposed, see, *e.g.*, (Hasse *et al.* 2009; Pinskier *et al.* 2019, 2020; Wang 2009a; Zhu *et al.* 2014). Section 2.2 provides a comparison of previously proposed TO problem formulations for flexure design and addresses the remaining challenges in the field with respect to simplicity, versatility and computational effort.

To address the challenges, in Section 2.3 we propose a novel and intuitive topology optimization formulation to design flexures. The basic idea is to maximize the stiffness of *a priori* defined constrained MPs, whilst imposing an upper bound on the stiffness of *a priori* defined free MPs. Motion pattern stiffnesses are evaluated via strain energies under prescribed movements of the rigid links. The contribution of this work is thus on the use of energy-based response functions under *prescribed displacement conditions* and the manner in which these response functions are combined to form the optimization *problem formulation*. The contribution, advantageous properties of the problem formulation and consequential simplicity, versatility and computational efficiency is elaborated on in Section 2.3.2.

The basic formulation proposed in this work focuses on the primary design requirement, that is maximization of the relative stiffness between the free and constraint MPs. Implementation specific considerations are discussed in Section 2.4, followed by numerical examples in Section 2.5. We will additionally demonstrate the ease and influence of taking into account stress considerations as well as manufacturing robustness in Section 2.5, all within the limits of linear elasticity theory. The manuscript is completed with stating the limitations of the proposed formulation, providing recommendations for future work and concluding remarks.

2.2 Comparison of existing formulations

Currently, a good comparison between different TO formulations to synthesize flexures is absent. To compare different formulations, we define the following three quality criteria:

- simplicity,
- versatility, and
- computational effort.

We define simplicity as the ease of understanding, implementation and use of the formulation. This includes the number of parameters required to define the optimization problem and the ease of assigning an appropriate value to those parameters. Versatility is

the applicability of the method to a wide range of uses *e.g.* planar to three-dimensional or single-axis to multi-axis flexures. The total computational effort to obtain an optimized design in a nested analysis and design process depends on the number of design iterations and the effort per design iteration. The number of design iterations is highly dependent on the ease of solving the resulting optimization problem and, thereto, the complexity (from an optimization point of view) of the optimization problem formulation. The main contribution to the computational effort per design iteration is the number of analyses and their expense, such as a Finite Element Analysis (FEA). The effort of an analysis can be predominantly separated in the effort of the preconditioning/factorization (most expensive) and the iterative solve(s)/back-substitution(s), for iterative and direct solution approaches, respectively. By focusing on the chosen quality criteria, relevant criteria such as the control of range of motion, feature size, compliance distribution, stress levels and parasitic motion (*e.g.* change of rotational center), although relevant, are not considered in this comparison.

Below different approaches to flexure design using TO are discussed from the perspective of the aforementioned quality criteria. The aim of the discussion is to provide a concise overview of the field and build the argumentation for the present work. Thereto, in-depth review and/or comparison is out of the scope of this work. For a detailed description of the formulations the reader is referred to the relevant contributions, as presented in the first column of Table 2.1. This table presents quantifiable measures of the quality criteria for a set of distinct topology optimization problem formulations. The following discussion adopts a categorization in kinetostatic and kinetoelastic formulations as proposed by Wang (2009a).

Kinetostatic formulations

Naturally, TO problem formulations for flexure design find their origin in the field of compliant mechanism design. Kinetostatics[4] deals with the determination of forces that act upon the elements of a mechanism, given the mechanical system acts as a *static* construction (Burns *et al.* 1968). The so-called kinetostatic formulations, in one form or another, simultaneously aim to maximize the energy transmission between the input and output ports and mechanism's structural stiffness (Wang 2009a). The mechanism performance is generally quantified using the concepts of mechanical advantage, geometric advantage, mechanical efficiency, flexibility-stiffness or mutual potential energy (Cao *et al.* 2013). Although there is no single universally accepted formulation, it has been shown that these formulations produce almost similar topologies for the optimized compliant mechanisms (Deepak *et al.* 2009), *viz.* these topologies tend to emulate their rigid-body counterpart (Wang 2009a).

Derived from these kinetostatic formulations for compliant mechanism design, Zhu *et al.* (2014) and Pinskier *et al.* (2020) independently proposed straightforward approaches for designing planar single-axis (prismatic and/or revolute) joints, taking into account axis drift. The formulation of Zhu *et al.* (2014) has been extended to account for geometric nonlinearity (Liu *et al.* 2015), stress constraints (Liu *et al.* 2017b), stress and compliance distribution (Liu *et al.* 2020, 2018), and prescribed stiffness characteristics (Liu *et al.* 2021b). The formulation is utilized in several studies on the topology optimization of application specific flexures (Liu *et al.* 2016, 2017a; Qiu *et al.* 2019, 2020).

[4] Often referred to as 'inverse dynamics'.

Table 2.1: Topology optimization problem formulations for flexure design versus quantifiable measures of the quality criteria. The papers are ordered by year, ending with the present contribution. Versatility is expressed by the demonstrated range of applications (types of joints, single or multi-axis) and dimensionality (2D or 3D). The Parameters column denotes the *minimum* number and type of parameters required to set up the formulation. Implementation includes notable features, such as the type of analyses and responses. The computational effort of a single design iteration is dominated by the effort of finite element and sensitivity analyses. The last column indicates, subsequently the number of (i) preconditioning/factorization steps, (ii) physical loads, and (iii) additional adjoint loads per design iteration. The sum of the loads indicates the number of iterative solves/back-substitutions required. For fair comparison all listed numbers of parameters and loads are for a single-axis flexure formulation.

Paper	Dim	Versatility	Parameters	Implementation	Effort
Hasse *et al.* 2009	2D	any single-axis joint	max volume eigenmode	static condensation orthogonalization eigensystem analysis	1, 2, 0
Wang 2009a	2D	any single or multi-axis joint	max volume eigenmode	static condensation eigensystem analysis	1, 2, 0
Zhu *et al.* 2014	2D	prismatic and revolute joint	max volume non-design domain size max axis drift spring stiffness	non-design domain exotic responses additional spring	1, 2, 1
Pinskier *et al.* 2019	3D	leaf flexure	max volume max strain energy	strain energy based	1, 2, 0
Pinskier *et al.* 2020	2D	revolute joint	max volume non-design domain size max displacement	non-design domain exotic responses	1, 3, 3
Present	2D,3D	any single or multi-axis joint	max strain energy	strain energy based	1, 2, 0

Pinskier *et al.* (2019) additionally proposed a simple and intuitive TO formulation aimed at the synthesis of leaf-springs using only strain-based measures. As a result, the formulation is simple and computationally efficient.

A variety of studies derive the topology optimization response functions from design requirements of their application, such as the design of a structural flexure for force sensing in a wind tunnel balance (Sung *et al.* 2022), the design of flexures for mounting of mirrors (Hu *et al.* 2017; Koppen *et al.* 2018; Liu *et al.* 2021a), and the redesign of flexural hinges for compliant mechanisms (Koppen *et al.* 2021a).

Kinetoelastic formulations

As opposed to kinetostatic formulations, kinetoelastic formulations consider the mechanism's kinematic functions as an integral part of the *elastic* properties of the continuum structure and seek to find compliant mechanisms with desirable intrinsic properties (Wang 2009a). This is, thus far, accomplished by shaping the mechanism stiffness matrix entries. The mechanism stiffness matrix is obtained by static condensation of the global stiffness matrix to a small set of nodal displacementsthat can describe the MPs (Guyan 1965; Irons 1965). The formulation was effectively applied to the design of planar prismatic joints (Wang 2009a,b; Wang *et al.* 2009), and revolute joints (Li *et al.* 2019).

From a shape-morphing design philosophy, Hasse *et al.* (2009) propose a kinetoelastic formulation to design compliant mechanisms with *selective compliance* by shaping the modal properties of the mechanism stiffness matrix (*i.e.* eigenmodes and eigenvalues). Compliant mechanisms with selective compliance combine the advantages of both lumped and distributed compliance, that is reduced stress concentrations and a distributed deformation pattern, while preserving defined kinematics (Hasse *et al.* 2017). The method has been improved upon (Hasse *et al.* 2017; Kirmse *et al.* 2021a), and extended to multiple mechanism DOFs (Kirmse *et al.* 2021c).

The kinetoelastic formulations of Hasse *et al.* (2009) and Wang (2009a) use static condensation to obtain the mechanism stiffness matrix. This procedure requires an expensive analysis which scales with the number of nodal displacements required to describe the MPs. Thereto, this is highly efficient for problems like single-input-single-output compliant mechanisms, for which the MP can, generally, be described using only two nodal displacements (Koppen *et al.* 2022c). However, for the aforementioned problem formulations, a vast number of nodal displacements are required to describe the MPs. As a result, applying static condensation (without further adaptation) to flexure design would generally require substantial high computational effort.

Concluding remarks

Despite the attention devoted to TO of flexures, the previously proposed formulations have disadvantages and pose challenges, see also Table 2.1. The kinetostatic formulations are straightforward but tend to be specific for a small set of flexures (Pinskier *et al.* 2019, 2020; Zhu *et al.* 2014). In contrast, the kinetoelastic formulations are versatile, however are generally more complex to implement (Hasse *et al.* 2009; Wang 2009a). Several formulations include responses that depend highly nonlinear on the nodal displacements (Pinskier *et al.* 2020; Wang 2009a) or make use of artificial stiffness and additional user-defined parameters (Zhu *et al.* 2014), that can make application difficult. Some show inferior convergence properties (many iterations or oscillatory behaviour) and/or deliver non-binary (and hence non-manufacturable) topologies due to absence of conflicting requirements (Pinskier *et al.* 2020; Zhu *et al.* 2014). Finally, some formulations require substantial computational effort, which makes application of the method unpracticable. To conclude; none of the previously proposed formulations is simple to understand, implement and use *as well as* versatile *and* computationally efficient.

2.3 Method

Consider a structure within a bounded domain Ω, made of an isotropic linear elastic material. For simplicity of explanation, we discretize the domain in a structured grid of N finite elements ($\texttt{nelx} \times \texttt{nely}$)[5] with a total of n nodal displacements, as sketched in Figure 2.2. Let us define a set $\mathbb{M}$, consisting of unique free and constrained MPs. For example, consider the set $\mathbb{M} = \{\texttt{tx}, \texttt{ty}, \texttt{rz}\}$, in line with the 2D problem depicted in Figure 2.2. These MPs define prescribed nodal displacements at the interfaces between rigid link and flexure (*e.g.* a unit displacement in x-direction between top and bottom interfaces for mechanism degree $\texttt{tx}$). The assumption that these interfaces are rigid is valid if the links can be considered

[5] Typewriter font is used to indicate the names as used in the code associated with this work to replicate the results, see ??.

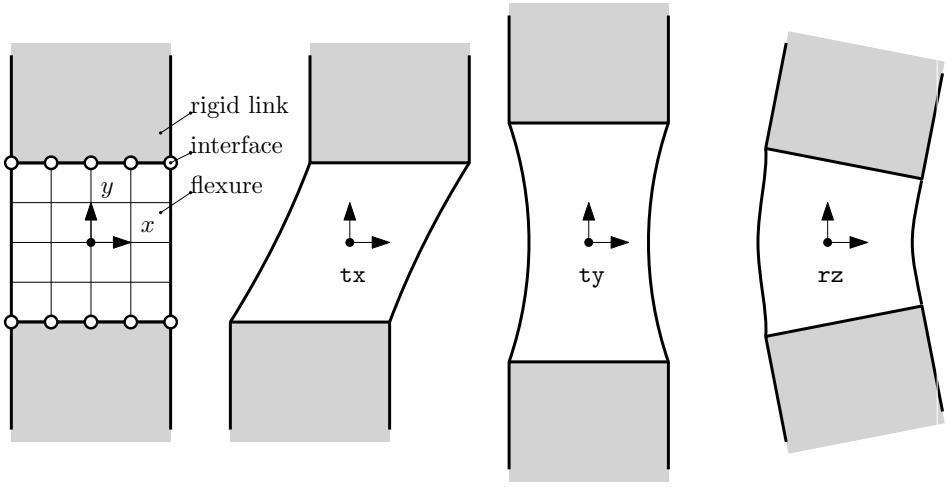


Figure 2.2: On the left: two rigid links (gray) connected via a flexure (white). The flexure geometry is discretized using `nelx`×`nely` finite elements. The interface nodes used to prescribe the MPs are here denoted by circles (◦). Those nodal displacements are used to prescribe different MPs. On the right: three different MPs commonly used in 2D flexure design; relative translation along the x-axis (`tx`) and y-axis (`ty`) and rotation about the z-axis (`rz`).

much stiffer compared to the flexure. As such, the MPs correspond to the relative rigid body motions of the interfaces.

We define subset $\mathbb{C} \subset \mathbb{M}$ that contains the constrained MPs, and subset $\mathbb{F} = \mathbb{M} \setminus \mathbb{C}$ that contains the free MPs. In line with the primary design requirement for short-stroke flexures, we aim to maximize the stiffness of the MPs in $\mathbb{C}$, while constraining the maximum stiffness of the MPs in $\mathbb{F}$.

The present work uses the strain energy of the MPs as a measure for stiffness. In contrast to the traditional compliance minimization under *applied loads*, minimization of strain energy under *prescribed displacements* results in *minimization* of corresponding stiffness (Klarbring *et al.* 2013). That is, stiffness maximization under *applied loads* equates to *minimization* of corresponding displacements, whereas under *prescribed displacements* this equates to *maximization* of corresponding reaction loads.

2.3.1 Optimization problem formulation

The proposed constrained nonlinear optimization problem formulation for flexure synthesis now simply reads

$$\mathcal{P} = \begin{cases} \underset{\mathbf{x}}{\text{minimize}} & -f\left[\mathcal{E}_i[\mathbf{x}]\right], \quad i \in \mathbb{C} \\ \text{subject to} & \mathcal{E}_j[\mathbf{x}] \leq \overline{\mathcal{E}}_j, \quad j \in \mathbb{F}\,, \\ & \mathbf{x} \in \mathcal{X}^N \end{cases} \tag{2.1}$$

where the dimensionless objective $f \in \mathbb{R}^+$ is a monotonically increasing function of strain energies $\mathcal{E}_i \in \mathbb{R}^+$ and $\overline{\mathcal{E}}_j \in \mathbb{R}^+$ is the maximum allowable strain energy of MP j. The topology

is described by N continuous differentiable design variables $\mathbf{x}$ with its components in $\mathcal{X} := \{x \in \mathbb{R} \mid 0 < x \leq 1\}$.

The strain energy of MP i in a discretized setting is defined as

$$\mathcal{E}_i[\mathbf{x}] = \frac{1}{2}\mathbf{u}_i \cdot \mathbf{K}[\mathbf{x}]\,\mathbf{u}_i, \tag{2.2}$$

where $\mathbf{K}[\mathbf{x}] \in \mathbb{R}^{n\times n}$ is the design dependent symmetric stiffness matrix and $\mathbf{u}_i \in \mathbb{R}^n$ contains the nodal displacements of MP i. These nodal displacements are obtained by analysis of the structural behaviour, described by n linear governing equations

$$\mathbf{K}[\mathbf{x}]\,\mathbf{u}_i = \mathbf{f}_i, \quad \forall\ i \in \mathbb{M}, \tag{2.3}$$

where $\mathbf{f}_i \in \mathbb{R}^n$ are the nodal loads of MP i. To calculate $\mathbf{u}_i$, we partition Equation (2.3) as

$$\begin{bmatrix} \mathbf{K}_{\mathrm{ff}} & \mathbf{K}_{\mathrm{fp}} \\ \mathbf{K}_{\mathrm{fp}}^{\mathsf{T}} & \mathbf{K}_{\mathrm{pp}} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{\mathrm{f},i} \\ \mathbf{u}_{\mathrm{p},i} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\mathrm{f},i} \\ \mathbf{f}_{\mathrm{p},i} \end{bmatrix}, \tag{2.4}$$

where $\mathbf{u}_{\mathrm{f},i}$ are the free nodal displacements, $\mathbf{u}_{\mathrm{p},i}$ the prescribed nodal displacements, $\mathbf{f}_{\mathrm{f},i}$ the applied nodal loads and $\mathbf{f}_{\mathrm{p},i}$ the nodal reaction loads of MP i. As mentioned, the MPs are defined purely in terms of prescribed nodal displacements at the interfaces, without additional applied loads. Hence, the applied loads $\mathbf{f}_{\mathrm{f}} = \mathbf{0}$ in all cases. The solutions to Equation (2.4), $\mathbf{u}_{\mathrm{f},i}$, can be obtained by solving the system of linear equations

$$\mathbf{K}_{\mathrm{ff}}\mathbf{u}_{\mathrm{f},i} = -\mathbf{K}_{\mathrm{fp}}\mathbf{u}_{\mathrm{p},i}, \quad \forall\ i \in \mathbb{M}. \tag{2.5}$$

The allowable strain energies of the free MPs are the primary design requirement. These energies are generally known from system requirements, either directly as energy term or indirectly as stiffness. The maximum allowable strain energy can be approximated by the desired free MP stiffness via simple one-dimensional equivalent models. For example, consider the MP `ty` from Figure 2 with a known desirable stiffness k and a prescribed relative displacement u between the interfaces. Then, the maximum allowable strain energy may be approximated by $\overline{\mathcal{E}}_{\mathtt{ty}} \approx \frac{1}{2}ku^2$. The desired stiffness can, if unknown, be derived from the required stroke for a given maximum actuation force or *vice versa* from the required actuation force for a given stroke.

Sensitivity analysis

TO generally requires the consecutive calculation of structural responses (objectives or constraints) and their sensitivity to the design variables. Both generally involve one or multiple computationally expensive FEA. For specific optimization responses—for example strain-energy—the problem becomes so-called 'self-adjoint' (Rozvany *et al.* 1993). In self-adjoint problems, the loading terms of the analyses required to obtain the structural response and sensitivity information are linearly dependent. As a result, the computational cost of the sensitivity analysis reduces dramatically. Note that this advantage is only applicable to the linear case, which is the focus of this study. All responses in $\mathcal{P}$ are self-adjoint. As a result, the sensitivities can be calculated based on available information.

In addition, the sensitivities are separable, *i.e.* each design variable contributes solely via its elemental strain energy. Thereto, one may write

$$\frac{\mathrm{d}\mathcal{E}_i}{\mathrm{d}x_j} = \gamma_{i,j}\left[x_j\right]\varepsilon_{i,j}, \tag{2.6}$$

with $\varepsilon_{i,j} \in \mathbb{R}^+$ and $\gamma_{i,j} \in \mathbb{R}$ the elemental strain energy and multiplication factor of element j due to degree i. The interpretation and derivation of $\gamma_{i,j}$ is further explained in Section 2.4.

As a consequence of the *prescribed displacements* scenarios—in contrast to 'classic' compliance minimization under *applied loads*—the sensitivities of strain energy, and thus the constraints, have a strictly positive sign. Intuitively this can be understood by the increase of an elastic body's strain energy under prescribed deformation upon increase of the Young's modulus. The sensitivities of the objective have a strictly negative sign due to the reformulation of a maximization problem to a minimization problem (max f equates to min $-f$).

2.3.2 Formulation properties

The contribution of this work is on the use of energy-based response functions under *prescribed displacement conditions* and the manner in which these response functions are combined to form the optimization *problem formulation*. The resulting advantageous properties of the optimization problem formulation pose benefits in terms of simplicity, versatility and computational efficiency.

Simplicity

The simplicity and effectiveness of the formulation is directly related to the similarity with the primary and most commonly used 'compliance minimization' TO problem formulation by Bendsøe (1989). The objective and constraints are monotonic functions with strictly opposite sign of design sensitivities, which proves a well-defined optimization problem. This results in a, relative to the state-of-the-art, easy to solve optimization problem with good convergence properties if a standard optimization is applied. The constraint(s) take over the 'role' of the volume constraint in Bendsøe (1989) to provide auto-penalization of design variables with intermediate values, which is evidenced in binary topologies.

The formulation requires a minimal number of independent parameters to define the optimization problem (only maximum strain energies of the free MPs), simplifying its use and circumventing the common 'trial-and-error' approach towards parameter value selection.

The formulation is uniquely based on strain energy measures. This makes implementation in/in combination with commercial FEA software packages simple, as such packages generally make this data accessible for the user. Since element strain energies (or elemental stiffness matrices in combination with the nodal displacements) are common output data in commercial finite element analysis software, also the sensitivity analysis is straightforward to implement, even when using software packages that do not already provide sensitivity information.

Most of the existing formulations share one of the above advantageous properties, see Table 2.1. However, none of the formulations simultaneously show ease of implementation and use, a minimal number of parameters, and a well-defined and easy to solve optimization problem (demonstrated by fast and smooth convergence).

Versatility

All response functions in the problem formulation are of the same form; that is strain energy measured under prescribed displacement conditions. Note that, due to this generality of the method, the problem formulation can include one or multiple constrained MPs in the objective while constraining the stiffness of one or multiple free MPs. As such, the formulation can be used to design both single-axis as well as multi-axis flexures.

What is more, the MPs are defined by the user, and are not restricted to specific geometries, design domains or applications. Although the proposed optimization problem formulation is relatively simple and only involves strain energy contributions from the considered MPs, it is thus effective in generating many types of flexures, as will be shown in Section 2.5.

Computational effort

Independent of the number of MPs, the formulation requires a single factorization/preconditioning step plus one back-substitution/iterative solve per MP. Thus the majority of the computational effort does not scale with the number of MPs. What is more, the optimization problem is self-adjoint and obtaining the sensitivity information requires negligible computational effort. Note that multiple formulations share this computational efficiency, as shown by the comparison in Table 2.1.

2.4 Implementation

Independent of the problem formulation as presented, the user has to consider, select and implement a variety of methods to effectively use the formulation in a TO setting. Without loss of generality, the following aids in the consideration and implementation of design parametrization, filtering, material interpolation, response formulation and gradient-based optimization. All numerical examples employ the implementation choices described here. The default constants used in the examples, as implemented in the attached code, are listed in Table 2.2.

For the Finite Element Analysis (FEA), we opt for standard 4-node quadrilateral (2D) and 8-node hexahedral elements (3D) in structured meshes. The design domain is parametrized by assignment of a design variable $x_i \in \mathcal{X}$ to each finite element i, which allows for local control of the material properties (Bendsøe *et al.* 2004).

It is generally recognized that both final design and performance are sensitive to the initial design $\mathbf{x}^{\{0\}}$. This is especially the case for compliant mechanisms, and thus also for flexure optimization (Chen *et al.* 2017; Sigmund *et al.* 1998). We consider this influence out of the scope of this paper and thereto opt for the commonly used homogeneous initial design.

To eliminate modeling artifacts, the design variable field is generally blurred as to obtain the filtered field $\tilde{\mathbf{x}} \in \mathcal{X}^N$ using a linear filtering operation $\mathbf{H}[\mathbf{x}] : \mathcal{X}^N \rightarrow \mathcal{X}^N$ with relative filter radius $r \in \mathbb{R}^+$, see *e.g.* Bruns *et al.* (2001). This operation is also accounted for in the sensitivity calculation, as described in the cited reference.

Asymmetric topologies resulting from problems with symmetric boundary conditions is, although not often explicitly reported, common and inevitable. The gradient-based optimizer solves many independent convex problems until a finite convergence criterion is met. As a result, round-off errors are inevitable. This leads, in most cases, to divergence

Table 2.2: Constant parameters and assigned values.

Symbol	Description	Value
ε	Stiffness ratio	10^{-6}
ν	Poisson ratio	0.3
p	SIMP penalty	3.0
r	filter radius (no. elements)	2.0
ϵ	maximum design change	10^{-3}
x^0	homogeneous initial design	0.5

from the symmetric local optimum. One may easily enforce symmetry by linking design variables over one or multiple axes; either by creating a dependency or by averaging.

The Young's modulus of an element is related to the filtered design variable via a element-wise composite rule, that is

$$\frac{E_i[\tilde{x}_i]}{\overline{E}} = \varepsilon + (1-\varepsilon)R[\tilde{x}_i], \tag{2.7}$$

with $\overline{E}$ the material Young's modulus, ε the relative stiffness between solid and void and R the material interpolation function. We apply the commonly used modified Solid Isotropic Material with Penalization (SIMP) interpolation function proposed by Bendsøe (1989), that is

$$R[x] = x^p, \tag{2.8}$$

with $p \in \mathbb{R}^+$ a user definable parameters. It is commonly known that this interpolation function increases the probability to obtain a 0/1 solution of a strain-based optimization problem. Note that, as a result, the elemental multiplication factor is simply obtained via

$$\gamma_{i,j} := (1-\varepsilon)\frac{\partial R_i}{\partial x_j}. \tag{2.9}$$

It is generally beneficial to scale the objective such that it holds a reasonable value (as compared to the constraints) Svanberg 1987. We opt to normalize the strain energy of degree i to its strain energy at the first optimization iteration, that is

$$\alpha_i^{\{k\}} := \frac{\mathcal{E}_i^{\{k\}}}{\mathcal{E}_i^{\{0\}}} \tag{2.10}$$

with α the relative strain energy and k the optimization iteration counter. Note that, as a result of Equation (2.10), the normalized strain energy α is a dimensionless positive scalar value by definition and $\alpha_i^{\{0\}} = 1$ for all i.

In order to simultaneously maximize the stiffness of multiple MPs, the corresponding normalized strain energies are combined in a monotonically increasing function. We opt here for a simple summation, that is the objective at iteration count k yields

$$f^{\{k\}}[\mathcal{E}_i] := \sum_i^{|\mathbb{C}|} \alpha_i^{\{k\}}[\mathcal{E}_i] \tag{2.11}$$

with $|\mathbb{C}|$ the DOCs. One might, in addition, add weight factors to the individual strain energy measures to control relative importance or opt for a smooth minimum function (Ma *et al.* 1994). Note that, as a consequence of Equation (2.10), the magnitude of the prescribed displacements for different constrained MPs become irrelevant.

The gradient-based inequality-constrained nonlinear optimization problem $\mathcal{P}$ is solved in a nested analysis and design setting. The design variables are iteratively updated by a sequential approximate optimization scheme, as is common in the topology optimization field. We use the Method of Moving Asymptotes (MMA) by Svanberg (1987). The resulting convex sub-problems are solved using a primal-dual interior point method. The optimization is terminated when the maximum design change is smaller than ϵ.

This work includes a MATLAB code to design 2D single and multi-axis flexures, which can be found on Koppen *et al.* (2022d). We provide a briefly introduction here to allow the reader to understand how to replicate the results in upcoming sections. The code is an extension of the commonly used `top71.m` code by Andreassen *et al.* (2010), and can be called using a similar syntax, that is

```
flexure(nelx, nely, doc, dof, emax)
```

where `doc` and `dof` are lists of strings of the names of the desired constrained and free MPs, respectively. Parameter `emax` is a list of maximum allowable strain energies corresponding to the free MPs in `dof`.

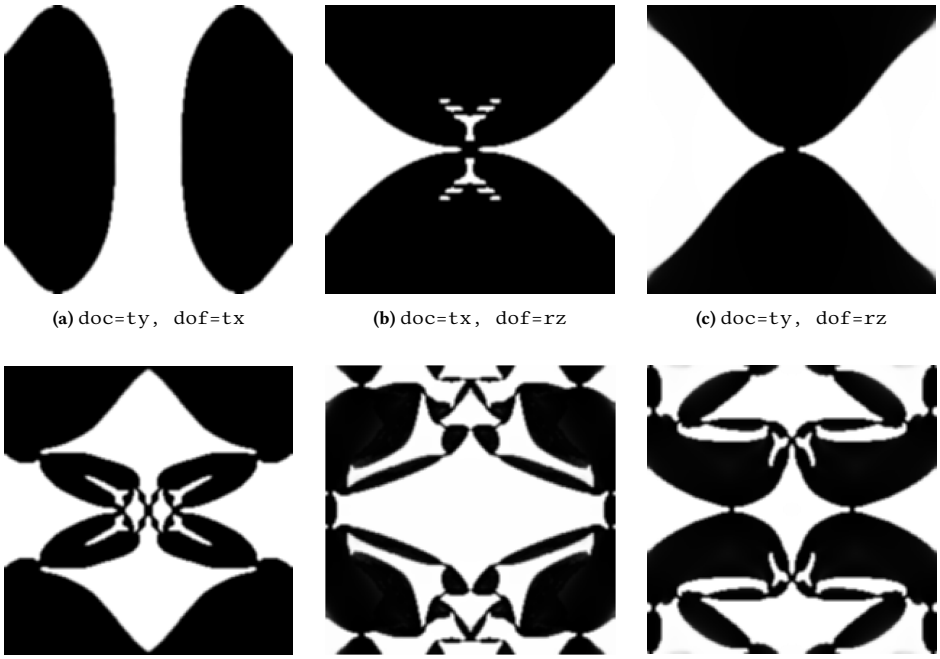
2.5 Numerical examples

In order to demonstrate the method proposed in Section 2.3, we apply it to common problems for which results have been reported in literature. Thereto, we introduce a set $\mathbb{D} = \{\texttt{tx}, \texttt{ty}, \texttt{rz}\}$ consisting of the three rigid body MPs of the rigid links; two relative translations and rotation around the center of the flexure, see Figure 2.2. A sketch of the deformed structure resulting from the prescribed MPs for $\mathbf{x}^{\{0\}}$ are shown in Figure 2.2. Note that, without adjusting the formulation, any other set of unique MPs may be used.

Figure 2.3 show the resulting topologies for a variety of planar design cases. Primitive topologies for the design of a compliant prismatic and revolute joint are shown in Figure 2.3a and Figure 2.3c respectively. The results are as expected and fully in accordance to the results obtained by both synthesis methods (Gallego *et al.* 2009; Howell 2001) and topology optimization formulations (Wang 2009a; Zhang *et al.* 2018).

Complex topologies appear for different less common design cases, such as those shown in Figures 2.3d to 2.3f. These results, to the best knowledge of the authors, have not been reported in literature. Increasing the DOCs and/or DOFs generally results in more complex (number of bodies and rotation points) and innovative topologies, see Figures 2.3e and 2.3f. The convergence history of the responses for a planar optimization problem with representative set of input parameters is shown in Figure 2.4. For feasible input parameters, the convergence history of problem $\mathcal{P}$ is characterized by a quickly active and satisfied constraints and a steadily and smoothly increasing performance, converging within a limited number of iterations.

High-resolution 3D topologies are presented in Figure 2.5. Those topologies are examples of the high variety of designs that can be obtained based on the set of rigid body MPs in a 3D space.



(a) `doc=ty, dof=tx` **(b)** `doc=tx, dof=rz` **(c)** `doc=ty, dof=rz`

(d) `doc=tx, dof=ty` **(e)** `doc=rz, dof=[tx,ty]` **(f)** `doc=[ty,rz], dof=tx`

Figure 2.3: Generated designs for various varieties of problem formulations based on $\mathcal{P}$. This includes both single and multi-axis mechanisms for 2D topologies. These designs are generated by `flexure(200,200,doc,dof,emax)` (Koppen *et al.* 2022d), with `doc` and `dof` given in the subcaptions. A variety of `emax` has been chosen. These results can be replicated using the attached code. Note that we have omitted the string signature (*e.g.* `"tx"`) here for simplicity.

The resulting topologies validate the correct working principle of the proposed formulation. In addition it shows optimized flexures have relatively small features with highly lumped strain energy, see *e.g.* Figure 2.3c. As a result, those flexures have a small range of motion limited by the critical stress and their performance is sensitive to manufacturing errors.

In order to practically use the resulting designs, the maximum allowable stress as well as manufacturing uncertainties should be taken into consideration. In the following we will show the possibilities of limiting stress levels and/or introducing manufacturing robustness in the formulation, without aiming to provide a thorough investigation of design parameters. To this end, we use the resulting design from Figure 2.3d as a reference. We denote its objective by f^0 and corresponding maximum stress by $\overline{\sigma}^0$.

Fault-tolerant design

The desired kinematics of a flexure are sensitive to both uniform and spatially varying geometric deviations. However, in classical deterministic topology optimization, the effect of such uncertain parameters on the performance of the structure is not taken into account. This may lead to a design that is very sensitive to manufacturing errors. As a consequence,

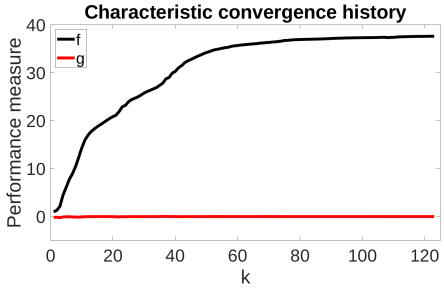


Figure 2.4: Characteristic convergence history of problem formulation $\mathscr{P}$: objective f and constraint g as a function of design iteration k. Note the objective is relative with respect to the first iteration, *i.e.* $f^{\{0\}} = 1$.

the performance of the actual structure may be far from optimal. Sigmund (2009) and Wang *et al.* (2011) propose a robust approach to topology optimization where the effect of uniform manufacturing errors is taken into account. Uniform erosion and dilation effects, from here on denoted by superscripts (e) and (d), are simulated by means of a projection method: the filtering of the design variable field is followed by a differentiable Heaviside projection using a high projection threshold $\eta^{\mathrm{e}} = \eta + \Delta\eta$ to simulate an erosion and a low projection threshold $\eta^{\mathrm{d}} = \eta - \Delta\eta$ to simulate a dilation. An additional advantage of the robust formulation is the direct control of the minimum feature size of both solid and void.

For the robust design of flexures, only a slight difference of Equation (2.1) is required, that is[6]

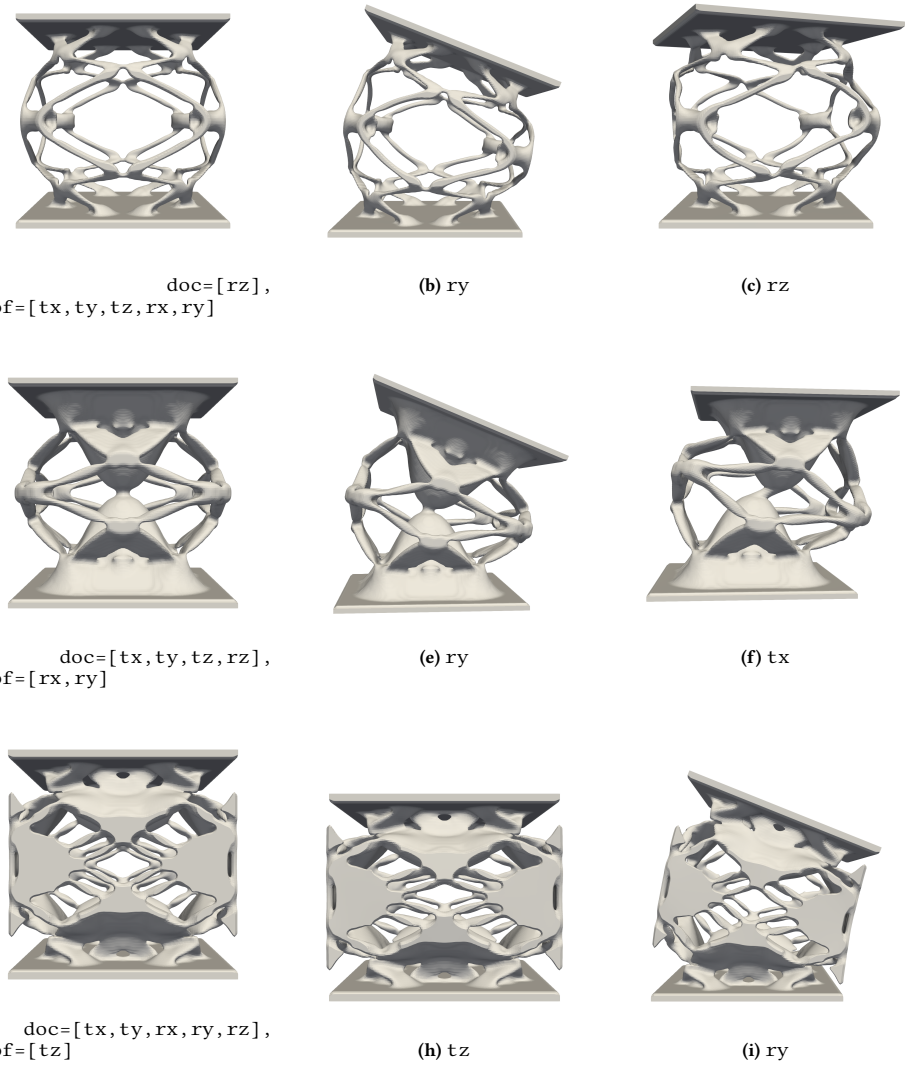
$$\mathscr{P}_\eta = \begin{cases} \underset{\mathbf{x}}{\text{minimize}} & -f\left[\mathcal{E}_i\left[\mathbf{x}^{\mathrm{e}}\right]\right], \quad i \in \mathbb{C} \\ \text{subject to} & \mathcal{E}_i\left[\mathbf{x}^{\mathrm{d}}\right] \leq \overline{\mathcal{E}}_i, \quad i \in \mathbb{F}, \\ & \mathbf{x} \in \mathcal{X} \end{cases} \tag{2.12}$$

with $\mathcal{E}\left[\mathbf{x}^{\mathrm{e}}\right]$ and $\mathcal{E}\left[\mathbf{x}^{\mathrm{d}}\right]$ strain energies based on the eroded and dilated fields, respectively.

Since the eroded and dilated designs will always hold the maximum and minimum strain energies, respectively, the intermediate design can be excluded from the optimization formulation without compromising robustness in terms of length scale control (Lazarov *et al.* 2016). This allows to, partially, reduce the added cost of the robust formulation. Note that all responses still only involve self-adjoint strain energy terms.

Figure 2.6a shows the resulting designs of an optimization problem with filter radius $r = 4$ finite elements, $\eta = 0.5$ and $\Delta\eta = 0.2$. The robustness poses a heavy restriction on the achievable performance, as can be observed by the decrease in performance, *i.e.* $f = 0.32 \times f^0$. The hinges are clearly lengthened, thus distributing the strain energy over larger areas of the topology. In line with this observation, Amir *et al.* (2018) shows it is possible to indirectly achieve stress-constrained topological design via length scale control. Note the non-intuitive presence of protrusions along the center horizontal axis. Upon further investigation, it is observed that those do not add stiffness to the free MPs, whilst

[6] We omit further explanation of this formulation, as arguments and implications are discussed extensively in Wang *et al.* (2011).



(a) `doc=[rz], dof=[tx,ty,tz,rx,ry]` **(b)** `ry` **(c)** `rz`

(d) `doc=[tx,ty,tz,rz], dof=[rx,ry]` **(e)** `ry` **(f)** `tx`

(g) `doc=[tx,ty,rx,ry,rz], dof=[tz]` **(h)** `tz` **(i)** `ry`

Figure 2.5: High-resolution 3D flexure designs generated using a C++ implementation. The number of MPs allows for a high number of variations. From left to right: the topology as a result of solving $\mathcal{P}$ for the given set of constrained and free MPs, and corresponding deformed topology for a free and constrained MP. Note that the deformations are highly scaled for visualization purposes. Corresponding STL files can be found in Koppen *et al.* (2022d).

contributing some (although little) stiffness to the constrained MPs. Considering this lack of sensitivity, those are expected to be removed first upon, for example, introduction of a volume constraint.

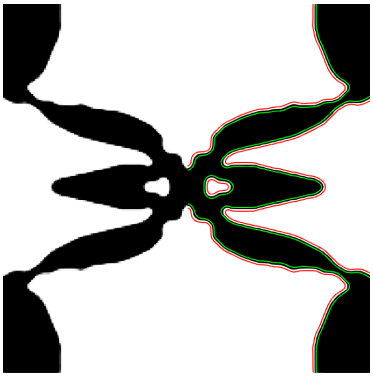
Stress-based design

In order to limit the maximum stress for a given range of motion, or similarly extend the range of motion for a given maximum stress, one can simply extend the problem formulation $\mathcal{P}$ with stress constraints on the free MPs, which yields

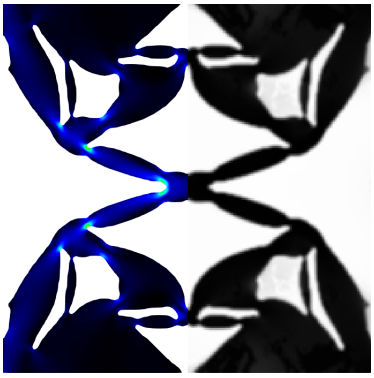
$$\mathcal{P}_\sigma = \begin{cases} \min\limits_{\mathbf{x}} & -f[\mathcal{E}_i], & i \in \mathbb{C} \\ \text{s.t.} & \mathcal{E}_i[\mathbf{x}] \leq \overline{\mathcal{E}}_i, & i \in \mathbb{F} \\ & g_{\sigma_i}[\boldsymbol{\sigma}_i] \leq \overline{\sigma}, & i \in \mathbb{F} \\ & \mathbf{x} \in \mathcal{X} \end{cases}, \tag{2.13}$$

where $\boldsymbol{\sigma}_i$ are the elemental stresses obtained by prescribing free MP i, and $\overline{\sigma}$ the maximum allowable stress, based on some theory of failure. To evaluate stress constraints, elemental strain energies are no longer usable. Many different formulations of g_σ are available Silva *et al.* 2019. We use the unified aggregation and relaxation approach as proposed by Verbart *et al.* (2017).

Figure 2.6b shows the resulting design of an optimization problems with $\overline{\sigma} = 0.4 \times \overline{\sigma}^0$. The stress constraints are satisfied by introduction of (more) hinges with a more distributed deformation energy. Although the maximum stress is drastically reduced, the introduction of stress constraints have a relative limited impact on the performance decreases, namely $f = 0.92 \times f^0$. This demonstrates that the proposed formulation can effectively be extended with stress constraints, yet a thorough investigation thereof is considered out of the scope of this work.



(a) Design obtained by solving $\mathcal{P}_\eta$, see Equation (2.12). The eroded and dilated designs are indicated by, respectively, green and red contour lines. The design is robust with respect to uniform manufacturing errors *and* satisfies a minimum feature size (both solid and void). However, the constrained MP stiffness is decreased with 68% as compared to the design in Figure 2.3d.



(b) Design obtained by solving $\mathcal{P}_\sigma$, see Equation (2.13). Major change in topology can be observed as compared to Figure 2.3d. As a result of the applied stress constraints the range of motion is extended, with just 8% decrease of the constrained MP stiffness as compared to the design in Figure 2.3d.

Figure 2.6: Resulting design of the topology optimization problem from Figure 2.3d extended with (a) the robust formulation, and (b) stress constraints.

2.6 Discussion

Before concluding this article, a reflection on the formulation in light of existing methods, the limitations and related future work and potential applications is in order. Although dissimilar in formulation and implementation, the works of both Hasse *et al.* (2009) and Wang (2009a) share the same kinetoelastic design philosophy, resulting in designs with *selective compliance*. Thereto, this work can be considered both a simplification and generalization of Hasse *et al.* (2009) and Wang (2009a). Also—although the proposed method does not use the separation of scales and periodicity of numerical homogenization (Bensoussan *et al.* 1978; Sanchez-Palencia 1980)—it bears some resemblance to TO of tailored materials with prescribed elastic properties by inverse homogenization (Sigmund 1994, 1995). Similarly to those approaches, in the proposed formulation independent MPs to optimize the structure's intrinsic properties. The introduction of these MPs allows to easily perform design variations such as multi-axis flexures (*e.g.* as demonstrated in Figure 2.3 and Figure 2.5) *and* straightforward adaptation to specific applications.

While clear advantages can be identified, the proposed formulation in not without limitations. As presented here, it is intended for and limited to the design of short-stroke flexures, *i.e.* satisfying linear strain-displacement *and* stress-strain relationships. However, the prototyped samples indicate many of the resulting topologies can be used effectively in a finite range of motion. Duenser *et al.* (2021) analyzed the finite range of motion behavior for a subset of the prototyped flexures using a novel nonlinear eigenmode analysis technique. Results indeed indicate the flexures retain their predicted properties at least for a small finite range. As expected, for a large range of motion, stiffness properties deviate substantially. Considering both the need for and interest in large-range compliant joints, *e.g.* compliant implants, the authors pursue a study to extend the proposed formulation to design for

long-stroke flexures.

It is beyond the scope of this study to include and discuss all possible variations building on this formulation, *e.g.* change of geometry, MPs or objective function. To that, we stress that, in our experience, the formulation does not pose limitations for shape-morphing or compliant mechanism design. As such, design studies and variations of the formulation to applications mentioned above are considered valuable future work, and the provided source code intents to facilitate this.

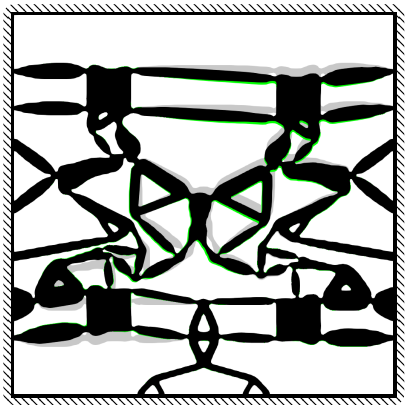
The ease of implementation and use, versatility and modest computational effort of the proposed optimization problem formulation poses a high potential of the formulation to be used as a design tool for practically relevant applications. As such, this work contributes to a more widespread use of topology optimization as a design tool to design flexures in the high-tech industry.

2.7 Conclusion

Despite the extensive efforts devoted to research in the field of TO and the need for a effective tool to synthesize flexures for high-precision applications, a generally accepted TO problem formulation for short-stroke flexure synthesis has been absent so far. Motivated by this, we propose a simple, versatile and computationally efficient topology optimization problem formulation. Using motion patterns this strain energy based formulation simplifies understanding and implementation of TO synthesis of flexures, and features low computation cost, smooth convergence to well-defined designs, and a minimum of tuning parameters. The base formulation is easily extended with additional design requirements and maintains its favorable properties. Although designed for short-stroke applications, the resulting 3D designs prove practically useful within a small finite range of motion. With source code provided to replicate the demonstrated results, this formulation is ready to be further explored and applied in academia and industry.

3

Topology optimization of multi-DOF compliant mechanisms



This chapter proposes a simple yet versatile energy-based topology optimization problem formulation for designing compliant mechanisms, flexures and shape-morphing structures with multiple mechanism degrees of freedom for use in applications such as precision manipulation and adjustment devices.

This chapter is based on a manuscript intended to be published in:
Structural and Multidisciplinary Optimization

3 Synthesis of multi-degree of freedom compliant mechanisms via the kinetoelastic energy-based topology optimization method: principles and applications

Abstract *Short-stroke multi-degree-of-freedom compliant mechanisms provide highly repeatable motions, making them ubiquitous in precise manipulation and adjustment devices. This work proposes a simple yet versatile energy-based topology optimization problem formulation for designing compliant mechanisms, flexures and shape-morphing structures with multiple mechanism degrees of freedom. The proposed optimization problem formulation can be seen as a natural extension of the classic compliance minimization problem and comes with comparable advantageous optimization properties. Numerical examples implemented in academic code and commercial software demonstrate the ease of implementation, use, and applicability to various problems.*

3.1 Introduction

Compliant Mechanisms (CMs) are monolithic and, as opposed to rigid-body mechanisms, attain motion by elastic deformation of the material that constitutes the structure (Midha *et al.* 1994). This flexibility allows CMs to transfer loads and motions,[1] morph a continuous shape or absorb and release energy, see *e.g.* Jovanova *et al.* (2019), Lu *et al.* (2003), and Machekposhti *et al.* (2018). Due to the absence of joints, CMs do not suffer from friction and backlash. As a result, CMs facilitate precise and reliable movement (Howell 2013). However, material failure limits the attainable range of motion (Leon *et al.* 2015). In the present work, we focus on CMs with a range of motion considerably smaller than their dimension, so-called short-stroke CMs. We assume small rotations and strain, which permits the use of linear-elastic theory. The stated advantages make short-stroke CMs ubiquitous in high-precision technology, such as nano-manipulation and manufacturing, MEMS sensors and actuators, adjustable (optical) mounts and alignment devices (Kota *et al.* 2001; Vukobratovich *et al.* 1988; Wu *et al.* 2018).

A vital characteristic of any mechanism is the *mechanism Degrees of Freedom* (DOFs), herein defined as the number of independent variables that define its configuration. From a classical mechanism design point of view, a CM with multiple mechanism DOFs, or

[1] Note we use the terms load and motion as opposed to force and displacement to emphasize the load can be of any form, concentrated such as a force and torque or distributed, *e.g.* pressure. Likewise, a motion can refer to a displacement, rotation or shape-change.

simply multi-DOF CM, allows for the control of multiple DOFs of a single end-effector (Hao *et al.* 2015).[2] In the present work, we generalize the definition of a multi-DOF CM to *any monolithic structure with multiple mechanism DOFs* (and potentially multiple end-effectors). Our definition includes shape-morphing structures, flexures and springs, and other CMs with multiple mechanism DOFs.

In contrast to their serial counterparts, parallel multi-DOF mechanisms generally have advantageous static and dynamic properties, essential for precise, reliable and fast manipulation (Patel *et al.* 2012).[3] However, these advantages come with limited motion ranges that are often inherently coupled (Awtar *et al.* 2012). Albeit not crucial for actuator-controlled mechanisms, precise manual adjustment of coupled mechanisms is a cumbersome iterative process (Ahmad 2017). One of the many challenges characteristic of the design of multi-DOF mechanisms is motion decoupling (Awtar *et al.* 2012). Coupled systems will—upon actuation of an input port—generate parasitic loads, motions or energy (combination of load and motion), including those at unintended output ports (output-coupling) or other input ports (input-coupling).[4][5] The present work focuses on designing decoupled multi-DOF CMs with predefined short-stroke kinematics.

Critical short-stroke CM design requirements include the desired transmission of loads and motion between input and output port(s), the suppression of undesired motion and loads, stiffness ranges at the ports, strength and fatigue limitations, dynamic properties, and robustness of such performance measures considering manufacturing tolerances and uncertainties, to name a few (Howell 2013). The present work focuses on the design for motion transmission, motion decoupling, and (quasi-) static stiffness of the compliant structure.

The load-dependent kinematics of CMs complicates their accurate analytical modelling and purposeful design (Linss *et al.* 2019). Structural optimization methods, and in particular topology optimization (TO), have proven to enable and facilitate the design of high-performance structures without a priori knowledge of the topology while handling diverse stringent design requirements (Bendsøe 1989; Sigmund 2001). Given the typically large number of conflicting design requirements and loading conditions in the design of multi-DOF CM, TO shows excellent potential for the basis of a generalized[6] formulation dedicated to the design of parallel short-stroke multi-DOF CMs.

[2] An end-effector is a device or object at the end of a (kinematic chain of) mechanism(s) designed to interact with the environment.

[3] Serial multi-DOF mechanisms, such as robotic arms, consist of several links in a kinematic chain. Serial mechanisms generally have an extensive range of motion but disadvantageous static and dynamic properties. In contrast, a parallel multi-DOF mechanism, such as a hexapod, connects the end-effector directly to its base by multiple separate and independent links. Parallel mechanisms generally offer fast and precise motion but come with a limited range of motion.

[4] A *port* is a commonly used term loosely defined as a point, line, area or volume that acts as a connection between the mechanism and the environment, such as the location where an actuator or end-effector is attached (Saxena 2005).

[5] The type of parasitic effect (load, motion or energy) depends on the boundary conditions at corresponding ports; infinite (blocked motion), zero (free motion) or finite stiffness (*e.g.* workpiece or actuator stiffness), respectively.

[6] Such as arbitrary design volumes, mechanism DOFs and corresponding kinematics.

3.1.1 Topology optimization of multi-DOF compliant mechanisms

In what follows, we give a concise description of previously proposed TO problem formulations for designing multi-DOF CMs. The reader is referred to Cao *et al.* (2013), Deepak *et al.* (2009), Koppen *et al.* (2021b), and Zhu *et al.* (2020) for in-depth comparison and review of single-DOF CM and flexure TO formulations.

Surprisingly, one of the earliest works on TO of CMs by Larsen *et al.* (1997) introduces an example of a multi-DOF CM. The formulation is based on a least-square-error minimization of prescribed load or motion transmission while maximizing energy transmission. In line with this work, Sigmund (2001) proposes a problem formulation for designing (electro-)thermally driven actuators with up to three mechanism DOFs. The work includes *explicit* constraints on parasitic motions by limiting these motions relative to, for example, the input motion.

Subsequent contributions propose variations on the multi-criteria formulation proposed by Frecker *et al.* (1997), also referred to as the 'unit dummy load' or 'Mutual Potential Energy (MPE)' method (Shield *et al.* 1970b), see ,*e.g.*, Alonso *et al.* (2014). These formulations maximize the displacements at output ports for given inputs loads while simultaneously maximizing the output stiffness. Motion decoupling is (partly) handled by minimization of the MPE related to the parasitic motion (Zhan *et al.* 2010), or by the addition of numerous explicit displacement-based constraints (Rong *et al.* 2021; Zhu *et al.* 2018) in a similar form as proposed by Sigmund (2001). Derived formulations are applied to the design of translational positioning stages (Dinesh *et al.* 2007; Ramesh *et al.* 2014; Zhu *et al.* 2019).

From a kinematic design philosophy, Hasse *et al.* (2009) and Wang (2009a) independently propose formulations using a condensed 'mechanism stiffness model', which later formed the basis for multi-DOF formulations (Jin *et al.* 2018; Kirmse *et al.* 2021b). The condensed stiffness matrices are independent of externally applied loads and are considered intrinsic properties of the CM. See Appendix 3.B for an elaboration on these properties. Based on such a condensed (but exact) model, Jin *et al.* (2018) propose a multi-objective formulation to simultaneously minimize a least-square-error of prescribed motion transmissions (including output coupling terms) *and* maximize the stiffness at the input and output ports. Originating from shape-morphing design philosophy and based upon the work of Hasse *et al.* (2009), Kirmse *et al.* (2021b) propose a method to design multi-DOF CMs based on an eigensystem analysis of the condensed mechanism stiffness model that considers mode orthogonality.

3.1.2 Challenges

Current TO problem formulations for multi-DOF CM design face challenges in terms of *simplicity* and *versatility*. The formulations tend to be complex to *understand*, *implement* or *use* and/or are *applicable* to a limited range of problems.

A recurring problem is that many formulations show undesirable convergence properties (oscillatory behaviour and, not uncommonly, over 1000 design iterations) or deliver non-binary (and hence non-manufacturable) topologies (Kirmse *et al.* 2021b; Larsen *et al.* 1997; Rong *et al.* 2021; Sigmund 2001; Zhu *et al.* 2018). These issues are typically a consequence of (i) the presence or absence of conflicting requirements, (ii) the introduction of highly restrictive constraints, or (iii) the use of responses that depend highly nonlinearly on the displacements. As the desired mechanism DOFs increases, the number of explicit

constraints on parasitic motions rises, negatively impacting optimization convergence. Many formulations use complicated optimization schemes or extensive use of filters, projection schemes and continuation approaches (accompanied with many parameters) to obtain a 'meaningful' solution (Alonso *et al.* 2014; Kirmse *et al.* 2021b; Rong *et al.* 2021; Zhu *et al.* 2018). It appears to be challenging to make the feasible space attainable to the optimizer and obtain a 'meaningful' solution simply and effectively.

The versatility of an optimization problem formulation is closely related to the ability to optimize for the first and foremost requirement in mechanism design—its function—expressed in terms of the intended kinematics. "Although the kinematics generally forms the basis for the design of conventional mechanisms, most TO problem formulations do not explicitly include the kinematics; it is more or less an outcome of the optimization process. (Hasse *et al.* (2017))" Most formulations (either explicitly or implicitly) maximize the energy transmission between input and output port(s) or—equivalently—minimize the energy stored (Alonso *et al.* 2014; Larsen *et al.* 1997; Rong *et al.* 2021; Sigmund 2001; Zhan *et al.* 2010; Zhu *et al.* 2018). These formulations are thus limited to solving problems that require maximum energy transmission. Such formulations produce similar topologies for the optimized CMs (Deepak *et al.* 2009), which tend to emulate their rigid-body counterpart (Wang 2009a).[7] These optimized solutions consist of stiff bodies connected by single-noded hinges, generally considered undesirable due to their sensitivity to manufacturing errors and stress localization. Various formulation adaptations, constraints and filters have been proposed to eliminate single-noded hinges. Although these remedies successfully eliminate single-noded hinges, the resulting topologies still exhibit hinges with highly localized deformation as these are favoured to minimize energy storage. What is more, the correct working principle of a substantial number of formulations critically relies upon the use of 'artificial' stiffness attached to input and output ports (Alonso *et al.* 2014; Rong *et al.* 2021; Sigmund 2001; Zhan *et al.* 2010; Zhu *et al.* 2018). Albeit relevant for specific applications, *e.g.* when the end-effector is attached to a deformable object, this limits the applicability and, in the general case, makes these parameters somewhat arbitrary and finding appropriate values requires tuning. In addition, the use of such 'artificial' stiffness is undefined in structures without a clear definition of input and output ports, such as for adaptive structures, springs or flexures.

3.1.3 Aim, contributions and outline

The present work aims to provide a *simple* and *versatile* method for synthesizing short-stroke parallel decoupled multi-DOF CMs. To do so, we extend the work presented in Koppen *et al.* (2021b) on the TO of flexures to the design of CMs with multiple mechanism DOFs. We propose a novel kinetoelastic energy-based TO problem formulation, much in line with the classic compliance minimization formulation (Bendsøe 1989). The optimization objective is to maximize the stiffness of undesired kinematics whilst constraining the maximum stiffness of *a priori* defined desired kinematics, as will be motivated and explained in depth in Section 6.2 (Method). The main contribution of the present work is to address the posed challenges in Section 3.1.2 by the proposition of this novel TO problem formulation for the design of multi-DOF CMs that is (i) easy to understand and use, (ii) straightforward to

[7] Actually, any optimal solution would store zero energy and, as such, consist of rigid-body motions. Problems arise upon analysis of such a singular structure.

implement, (iii) widely applicable, and (iv) holds advantageous optimization properties. In line with the design philosophy, we propose the concepts of generalized degrees of freedom, and motion and load patterns, along with corresponding magnitudes and characteristic stiffness measures. Furthermore, we provide a new symbolic language to describe problem settings, use cases, load cases, and boundary conditions specifically focused on TO of multi-DOF CM problems. The formulation is widely applicable since many design problems can be (re)formulated in the proposed format, including but not limited to the design of CMs, flexures, springs, and adaptive structures, as will be demonstrated in Section 5.5 (Numerical Examples). What is more, we introduce new single-DOF and multi-DOF benchmark TO CM problems including corresponding codes. We encourage future work on this topic to use the same or variations of the proposed problem settings. We provide our insights regarding the key advantages, limitations and future work in Section 5.6 (Discussion), followed by conclusions on the method and its application potential in Section 5.7 (Conclusion).

3.2 Method

This section is organized as follows: First, we provide a recap of concepts that lie at the core of the formulation. We discuss the relations between potential energy, strain energy, work and compliance and emphasize critical properties of these energy terms under so-called load-based and motion-based scenarios. Next, we introduce the concepts of generalized DOFs of interest, motion patterns and load patterns, that are critical concepts to measure the characteristic stiffness of the (un)desired kinematics. Based on these concepts we propose an alternative view on the synthesis of CMs using characteristic stiffnesses in Section 3.2.2. Finally, we propose a novel optimization problem formulation for the design of multi-DOF CMs in Section 3.2.3, and discuss the working principle, optimization problem properties, forward and sensitivity analyses, computational effort and implementation considerations. We finish this section with a description of the advantages and use of coordinate transformations in combination with the proposed optimization problem formulation.

3.2.1 Energy principles

Consider a continuum solid made of a *linear elastic* material subjected to *conservative*[8] loading conditions independent of the motion. By conservation of energy, the *external work* $\mathcal{W}_{\text{ext}} \in \mathbb{R}$, that is the work performed by the external loads on the elastic solid increasing as they pass through the motions, is stored completely as strain energy $\mathcal{E} \in \mathbb{R}$ within the solid.

We consider the *total potential energy* $\Pi \in \mathbb{R}$ of this solid as the summation of elastic potential or strain energy and the load potential with respect to the *unstrained* position; that is the stress-free and unloaded condition.

The efficient analysis of an arbitrary complex continuum requires approximated solutions via discretization by the application of, for example, the principle of weighted residuals of Galerkin, such as common in Finite Element Analysis (FEA) (Cook *et al.* 2007). Such methods approximate the continuous solution (*e.g.* displacement field) using a finite number of nodal DOFs, hereafter referred to as *structural* DOFs. The total potential energy

[8] For a conservative loading the total work done is independent of the path taken. The load changes the potential energy by an amount independent of the path taken, satisfying conservation of energy.

of a solid consisting of linear elastic material in such a discretized setting reads as

$$\Pi = \frac{1}{2}\mathbf{u}\cdot\mathbf{K}\mathbf{u} - \mathbf{f}\cdot\mathbf{u}, \tag{3.1}$$

with, $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{f} \in \mathbb{R}^n$ a kinematically admissible motion and external loads at the $n \in \mathbb{Z}^+$ structural DOFs and $\mathbf{K} \in \mathbb{R}^{n\times n}$ the symmetric and positive-definite stiffness matrix. That is we consider a properly constrained system in which rigid-body motions are suppressed. The *principle of stationary potential energy* states that the total potential energy Π is stationary in a state of static equilibrium, that is,

$$\frac{\mathrm{d}\Pi}{\mathrm{d}\mathbf{u}} = \mathbf{0} \quad \longrightarrow \quad \mathbf{K}\mathbf{u} = \mathbf{f}, \tag{3.2}$$

assuming (possibly non-zero) prescribed motions are satisfied. The total potential energy *at equilibrium* thus equates to

$$\Pi = -\frac{1}{2}\mathbf{u}\cdot\mathbf{K}\mathbf{u} = -\frac{1}{2}\mathbf{f}\cdot\mathbf{u}. \tag{3.3}$$

Thus, for a linear system the minimum value of Π is the negative of strain energy or *complementary strain energy* $\mathcal{E}^*$. By conservation of energy this furthermore equates to the negative of external work. Twice this value is typically referred to as the (static elastic) *compliance*[9] $\mathcal{C} := \mathbf{f}\cdot\mathbf{u}$ within the structural optimization community; one of the most commonly used performance measures for stiffness design (Bendsøe 1989; Prager 1968). Thus, given the assumptions of conservatism, elasticity and linearity, the *value* of minimum total potential energy at equilibrium may equivalently be given by

$$\Pi = \mathcal{E}^* = -\mathcal{E} = -\mathcal{W}_{\text{ext}} = -\tfrac{1}{2}\mathcal{C}. \tag{3.4}$$

Consequently, one may trivially switch between energy terms as fits the design problem at hand by taking into account the constant conversation factor. In what follows we deliberately switch between the use of strain energy and compliance as fits best to elucidate the topic of interest.

Load-based and motion-based scenarios

The design of multi-DOF CMs typically involves multiple *scenarios*, often referred to as loading conditions or load cases. We define a *scenario* as a quantitative description of specific loading *and* boundary conditions. This thus involves the partitioning of degrees of freedom into 'free' and 'prescribed' *and* the value of the prescribed motions and applied loads. Each scenario involves a unique partitioning of the discretized system of governing equations from Equation (3.2) into free (subscript 'f') and prescribed (subscript 'p') structural DOFs, that is

$$\begin{bmatrix}\mathbf{K}_{\text{ff}} & \mathbf{K}_{\text{fp}}\\ \mathbf{K}_{\text{pf}} & \mathbf{K}_{\text{pp}}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{\text{f}}\\ \mathbf{u}_{\text{p}}\end{bmatrix} = \begin{bmatrix}\mathbf{f}_{\text{f}}\\ \mathbf{f}_{\text{p}}\end{bmatrix}, \tag{3.5}$$

[9] Note the difference with respect to the—at least in the field of mechanism design—more commonly used definition of compliance as the inverse of stiffness.

or, equivalently

$$\mathbf{K}_{\mathrm{ff}}\mathbf{u}_{\mathrm{f}} = \mathbf{f}_{\mathrm{f}} - \mathbf{K}_{\mathrm{fp}}\mathbf{u}_{\mathrm{p}}, \quad \text{and} \tag{3.6}$$

$$\mathbf{f}_{\mathrm{p}} = \mathbf{K}_{\mathrm{pf}}\mathbf{u}_{\mathrm{f}} + \mathbf{K}_{\mathrm{pp}}\mathbf{u}_{\mathrm{p}}, \tag{3.7}$$

with the applied loads $\mathbf{f}_{\mathrm{f}} \in \mathbb{R}^{f}$ and prescribed motions $\mathbf{u}_{\mathrm{p}} \in \mathbb{R}^{p}$.

We distinguish between two types of scenarios, namely (i) load-based scenarios in which the applied loads $\mathbf{f}_{\mathrm{f}} \neq \mathbf{0}$ and prescribed motion $\mathbf{u}_{\mathrm{p}} = \mathbf{0}$, and (ii) motion-based scenarios in which the applied loads $\mathbf{f}_{\mathrm{f}} = \mathbf{0}$ and prescribed motion $\mathbf{u}_{\mathrm{p}} \neq \mathbf{0}$. Theoretically, mixtures of these are also possible, however for the purpose of this investigation we limit ourselves to these two types. To study the influence of the type of scenario on the compliance (or similarly on the strain energy or external work), consider first the compliance under a load-based scenario, that is

$$\mathcal{C}_{\mathrm{f}} := \mathbf{f}_{\mathrm{f}} \cdot \mathbf{u}_{\mathrm{f}}, \tag{3.8}$$

with $\mathcal{C}_{\mathrm{f}} \in \mathbb{R}^{+}$ the work performed by the constant applied loads in combination with the deformation response of the structure. In contrast, under a motion-based scenario

$$\mathcal{C}_{\mathrm{p}} := \mathbf{f}_{\mathrm{p}} \cdot \mathbf{u}_{\mathrm{p}}, \tag{3.9}$$

with $\mathcal{C}_{\mathrm{p}} \in \mathbb{R}^{+}$ the work performed by the prescribed motions, in combination with the reaction loads of the structure.

It is well known that to achieve a stiff structure under *applied loads* $\mathbf{f}_{\mathrm{f}}$, we should find a design that minimizes the corresponding displacements $\mathbf{u}_{\mathrm{f}}$, *i.e. minimization* of the compliance $\mathcal{C} = \mathcal{C}_{\mathrm{f}}$ (Bendsøe 1989). In contrast, to achieve a stiff structure under *prescribed motions* $\mathbf{u}_{\mathrm{p}}$, we should find a design that maximizes the corresponding reaction loads $\mathbf{f}_{\mathrm{p}}$, *i.e. maximization* of the compliance $\mathcal{C} = \mathcal{C}_{\mathrm{p}}$ (Klarbring *et al.* 2012). This conclusion also holds for the total potential energy, strain energy or external work under the given assumptions of conservatism and linearity.

This observation of the influence of the type of scenario is crucial for functioning of the optimization problem formulation that will be presented in Section 3.2.3. An one-dimensional example further elucidating this observation is provided in Appendix 3.A, based on the Appendix of Pedersen *et al.* (2011).

Generalized DOFs of interest

Despite the continuously deformable nature of CMs, there typically exists a limited number of $m \in \mathbb{Z}^{+}$ DOFs whose behaviour is *crucial* for the intended function of the mechanism. The kinematics of a CM is *predominantly* defined by those DOFs and will hereafter be referred to as the *Generalized DOFs of Interest* (GDIs). These GDIs are typically the DOFs representing the input and output port(s), as ,*e.g.*, in Figure 3.1. The prefix 'generalized' indicates the GDIs may also include a rotation, relative motion between two subregions of the continuum, actuator displacement or shape change, to name a few.

Although not strictly necessary, but useful for elucidating purposes, we next describe the behaviour of the CM in terms of the GDIs only. For that purpose, we rewrite Equation (3.5) using oversets ($\wedge$) and ($\vee$) to represent the properties related to GDIs and (generalized)

DOFs not of interest, respectively, that is

$$\begin{bmatrix} \overset{\wedge}{\mathbf{K}} & \overset{\wedge}{\mathbf{K}}_{\mathrm{f}} & \overset{\wedge}{\mathbf{K}}_{\mathrm{p}} \\ \overset{\vee}{\mathbf{K}}_{\mathrm{f}} & \overset{\vee}{\mathbf{K}}_{\mathrm{ff}} & \overset{\vee}{\mathbf{K}}_{\mathrm{fp}} \\ \overset{\vee}{\mathbf{K}}_{\mathrm{p}} & \overset{\vee}{\mathbf{K}}_{\mathrm{pf}} & \overset{\vee}{\mathbf{K}}_{\mathrm{pp}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \overset{\vee}{\mathbf{u}}_{\mathrm{f}} \\ \overset{\vee}{\mathbf{u}}_{\mathrm{p}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}} \\ \overset{\vee}{\mathbf{f}}_{\mathrm{f}} \\ \overset{\vee}{\mathbf{f}}_{\mathrm{p}} \end{bmatrix}, \tag{3.10}$$

with $\hat{\mathbf{u}} \in \mathbb{R}^m$ and $\hat{\mathbf{f}} \in \mathbb{R}^m$ the motion and loads at the GDIs, respectively. Upon condensation of the DOFs that are not of interest via static condensation (Hasse *et al.* 2009; Jin *et al.* 2018; Koppen *et al.* 2022c; Wang 2009a) the motion at the GDIs is equivalently governed by the system of equations

$$\tilde{\mathbf{K}}\hat{\mathbf{u}} = \hat{\mathbf{f}}, \tag{3.11}$$

with $\tilde{\mathbf{K}} \in \mathbb{R}^{m \times m}$ the reduced-order stiffness matrix exactly describing the stiffness of the GDIs. A more detailed description of the condensation process is provided in Section 3.2.4, and the works of Guyan (1965), Irons (1965), and Koppen *et al.* (2022c).

For the remainder of this work we assume that all applied loads *and* prescribed motions related to the DOFs that are not of interest have zero magnitude, that is $\overset{\vee}{\mathbf{f}}_{\mathrm{f}} = \mathbf{0}$ and $\overset{\vee}{\mathbf{u}}_{\mathrm{p}} = \mathbf{0}$, such that the compliance

$$\overset{\vee}{\mathcal{C}} := \overset{\vee}{\mathbf{f}} \cdot \overset{\vee}{\mathbf{u}} = 0. \tag{3.12}$$

and, consequently $\mathcal{C} = \hat{\mathcal{C}} = \hat{\mathbf{f}} \cdot \hat{\mathbf{u}}$. This is a reasonable assumption since any non-zero contribution (load or motion) would likely make the DOF interesting, and, as such be contained in the set of GDIs. What follows next is a description of the behaviour of an arbitrary CM under motion *or* load-based scenarios, and definitions of characteristic stiffness based on the energy involved in the system upon deformation.

Motion patterns, energy and characteristic stiffness

Consider a motion-based scenario; the motion at all GDIs is prescribed. We separate the magnitude and pattern of the prescribed motion at the GDIs, by rewriting

$$\hat{\mathbf{u}} := \gamma \mathbf{v}, \tag{3.13}$$

with $\gamma \in \mathbb{R}$ the magnitude of non-dimensional unit length Motion Pattern (MP) $\mathbf{v} \in \mathbb{R}^m$ that expresses a pattern of motion through a combination of relative values at the GDIs. The dimension of the MP magnitude depends on the type of motion expressed by the MP, typically length or angle to describe displacements and rotations.

In contrast to rigid-body mechanisms, CMs store energy upon actuation. The strain energy stored (or external work done), to prescribe an MP $\mathbf{v}$ with an amplitude of γ is given by

$$\mathcal{E}_{\mathrm{p}} = \frac{1}{2}\hat{\mathbf{f}} \cdot \hat{\mathbf{u}} = \frac{1}{2}\hat{\mathbf{u}} \cdot \tilde{\mathbf{K}}\hat{\mathbf{u}} = \frac{1}{2}\gamma^2 \mathbf{v} \cdot \tilde{\mathbf{K}}\mathbf{v}, \tag{3.14}$$

with $\hat{\mathbf{f}} \in \mathbb{R}^m$ now the reaction loads at the GDIs upon the prescribed motion. Similar to Equation (3.9), the subscript 'p' indicates a setting in which the motion is prescribed. This energy term is a measure of resistance of the compliant system to the enforced MP.

The stiffness of simple linear spring in a motion-based scenario, as a function of the strain energy, reads

$$\mathcal{E}_\mathrm{p} = \frac{1}{2} k \gamma^2 \quad \longleftrightarrow \quad k = \frac{2\mathcal{E}_\mathrm{p}}{\gamma^2}, \tag{3.15}$$

with γ the magnitude of the extension of the spring. We generalize this to a multi-dimensional setting by introduction of the *motion-based characteristic stiffness* of MP $\mathbf{v}$ as

$$\mathcal{K}_\mathrm{p} := \frac{2\mathcal{E}_\mathrm{p}}{\gamma^2} = \frac{C_\mathrm{p}}{\gamma^2} = \mathbf{v} \cdot \tilde{\mathbf{K}} \mathbf{v}. \tag{3.16}$$

3

This characteristic stiffness is an magnitude independent intrinsic property of the compliant system. Note that the characteristic stiffness $\mathcal{K}_\mathrm{p}$ has the same unit as the entries of the (reduced-order) stiffness matrix, assuming the entries of this matrix have the same units. The motion-based characteristic stiffnesses can be used to differentiate between 'stiff' and 'compliant' MPs.

Load patterns, energy and characteristic stiffness

Next consider a load-based scenario, *i.e.* a load is applied to all GDIs, and the motion is thus *not* prescribed. Let us now similarly redefine the load by separation of magnitude and pattern, such that

$$\hat{\mathbf{f}} := \mu \mathbf{g}, \tag{3.17}$$

with μ the magnitude (typically of dimension of force or moment) of non-dimensional unit length Load Pattern (LP) $\mathbf{g} \in \mathbb{R}^m$. The strain energy in a load-based scenario now reads

$$\mathcal{E}_\mathrm{f} = \frac{1}{2} \mu \mathbf{g} \cdot \hat{\mathbf{u}} = \frac{1}{2} \mu^2 \mathbf{g} \cdot \tilde{\mathbf{K}}^{-1} \mathbf{g}. \tag{3.18}$$

The stiffness of a linear spring under a load-based scenario is given by

$$\mathcal{E}_\mathrm{f} = \frac{1}{2} k u^2 = \frac{1}{2} \frac{\mu^2}{k} \quad \longleftrightarrow \quad k = \frac{\mu^2}{2\mathcal{E}_\mathrm{f}} \tag{3.19}$$

with μ the magnitude of the applied load to (and in) the spring. In line with this definition of stiffness, we generalize to a multi-dimensional setting by definition of the *load-based characteristic stiffness* of LP $\mathbf{g}$ as

$$\mathcal{K}_\mathrm{f} := \frac{\mu^2}{2\mathcal{E}_\mathrm{f}} = \frac{1}{\mathbf{g} \cdot \mathbf{K}^{-1} \mathbf{g}}, \tag{3.20}$$

and its inverse, the *load-based characteristic compliance* as

$$\mathcal{K}_\mathrm{f}^{-1} := \frac{1}{\mathcal{K}_\mathrm{f}} = \mathbf{g} \cdot \tilde{\mathbf{C}} \mathbf{g}, \tag{3.21}$$

with $\tilde{\mathbf{C}} := \tilde{\mathbf{K}}^{-1} \in \mathbb{R}^{m \times m}$ the reduced-order but exact compliance matrix. Similar to $\mathcal{K}_\mathrm{p}$, the characteristic stiffness $\mathcal{K}_\mathrm{f}$ has the same unit as the entries of the reduced-order stiffness matrix. These measures of stiffness are fully independent, since those are defined under

different conditions *and* we do not consider mixtures of load-based and motion-based scenarios.

One special pattern is a basis pattern; an MP or LP with a single non-zero component. Such pattern can be used to measure the characteristic stiffness of *individual* GDIs. In general, this stiffness is defined as the applied load at a GDI required to attain a unit motion at the same GDI. However, such definition is incomplete if not considering the specific boundary conditions. We define the *GDI stiffness* as the load-based characteristic stiffness of a basis LP at the GDI. This characteristic stiffness thus corresponds to the inverse of the motion at the GDI per unit load applied to the GDI, considering all other GDIs are *unconstrained*.

3.2.2 Synthesis of compliant mechanisms using characteristic stiffnesses

Compliant mechanisms *predominantly* deform according to a deformation field given by a linear combination of deformation fields corresponding to MPs with low characteristic stiffness, since this requires the least amount of energy (Hasse *et al.* 2009). This is even more so if the loading location and direction (partly) coincides with those MPs, as is typically the case in the controlled environment typical to high-precision mechanism usage. Therefore, we have the vision that the aforementioned concepts of characteristic stiffnesses can be utilized to synthesize a geometry (topology, shape, size) to be compliant for specific MPs associated with the mechanism DOFs, hereafter denoted the *free MPs*, while stiff in other MPs, hereafter denoted the *constrained MPs*.[10]

The free MPs and corresponding desired characteristic stiffness (or strain energy for a known MP and magnitude) are known desired user parameters; the kinematics and corresponding stiffness define the functionality of a CM. In contrast, the constrained MPs are typically unknown, apart from the desirable property to be orthogonal to the free MPs (otherwise partly describing the free MPs) and have an as high as possible (compared to the free MPs) characteristic stiffness (Hasse *et al.* 2009).

Since the constrained MPs are unknown, we alternatively propose to maximize the characteristic stiffness of *all* MPs (including the free MPs), while constraining the characteristic stiffness (or energy for known MP and magnitude) of the free MPs from above. The resulting designs will consequently have a large difference in characteristic stiffness between the free MPs and the constrained MPs, whilst the user is able to control both kinematics and corresponding stiffness. As such, we envision, one obtains the stiffest structure that only allows for the motion in the direction of the free MP(s). As such, we will utilize the proposed definitions of energy (derived) measures of stiffness from Equations (3.16) and (3.20) to form the basis for a novel compliant mechanism optimization problem formulation.

3.2.3 Optimization problem formulation

In line with the idea put forward in Section 3.2.2, we propose a topology optimization problem formulation that aims to find an optimal solution of bounded design variables

[10] The concept of free and constrained MPs with low, respectively high, characteristic stiffness was previously introduced for the synthesis of flexures (Koppen *et al.* 2021b), based on the concept of 'deformation modes' by Hasse *et al.* (2009).

Table 3.1: Definition of symbols as used in the description of problem setting and scenarios.

Symbol	Description
?	Design domain
	Non-design domain (solid)
	Prescribed zero motion (ground)
	Prescribed non-zero motion
	Prescribed zero motion
	Desired motion
	Undesired motion
	Applied non-zero load
	Desired reaction load
	Undesired reaction load

that, for a given amount of material usage, maximizes the equivalent serial load-based characteristic stiffness of all GDIs, while satisfying a maximum motion-based characteristic stiffness of the free MPs corresponding to the mechanism DOFs. Maximization of the equivalent stiffness of characteristic stiffnesses in series equates to minimization of the summation of corresponding characteristic compliances. The general gradient-based constrained non-linear optimization problem formulation in negative null form reads as

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] : \sum_i \alpha_i \mathcal{K}_{\mathrm{f},i}^{-1}[\mathbf{x}] \\ \text{s.t.} \quad & g_{\mathcal{K},j}[\mathbf{x}] : \alpha_j \left(\frac{\mathcal{K}_{\mathrm{p},j}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p},j}} - 1 \right) \leq 0 \\ & g_v[\mathbf{x}] : \alpha_v \left(\frac{v[\mathbf{x}]}{\overline{v}} - 1 \right) \leq 0 \end{aligned} \tag{3.22}$$

with $i \in \mathcal{S}_\mathrm{f}$ and $j \in \mathcal{S}_\mathrm{p}$. The design variables $\mathbf{x} \in \mathbb{X}^N \subseteq \mathbb{R}^N$ are bounded from below and above, objective $f[\mathbf{x}] \in \mathbb{R}^+$ is a summation of the *load-based characteristic compliances* of the load-based scenarios collected in set $\mathcal{S}_\mathrm{f}$, the constraint $g_{\mathcal{K},j}[\mathbf{x}] \in \mathbb{R}$ limits the *motion-based characteristic stiffness* $\mathcal{K}_{\mathrm{p},j}[\mathbf{x}]$ of the jth motion-based scenario in set $\mathcal{S}_\mathrm{p}$ to the upper bound $\overline{\mathcal{K}}_{\mathrm{p},j} \in \mathbb{R}^+$, and constraint $g_v[\mathbf{x}] \in \mathbb{R}$ limits the material usage $v[\mathbf{x}]$ to the maximum volume fraction $\overline{v}$ in the closed unit interval. The set of load-based scenarios $\mathcal{S}_\mathrm{f}$ consist of m basis LPs at the GDIs to measure the individual GDI stiffnesses as defined in Section 3.2.1. The set of motion-based scenarios $\mathcal{S}_\mathrm{p}$ consists of d free MPs corresponding to the mechanism DOFs. For an example of load-based and motion-based scenarios involved in the problem formulation of a single-DOF and multi-DOF CM see Figures 3.1 and 3.2 and corresponding definitions of the symbols used in Table 3.1. The objective function and constraints should be properly scaled,[11] using constant scaling coefficients α_i, α_j and α_v, respectively. The

[11] Proper scaling depends on the optimizer of choice. However, as a rule of thumb, the response values should be in the same order of magnitude in the first design iteration avoiding very small or large values (Svanberg 1987).

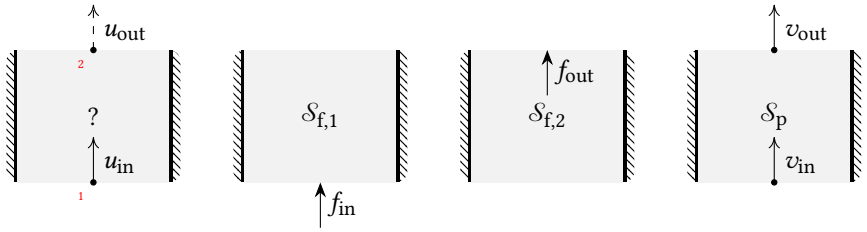


Figure 3.1: Problem setting and motion-based use case scenario (left), and corresponding load-based scenarios (middle) and motion-based scenario (right). The latter are used to set up the proposed problem formulation for designing a compliant mechanism with a single mechanism degree of freedom. The numbering of the generalized degrees of freedom is indicated in red (left). The aim for this use case is to design a (yet unknown) structure that achieves a desired target output motion u_{out} for a prescribed input motion u_{in}. The objective requires the calculation of the load-based characteristic stiffnesses of the two load-based scenarios in $\mathcal{S}_f$. The motion-based characteristic stiffness is measured based on the motion-based scenario $\mathcal{S}_p$ that is defined by the free MP given by non-dimensional unit vectors $\mathbf{v} = \begin{bmatrix} v_{in} & v_{out} \end{bmatrix}$. Symbols as per defined in Table 3.1.

individual entries of the objective function can be scaled with respect to each other to circumvent or generate bias.

Working principle and formulation properties

Maximization of the equivalent stiffness of characteristic stiffnesses in series necessitates a proper connection between *all* the GDIs and ground, effectively circumventing non-properly connected or floating mechanisms; a finding in line with the conclusion of Jin *et al.* (2018). The *global optimum* of the unconstrained optimization problem is a fully solid design, since that is the structure with the highest stiffness at the GDIs (think of unconstrained compliance minimization). The working principle of the proposed constrained optimization problem may be perceived as the selective removal of material to exactly satisfy the characteristic stiffness of the free MP(s), while retaining maximum characteristic stiffness between the GDIs and the ground. The volume constraint is further responsible for removal of material at the locations where the material adds least to the equivalent serial GDI stiffness. By enforcement of a low characteristic stiffness of the free MP(s), the optimization process is limited to maximizing all but the free MPs; that is the constrained MPs. This effectively maximizes the ratio of characteristic stiffnesses between free and constrained MPs. It is this high ratio that is responsible for the desired kinematics, which is only introduced when *both* (i) increasing the characteristic stiffness of the constrained MPs, and (ii) enforcing a low characteristic stiffness of the free MPs. Thus the desired effect is solely achieved by this unique combination of objective and constraint(s).

The kinematics upon actuation of an optimized design, that is the motion-based use case scenario, will be precisely equal to the free MP(s) when the characteristic stiffness ratio between constrained and free MPs approaches infinity. For any practical CM this is not achievable. However, for sufficiently large ratios, the proposed problem formulation produces structures whose motion approaches the imposed free MPs, as will be demonstrated in Section 5.5.

The formulation requires only a *single* constraint per mechanism DOF, irrespective of the complexity of the desired kinematics (see Section 3.3.4). Using a more conventional CM optimization problem formulation (Koppen *et al.* 2022a; Rong *et al.* 2021) typically requires

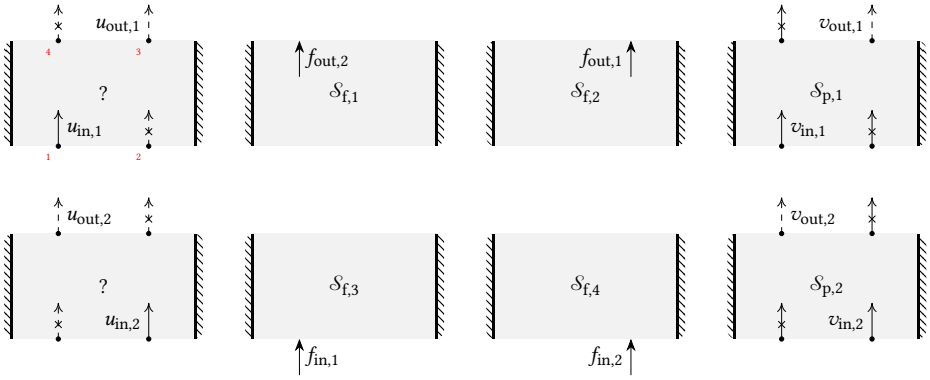


Figure 3.2: Problem setting and use case scenarios (left), and corresponding load-based scenarios (middle) and motion-based scenario (right). These scenarios are used to set up the proposed problem formulation to design a compliant mechanism with two decoupled mechanism degrees of freedom. The aim is to design a (yet unknown) structure that achieves a desired target output motion $u_{\text{out},1}$ for a prescribed input motion $u_{\text{in},1}$ *and* a desired target output motion $u_{\text{out},2}$ for a prescribed input motion $u_{\text{in},2}$ without input and output coupling. The objective requires the calculation of the load-based characteristic stiffnesses of the four load-based scenarios in $\mathcal{S}_{\text{f}}$. The motion-based characteristic stiffnesses is measured based on the motion-based scenarios $\mathcal{S}_{\text{p},1}$ and $\mathcal{S}_{\text{p},2}$ respectively. Those scenarios are defined by the free MPs given by non-dimensional unit vectors $\mathbf{v}_1 = \begin{bmatrix} v_{\text{in},1} & 0 & v_{\text{out},1} & 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 & v_{\text{in},2} & 0 & v_{\text{out},2} \end{bmatrix}$ following the numbering of the generalized degrees of freedom as indicated in red (left).

tens of constraints for the design of a two-DOF mechanism, as all undesired motions are individually and explicitly controlled. Next to increasing optimization complexity and cost, such explicit control leaves less room for design changes compared to the implicit approach of the proposed formulation.

The independent parameters in the proposed problem formulation are the free MPs and corresponding maximum characteristic stiffnesses (the constraint values). While the free MPs originates from the desired kinematics, corresponding characteristic stiffness is typically retrieved from (a combination of) common mechanism design requirements such as desired actuator force, range-of-motion, external work, load transmission or energy efficiency. Such requirements can be relatively trivially rewritten in terms of characteristic stiffness, as per Section 3.2.1. Alternatively, the proposed optimization problem formulation can be fully written in terms of strain energy, external work, potential energy and compliance (under the given assumptions of conservatism and linearity).

The proposed optimization problem formulation holds a remarkable resemblance with the classic compliance minimization problem; it may be considered a generalization/extension thereof. The proposed problem formulation in fact equates to a compliance constrained compliance minimization problem from an implementation point of view. This resemblance not only makes the formulation comprehensible, but also readily usable in commercial software.

Forward analysis

Both the motion and load-based characteristic stiffness are a function of the motions and loads at the GDIs via

$$f[\mathbf{x}] = \sum_i \frac{1}{\mu_i^2} \hat{\mathbf{f}}_i \cdot \hat{\mathbf{u}}_i[\mathbf{x}], \tag{3.23}$$

with

$$\tilde{\mathbf{K}}[\mathbf{x}]\hat{\mathbf{u}}_i = \mu_i \mathbf{g}_i, \quad i = 1, ..., m \tag{3.24}$$

and

$$g_{\mathcal{K},i}[\mathbf{x}] = \frac{1}{\gamma_i^2} \hat{\mathbf{f}}_i[\mathbf{x}] \cdot \hat{\mathbf{u}}_i, \tag{3.25}$$

with

$$\tilde{\mathbf{K}}[\mathbf{x}]\,\gamma_i \mathbf{v}_i = \hat{\mathbf{f}}_i, \quad i = 1, ..., d. \tag{3.26}$$

As will become apparent in Section 3.2.3, to obtain the sensitivities to all variables the motions at all structural DOFs is required. Because the compliance of GDIs that are not of interest is zero, see Equation (3.12), we may write

$$\hat{\mathbf{f}} \cdot \hat{\mathbf{u}} = \mathbf{f} \cdot \mathbf{u}, \tag{3.27}$$

for both motion and load-based scenarios. The motions can be obtained by finding the solution to the system of equations in Equation (3.10) without generation of the reduced-order stiffness matrix.

Each motion-based scenario $\mathcal{S}_{\mathrm{p},i}$ is entirely governed by the prescribed motion of one free MP $\mathbf{v}_i$ and corresponding magnitude γ_i. The unknown free motions $\check{\mathbf{u}}_{\mathrm{f},i}$ can be determined as a function of the prescribed MPs $\hat{\mathbf{u}}$ by solving the design dependent system of equations

$$\check{\mathbf{K}}_{\mathrm{ff}}[\mathbf{x}]\check{\mathbf{u}}_{\mathrm{f},i} = -\check{\hat{\mathbf{K}}}_{\mathrm{f}}[\mathbf{x}]\hat{\mathbf{u}}_i, \tag{3.28}$$

with

$$\hat{\mathbf{u}}_i = \gamma_i \mathbf{v}_i, \quad \forall\ i = 1, ..., m. \tag{3.29}$$

In each of the load-based scenarios $\mathcal{S}_{\mathrm{f},i}$ a load μ_i is applied to the ith GDI, such that $\hat{\mathbf{f}}_i = \mu_i \mathbf{g}_i$, with $\mathbf{g}_i$ the basis LP at the ith GDI. Thus the motion at the GDIs $\hat{\mathbf{u}}$ is *not* prescribed for these load-based scenarios. The unknown free motions for the load-based scenarios can be determined as a function of the applied loads at the GDIs $\hat{\mathbf{f}}$ by solving the design dependent system of equations

$$\begin{bmatrix} \hat{\hat{\mathbf{K}}} & \hat{\check{\mathbf{K}}}_{\mathrm{f}} \\ \check{\hat{\mathbf{K}}}_{\mathrm{f}} & \check{\mathbf{K}}_{\mathrm{ff}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_i \\ \check{\mathbf{u}}_{\mathrm{f},i} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_i \\ \mathbf{0} \end{bmatrix}, \tag{3.30}$$

[12]

with

$$\hat{\mathbf{f}}_i = \mu_i \mathbf{g}_i, \quad \forall\ i = 1, ..., d. \tag{3.31}$$

The dominant contributor to the computational effort is finding the solutions to the discretized governing equations Equations (3.28) and (3.30). Independent of the number of mechanism DOFs and GDIs, the proposed optimization problem formulation requires

[12] Design dependency is omitted here for brevity.

two partitions, namely one for (i) the load-based scenarios in which the motion of the GDIs is free, and one for (ii) the motion-based scenarios in which the motion at the GDIs is prescribed. Thus, to find the required solutions to the governing equations, *two* factorization or preconditioning steps have to be performed, one for each partition. The number of solves is the sum of the number of GDIs m (for the objective) and the mechanism DOFs d (for the constraints).

Sensitivity analysis

The optimization problem formulation consists uniquely on energy (derived) measures, similar to the compliance minimization problem. The sensitivities of the characteristic stiffness measures with respect to design variable x_i read as

$$\begin{aligned}\frac{\mathrm{d}\mathcal{K}_{\mathrm{p},j}[\mathbf{x}]}{\mathrm{d}x_i} &= \mathbf{v}_j \cdot \frac{\mathrm{d}\tilde{\mathbf{K}}[\mathbf{x}]}{\mathrm{d}x_i}\mathbf{v}_j \\ &= \frac{1}{\gamma_j^2}\hat{\mathbf{u}}_j \cdot \frac{\mathrm{d}\tilde{\mathbf{K}}[\mathbf{x}]}{\mathrm{d}x_i}\hat{\mathbf{u}}_j \\ &= \frac{1}{\gamma_j^2}\mathbf{u}_j \cdot \frac{\mathrm{d}\mathbf{K}[\mathbf{x}]}{\mathrm{d}x_i}\mathbf{u}_j,\end{aligned} \tag{3.32}$$

and

$$\begin{aligned}\frac{\mathrm{d}\mathcal{K}_{\mathrm{f},j}^{-1}[\mathbf{x}]}{\mathrm{d}x_i} &= \mathbf{g}_j \cdot \frac{\mathrm{d}\tilde{\mathbf{C}}[\mathbf{x}]}{\mathrm{d}x_i}\mathbf{g}_j \\ &= -\mathbf{g}_j\tilde{\mathbf{K}}^{-1} \cdot \frac{\mathrm{d}\tilde{\mathbf{K}}[\mathbf{x}]}{\mathrm{d}x_i}\tilde{\mathbf{K}}^{-1}\mathbf{g}_j \\ &= -\frac{1}{\mu_j^2}\hat{\mathbf{u}}_j \cdot \frac{\mathrm{d}\tilde{\mathbf{K}}[\mathbf{x}]}{\mathrm{d}x_i}\hat{\mathbf{u}}_j \\ &= -\frac{1}{\mu_j^2}\mathbf{u}_j \cdot \frac{\mathrm{d}\mathbf{K}[\mathbf{x}]}{\mathrm{d}x_i}\mathbf{u}_j.\end{aligned} \tag{3.33}$$

The sensitivities can thus be calculated based on elemental stiffness matrix and of the displacement field only (or otherwise the elemental strain energy), that are common commercial FEA output data, via

$$\mathbf{u} \cdot \frac{\mathrm{d}\mathbf{K}[\mathbf{x}]}{\mathrm{d}x_i}\mathbf{u} = \mathbf{u}_i \cdot \frac{\mathrm{d}\mathbf{K}_i[x_i]}{\mathrm{d}x_i}\mathbf{u}_i, \tag{3.34}$$

for all $i = 1, ..., N$ and where $\mathbf{u}_i$ and $\mathbf{K}_i$ denote the motion and stiffness at the structural DOFs corresponding to the ith finite element. Similar to the classic compliance minimization problem, no adjoint equations are to be solved to obtain the sensitivity information because of the self-adjoint property of the energy-derived responses. Hence, the computational effort of the sensitivity analysis is negligible.

Note the difference between the types of scenarios used in objective and constraints in relation to the conclusion of the study presented in Section 3.2.1. The sensitivities of an energy derived stiffness measure from a load-based scenario are *strictly negative*, whereas the sensitivities of an energy measure from a motion-based scenario are *strictly positive*

for all design variables throughout all design iterations. As a consequence of this strict sign difference and the type of responses, the optimization problem formulation has an self-penalizing character thus favouring black-and-white solutions, as will be visually clear from the numerical examples in Section 5.5.

3.2.4 Coordinate transformations

In this section we elaborate on why and how to use coordinate transformations in the context of the proposed topology optimization problem formulation. Coordinate transformations can be a useful tool to (i) apply constraints on the relative motion between structural DOFs, (ii) introduce new GDIs such as rotational DOFs, and (iii) build a reduced 'mechanism stiffness model' to describe the mechanism behaviour in terms of GDIs without loss of accuracy. Thus, coordinate transformations (i) allow for trivial and natural description of GDIs, and (ii) can enhance computational efficiency.

We emphasize that these methods may enhance the applicability and computational efficiency, but the working principle of the proposed optimization problem formulation does not depend on these extensions. In what follows we have for brevity omitted the dependency of the stiffness matrix and derived measures on the design variables.

Introduction of new generalized degrees of freedom

As previously noted in Section 3.2.1, the kinematics of a CM are predominantly defined by GDIs. This includes structural DOFs, but also derived generalized DOFs such as a rigid-body rotation. Such GDIs are, typically, not present in the original set of structural DOFs, thus new DOFs have to be introduced. This can be accomplished via multi-point constraints, such as primary-secondary, penalty augmentation or Lagrange multiplier adjunction methods (Abel *et al.* 1979). For the present linear setting, we opt to apply one of the most basic and well known primary-secondary multi-point constraint methods to elucidate on the introduction of new and generalized DOFs. Consider thereto the coordinate transformation

$$\mathbf{u} = \mathbf{T}\overline{\mathbf{u}}, \tag{3.35}$$

with the linear transformation matrix $\mathbf{T} \in \mathbb{R}^{n \times c}$ (generally highly sparse) and the motion of the *generalized DOFs* $\overline{\mathbf{u}} \in \mathbb{R}^{c}$. Note that in addition to the new DOFs, the set of generalized DOFs may, and typically does, contain part of the structural DOFs. Thus, the number of generalized DOFs c is typically smaller than the number of structural DOFs n. Substitution of Equation (3.35) in Equation (3.2) and subsequent pre-multiplication of both sides with the transpose of the transformation matrix, yields the *generalized system of equations*, that is

$$\overline{\mathbf{K}}\overline{\mathbf{u}} = \overline{\mathbf{f}}, \tag{3.36}$$

with the *generalized stiffness matrix* $\overline{\mathbf{K}} \in \mathbb{R}^{c \times c}$ given by

$$\overline{\mathbf{K}} = \mathbf{T}^{\mathsf{T}}\mathbf{K}\mathbf{T}, \tag{3.37}$$

and $\overline{\mathbf{f}} := \mathbf{T}^{\mathsf{T}}\mathbf{f} \in \mathbb{R}^{c}$ the generalized loads. Subsequently, the sensitivities of the generalized stiffness matrix any of the design variables simply read

$$\frac{\mathrm{d}\overline{\mathbf{K}}}{\mathrm{d}x_i} = \mathbf{T}^{\mathsf{T}}\frac{\mathrm{d}\mathbf{K}}{\mathrm{d}x_i}\mathbf{T}. \tag{3.38}$$

As such, the transformation does not introduce significant computational effort nor complexity. Next, the generalized system of equations from Equation (3.36) can be partitioned similar to Equation (3.10), and the solution(s) may, following the description in Section 3.2.3, form the basis for the energy derived stiffness measures as required for the response functions. An illustrative example of the introduction of new GDIs will be presented in Section 3.3.3.

Exact reduced-order mechanism stiffness model

In the design of CMs, we are predominantly interested in the behaviour of the GDIs. Thereto, we desire to describe the mechanism behaviour in a natural manner as a function of the GDIs only.

In what follows, we continue on the system of equations from Equation (3.2). However, without loss of generality, the following may equivalently be applied to the generalized system of equations of Equation (3.36) if new and generalized DOFs have been introduced.

Similar to the partitioning as per Equation (3.10), yet omitting the additional partitioning in free and prescribed DOFs, we use oversets ($\wedge$) and ($\vee$) to represent the properties related to GDIs and DOFs not of interest, respectively, as

$$\begin{bmatrix} \overset{\wedge\wedge}{\mathbf{K}} & \overset{\wedge\vee}{\mathbf{K}} \\ \overset{\vee\wedge}{\mathbf{K}} & \overset{\vee\vee}{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \overset{\vee}{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}} \\ \overset{\vee}{\mathbf{f}} \end{bmatrix}, \tag{3.39}$$

where the principal sub-matrix $\overset{\vee\vee}{\mathbf{K}}$ is *nonsingular*, such that $\overset{\vee\vee}{\mathbf{K}}$ is invertible.

Upon application of static condensation (Guyan 1965; Irons 1965; Koppen *et al.* 2022c), one may condense the DOFs that are *not* of interest and obtain the exact reduced-order system of equations, or similar the 'mechanism stiffness model' (Hasse *et al.* 2009; Jin *et al.* 2018; Wang 2009a) as presented in Equation (3.11). Herein, the reduced but exact mechanism system matrix is now defined as

$$\tilde{\mathbf{K}} := \overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge\vee}{\mathbf{K}}\overset{\vee\wedge}{\mathbf{X}}, \tag{3.40}$$

with $\overset{\vee\wedge}{\mathbf{X}} \in \mathbb{R}^{(n-m)\times m}$ the solution of the system of equations

$$\overset{\vee\vee}{\mathbf{K}}\overset{\vee\wedge}{\mathbf{X}} = \overset{\vee\wedge}{\mathbf{K}}. \tag{3.41}$$

This may equivalently be written as a coordinate transformation as per Equation (3.35) with $\overline{\mathbf{u}} = \hat{\mathbf{u}}$ and

$$\mathbf{T} := \begin{bmatrix} \mathbf{I} \\ -\overset{\vee\wedge}{\mathbf{X}} \end{bmatrix}. \tag{3.42}$$

For a detailed elaboration of corresponding sensitivity analysis, and the impact of static condensation on optimization problem formulations involving multiple loading conditions in general the reader is referred to Koppen *et al.* (2022a,c).

Upon application of static condensation as described above, the computational effort, in contrast to the non-condensed version as described in Section 3.2.3, lies mainly in finding the solution to Equation (3.41). Note that, this requires *only* a single factorization or preconditioning step and m solves requiring considerable computational effort, as opposed

to two factorization or preconditioning steps and $m + d$ solves. Similar to what is outlined in Section 3.2.3,the sensitivity information only requires readily available information obtained in the forward analysis, *i.e.* the solution of Equation (3.41), thus does not introduce additional computational effort. Consequently, one may reasonably expect approximately 50% reduction of computational effort in the evaluation of response function and design sensitivities values of the proposed problem formulation for solution methods that spend the majority of effort in factorization/preconditioning. Thus, the condensation allows for a natural description of the mechanism behaviour in terms of GDIs with negligible additional complexity and reduced computational effort.

3.3 Numerical examples

In this section we demonstrate the implementation, working principle and use of the proposed optimization problem formulations and highlight the advantages and limitations. This section includes the synthesis of a single-DOF CM inverter, to elucidate the use and working principle of the proposed optimization problem formulation in Section 3.3.2 and parameter study to investigate the influence of the free MP and constraint values on the motion transmission accuracy in Section 3.3.2.[13] Next we propose and briefly investigate an adaptive formulation that allows to enhance the motion transmission accuracy in Section 3.3.2 and propose a reduced robust problem formulation in Section 3.3.2. This is followed by demonstration of the introduction of new and generalized DOFs and the use of the exact reduced-order mechanism stiffness model on the design of several relevant design problems in Sections 3.3.2 to 3.3.4. Special emphasis is put on the demonstration of application of the formulation to the design of a decoupled multi-DOF CM and shape-morphing structure with multiple decoupled mechanism DOFs in Sections 3.3.4 and 3.3.5. Prior to these demonstrations, we provide a concise description of parametrization, discretization, analysis and optimization considerations of the implementation of the numerical examples and corresponding codes provided with this paper.

3.3.1 Topology optimization considerations

Independent of the optimization problem formulation as presented, the user has to consider, select and implement a variety of methods to effectively use the formulation. The following aids in the consideration and implementation of design discretization and parametrization, design variable filtering, topology optimization material interpolation, and gradient-based optimization. Unless otherwise indicated, all numerical examples employ the implementation choices described here. The default constants used in the examples, as implemented in the provided codes, are listed in Table 3.2. We have opted to use practical values (as opposed to non-dimensional values) for the material properties, design volume dimensions and constraint values to emphasize the practical use of the proposed formulation.

For the numerical analysis in all 2D examples, we opt to use standard bi-linear four-node quadrilateral finite elements in structured meshes. The design domain is parametrized by assignment of one design variable $x_i \in \mathbb{X} := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ to each finite element $i = 1, ..., N$, which allows for local control of the material properties, such as density and Young's modulus (Bendsøe 1989).

[13] The definition of the motion transmission accuracy will be given in Section 3.3.2.

Table 3.2: Constant parameters and assigned values.

Symbol	Description	Value
x^0	Homogeneous initial design	0.25
$\overline{v}$	Maximum volume fraction	0.25
ϵ	Design change threshold	5×10^{-4}
$\alpha_{\mathcal{K}}$	Scaling stiffness constraints	1
α_v	Scaling volume constraint	1
r	Normalized filter radius	2.0
ε	Stiffness ratio	10^{-9}
p	SIMP penalty	3.0
E	Young's modulus	1 GPa
ν	Poisson ratio	0.3
l	Length	100 mm
t	Thickness	10 mm

To eliminate modeling artifacts, the design variable field is generally blurred to obtain the filtered field $\tilde{\mathbf{x}} \in \mathbb{X}^N$ using a linear convolution operator with normalized (with respect to the element size) filter radius $r \in \mathbb{R}^+$ (Bruns *et al.* 2001). This operation is also accounted for in the sensitivity calculation, as described in the cited reference.

To locally control the stiffness of the material, the Young's modulus of each finite element is related to the filtered design variable via an element-wise composite rule, that is

$$\frac{E_i[\tilde{x}_i]}{\overline{E}} = \varepsilon + (1-\varepsilon) R[\tilde{x}_i], \tag{3.43}$$

with $\overline{E}$ the material Young's modulus, ε the relative stiffness between 'solid' and 'void' material and R the material interpolation function. We opt to use the most commonly used 'modified Solid Isotropic Material with Penalization' (SIMP) interpolation function (Bendsøe 1989; Sigmund 2007). This function ensures correct modeling of material with intermediate Young's modulus, and stimulates solutions at which variables are at the lower or upper bound, so-called black-and-white solutions, generally preferred for manufacturability considerations.

It is generally recognized that the optimized solution found is sensitive to the initial design, hereafter denoted by $\mathbf{x}^{\{0\}}$, in which the superscript indicates the design iteration number. This sensitivity is particularly high for CM design optimization problems (Chen *et al.* 2017; Sigmund *et al.* 1998). We consider the impact of the initial design on the optimized solution and performance out of the scope of this paper, and opt for the commonly used homogeneous initial design $x_i^{\{0\}} = x^0 = \overline{v}$ for all $i = 1, ..., N$.

Even for design optimization problems with perfectly symmetric boundary and loading conditions it is, although not often explicitly reported, common and inevitable to obtain asymmetric optimized solutions. One may enforce symmetry by removal of duplicate variables or by linking design variables over one or multiple axes; either by creating a dependency or by averaging, as we utilize in the numerical examples Sections 3.3.3 and 3.3.4.

As mentioned in Section 3.2.3, it is generally beneficial to (re)scale the objective to a reasonable value, as compared to the constraints. We opt to normalize the individual load-based characteristic stiffnesses to their value at the first optimization iteration. As a consequence of this scaling, the magnitude of the applied loads becomes irrelevant and the individual terms hold the same value at the first design iteration.

The inequality-constrained nonlinear optimization problem is solved in a nested analysis and design setting. The design variables are iteratively updated by a sequential approximate optimization scheme, as is common in the topology optimization field. As standard within the TO community, and without loss of generality, we opt to approximate the optimization problem using the popular 'Method of Moving Asymptotes', use default parameter values for the update scheme as proposed by Svanberg (1987), and solve the resulting convex and separable subproblems using a gradient-based primal-dual interior point method. The optimization process is terminated when the mean absolute design change is smaller than a tolerance, hereafter denoted ϵ, normalized by the number of design variables.

3.3.2 Design of a single-DOF compliant inversion mechanism

We demonstrate the design of a single-DOF CM inverter to verify and elucidate the working principle of the proposed formulation. Consider the problem setting and motion-based use case scenario sketched in Figure 3.3a, and corresponding scenarios in Figure 3.1 required to set up the problem formulation. Discretization according to the considerations in Section 3.3.1, followed by stationary potential energy yields the equilibrium equations as per Equation (3.2). The GDIs are the displacements at the input and output ports are given in Figure 3.3a. Application of the exact reduced-order modelling technique discussed in Section 3.2.4, followed by partitioning in input (i) and output (o) GDIs yields the generalized reduced-order system of equations of interest

$$\begin{bmatrix} k_{\mathrm{ii}} & k_{\mathrm{io}} \\ k_{\mathrm{io}} & k_{\mathrm{oo}} \end{bmatrix} \begin{bmatrix} u_{\mathrm{i}} \\ u_{\mathrm{o}} \end{bmatrix} = \begin{bmatrix} f_{\mathrm{i}} \\ f_{\mathrm{o}} \end{bmatrix}, \tag{3.44}$$

with u_{i} and u_{o} the displacements at input and output ports, respectively.[14] The behaviour of the single-input-single-output compliant mechanism is expressed in terms of the motion and loads at the GDIs and fully governed by the intrinsic properties $k_{\mathrm{ii}}[\mathbf{x}] \in \mathbb{R}^+$, $k_{\mathrm{io}}[\mathbf{x}] \in \mathbb{R}$, and $k_{\mathrm{oo}}[\mathbf{x}] \in \mathbb{R}^+$.

Next we set up the optimization problem formulation Equation (3.22) for this design problem, which reads

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{X}^N} \quad & f[\mathbf{x}] : \mathcal{K}_{\mathrm{f,i}}^{-1}[\mathbf{x}] + \mathcal{K}_{\mathrm{f,o}}^{-1}[\mathbf{x}] \\ \text{s.t.} \quad & g_{\mathcal{K}}[\mathbf{x}] : \frac{\mathcal{K}_{\mathrm{p}}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}} - 1 \leq 0 \\ & g_{\mathrm{v}}[\mathbf{x}] : \frac{v[\mathbf{x}]}{\overline{v}} - 1 \leq 0 \end{aligned} \tag{3.45}$$

[14] Design dependency is omitted here for brevity.

with the objective function defined as the sum of load-based characteristic input and output compliances. The constraint $g_{\mathcal{K}}$ limits the maximum motion-based characteristic stiffness of the desired (non dimensional unit length) free MP $\mathbf{v} = \begin{bmatrix} v_\mathrm{i} & v_\mathrm{o} \end{bmatrix}^\mathsf{T}$ to $\overline{\mathcal{K}}_\mathrm{p}$.

For the remainder of this example we choose, the entries of the free MP such that

$$\frac{v_\mathrm{o}}{v_\mathrm{i}} = J^*, \tag{3.46}$$

with J^* the desired motion transmission, also referred to as the geometric advantage. A topology, defined by design variables $\mathbf{x}$ of optimization problem Equation (3.45) will have a true motion transmission ratio defined by

$$J[\mathbf{x}] := \frac{u_\mathrm{o}[\mathbf{x}]}{u_\mathrm{i}} = -\frac{k_\mathrm{io}[\mathbf{x}]}{k_\mathrm{oo}[\mathbf{x}]}, \tag{3.47}$$

that is the magnitude of the output motion for a prescribed unit input motion. This can be trivially obtained using the second line of Equation (3.44) and the given use case for which $f_\mathrm{o} = 0$. Note that the motion transmission is independent of the input stiffness k_ii following the motion-based use case scenario.

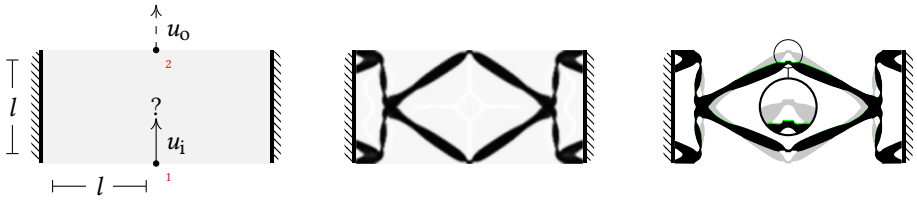
The filtered design variable field $\tilde{\mathbf{x}}^*$ of a solution to Equation (3.44) for a desired motion transmission $J^* = -1$ and maximum characteristic stiffness of corresponding free MP of $\overline{\mathcal{K}}_\mathrm{p} = 10\,\mathrm{N\,mm^{-1}}$ is displayed in Figure 3.3b. Deformed configurations of a post-processed (simple thresholding at $\tilde{\mathbf{x}}^* = 0.5$) version of the solution posed in Figure 3.3b are displayed in Figure 3.3c. This figure contains the topology in (i) undeformed configuration (gray), (ii) deformation as obtained by imposing the free MP (green), and (iii) deformation obtained by the use case as given in Figure 3.3a.

The optimization process and mechanism properties as a function of design iteration k are displayed in Figure 3.4. We can roughly distinguish two phases of the optimization process. In the first phase (typically around 10–20 iterations) a feasible solution is sought and the topology is roughly defined, see Figures 3.4a and 3.4b. Subsequently, in the second phase, the solution is gradually modified while steadily increasing the ratio between the characteristic stiffnesses of GDIs versus free MP, therewith the true motion transmission ratio $J[\mathbf{x}]$ approaches the intended motion transmission J^*, see Figures 3.4c and 3.4d. Furthermore, as displayed in Figure 3.4d, the chosen objective function increases the GDI stiffnesses, while satisfying the maximum characteristic stiffness of the free MP. The obtained solution is vertically symmetric, while this is not enforced. Thus, for this design problem the input and output stiffnesses are equal, that is $\mathcal{K}_{\mathrm{f,i}}[\mathbf{x}^*] = \mathcal{K}_{\mathrm{f,o}}[\mathbf{x}^*]$.

In line with the working principle as discussed in Section 3.2.3, we observe a difference between the intended motion transmission J^* and the motion transmission of the optimized solution $J[\mathbf{x}^*]$. This deviation is clearly visible by the difference between the output motion for imposed free MP and obtained output motion as per use case, see the green and black topologies and deviation in the zoom-in of Figure 3.3c. To gain more insight in this, the next subsection presents a study of this deviation for a variety of desired transmissions and stiffnesses.

Parametric study

As observed, the kinematics of the optimized solution typically differ from the imposed free MP (Figure 3.3c); an error between the desired and true transmission is present, see also

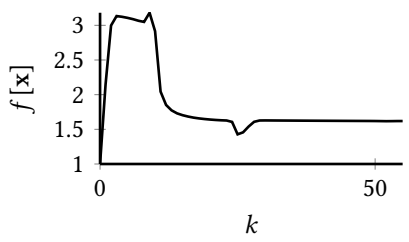


(a) Problem setting and motion-based use case scenario indicating the design domain of size $2l \times l \times t$, boundary and loading conditions of the use case. Both left and right sides are fully clamped. The problem involves two generalized DOFs of interest, that are the prescribed input motion and desired output motions.

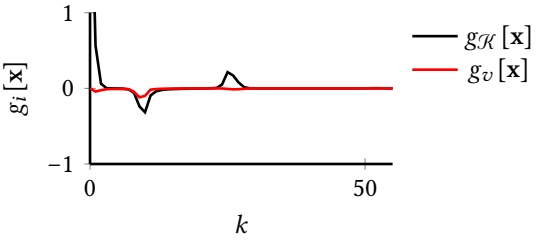
(b) Filtered design variable field $\tilde{\mathbf{x}}^*$ of a topology optimized single-degree-of-freedom compliant mechanism. This mechanism inverts the direction while preserving the magnitude of motion ($J^* = -1$).

(c) Post-processed deformed geometries of the optimized topology as a result of (i) the use case (black), that is a prescribed input motion u_i (as per motion-based scenario in Figure 3.3a), and (ii) the imposed free MP (green) with $v_i = u_i$ as per Figure 3.1 (right).

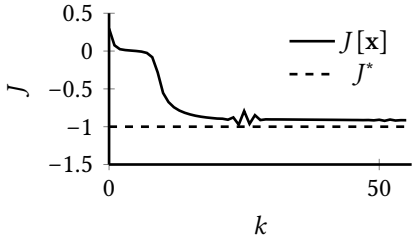
Figure 3.3: Problem setting and motion-based use case scenario, optimized topology and corresponding deformed configurations for the design of a single mechanism degree-of-freedom compliant mechanism inverter. The use case involves a desired motion transmission, that is a desired output motion u_o for a prescribed input motion u_i.



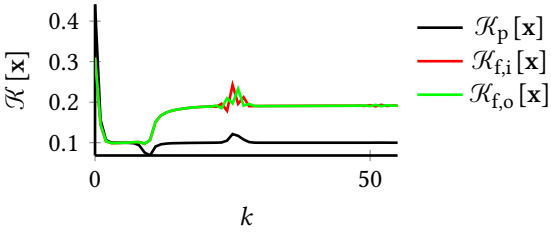
(a) Objective scaled relative to objective value at first design iteration.



(b) Constraint values.



(c) Desired and true motion transmission.



(d) Load and motion-based characteristic stiffnesses.

Figure 3.4: Optimization properties (objective, constraints) and mechanism properties (motion transmission and characteristic stiffness) as a function of design iteration number.

Table 3.3: Motion transmission accuracy $\eta\left[\mathbf{x}^*, J^*, \overline{\mathcal{K}}_\mathrm{p}\right]$ (%) of topology optimized compliant mechanism inverters obtained by solving optimization problem defined by Equation (3.45) for a variety of desired motion transmission J^* and constraint values $\overline{\mathcal{K}}_\mathrm{p}$ $\left(\mathrm{N\,mm^{-1}}\right)$.

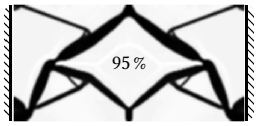
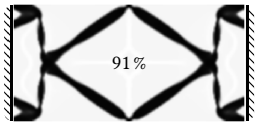
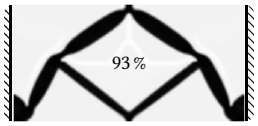
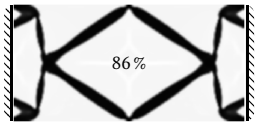
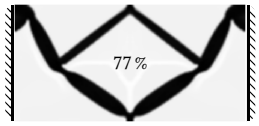
$\overline{\mathcal{K}}_\mathrm{p}$ \ J^*	-0.5	-1.0	-2.0
5	 97 %	 95 %	 90 %
10	 95 %	 91 %	 85 %
20	 93 %	 86 %	 77 %

Figure 3.4c. We will hereafter define the relative motion transmission accuracy $\eta[\mathbf{x}] \in \mathbb{R}^+$ as

$$\eta[\mathbf{x}] := 100\,\% - \left\| \frac{J[\mathbf{x}]}{J^*} - 1 \right\| \cdot 100\,\%. \tag{3.48}$$

Using numerical examples, we now further study the influence of the value of the desired motion transmission J^* and constraint value $\overline{\mathcal{K}}_\mathrm{p}$ on the motion transmission accuracy; see the results of this parametric study in Table 3.3. The relative transmission accuracy $\eta[\mathbf{x}^*]$ tends to increase with magnitude of the target motion transmission J^* and characteristic stiffness of the free MP $\overline{\mathcal{K}}_\mathrm{p}$. For any non-singular system, forcing the system in the free MP requires work. Upon release of the output port, *i.e.* changing the boundary condition of the DOF related to the output motion from prescribed to free, part of the stored energy is released. This reduction of energy comes with a decrease of motion transmission and thus a loss of accuracy with respect to the desired motion transmission. One may reasonably expect a larger reduction of energy—and thus decrease of motion transmission—for larger magnitudes of the desired motion transmission. Larger values of the characteristic stiffness of the free MP(s) come with large release of stored energy upon release of the output port, and, thus, larger decrease of motion transmission accuracy.

On the topologies we observe that there is vertical symmetry for $J^* = -1$ (which is not forced in the optimization process), and that the topologies for J^* and $\frac{1}{J^*}$ are their vertically symmetric counterparts, but their corresponding relative transmission accuracies differ largely. What is more, there seems to be *no* strong tendency for the creation of very short hinges with highly localized deformation; nor is there a clear tendency for the creation of long hinges with distributed deformation.

Adaptive formulation

The results from the previous section Table 3.3 clearly indicate the magnitudes of both the free MP and characteristic stiffness influence the motion transmission accuracy. For many practical problems, limiting parasitic coupling is more important than reaching an exact transmission, as this can be easily compensated by the input motion. In case an exact transmission is required, one may opt for adaptations or variations to closer satisfy the desired kinematics. One such possible solution is to *adaptively* adjust the motion transmission of the imposed free MP throughout the optimization process. We define the imposed motion transmission ratio of the non-dimensional unit length free MP $\mathbf{v}^{\{k\}}$, at design iteration k, as

$$\phi^{\{k\}} := \frac{v_{\mathrm{o}}^{\{k\}}}{v_{\mathrm{i}}^{\{k\}}}. \tag{3.49}$$

For demonstration purposes we introduce the simple update scheme given by

$$\phi^{\{k\}} = \phi^{\{k-1\}} + \beta \left(J\left[\mathbf{x}^{\{k-1\}}\right] - J^{*} \right), \tag{3.50}$$

with $\beta \in \mathbb{R}^{+}$ an algorithmic damping parameter.[15] The results are quite insensitive to the value of the damping parameter; values of $0.01 \leq \beta \leq 0.1$ typically provide desirable results.

The result of applying such scheme to the design of the CM inverter with $J^{*} = -1$ for a variety of damping values is given in Table 3.4. The graphs in Table 3.4 display the values of $\phi^{\{k\}}$ and $J\left[\mathbf{x}^{\{k\}}\right]$ throughout the optimization process for a variety of algorithmic damping values β. Comparison of both topology and attained transmission accuracy to the design obtained *without* adaptive scheme in Figure 3.3b indicates the performance can be clearly improved and vertically non-symmetric solutions are found. This simple scheme clearly provides the desired result without negative effects on topology or optimization convergence properties. We observe that, to retain a similar optimization convergence behaviour, CMs with higher desired motion-based characteristic stiffness values require larger damping. With such an adaptive scheme, we thus effectively eliminate the dependency of the transmission accuracy on both the imposed motion transmission ratio and characteristic stiffness constraint value.

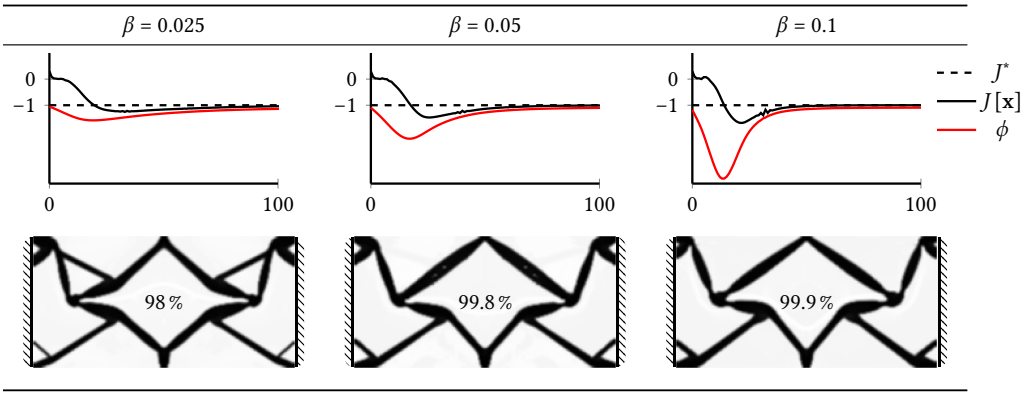
Robust design

The desired kinematics and stiffness of a compliant mechanism, and thus its performance, are sensitive to varying geometric deviations of the nominal design arising, for example, from the finite precision of the manufacturing process. In classical deterministic topology optimization, the effect of such uncertain parameters on the performance of the structure is not taken into account. This may lead to a design that is very sensitive to manufacturing errors; consequently, the performance of the actual structure may be far from optimal.

Sigmund (2009) and Wang *et al.* (2011) propose a robust approach to topology optimization where the effect of uniform manufacturing errors is taken into account. Uniform erosion and dilation effects, from here on denoted by superscripts (e) and (d), are simulated by means of a projection method: the filtering of the design variable field is followed by a differentiable Heaviside projection using a high projection threshold $\eta^{\mathrm{e}} = \eta^{\mathrm{i}} + \Delta\eta$ to simulate

[15] We emphasize this is 'just' one of the many possible adaptive schemes.

Table 3.4: Target J^*, true $J\left[\mathbf{x}^{\{k\}}\right]$ and adaptively imposed $\phi^{\{k\}}$ motion transmission and optimized solution (filered design variable field) for the proposed optimization problem formulation with *adaptive* scheme applied to the optimization problem posed in Figure 3.3a with $J^* = -1$ for different values of algorithmic damping β.



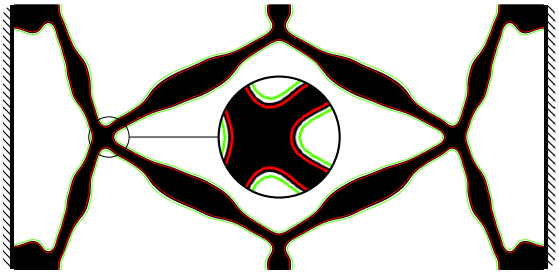


Figure 3.5: Topology optimized compliant inverter mechanism designed using the proposed 'reduced' robust problem formulation from Equation (3.51) applied to the optimization problem Equation (3.44) and corresponding problem setting displayed in Figure 3.1. This solution corresponds to a maximum load-based characteristic stiffness of $\overline{\mathcal{K}}_{\mathrm{p}} = 3\,\mathrm{N\,mm^{-1}}$, filter radius $r = 6$ and constant projection threshold variation of $\Delta\eta = 0.1$. The contour lines indicate corresponding eroded (red) and dilated (green) projections. The projection intensity factor β is continuously increased from linear projection to close to a step-wise projection using as to enforce minimum feature size, in this case via $\beta^{\{k\}} = \max\left(1.05\beta^{\{k-1\}}, 20\right)$ with $\beta^{\{0\}} = 0.5$.

an erosion and a low projection threshold $\eta^{\mathrm{d}} = \eta^{\mathrm{i}} - \Delta\eta$ to simulate a dilation. Here $0 < \eta^{\mathrm{i}} < 1$ is the projection threshold of the intermediate or blue-print design, typically $\eta^{\mathrm{i}} = 0.5$, and $\Delta\eta$ the projection threshold variation. Additional advantages of the robust formulation are the direct control of the minimum feature size of both solid and void, circumventing designs with very thin hinges that typically exceed the manufacturing resolution, are prone to failure during manufacturing, post-processing or use, and exhibit high stress peaks. Consequently the resulting designs do not consist of one-node connected hinges; in fact (if the filter size is selected large enough) the optimized topologies tend to consist of flexures with distributed compliance, typically more desirable compared to their lumped compliance alternative. The value of the minimum feature size on both solid and void is a function of the chosen filter radius and projection threshold values (Fernández *et al.* 2020; Sigmund 2009; Wang *et al.* 2011). The increase in robustness typically comes with a loss of performance, see *e.g.* Fernández *et al.* (2020), Qian *et al.* (2013), and Silva *et al.* (2019).

To obtain such manufacturing-tolerant designs in the present setting, the optimization problem is typically formulated as a 'worst-case' formulation; minimizing the maximum objective value of the eroded, intermediate and dilated designs, that is $\min\left(\max\left(f\left[\mathbf{x}^{\mathrm{e}}\right], f\left[\mathbf{x}^{\mathrm{i}}\right], f\left[\mathbf{x}^{\mathrm{d}}\right]\right)\right)$, with $\mathbf{x}^{\mathrm{i}}$ the projection with intermediate threshold typically serving as blue-print or physical design. Without reformulation such min-max formulation are non-differentiable. Thereto, one typically resorts to smooth differentiable maximum approximation functions that introduce a large amount of non-linearities. The robust formulation, in its original form, requires to solve the finite element problem three times for each optimization iteration, as opposed to a single solve for the deterministic problem.[16]

The similarity of the proposed compliant mechanism optimization problem formulation to the classical 'volume constrained compliance minimization problem' and corresponding similarity in optimization problem properties poses similar advantages for robust design. Due to monotonic dependence of the compliance objective on the design variables the eroded design will always have the largest objective function value; the min-max formulation equates to minimization of the objective value of the eroded design (Sigmund 2009), hence requiring only a single finite element problem to be solved.

The proposed compliant mechanism optimization problem formulation as per Equation (3.22) holds the same monotonicity property for the objective. In addition, *all* constraint functions—equivalent to the volume constraint in the compliance minimization problem—are monotonic functions as well. The objective and constraints have strictly[17] opposite sign of design sensitivities. Thus, the dilated design will always have the largest constraint function values. Consequently, the constraints on the intermediate and eroded design can be omitted.

The proposed optimization problem formulation in a robust setting, much in line with the 'reduced' compliance minimization robust formulation (Lazarov *et al.* 2016; Sigmund

[16] These finite element problems are fully decoupled; corresponding solves can be executed in parallel.
[17] 'Strict' meaning here: for all variables and variable values throughout all design iterations.

2009), now reads

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f\left[\mathbf{x}^{\mathrm{e}}\right] : \sum_i \mathcal{K}_{\mathrm{f},i}^{-1}\left[\mathbf{x}^{\mathrm{e}}\right], && i \in \mathcal{S}_{\mathrm{f}} \\ \text{s.t.} \quad & g_{\mathcal{K},j}\left[\mathbf{x}^{\mathrm{d}}\right] : \mathcal{K}_{\mathrm{p},j}\left[\mathbf{x}^{\mathrm{d}}\right] \leq \overline{\mathcal{K}}_{\mathrm{p},j} && j \in \mathcal{S}_{\mathrm{p}} \\ & g_v\left[\mathbf{x}^{\mathrm{d}}\right] : v\left[\mathbf{x}^{\mathrm{d}}\right] \leq \overline{v} \end{aligned} \tag{3.51}$$

3

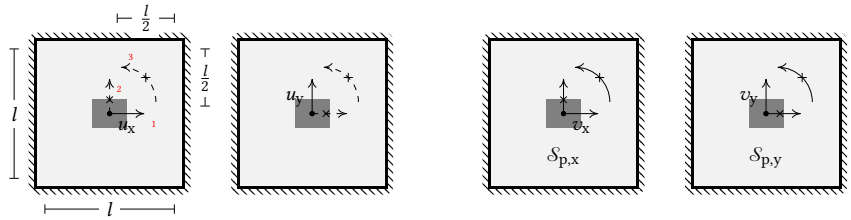
with $\mathbf{x}^{\mathrm{e}}[\mathbf{x}]$ and $\mathbf{x}^{\mathrm{d}}[\mathbf{x}]$ the eroded and dilated variable projections. The objective $f[\mathbf{x}^{\mathrm{e}}]$ is a summation of the *load-based characteristic compliances* based upon the eroded design, the constraint $g_{\mathcal{K},j}\left[\mathbf{x}^{\mathrm{d}}\right]$ limits the *motion-based characteristic stiffness* of the jth motion-based scenario based upon the dilated design, and constraint $g_v\left[\mathbf{x}^{\mathrm{d}}\right]$ limits the volume fraction of the dilated design. Equivalent to solving optimization problem Equation (3.22) (assuming no model reduction is used), solving the optimization problem Equation (3.51) requires two factorization or preconditioning steps. That is, the robust problem formulation adds little additional computational effort compared to its deterministic counterpart while providing valuable advantages as described, apart from the typically increase of design iterations due to the employed continuation scheme for the smooth Heaviside projection, which is commonly applied for robust formulations (Lazarov *et al.* 2016; Sigmund 2009; Wang *et al.* 2011). Furthermore, as the constraint on characteristic stiffness of free MPs are always active and the values of the dilated variable projection are strictly larger than the values of the eroded variable projection, the problem is self-penalizing and the penalty parameter in the material interpolation can be set to one, that is a linear interpolation suffices (Lazarov *et al.* 2016).

A solution to the proposed 'reduced' robust problem formulation from Equation (3.51) applied to the optimization problem Equation (3.44) and corresponding problem setting displayed in Figure 3.1 is provided in Figure 3.5; details of the chosen parameter values are given in the caption. Note the change of design with respect to the solutions from Table 3.3 (middle column), the contour lines indicating the eroded (red) and dilated (green) projections, and corresponding minimum feature size imposed on solid and void (see zoom-in). The loss of performance with respect to the deterministic variation is expressed in a loss of input and output stiffness, consequently decrease of contrast with respect to the motion-based characteristic stiffness, and thus decreased motion transmission accuracy.

Despite the advantages in terms of manufacturability and compliance distribution of using the robust formulation Equation (3.51) versus the deterministic formulation Equation (3.22), we opt to not further demonstrate the use of the robust formulation in upcoming examples, but rather continue with the core deterministic formulation and demonstrate the properties and results without interference from additional methods.

3.3.3 Design of a multi-DOF compliant spring

In many design problems the GDIs are not readily available from the set of structural DOFs, *e.g.* when considering rotations of rigid components. In what follows we (i) show why and how to introduce GDIs, (ii) set up the exact reduced-order mechanism stiffness model, and (iii) apply the proposed optimization problem formulation to the design of a multi-DOF spring-like compliant structure.



(a) Problem setting and motion-based use case scenarios, indicating the design domain of size $l \times l$ and boundary conditions. All sides are fully clamped. The problem involves three generalized DOFs of interest related to the translations and rotation of a rigid component in the centre of the design domain. The use cases are decoupled planar motion.

(b) Motion-based scenarios corresponding to the use cases in Figure 3.6a. Symbols according to previous examples and Table 3.1.

Figure 3.6: Topology optimization of a multi-DOF spring-like structure; problem setting, use cases and scenarios.

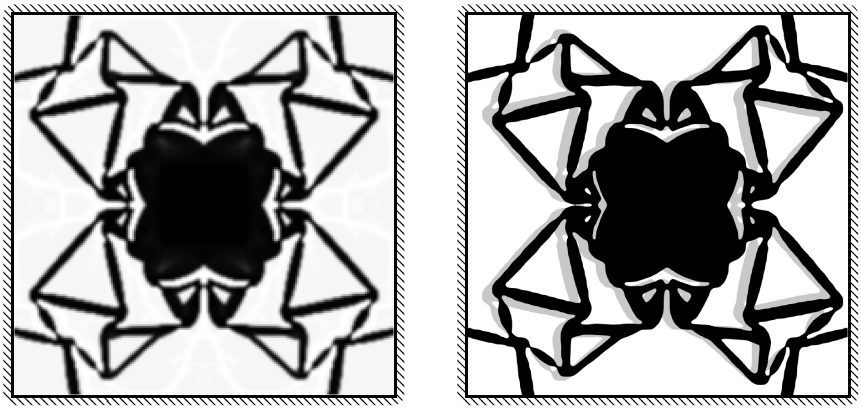


Figure 3.7: Topology optimized solution (filtered design variable field) and post-processed (thresholded at $\tilde{\mathbf{x}}^* = 0.4$) deformed structures of a multi-degree of freedom spring-like structure with high rotational stiffness. Since the input and output ports coincide (both centre), the deformed geometry for a prescribed u_x motion coincides exactly with the corresponding free MP, and vice versa for the other use case scenario.

Consider the problem setting and motion-based use case scenarios sketched in Figure 3.6a. This problem involves three GDIs, of which one is the rotational motion of the rigid component in the centre. To be able to describe the behaviour of the rotational motion we *first* describe the displacements of the rigid component as a function of newly introduced GDIs, that are the x and y displacements, as well as rotation θ of the rigid component, by application of multi-point constraints as described in Section 3.2.4. The rotational motion of the rigid component is described as function of the structural DOFs by linearized rigid-body kinematics, their relationship follows from basic use of a rotation matrix. Next, we apply static condensation, as described in Section 3.2.4 to obtain a reduced-order mechanism stiffness model exactly describing the behaviour of the GDIs involved in the design of this multi-DOF spring-like structure. Note the order of these steps is crucial for the computational efficiency; since first the new GDIs are introduced, only three solves are required to build the reduced-order mechanism stiffness model. Using the reverse order would in this case require as many solves as structural DOFs contained in the rigid body.

Since the units of rotational and translational DOFs differ, one cannot directly compare the magnitude of corresponding motions. What is more, the motion and load patterns need to consist of DOFs with the same units. Therefore, we unify the units of the rotational GDI by introduction of a characteristic length l_c, with its magnitude equal to the diagonal length of the rigid-body component (Section 2.4 of Jin *et al.* (2018)). Using this characteristic length, we define the equivalent displacement $u_\theta = l_c\theta$, with θ the rotation of the rigid-body component and corresponding equivalent force $f_\theta = \frac{1}{l_c}\tau$, with τ the torque. Such equivalent displacement corresponds to the magnitude of the displacement at the corner of the rigid component for a unit rotation. Although described here as a subsequent step, this unification can be written as a coordinate transformation and thus be included in the introduction of a new GDI using multi-point constraints. It is therefore, from an implementation point of view, simpler to directly introduce the new GDI u_θ as a function of the displacements of the rigid-body, as opposed to first introducing the new GDI θ, then applying static condensation and finally introducing the new GDI u_θ. After introduction of the unified new GDI and application of static condensation we obtain the generalized reduced-order system of equations that reads

$$\begin{bmatrix} k_{xx} & k_{xy} & k_{x\theta} \\ k_{xy} & k_{yy} & k_{y\theta} \\ k_{x\theta} & k_{y\theta} & k_{\theta\theta} \end{bmatrix} \begin{bmatrix} u_x \\ u_y \\ u_\theta \end{bmatrix} = \begin{bmatrix} f_x \\ f_y \\ f_\theta \end{bmatrix}. \tag{3.52}$$

The aim of the design of this multi-DOF spring-like structure is to synthesize it such that the rigid body has a specified stiffness in x and y translation while the rotational characteristic stiffness is maximized. Application of the proposed optimization problem

formulation Equation (3.22), gives

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] : \mathcal{K}_{\mathrm{f,x}}^{-1}[\mathbf{x}] + \mathcal{K}_{\mathrm{f,y}}^{-1}[\mathbf{x}] + \mathcal{K}_{\mathrm{f},\theta}^{-1}[\mathbf{x}] \\ \text{s.t.} \quad & g_{\mathcal{K},\mathrm{x}}[\mathbf{x}] : \frac{\mathcal{K}_{\mathrm{p,x}}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}} - 1 \leq 0 \\ & g_{\mathcal{K},\mathrm{y}}[\mathbf{x}] : \frac{\mathcal{K}_{\mathrm{p,y}}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}} - 1 \leq 0 \\ & g_{\mathrm{v}}[\mathbf{x}] : \frac{v[\mathbf{x}]}{\overline{v}} - 1 \leq 0 \end{aligned} \tag{3.53}$$

with $\mathcal{K}_{\mathrm{f},\theta}[\mathbf{x}]$ the load-based characteristic stiffness under for a basis equivalent force LP and $\mathcal{K}_{\mathrm{p,x}}[\mathbf{x}]$ and $\mathcal{K}_{\mathrm{p,y}}[\mathbf{x}]$ the motion-based characteristic stiffnesses under prescribed non-dimensional unit length free MPs $\mathbf{v}_{\mathrm{x}} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^{\mathsf{T}}$ and $\mathbf{v}_{\mathrm{y}} = \begin{bmatrix}0 & 1 & 0\end{bmatrix}^{\mathsf{T}}$ (order of GDIs as per Figure 3.6a).

Figure 3.7 displays the optimized solution of Equation (3.53) for a design that requires a motion-based characteristic stiffness of $\mathcal{K}_{\mathrm{p}} = 10\,\mathrm{N\,mm^{-1}}$. The solution has a reduced-order stiffness matrix that reads

$$\tilde{\mathbf{K}}[\mathbf{x}^*] = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.16 \end{bmatrix}. \tag{3.54}$$

As a result of symmetry, which is imposed in this example, any coupling is absent; all off-diagonal entries have zero stiffness. Because of this, the desired motion-based characteristic stiffness of the free MPs equates exactly to the corresponding diagonal entries of the reduced-order stiffness matrix. The translational stiffnesses are exactly as desired, and the rotational stiffness is substantially higher (more than one order); there is a clear separation between the characteristic stiffnesses of the constrained and free MPs.

3.3.4 Design of a decoupled multi-DOF positioning stage

Next we will demonstrate the use of the proposed problem formulation to the design of a multi-DOF CM with suppression of input coupling, output coupling and parasitic motions. To this end, consider the two-DOF CM design problem sketched in Figure 3.8 displaying the two motion-based use case scenarios (left, middle) and the motion-based scenario used for the setting up the problem formulation. Such design may for example be used as a precision positioning stage. The problem involves two desired mechanism DOFs that both include the motions of four ports, see the left and middle motion-based use case scenarios in Figure 3.8. The motion of each port is described by three DOFs, that are the translational motions in x and y direction, as well as rotation θ. The aim of this design optimization problem is to find a solution that allows for decoupled motion (translation in y direction) between the two mechanism DOFs, that is (i) motion from bottom-left to top-right, and (ii) motion from bottom-right to top-left. Apart from the desired decoupling of both input and output motions, we desire a minimum of parasitic motion—both translation in x direction and rotation—of the four ports. Despite the design for two mechanism DOFs, due to horizontal

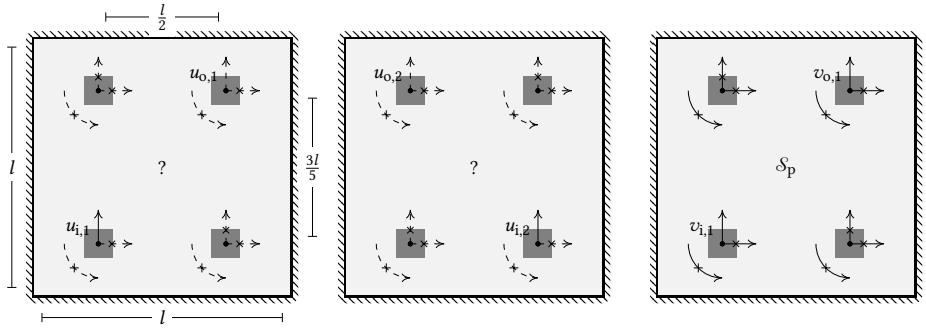


Figure 3.8: Problem setting, motion-based use case scenarios and the motion-based scenario used for the design of a multi-DOF positioning stage. The design domain is of size $l \times l$. All sides are fully clamped. The problem involves twelve generalized DOFs of interest related to the translations and rotation of the four ports. The left and middle figure display the use cases scenarios corresponding to the two mechanism DOFs. The right figure displays the motion-based scenario and corresponding free MP as used to measure the motion-based characteristic stiffness. Due to symmetry only a single motion-based scenario is required to set up the problem formulation.

symmetry *and* symmetry of the free MPs, only a single energy constraint is required; thus a single motion-based scenario suffices (see right figure of Figure 3.8).

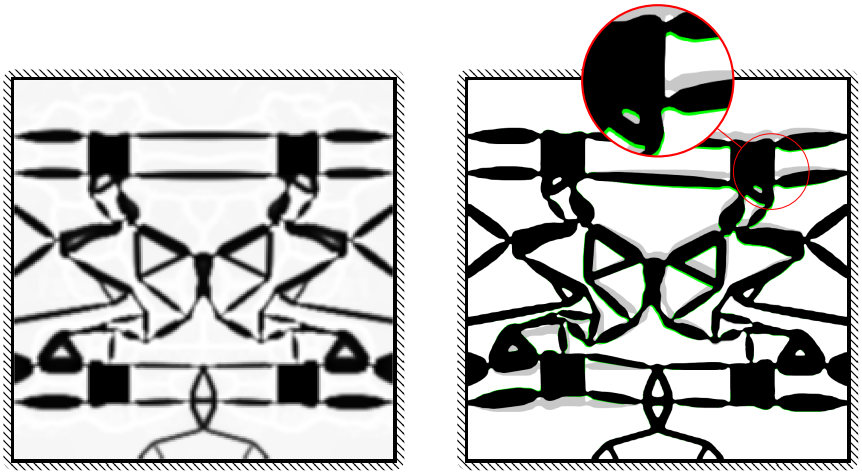
Note that this minimum working example is relatively complex, as the force paths between the two desired mechanism DOFs have to cross, making the structure inherently coupled; *i.e.* no trivially decoupled solution exists. A similar design optimization problem and solution using a conventional optimization problem formulation can be found in Koppen *et al.* (2022a). As indicated in Section 6.1, the design of such compliant structure typically involves a large number of optimization constraints, posing a large restriction on possible optimized solutions as well as design changes during the optimization process. In what follows, we intent to demonstrate the working principle and performance of replacing this set of constraints with a single constraint on the characteristic stiffness of the desired free MP, as per proposed problem formulation Equation (3.22).

Following the methods to introduce new generalized DOFs in Section 3.2.4, the introduction of characteristic lengths as per Section 3.3.3, followed by the reduction of dimensionality in Section 3.2.4, we obtain a mechanism stiffness model of reduced dimension that exactly describes the behaviour of the resulting twelve GDIs (two translational DOFs and one rotational DOF per port).

Application of the proposed optimization problem formulation gives

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] : \sum_i^{12} \mathcal{K}_{\mathrm{f},i}[\mathbf{x}] \\ \text{s.t.} \quad & g_{\mathcal{K}}[\mathbf{x}] : \frac{\mathcal{K}_{\mathrm{p}}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}} - 1 \leq 0 \\ & g_{\mathrm{v}}[\mathbf{x}] : \frac{v[\mathbf{x}]}{\overline{v}} - 1 \leq 0 \end{aligned} \tag{3.55}$$

with $\mathcal{K}_{\mathrm{f},i}[\mathbf{x}]$ the load-based characteristic stiffness of the ith basis LP obtained as per



(a) Topology optimized solution (filtered design variable field) of Equation (3.55).

(b) Undeformed (grey) and deformed (green and black) configurations of the solution of Figure 3.9a. The green deformation field is obtained by the imposed free MP, and the black deformation field corresponds to the use case. The imposed free MP is scaled such that $v_{i,1} = u_{i,1}$, and both are scaled for visualization purposes. Note the close correspondence between the deformed configurations.

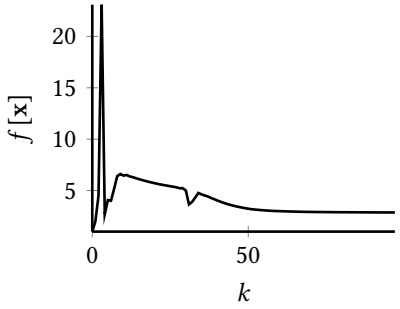
Figure 3.9: Solution and deformation of a topology optimized decoupled multi-DOF positioning stage. The design problem considers motion transmission, parasitic motion, as well as input and output coupling.

load-based scenario $\mathcal{S}_{f,i}$ and $\mathcal{K}_p[\mathbf{x}]$ the motion-based characteristic stiffness of the imposed free MP (right figure of Figure 3.8).

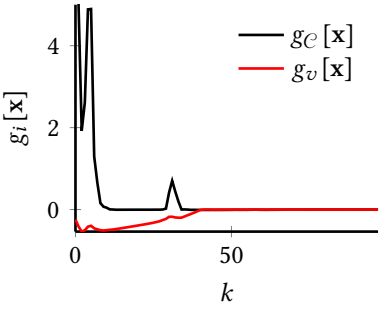
A topology optimized solution of Equation (3.55) for a desired motion transmission of $\frac{v_o}{v_i} = -1.1$ is displayed in Section 6.5.1. The left figure (Figure 3.9a) displays the filtered design variable field of the optimized solution. The right figure (Figure 3.9b) displays the undeformed and deformed configurations of a post-processed (thresholded at $\tilde{\mathbf{x}}^* = 0.4$) version of the solution; colorschemes as per previous numerical examples.

The deformations of the optimized topology for the use case closely coincides with the free MP; the difference between the green and black is barely visible. This particular design has an input and output coupling of less than 2 % and a maximum relative parasitic motion in x and θ of less than 1 %. Figure 3.10e displays the mean and standard deviation of parasitic motion throughout the optimization process and Table 3.5 gives the values of the optimized design. Thus, the error on coupling and parasitic motions is relatively negligible (for most practical applications). The motion transmission accuracy is 87.8 %, much in line with the findings in Section 3.3.2. Note that the suppression of parasitic effects is typically much better achieved compared to the transmission accuracy.

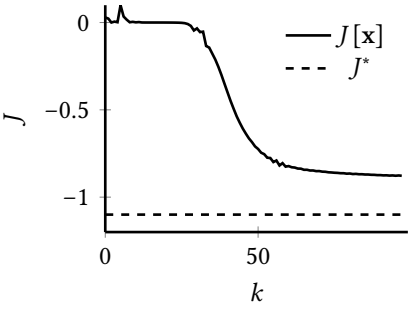
See Figure 3.10 for the properties of the optimization and design as a function of the design iteration number. The optimization process can, similarly to Section 3.3.2, roughly be divided in two phases. In the first phase a feasible solution is sought, typically taking 10–30 design iterations. In this phase the design changes are large and constraints are violated. The constraints rapidly decrease in value, typically accompanied with an increase



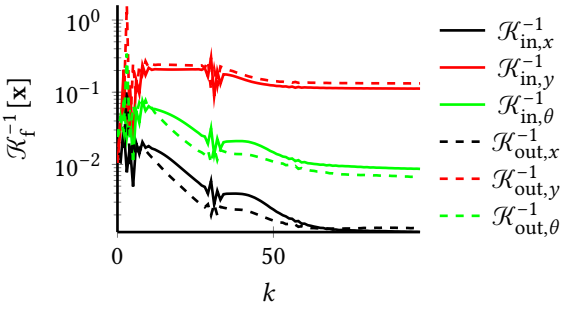
(a) Objective scaled relative to objective value at first design iteration.



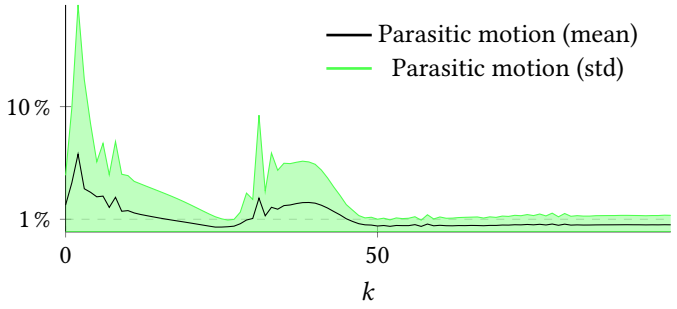
(b) Constraint values.



(c) Desired and true motion transmission.



(d) Load-based characteristic compliances. Considering symmetry, six out of twelve unique values are displayed.



(e) Mean and standard deviation of the absolution parasitic motion, normalized with respect to the input motion. After approximately 50 design iterations the mean parasitic motion passes the 1 % threshold.

Figure 3.10: Optimization properties (objective, constraints) and mechanism properties (motion, and stiffness) as a function of design iteration number.

Table 3.5: Absolute transmission accuracy of all twelve GDI for a unit y-motion at input 1 (in %) compared to free MP. The highlighted cells are (from left to right) the maximum parasitic motion, input coupling, motion transmission accuracy and output coupling.

	Input 1	Input 2	Output 1	Output 2
y	100	98	87.8	98.2
x	99.8	99.99	99.8	99.96
θ	99.3	99.98	99.7	99.9

of the objective, see Figure 3.10a. Starting with the feasible solution of phase one, in the second phase the objective is gradually improved without constraint violations, typically taking around 50–100 design iterations. In this phase the design takes clear shape (towards a black-and-white design), motion transmission tends towards the desired value, parasitic motions are reduced and stiffness in undesired directions increases, see Figures 3.10c and 3.10d.

Albeit a full interpretation of the design and working principle is out of the scope of the present work, we will hereafter provide some observations in the context of the problem formulation. All input and output blocks are attached to the ground with two leaf springs in parallel that act as a (quasi-)linear guidance.[18] These are beneficial for minimization of the objective function without storing much energy in the free MPs (involves only y-motion of the ports), see Figure 3.10d. The connectivity and location of bodies and hinges connecting the four blocks is responsible for the kinematic relationship that is very close to the imposed free MPs as is clear from Figure 3.9b.

3.3.5 Design of a deformable mirror

Apart from translations and rotations of rigid bodies, precise and smooth deformation of surfaces are of widespread interest. Such morphing structures are capable of changing shape across multiple operating or environmental conditions to (i) improve the performance, or (ii) enable new functionalities (Kota *et al.* 2001). Such shape-morphing structures are of great interest to aircraft and rotorcraft applications in the form of variable camber morphing wings, and to space applications in the form of deformable mirrors.

In this example, we demonstrate application of the proposed problem formulation to the design of a multi-DOF deformable mirror with two decoupled actuation modes. Consider the design problem and use cases sketched in Figure 3.11, following the drawing convention of Table 3.1. The problem involves two mechanism DOFs that include the motion of two input ports and the out of plane motion of the mirror surface (top layer). The aim of this optimization design problem is to find a solution that allows for decoupled deformation of the mirror surface in a sine shape, respectively cosine shape by two decoupled input ports (translation in y direction of the input ports).

The GDIs involved in this problem are the two structural DOFs describing the input motion of the two ports hereafter denoted $u_{y,1}$ and $u_{y,2}$, *and* all the structural DOFs used to describe the vertical motion of the surface. Considering the large number of structural

[18] Such compliant guidance systems are known for their decoupling property for very short range-of-motions. One may directly observe a cosine-like non-linear error will be present for larger range-of-motions.

DOFs describing the surface deformation, we opt here for computational efficiency to not reduce the dimensionality of the system a priori to application of the formulation. Also, there is no need to introduce additional (generalized) DOFs. Application of the proposed problem formulation now reads

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] : \ \alpha\left(\mathcal{K}_{\mathrm{f},1}^{-1}+\mathcal{K}_{\mathrm{f},2}^{-1}\right)+(1-\alpha)\mathcal{K}_{\mathrm{f,top}}^{-1} \\ \text{s.t.} \quad & g_{\mathcal{K},1}[\mathbf{x}] : \ \frac{\mathcal{K}_{\mathrm{p},1}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}}-1\leq 0 \\ & g_{\mathcal{K},1}[\mathbf{x}] : \ \frac{\mathcal{K}_{\mathrm{p},2}[\mathbf{x}]}{\overline{\mathcal{K}}_{\mathrm{p}}}-1\leq 0 \\ & g_{v}[\mathbf{x}] : \ \frac{v[\mathbf{x}]}{\overline{v}}-1\leq 0 \end{aligned} \tag{3.56}$$

with $\mathcal{K}_{\mathrm{f},1}$ and $\mathcal{K}_{\mathrm{f},2}$ the load-based characteristic stiffnesses of the input ports, $\mathcal{K}_{\mathrm{f,top}}$ the load-based characteristic stiffness of the mirror surface, α a coefficient to control the relative influence of these stiffnesses and the $\mathcal{K}_{\mathrm{p},1}$ and $\mathcal{K}_{\mathrm{p},2}$ the motion-based characteristic stiffnesses of the first, respectively second, free MP. This design problems thus involves three load-based scenarios and two motion-based scenarios. The load-based scenarios are used to measure the characteristic stiffness of the two input motions (described by basis load LPs) and the characteristic stiffness of the surface is described by a distributed load LP on the mirror surface (Figure 3.12 right). The motion-based scenarios involve the two MPs corresponding to the desired decoupled use cases visualized in Figure 3.12 (left and middle); the first MP consists of a prescribed cosine at the surface and nonzero motion at the first input port, the second MP consists of a prescribed sine (1.5 period) and nonzero motion at the second output port. Modification of the MPs allows for generation of a transmission between the input ports and the magnitude of the shape change.

Figure 3.13 displays a topology optimized solution (filtered design variable field) of Equation (3.56) with a target maximum magnitude of the shape-change of 0.1, and corresponding deformed configurations for both use case scenarios. Note that both input ports and mirror surface are firmly connected to the ground, while retaining the desired functionality. In line with previous numerical examples, the undeformed configuration is displayed in gray. The deformed configuration are a result of the imposed MPs (green) and use cases (black). Note how extremely well the black and green deformed structures overlap (the green deformed structure is barely visible). Especially the lack of output coupling is remarkable, demonstrating the performance of the proposed formulation for shape-morphing applications. The performance of this solution is further highlighted in Figure 3.14; Figure 3.14a displays the vertical surface motion as a function of location on the surface for the imposed MPs (solid lines) and for corresponding use case (dashed lines) for a unit input motion. Corresponding absolute error as a function of location on the surface is provided in Figure 3.14b. The maximum absolute error normalized with respect to the desired output motion magnitude is less than 0.15 % located at the centre of the mirror surface.

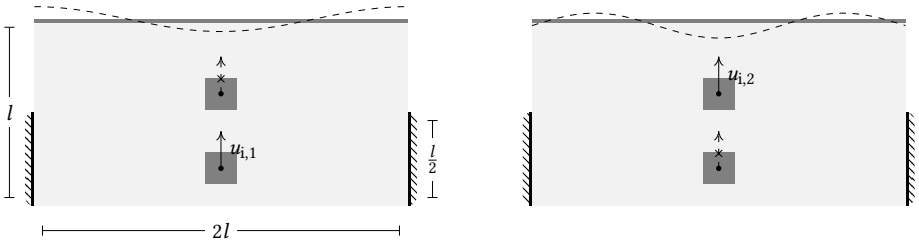


Figure 3.11: Problem setting and motion-based use case scenarios for the design of a multi-DOF decoupled deformable mirror. The design domain is of size $2l \times l$. The non-design domain at the top represents the mirror surface. The bottom-half of both sides are fully clamped. The problem involves the two generalized DOFs of interest related to the vertical motion of the input ports and all structural DOF describing the out of plane motion of the mirror surface. Considering horizontal symmetry of the design domain and loading conditions only half of the domain is modelled; thus the resulting design is expected to be symmetric.

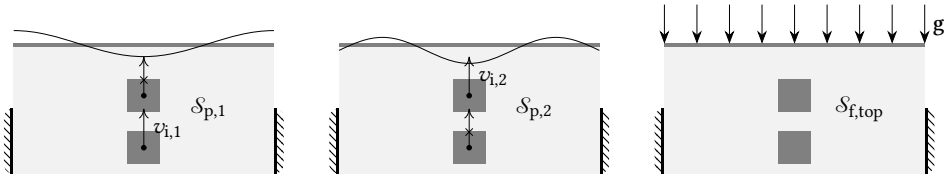


Figure 3.12: Two motion-based scenarios (left, middle) and one of the load-based scenarios (right) used in the design of the multi-DOF decoupled deformable mirror. This load-based scenario is used to measure the characteristic stiffness of the mirror surface. Note that, as opposed to what is proposed in Section 3.2.3, the distributed loading is not a basis load pattern. The other two load-based scenarios (basis load patterns on the input ports) are trivial and not shown.

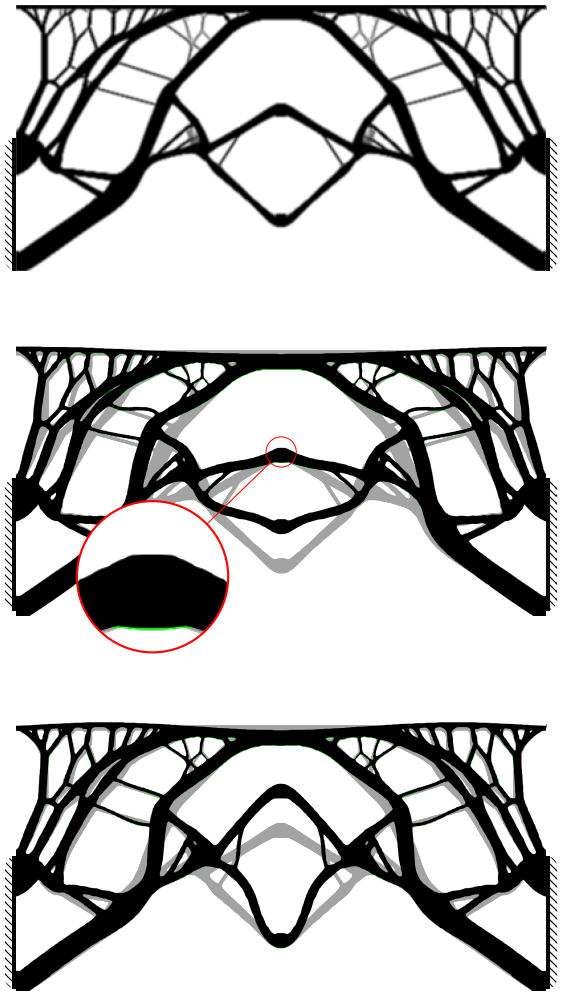
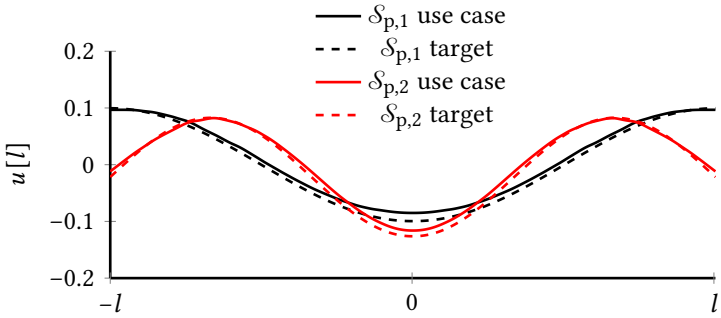
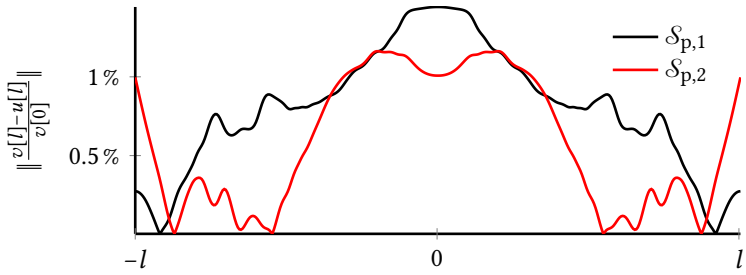


Figure 3.13: Topology optimized solution (filtered design variable field) and deformed configurations for the two motion-based use case scenarios and enforced MPs. Those MPs (displayed in green) are barely visible due to the high transmission accuracy and decoupled performance. Note that the deformations are scaled for visualization purposes.



(a) Absolute out of plane motion of the mirror surface for a variety of scenarios. Dashed lines are the target motions as prescribed in the motion-based scenarios. The solid lines indicate the displacement due to the imposed unit input motions, as per use case scenarios.



(b) Error of the out of plane motion of the mirror surface between the use case and imposed motion patterns relative to the maximum target out of plane motion. Note the error can roughly be halved by compensation of input motion.

Figure 3.14: Displacement of the mirror surface for different scenarios and corresponding relative error as a function of the position on the surface. The middle of the surface is taken as reference.

3.4 Discussion

Clear advantages of the proposed formulation can be identified, namely in terms of (i) versatility of application, (ii) simplicity of implementation, use and solving the resulting optimization problem, (iii) computational efficiency, and (iv) effectiveness in suppression of parasitic motion effects. However, the proposed formulation is not without limitations. These are discussed below, also to provide inspiration for its further development.

Firstly, the motion transmission accuracy is by definition limited. As observed by the parametric study in Section 3.3.2, the accuracy depends on both the imposed free MP and constraint value(s), see Table 3.3. For many practically relevant cases however, the lack of transmission accuracy is of much less relevance compared to the suppression of parasitic motion; one may trivially compensate for the transmission mismatch by slightly altering the magnitude of input motion. In cases where a very high transmission accuracy is required, one may opt for using an adaptive formulation, such as demonstrated in Section 3.3.2.

Furthermore, parasitic motions at the GDIs are indirectly controlled; one cannot ensure exact values of these motions prior to the optimization process. In practice, maximum parasitic and coupling motion of less than 10 % is typically perceived sufficient. Although impressive decoupling is demonstrated, see the results of the numerical examples in Sections 3.3.4 and 3.3.5, explicit constraints can (and should) be added if direct control of individual motions is desired.

The simplicity of the proposed formulation furthermore comes with loss of explicit control of individual GDI stiffness. Control of the individual stiffness can (to some extent) be regained by (i) constant or adaptive modification of the scaling coefficients of the multi-objective function, *e.g.* similar to the proposed adaptive formulation (ii) use of a smooth minimum function (Ma *et al.* 1994), or (iii) application of explicit constraint(s). While there is little doubt that the relevant constraints can be added, an open question is whether the favourable convergence behaviour of the proposed formulation can be preserved. Instead of overextending this paper, the authors encourage future work on modification of the formulation to increase the motion transmission accuracy, and have direct control over the kinematics and stiffness of GDIs.

The energy-derived stiffness measures used in the present work are closely related to the eigenvalues of the exact reduced-order stiffness matrix, as explained and used in Hasse *et al.* (2009), Li *et al.* (2019), and Wang (2009a); eigenmodes with low eigenvalues require little energy to reach a given amplitude level, and vice versa for high eigenvalues (Hasse *et al.* 2009). Despite the implicit relationship between the characteristic stiffness of MPs and the eigenvalues of the reduced stiffness matrix, working with characteristic stiffness (as opposed to eigenvalues) poses advantages in terms of the simplicity of understanding, use and implementation, and the properties of the resulting optimization problem.

The problem formulation as proposed leaves room for modifications to fit the design problem at hand. Reconsider for example the use of the load-based scenario with distributed LP, as opposed to the suggested basis LP, in the design of the deformable mirror in Section 3.3.5 to reduce the computational effort. Another example of such modification is the use of motion-based scenarios in the objective as used in the design of flexures, see Koppen *et al.* (2021b). Note that the numerical examples displayed in Sections 3.3.2 and 3.3.3 can equivalently be solved applying this flexure formulation.

Practical use of topology optimized CMs in high-tech applications requires analysis of

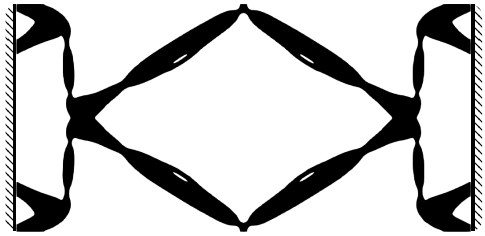


Figure 3.15: Topology optimized compliant mechanism inverter designed using the proposed problem formulation Equation (3.22) with an explicit aggregated constraint on the maximum Von Mises stress. The constraint is imposed on the stress field of the motion-based scenario. Implementation of the stress constraint is similar to Koppen *et al.* (2021b) as per Verbart *et al.* (2017).

and design for (at least) strength, dynamic properties, manufacturability and robustness of these measures with respect to environmental changes, uncertainties and manufacturing tolerances. As far as the authors have experienced, the proposed problem formulation does not pose any obstruction for such additional design requirements. The proposed 'reduced' robust optimization problem formulation allows for effective control of manufacturability and robustness with negligible addition computational effort, see Section 3.3.2 and the optimized solution in Figure 3.5. Design requirements on strength may be incorporated by addition of explicit constraints on the maximum Von Mises stress as described in more detail in the related work of Koppen *et al.* (2021b) (Section 5 paragraph 'Stress-based design'). Examples of optimized solutions of the single-DOF mechanism inverter with explicit constraints on the maximum Von Mises stress and minimum fundamental eigenfrequency are displayed in Figures 3.15 and 3.16, respectively.

As presented here, the optimization problem formulation is intended for, and limited to, the design of short-stroke CMs, such that the assumptions of linear strain-displacement and stress-strain relationships are valid. The optimized solutions to the numerical examples consist of features that are known to induce parasitic motions for larger ranges of motion, such as the double leaf flexures for short-stroke linear guiding in the solution of Section 3.3.4, see Section 6.5.1. The presence of structural nonlinearities is by no means controlled in the design optimization process; the accuracy of motion, magnitude of parasitic effects and variation of stiffness for larger ranges of motion is highly dependent on the design problem and optimized solution, and can thus only be guaranteed using a geometrically nonlinear post-analysis. Using the proposed optimization problem formulation to design CMs with large deflections and rotations, to guarantee the intended functionality in a larger range of motion, offers a promising direction for future research.

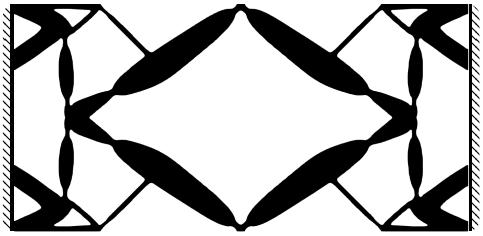


Figure 3.16: Topology optimized compliant mechanism inverter designed using the proposed problem formulation Equation (3.22) with an explicit constraint on the minimum fundamental eigenfrequency. The constraint is imposed on the load-based scenario conditions, that is both GDIs are unconstrained. The eigenfrequency constraint is implemented as per Ma *et al.* (1994) considering the lowest three eigenfrequencies.

3.5 Conclusions

The design of short-stroke multi-degree of freedom compliant mechanisms is characterized by a large number of loading scenarios, numerous stringent design requirements and inherent coupling issues. The development of effective topology optimization problem formulations for the synthesis of multi-degree of freedom compliant mechanisms faces the challenge of general applicability while keeping the formulation comprehensible, easy and efficient to implement, and use.

This work proposes a kinetoelastic energy-based topology optimization method for the synthesis of short-stroke decoupled multi-degree of freedom compliant mechanisms. At the core of the proposed formulation lie the concepts of generalized degrees of freedom as well as motion and load patterns and their characteristic stiffness. These concepts allow for a natural description of the compliant mechanism design problem, while drastically reducing the amount and strictness of constraint functions. The formulation is found to require only *a single* constraint per mechanism DOF, as opposed to numerous independent stringent and often highly nonlinear displacement-based constraints. From these properties we conclude that this unique composition of objectives and constraints results in favourable optimization convergence and computational efficiency. The proposed formulation can be seen as an extension of the classical compliance minimization problem formulation and can be written in terms of compliances, making the formulation compatible with commercial software. Furthermore, the responses are likewise self-adjoint, consequently reducing the computational effort required to obtain design sensitivities to a minimum. These favourable properties come at the cost of relaxing the desired kinematics, which may nevertheless prove acceptable in many practical use cases. If crucial, additional measures can enforce kinematics strictly.

From the wide range of numerical examples we conclude that the formulation is easy to implement in both academic code and commerical software, easy to use requiring a minimum of independent parameters, and effective in solving a large variety of compliant mechanism design optimization problems, next to multi-degree of freedom compliant mechanisms also including flexures and shape-morphing structures. The optimized solutions are found to have good decoupling properties and effectively suppress parasitic effects, even for problems with complex desired kinematics.

This work in addition provides a new symbolic language to describe compliant mechanism problem settings for numerical examples and use cases, and introduces a set of single and multi-degree of freedom benchmark problems. Freely available codes and implementation in commercial software stimulate replication of results and use of the formulation in academia and industry. Future work aims to explore the potential of the formulation for use in the design of compliant mechanisms exhibiting larger deflections and rotations.

3.A Compliance under prescribed motion: a one-DOF example

The following is much in line with the appendix of the work of Pedersen *et al.* (2011), but rewritten here to fit the context, nomenclature and use of symbols.

Consider a linear design dependent spring (or a condensed structure) with stiffness $k[\mathbf{x}]$ (here $\mathbf{x}$ are the design variables), load f and motion u. Upon equilibrium it holds that

$$f = k[\mathbf{x}]\, u \quad \longleftrightarrow \quad u = \frac{f}{k[\mathbf{x}]}. \tag{3.57}$$

Upon application of a design independent load $f = f_0$, the compliance

$$\mathcal{C}[\mathbf{x}]|_{f=f_0} := f \cdot u[\mathbf{x}] = f_0 \cdot \frac{f_0}{k[\mathbf{x}]} \propto \frac{1}{k[\mathbf{x}]}, \tag{3.58}$$

and consequently

$$\max_{\mathbf{x}} \; k[\mathbf{x}]|_{f=f_0} \quad \longrightarrow \quad \min_{\mathbf{x}} \; \mathcal{C}[\mathbf{x}]|_{f=f_0} \tag{3.59}$$

However, upon application of a design independent motion $u = u_0$, the compliance

$$\mathcal{C}[\mathbf{x}]|_{u=u_0} := f[\mathbf{x}] \cdot u = k[\mathbf{x}] \cdot u_0^2 \propto k[\mathbf{x}], \tag{3.60}$$

and

$$\max_{\mathbf{x}} \; k[\mathbf{x}]|_{u=u_0} \quad \longrightarrow \quad \max_{\mathbf{x}} \; \mathcal{C}[\mathbf{x}]|_{u=u_0}. \tag{3.61}$$

As a result, the compliance $\mathcal{C}|_{u=u_0}$ is proportional to the stiffness, whereas the compliance $\mathcal{C}|_{f=f_0}$ is inversely proportional to the stiffness. The compliance is also defined as the inverse of stiffness, as commonly used in mechanism design. Thus, under prescribed motion conditions, minimizing compliance (as defined by the mechanism design community, *i.e.* $c := \frac{1}{k}$) requires maximization of compliance (as defined by the structural optimization community as $\mathcal{C} := \mathbf{f} \cdot \mathbf{u}$).

3.B Properties of the (exact reduced-order) mechanism model

Several metrics can be deducted from the generalized system of equations of interest. First of all, the entries of the generalized reduced stiffness matrix $\tilde{\mathbf{K}}$ and its inverse $\tilde{\mathbf{C}}$, the 'mechanism compliance matrix', that is

$$\tilde{\mathbf{C}} := \tilde{\mathbf{K}}^{-1}. \tag{3.62}$$

Note that pre-multiplication of both sides of Equation (3.11) with $\tilde{\mathbf{K}}^{-1}$ yields

$$\tilde{\mathbf{C}}\hat{\mathbf{f}} = \hat{\mathbf{u}}. \tag{3.63}$$

The GDIs can be partitioned in generalized input (i) and output (o) DOFs (of interest), which yields

$$\begin{bmatrix} \tilde{\mathbf{K}}_{ii} & \tilde{\mathbf{K}}_{io} \\ \tilde{\mathbf{K}}_{oi} & \tilde{\mathbf{K}}_{oo} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_i \\ \hat{\mathbf{u}}_o \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_i \\ \hat{\mathbf{f}}_o \end{bmatrix}. \tag{3.64}$$

3

Consider unconstrained output GDIs $\hat{\mathbf{u}}_o$, then $\hat{\mathbf{f}}_o = \mathbf{0}$ and we may write

$$\hat{\mathbf{u}}_o = \mathbf{J}\hat{\mathbf{u}}_i, \tag{3.65}$$

with the mechanisms 'motion transmission matrix' $\mathbf{J} \in \mathbb{R}^{o \times i}$ defined as

$$\mathbf{J} = -\tilde{\mathbf{K}}_{oo}^{-1}\tilde{\mathbf{K}}_{io}^{\mathsf{T}}. \tag{3.66}$$

Now consider the constrained generalized mechanism output DOFs $\hat{\mathbf{u}}_o = 0$, then we may write

$$\hat{\mathbf{f}}_o = \mathbf{Y}\hat{\mathbf{f}}_i, \tag{3.67}$$

with the mechanisms 'load transmission matrix' $\mathbf{Y} \in \mathbb{R}^{o \times i}$ defined as

$$\mathbf{Y}^{\mathsf{T}} = \tilde{\mathbf{K}}_{ii}^{-1}\tilde{\mathbf{K}}_{io}. \tag{3.68}$$

Finally, we get a relationship between $\hat{\mathbf{u}}_o$ and $\hat{\mathbf{f}}_i$ (and $\hat{\mathbf{f}}_o$) that yields

$$\hat{\mathbf{u}}_o = \mathbf{W}^{-1}\left(\hat{\mathbf{f}}_o - \mathbf{Y}\hat{\mathbf{f}}_i\right), \tag{3.69}$$

with

$$\mathbf{W} = \tilde{\mathbf{K}}_{oo} - \mathbf{Y}\tilde{\mathbf{K}}_{io}. \tag{3.70}$$

Note that this last step is basically condensation of the input GDIs. For an in-depth analysis of the use of condensation in the context of topology optimization the reader is referred to Koppen *et al.* (2022c).

4

Local redesign for additive manufacturability

Design for additive manufacturing constraints tend to drastically decrease the performance of topology optimized compliant mechanisms. It is crucial to impose these constraints on the flexure regions, while in others part of the compliant mechanism design, these constraints can be relaxed. This chapter proposes a method to redesign selected refined local domains with design for additive manufacturing constraints, whilst simultaneously considering the global mechanism performance.

This chapter is based on conference paper:
Koppen, S., Hoes, E., Langelaar, M., and Frecker, M. I. (2021). *Local redesign for additive manufacturability of compliant mechanisms using topology optimization.* IDETC and CIE (Vol. 85444). ASME.

Local redesign for additive manufacturability of compliant mechanisms using topology optimization

Abstract *Compliant mechanisms are crucial components in current and future high-precision applications. Topology optimization and additive manufacturing offer freedom to design complex compliant mechanisms that were impossible to realize using conventional manufacturing. Design for additive manufacturing constraints, such as the maximum overhang angle and minimum feature size, tend to drastically decrease the performance of topology optimized compliant mechanisms. It is observed that, among others, design for additive manufacturing constraints are only dominant in the flexure regions. Flexures are most sensitive to manufacturing errors, experience the highest stress levels and removal of support material carries the highest risk of failure. It is crucial to impose these constraints on the flexure regions, while in others part of the compliant mechanism design, these constraints can be relaxed. We propose to first design the global compliant mechanism layout in the full domain without imposing any design for additive manufacturing constraints. Subsequently we redesign selected refined local redesign domains with design for additive manufacturing constraints, whilst simultaneously considering the mechanism performance. The method is applied to a single-input-multi-output compliant mechanism case study, limiting the maximum overhang angle, introducing manufacturing robustness and limiting the maximum stress levels of a selected refined redesign domain. The high resolution local redesigns are detailed and accurate, without a large additional computational effort or decrease in mechanism performance. Thereto, the method proves widely applicable, computationally efficient and effective in its purpose.*

4.1 Introduction

Compliant mechanisms (CMs) achieve force, motion or energy transmission through elastic deformation. Due to their intrinsic high repeatability CMs are crucial components in current and future high-precision applications, among others (Herder 2017). In addition to their transmission function, common CM design requirements include range of motion, volume and/or mass, cross-talk and parasitic motion, stress levels and fatigue life as well as the sensitivity of those factors to, *e.g.*, manufacturing errors (Howell 2001).

CMs are traditionally synthesized by kinematic or building block approaches (Danun *et al.* 2021; Gallego *et al.* 2009). The use of Topology Optimization (TO) techniques to design CMs has recently gained increasing interest. TO provides maximum design freedom to create optimal mechanisms satisfying application-specific conflicting requirements whilst

requiring minimal designer input regarding mechanism kinematics (Bendsøe *et al.* 2004; Cao *et al.* 2013; Deepak *et al.* 2009; Frecker *et al.* 1997; Larsen *et al.* 1997; Sigmund *et al.* 1997).

Alongside the increasing popularity of TO, Additive Manufacturing (AM) shows promise as the go-to technology for realization of CMs—particularly that of complex topology optimized CMs (Khurana *et al.* 2018; Langelaar 2016; Mirth 2016). The constructive freedom that AM offers allows fabrication of CM designs that were impossible to realize using conventional manufacturing. But also AM capabilities have their limits, which has resulted in the development of specific Design for AM (DfAM) methodologies (Rosen 2014).

The majority of DfAM restrictions are a consequence of the intrinsic layer-wise and temperature intensive nature of AM processes (Diegel *et al.* 2019; Gao *et al.* 2015). The most dominant DfAM constraints are the maximum overhang angle and minimum feature size (Diegel *et al.* 2019). Furthermore, the kinematic behaviour of CMs is highly sensitive to geometrical deviations arising from manufacturing errors. Additionally, limiting peak stress levels is important for any CM to prevent failure. At the design stage, this requires accurate Finite Element (FE) modeling.

Much attention has been devoted to incorporating DfAM constraints in the TO process, including overhang angle (self-supporting structures) filters/constraints (Garaigordobil *et al.* 2019; Gaynor *et al.* 2016; Langelaar 2016, 2017; Pellens *et al.* 2019; Ven *et al.* 2018, 2020) as well as support structures (Kuo *et al.* 2018; Langelaar 2018; Ven *et al.* 2001) and (combined) building orientation (Langelaar 2018) optimization. Both robustness of performance with respect to uniform AM errors and feature size control can be achieved using the robust TO formulation (Lazarov *et al.* 2011; Schevenels *et al.* 2011; Sigmund 2009; Silva *et al.* 2019; Wang *et al.* 2011). Inclusion of stress constraints in the TO process of CMs has also been reported (Assis Pereira *et al.* 2018; Leon *et al.* 2015; Saxena *et al.* 2001; Silva *et al.* 2019).

However, global DfAM constraints may drastically decrease the performance of CMs designed using TO (Langelaar 2016). In addition, accurate FE modeling requires fine meshes, which—although increasing both design space and modeling accuracy—drastically increase the required computational effort, especially when used within an iterative process such as TO.

Both from analysis and experiments it is observed that DfAM constraints—when applied to lumped CMs—are (only) dominant in the most flexible regions, hereafter simply called flexures. The kinematic behaviour of these flexures is most sensitive[1] to manufacturing errors. In addition, flexures are generally the regions experiencing high stress levels. As opposed to other regions of the design, removal of support material in flexures regions carries high risk of damaging flexures with obvious negative effects. Hence, it is crucial to ensure flexure regions are self-supporting, while in other parts of the CM design, this AM constraint can be relaxed.

To address these challenges, we propose a two-step CM TO method:

(i) first design the global CM layout in the full domain (using TO) without DfAM considerations, and subsequently

(ii) redesign selected refined local redesign domain(s) (flexures) with DfAM considerations, whilst simultaneously considering the mechanism performance.

[1] Bending stiffness scales cubically with thickness.

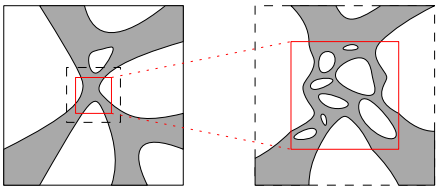


Figure 4.1: A two-step approach towards DfAM using TO.

Figure 4.1 shows a schematic representation of the proposed idea. The topology of the global design on the left is analyzed and a selection of local redesign domain(s) is made. The right schematic shows the topology of the high-resolution flexure redesign—possibly taking into account DfAM constraints. An additional advantage of the proposed method is that through employing static condensation, as detailed in the next section, the second refined design stage has low computational cost while still accounting for the performance of the entire mechanism.

In general, a global design may have multiple redesign domains that should be reconsidered simultaneously. Also, a larger redesign region provides more design flexibility, but comes at a higher computational cost. For simplicity and without loss of generality, we consider a single redesign domain throughout the remainder of the present work. We also assume that the designer manually selects the redesign region(s). As an extension these could be identified automatically using geometrical and strain-based indicators, since flexures typically exhibit higher strain levels as compared to other parts of the structure.

The manual post-processing method of Shih *et al.* (2006) allows for local redesign, however it does not facilitate the use of topology optimization nor the local application of constraints. The method proposed in this work has—to the best knowledge of the authors—not been explored yet.

The expected advantages of this approach are twofold, namely:

(i) detail and accuracy where it matters, with low additional computational effort, and

(ii) DfAM where it is required, without drastically decreasing the mechanism performance.

Hereafter the method is further outlined. This is followed by implementation details, after which we apply the method to a case study with different DfAM and performance constraints.

4.2 Method

The method consists of three subsequent phases, visualized in Figure 4.2:

(i) Generation of the global CM design—possibly via TO—without taking into account DfAM constraints. This generally is an iterative process.

(ii) Pre-processing of the redesign phase including stiffness mapping and mesh coupling.

(iii) Redesign of the local domain(s) using TO taking into account DfAM constraints. Similar to the global design phase, this is an iterative process.

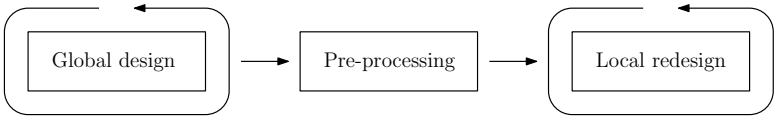


Figure 4.2: Flowchart of the three subsequent phases required to locally redesign for AM of CMs using TO.

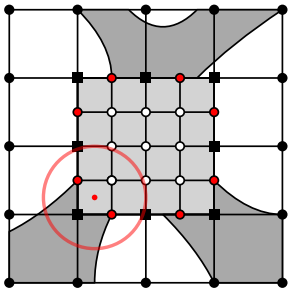


Figure 4.3: Reinitialization of redesign domain. The structural DOFs are separated in three sets, the DOFs in the redesign domain (○), non-design domain (•) and interface (■) regions. Some redesign DOFs (•) are coupled to the interface DOFs using MPCs. Note that for some design variables the filter radius (both for density and overhang filter) exceeds outside of the redesign domain, as indicated by the red circle.

Whereas the first phase relies on established state-of-the-art, both the pre-processing and local redesign phases require further explanation. In what follows, both phases are further elaborated.

4.2.1 Pre-processing

The pre-processing phase subsequently consists of

(i) definition of local redesign domain(s),

(ii) partitioning of structural Degrees of Freedom (DOFs) in three disjoint sets: non-design, interface and redesign,

(iii) mapping of the non-design domain stiffness to the interface DOFs,

(iv) re-initialization of the local redesign domain(s) (with refined mesh),

(v) coupling of non-matching meshes between the interface and redesign domain(s), and

(vi) adjustment of filters (density and/or overhang) to take into account non-design topology within filter radius.

Consider a global design as visualized on the left of Figure 4.1. After locating the local redesign domain(s), in this case the flexure, we split up the structural DOFs of the discretized partial differential equation in three sets, consecutively containing the DOFs in the redesign domain (○), non-design domain (•) and interface (■) regions, as visualized

in Figure 4.3. Some redesign DOFs (•) are coupled to the interface DOFs to ensure mesh continuity.

Since only a small part of the structure is subject to change, the computational effort can be significantly reduced if the dimensionality of the problem is reduced before the redesign phase, such that repetitive analysis of the non-design domain(s) is avoided. We propose to a priori apply exact model-order reduction in the form of static condensation to map the stiffness of the non-design domain to the interface without loss of information (Botkin *et al.* 1989; Gangadharan *et al.* 1990; Guyan 1965; Irons 1965; Yang *et al.* 1996). In addition, this highly reduces the computational effort of the sensitivity analysis (Koppen *et al.* 2022c). To this end, consider the discretized governing system of equations $\mathbf{Bu} = \mathbf{f}$, with $\mathbf{B}$, $\mathbf{u}$ and $\mathbf{f}$ the stiffness matrix, displacement field and force field of the non-design domain(s), respectively. We partition the system of equations in DOF sets of the interface (i) and non-design (n) domain, *i.e.*

$$\begin{bmatrix} \mathbf{B}_{ii} & \mathbf{B}_{in} \\ \mathbf{B}_{ni} & \mathbf{B}_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{u}_i \\ \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{f}_i \\ \mathbf{f}_n \end{bmatrix}. \tag{4.1}$$

Via the condensation process we obtain the reduced system of discretized governing equations, solely in terms of the interface DOFs, *i.e.*

$$\tilde{\mathbf{K}}\mathbf{u}_i = \mathbf{f}_i + \tilde{\mathbf{f}}, \tag{4.2}$$

with the reduced system matrix

$$\tilde{\mathbf{K}} := \mathbf{B}_{ii} - \mathbf{B}_{in}\mathbf{B}_{nn}^{-1}\mathbf{B}_{ni}, \tag{4.3}$$

and the reduced load

$$\tilde{\mathbf{f}} := -\mathbf{B}_{in}\mathbf{B}_{nn}^{-1}\mathbf{f}_n. \tag{4.4}$$

Here we assume the principal sub-matrix $\mathbf{B}_{nn}$ to be non-singular, such that it is invertible (Benscoter 1948). Note that the preconditioning of the—generally high dimensional—matrix $\mathbf{B}_{nn}$ can be reused. The reduced system matrix and load are independent of the design variables. Thereto, this—relatively expensive—process only needs to be carried out once prior to the redesign.

Note that the interface DOFs generally are the geometrical interface between non-design and redesign domain, however may also include other DOFs, for example the DOFs at input/output of a mechanism. By doing so, one does not have to recompute the displacement field of condensed DOFs via the displacement field of the interface DOFs via

$$\mathbf{u}_n = \mathbf{B}_{nn}^{-1}\left(\mathbf{f}_n - \mathbf{B}_{ni}\mathbf{u}_i\right). \tag{4.5}$$

Note that $\mathbf{B}_{nn}^{-1}\mathbf{f}_n$ has previously been calculated in Equation (4.4). In the remainder of the present work we assume all relevant DOFs are included as interface DOFs, including the DOFs with applied loads. As a result, we assume $\tilde{\mathbf{f}} = \mathbf{0}$.

After mapping of the non-design domain stiffness, the local redesign domain is remeshed with a finer mesh and reinitialized. The mesh of the redesign domain is non-conforming with respect to the interface mesh (Figure 4.3). To circumvent discontinuities in the displacement field, mesh coupling is required to enforce continuity. One can introduce

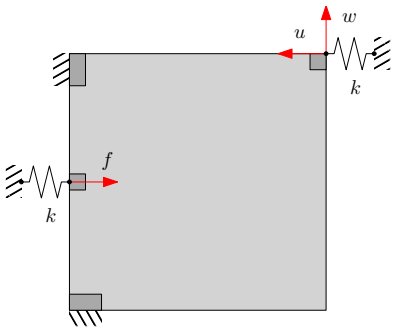
weak geometric compatibility via Multi-Point Constriant (MPC) methods such as, *e.g.*, main-secondary elimination, penalty function augmentation or Lagrange multiplier adjunction (Boer *et al.* 2007). We opt to linearly couple the interfacing redesign DOFs (•) to adjacent interface DOFs (■) via the main-secondary elimination technique. Consider the uncoupled discretized governing system of equations $\mathbf{A}[\mathbf{x}]\,\mathbf{u} = \mathbf{f}$, partitioned in DOF sets of the interface (i) and redesign (r) domain. The design-dependent stiffness matrix $\mathbf{A}[\mathbf{x}]$ has not yet been assembled with the non-design domain stiffness $\tilde{\mathbf{K}}$. A transformation matrix $\mathbf{T}$ is constructed[2], coupling the dependent redesign DOFs to the interface DOFs via

$$\mathbf{u} = \mathbf{T}\mathbf{v}, \tag{4.6}$$

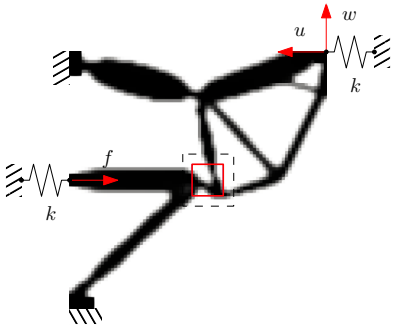
with $\mathbf{v}$ the independent main DOFs. Thereto, after pre-multiplication of the discretized system of equations by $\mathbf{T}^\mathsf{T}$, the constrained stiffness matrix $\mathbf{K}$ is simply calculated via

$$\mathbf{K}[\mathbf{x}] = \mathbf{T}^\mathsf{T}\mathbf{A}[\mathbf{x}]\mathbf{T}. \tag{4.7}$$

The mesh refinement also influences any filtering techniques used in the TO process, *e.g.* the common density filter (Bruns *et al.* 2001) or overhang filter (Langelaar 2016). As shown in Figure 4.3, if the filter radius exceeds the interface, the filter operator has to be adapted to include the non-design topology in order to ensure a smooth transition between the redesign and non-design domains. In the present work, this is simply handled by refining the density description in these boundary regions of non-design elements.



(a) Design domain and boundary conditions. The domain is discretized in 80 by 80 square quadrilateral finite elements.



(b) Topology resulting from solving Equation (4.11).

Figure 4.4: Single-input-multi-output CM design problem.

4.2.2 Local redesign

After pre-processing one can rewrite the original optimization problem formulation, in terms of the redesign variables and DOFs, *i.e.*

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & g_0[\mathbf{x}, \mathbf{v}[\mathbf{x}]] \\ \text{subject to} \quad & g_i[\mathbf{x}, \mathbf{v}[\mathbf{x}]] \leq 0, \quad i = 1, \ldots, m \\ & x_i \in \mathbb{X}, \quad\quad\quad\quad i = 1, \ldots, N \end{aligned} \tag{4.8}$$

[2]This is considered common knowledge and hence not elaborated.

with $\mathbf{x} \in \mathbb{R}^N$ the design variables, each satisfying the bound constraints, that is $\mathbb{X} := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$, $g_0[\mathbf{x}, \mathbf{v}[\mathbf{x}]]$ the objective function as used for the global CM design, now written in terms of the displacement field of the redesign and/or interface $\mathbf{v}[\mathbf{x}]$, and $g_i[\mathbf{x}, \mathbf{v}[\mathbf{x}]]$ the CM and/or DfAM constraints imposed on the redesign domain.

To solve the optimization problem in Equation (4.8), we follow—with the exception of some small adaptations—the standard procedure for "density based sequential approximate TO" (Bendsøe *et al.* 2004). This entails iteratively performing the following steps until convergence:

(i) Density filter (and possibly AM and/or Heaviside filter)

(ii) Material interpolation

(iii) Redesign stiffness matrix assembly

(iv) Coupling of redesign and interface meshes using MPCs

(v) Assembly of redesign and non-design domain stiffnesses

(vi) Finite element analysis

(vii) Response evaluation and sensitivity analysis

(viii) Optimization of convexified subproblem

The adjustment of filter and transformation operators have been explained in the pre-processing phase. The largest deviation from the standard process is step (v): assembly of redesign and non-design domain stiffnesses.

In order to incorporate the full structural behaviour into the redesign process, the reduced system matrix $\tilde{\mathbf{K}}$ is combined with the constrained stiffness matrix $\mathbf{K}[\mathbf{x}]$ and load properties of the redesign domain via

$$\begin{bmatrix} \mathbf{K}_{\mathrm{rr}}[\mathbf{x}] & \mathbf{K}_{\mathrm{ri}}[\mathbf{x}] \\ \mathbf{K}_{\mathrm{ir}}[\mathbf{x}] & \mathbf{K}_{\mathrm{ii}}[\mathbf{x}] + \tilde{\mathbf{K}} \end{bmatrix} \mathbf{v} = \mathbf{T}^{\top} \begin{bmatrix} \mathbf{f}_{\mathrm{r}} \\ \mathbf{f}_{\mathrm{i}} \end{bmatrix}. \tag{4.9}$$

The low dimensional discretized governing equations of Equation (4.9) can now be used to analyze the structural response at a highly reduced computational effort. As such, it is relatively cost effective to carry out redesign at specifically chosen local domains. Since the reduced stiffness matrix is design-independent, it does not influence the sensitivity analysis, which is omitted here for brevity.

4.3 Implementation

Independent of the problem formulation as presented, the user has to consider, select, and implement a variety of methods to effectively use the formulation in a topology optimization setting. Without loss of generality, the case study employs the implementation choices described here.

For the finite element analysis, we opt for standard 4-node quadrilateral (2D) elements in structured meshes. The domain is parametrized by assignment of a design variable

$x_i \in \mathbb{X}$ to each finite element i, which allows for local control of the material properties (Bendsøe *et al.* 2004).

It is generally recognized that both final topology and performance are sensitive to the initial design. This is especially the case for CM TO. We consider this influence out of the scope of this paper and thereto opt for a homogeneous initial design, both for the global design and local redesign. The volume fraction of the local redesign initial design is set equal to the volume fraction of the redesign domain of the global design.

To eliminate modeling artifacts, the design variable field $\mathbf{x}$ is generally blurred as to obtain the filtered field $\tilde{\mathbf{x}} \in \mathbb{R}^N$ using a linear filtering operation with radius $r \in \mathbb{R}^+$, see *e.g.* (Bruns *et al.* 2001). The modification of filter operator has further been explained under pre-processing.

The Young's modulus of an element is related to the filtered design variable $\tilde{x}_i$ via an element-wise composite rule. We apply the commonly used modified solid isotropic material with penalization interpolation function (Sigmund *et al.* 1997), that is

$$E_i[\tilde{x}_i] = \underline{E} + \left(\overline{E} - \underline{E}\right) \tilde{x}_i^p, \tag{4.10}$$

with $\underline{E}$ and $\overline{E}$ the Young's moduli of void and solid and $p \in \mathbb{R}^+$ a user definable parameter. It is known that this interpolation function stimulates a 0/1 solution of a compliance-based optimization problem with volume constraint.

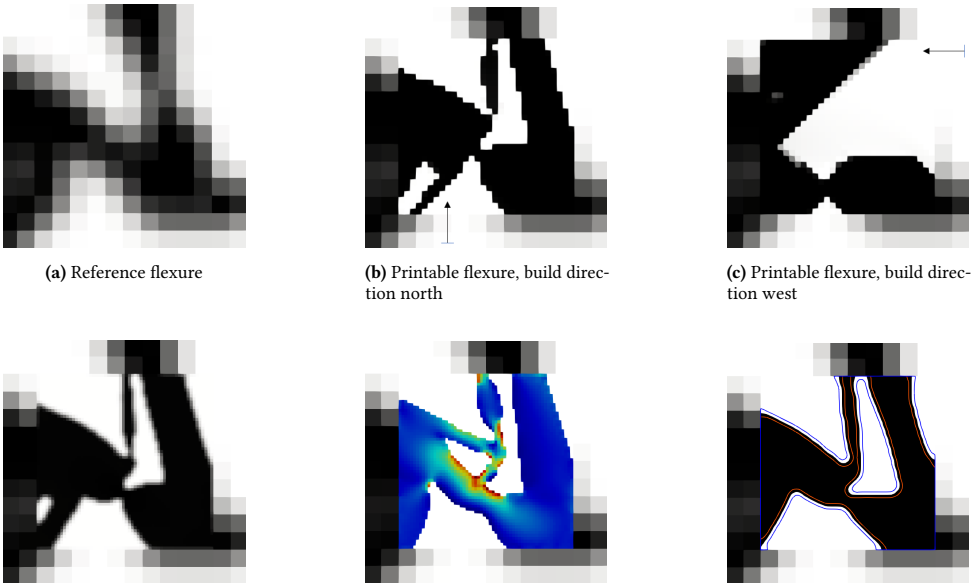
The gradient-based inequality-constrained nonlinear optimization problem in Equation (4.8) is solved in a nested analysis and design setting. The design variables are iteratively updated by a sequential approximate optimization scheme, as is common in the topology optimization field. We use the method of moving asymptotes by Svanberg (1987), including the parameter settings provided therein. The resulting convex sub-problems are solved using a primal-dual interior point method. The optimization is terminated when the maximum design change is smaller than a termination value.

4.4 Case study

To illustrate the effectiveness and versatility of the proposed method, it is applied to a case study, based on the single-input-multi-output CM design example used in (Bendsøe *et al.* 2004). Consider the global design problem posed in Figure 4.4a and corresponding optimization problem formulation

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & g_0[\mathbf{v}[\mathbf{x}]] := u[\mathbf{x}] \\ \text{subject to} \quad & \frac{w^2[\mathbf{x}]}{u^2[\mathbf{x}]} \leq \overline{\epsilon} \\ & V[\mathbf{x}] \leq \overline{V} \\ & x_i \in \mathbb{X}, \quad i = 1, ..., N \end{aligned} \tag{4.11}$$

with $u[\mathbf{x}]$ the output displacement, $w[\mathbf{x}]$ the parasitic motion, $\overline{\epsilon}$ the maximum allowable relative displacement, $V[\mathbf{x}]$ the volume fraction of the redesign and $\overline{V}$ the maximum allowable volume fraction (set to 0.2 for the global design problem). The objective is to maximize the output displacement u, while constraining the parasitic motion w both due to an imposed unit load f at the input. The topology of the global design obtained by



(a) Reference flexure

(b) Printable flexure, build direction north

(c) Printable flexure, build direction west

(d) Refined flexure redesign

(e) Stress constrained flexure redesign

(f) Robust flexure redesign

Figure 4.5: A variety of refined redesigned flexures with varying refinement ratios and DfAM constraints.

solving Equation (4.11), with parameters set as given in Table 4.1 is shown in Figure 4.4b. The volume fraction of the redesign domain is $\tilde{V}$. The remaining redesigns are subjected to $\overline{V} = \tilde{V}$ for fair comparison. The applied refinement factors for each redesign case are listed in Table 4.2.

The reference redesign, that is the redesigned flexure without additional refinement or constraints, is shown in Figure 4.5a and we define its reference performance as $g_0 = \tilde{g} = 1$ (normalized output displacement). The refined (6 times) redesign is shown in Figure 4.5d and has performance $g_0 = 1.03\tilde{g}$. The increased design space allows for minor improved performance.

In the following we will show the possibilities of introducing additional DfAM or performance constraints: limiting the maximum allowable overhang angle, introducing manufacturing robustness, or limiting the maximum stress levels in the formulation as posed in Equation (4.11), without providing a thorough investigation of design parameters.

4.4.1 Self-supporting flexure

This variation includes a simplified AM fabrication model to exclude unprintable geometries from the design space, resulting in fully self-supporting (maximum 45°overhang angle) optimized designs (Langelaar 2016). The problem formulation as posed in Equation (4.11) is unaltered. However, the TO process includes an additional filtering step between the density filter and material interpolation. The resulting topologies corresponding to a north and west building direction are shown in Figures 4.5b and 4.5c, respectively. A redesign with north building direction can easily be obtained without any loss of performance

($g_0 = 1.00\tilde{g}$), see the results summarized in Table 4.2. However, the topology of the west building direction highly deviates from the reference and a loss of performance is observed ($g_0 = 0.87\tilde{g}$).

Whilst these redesigned domains are self-supporting, the remainder of the design is not (per se). When including the overhang filter on the full design domain, no feasible solutions are found for any of the four building directions (south, east, north, west). These solutions do not satisfy the constraints and/or the design variables take on intermediate values.

4.4.2 Stress-optimized flexure

In order to limit the maximum stress for a given range of motion, or similarly extend the range of motion for a given maximum stress, one can extend the problem formulation in Equation (4.11) with stress constraints, *i.e.*

$$g_\sigma := \max(\boldsymbol{\sigma}[\mathbf{x}]) \leq \overline{\sigma}, \tag{4.12}$$

with $\boldsymbol{\sigma}$ the stress field, for example using the Von Mises stress criterion, and $\overline{\sigma}$ the maximum allowable stress. Many different formulations of g_σ are available, see *e.g.* (Assis Pereira *et al.* 2018; Leon *et al.* 2015; Silva *et al.* 2019) for CM specific formulations. Without loss of generality, we use the unified aggregation and relaxation approach as proposed by Verbart et al. (Verbart *et al.* 2017).

Figure 4.5e depicts the stress constrained redesign. Here $\overline{\sigma}$ is set to 0.3 times the maximum stress of the topology as obtained in Figure 4.5d. Despite the strict constraint, the topology has a similar performance to the refined redesign; see Table 4.2.

If the problem formulation includes stress constraints on the full design domain, a performance reduction of at least 15% is observed.

4.4.3 Robust flexure

The desired kinematics and stress field of a flexure is sensitive to geometric deviations. However, in classical deterministic topology optimization, the effect of such uncertain parameters on the performance of the structure is not taken into account. This may lead to a design that is very sensitive to manufacturing errors. As a consequence, the performance of the actual structure may be far from optimal.

The robust approach to topology optimization (Lazarov *et al.* 2011; Sigmund 2009; Silva *et al.* 2020; Wang *et al.* 2011) takes into account uniform manufacturing errors. Uniform erosion and dilation effects, from here on denoted by superscripts (e) and (d), are simulated by means of a projection method: the filtering of the design variable field is followed by a differentiable Heaviside projection using a high projection threshold η^{e} to simulate an erosion and a low projection threshold η^{d} to simulate a dilation. For further details on the robust formulation the reader is referred to (Wang *et al.* 2011).

Considering the robustness one can reformulate the problem formulation in Equa-

Table 4.1: Case study constant parameters

Parameter	Value	Description
$\overline{E}$	$1.0\,\mathrm{Pa}$	solid Young's modulus
$\underline{E}$	$1.0\times10^{-9}\,\mathrm{Pa}$	void Young's modulus
r	2.0	filter radius (no. of elements)
p	3.0	SIMP penalty
k	$0.1\,\mathrm{N\,m^{-1}}$	actuator stiffness
$\overline{\epsilon}$	0.01	relative crosstalk

Table 4.2: Case study results

Result	g_0	Refinement	Comment
Figure 4.5a	1.00	1	Reference
Figure 4.5b	1.00	3	Overhang filter (north)
Figure 4.5c	0.87	3	Overhang filter (west)
Figure 4.5d	1.03	6	Refined
Figure 4.5e	1.03	6	Stress constrained
Figure 4.5f	0.80	6	Robust

tion (4.11) as

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & g_0\left[\mathbf{v}\left[\mathbf{x}\right]\right] := \min\left(u\left[\mathbf{x}^{\mathrm{e}}\right], u\left[\mathbf{x}^{\mathrm{d}}\right]\right) \\ \text{subject to} \quad & \frac{w^2\left[\mathbf{x}^i\right]}{u^2\left[\mathbf{x}^i\right]} \leq \overline{\epsilon}, \quad i \in \{\mathrm{e},\mathrm{d}\} \\ & v\left[\mathbf{x}^{\mathrm{d}}\right] \leq \overline{v} \\ & x_i \in \mathbb{X}, \quad i = 1,\ldots,N \end{aligned} \tag{4.13}$$

Note the objective function now is a "max-min" formulation between the output displacement of eroded and dilated topologies. The constraint on parasitic motion is applied to both the eroded and dilated topologies, whereas the volume constraint is applied to the worst-performing field, that is the dilated topology.

The resulting flexure topology is shown in Figure 4.5f. Here the blue and red contours indicate the dilated and eroded geometry boundaries, respectively. It is observed that the robustness requirement has a large impact on performance, see Table 4.2. However, apart from the robustness, one guarantees both a minimum feature size on solid and void. In addition and opposed to foregoing results, the robust formulation also provides perfect 0/1 solutions, which lowers the probability of performance decrease upon design interpretation (*e.g.* post-processing conversion to CAD model).

4.5 Conclusions

The combination of topology optimization and additive manufacturing has great potential for compliant mechanism design. However, global design for additive manufacturing constraints have a large impact on both topology and performance. As found in the present work, for some cases, a feasible solution is unreachable or even non-existent. Therefore, instead of global design, a two-step local redesign approach is proposed, targeting performance-critical flexure regions. While performing a local redesign, the method is nevertheless based on performance evaluations of the entire mechanism.

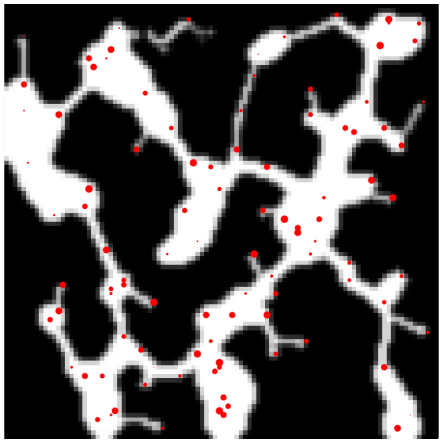
The presented compliant mechanism case study demonstrates the possibility to locally control printability, stress or robustness. Although applied to CM design with DfAM constraints, the generality of the method allows application to any problem in which constraints are locally dominant. The simultaneous consideration of these constraints is a direction for future work. Redesigning multiple regions simultaneously does not fundamentally change the procedure; however it increases computational cost. Sequential redesign of individual flexures is also possible, but the simultaneous approach has the highest design freedom and is expected to yield the best results.

The computational effort of the proposed local redesign method is relatively small as compared to the global design phase. Although not presented in the the present work, the method can be applied to 3D problems without significant adjustments. For 3D problems, the computational efficiency of the method is expected to increase further. The exact additional effort depends on the number, size and refinement ratios of local design domains.

The method allows to apply constraints locally, at those regions it matters most, without drastically decreasing the global performance. In addition, the high resolution local redesign domains are detailed and accurate, without a large additional computational effort. Thereto, the method proves versatile, computationally efficient and overall effective.

5

Efficient multi-partition topology optimization



A multi-partition problem involves multiple partitions of the same discretization, typically corresponding to different loading scenarios. Solving such problems involves multiple factorization/preconditionings of the system matrix, requiring a high computational effort. In this paper, a novel method is proposed to efficiently calculate the responses and accompanying design sensitivities in such multi-partition problems using static condensation for use in gradient-based topology optimization.

This chapter is based on peer-reviewed journal paper:
Koppen, S., Langelaar, M., and van Keulen, F. (2022). *Efficient multi-partition topology optimization.* Computer Methods in Applied Mechanics and Engineering, 393.

Efficient multi-partition topology optimization

Abstract *In topology optimization, the state of structures is typically obtained by numerically evaluating a discretized PDE-based model. The degrees of freedom of such a model can be partitioned in free and prescribed sets to define the boundary conditions. A multi-partition problem involves multiple partitions of the same discretization, typically corresponding to different loading scenarios. As a result, solving multi-partition problems involves multiple factorization/preconditionings of the system matrix, requiring a high computational effort. In this paper, a novel method is proposed to efficiently calculate the responses and accompanying design sensitivities in such multi-partition problems using static condensation for use in gradient-based topology optimization. A main problem class that benefits from the proposed method is the topology optimization of small-displacement multi-input-multi-output compliant mechanisms. However, the method is applicable to any linear problem. We present its formulation and an algorithmic complexity analysis to estimate computational advantages for both direct and iterative solution methods to solve the system of equations, verified by numerical experiments. It is demonstrated that substantial gains are achievable for large-scale multi-partition problems. This is especially true for problems with both a small set of number of degrees of freedom that fully describes the performance of the structure and with large similarities between the different partitions. A major contribution to the gain is the lack of large adjoint analyses required to obtain the sensitivities of the performance measure.*

5.1 Introduction

Finding a solution to large scale gradient-based topology optimization problems in a nested analysis and design approach requires a significant amount of computational effort due to the need to perform one or multiple expensive simulations per design iteration (Amir *et al.* 2009). These simulations are often based on solving discretized governing partial differential equations with many Degrees of Freedom (DOFs). A solution to such a linear system of equations is an assignment of values to the unknown variables at each DOF for given Boundary Conditions (BCs), such that all the equations are satisfied simultaneously. In this setting, we call this set of variables the state, which may represent nodal values such as, *e.g.*, temperature, displacement, pressure or velocity.

One may opt for a direct or iterative solution method to solve the involved system of linear equations; both methods consist of two subsequent steps. Direct methods require the (generally expensive) construction of factorization, and subsequently find the exact solution via comparatively inexpensive back-substitutions (Davis 2006). In contrast, iterative methods require the construction of a preconditioner, and subsequently generate a sequence of

improving approximate solutions until convergence (Amir *et al.* 2014; Saad 2003). Relative cost of preconditioner construction and the iterative solution process depends on many factors, such as the type of preconditioner and condition number. For conciseness, we will hereafter denote the construction of factorization/preconditioner simply by *preprocessing* and the back-substitution/iterative method simply by the *solve*.

A priori to finding the solution to the governing equations, the DOFs are generally partitioned in two unique sets based on the applied BCs. One may either specify (i) the magnitude of the solution at a DOF, that is the state is prescribed, or (ii) the magnitude of the applied discretized load, and thus the state is free. These types are complementary—all DOFs for which the state is free will have an applied load (although possibly with zero magnitude), and vice versa.

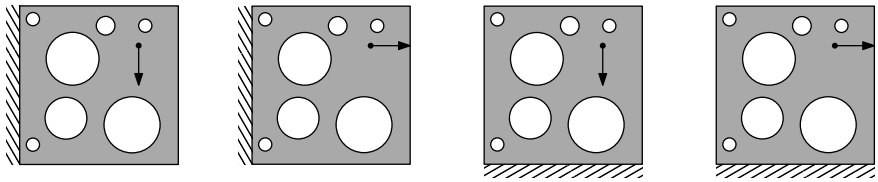
As Rozvany *et al.* (1993) pointed out almost three decades ago, real-world optimization problems generally involve multiple scenarios. We define a scenario as an engineering description of a specific loading condition, see for example the four scenarios depicted in Figure 5.1. Examples of optimization problems that typically involve multiple scenarios are problems with (i) uncertainties in the intended use of the product, (ii) multiple use cases, such as multi-input-multi-output compliant mechanisms, (iii) a strict set of requirements (as obtained for example from the user), see *e.g.* Cavazzuti *et al.* (2011), and (iv) problems that consider—in addition to the critical use case scenarios—variations in boundary and loading conditions originating from manufacturing, assembly, transportation and testing of the product. Each scenario is translated to BCs, *i.e.* prescribed state(s) and applied load(s), that represent the scenario as good as possible. In principle, each scenario comprise a unique partitioning of DOFs. However, scenarios of which only the magnitude of the applied BCs vary, but the type remains the same, share equal partitioning.

In this work, we define an analysis as the process of finding a solution to a properly constrained set of discretized governing equations—that is finding the state $\mathbf{u} \in \mathbb{R}^n$, with $n \in \mathbb{N}$ the number of nodal DOFs. Each analysis relates to a single scenario, and thereto also comprises a unique partitioning of DOFs. A combination of analyses with equal partitioning is hereafter called an analysis set. A single analysis set thus requires finding the states to multiple loads (possibly with varying magnitude) with the same partitioning of DOFs. These different states can generally be efficiently found, since only a single preprocessing process is required (Diaz *et al.* 1992).[1]

Topology optimization problem formulations involving responses (objective and constraints) that depend on the states of multiple analyses sets are called Multi-Partition Optimization Problems (MPOPs). In contrast to problems involving a single partition, in MPOPs the preprocessing cannot be reused. This is evident in the design of, for example, multi-input-multi-output compliant mechanisms (see *e.g.* Jin *et al.* (2017) and Liu *et al.* (2009)).

For clarification, consider a response, herafter denoted by $g \in \mathbb{R}$, that depends on the state(s) of $a \in \mathbb{N}$ analysis sets. Thus, $g = g[\mathbf{U}^{(1)}, ..., \mathbf{U}^{(a)}]$, with $\mathbf{U}^{(i)} \in \mathbb{R}^{n \times l^{(i)}}$ and $l^{(i)} \in \mathbb{N}$ a positive natural number of analyses contained in analysis set i, thus $\mathbf{U}^{(i)} := \left[\mathbf{u}_1^{(i)}, ..., \mathbf{u}_{l^{(i)}}^{(i)}\right]$ with each state $\mathbf{u}^{(i)} \in \mathbb{R}^n$. The states of an analysis set $\mathbf{U}^{(i)}$ implicitly depend on the design

[1] For very large number of loads one can opt for efficient sampling schemes, see *e.g.* Zhang *et al.* (2017, 2020).



(a) First scenario, providing state $\mathbf{u}_1^{(1)}$. The state is prescribed on the left side and the body is subjected to a load $\mathbf{f}_1^{(1)}$.

(b) Second scenario, providing state $\mathbf{u}_2^{(1)}$. The state is prescribed on the left side and the body is subjected to a load $\mathbf{f}_2^{(1)}$.

(c) Third scenario, providing state $\mathbf{u}_1^{(2)}$. The state is prescribed on the bottom and the body is, in this case a similar to Figure 5.1a, point load $\mathbf{f}_1^{(2)}$.

(d) Fourth scenario, providing state $\mathbf{u}_2^{(2)}$. The state is prescribed on the bottom and the body is, in this case a similar to Figure 5.1b, point load $\mathbf{f}_2^{(2)}$.

Figure 5.1: A structural topology to be optimized for four different scenarios, as shown in Figures 5.1a to 5.1d. As a result, this Multi-Partition Optimization Problem (MPOP) consists of two unique analysis sets, both comprising two analyses.

variables $\mathbf{x} \in \mathbb{R}^N$ via the discretized governing equations, that is

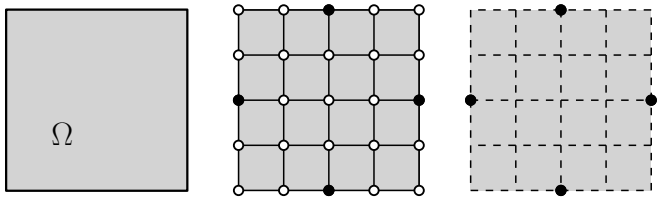
$$\mathbf{K}[\mathbf{x}]\,\mathbf{U}^{(i)} = \mathbf{F}^{(i)}, \quad \forall\, i \in \mathbb{A} \tag{5.1}$$

with the index set of all analyses sets $\mathbb{A} := \{i \in \mathbb{N} \mid i \leq a\}$, the design dependent system matrix $\mathbf{K}[\mathbf{x}] \in \mathbb{R}^{n \times n}$ and loads $\mathbf{F}^{(i)} \in \mathbb{R}^{n \times l^{(i)}}$. Hereafter we assume the system matrix is symmetric and, without loss of generality, assume the applied loads are design independent. Note that, in Equation (5.1), the states and loads of analysis set i contain all DOFs; both free and prescribed states as well as applied and reaction loads.

A simple example of an MPOP is the optimization of a structure considering four scenarios, schematically shown in Figure 5.1. Note that between the four scenarios the set of DOFs for which the state is prescribed differs. Therefore, in this example, two unique analysis sets exist, comprising of two analyses each. The response function g of this artificial problem depends on the states of these two unique analysis sets, *i.e.* $g[\mathbf{U}^{(1)}, \mathbf{U}^{(2)}]$, where for this problem $\mathbf{U}^{(i)} = \left[\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}\right]$ are the states (in this case the displacement) of analysis set i.

In spite of the fact that MPOPs are relevant for many applications, the computational efficiency of both the forward and sensitivity analyses involved in such problems has, to the best knowledge of the authors, not been investigated yet. With an increasing number of analysis sets, the computational effort to solve MPOPs increases quickly. Since the preprocessing used in the analysis occurring in MPOPs is partition-specific, in a straightforward approach the computational effort of solving these problems scales with the number of analysis sets considered. On the contrary, within reasonable limits, the number of analyses per set is of less importance, given the fact that a solve requires negligible computational effort as compared to the preprocessing.

The computational efficiency of an analysis in the context of structural optimization can be increased by efficient iterative solution strategies (see *e.g.* Aage *et al.* (2015), Amir *et al.* (2014), Evgrafov *et al.* (2008), and Wu *et al.* (2016)), or approximated reanalysis, see *e.g.* Amir (2015), Amir *et al.* (2009), and Kirsch (1993). Alternatively, one can opt for model-order reduction techniques, which reduce the computational effort of both analysis and



(a) Bounded domain Ω occupied by a material with relevant physical properties, modeled as a continuum.

(b) Domain from Figure 5.2a discretized into 4 × 4 finite elements. The resulting nodes are split into primary (•) and secondary (○) nodes.

(c) Superelement as obtained after static condensation. The information of the secondary DOFs is condensed into the primary DOFs.

Figure 5.2: Step-wise illustration of static condensation. For simplicity, and without loss of generality, the domain is chosen to be a 2D square and discretized using a structured mesh of 4 × 4 quadrilateral finite elements. For simplicity of visualization, we consider all DOFs of a single node to be either primary or secondary DOFs. Primary nodes are indicated by a disk (•), whereas secondary nodes are indicated by a circle (○).

optimization, see *e.g.* Choi *et al.* (2019) and Yoon (2010). Typically, however, model-order reduction techniques introduce a loss of accuracy. One special reduction method is what is called static condensation amongst engineers, for which the reduction is *exact* (Guyan 1965; Irons 1965). As will be shown hereafter, in particular in MPOPs the use of static condensation is attractive. Moreover, as static condensation, similar to an analysis as previously defined, relies on finding the solution to a set of linear system of equations, it can also be combined with efficient iterative solution techniques and/or approximated reanalysis.

Static condensation was originally developed by Guyan (1965) and Irons (1965), hence the common nomenclature Guyan-Irons reduction scheme. The term 'static' was added to differentiate from its application in dynamic analysis, for which the reduction scheme is no longer exact. Static condensation is regularly applied in the process of substructuring, also known as the Schur complement domain decomposition method (Noor *et al.* 1978; Przemieniecki 1963; Schmit *et al.* 1976). This technique is based on the decomposition of the original structure into one or multiple sub-structures, often referred to as super-elements. The dimensionality reduction is achieved by a procedure that is closely related to Kron's method of tearing and Gaussian elimination (Kron 1955; Wilson 1974).

Consider the continuum from Figure 5.2a. After discretization of the governing equations in n DOFs, as shown in Figure 5.2b, the behaviour of the system can be described by the set of linear equations given by

$$\mathbf{K}[\mathbf{x}]\,\mathbf{u} = \mathbf{f}, \tag{5.2}$$

where, for simplicity, we consider a single analysis set consisting of a single analysis and thus $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{f} \in \mathbb{R}^n$.

The DOF index set $\mathbb{D} := \{i \in \mathbb{N} \mid i \leq n\}$ can be split in subset $\mathbb{M} \subset \mathbb{D}$ containing all indices of the primary DOFs, and its complement $\mathbb{S} := \mathbb{D} \setminus \mathbb{M}$ containing all indices of the secondary DOFs. The primary DOFs are those DOFs contained in the condensed model, whereas the secondary DOFs are eliminated. After partitioning accordingly, one may

rewrite the governing equations from Equation (6.23) using oversets ($\wedge$) and ($\vee$) to represent the properties related to primary and secondary DOFs, respectively, as

$$\begin{bmatrix} \overset{\wedge\wedge}{\mathbf{K}} & \overset{\wedge\vee}{\mathbf{K}} \\ \overset{\vee\wedge}{\mathbf{K}} & \overset{\vee\vee}{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \check{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}} \\ \check{\mathbf{f}} \end{bmatrix}, \tag{5.3}$$

where the principal sub-matrix $\overset{\vee\vee}{\mathbf{K}}$ is *nonsingular*, such that $\overset{\vee\vee}{\mathbf{K}}$ is invertible (Benscoter 1948). This partitioning is completely independent of the partitioning in free (f) and prescribed (p) DOFs. In case we are solely interested in the state of the primary DOFs, we can write the secondary state $\check{\mathbf{u}}$ in terms of the primary state $\hat{\mathbf{u}}$ using the second row of Equation (5.3), which yields

$$\check{\mathbf{u}} = -\overset{\vee\vee}{\mathbf{K}}^{-1}\overset{\vee\wedge}{\mathbf{K}}\hat{\mathbf{u}}, \tag{5.4}$$

assuming $\check{\mathbf{f}} = \mathbf{0}$ for simplicity (the general case will be treated in Section 6.2). Subsequently, $\check{\mathbf{u}}$ can be eliminated by substitution of Equation (5.4) in the first row of Equation (5.3), which gives

$$\left(\overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge\vee}{\mathbf{K}}\overset{\vee\vee}{\mathbf{K}}^{-1}\overset{\vee\wedge}{\mathbf{K}}\right)\hat{\mathbf{u}} = \hat{\mathbf{f}}. \tag{5.5}$$

The result of this condensation is a reduced system matrix $\tilde{\mathbf{K}} \in \mathbb{R}^{m\times m}$, with the number of primary DOFs $m := |\mathbb{M}|$, that exactly describes the behaviour of the full-order model. This reduced system matrix follows as

$$\tilde{\mathbf{K}} := \overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge\vee}{\mathbf{K}}\overset{\vee\vee}{\mathbf{K}}^{-1}\overset{\vee\wedge}{\mathbf{K}}, \tag{5.6}$$

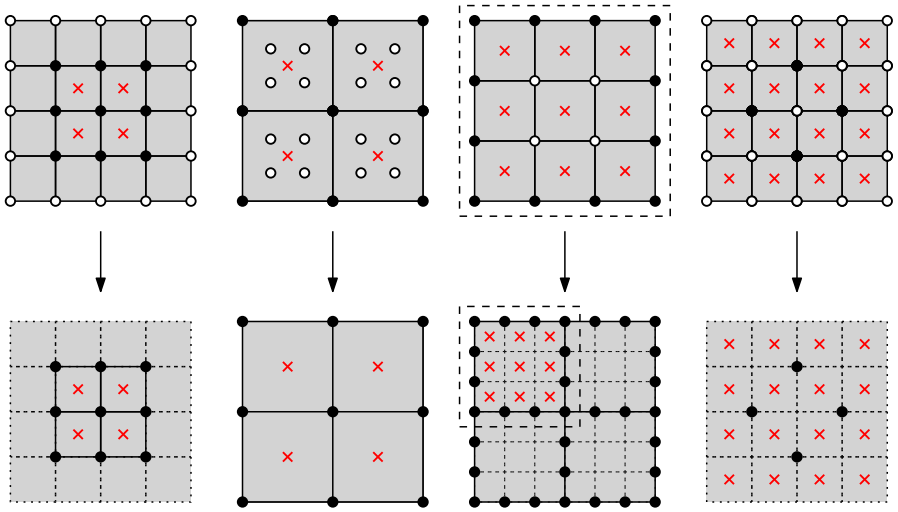
and was originally proposed as the "Schur complement of $\overset{\vee\vee}{\mathbf{K}}$ in $\mathbf{K}$", first introduced in this form and notation by Haynsworth (1968). The eponymous adjective "Schur" was chosen by Haynsworth because of the first use of the complement in the"Schur determinant lemma" published by Schur (1917).

Static condensation has been applied in the field of gradient-based structural optimization in a variety of settings. Contributions include formulations that condense the

- DOFs within non-design domain(s) (Botkin *et al.* 1989; Gangadharan *et al.* 1990; Yang *et al.* 1996),
- internal DOFs of (complex) finite elements or components (Amir *et al.* 2019; Groen *et al.* 2017; Xia *et al.* 2012; Yoon *et al.* 2007),
- internal DOFs of repetitive unit cell(s) (Wu *et al.* 2019; Zhang *et al.* 2006), and/or
- DOFs of a structure that do not occur in the responses, *e.g* all DOFs except input and output in a compliant mechanism (Hasse *et al.* 2009; Wang 2009a).

Figure 5.3 schematically illustrates the use of static condensation in each of these applications, which are discussed in more detail hereafter.

Early studies by Botkin *et al.* (1989), Gangadharan *et al.* (1990), and Yang *et al.* (1996) aim to reduce the computational effort of both the Finite Element Analysis (FEA) and/or the sensitivity analysis of 3D structural optimization problems by static condensation of specific



(a) Condensation of the DOFs in non-design domain(s), noticeable by the lack of influence by design variables in those regions, reducing the dimensionality of the resulting design dependent system matrix.

(b) Condensation of four internal DOFs of four structural or finite elements, effectively reducing the dimensionality of each element.

(c) Condensation of the internal DOFs of a periodic unit cell. The resulting equivalent superelements of lower dimensionality are subsequently assembled in a 2 × 2 grid.

(d) Condensation of all DOFs of a structure that do not *directly* influence any of the responses, resulting in a reduced system matrix consisting solely of the primary DOFs.

Figure 5.3: Four different settings in which static condensation can be used in structural optimization. Top and bottom figures illustrate discretized domains before and after the condensation process, respectively. Similar to Figure 5.2, disks (•) and circles (○) denote primary and secondary nodes, respectively. Red crosses (×) indicate the location of design variables, which only locally influence the design.

sub-domains. The topology in those domains does not depend on the design variables and hence remain unchanged during optimization, as shown in Figure 5.3a. Since only a small part of the structure is subject to change, significant computational efficiency increase can be achieved if the dimensionality of the problem is reduced before the optimization, such that repetitive analysis of the non-design domain(s) is avoided. Since these early studies, this use of condensation in optimization has become commonplace and is also available in commercial software packages.

In the fields of geometrically non-linear, high-resolution and multi-component layout structural optimization, static condensation is effectively applied to eliminate internal DOFs from finite elements before global assembly, as schematically shown in Figure 5.3b. This process reduces element complexity and decreases the dimensionality of the assembly system matrix, see *e.g.* Amir *et al.* (2019), Groen *et al.* (2017), Xia *et al.* (2012), and Yoon *et al.* (2007).

In addition to the aforementioned, the computational effort can be further reduced in case the structure is periodic. Since all repeated cells share the same topology, only static condensation of a single cell is required, as shown in Figure 5.3c. The system matrix of the cell is condensed to reduce the dimensionality of the assembled system matrix that forms the periodic structure, see *e.g.* Wu *et al.* (2019) and Zhang *et al.* (2006).

A rather underexposed application of the static condensation procedure is its use in a unique set of optimization problem formulations of which the responses target a small set of DOFs, see Figure 5.3d. The state itself will generally depend on the full set of design variables. Representative examples are studies proposing problem formulations for the synthesis of compliant mechanisms with desired kinematics (Hasse *et al.* 2017; Li *et al.* 2019; Wang 2009a). The added value of using static condensation in these formulations is the natural way of describing the performance of the structure as a function of a limited set of primary DOFs. In these formulations, the scenarios generally consist of a large number of DOFs that share the same type of BCs.

Although the usefulness of static condensation in FEA has been common knowledge, the usefulness in sensitivity analysis and the impact on the computational efficiency of structural optimization problem formulations has been unexplored so far. Gangadharan *et al.* (1990) acknowledge and demonstrate the potential usefulness of static condensation to reduce the computational effort in the sensitivity analysis for problems in which the design variables only affect a small part of the domain. However, the usage of static condensation shows high potential in terms of computational effort for solving MPOPs as well, as it could reduce the cost of both the forward and sensitivity analysis, as will be shown in Section 6.2. The findings in the present work explore the potential in this field, ready to be applied to real-world problems.

This paper presents a novel method to efficiently solve MPOPs of which the responses involve a relatively small set of DOFs, see Figure 5.3d. The method exploits the concept of static condensation in a way not reported before. We analyse and discuss the method's characteristics and limitations, provide a comparison between the proposed method and the state-of-the-art, and demonstrate the computational efficiency by numerical examples.

This work is focused on gradient-based topology optimization problems, involving linear state equations, solved via nested analysis and design. In addition, we assume (quasi-)static design independent loading conditions. The method operates on any square and

symmetric system matrix, *e.g.* as obtained by assembly of 2D or 3D finite elements.

Note that it is not per se true that analyses of the reduced-order model is cheaper than analyses of the full-order model under all conditions. The reduced system matrix is not only of reduced dimensionality, but also non-sparse. This loss of sparsity may impair any efficiency increase due to dimension reduction. Hence, it is critical to understand under what conditions to apply static condensation in the context of gradient-based structural optimization.

A generalized mathematical description of the method and corresponding sensitivity analyses will be given in Sections 5.3 and 6.2. Both direct and adjoint sensitivity analysis aspects will be discussed. The description forms the basis for the formulation of a computational effort estimate. This estimate is based on the algorithmic complexity as a function of both the number of required analyses and the reduction of number of DOFs. The resulting estimation of computational efficiency with respect to a straightforward approach, *i.e.* performing all analyses without condensation, is outlined in Section 5.4. The computational effort estimate is verified by numerical experiments. Section 5.5 describes a comparison of computational efficiency for a variety of relevant examples. For simplicity, the proposed method is demonstrated by, but by no means limited to, 2D single-physics static topology optimization problems. The description of the method fully complies with any square, positive definite and symmetric system matrix. This is followed by a description of an example implementation and closed by discussions and conclusions.

5.2 Method

This section describes two different approaches to solve a generalized MPOP, namely (i) an elementary approach, and (ii) the novel work presented herein—the condensation approach. We start the description of the elementary approach of solving MPOPs, which will act as a reference.

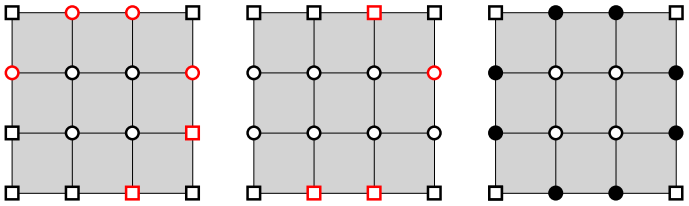
5.2.1 Elementary approach

In the elementary approach one first defines the set of prescribed DOFs for each scenario. For each of the resulting sets of prescribed DOFs one can partition the system of linear equations accordingly. Subsequently, as in standard FEA one solves for the free state(s) (Cook *et al.* 2001).

The index set of all DOFs $\mathbb{D}$ can be split in subset $\mathbb{P}^{(i)} \subset \mathbb{D}$ containing the indices of all prescribed, and its complement $\mathbb{F}^{(i)} := \mathbb{D} \setminus \mathbb{P}^{(i)}$ containing the indices of all free DOFs of analysis set i, hereafter denoted by subscripts 'p' and 'f', respectively. After partitioning accordingly, one may write the governing equations as

$$\begin{bmatrix} \mathbf{K}_{\mathrm{ff}}^{(i)} & \mathbf{K}_{\mathrm{fp}}^{(i)} \\ \mathbf{K}_{\mathrm{pf}}^{(i)} & \mathbf{K}_{\mathrm{pp}}^{(i)} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\mathrm{f}}^{(i)} \\ \mathbf{U}_{\mathrm{p}}^{(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathrm{f}}^{(i)} \\ \mathbf{F}_{\mathrm{p}}^{(i)} \end{bmatrix}, \quad \forall\, i \in \mathbb{A}, \tag{5.7}$$

where $\mathbf{U}_{\mathrm{f}}^{(i)}$ are the free state(s), $\mathbf{U}_{\mathrm{p}}^{(i)}$ the prescribed state(s), $\mathbf{F}_{\mathrm{f}}^{(i)}$ the applied load(s) and $\mathbf{F}_{\mathrm{p}}^{(i)}$ the reaction load(s) of analysis set i. Note that $\mathbf{U}_{\mathrm{f}}^{(i)} \in \mathbb{R}^{f^{(i)} \times l^{(i)}}$, where $f^{(i)} := |\mathbb{F}^{(i)}|$. Where subscript 'f' indicates the subset of the state and italic 'f' indicates the corresponding dimensionality of this set. The solutions to Equation (5.7), $\mathbf{U}_{\mathrm{f}}^{(i)}$, now can be obtained by solving the system



(a) First example analysis. Note the unique subdivision of DOFs in free and prescribed, visualized by shape (○ vs. □) as well as the DOFs of interest, visualized by color (black vs. red)

(b) Second example analysis. Note the change in partitioning (free vs. prescribed) of DOFs according to the boundary conditions of this scenario as well as the change in DOFs of interest.

(c) Subdivision of the set of DOFs $\mathbb{D}$ in disjoint sets $\overset{\vee}{\mathbb{P}}$ (□), $\overset{\vee}{\mathbb{F}}$ (○) and $\mathbb{M}$ (•), according to the definitions in Equations (5.10) and (5.12) and the partitioning as visualized in Figures 5.4a and 5.4b.

Figure 5.4: Example of the discretized domain from Figure 5.2a subjected to two distinct analyses; as shown in Figures 5.4a and 5.4b. The DOFs with prescribed states are denoted by a square (□), the DOFs with free states by a circle (○) and primary DOFs by a black dot (•) (analogously to Figures 5.2 and 5.3). DOFs of interest are coloured in red (○, □). The combination of analyses leads to a subdivision in sets following Equations (5.10) and (5.12), as shown in Figure 5.4c.

of linear equations

$$\mathbf{K}_{\mathrm{ff}}^{(i)}\mathbf{U}_{\mathrm{f}}^{(i)} = \mathbf{F}_{\mathrm{f}}^{(i)} - \mathbf{K}_{\mathrm{fp}}^{(i)}\mathbf{U}_{\mathrm{p}}^{(i)}, \quad \forall\, i \in \mathbb{A}. \tag{5.8}$$

Subsequently, if required, the reaction loads are obtained via

$$\mathbf{F}_{\mathrm{p}}^{(i)} = \mathbf{K}_{\mathrm{pf}}^{(i)}\mathbf{U}_{\mathrm{f}}^{(i)} + \mathbf{K}_{\mathrm{pp}}^{(i)}\mathbf{U}_{\mathrm{p}}^{(i)}, \quad \forall\, i \in \mathbb{A}. \tag{5.9}$$

Note that because of the uniqueness of each analysis set, following our definition, $\mathbb{F}^{(i)} \neq \mathbb{F}^{(j)}$ and $\mathbb{P}^{(i)} \neq \mathbb{P}^{(j)}$ for all $i \neq j \in \mathbb{A}$. Therefore, the number of times one has to preprocess a large system of equations equals a.

5.2.2 Condensation approach

We propose a novel method, aimed at solving MPOPs, consisting of two subsequent steps: condensation and analyses. First, one defines a set of primary DOFs that can fully describe the response function(s) and one condenses the remaining secondary DOFs to obtain a reduced system matrix. With this step, one basically solves for the secondary DOFs applying the same BCs for all analysis sets. Subsequently, a variety of analyses is performed based on this reduced system matrix with varying partitions, originating from the BCs of the scenarios. The latter step is fully equivalent to the elementary approach, however, as a result of the condensation, involves a much smaller system of linear equations. As such, one basically circumvents repeatedly solving for the secondary DOFs. Key is to define the set of primary DOFs in an adequate manner to omit repetitive analyses of large systems of equations, by restricting the size of the set of primary DOFs.

Definition of primary DOFs

Consider again the structure as visualized in Figure 5.2a. For the purpose of explanation, after discretization, we subject this arbitrary structure to two different scenarios as illus-

trated in Figures 5.4a and 5.4b. For each scenario one defines the DOFs with a prescribed state, indicated by a square in Figure 5.4. In addition, one defines per scenario which DOFs are of interest for the response function of the MPOP, in Figure 5.4 indicated in red. Note that both free and prescribed DOFs may be of interest. In what follows we explain how to find the smallest possible set of primary DOFs that can fully represent the behaviour of all scenarios and contains all DOFs of interest for a given MPOP.

For the condensation approach the index set of all DOFs $\mathbb{D}$ is split in three disjoint parts, namely $\mathbb{D} = \overset{\vee}{\mathbb{P}} \sqcup \overset{\vee}{\mathbb{F}} \sqcup \mathbb{M}$ as shown in Figure 5.4c. We define the proper index subset of prescribed secondary DOFs, $\overset{\vee}{\mathbb{P}} \subsetneq \mathbb{D}$, to consist of all indices of DOFs for which the state is prescribed in all analyses excluding those that are a DOF of interest for any of the analyses. Mathematically, this can be written as the intersection between all index sets of prescribed DOFs $\mathbb{P}^{(i)}$ for which the DOFs are not a DOF of interest, *i.e.*

$$\overset{\vee}{\mathbb{P}} := \left\{ x \in \left\{ \cap_{i \in \mathbb{A}} \mathbb{P}^{(i)} \right\} \mid x \notin \left\{ \cup_{i \in \mathbb{A}} \mathbb{M}^{(i)} \right\} \right\}. \tag{5.10}$$

Here $\mathbb{M}^{(i)}$ consists of the indices of all DOFs of interest of analysis i. Fully analogously, the index set of free secondary DOFs, $\overset{\vee}{\mathbb{F}} \subsetneq \mathbb{D}$, is defined as all indices of DOFs for which the state is free in all analyses, excluding those that are a DOF of interest for any of the analyses. Based on these chosen definitions, the index set of primary DOFs, $\mathbb{M}$, is simply defined as the union between all indices of DOFs that are of interest in any of the analysis sets and the indices of DOFs that change freedom between analysis sets, *i.e.*

$$\mathbb{M} := \left(\cup_{i \in \mathbb{A}} \mathbb{M}^{(i)} \right) \cup \left(\left(\cup_{i \in \mathbb{A}} \mathbb{P}^{(i)} \right) \setminus \left(\cap_{i \in \mathbb{A}} \mathbb{P}^{(i)} \right) \right). \tag{5.11}$$

Note this is also the complement of all secondary DOF indices, that is

$$\mathbb{M} := \mathbb{D} \setminus \left(\overset{\vee}{\mathbb{P}} \cup \overset{\vee}{\mathbb{F}} \right). \tag{5.12}$$

In many relevant applications, $1 < m \ll n$, meaning the number of primary DOFs is much lower than the total number of DOFs.

Static condensation

Considering a general MPOP, the response can now be rewritten as $g\left[\hat{\mathbf{U}}^{(1)}, \hat{\mathbf{U}}^{(2)}, ..., \hat{\mathbf{U}}^{(a)}\right]$, with $\hat{\mathbf{U}}^{(i)}$ the state of the primary DOFs of analysis set i. From the system matrix $\mathbf{K} \in \mathbb{R}^{n \times n}$ one can obtain a reduced, but dense, system matrix $\tilde{\mathbf{K}} \in \mathbb{R}^{m \times m}$ through static condensation, which exactly represents the system behaviour in terms of the primary DOFs. We can partition the system of linear equations in line with the subdivision in sets as proposed in Section 5.2.2, which gives

$$\begin{bmatrix} \hat{\hat{\mathbf{K}}} & \hat{\overset{\vee}{\mathbf{K}}}_{\mathrm{f}} & \hat{\overset{\vee}{\mathbf{K}}}_{\mathrm{p}} \\ \overset{\vee}{\hat{\mathbf{K}}}_{\mathrm{f}} & \overset{\vee}{\overset{\vee}{\mathbf{K}}}_{\mathrm{ff}} & \overset{\vee}{\overset{\vee}{\mathbf{K}}}_{\mathrm{fp}} \\ \overset{\vee}{\hat{\mathbf{K}}}_{\mathrm{p}} & \overset{\vee}{\overset{\vee}{\mathbf{K}}}_{\mathrm{pf}} & \overset{\vee}{\overset{\vee}{\mathbf{K}}}_{\mathrm{pp}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}} \\ \overset{\vee}{\mathbf{U}}_{\mathrm{f}} \\ \overset{\vee}{\mathbf{U}}_{\mathrm{p}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{F}} \\ \overset{\vee}{\mathbf{F}}_{\mathrm{f}} \\ \overset{\vee}{\mathbf{F}}_{\mathrm{p}} \end{bmatrix}, \tag{5.13}$$

where $\overset{\vee}{\mathbf{F}}_{\mathrm{f}} \in \mathbb{R}^{\overset{\vee}{f} \times l}$ are the loads on the secondary DOFs and $\overset{\vee}{\mathbf{U}}_{\mathrm{p}} \in \mathbb{R}^{\overset{\vee}{p} \times l}$ the prescribed states of the secondary DOFs. Note that the primary states $\hat{\mathbf{U}}$ contain both free and prescribed

DOFs in accordance with the scenarios imposed, as will become evident from Section 5.2.2. Here we have defined l to be the sum of the loads in all analysis sets, *i.e.*

$$l := \sum_{i \in \mathbb{A}} l^{(i)}. \tag{5.14}$$

Hence, one can write for all $i \in \mathbb{A}$

$$\overset{\vee}{\mathbf{F}}_{\mathrm{f}} = \left[\overset{\vee}{\mathbf{F}}_{\mathrm{f}}^{(1)}, ..., \overset{\vee}{\mathbf{F}}_{\mathrm{f}}^{(a)}\right], \quad \text{with} \quad \overset{\vee}{\mathbf{F}}_{\mathrm{f}}^{(i)} = \left[\overset{\vee}{\mathbf{f}}_{\mathrm{f},1}^{(i)}, ..., \overset{\vee}{\mathbf{f}}_{\mathrm{f},l^{(i)}}^{(i)}\right] \in \mathbb{R}^{\overset{\vee}{f}^{(i)} \times l^{(i)}}. \tag{5.15}$$

Analogously to Equation (5.4), from the second row of Equation (5.13), we obtain

$$\overset{\text{w}}{\mathbf{K}}_{\mathrm{ff}} \overset{\vee}{\mathbf{U}}_{\mathrm{f}} = \overset{\vee}{\mathbf{F}}_{\mathrm{f}} - \overset{\text{n}}{\mathbf{K}}_{\mathrm{f}} \hat{\mathbf{U}} - \overset{\text{w}}{\mathbf{K}}_{\mathrm{fp}} \overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.16}$$

Substitution of Equation (5.16) into the first row of Equation (5.13) gives

$$\tilde{\mathbf{K}} \hat{\mathbf{U}} = \hat{\mathbf{F}} + \tilde{\mathbf{F}}. \tag{5.17}$$

Here $\tilde{\mathbf{K}}$ is the reduced system matrix defined as

$$\tilde{\mathbf{K}} := \overset{\text{m}}{\mathbf{K}} - \overset{\text{n}}{\mathbf{K}}_{\mathrm{f}} \overset{\text{n}}{\mathbf{X}}_{\mathrm{f}}, \quad \tilde{\mathbf{K}} \in \mathbb{R}^{m \times m}, \tag{5.18}$$

and $\tilde{\mathbf{F}}$ is the *reduced load*, defined as

$$\tilde{\mathbf{F}} := \overset{\text{n}}{\mathbf{K}}_{\mathrm{f}} \overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\text{n}}{\mathbf{K}}_{\mathrm{p}} \overset{\vee}{\mathbf{U}}_{\mathrm{p}}, \quad \tilde{\mathbf{F}} \in \mathbb{R}^{m \times l}. \tag{5.19}$$

In order to obtain this reduced system of equations one has to solve the linear systems of equations

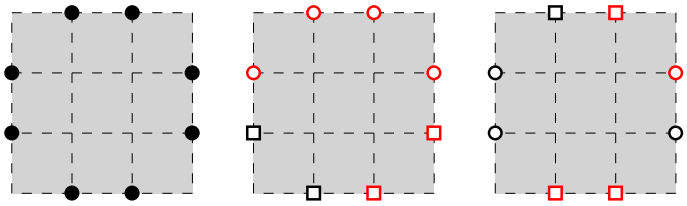
$$\overset{\text{w}}{\mathbf{K}}_{\mathrm{ff}} \overset{\text{n}}{\mathbf{X}}_{\mathrm{f}} = \overset{\text{n}}{\mathbf{K}}_{\mathrm{f}}, \tag{5.20}$$

and

$$\overset{\text{w}}{\mathbf{K}}_{\mathrm{ff}} \overset{\vee}{\mathbf{V}}_{\mathrm{f}} = \overset{\text{w}}{\mathbf{K}}_{\mathrm{fp}} \overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\vee}{\mathbf{F}}_{\mathrm{f}}, \tag{5.21}$$

where we have introduced the 'condensation states' $\overset{\text{n}}{\mathbf{X}}_{\mathrm{f}} \in \mathbb{R}^{\overset{\vee}{f} \times m}$ and $\overset{\vee}{\mathbf{V}}_{\mathrm{f}} \in \mathbb{R}^{\overset{\vee}{f} \times l}$. Note that the total number of analyses thus equals $m + l$. The method is thus capable of handling loads on the secondary DOFs, such as body loads. However, this will have an impact on the computational effort. The additional effort will—in practice—be modest, since (i) the analysis in Equations (5.20) and (5.21) belong the the same analysis set, and (ii) the majority of variations in loading conditions will only affect the primary DOFs. As such, although l might be large, the number of unique loads will generally be a few.

The reduced system matrix $\tilde{\mathbf{K}}$ consists of two parts: $\overset{\text{m}}{\mathbf{K}}$ containing the contributions of the primary DOFs, and $\overset{\text{n}}{\mathbf{K}}_{\mathrm{f}} \overset{\text{n}}{\mathbf{X}}_{\mathrm{f}}$ containing the contributions of the secondary DOFs reflected onto the primary DOFs. Moreover, by definition of Equation (5.6), $\tilde{\mathbf{K}}$ is square and symmetric. One should add the appropriate reduced loads to the loads on the primary DOFs in all subsequent analyses based upon the reduced system matrix. If the loads on the secondary DOFs $\overset{\vee}{\mathbf{F}}_{\mathrm{f}} = \mathbf{0}$ and the prescribed state on the secondary DOFs $\overset{\vee}{\mathbf{U}}_{\mathrm{p}} = \mathbf{0}$, then the reduced load vanishes.



(a) Superelement as obtained after condensation of the system in Figure 5.4c.

(b) First partitioning of the primary DOFs. The partitioning is in line with the partitioning of the full system, see Figure 5.4a.

(c) Second partitioning, in accordance with the scenario in Figure 5.4b.

Figure 5.5: Analyses on using the reduced system matrix, based on the example in Figure 5.4. Figure 5.5a shows the condensed domain of Figure 5.4c, resulting in a reduced system matrix. Figures 5.5b and 5.5c illustrate the analyses based upon the reduced system matrix, in accordance with the scenarios of Figures 5.4a and 5.4b.

After solving Equation (5.17), as will be described in the next section, one can retrieve the secondary states via

$$\overset{\vee}{\mathbf{U}}_{\mathrm{f}} = -\left(\overset{\vee}{\mathbf{X}}_{\mathrm{f}}\hat{\mathbf{U}} + \overset{\vee}{\mathbf{V}}_{\mathrm{f}}\right), \tag{5.22}$$

and subsequently the corresponding reaction loads on the secondary DOFs via

$$\overset{\vee}{\mathbf{F}}_{\mathrm{p}} = \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\hat{\mathbf{U}} + \overset{\vee}{\mathbf{K}}_{\mathrm{pf}}\overset{\vee}{\mathbf{U}}_{\mathrm{f}} + \overset{\vee}{\mathbf{K}}_{\mathrm{pp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.23}$$

Analyses based on the reduced system matrix

After condensation, the reduced system matrix $\tilde{\mathbf{K}}$ can be used to repeatedly analyse the reduced governing equations, and obtain the required states and/or reaction loads for the response(s) of the considered MPOP.

Reconsider the system of equations given in Equation (5.17), *i.e.*

$$\tilde{\mathbf{K}}\hat{\mathbf{U}} = \hat{\mathbf{F}} + \tilde{\mathbf{F}}. \tag{5.24}$$

Note that the dimensionality of this system of equations is m. Similarly to the elementary approach, Section 5.2.1, one can split the indices in $\mathbb{M}$ according to the type of boundary conditions, in the disjoint sets $\mathbb{M} = \hat{\mathbb{F}}^{(i)} \sqcup \hat{\mathbb{P}}^{(i)}$ for each analysis set i. We define

$$\begin{aligned} \hat{\mathbb{F}}^{(i)} &:= \mathbb{M} \cap \mathbb{F}^{(i)}, \\ \hat{\mathbb{P}}^{(i)} &:= \mathbb{M} \cap \mathbb{P}^{(i)}. \end{aligned} \tag{5.25}$$

Partitioning of the system of linear equations accordingly gives

$$\begin{bmatrix} \tilde{\mathbf{K}}^{(i)}_{\mathrm{ff}} & \tilde{\mathbf{K}}^{(i)}_{\mathrm{fp}} \\ \tilde{\mathbf{K}}^{(i)}_{\mathrm{pf}} & \tilde{\mathbf{K}}^{(i)}_{\mathrm{pp}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}^{(i)}_{\mathrm{f}} \\ \hat{\mathbf{U}}^{(i)}_{\mathrm{p}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{F}}^{(i)}_{\mathrm{f}} \\ \hat{\mathbf{F}}^{(i)}_{\mathrm{p}} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{F}}^{(i)}_{\mathrm{f}} \\ \tilde{\mathbf{F}}^{(i)}_{\mathrm{p}} \end{bmatrix}, \quad \forall\, i \in \mathbb{A}. \tag{5.26}$$

Note again, that this division of DOFs is unique for each set of analyses based upon $\tilde{\mathbf{K}}$, since $\hat{\mathbb{F}}^{(i)} \neq \hat{\mathbb{F}}^{(j)}$ for all $i \neq j \in \mathbb{A}$. In addition to the partitioning, for each analysis set the

Table 5.1: Commonly used operators, dimensionality and first appearance equation number.

Op	Dim ($\mathbb{R}$)	Eq.
$\mathbf{K}$	$n \times n$	6.23
$\mathbf{K}_{\mathrm{ff}}^{(i)}$	$f^{(i)} \times f^{(i)}$	5.8
$\mathbf{U}_{\mathrm{f}}^{(i)}$	$f^{(i)} \times l^{(i)}$	5.8
$\tilde{\mathbf{K}}$	$m \times m$	5.18
$\tilde{\mathbf{F}}$	$m \times l$	5.19
$\overset{\vee}{\mathbf{K}}_{\mathrm{ff}}$	$\overset{\vee}{f} \times \overset{\vee}{f}$	5.20
$\overset{\vee}{\mathbf{X}}_{\mathrm{f}}$	$\overset{\vee}{f} \times m$	5.20
$\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$	$\overset{\vee}{f} \times l$	5.20
$\tilde{\mathbf{K}}_{\mathrm{ff}}^{(i)}$	$\hat{f}^{(i)} \times \hat{f}^{(i)}$	5.27
$\hat{\mathbf{U}}_{\mathrm{f}}^{(i)}$	$\hat{f}^{(i)} \times l^{(i)}$	5.27

appropriate loads are selected. Thus, whereas $\tilde{\mathbf{F}} \in \mathbb{R}^{m \times l}$, $\tilde{\mathbf{F}}_{\mathrm{f}}^{(i)} \in \mathbb{R}^{\hat{f}^{(i)} \times l^{(i)}}$. The subdivision is schematically visualized in Figure 5.5, in line with the representation in Figure 5.4.

The states of a variety of analyses sets can be obtained by solving the set of linear equations similar to the elementary approach, *i.e.*

$$\tilde{\mathbf{K}}_{\mathrm{ff}}^{(i)} \hat{\mathbf{U}}_{\mathrm{f}}^{(i)} = \hat{\mathbf{F}}_{\mathrm{f}}^{(i)} - \tilde{\mathbf{K}}_{\mathrm{fp}}^{(i)} \hat{\mathbf{U}}_{\mathrm{p}}^{(i)} + \tilde{\mathbf{F}}_{\mathrm{f}}^{(i)}, \quad \forall\, i \in \mathbb{A}. \tag{5.27}$$

Note that system matrix $\tilde{\mathbf{K}}_{\mathrm{ff}}^{(i)}$ is square and non-singular for all i. If required, the reaction loads are obtained via

$$\hat{\mathbf{F}}_{\mathrm{p}}^{(i)} = \tilde{\mathbf{K}}_{\mathrm{pf}}^{(i)} \hat{\mathbf{U}}_{\mathrm{f}}^{(i)} + \tilde{\mathbf{K}}_{\mathrm{pp}}^{(i)} \hat{\mathbf{U}}_{\mathrm{p}}^{(i)} - \tilde{\mathbf{F}}_{\mathrm{p}}^{(i)}, \quad \forall\, i \in \mathbb{A}. \tag{5.28}$$

5.2.3 Comparison of response evaluation

Although the approaches obtain *exactly* the same state and thereto response value(s), they differ in implementation, with critical consequences for the required computational effort. In order to keep an overview of the large number of sets, subscripts and different operators presented, the most important operators, their dimensionality and reference to their first use can be found in Table 5.1.

To highlight the differences, consider the pseudo codes in Algorithms 1 and 2, that describe the procedures to obtain the same response function using the elementary and condensation approach, respectively. Note the *critical* difference between the solves in Algorithm 1: Line 2 and Algorithm 2: Line 10. To obtain the state(s) of the DOFs of interest, in the condensation approach, one solves for the state using the partitioned reduced system matrices $\tilde{\mathbf{K}}_{\mathrm{ff}}^{(i)} \in \mathbb{R}^{\hat{f}^{(i)} \times \hat{f}^{(i)}}$, where the number of free primary DOFs $\hat{f}^{(i)} := m - \hat{p}^{(i)}$ and $\hat{p}^{(i)} := |\hat{\mathbb{P}}^{(i)}|$. In the elementary approach, one solves for the state using system matrices $\mathbf{K}_{\mathrm{ff}}^{(i)} \in \mathbb{R}^{f^{(i)} \times f^{(i)}}$, where the number of free DOFs $f^{(i)} := n - p^{(i)}$ and $p^{(i)} := |\mathbb{P}^{(i)}|$. Since $f^{(i)} \approx n$ and $\hat{f}^{(i)} \approx m$, for sufficiently large systems, the dimensionality of the constrained system matrix $\mathbf{K}_{\mathrm{ff}}^{(i)}$ is much

larger than the constrained reduced system matrix $\tilde{\mathbf{K}}_{\mathrm{ff}}^{(i)}$, that is $n \gg m$ for all analysis sets i. The difference in computational effort of preprocessing a system matrix of dimensionality $n \times n$ as compared to $m \times m$ is large. The condensation approach only requires a single preprocessing of a large system matrix (Algorithm 2: Line 4), compared to a preprocessing steps of a large system matrix for the elementary approach (Algorithm 1: Line 2). As a result, the computational effort to evaluate the response(s) can be highly reduced by using the condensation approach, when multiple scenarios are considered.

Upon close examination of the approaches, it becomes apparent the elementary approach contains the differences between the BCs of analysis sets in the large system of equations (Algorithm 1: Line 2), effectively solving *repeatedly* for the large amount of secondary DOFs. In contrast, the condensation approach basically uses some pre-processing (the static condensation in Algorithm 2: Line 4) to solve for the secondary DOFs. The different partitionings are 'moved' to the system of equations of reduced dimensionality, see Algorithm 2: Line 10.

In terms of implementation, little is required to extend the elementary approach to the condensation approach. Apart from additional 'administrative tasks' involving index sets, *i.e.* Algorithm 2: Lines 1—3 and 8—9, the extension requires additional matrix-vector multiplications (Lines 5—6) and solves (Lines 4 and 10). Irrespective of the programming language of choice, those operations are also required for the elementary approach. The similarity between approaches in terms of implementation becomes all the more evident upon examination of the attached MATLAB code.

Algorithm 1 elementary approach: response evaluation

1: **for all** $i \in \mathbb{A}$ **do**
2: solve $\mathbf{K}_{\mathrm{ff}}^{(i)}\mathbf{U}_{\mathrm{f}}^{(i)} = \mathbf{F}_{\mathrm{f}}^{(i)} - \mathbf{K}_{\mathrm{fp}}^{(i)}\mathbf{U}_{\mathrm{p}}^{(i)}, \quad \mathbf{K}_{\mathrm{ff}}^{(i)} \in \mathbb{R}^{f^{(i)} \times f^{(i)}}$
3: **end for**
4: **return** $g \longleftarrow g\left[\mathbf{U}^{(1)}, ..., \mathbf{U}^{(a)}\right]$

Algorithm 2 Condensation approach: response evaluation

$\overset{\vee}{\mathbb{P}} := \{x \in \{\cap_{i\in\mathbb{A}}\, \mathbb{P}^{(i)}\} \mid x \notin \{\cup_{i\in\mathbb{A}}\, \mathbb{M}^{(i)}\}\}$
$\overset{\vee}{\mathbb{F}} := \{x \in \{\cap_{i\in\mathbb{A}}\, \mathbb{F}^{(i)}\} \mid x \notin \{\cup_{i\in\mathbb{A}}\, \mathbb{M}^{(i)}\}\}$
$\mathbb{M} \longleftarrow \mathbb{D} \setminus \left(\overset{\vee}{\mathbb{P}} \cup \overset{\vee}{\mathbb{F}}\right)$
solve $\quad \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}} \left[\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} \quad \overset{\vee}{\mathbf{V}}_{\mathrm{f}}\right] = \left[\overset{\wedge}{\mathbf{K}}_{\mathrm{f}} \quad \overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}} \overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\vee}{\mathbf{F}}_{\mathrm{f}}\right], \quad \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}} \in \mathbb{R}^{\overset{\vee}{f}\times\overset{\vee}{f}}$
$\tilde{\mathbf{K}} \longleftarrow \overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge\vee}{\mathbf{K}}_{\mathrm{f}} \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}$
$\tilde{\mathbf{F}} \longleftarrow \overset{\wedge\vee}{\mathbf{K}}_{\mathrm{f}} \overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\wedge\vee}{\mathbf{K}}_{\mathrm{p}} \overset{\vee}{\mathbf{U}}_{\mathrm{p}}$
for all $i \in \mathbb{A}$ **do**
  $\overset{\wedge}{\mathbb{F}}^{(i)} \longleftarrow \mathbb{M} \cap \mathbb{F}^{(i)}$
  $\overset{\wedge}{\mathbb{P}}^{(i)} \longleftarrow \mathbb{M} \cap \mathbb{P}^{(i)}$
  solve $\quad \tilde{\mathbf{K}}^{(i)}_{\mathrm{ff}} \overset{\wedge}{\mathbf{U}}^{(i)}_{\mathrm{f}} = \overset{\wedge}{\mathbf{F}}^{(i)}_{\mathrm{f}} - \tilde{\mathbf{K}}^{(i)}_{\mathrm{fp}} \overset{\wedge}{\mathbf{U}}^{(i)}_{\mathrm{p}} + \tilde{\mathbf{F}}^{(i)}_{\mathrm{f}}, \quad \tilde{\mathbf{K}}^{(i)}_{\mathrm{ff}} \in \mathbb{R}^{\hat{f}^{(i)}\times\hat{f}^{(i)}}$
end for
return $g \longleftarrow g\left[\overset{\wedge}{\mathbf{U}}^{(1)}, ..., \overset{\wedge}{\mathbf{U}}^{(a)}\right]$

5

5.3 Sensitivity analysis

Design sensitivities are essential for gradient-based structural optimization schemes, as they provide the basis for the design update. This section describes the sensitivity analysis of all relevant entries used in the condensation approach, as described in Section 5.2.2. Since the second part of the condensation approach, as described in Section 5.2.2, also covers the elementary approach we omit a separate description of the sensitivities for the elementary approach.

Making use of the chain rule of differentiation, the full derivative of a response function $g\left[\overset{\wedge}{\mathbf{U}}\left[\tilde{\mathbf{K}}[\mathbf{K}[\mathbf{x}]], \tilde{\mathbf{F}}[\mathbf{K}[\mathbf{x}]]\right]\right]$ to design variable x_k reads as

$$\frac{\mathrm{d}g}{\mathrm{d}x_k} = \frac{\mathrm{d}g}{\mathrm{d}\overset{\wedge}{\mathbf{U}}} : \left(\frac{\partial \overset{\wedge}{\mathbf{U}}}{\partial \tilde{\mathbf{K}}} : \frac{\partial \tilde{\mathbf{K}}}{\partial \mathbf{K}} + \frac{\partial \overset{\wedge}{\mathbf{U}}}{\partial \tilde{\mathbf{F}}} : \frac{\partial \tilde{\mathbf{F}}}{\partial \mathbf{K}}\right) : \frac{\mathrm{d}\mathbf{K}}{\mathrm{d}x_k}. \tag{5.29}$$

Herein, we assume $\frac{\mathrm{d}g}{\mathrm{d}\overset{\wedge}{\mathbf{U}}}$ and $\frac{\mathrm{d}\mathbf{K}}{\mathrm{d}x_k}$ are known, since they are problem and response dependent. Although not a restriction of the method, for simplicity of the discussion in this section, we assume design independent loads. This section is subdivided into the separate sensitivity analysis of

1. the reduced system matrix $\tilde{\mathbf{K}}$ to the system matrix $\mathbf{K}$,
2. the reduced load(s) $\tilde{\mathbf{F}}$ to the system matrix $\mathbf{K}$, and
3. the primary state(s) $\overset{\wedge}{\mathbf{U}}$ to the reduced system matrix $\tilde{\mathbf{K}}$ and load(s) $\tilde{\mathbf{F}}$.

The remaining sensitivity analyses, *e.g* the sensitivities of the reaction loads $\overset{\wedge}{\mathbf{F}}_{\mathrm{p}}$ with respect to the reduced system matrix, are not considered critical for understanding of the method

and subsequent conclusions. However, for completeness those and additional sensitivities can be found in Section 5.A.

In the following, the sub-matrices of the system matrices in Equations (5.13) and (5.26) are written using selection matrices. Some examples are:

$$\begin{aligned} \mathbf{K}_{\mathrm{pf}} &:= \mathbf{S}_{\mathrm{p}}^{\mathsf{T}} \mathbf{K} \mathbf{S}_{\mathrm{f}} \\ \overset{\wedge\wedge}{\mathbf{K}} &:= \overset{\wedge}{\mathbf{S}}^{\mathsf{T}} \mathbf{K} \overset{\wedge}{\mathbf{S}} \\ \overset{\wedge\vee}{\mathbf{K}}_{\mathrm{f}} &:= \overset{\wedge}{\mathbf{S}}^{\mathsf{T}} \mathbf{K} \overset{\vee}{\mathbf{S}}_{\mathrm{f}} \\ \tilde{\mathbf{K}}_{\mathrm{pf}} &:= \tilde{\mathbf{S}}_{\mathrm{p}}^{\mathsf{T}} \tilde{\mathbf{K}} \tilde{\mathbf{S}}_{\mathrm{f}} \end{aligned} \tag{5.30}$$

These highly sparse rectangular selection matrices $\overset{(.)}{\mathbf{S}}_{(.)}$ can be used to select parts from the system matrix, which will prove useful during the sensitivity analysis.

5.3.1 Sensitivities of the reduced system matrix

Consider a response $g\left[\hat{\mathbf{U}}\left[\tilde{\mathbf{K}}[\mathbf{K}[\mathbf{x}]]\right]\right]$, with

$$\tilde{\mathbf{K}} := \overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge\vee}{\mathbf{K}}_{\mathrm{f}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}}. \tag{5.18 revisited}$$

Here the system matrix is a function of design variables $\mathbf{x}$, *i.e.* $\mathbf{K}[\mathbf{x}]$. However, for demonstration of the method we are solely interested in the sensitivities of the reduced system matrix $\tilde{\mathbf{K}}$ with respect to the system matrix $\mathbf{K}$. Therefore, this dependency is omitted from here onward. For conciseness, we use the full differential. The full derivative of the response function can be written as

$$\mathrm{d}g\left[\tilde{\mathbf{K}}[\mathbf{K}]\right] = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}} : \mathrm{d}\mathbf{K}, \tag{5.31}$$

where $\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}} : \frac{\mathrm{d}\hat{\mathbf{U}}}{\mathrm{d}\tilde{\mathbf{K}}}$ is the sensitivity of the response with respect to $\tilde{\mathbf{K}}$ and $\mathrm{d}\mathbf{K}$ the sensitivity of the system matrix. In the following we assume both are known. The aim is to find $\frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}}$, *i.e.* the sensitivities of the reduced system matrix with respect to the system matrix.

The total derivative of $\tilde{\mathbf{K}}$ is given by

$$\mathrm{d}\tilde{\mathbf{K}}[\mathbf{K}] = \mathrm{d}\overset{\wedge\wedge}{\mathbf{K}} - \mathrm{d}\overset{\wedge\vee}{\mathbf{K}}_{\mathrm{f}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\wedge\vee}{\mathbf{X}}_{\mathrm{f}} \mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge\vee}{\mathbf{X}}_{\mathrm{f}} \mathrm{d}\overset{\vee\wedge}{\mathbf{K}}_{\mathrm{f}}. \tag{5.32}$$

Here we have made use of

$$\mathrm{d}\overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}} = -\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}^{-1} \mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}^{-1} \mathrm{d}\overset{\vee\wedge}{\mathbf{K}}_{\mathrm{f}}. \tag{5.33}$$

Substitution of Equation (5.30) into Equation (5.32) yields

$$\mathrm{d}\tilde{\mathbf{K}}[\mathbf{K}] = \left(\hat{\mathbf{S}} - \overset{\vee}{\mathbf{S}}_{\mathrm{f}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}}\right)^{\mathsf{T}} \mathrm{d}\mathbf{K} \left(\hat{\mathbf{S}} - \overset{\vee}{\mathbf{S}}_{\mathrm{f}} \overset{\vee\wedge}{\mathbf{X}}_{\mathrm{f}}\right). \tag{5.34}$$

By rewriting[2,3] Equation (5.34) we obtain the full derivative

$$\mathrm{d}\tilde{\mathbf{K}}[\mathbf{K}] = (\mathbf{A} \otimes \mathbf{A}) : \mathrm{d}\mathbf{K}, \tag{5.35}$$

with

$$\mathbf{A} := \hat{\mathbf{S}} - \check{\mathbf{S}}_\mathrm{f} \check{\mathbf{X}}_\mathrm{f}. \tag{5.36}$$

Thus

$$\frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}} = \mathbf{A} \otimes \mathbf{A}, \tag{5.37}$$

which consists only of terms that have been previously computed in the condensation. The following section further explores this property of the reduced system matrix. The full step-by-step derivation, both for the direct and adjoint approach, can be found in Section 5.A.1.

On the self-adjointness of the reduced system matrix

Following the adjoint sensitivity approach (Arora *et al.* 1979; Belegundu 1986) we define

$$\mathcal{L}\left[\tilde{\mathbf{K}}[\mathbf{K}], \hat{\boldsymbol{\Lambda}}_\mathrm{f}\right] = g\left[\tilde{\mathbf{K}}[\mathbf{K}]\right] + \hat{\boldsymbol{\Lambda}}_\mathrm{f} : \left(\check{\mathbf{K}}_\mathrm{ff} \hat{\mathbf{X}}_\mathrm{f} - \check{\mathbf{K}}_\mathrm{f}\right), \tag{5.38}$$

where $\hat{\boldsymbol{\Lambda}}_\mathrm{f}$ are the Lagrange multipliers. The sensitivities of the Lagrangian (extensive derivation can be found in Section 5.A.1) are obtained via

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \left(\hat{\mathbf{S}}^\mathsf{T} \mathrm{d}\mathbf{K}\mathbf{A}\right) + \hat{\boldsymbol{\Lambda}}_\mathrm{f} : \left(\check{\mathbf{S}}_\mathrm{f}^\mathsf{T} \mathrm{d}\mathbf{K}\mathbf{A}\right). \tag{5.39}$$

To find $\hat{\boldsymbol{\Lambda}}_\mathrm{f}$, one has to solve the following system of equations:

$$\check{\mathbf{K}}_\mathrm{ff} \hat{\boldsymbol{\Lambda}}_\mathrm{f} = \hat{\mathbf{K}}_\mathrm{f} \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}}. \tag{5.40}$$

By suitable choice of calculation order (Vanderplaats 1980), this can be rewritten as

$$\hat{\boldsymbol{\Lambda}}_\mathrm{f} = \check{\mathbf{K}}_\mathrm{ff}^{-1} \left(\hat{\mathbf{K}}_\mathrm{f} \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}}\right) = \hat{\mathbf{X}}_\mathrm{f} \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}}. \tag{5.41}$$

In contrast to the general adjoint approach, where the calculation order is determined by the number of responses versus the number of design variables, the most suitable choice in this case is the same under all conditions, since $\hat{\mathbf{X}}_\mathrm{f}$ has already been calculated. Thus, the sensitivities of the reduced system matrix can be calculated without solving an additional system of linear equations. Since both the state and adjoint fields are exactly identical we can conclude that—for a symmetric system matrix—an optimization problem with response $g\left[\tilde{\mathbf{K}}\right]$ has a *self-adjoint* operator (Belegundu 1986; Rozvany *et al.* 1993) independent of the self-adjointness of the operator(s) $\tilde{\mathbf{K}}$ is acting upon. As a result, no extra preprocessing is required to compute the sensitivities of $\tilde{\mathbf{K}}$. This has critical consequences with respect to the computational efficiency of the condensation approach.

[2] Here we introduce the dyadic product, also known and equivalent to the outer or tensor product. The dyadic product of two vectors $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ can be represented as $\mathbf{a} \otimes \mathbf{b} \equiv \mathbf{a}\mathbf{b}^\mathsf{T} \equiv a_i b_j$. For matrices $\mathbf{A} \in \mathbb{R}^{n \times c}$ and $\mathbf{B} \in \mathbb{R}^{m \times c}$, the dyadic product is defined as $\mathbf{A} \otimes \mathbf{B} \equiv \mathbf{A}\mathbf{B}^\mathsf{T} \equiv A_{ij} B_{kj}$.

[3] $\mathbf{A}^\mathsf{T}\mathbf{B}\mathbf{C} \equiv A_{ij} B_{jk} C_{ki} = A_{ji} C_{ik} B_{jk} = \left(\mathbf{A}\mathbf{C}^\mathsf{T}\right) : \mathbf{B} \equiv (\mathbf{A} \otimes \mathbf{C}) : \mathbf{B}$

5.3.2 Sensitivities of the reduced load

Since the required steps to obtain the sensitivities of the reduced load are very similar to those in the previous section we solely present the result here. The full derivation—both for the direct and adjoint approach—can be found in Section 5.A.2.

Recall the definition of the reduced load, namely

$$\tilde{\mathbf{F}}\left[\mathbf{K}, \overset{\vee}{\mathbf{F}}_\mathrm{f}, \overset{\vee}{\mathbf{U}}_\mathrm{p}\right] := \overset{\wedge}{\mathbf{K}}_\mathrm{f} \overset{\vee}{\mathbf{V}}_\mathrm{f} - \overset{\wedge}{\mathbf{K}}_\mathrm{p} \overset{\vee}{\mathbf{U}}_\mathrm{p}. \qquad \text{(5.19 revisited)}$$

Let us consider a response function $g\left[\tilde{\mathbf{F}}[\mathbf{K}]\right]$, hence omitting the dependency on the free secondary load(s) $\overset{\vee}{\mathbf{F}}_\mathrm{f}$ and prescribed secondary state(s) $\overset{\vee}{\mathbf{U}}_\mathrm{p}$ here. Substitution of Equation (5.30) into Equation (5.19) and application of either direct or adjoint sensitivity analysis yields

$$\frac{\partial \tilde{\mathbf{F}}}{\partial \mathbf{K}} = \mathbf{A} \otimes \mathbf{B} \tag{5.42}$$

with

$$\mathbf{B} := \overset{\vee}{\mathbf{S}}_\mathrm{f} \overset{\vee}{\mathbf{V}}_\mathrm{f} - \overset{\vee}{\mathbf{S}}_\mathrm{p} \overset{\vee}{\mathbf{U}}_\mathrm{p}. \tag{5.43}$$

Note that Equation (5.42) solely involves quantities that have been calculated in the condensation process and requires therefore negligible additional computational effort.

5.3.3 Sensitivities of the state

In an MPOP the response function, often $g\left[\hat{\mathbf{U}}_\mathrm{f}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}, \hat{\mathbf{F}}_\mathrm{f}, \hat{\mathbf{U}}_\mathrm{p},\right]\right]$, generally depends on multiple states, *i.e.* $g\left[\hat{\mathbf{U}}_\mathrm{f}^{(1)},,,...,\hat{\mathbf{U}}_\mathrm{f}^{(a)}\right]$ is obtained by solving

$$\tilde{\mathbf{K}}_\mathrm{ff}^{(i)} \hat{\mathbf{U}}_\mathrm{f}^{(i)} = \hat{\tilde{\mathbf{F}}}_\mathrm{f}^{(i)} - \tilde{\mathbf{K}}_\mathrm{fp}^{(i)} \hat{\mathbf{U}}_\mathrm{p}^{(i)} + \tilde{\mathbf{F}}_\mathrm{f}^{(i)}, \quad \forall\, i \in \mathbb{A}. \qquad \text{(5.27 revisited)}$$

Therefore, one is required to determine the total derivative $\mathrm{d}g\left[\hat{\mathbf{U}}_\mathrm{f}^{(i)}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}^{(i)}, \hat{\mathbf{F}}_\mathrm{f}^{(i)}, \hat{\mathbf{U}}_\mathrm{p}^{(i)}\right]\right]$ for all $i \in \mathbb{A}$. For clarity, the superscript i is omitted from here on. Moreover, for conciseness we omit the dependency on the free loads $\hat{\mathbf{F}}_\mathrm{f}$ and prescribed states $\hat{\mathbf{U}}_\mathrm{p}$ in this section. Although this section focuses on the condensation approach, the following is similar for the elementary approach, where $g\left[\mathbf{U}_\mathrm{f}[\mathbf{K}]\right]$. Complete and step-by-step derivations can be found in Section 5.A.3.

The full derivative of the response function is given by

$$\mathrm{d}g\left[\hat{\mathbf{U}}_\mathrm{f}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}\right]\right] = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} : \left(\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{K}}} : \mathrm{d}\tilde{\mathbf{K}} + \frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{F}}} : \mathrm{d}\tilde{\mathbf{F}}\right). \tag{5.44}$$

We aim to find $\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{K}}}$ and $\frac{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}}{\mathrm{d}\tilde{\mathbf{F}}}$, while assuming $\mathrm{d}\tilde{\mathbf{K}}$, $\mathrm{d}\tilde{\mathbf{F}}$ and $\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}}$ to be known. The full derivative $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ can be obtained by differentiation of Equation (5.27), which gives

$$\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}\right] = \tilde{\mathbf{K}}_\mathrm{ff}^{-1}\left(\mathrm{d}\tilde{\mathbf{F}}_\mathrm{f} - \mathrm{d}\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{U}}_\mathrm{f} - \mathrm{d}\tilde{\mathbf{K}}_\mathrm{fp}\hat{\mathbf{U}}_\mathrm{p}\right). \tag{5.45}$$

Substitution of Equation (5.30) and simplification yields

$$\frac{\partial g}{\partial \tilde{\mathbf{K}}} = -\tilde{\mathbf{S}}_{\mathrm{f}} \hat{\mathbf{\Lambda}}_{\mathrm{f}} \otimes \hat{\mathbf{U}} \tag{5.46}$$

and

$$\frac{\partial g}{\partial \tilde{\mathbf{F}}} = \tilde{\mathbf{S}}_{\mathrm{f}} \hat{\mathbf{\Lambda}}_{\mathrm{f}}, \tag{5.47}$$

where

$$\tilde{\mathbf{K}}_{\mathrm{ff}} \hat{\mathbf{\Lambda}}_{\mathrm{f}} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_{\mathrm{f}}} \tag{5.48}$$

is the adjoint equation and $\hat{\mathbf{\Lambda}}^{(i)} \in \mathbb{R}^{m \times l^{(i)}}$ is the state of the adjoint problem of analysis set i. Thus, to obtain $\mathrm{d}\hat{\mathbf{U}}_{\mathrm{f}}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}\right]$ one requires one additional system of linear equations to be solved per analysis set.

The total number of adjoint loads to solve an MPOP theoretically equals the number of responses times the number of states, that is the total number of adjoint loads

$$b = hl = h \sum_{i=1}^{a} l^{(i)}, \tag{5.49}$$

with h the number of responses of the MPOP. However, if the right-hand-side is linear dependent on the applied load, the problem is self-adjoint and solving the corresponding adjoint equation can be omitted (Rozvany *et al.* 1993). This is for example the case in classical compliance minimization (Bendsøe *et al.* 1988). What is more, in practice, the majority of responses depend on a single state or multiple states of a single analysis set, that is many adjoint loads $\frac{\mathrm{d}g_k}{\mathrm{d}\hat{\mathbf{u}}_j^{(i)}}$ are zero. Thereto, generally the total number of adjoint loads satisfies

$$b = \sum_{j=1}^{h} \sum_{i=1}^{a} b_j^{(i)}, \tag{5.50}$$

with $b_j^{(i)}$ the number of non-zero adjoint loads of response function j to the states of analysis set i, and thus $b^{(i)} \leq l^{(i)}$.

5.3.4 Comparison of sensitivity analysis

To compare the sensitivity analyses, consider the pseudo-codes Algorithms 3 and 4, respectively for the elementary and condensation approach. Herein, it is assumed both applied loads and prescribed states are design independent. Note the difference in dimensionality of the systems of equations to be solved in the second step of the algorithms, that is Algorithm 3: Line 2 and Algorithm 4: Line 2. It may be assumed here no additional preprocessing is required, as the preprocessing of the response evaluation can be reused. As a result, in the elementary approach, for each response g_j, b_j adjoint states of substantial dimensionality are to be solved for. In contrast, assuming $m \ll n$, the computational effort of the sensitivity analysis in the condensation approach is negligible. Most notably is the fact that *zero* solves of substantial dimensionality are required, see Section 5.3.1.

Apart from some additional matrix-vector operations, see Algorithm 4: Lines 4—5, implementation of the sensitivity analysis of the condensation approach is straightforward. Also note the similarity between the structure of Algorithm 3: Line 4 and Algorithm 4: Line 6.

Full step-by-step sensitivity analysis of $g\left[\hat{\mathbf{U}}_\mathrm{f}\left[\mathbf{K},\check{\mathbf{F}}_\mathrm{f},\check{\mathbf{U}}_\mathrm{p},\hat{\mathbf{F}}_\mathrm{f},\hat{\mathbf{U}}_\mathrm{p}\right]\right]$ and $g\left[\hat{\mathbf{F}}_\mathrm{p}\left[\mathbf{K},\check{\mathbf{F}}_\mathrm{f},\check{\mathbf{U}}_\mathrm{p},\hat{\mathbf{F}}_\mathrm{f},\hat{\mathbf{U}}_\mathrm{p}\right]\right]$, without those assumptions, can be found in the appendices of Koppen *et al.* (2022c). In addition, the derivation of sensitivities for responses involving the state or reaction load of secondary DOFs, that is $g\left[\check{\mathbf{U}}_\mathrm{f}\left[\mathbf{K},\check{\mathbf{F}}_\mathrm{f},\check{\mathbf{U}}_\mathrm{p},\hat{\mathbf{F}}_\mathrm{f},\hat{\mathbf{U}}_\mathrm{p}\right]\right]$ and $g\left[\check{\mathbf{F}}_\mathrm{p}\left[\mathbf{K},\check{\mathbf{F}}_\mathrm{f},\check{\mathbf{U}}_\mathrm{p},\hat{\mathbf{F}}_\mathrm{f},\hat{\mathbf{U}}_\mathrm{p}\right]\right]$, and a summary of all sensitivities can also be found in the appendices of Koppen *et al.* (2022c).

Algorithm 3 elementary approach: sensitivity analysis

1: **for all** $i \in \mathbb{A}$ **do**
2: solve $\quad \mathbf{K}_\mathrm{ff}^{(i)}\boldsymbol{\Lambda}_\mathrm{f}^{(i)} = \frac{\partial g}{\partial \mathbf{U}_\mathrm{f}^{(i)}}, \quad \mathbf{K}_\mathrm{ff}^{(i)} \in \mathbb{R}^{f^{(i)}\times f^{(i)}}$
3: **end for**
4: **return** $\frac{\partial g}{\partial x_k} \longleftarrow -\sum_{i\in\mathbb{A}}\sum_{j=1}^{l^{(i)}} \boldsymbol{\lambda}_j^{(i)} \cdot \frac{\partial \mathbf{K}}{\partial x_k}\mathbf{u}_j^{(i)}$

Algorithm 4 Condensation approach: sensitivity analysis

1: **for all** $i \in \mathbb{A}$ **do**
2: solve $\quad \tilde{\mathbf{K}}_\mathrm{ff}^{(i)}\hat{\boldsymbol{\Lambda}}_\mathrm{f}^{(i)} = \frac{\partial g}{\partial \hat{\mathbf{U}}_\mathrm{f}^{(i)}}, \quad \tilde{\mathbf{K}}_\mathrm{ff}^{(i)} \in \mathbb{R}^{\hat{f}^{(i)}\times \hat{f}^{(i)}}$
3: **end for**
4: $\mathbf{A} \longleftarrow \hat{\mathbf{S}} - \check{\mathbf{S}}_\mathrm{f}\check{\mathbf{X}}_\mathrm{f}$
5: $\mathbf{B} \longleftarrow \check{\mathbf{S}}_\mathrm{f}\check{\mathbf{V}}_\mathrm{f} - \check{\mathbf{S}}_\mathrm{p}\check{\mathbf{U}}_\mathrm{p}$
6: **return** $\frac{\partial g}{\partial x_k} \longleftarrow \sum_{i\in\mathbb{A}}\sum_{j=1}^{l^{(i)}} \left(\mathbf{A}\hat{\boldsymbol{\lambda}}_j^{(i)}\right) \cdot \frac{\partial \mathbf{K}}{\partial x_k}\left(\mathbf{b}_j^{(i)} - \mathbf{A}\hat{\mathbf{u}}_j^{(i)}\right)$

5.4 Computational effort

This section describes the computational effort involved in using the condensation approach compared to the elementary approach. First, predominant factors that influence the computational effort are discussed. Subsequently, the computational effort of both approaches is estimated using algorithmic complexity analysis, which yields an estimate of the computational efficiency of using the condensation approach with respect to the elementary approach as a function of the predominant factors.

5.4.1 Predominant factors

A critical factor influencing the computational effort of the condensation approach is the number of free secondary DOFs versus number of primary DOFs. The number of free secondary DOFs $\check{f}$ depends on the system size n, the number of prescribed secondary DOFs $\check{p}$ and the number of primary DOFs m via $\check{f} = n - \check{p} - m$. The number of free secondary DOFs $\check{f}$ is the primary contributor to the total effort of calculating the Schur complement,

as it directly influences the cost of the preprocessing of $\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}$. The number of primary DOFs m determines the number of right-hand-sides of the system of linear equations solved in the static condensation process. If the number of primary DOFs increases, the number of secondary DOFs decreases with equal amount. In addition, the number of primary DOFs m influences the computational effort of the subsequent analysis sets based upon $\tilde{\mathbf{K}}$. The computational effort of subsequent analysis sets increases rapidly as m increases, since $\tilde{\mathbf{K}}$ is a dense system matrix as compared to the generally sparse system matrix $\mathbf{K}$.

The similarity between analysis sets or, equivalently, the amount of overlap of sets $\mathbb{F}^{(i)}$ for all analysis sets is crucial. The higher the overlap of those sets, the smaller $\mathbb{M}$ can be, and hence, the lower becomes the computational effort of both the static condensation and subsequent analysis sets. The presence of non-zero prescribed values on the secondary DOFs, *i.e.* $\overset{\vee}{\mathbf{U}}_{\mathrm{p}} \neq \mathbf{0}$ and/or non-zero applied loads on the secondary DOFs $\overset{\vee}{\mathbf{F}}_{\mathrm{f}} \neq \mathbf{0}$ introduces l extra right-hand-sides for the static condensation and introduces a reduced load term. In that case, the number of prescribed secondary DOFs $\overset{\vee}{p}$ influences the effort of matrix-vector product $\overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}$ and $\overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}$ if $\overset{\vee}{\mathbf{U}}_{\mathrm{p}} \neq \mathbf{0}$, and influences $\overset{\vee}{f}$ via $\overset{\vee}{f} = n - m - \overset{\vee}{p}$. However, for the majority of large scale problems $\overset{\vee}{p} \ll n$ and the additional effort becomes negligible.

5

Another clearly important factor influencing the computational effort is the number of analysis sets a. The computational effort of the elementary approach increases linearly with a. In contrast, in the condensation approach a has little influence on the computational effort if $m \ll n$, as analysis sets based upon $\tilde{\mathbf{K}}$ are—in that case—of negligible effort. In contrast, when $m \longrightarrow n$ both the static condensation and all analysis sets based upon $\tilde{\mathbf{K}}$ become expensive, due to the high number of right-hand-sides in the static condensation and the large dense matrix $\tilde{\mathbf{K}}$. For the elementary approach, most important is the number of free DOFs for each analysis set $f^{(i)}$, which determines the effort of the preprocessing of $\mathbf{K}^{(i)}_{\mathrm{ff}}$. Note that generally $f^{(i)} \approx f^{(j)}$ for all $i, j \in \mathbb{A}$. The number of loads $l^{(i)}$ for each analysis set will have a positive influence on the efficiency of the condensation approach with respect to the elementary approach for $m \ll n$, but *vice versa* if $m \longrightarrow n$.

Also the sparsity, condition number, bandedness and band structure of $\mathbf{K}$ influence the type and cost of preprocessing and, hence, to a lesser extent impact the computational effort of both approaches. Such properties mostly depend on the type of governing equations, but are also influenced by the discretization (*e.g.* node numbering and reordering) and dimension (*e.g.* band structure and sparsity are different for 2D and 3D problems). Lastly, implementation-related factors play a role, *e.g.* type of solver, compiled vs. interpreted code, vectorization and parallelization. Note that some of the dominant parameters mostly affect the computational effort of the condensation approach (such as the number of analyses a or dimensionality of $\tilde{\mathbf{K}}$), whereas others will have a comparable influence on both approaches, *e.g.* the type of solver or specific implementation.

5.4.2 Prognosis

Based on the previous discussion we expect the computational efficiency of the condensation approach with respect to the elementary approach to be high when $m \ll n$ and to increase linearly with a. Assuming $m \ll n$ and $a \geq 1$, we expect the source of computational efficiency of the condensation approach with respect to the elementary approach to be twofold:

1. The condensation approach requires a single preprocessing of the system matrix, as opposed to a for the elementary approach.[4] All subsequent analysis sets based on the reduced system matrix have relatively negligible computational effort. Thus, the more analysis sets are based upon $\tilde{\mathbf{K}}$, the higher will be the computational efficiency with respect to the elementary approach.

2. As described in Section 5.3.1, the sensitivities of the reduced system matrix are self-adjoint, independent of the response. All information required to compute those sensitivities has been calculated during the condensation process. What remains is the computational effort of the adjoint problems based on the reduced system matrix. The problem may still consist of multiple non-self-adjoint responses, but when $m \ll n$, the associated adjoint problems are negligibly small compared to the condensation process required for the response evaluation. The sensitivity analysis of any response depending solely on the reduced system matrix is therefore of negligible computational effort.

For a fixed number of analysis sets a, an increase in m causes an increase in computational efficiency due to the low cost of repetitive analysis sets based upon the reduced system matrix in contrast to the high cost of each extra analysis set in the elementary approach. However, when m becomes large, the static condensation becomes computationally more demanding. Of even more importance is the increase in dimensionality of the reduced system matrix, which is a dense matrix. Based on the above, it is expected that the gain is higher than one (meaning computational effort of condensation approach is less than the effort of the elementary approach) for any problem with more than one analysis a and any ratio of $\eta := \frac{m}{n}$ that satisfies $0 < \underline{\eta} \leq \eta \leq \overline{\eta} < 1$, with $\underline{\eta}$ and $\overline{\eta}$ the bounds on η for which positive gain is expected.

5.4.3 Algorithmic complexity analysis

To fairly compare the computational efficiency of the condensation approach to the elementary approach, we investigate the computational gain Ξ, defined as the relative decrease in computational effort through the use of the condensation approach compared to the elementary approach, as a function of the predominant parameters. We define the theoretical computational effort as the number of floating point operations (FLOPs) β required and assume that the relative FLOP count of the two algorithms is an approximation that gives sufficient accuracy to estimate the resulting relative Central Processing Unit (CPU) time t. In what follows it will be shown that this assumption is sufficiently valid under the chosen conditions to get a simple order of magnitude estimator. The computational gain can be based upon the theoretical estimated FLOP count, denoted by $\Xi_{\tilde{\beta}}$ or, alternatively, based on experimentally measured run time, hereafter denoted by Ξ_t. First, this section presents an algorithmic complexity analysis, followed by a numerical study in Section 5.5. The overset of $\tilde{\beta}$ with a tilde emphasizes, in contrast to the exact FLOP count β, that the FLOP count is an estimation. The goal of the theoretical estimation is to have a 'back-of-the-envelope' measure of the *order of magnitude gain one may expect.* Considering the number and severity of the assumptions that will be required to get a reasonably simple analytical

[4] Note that in the elementary approach, if using a direct solution method, one also has to store a system matrices of notable size in memory, as opposed to a single matrix in the condensation approach.

formula for the estimation, the resulting formulae is by no means expected to be an exact computational effort predictor.

Note that computational gain is measured in relative FLOPs or relative runtime, whereas—in some cases—storage considerations might be critical. A thorough investigation of memory storage (as well as parallelization) of the method is considered out of scope of the present work, nonetheless some critical observations can be found in Section 5.6.

Assumptions on FLOP count of predominant operations

The FLOP count for the operations involved in the condensation and elementary approach are based upon the following assumptions:

- Arithmetic operations with individual elements have computational complexity $\mathcal{O}(1)$, as is the case with fixed-precision floating-point arithmetic. CPU time is equal for all individual operations (addition, multiplication etc.).
- Linear systems of equations are either solved using direct (factorization plus back-substitution) or iterative (preconditioning plus iterative solve) solution methods.
- The computational effort is dominated by the preprocessing and solve. Thus, all other operations are of negligible cost. In other words, we assume problems to be sufficiently large, that is $n \gg 1$.
- The properties (sparsity, symmetry, band density, definiteness, band structure etc.) of all system matrices are known. Based on this the solution method is chosen, as well as specific implementation(s).
- To calculate the gain $\Xi_{\tilde{\beta}}$ the following are assumed to be known: problem size n and number of analysis sets a as well as the index sets $\mathbb{F}^{(i)}$ and $\mathbb{P}^{(i)}$, indices of the DOFs of interest $\mathbb{M}^{(i)}$, prescribed state $\mathbf{U}_{\mathrm{p}}^{(i)}$ and applied load $\mathbf{F}_{\mathrm{f}}^{(i)}$ for each analysis set $i \in \mathbb{A}$.

The accuracy of $\Xi_{\tilde{\beta}}$ with respect to Ξ_{t} under the given assumptions, as well as the range of validity will be shown by numerical experiments in Section 5.5.1.

Theoretical approximated gain

An estimation of the FLOP count for solving a system of equations can, in general, be described by a function of the form

$$\tilde{\beta}[n, l] = \gamma[n] + \epsilon[n] \times l, \tag{5.51}$$

with n the system size, l the number of analyses, and $\gamma[n]$ and $\varepsilon[n]$ system size dependent functions that are unique to the solution method. Here $\gamma[n]$ is the FLOP count of the preprocessing, and $\varepsilon[n]$ the count per solve. Using such functions, we can approximate the expected gain of using the condensation approach compared to the elementary approach for a MPOP by

$$\Xi_{\tilde{\beta}}[n, m, a, l, b] \approx \frac{\sum_i^a \tilde{\beta}_{\mathrm{s}}[n, l^{(i)} + b^{(i)}]}{\tilde{\beta}_{\mathrm{s}}[n - m, m] + \sum_i^a \tilde{\beta}_{\mathrm{d}}[m, l^{(i)} + b^{(i)}]}, \tag{5.52}$$

where the subscripts 's' and 'd' refer to the solution method specific functions for solving a sparse and dense system of equations, respectively. The parameter $b^{(i)}$ denotes the number of non-zero right-hand-sides for the adjoint solve of analysis set i. The gain is based upon the response evaluation *and* sensitivity analysis, where we assume the preprocessing is reused where possible. In addition, we assumed that for a sufficiently large problems the ratio of prescribed DOFs to free DOFs becomes so small, that it is valid to assume the size of the analysis equals the system size, that is $f^{(i)} \approx n$, $\hat{f}^{(i)} \approx m$ for all $i \in \mathbb{A}$.

The ratio between the coefficients, that is $\chi[n] := \frac{\gamma[n]}{\epsilon[n]}$, is a predominant parameter that differentiates different solution methods, in the context of estimating the computational effort of problems for which the effort is dominated by linear solves. Direct solution methods are generally characterized with a high χ. For extremely high χ, such as direct computation of the inverse system matrix, the gain estimation becomes independent of the number of analyses. In that case, the gain is fully defined by the number and size of the analysis sets, that is Equation (5.52) is approximated as

$$\Xi_{\tilde{\beta}}[n, m, a] \approx \frac{a \times \gamma_{\mathrm{s}}[n]}{\gamma_{\mathrm{s}}[n-m] + a \times \gamma_{\mathrm{d}}[m]}, \tag{5.53}$$

with γ_{s} and γ_{d} the coefficients determining the preprocessing FLOP count for solving a sparse and dense system of equations, respectively.

Preconditioned iterative methods characteristically have a substantial lower χ. For solution methods with an extremely low χ, such as Jacobi preconditioned iterative methods, the gain estimation is dominated by the number of analyses irrespective of the number of analysis sets. That is, Equation (5.52) may reasonably be approximated by

$$\Xi_{\tilde{\beta}}[n, m, l, b] \approx \frac{(l+b) \times \epsilon_{\mathrm{s}}[n]}{m \times \epsilon_{\mathrm{s}}[n-m] + (l+b) \times \epsilon_{\mathrm{d}}[m]}, \tag{5.54}$$

with ϵ_{s} and ϵ_{d} the coefficients determining the FLOP count of solving a sparse and dense system of equations, respectively. For illustrational purpose, we assumed in both cases that the extreme values of χ pertained to the solution methods of both sparse and dense solvers.

Theoretical complexity of solution methods

The Cholesky factorization (CHOL) with subsequent back-substitutions is the most commonly applied direct solution method (both for sparse and dense matrices) to solve symmetric positive definite system matrices of small to moderate size. The FLOP count of a CHOL (without reordering) and subsequent back-substitutions of a symmetric sparse banded system matrix depends on dimensionality n and bandwidth k. Note that the bandwidth for discretized PDEs depends on the number of dimensions, *e.g.* in 2D $k = \sqrt{n}$, whereas in 3D $k = n^{\frac{2}{3}}$ (Yano *et al.* 2012). The dimensionality has a slight impact on the cost of the factorization; the factorization is relatively more expensive for 3D problems. Given the difference in number of factorizations between the condensation and elementary approach—at least for the the solution method presented—higher estimated gains are expected for 3D problems. In this work we will—without loss of generality—study 2D examples, and hence use the function coefficients of CHOL as provided in Table 5.2.

One of the most common iterative solution methods to solve symmetric positive definite system of equations is to first construct an incomplete Cholesky factorization (ICHOL)

Table 5.2: Characteristic function coefficients of a variety of commonly applied solution methods (Amir *et al.* 2014; Boyd *et al.* 2004; Davis 2006; Saad 2003; Yano *et al.* 2012). The subscripts 'd' and 's' refer to the dense and sparse implementations, respectively. Both iCHOL-CG and MG-CG have negligible preprocessing effort.

	$\gamma[n]$	$\epsilon[n]$	$\chi[n]$
CHOL_d	$\frac{1}{3}n^3$	$2n$	$\frac{1}{6}n^2$
CHOL_s	n^2	$2n^{\frac{3}{2}}$	$\frac{1}{2}\sqrt{n}$
iCHOL-PCG		$2n^2$ (2D)	
MG-PCG		$\propto n$	

preconditioner and then find the solution using a Conjugate Gradient (CG) algorithm (Hestenes *et al.* 1952). The computational cost of the preconditioning (for simple preconditioners) is assumed negligible compared to the cost of the iterative solve. Therefore, in what follows, this contribution is neglected. This assumption also emphasizes the usability of the method for a wide range of solution methods that meet this characteristic. Each CG iteration involves multiple operations, of which one matrix-vector product is dominating. For a k-banded matrix, the total FLOP count per iteration is thus $2nk$ (Saad 2003). Thus, for the same system size n, 3D problems will have a higher FLOP count per iteration. The resulting additional computational effort is expected to have a slight impact on the computational gain for high ratios of η. The required number of CG iterations is approximately proportional to $\sqrt{n}$, see Table 5.2.

In modern large-scale topology optimization implementations Krylov subspace iterative solvers, in particular MultiGrid Preconditioned Conjugate Gradient (MG-PCG) methods, are commonly applied. Such solvers provide mesh-independent convergence as well as good parallel scalability (Amir *et al.* 2014). Similar to iCHOL-CG methods, the preprocessing costs are relatively negligible. However, in contrast, the number of CG iterations is independent of the system size. As a result, MG-PCG can provide a solution of a linear system for a computational cost proportional to n, see Table 5.2. Assuming the same solution method to solve the dense problem, the MG-PCG is will have similar gain for low to moderate η (as the cost of solving the dense system is negligible). For high η the advantageous scaling of MG-PCG (relative to iCHOL-CG) is expected to have a slight negative impact on the gain of MG-PCG compared to iCHOL-CG, as the increasing cost of solving the dense system will have a larger influence.

Note that in both functions $\tilde{\beta}_\text{s}[n,l]$ and $\tilde{\beta}_\text{d}[n,l]$ and coefficients therein depend on the type of solver, implementation and computer hardware, and are thus subjected to change. As such, the provided coefficients in Table 5.2 are not expected to be accurate; the values are intended to provide insight on the influence of the ratio η to the gain estimation for different solution methods.

In order to show the general applicability, but keep the number of numerical examples to a minimum, the estimation is verified on numerical examples using two solution strategies with substantially different χ to solve the system of equations. To represent problems of moderate size we first consider solving both sparse and dense system of equations using a direct solution method. To represent large-scale to very large-scale problems (in which memory storage and communication requirements are dominant), we employ an approach

in which the sparse system of equations are solved using an iterative solution method. Considering the simplicity of implementation and ease of replication of results we use the iICHOL-CG. For the dense systems, the direct solution method remains the preferred choice.

5.5 Numerical examples

This section demonstrates the computational efficiency of the condensation approach for two different example problems, each representative for a specific type of MPOP.

We consider the following distinct examples:

1. A self-adjoint thermal conductivity MPOP where, by definition of the problem, the number of analysis sets equals the number of DOFs of interest, that is $a = m$, and each additional DOF of interest introduces an additional analysis set based on the reduced system matrix. This problem is representative for structural optimization problems with uncertainties in the use case or boundary conditions.

2. A non-self-adjoint displacement-based compliant mechanism design MPOP. The problem emphasizes the effect of reduction of computational effort of the sensitivity analysis. This example is representative for problems in which multiple response functions (requirements on behaviour from different use cases) all depend on a limited set of DOFs of interest.

These examples are defined to easy allow variations of parameters to emphasize the validation of the efficiency improvements. Despite the conceptual description, the examples are representative for practical structural optimization problems.

Both examples are topology optimization problems, parametrized in a 2D structured grid of N bilinear quadrilaterals with a single design variable per element. The constrained nonlinear optimization problem is defined as

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & g_0[\mathbf{x}] \\ \text{subject to} \quad & g_j[\mathbf{x}] \leq 0, \quad j = 1, \ldots, m \\ & \mathbf{x} \in \mathbb{X}^N \end{aligned} \tag{5.55}$$

where $\mathbf{x}$ is the field of design variables, $\mathbb{X} := \{x_i \in \mathbb{R} \mid 0 < x_i \leq 1\}$. The design variable field is blurred using a filter (Bruns *et al.* 2001) with a radius of $r = 2.0$ finite elements. The element conductivity/stiffness is related to the filtered design variable via the modified SIMP[5] interpolation function (Sigmund 2007) with a constant penalty value of $p = 3.0$.

5.5.1 Problem 1

Consider the well known heat conduction topology optimization problem in a two-dimensional domain (Bendsøe *et al.* 2004). We investigate a variation of this problem with multiple analysis sets, by considering efficient heat conduction between m randomly distributed DOFs. In each of the a analysis sets, a single DOF acts as a heat sink (temperature prescribed

[5] We omit further explanation as these terms, arguments and implications are considered common knowledge within the field.

to zero), and a heat load of random magnitude (between $0\,\mathrm{J\,s^{-1}}$ and $1\,\mathrm{J\,s^{-1}}$) is applied to the $m-1$ remaining primary DOFs.

Following the drawing conventions as previously used, a representation of this problem is sketched in Figure 5.6. For this MPOP, the number of primary DOFs is equal to the number of distinct analysis sets, *i.e.* $m = a$. The aim is to find a design that efficiently conducts heat between all primary DOFs, for each considered scenario. Thereto, we minimize the overall conductive resistance for all considered analyses, that is

$$g_0\left[\mathbf{U}^{(1)}[\mathbf{x}],...,\mathbf{U}^{(a)}[\mathbf{x}]\right] = \sum_{i\in\mathbb{A}}\sum_{j}^{l^{(i)}} \mathbf{u}_j^{(i)}\cdot\mathbf{K}[\mathbf{x}]\,\mathbf{u}_j^{(i)} \tag{5.56}$$

for the elementary approach and

$$g_0\left[\hat{\mathbf{U}}^{(1)}[\mathbf{x}],...,\hat{\mathbf{U}}^{(a)}[\mathbf{x}]\right] = \sum_{i\in\mathbb{A}}\sum_{j}^{l^{(i)}} \hat{\mathbf{u}}_j^{(i)}\cdot\tilde{\mathbf{K}}[\mathbf{x}]\,\hat{\mathbf{u}}_j^{(i)} \tag{5.57}$$

for the condensation approach. The material usage is constrained as

$$g_1 = \sum_{i}^{N}\frac{\tilde{x}_i}{N\overline{v}} - 1, \tag{5.58}$$

where $\overline{v}$ is the maximum ratio of solid to void material based upon the filtered field $\tilde{\mathbf{x}}$. An example of a resulting topology for $m = 100$ is shown in Figure 5.7, which can be reproduced using the attached MATLAB code.

Computational efficiency

Although the choice of approach, *i.e.* elementary or condensation-based, has no influence on the resulting topology or design performance, it does influence the computational effort. For this specific MPOP $l^{(i)} = m-1$ for all $i \in \mathbb{A}$ and $\frac{\mathrm{d}g_0}{\mathrm{d}\mathbf{U}^{(i)}} = \mathbf{F}^{(i)}$. As a result, we can simply write $\mathbf{\Lambda}^{(i)} = \mathbf{U}^{(i)}$ and fully equivalent for the condensation approach $\hat{\mathbf{\Lambda}}^{(i)} = \hat{\mathbf{U}}^{(i)}$. Hence, the problem is self-adjoint and there is no need to solve the additional system of equations of Equation (5.48). Therefore, $b^{(i)} = 0$ for all $i \in \mathbb{A}$. For the proposed optimization problem, Equation (5.52) can thereto be simplified to

$$\Xi_{\tilde{\beta}}[n,m] \approx \frac{\sum_i^m \tilde{\beta}_\mathrm{s}[n,m-1]}{\tilde{\beta}_\mathrm{s}[n-m,m] + \sum_i^m \tilde{\beta}_\mathrm{d}[m,m-1]}. \tag{5.59}$$

The computational gain $\Xi_{\tilde{\beta}}$ as a function of n and m is plotted in Figures 5.8a and 5.8b for direct and iterative solution methods respectively. The contour graphs show the condensation approach clearly outperforms the elementary approach, especially for large systems and small to moderate m. For this type of MPOP, a problem with many DOFs and a low ratio of m to n can be solved up to *1000 times faster* using the condensation approach independent of the chosen solution method. Figures 5.9a and 5.9b give a clear view of the gain as a function of m for different system sizes using direct and iterative solution methods, respectively. In addition to the gain as a function of FLOP count $\Xi_{\tilde{\beta}}$, the gain obtained via

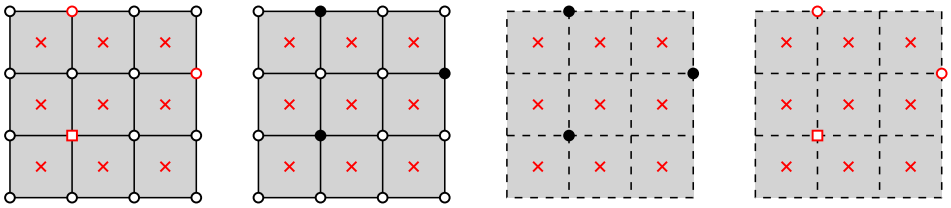


Figure 5.6: Schematic representation of Problem 1 for $N = 9$ and $m = 3$. The DOFs of interest are chosen arbitrarily. From left to right: one of the three analyses as solved for in the elementary approach, subdivision into primary and secondary DOFs, resulting system after static condensation and, the equivalent analysis as solved for in the condensation approach. The DOFs with prescribed states are denoted by a square (□), the DOFs with free states by a circle (○) and primary DOFs by a black dot (•) (analogously to Figures 5.2 and 5.3). DOFs of interest are coloured in red (○, □).

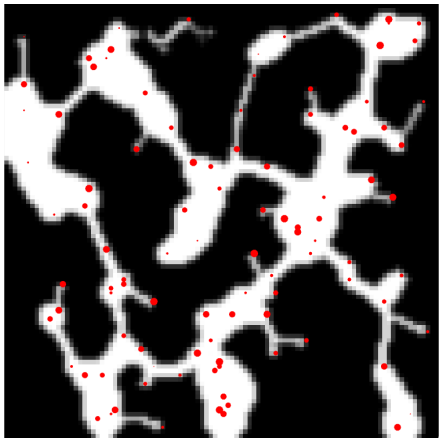
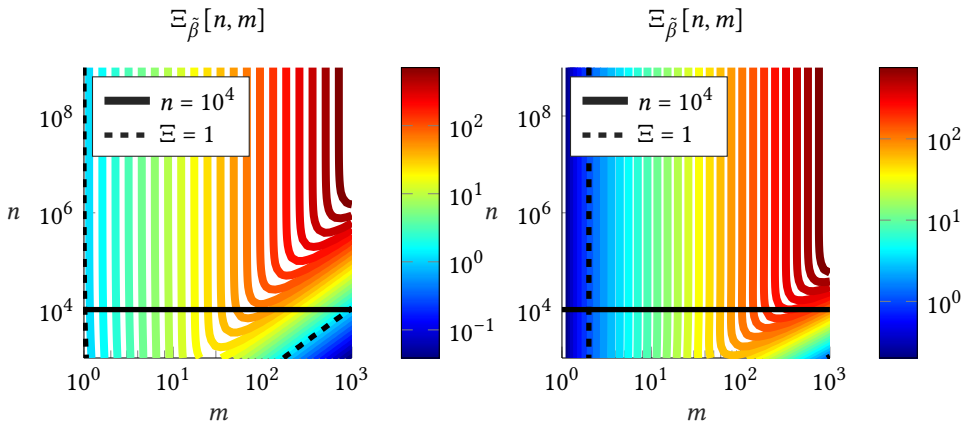


Figure 5.7: Resulting topology of Problem 1 as schematically sketched in Figure 5.6. The white material is highly conductive, whereas the black material has low conductivity. For this MPOP the number of variables $N = 100 \times 100$, the number of structural DOFs $n = 10000$, and the number of primary DOFs and analyses $m = a = 100$. The maximum amount of material usage $\overline{v} = 0.2$. The location and magnitude (size of red dot) of the primary DOFs and applied loads are indicated in red.



(a) Direct solution method.

(b) Iterative solution method.

Figure 5.8: Theoretical estimated gain of the condensation approach compared to the elementary approach for Problem 1. Figures 5.8a and 5.8b show a triple-log contour plot of the gain $\Xi_{\tilde{\beta}}$ as a function of system size n and number of DOFs of interest m for the direct (a) and iterative (b) solution methods respectively.

5

a time measurement[6] for $n = 10^4$ is plotted[7]. One may observe that Ξ_t slightly deviates from $\Xi_{\tilde{\beta}}$. Such deviation may reasonably be expected considering the severity and number of assumptions, as well as the uncertainties on the efficiency of the software, compilers and hardware used. As such, the analytical formulae, in its turn, are not expected to be an exact predictor of the results obtained by the time measurements. Despite the deviations, the trend is captured. In line with the prognosis; there is an increasing gain up to some point, after which the gain decreases. This confirms that Equation (5.52) can be used as a 'order of magnitude' estimator for the expected gain of using the condensation approach. In the following example, we will, thereto, utilize the analytical formulae without further verification.

5.5.2 Problem 2

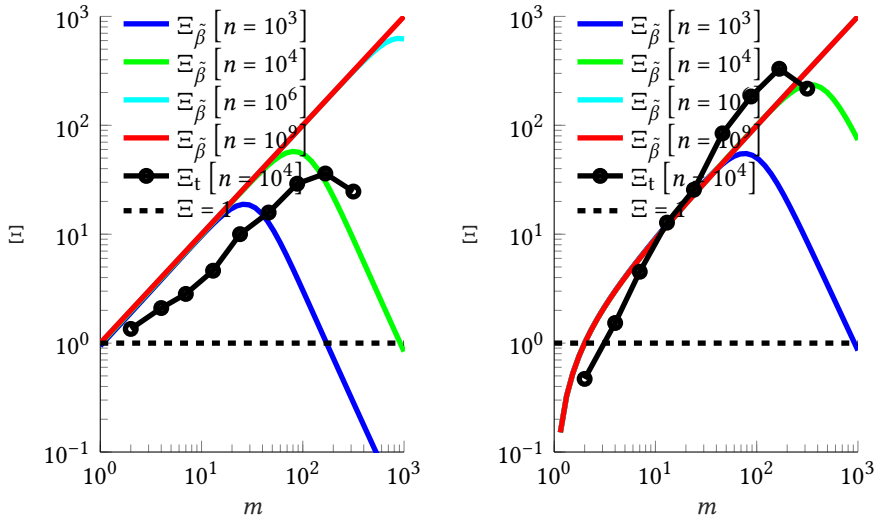
The second problem aims to maximize the amount of material required to achieve a displacement-based x-input x-output compliant mechanism satisfying prescribed input-output transmission ratios, see *e.g.* (Larsen *et al.* 1997). The problem is sketched in Figure 5.10. This second problem emphasizes the effect of reduction of computational effort of the sensitivity analysis.

The objective function for this MPOP is

$$g_0 = \sum_i^N \tilde{x}_i. \tag{5.60}$$

[6] The gain based on time measurements, equivalently to the gain based on FLOP count, only considers the time of solving the systems of equations.

[7] This problem size is considered sufficiently large to give a representative measure, without requiring unnecessary computation cost of the verification.



(a) Direct solution method.

(b) Iterative solution method.

Figure 5.9: Theoretical estimated and experimental gain of the condensation approach compared to the elementary approach for Problem 1. Figures 5.9a and 5.9b show double-log graphs of the gain as a function of the number of primary DOFs m for a variety of system sizes. The gain based on FLOP count is verified by calculating the gain based on CPU time as shown in Figures 5.9a and 5.9b.

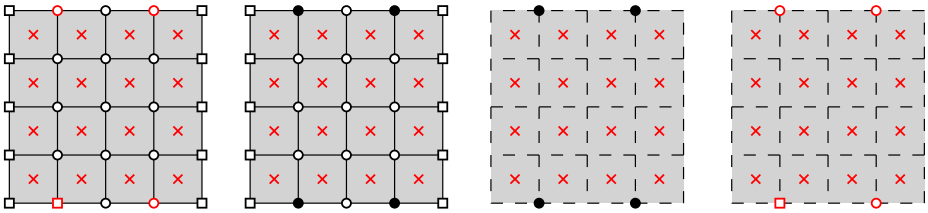
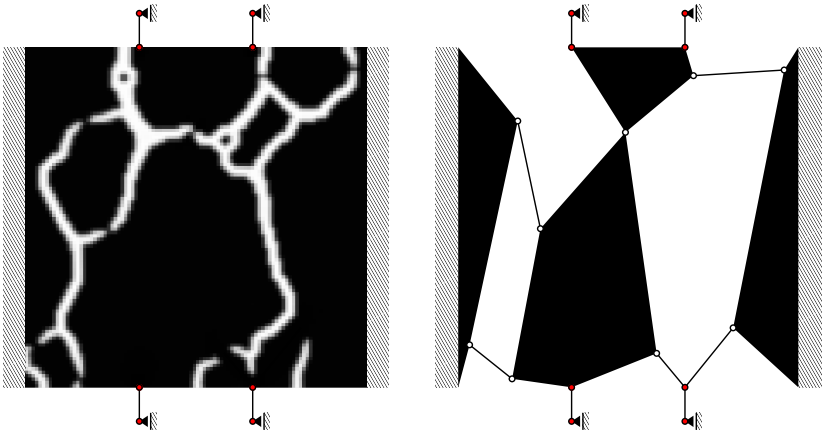


Figure 5.10: Schematic representation of Problem 2; a multi-input-multi-output compliant mechanism with $N = 16$, the number of inputs $x = 2$ and $m = 2x = 4$. The number of DOFs per node is 2, however, for simplicity of visualization, only a single DOF is shown. From left to right (analogous to Figure 5.6): one of the two analyses as solved for in the elementary approach, subdivision into primary and secondary DOFs, resulting system after static condensation and, the equivalent analysis as solved for in the condensation approach. Note that the displacements of both sides are fixed.



(a) Topology optimization material distribution. The white material is void, whereas the black material is solid.

(b) Rigid-body mechanism interpretation of the material distribution of Figure 5.11a. Rigid bodies are indicated by black bodies, rigid links by black lines and joints by circles.

Figure 5.11: Resulting topology and rigid-body mechanism interpretation of the MPOP posed in Section 5.5.2 as schematically sketched in Figure 5.10. For this MPOP the number of variables $N = 100 \times 100$, the number of structural DOFs $n = 10000$, the number of analysis sets $a = 2$ and the number of primary DOFs $m = 2x = 4$. The location of primary DOFs and applied loads are indicated in red. This multi-input-multi-output compliant mechanism is optimized for a Jacobian matrix $\bar{\mathbf{J}} = [0.5\ 2.0;\ 1.0\ -1.0]$.

One can write the mechanism output displacements as a function of the input displacements as

$$\mathbf{u}_i = \mathbf{J}_{ij}\mathbf{u}_j, \quad \forall\ i,j = 1,...,x \tag{5.61}$$

where $\mathbf{J}$ is the so-called Jacobian matrix, i denotes the output and j the input. For example, $\mathbf{J}_{12}$ is the displacement of Output 1 due to an input displacement on Input 2. One requires $a = x$ analyses with prescribed unit input displacements to obtain $\mathbf{J}$. In order to control all entries in the Jacobian matrix, we introduce $2x$ transmission constraints, *i.e.*

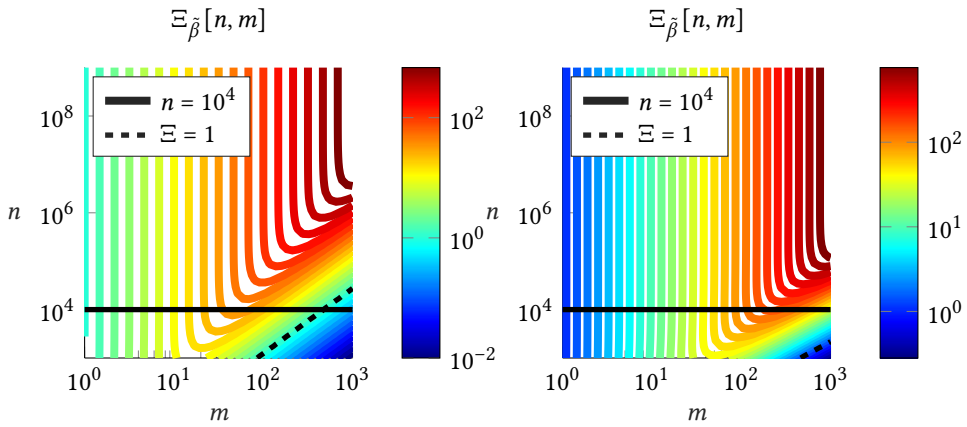
$$g_{ij} = \frac{\mathbf{J}_{ij}}{\bar{\mathbf{J}}_{ij}} + 1, \tag{5.62}$$

with $\bar{\mathbf{J}}_{ij} \neq 0$ the minimum transmission between input i and output j. Note this response can be used for both positive and negative ratios. If one assigns $\bar{\mathbf{J}}_{ij} = u^*$, then the required $\mathbf{J}_{ij} \leq -u^*$ and if $\bar{\mathbf{J}}_{ij} = -u^*$, then $\mathbf{J}_{ij} \geq u^*$, for any target displacement $u^* \in \mathbb{R}^+$.

An example of the resulting design for $x = 2$ and

$$\bar{\mathbf{J}} = \begin{bmatrix} 0.5 & 2.0 \\ 1.0 & -1.0 \end{bmatrix}$$

is shown in Figure 5.11. Due to the conflicting behaviour between volume maximization and minimum transmission ratio the resulting topology almost exactly satisfies the target Jacobian with a minimum amount of holes. To get an idea of the working principle,



(a) Direct solution method.

(b) Iterative solution method.

Figure 5.12: Theoretical estimated gain of the condensation approach compared to the elementary approach for Problem 2. Figures 5.12a and 5.12b show a triple-log contour plot of the gain $\Xi_{\tilde{\beta}}$ as a function of system size n and number of DOFs of interest m for direct (a) and iterative (b) solution methods respectively.

Figure 5.11b provides a rigid-body mechanism interpretation of the material distribution of Figure 5.11a. One may vastly observe the desired behaviour is restricted to small displacements.

Computational efficiency

For this specific MPOP $l^{(i)} = 1$ for all $i \in \mathbb{A}$. Since the problem is non-self-adjoint an additional x right-hand-sides per analysis has to be solved for, that is $b^{(i)} = x$ for all $i \in \mathbb{A}$. Note that the number of DOFs of interest equals twice the number of inputs[8], *i.e.* $m = 2x$. For the proposed optimization formulation, Equation (5.52) can therefore be simplified to

$$\Xi_{\tilde{\beta}}[n, m] \approx \frac{\sum_i^{\frac{m}{2}} \tilde{\beta}_{\mathrm{s}}[n, m]}{\tilde{\beta}_{\mathrm{s}}[n - m, m] + \sum_i^{\frac{m}{2}} \tilde{\beta}_{\mathrm{d}}[m, m]}, \tag{5.63}$$

with $m = 2x$ and x the number of inputs and outputs. Note the computational gain for this type of problem is twofold. One requires a single analysis set with $2x$ analyses, as opposed to x analysis sets to obtain the responses plus an additional x analysis sets to obtain the sensitivities. Simple reasoning as well as the FLOP count analysis learns that for this type of problem the gain approximately equals the number of inputs (and thus the number of analysis sets) for low to moderate number of inputs, see Figure 5.12.

5.6 Discussion

The proposed method is applicable to any linear(ized) PDE as well as multi-physical coupled problems and can be extended or adapted to a range of other problems. The method is

[8] Assuming the number of inputs and outputs are equal

still applicable to asymmetric system matrices. The advantageous property of inexpensive sensitivities vanishes for asymmetric matrices since the self-adjointness property no longer holds, hence the method is expected to reduce efficiency. The method is applicable to reduce the dimensionality of eigen-problems, with tremendous reduction in computational cost. However, in the reduced mass matrix, combinations of stiffness and mass elements appear. The result is that the eigenvalue-eigenvector problem is closely but not exactly preserved, with scenario-dependent accuracy issues (Guyan 1965).

A response function may depend on secondary DOFs, in which case an additional large system of linear equations has to be solved to obtain the sensitivities, see Section 5.A.5. Although the additional effort is marginal since the preconditioner can be reused, it makes the use of responses dependent on secondary DOFs uneconomical. As opposed to the general guidelines presented in this work, one may want to put secondary DOFs with varying freedom per analysis set subjected to loading into the set of primary DOFs. This will increase the number of primary DOFs, however circumvents the necessity to calculate and use the reduced load. This will reduce the computational effort if the number of loaded secondary DOFs are small.

For scenarios with multiple DOFs following the same freedom *and* magnitude throughout each analysis set (for example the DOFs contained in a non-design domain), preprocessing of the system matrix by multi-point constraints[9] is recommended to limit the dimensionality of the reduced system matrix. This has no further consequences apart from implementation, possibly reduction in global dimensionality and change of sparsity pattern / band density.

For large-scale 3D problems, the memory requirements and paralellization generally become critical aspects. Using the elementary method requires storage of multiple (equal to the number of analysis sets a) large system matrices (either factorizations or preconditioners), whereas the condensation approach only has to store a single system matrix of substantial dimension. Just as the elementary method, the condensation method consists of linear solves, matrix-vector products and dot-products. Thereto, no substantial differences in terms of 'ease of parallelization' are expected. For very large-scale problems it is recommended to distribute the condensation solves, Equations (5.20) and (5.21), over multiple computing nodes (as opposed to distribute each solve over multiple nodes) to reduce the amount of communication.

5.7 Conclusion

This study aims to find a solution to efficiently solve for large-scale structural optimization problems involving multiple analyses with varying boundary conditions. With an increasing number of analyses, the computational effort to solve such problems quickly increases. Using static condensation to eliminate the DOFs that do not change freedom between analyses and are not of interest to any of the optimization responses, the effort can be highly reduced. In contrast to solving for multiple systems of equations with different sparse system matrices of high dimensionality, this allows for solving multiple systems of equations with the same sparse system matrix of high dimensionality and multiple systems of equations with different dense system matrices of low dimensionality. A profound

[9] Primary-secondary, penalty augmentation or Lagrange multiplier adjunction methods.

finding is that the sensitivities of the static condensation procedure can be calculated without solving additional systems of equations, even if the response functions are not self-adjoint. Consequently, the dimensionality of the adjoint equations is highly reduced.

Numerical examples indicate an algorithmic complexity analysis can be used to roughly estimate the expected computational effort of a broad range of problems, even if solely the major contributors are taken into account. In line with the prognosis the numerical examples demonstrate that substantial computational gains can be achieved. However, the presented examples only cover a small part of the total applicability.

It is expected that future research will reveal more applications that benefit from this method. We expect at least applications in structural and multidisciplinary optimization with many conflicting requirements based on a variety of load cases. Given the limited additional implementation effort, future work on—or including—multi-partition structural optimization problem(s) should consider a back-of-the-envelope estimation of computational gain by using the condensation approach.

5.A Sensitivity analyses

Dependencies will, after first mention, be omitted for brevity without further notice.

5.A.1 Sensitivities of the reduced system matrix

Consider a response function $g\left[\tilde{\mathbf{K}}\left[\mathbf{K}\right]\right]$, with

$$\tilde{\mathbf{K}} := \overset{\wedge\wedge}{\mathbf{K}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{f}} \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}. \tag{5.18 revisited}$$

Here the system matrix is a function of the design variables, *i.e.* $\mathbf{K}\left[\mathbf{x}\right]$. For now we are solely interested in the sensitivities of the reduced system matrix with respect to the system matrix and therefore omit this dependency from here onward. The full derivative can be written as

$$\mathrm{d}g\left[\tilde{\mathbf{K}}\left[\mathbf{K}\right]\right] = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}} : \mathrm{d}\mathbf{K}, \tag{5.31 revisited}$$

where $\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}}$ is the sensitivity of the response function with respect to $\tilde{\mathbf{K}}$ and $\mathrm{d}\mathbf{K}$ the sensitivity of the system matrix. In the following we assume both are known. The aim is to find the fourth order tensor $\frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}}$, *i.e.* the sensitivities of the reduced system matrix with respect to the system matrix.

Direct method

Direct differentiation of Equation (5.18) yields

$$\mathrm{d}\tilde{\mathbf{K}} = \mathrm{d}\overset{\wedge\wedge}{\mathbf{K}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}} \overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{f}} \mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}. \tag{5.32 revisited}$$

Here $\mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}$ can be obtained by direct differentiation of Equation (5.20), that is

$$\mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} = -\overset{\vee}{\mathbf{K}}_{\mathrm{ff}}^{-1} \mathrm{d}\overset{\vee}{\mathbf{K}}_{\mathrm{ff}} \overset{\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\vee}{\mathbf{K}}_{\mathrm{ff}}^{-1} \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}. \tag{5.33 revisited}$$

Substitution of Equation (5.33) into Equation (5.32) gives

$$\mathrm{d}\tilde{\mathbf{K}} = \mathrm{d}\overset{\wedge\wedge}{\mathbf{K}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\vee\wedge}{\mathbf{K}}_{\mathrm{f}}. \tag{5.64}$$

Substitution of Equation (5.30) into Equation (5.64) yields

$$\mathrm{d}\tilde{\mathbf{K}} = \overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\wedge}{\mathbf{S}} - \overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\wedge}{\mathbf{S}}. \tag{5.65}$$

Subsequent substitution of the result into Equation (5.31) gives the expression for the sensitivities of the response function, *i.e.*

$$\begin{aligned}\mathrm{d}g &= \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \left(\left(\overset{\wedge}{\mathbf{S}}^{\mathsf{T}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\right)\mathrm{d}\mathbf{K}\left(\overset{\wedge}{\mathbf{S}} - \overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\right)\right) \\ &= \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \left(\mathbf{A}^{\mathsf{T}}\mathrm{d}\mathbf{K}\mathbf{A}\right),\end{aligned} \tag{5.34 revisited}$$

where

$$\mathbf{A} := \overset{\wedge}{\mathbf{S}} - \overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}. \tag{5.36 revisited}$$

We can simplify[10] this to obtain

$$\mathrm{d}g = \left(\mathbf{A}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} \otimes \mathbf{A}\right) : \mathrm{d}\mathbf{K}, \tag{5.66}$$

In addition, rewriting[3] of Equation (5.34) gives

$$\frac{\mathrm{d}\tilde{\mathbf{K}}}{\mathrm{d}\mathbf{K}} = \mathbf{A} \otimes \mathbf{A}. \tag{5.37 revisited}$$

Adjoint method

Alternatively to the direct method, one can use the adjoint method (Arora *et al.* 1979; Vanderplaats 1980). After augmentation of the response function as defined in Equation (5.31) with Equation (5.20), we get

$$\mathcal{L}\left[\tilde{\mathbf{K}}[\mathbf{K}], \overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}}\right] := g\left[\tilde{\mathbf{K}}[\mathbf{K}]\right] + \overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\overset{\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\right), \tag{5.38 revisited}$$

where $\mathcal{L}\left[\tilde{\mathbf{K}}[\mathbf{K}], \overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}}\right]$ is the Lagrangian and $\overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}} \in \mathbb{R}^{\overset{\vee}{f} \times m}$ the Lagrange multipliers related to $\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}$. Full differentiation of both sides gives

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \mathrm{d}\tilde{\mathbf{K}} + \overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\mathrm{d}\overset{\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\vee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\right). \tag{5.67}$$

Substitution of Equation (5.32) into Equation (5.67) yields

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}} : \left(\mathrm{d}\overset{\wedge\wedge}{\mathbf{K}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\right) + \overset{\wedge}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\mathrm{d}\overset{\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} + \overset{\vee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\wedge}{\mathbf{X}}_{\mathrm{f}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\right). \tag{5.68}$$

[10] $\mathbf{A} : (\mathbf{BCD}) = \left(\mathbf{B}^{\mathsf{T}}\mathbf{A} \otimes \mathbf{D}\right) : \mathbf{C}$

Since $d\overset{\wedge}{\mathbf{X}}_f$ is computationally expensive to calculate, the aim is to circumvent this by a suitable choice of $\overset{\wedge}{\mathbf{\Lambda}}_f$. Collecting terms involving $d\overset{\wedge}{\mathbf{X}}_f$ gives

$$\overset{\wedge}{\mathbf{\Lambda}}_f : \left(\overset{\vee}{\mathbf{K}}_{ff} d\overset{\wedge}{\mathbf{X}}_f\right) - \frac{dg}{d\tilde{\mathbf{K}}} : \left(\overset{\wedge}{\mathbf{K}}_f d\overset{\wedge}{\mathbf{X}}_f\right). \tag{5.69}$$

This can be rewritten[11] into

$$\left(\overset{\vee}{\mathbf{K}}_{ff}\overset{\wedge}{\mathbf{\Lambda}}_f - \overset{\wedge}{\mathbf{K}}_f \frac{dg}{d\tilde{\mathbf{K}}}\right) : d\overset{\wedge}{\mathbf{X}}_f. \tag{5.70}$$

To make the terms involving $d\overset{\wedge}{\mathbf{X}}_f$ vanish, the term in parentheses must be zero. This can be achieved by choosing $\overset{\wedge}{\mathbf{\Lambda}}_f$ as

$$\overset{\wedge}{\mathbf{\Lambda}}_f = \overset{\vee}{\mathbf{K}}_{ff}^{-1}\overset{\wedge}{\mathbf{K}}_f \frac{dg}{d\tilde{\mathbf{K}}} = \overset{\wedge}{\mathbf{X}}_f \frac{dg}{d\tilde{\mathbf{K}}}. \tag{5.41 revisited}$$

Here the self-adjointness property appears; no additional solves are needed to compute $\overset{\wedge}{\mathbf{\Lambda}}_f$. With these multipliers, the terms involving $d\overset{\wedge}{\mathbf{X}}_f$ vanish from Equation (5.68), that is

$$d\mathcal{L} = \frac{dg}{d\tilde{\mathbf{K}}} : \left(d\overset{\wedge}{\tilde{\mathbf{K}}} - d\overset{\wedge}{\mathbf{K}}_f \overset{\wedge}{\mathbf{X}}_f\right) + \overset{\wedge}{\mathbf{\Lambda}}_f : \left(d\overset{\vee}{\mathbf{K}}_{ff}\overset{\wedge}{\mathbf{X}}_f - d\overset{\wedge}{\mathbf{K}}_f\right), \tag{5.71}$$

with $\overset{\wedge}{\mathbf{\Lambda}}_f$ as defined in Equation (5.41). Next, substitution of Equation (5.30) into Equation (5.71) yields

$$\begin{aligned} d\mathcal{L} &= \frac{dg}{d\tilde{\mathbf{K}}} : \left(\hat{\mathbf{S}}^T d\mathbf{K}\hat{\mathbf{S}} - \hat{\mathbf{S}}^T d\mathbf{K}\overset{\vee}{\mathbf{S}}_f\overset{\wedge}{\mathbf{X}}_f\right) + \overset{\wedge}{\mathbf{\Lambda}}_f : \left(\overset{\vee}{\mathbf{S}}_f^T d\mathbf{K}\overset{\vee}{\mathbf{S}}_f\overset{\wedge}{\mathbf{X}}_f - \overset{\vee}{\mathbf{S}}_f^T d\mathbf{K}\hat{\mathbf{S}}\right) \\ &= \frac{dg}{d\tilde{\mathbf{K}}} : \left(\hat{\mathbf{S}}^T d\mathbf{K}\mathbf{A}\right) - \overset{\wedge}{\mathbf{\Lambda}}_f : \left(\overset{\vee}{\mathbf{S}}_f^T d\mathbf{K}\mathbf{A}\right) \qquad \text{(5.39 revisited)} \\ &= \left(\hat{\mathbf{S}}\frac{dg}{d\tilde{\mathbf{K}}} - \overset{\vee}{\mathbf{S}}_f\overset{\wedge}{\mathbf{\Lambda}}_f\right)^T d\mathbf{K}\mathbf{A}. \end{aligned}$$

The reader can verify that substitution of Equation (5.41) into Equation (5.39) results in Equation (5.66).

5.A.2 Sensitivities of the reduced load

Consider a response function $g\left[\tilde{\mathbf{F}}\left[\mathbf{K}, \overset{\vee}{\mathbf{F}}_f, \overset{\vee}{\mathbf{U}}_p\right]\right]$, with

$$\tilde{\mathbf{F}}\left[\mathbf{K}, \overset{\vee}{\mathbf{F}}_f, \overset{\vee}{\mathbf{U}}_p\right] := \overset{\wedge}{\mathbf{K}}_f \overset{\vee}{\mathbf{V}}_f\left[\mathbf{K}, \overset{\vee}{\mathbf{F}}_f, \overset{\vee}{\mathbf{U}}_p\right] - \overset{\wedge}{\mathbf{K}}_p \overset{\vee}{\mathbf{U}}_p. \tag{5.19 revisited}$$

[11] $\mathbf{A} : (\mathbf{B}\mathbf{C}) = \left(\mathbf{B}^T\mathbf{A}\right) : \mathbf{C}$

By applying the chain-rule, the full derivative of this response function can be written as

$$\begin{aligned} \mathrm{d}g\left[\tilde{\mathbf{F}}\left[\mathbf{K},\overset{\vee}{\mathbf{F}}_{\mathrm{f}},\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right]\right] &= \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \mathrm{d}\tilde{\mathbf{F}}\left[\mathbf{K},\overset{\vee}{\mathbf{F}}_{\mathrm{f}},\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right] \\ &= \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \left(\frac{\partial\tilde{\mathbf{F}}}{\partial\mathbf{K}} : \mathrm{d}\mathbf{K} + \frac{\partial\tilde{\mathbf{F}}}{\partial\overset{\vee}{\mathbf{F}}_{\mathrm{f}}} : \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \frac{\partial\tilde{\mathbf{F}}}{\partial\overset{\vee}{\mathbf{U}}_{\mathrm{p}}} : \mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right). \end{aligned} \tag{5.72}$$

The aim is to find $\frac{\partial\tilde{\mathbf{F}}}{\partial\mathbf{K}}$, $\frac{\partial\tilde{\mathbf{F}}}{\partial\overset{\vee}{\mathbf{F}}_{\mathrm{f}}}$ and $\frac{\partial\tilde{\mathbf{F}}}{\partial\overset{\vee}{\mathbf{U}}_{\mathrm{p}}}$.

Direct method

Direct differentiation of both sides of Equation (5.19) yields

$$\mathrm{d}\tilde{\mathbf{F}} = \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.73}$$

Here $\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$ can be obtained by direct differentiation of Equation (5.21), that is

$$\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} = \overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}^{-1}\left(-\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right). \tag{5.74}$$

Substitution of Equation (5.74) into Equation (5.73), and using the definition of $\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$ from Equation (5.21) yields

$$\mathrm{d}\tilde{\mathbf{F}} = \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\left(-\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right) - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.75}$$

For conciseness, we define the variable

$$\mathbf{C}^{\mathsf{T}} := \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}, \tag{5.76}$$

and substitute this into Equation (5.75), which yields

$$\mathrm{d}\tilde{\mathbf{F}} = \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \mathbf{C}^{\mathsf{T}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.77}$$

Subsequent substitution of Equation (5.30) into Equation (5.77), and Equation (5.77) into Equation (5.72) gives an expression for the sensitivities of the response function, *i.e.*

$$\mathrm{d}g = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \left(\overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \mathbf{C}^{\mathsf{T}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right). \tag{5.78}$$

We can simplify[10] this to obtain

$$\mathrm{d}g = \left(\mathbf{A}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} \otimes \mathbf{B}\right) : \mathrm{d}\mathbf{K} - \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} + \mathbf{C}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}, \tag{5.79}$$

with

$$\mathbf{B} := \overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\vee}{\mathbf{S}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}. \tag{5.43 revisited}$$

And thus

$$\frac{\partial\tilde{\mathbf{F}}}{\partial\mathbf{K}} = \mathbf{A} \otimes \mathbf{B}. \tag{5.42 revisited}$$

Note that all terms in Equation (5.79) are calculated in the corresponding forward analysis, and thus no additional solves are required.

Adjoint method

After augmentation of the response function with the linear system of equations from Equation (5.21), we get

$$\mathcal{L}\left[\tilde{\mathbf{F}}\left[\mathbf{K},\overset{\vee}{\mathbf{F}}_{\mathrm{f}},\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right],\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}}\right] := g\left[\tilde{\mathbf{F}}\left[\mathbf{K},\overset{\vee}{\mathbf{F}}_{\mathrm{f}},\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right]\right] - \overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\vee}{\mathbf{F}}_{\mathrm{f}} - \overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right), \tag{5.80}$$

where $\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} \in \mathbb{R}^{\overset{\vee}{f}\times l}$ are the Lagrange multipliers related to $\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$. Differentiation of both sides gives

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \mathrm{d}\tilde{\mathbf{F}} - \overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} + \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} - \overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right). \tag{5.81}$$

Substitution of Equation (5.73) into Equation (5.81) gives

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \left(\mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right) - \overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} + \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \mathrm{d}\overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} + \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} - \overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right). \tag{5.82}$$

Since $\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$ is computationally expensive to calculate, aim is to circumvent this by a suitable choice of $\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}}$. Collecting terms involving $\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$ gives

$$\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \left(\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}\right) - \overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}\right). \tag{5.83}$$

This can be rewritten[11] into

$$\left(\overset{\wedge}{\mathbf{K}}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} - \overset{\vee\vee}{\mathbf{K}}\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}}\right) : \mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}. \tag{5.84}$$

To make the terms $\mathrm{d}\overset{\vee}{\mathbf{V}}_{\mathrm{f}}$ vanish we choose $\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}}$ as

$$\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} = \overset{\vee\vee}{\mathbf{K}}_{\mathrm{ff}}^{-1}\overset{\wedge}{\mathbf{K}}_{\mathrm{f}}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} = \overset{\wedge}{\mathbf{X}}_{\mathrm{f}}\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}}. \tag{5.85}$$

Note again that no adjoint solve is necessary. Substitution of Equation (5.30) into Equation (5.82) yields

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}} : \left(\overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\wedge}{\mathbf{S}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} - \overset{\wedge}{\mathbf{K}}_{\mathrm{p}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right) - \overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}} : \left(\overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{f}}\overset{\vee}{\mathbf{V}}_{\mathrm{f}} - \overset{\vee}{\mathbf{S}}_{\mathrm{f}}^{\mathsf{T}}\mathrm{d}\mathbf{K}\overset{\vee}{\mathbf{S}}_{\mathrm{p}}\overset{\vee}{\mathbf{U}}_{\mathrm{p}} + \mathrm{d}\overset{\vee}{\mathbf{F}}_{\mathrm{f}} - \overset{\vee\vee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right), \tag{5.86}$$

with $\overset{\vee}{\boldsymbol{\Lambda}}_{\mathrm{f}}$ as defined in Equation (5.85). The reader can verify that substitution of Equation (5.85) into Equation (5.86) results in Equation (5.79).

5.A.3 Sensitivities of the state

Often responses are a function of the state, *i.e.* $g\left[\hat{\mathbf{U}}_{\mathrm{f}}\left[\tilde{\mathbf{K}}\left[\mathbf{K}\right],\tilde{\mathbf{F}}\left[\mathbf{K},\overset{\vee}{\mathbf{F}}_{\mathrm{f}},\overset{\vee}{\mathbf{U}}_{\mathrm{p}}\right],\hat{\mathbf{F}}_{\mathrm{f}},\hat{\mathbf{U}}_{\mathrm{p}}\right]\right]$. Note that this may be the state of multiple analyses, that is $g\left[\hat{\mathbf{U}}_{\mathrm{f}}^{(1)},\hat{\mathbf{U}}_{\mathrm{f}}^{(2)},...,\hat{\mathbf{U}}_{\mathrm{f}}^{(a)}\right]$, where a is the total number of analyses based upon $\tilde{\mathbf{K}}$. Therefore, one is required to determine $\mathrm{d}\hat{\mathbf{U}}_{\mathrm{f}}^{(i)}\left[\tilde{\mathbf{K}},\tilde{\mathbf{F}},\hat{\mathbf{F}}_{\mathrm{f}}^{(i)},\hat{\mathbf{U}}_{\mathrm{p}}^{(i)}\right]$

5

for all $i \in \mathbb{A}$. For clarity, we omit the superscript i from here on (assume a single analysis set). Taking the full derivative of $g\left[\hat{\mathbf{U}}_\mathrm{f}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}, \hat{\mathbf{F}}_\mathrm{f}, \hat{\mathbf{U}}_\mathrm{p}\right]\right]$ yields

$$\mathrm{d}g = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{f} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} : \left(\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{K}}} : \mathrm{d}\tilde{\mathbf{K}} + \frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{F}}} : \mathrm{d}\tilde{\mathbf{F}} + \frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \hat{\mathbf{F}}_\mathrm{f}} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} + \frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \hat{\mathbf{U}}_\mathrm{p}} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{p} \right), \tag{5.44 revisited}$$

where we assume $\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}}$ to be known. The terms $\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{K}}}$, $\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \tilde{\mathbf{F}}}$, $\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \hat{\mathbf{F}}_\mathrm{f}}$ and $\frac{\partial \hat{\mathbf{U}}_\mathrm{f}}{\partial \hat{\mathbf{U}}_\mathrm{p}}$ are to be determined. To this end, we augment the response function with the linear system of equations from Equation (5.27), which yields

$$\mathcal{L}\left[\hat{\mathbf{U}}_\mathrm{f}, \hat{\mathbf{\Lambda}}_\mathrm{f}\right] := g\left[\hat{\mathbf{U}}_\mathrm{f}\right] - \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{F}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{fp}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{F}}_\mathrm{f} \right) \tag{5.87}$$

where $\mathcal{L}\left[\hat{\mathbf{U}}_\mathrm{f}, \hat{\mathbf{\Lambda}}_\mathrm{f}\right]$ is the Lagrangian and $\hat{\mathbf{\Lambda}}_\mathrm{f} \in \mathbb{R}^{\hat{f} \times l}$ the Lagrange multipliers related to $\hat{\mathbf{U}}_\mathrm{f}$. Full differentiation of both sides gives

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\mathrm{d}\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{ff}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f} - \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} + \mathrm{d}\tilde{\mathbf{K}}_\mathrm{fp}\hat{\mathbf{U}}_\mathrm{p} + \tilde{\mathbf{K}}_\mathrm{fp}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} - \mathrm{d}\tilde{\mathbf{F}}_\mathrm{f} \right). \tag{5.88}$$

Since $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ is computationally expensive to calculate, aim is to circumvent this by a suitable choice of $\hat{\mathbf{\Lambda}}_\mathrm{f}$. Collecting terms involving $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ gives

$$\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{\Lambda}}_\mathrm{f} : \tilde{\mathbf{K}}_\mathrm{ff}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}. \tag{5.89}$$

This can be rewritten[11] into

$$\left(\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}} - \tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{\Lambda}}_\mathrm{f} \right) : \mathrm{d}\hat{\mathbf{U}}_\mathrm{f}. \tag{5.90}$$

To make the terms $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ vanish we choose $\hat{\mathbf{\Lambda}}_\mathrm{f}$ as

$$\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{\Lambda}}_\mathrm{f} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}}. \tag{5.48 revisited}$$

Note that to obtain $\hat{\mathbf{\Lambda}}_\mathrm{f}$ one has to solve an additional system of equations. Subsequent substitution of Equation (5.30) into Equation (5.88) gives

$$\mathrm{d}\mathcal{L} = \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(-\tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\mathrm{d}\tilde{\mathbf{K}}\hat{\mathbf{U}} + \tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\mathrm{d}\tilde{\mathbf{F}} + \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} - \tilde{\mathbf{K}}_\mathrm{fp}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} \right), \tag{5.91}$$

where we also used

$$\hat{\mathbf{U}} = \tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{S}}_\mathrm{p}\hat{\mathbf{U}}_\mathrm{p}. \tag{5.92}$$

Note that Equation (5.91) can be rewritten as

$$\mathrm{d}\mathcal{L} = \left(-\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} \otimes \hat{\mathbf{U}} \right) : \mathrm{d}\tilde{\mathbf{K}} + \tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\tilde{\mathbf{F}} + \hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} - \tilde{\mathbf{K}}_\mathrm{pf}\hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{p}. \tag{5.93}$$

Substitution of Equations (5.35) and (5.77) into Equation (5.93) yields

$$\mathrm{d}\mathcal{L} = \left(\mathbf{A}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} \otimes \mathbf{D}\right) : \mathrm{d}\mathbf{K} - \overset{\vee}{\mathbf{X}}_\mathrm{f}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{f} + \mathbf{C}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{p} + \hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} - \tilde{\mathbf{K}}_\mathrm{pf}\hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{U}}_\mathrm{p}, \tag{5.94}$$

with

$$\mathbf{D} := \mathbf{B} - \mathbf{A}\hat{\mathbf{U}}. \tag{5.95}$$

Thus to obtain $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ one requires an extra solve or matrix-vector product (depending on the self-adjointness), Equation (5.48), per analysis based on the reduced system.

5.A.4 Sensitivities of the reaction load

From the solution of Equation (5.27) the reaction loads can be calculated via

$$\hat{\mathbf{F}}_\mathrm{p} = \tilde{\mathbf{K}}_\mathrm{pf}\hat{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{pp}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{F}}_\mathrm{p}, \tag{5.28 revisited}$$

Note that here we have again omitted the subscript *i*. Sensitivities of a response function $g\left[\hat{\mathbf{F}}_\mathrm{p}\left[\tilde{\mathbf{K}}, \tilde{\mathbf{F}}, \hat{\mathbf{U}}_\mathrm{f}, \hat{\mathbf{U}}_\mathrm{p}\right]\right]$ are defined as

$$\mathrm{d}g = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{p}. \tag{5.96}$$

To this end, we augment the response function with the linear system of equations from Equation (5.27), which yields

$$\mathcal{L}\left[\hat{\mathbf{F}}_\mathrm{p}, \hat{\mathbf{\Lambda}}_\mathrm{f}\right] := g\left[\hat{\mathbf{F}}_\mathrm{p}\right] - \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{F}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{fp}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{F}}_\mathrm{f}\right) \tag{5.97}$$

where $\mathcal{L}\left[\hat{\mathbf{F}}_\mathrm{p}, \hat{\mathbf{\Lambda}}_\mathrm{f}\right]$ is the Lagrangian and $\hat{\mathbf{\Lambda}}_\mathrm{f} \in \mathbb{R}^{\hat{f}\times l}$ the Lagrange multipliers related to $\hat{\mathbf{U}}_\mathrm{f}$. Full differentiation of both sides gives

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{p} - \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\mathrm{d}\tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{ff}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f} - \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} + \mathrm{d}\tilde{\mathbf{K}}_\mathrm{fp}\hat{\mathbf{U}}_\mathrm{p} + \tilde{\mathbf{K}}_\mathrm{fp}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} - \mathrm{d}\tilde{\mathbf{F}}_\mathrm{f}\right), \tag{5.98}$$

with

$$\mathrm{d}\hat{\mathbf{F}}_\mathrm{p} = \mathrm{d}\tilde{\mathbf{K}}_\mathrm{pf}\hat{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{K}}_\mathrm{pf}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f} + \mathrm{d}\tilde{\mathbf{K}}_\mathrm{pp}\hat{\mathbf{U}}_\mathrm{p} + \tilde{\mathbf{K}}_\mathrm{pp}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} - \mathrm{d}\tilde{\mathbf{F}}_\mathrm{p}. \tag{5.99}$$

Since $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ is computationally expensive to calculate, aim is to circumvent this by a suitable choice of $\hat{\mathbf{\Lambda}}_\mathrm{f}$. Collecting terms involving $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ gives

$$\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} : \tilde{\mathbf{K}}_\mathrm{pf}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{\Lambda}}_\mathrm{f} : \tilde{\mathbf{K}}_\mathrm{ff}\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}. \tag{5.100}$$

This can be rewritten[11] into

$$\left(\tilde{\mathbf{K}}_\mathrm{fp}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} - \tilde{\mathbf{K}}_\mathrm{ff}\hat{\mathbf{\Lambda}}_\mathrm{f}\right) : \mathrm{d}\hat{\mathbf{U}}_\mathrm{f}. \tag{5.101}$$

To make the terms $\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}$ vanish, the term in parentheses must be zero. This can be achieved by choosing $\hat{\mathbf{\Lambda}}_\mathrm{f}$ as

$$\tilde{\mathbf{K}}_{\mathrm{ff}}\hat{\mathbf{\Lambda}}_\mathrm{f} = \tilde{\mathbf{K}}_{\mathrm{fp}}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}}. \tag{5.102}$$

Note that to obtain $\hat{\mathbf{\Lambda}}_\mathrm{f}$ one has to solve a system of equations. Substitution of Equations (5.30), (5.92) and (5.99) into Equation (5.98) and rewriting yields

$$\mathrm{d}\mathcal{L} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} : \left(\tilde{\mathbf{S}}_\mathrm{p}^\mathsf{T}\mathrm{d}\tilde{\mathbf{K}}\hat{\mathbf{U}} + \tilde{\mathbf{K}}_{\mathrm{pp}}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{S}}_\mathrm{p}^\mathsf{T}\mathrm{d}\tilde{\mathbf{F}}\right) - \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\mathrm{d}\tilde{\mathbf{K}}\hat{\mathbf{U}} - \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} + \tilde{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\mathrm{d}\tilde{\mathbf{F}}\right) \tag{5.103}$$

Subsequent substitution of Equations (5.35) and (5.77) and simplifying using Equations (5.76) and (5.95) yields

$$\mathrm{d}g = (\mathbf{AE}\otimes\mathbf{D}) : \mathrm{d}\mathbf{K} - \check{\hat{\mathbf{X}}}_\mathrm{f}\mathbf{E} : \mathrm{d}\check{\hat{\mathbf{F}}}_\mathrm{f} + \mathbf{CE} : \mathrm{d}\check{\hat{\mathbf{U}}}_\mathrm{p} + \hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} + \left(\tilde{\mathbf{K}}_{\mathrm{pp}}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}} - \tilde{\mathbf{K}}_{\mathrm{pf}}\hat{\mathbf{\Lambda}}_\mathrm{f}\right) : \mathrm{d}\hat{\mathbf{U}}_\mathrm{p}, \tag{5.104}$$

with temporary[12] variable

$$\mathbf{E} := \tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} - \tilde{\mathbf{S}}_\mathrm{p}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}}. \tag{5.105}$$

Thus to obtain $\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}$ one requires an extra solve or matrix-vector product (depending on the self-adjointness), Equation (5.102), per analysis based on using the reduced system.

5.A.5 Sensitivities of the state on the secondary DOFs

After obtaining the solution to the state of the primary DOFs one can obtain the solution of the free secondary DOFs via

$$\check{\mathbf{U}}_\mathrm{f} = -\left(\check{\hat{\mathbf{X}}}_\mathrm{f}\hat{\mathbf{U}} + \check{\mathbf{V}}_\mathrm{f}\right). \tag{5.22 revisited}$$

We augment the response function with the linear systems of equations from Equations (5.20), (5.21) and (5.27), that is

$$\begin{aligned}\mathcal{L}\left[\check{\mathbf{U}}_\mathrm{f}, \check{\hat{\mathbf{\Lambda}}}_\mathrm{f}, \check{\mathbf{\Lambda}}_\mathrm{f}, \hat{\mathbf{\Lambda}}_\mathrm{f}\right] := g\left[\check{\mathbf{U}}_\mathrm{f}\right] &- \check{\hat{\mathbf{\Lambda}}}_\mathrm{f} : \left(\check{\tilde{\mathbf{K}}}_{\mathrm{ff}}\check{\hat{\mathbf{X}}}_\mathrm{f} - \check{\hat{\mathbf{K}}}_\mathrm{f}\right) \\ &- \check{\mathbf{\Lambda}}_\mathrm{f} : \left(\check{\tilde{\mathbf{K}}}_{\mathrm{ff}}\check{\mathbf{V}}_\mathrm{f} - \check{\tilde{\mathbf{K}}}_{\mathrm{fp}}\check{\mathbf{U}}_\mathrm{p} + \check{\mathbf{F}}_\mathrm{f}\right) \\ &- \hat{\mathbf{\Lambda}}_\mathrm{f} : \left(\tilde{\mathbf{K}}_{\mathrm{ff}}\hat{\mathbf{U}}_\mathrm{f} - \hat{\mathbf{F}}_\mathrm{f} + \tilde{\mathbf{K}}_{\mathrm{fp}}\hat{\mathbf{U}}_\mathrm{p} - \tilde{\mathbf{F}}_\mathrm{f}\right).\end{aligned} \tag{5.106}$$

[12] Only used within this section.

Full differentiation of both sides gives

$$\begin{aligned} \mathrm{d}\mathcal{L} = &- \frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}} : \left(\mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f}\overset{\wedge}{\mathbf{U}} + \overset{\wedge}{\mathbf{X}}_\mathrm{f}\mathrm{d}\overset{\wedge}{\mathbf{U}} + \mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f} \right) \\ &- \overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} : \left(\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{X}}_\mathrm{f} + \overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f} - \mathrm{d}\overset{\wedge}{\mathbf{K}}_\mathrm{f} \right) \\ &- \overset{\vee}{\mathbf{\Lambda}}_\mathrm{f} : \left(\mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{V}}_\mathrm{f} + \overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f} - \mathrm{d}\overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\overset{\vee}{\mathbf{U}}_\mathrm{p} - \overset{\curlyvee}{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{p} + \mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{f} \right) \\ &- \overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} : \left(\mathrm{d}\tilde{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{K}}_{\mathrm{ff}}\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f} - \mathrm{d}\overset{\wedge}{\mathbf{F}}_\mathrm{f} + \mathrm{d}\tilde{\mathbf{K}}_{\mathrm{fp}}\overset{\wedge}{\mathbf{U}}_\mathrm{p} + \tilde{\mathbf{K}}_{\mathrm{fp}}\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{p} - \mathrm{d}\tilde{\mathbf{F}}_\mathrm{f} \right). \end{aligned} \tag{5.107}$$

Since $\mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f}$, $\mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f}$ and $\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f}$ are computationally expensive to calculate, aim is to circumvent this by a suitable choices of $\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f}$, $\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}$ and $\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f}$. Collecting terms involving $\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f}$ gives

$$\left(-\tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\overset{\curlywedge}{\mathbf{X}}_\mathrm{f}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}} - \tilde{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} \right) : \mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f}, \tag{5.108}$$

where we used

$$\mathrm{d}\overset{\wedge}{\mathbf{U}} = \tilde{\mathbf{S}}_\mathrm{f}\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f} + \tilde{\mathbf{S}}_\mathrm{p}\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{p}. \tag{5.109}$$

To make the terms $\mathrm{d}\overset{\wedge}{\mathbf{U}}_\mathrm{f}$ vanish we define $\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f}$ as

$$\tilde{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} = -\tilde{\mathbf{S}}_\mathrm{f}^\mathsf{T}\overset{\curlywedge}{\mathbf{X}}_\mathrm{f}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}}, \tag{5.110}$$

which requires an additional solve using the condensed system. Collecting terms involving $\mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f}$ gives

$$\left(-\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}} - \overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f} \right) : \mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f}. \tag{5.111}$$

To make the terms $\mathrm{d}\overset{\vee}{\mathbf{V}}_\mathrm{f}$ vanish we define $\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}$ as

$$\overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f} = -\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}}, \tag{5.112}$$

for which an additional full system solve is needed. Finally, collecting terms involving $\mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f}$ gives

$$\left(-\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}}\overset{\wedge}{\mathbf{U}}^\mathsf{T} - \overset{\curlyvee}{\mathbf{K}}_{\mathrm{ff}}\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} \right) : \mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f}. \tag{5.113}$$

To make the terms $\mathrm{d}\overset{\wedge}{\mathbf{X}}_\mathrm{f}$ vanish we define $\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f}$ as

$$\overset{\curlyvee}{\mathbf{K}}\overset{\wedge}{\mathbf{\Lambda}}_\mathrm{f} = -\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}}\overset{\wedge}{\mathbf{U}}^\mathsf{T}. \tag{5.114}$$

Thus, with Equation (5.112)

$$\overset{\wedge}{\Lambda}_f \left[\overset{\vee}{\Lambda}_f\right] = \overset{\vee}{\Lambda}_f \hat{U}^T. \tag{5.115}$$

With all multipliers defined, substitution of Equations (5.30), (5.109) and (5.115) into Equation (5.107) and simplifying using Equation (5.95) gives

$$\begin{aligned} d\mathcal{L} = &- \left(\overset{\vee}{S}_f \overset{\vee}{\Lambda}_f \otimes \mathbf{D}\right) : d\mathbf{K} - \left(\tilde{S}_f \hat{\Lambda}_f \otimes \hat{U}\right) : d\tilde{\mathbf{K}} + \tilde{S}_f \hat{\Lambda}_f : d\tilde{F} \\ &- \overset{\vee}{\Lambda}_f : d\overset{\vee}{F}_f + \overset{\vee}{\overset{\vee}{K}}_{pf} \overset{\vee}{\Lambda}_f : d\overset{\vee}{U}_p + \hat{\Lambda}_f : d\hat{F}_f - \left(\tilde{K}_{pf} \hat{\Lambda}_f + \tilde{S}_p^T \hat{X}_f \frac{dg}{d\overset{\vee}{U}_f}\right) : d\hat{U}_p. \end{aligned} \tag{5.116}$$

Substitution of Equations (5.35) and (5.77) into Equation (5.116) yields

$$\begin{aligned} d\mathcal{L} = &\left[\left(A\tilde{S}_f \hat{\Lambda}_f - \overset{\vee}{S}_f \overset{\vee}{\Lambda}_f\right) \otimes \mathbf{D}\right] : d\mathbf{K} \\ &- \left(\hat{X}_f \tilde{S}_f \hat{\Lambda}_f + \overset{\vee}{\Lambda}_f\right) : d\overset{\vee}{F}_f + \left(A\tilde{S}_f \hat{\Lambda}_f + \overset{\vee}{\overset{\vee}{K}}_{pf} \overset{\vee}{\Lambda}_f\right) : d\overset{\vee}{U}_p + \hat{\Lambda}_f : d\hat{F}_f - \left(\tilde{K}_{pf} \hat{\Lambda}_f + \tilde{S}_p^T \hat{X}_f \frac{dg}{d\overset{\vee}{U}_f}\right) : d\hat{U}_p. \end{aligned} \tag{5.117}$$

5.A.6 Sensitivities of the reaction load on the secondary DOFs

Based on the solution of Equation (5.22) one can also analyse the reaction loads of the secondary DOFs via

$$\overset{\vee}{F}_p = \hat{K}_p \hat{U} + \overset{\vee}{\overset{\vee}{K}}_{pf} \overset{\vee}{U}_f + \overset{\vee}{\overset{\vee}{K}}_{pp} \overset{\vee}{U}_p. \tag{5.23 revisited}$$

We augment the response function with the linear systems of equations from Equations (5.20), (5.21) and (5.27), that is

$$\begin{aligned} \mathcal{L}\left[\overset{\vee}{F}_p, \overset{\wedge}{\Lambda}_f, \overset{\vee}{\Lambda}_f, \hat{\Lambda}_f\right] := g\left[\overset{\vee}{F}_p\right] &- \overset{\wedge}{\Lambda}_f : \left(\overset{\vee}{\overset{\vee}{K}}_{ff} \hat{X}_f - \hat{K}_f\right) \\ &- \overset{\vee}{\Lambda}_f : \left(\overset{\vee}{\overset{\vee}{K}}_{ff} \overset{\vee}{V}_f - \overset{\vee}{\overset{\vee}{K}}_{fp} \overset{\vee}{U}_p + \overset{\vee}{F}_f\right) \\ &- \hat{\Lambda}_f : \left(\tilde{K}_{ff} \hat{U}_f - \hat{F}_f + \tilde{K}_{fp} \hat{U}_p - \tilde{F}_f\right). \end{aligned} \tag{5.118}$$

Full differentiation of both sides gives

$$\begin{aligned} d\mathcal{L} = &\frac{dg}{d\overset{\vee}{F}_p} : \left(d\hat{K}_p \hat{U} + d\overset{\vee}{\overset{\vee}{K}}_{pf} \overset{\vee}{U}_f + d\overset{\vee}{\overset{\vee}{K}}_{pp} \overset{\vee}{U}_p\right) \\ &+ \frac{dg}{d\overset{\vee}{F}_p} : \left(\hat{K}_p d\hat{U} - \overset{\vee}{\overset{\vee}{K}}_{pf} \left(d\hat{X}_f \hat{U} + \hat{X}_f d\hat{U} + d\overset{\vee}{V}_f\right) + \overset{\vee}{\overset{\vee}{K}}_{pp} d\overset{\vee}{U}_p\right) \\ &- \overset{\wedge}{\Lambda}_f : \left(d\overset{\vee}{\overset{\vee}{K}}_{ff} \hat{X}_f + \overset{\vee}{\overset{\vee}{K}}_{ff} d\hat{X}_f - d\hat{K}_f\right) \\ &- \overset{\vee}{\Lambda}_f : \left(d\overset{\vee}{\overset{\vee}{K}}_{ff} \overset{\vee}{V}_f + \overset{\vee}{\overset{\vee}{K}}_{ff} d\overset{\vee}{V}_f - d\overset{\vee}{\overset{\vee}{K}}_{fp} \overset{\vee}{U}_p - \overset{\vee}{\overset{\vee}{K}}_{fp} d\overset{\vee}{U}_p + d\overset{\vee}{F}_f\right) \\ &- \hat{\Lambda}_f : \left(d\tilde{K}_{ff} \hat{U}_f + \tilde{K}_{ff} d\hat{U}_f - d\hat{F}_f + d\tilde{K}_{fp} \hat{U}_p + \tilde{K}_{fp} d\hat{U}_p - d\tilde{F}_f\right). \end{aligned} \tag{5.119}$$

Since $d\overset{\wedge\wedge}{\mathbf{X}}_f$, $d\overset{\vee}{\mathbf{V}}_f$ and $d\hat{\mathbf{U}}_f$ are expensive to calculate, aim is to circumvent this by a suitable choices of $\overset{\wedge\wedge}{\mathbf{\Lambda}}_f$, $\overset{\vee}{\mathbf{\Lambda}}_f$ and $\hat{\mathbf{\Lambda}}_f$. Using Equation (5.109) and subsequently collecting terms involving $d\hat{\mathbf{U}}_f$ gives

$$\left(-\tilde{\mathbf{S}}_f^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}-\tilde{\mathbf{K}}_{ff}\hat{\mathbf{\Lambda}}_f\right) : d\hat{\mathbf{U}}_f. \tag{5.120}$$

To make the terms $d\hat{\mathbf{U}}_f$ vanish we define $\hat{\mathbf{\Lambda}}_f$ as

$$\tilde{\mathbf{K}}_{ff}\hat{\mathbf{\Lambda}}_f=-\tilde{\mathbf{S}}_f^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}, \tag{5.121}$$

which requires an additional solve using the condensed system. Collecting terms involving $d\overset{\vee}{\mathbf{V}}_f$ gives

$$\left(-\overset{\vee\vee}{\mathbf{K}}_{fp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}-\overset{\vee\vee}{\mathbf{K}}_{ff}\overset{\vee}{\mathbf{\Lambda}}_f\right) : d\overset{\vee}{\mathbf{V}}_f. \tag{5.122}$$

To make the terms $d\overset{\vee}{\mathbf{V}}_f$ vanish we define $\overset{\vee}{\mathbf{\Lambda}}_f$ as

$$\overset{\vee\vee}{\mathbf{K}}_{ff}\overset{\vee}{\mathbf{\Lambda}}_f=-\overset{\vee\vee}{\mathbf{K}}_{fp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}, \tag{5.123}$$

for which an additional full system solve is needed. Finally, collecting terms involving $d\overset{\wedge\wedge}{\mathbf{X}}_f$ gives

$$\left(-\overset{\vee\vee}{\mathbf{K}}_{fp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}\hat{\mathbf{U}}^{\mathsf{T}}-\overset{\vee\vee}{\mathbf{K}}_{ff}\overset{\wedge\wedge}{\mathbf{\Lambda}}_f\right) : d\overset{\wedge\wedge}{\mathbf{X}}_f. \tag{5.124}$$

To make the terms $d\overset{\wedge\wedge}{\mathbf{X}}_f$ vanish we define $\overset{\wedge\wedge}{\mathbf{\Lambda}}_f$ as

$$\overset{\vee\vee}{\mathbf{K}}_{ff}\overset{\wedge\wedge}{\mathbf{\Lambda}}_f=-\overset{\vee\vee}{\mathbf{K}}_{fp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}\hat{\mathbf{U}}^{\mathsf{T}}. \tag{5.125}$$

Thus, again

$$\overset{\wedge\wedge}{\mathbf{\Lambda}}_f\left[\overset{\vee}{\mathbf{\Lambda}}_f\right]=\overset{\vee}{\mathbf{\Lambda}}_f\hat{\mathbf{U}}^{\mathsf{T}}. \tag{5.115 revisited}$$

Costs are comparable to those for the case in Section 5.A.5. With all multipliers defined, substitution of Equations (5.30), (5.109) and (5.115) into Equation (5.119) and simplifying using Equation (5.95) gives

$$\begin{aligned}\mathrm{d}\mathcal{L}=&-\left(\overset{\vee}{\mathbf{S}}_f\overset{\vee}{\mathbf{\Lambda}}_f\otimes\mathbf{D}\right) : \mathrm{d}\mathbf{K}-\left(\tilde{\mathbf{S}}_f\hat{\mathbf{\Lambda}}_f\otimes\hat{\mathbf{U}}\right) : \mathrm{d}\tilde{\mathbf{K}}+\tilde{\mathbf{S}}_f\hat{\mathbf{\Lambda}}_f : \mathrm{d}\tilde{\mathbf{F}}\\&-\overset{\vee}{\mathbf{\Lambda}}_f : \mathrm{d}\overset{\vee}{\mathbf{F}}_f+\left(\overset{\vee\vee}{\mathbf{K}}_{pp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}+\overset{\vee\vee}{\mathbf{K}}_{pf}\overset{\vee}{\mathbf{\Lambda}}_f\right) : \mathrm{d}\overset{\vee}{\mathbf{U}}_p+\hat{\mathbf{\Lambda}}_f : \mathrm{d}\hat{\mathbf{F}}_f-\left(\tilde{\mathbf{K}}_{pf}\hat{\mathbf{\Lambda}}_f+\tilde{\mathbf{S}}_p^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_p}\right) : \mathrm{d}\hat{\mathbf{U}}_p.\end{aligned} \tag{5.126}$$

Substitution of Equations (5.35) and (5.77) into Equation (5.126) yields

$$\begin{aligned} \mathrm{d}\mathcal{L} = &\left[\left(\overset{\vee}{\mathbf{S}}_\mathrm{f}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f} - \mathbf{A}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f}\right)\otimes\mathbf{D}\right] : \mathrm{d}\mathbf{K} \\ &- \left(\overset{\wedge}{\mathbf{X}}_\mathrm{f}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} + \overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}\right) : \mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{f} + \left(\overset{\curlyvee}{\mathbf{K}}_\mathrm{pp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}} + \mathbf{C}\tilde{\mathbf{S}}_\mathrm{f}\hat{\mathbf{\Lambda}}_\mathrm{f} + \overset{\curlyvee}{\mathbf{K}}_\mathrm{pf}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}\right) : \mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{p} \\ &+ \hat{\mathbf{\Lambda}}_\mathrm{f} : \mathrm{d}\hat{\mathbf{F}}_\mathrm{f} - \left(\tilde{\mathbf{K}}_\mathrm{pf}\hat{\mathbf{\Lambda}}_\mathrm{f} + \tilde{\mathbf{S}}_\mathrm{p}^\mathsf{T}\mathbf{C}^\mathsf{T}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}}\right) : \mathrm{d}\hat{\mathbf{U}}_\mathrm{p}. \end{aligned} \tag{5.127}$$

5.A.7 Summary

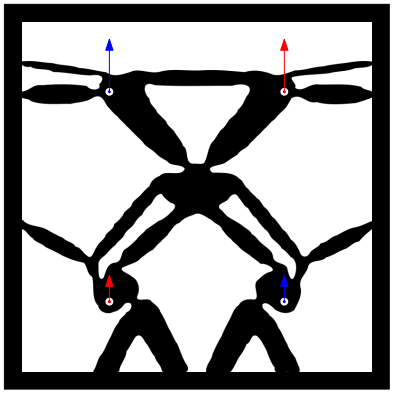
For clarity the resulting sensitivities as obtained in Sections 5.A.1 to 5.A.6 are summarized in Table 5.3. It also lists the additional solves, along with their dimensionality and number of right-hand-sides, required to calculate the sensitivities. Here we assumed a worst-case scenario, where a response g depends on *all* the $l^{(i)}$ states $\hat{\mathbf{U}}_\mathrm{f}^{(i)}$, $\hat{\mathbf{F}}_\mathrm{p}^{(i)}$, $\overset{\vee}{\mathbf{U}}_\mathrm{f}^{(i)}$ and/or $\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}$ of analysis i. In practice the number of right-hand-sides for a single response depends on the dependency of the response with respect to each state. Table 5.3 shows expensive additional solves are only required for responses depending on the state(s) or reaction loads of secondary DOFs.

Table 5.3: Sensitivities of common response dependencies with respect to the system matrix, applied loads and prescribed states. Format is $\mathrm{d}g_i\,[x_i] = \sum_j a_{ij} : \mathrm{d}b_j$, with i the row and j the column. Note the dimensionality of $\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f} \in \mathbb{R}^{\overset{\vee}{f}^{(i)}\times l^{(i)}}$ and $\hat{\mathbf{\Lambda}}_\mathrm{f} \in \mathbb{R}^{\hat{f}^{(i)}\times l^{(i)}}$.

x	$\mathrm{d}\mathbf{K}$	$\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{p}^{(i)}$	$\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{f}^{(i)}$	$\mathrm{d}\hat{\mathbf{U}}_\mathrm{p}^{(i)}$	$\mathrm{d}\hat{\mathbf{F}}_\mathrm{f}^{(i)}$	a	Solves
$\tilde{\mathbf{K}}$	$\mathbf{A}a\otimes\mathbf{A}$					$\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{K}}}$	
$\tilde{\mathbf{F}}^{(i)}$	$\mathbf{A}a\otimes\mathbf{B}^{(i)}$	$\mathbf{C}a$	$-\overset{\wedge}{\mathbf{X}}_\mathrm{f}a$			$\frac{\mathrm{d}g}{\mathrm{d}\tilde{\mathbf{F}}^{(i)}}$	
$\hat{\mathbf{U}}_\mathrm{f}^{(i)}$	$\mathbf{A}a\otimes\mathbf{D}^{(i)}$	$\mathbf{C}a$	$-\overset{\wedge}{\mathbf{X}}_\mathrm{f}a$	$-\tilde{\mathbf{K}}_\mathrm{pf}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\tilde{\mathbf{S}}_\mathrm{f}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\tilde{\mathbf{K}}_\mathrm{ff}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = \frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{U}}_\mathrm{f}^{(i)}}$
$\hat{\mathbf{F}}_\mathrm{p}^{(i)}$	$\mathbf{A}a\otimes\mathbf{D}^{(i)}$	$\mathbf{C}a$	$-\overset{\wedge}{\mathbf{X}}_\mathrm{f}a$	$-\tilde{\mathbf{K}}_\mathrm{pf}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$ $+\tilde{\mathbf{K}}_\mathrm{pp}^{(i)}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}^{(i)}}$	$\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\tilde{\mathbf{S}}_\mathrm{f}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} -$ $\tilde{\mathbf{S}}_\mathrm{p}^{(i)}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}^{(i)}}$	$\tilde{\mathbf{K}}_\mathrm{ff}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = \tilde{\mathbf{K}}_\mathrm{fp}^{(i)}\frac{\mathrm{d}g}{\mathrm{d}\hat{\mathbf{F}}_\mathrm{p}^{(i)}}$
$\overset{\vee}{\mathbf{U}}_\mathrm{f}^{(i)}$	$\left(\mathbf{A}a - \overset{\vee}{\mathbf{S}}_\mathrm{f}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}\right)\otimes\mathbf{D}^{(i)}$	$\mathbf{C}a + \overset{\curlyvee}{\mathbf{K}}_\mathrm{pf}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$-\overset{\wedge}{\mathbf{X}}_\mathrm{f}a - \overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$-\tilde{\mathbf{K}}_\mathrm{pf}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} -$ $\tilde{\mathbf{S}}_\mathrm{p}^{(i)\mathsf{T}}\overset{\wedge}{\mathbf{X}}_\mathrm{f}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}^{(i)}}$	$\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\tilde{\mathbf{S}}_\mathrm{f}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\overset{\curlyvee}{\mathbf{K}}_\mathrm{ff}^{(i)}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = -\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}^{(i)}}$ $\tilde{\mathbf{K}}_\mathrm{ff}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = -\tilde{\mathbf{S}}_\mathrm{f}^{(i)\mathsf{T}}\overset{\wedge}{\mathbf{X}}_\mathrm{f}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{U}}_\mathrm{f}^{(i)}}$
$\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}$	$\left(\mathbf{A}a - \overset{\vee}{\mathbf{S}}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}\right)\otimes\mathbf{D}$	$\mathbf{C}a + \overset{\curlyvee}{\mathbf{K}}_\mathrm{pf}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$ $+\overset{\curlyvee}{\mathbf{K}}_\mathrm{pp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}}$	$-\overset{\wedge}{\mathbf{X}}_\mathrm{f}a - \overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$-\tilde{\mathbf{K}}_\mathrm{pf}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} -$ $\tilde{\mathbf{S}}_\mathrm{p}^{(i)\mathsf{T}}\mathbf{C}^\mathsf{T}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}}$	$\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\tilde{\mathbf{S}}_\mathrm{f}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)}$	$\overset{\curlyvee}{\mathbf{K}}\overset{\vee}{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = -\overset{\curlyvee}{\mathbf{K}}_\mathrm{fp}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}}$ $\tilde{\mathbf{K}}_\mathrm{ff}^{(i)}\hat{\mathbf{\Lambda}}_\mathrm{f}^{(i)} = -\tilde{\mathbf{S}}_\mathrm{f}^{(i)\mathsf{T}}\mathbf{C}^\mathsf{T}\frac{\mathrm{d}g}{\mathrm{d}\overset{\vee}{\mathbf{F}}_\mathrm{p}^{(i)}}$

6

Efficient computation of states and sensitivities



Real-world structural optimisation problems typically involve multiple loading conditions and design constraints; coming with increased computational effort to solve for the state and adjoint equations. This chapter proposes using a Linear Dependency Aware Solver (LDAS) to detect and exploit linear dependencies between encountered physical and adjoint loads. The proposed algorithm can efficiently detect linear dependencies between all loads and obtain the exact solution while avoiding unnecessary solves entirely and automatically.

This chapter is based on peer-reviewed journal paper:
Koppen, S., van der Kolk, M., van den Boom, S., and Langelaar, M. (2022). *Efficient computation of states and sensitivities for compound structural optimisation problems using a Linear Dependency Aware Solver (LDAS).* Structural and Multidisciplinary Optimization, 65(9).

Efficient computation of states and sensitivities for compound structural optimisation problems using a Linear Dependency Aware Solver (LDAS)

Abstract *Real-world structural optimisation problems involve multiple loading conditions and design constraints, with responses typically depending on states of discretised governing equations. Generally, one uses gradient-based nested analysis and design approaches to solve these problems. Herein, solving both physical and adjoint problems dominates the overall computational effort. Although not commonly detected, real-world problems can contain linear dependencies between encountered physical and adjoint loads. Manually keeping track of such dependencies becomes tedious as design problems become increasingly involved. This work proposes using a Linear Dependency Aware Solver (LDAS) to detect and exploit such dependencies. The proposed algorithm can efficiently detect linear dependencies between all loads and obtain the exact solution while avoiding unnecessary solves entirely and automatically. Illustrative examples demonstrate the need and benefits of using an LDAS, including a run-time experiment.*

6.1 Introduction

In structural optimisation, particularly in topology optimisation, the *self-adjoint* compliance minimisation problem is often studied (Rozvany *et al.* 1989). One can obtain design sensitivities for gradient-based optimisation at a marginal computational cost due to the self-adjointness of the problem. This advantage has likely contributed to the popularity of studying the compliance minimisation problem. However, as Rozvany *et al.* (1993) pointed out almost three decades ago: "Self-adjoint problems, such as design for a single stress, a single compliance or single natural frequency constraint do not represent a real-world situation, because most practical structures are subject to several load conditions and design constraints." Almost three decades later, solving large-scale linear problems considering multiple physical loads and a large variety of responses—hereafter denoted by *compound* problems—is becoming increasingly attainable as available computational power increases. However, regardless of available computational power, efficient numerical implementations remain essential.

Typically, finding the state corresponding to a load, *i.e.* the solution to the governing equations dominates the overall computation time during optimisation. As Borrvall *et al.* (2001) report, the computational time of such procedures approaches 97% for minimum compliance problems considering a single physical load, where computation times increase further when considering compound problems.

Finding a solution to these systems of linear equations generally consists of two steps: preprocessing and solving (Amir *et al.* 2010a). The preprocessing for direct methods requires the (generally expensive) matrix factorisation, and solving requires finding the exact solution via comparatively inexpensive back-substitutions (Davis 2006). In contrast, iterative methods require the construction of a preconditioner, and they subsequently generate a sequence of approximate solutions until convergence (Saad 2003). The relative cost of preconditioner construction and the iterative solution process depends on many factors, such as the type of preconditioner and condition number. The preprocessing information can be repeatedly reused for subsequent solves within the same design iteration when this involves a system matrix with equivalent partitioning. This possibility holds for both solution methods.

Three strategies can be distinguished to lower the computational effort of solving large-scale linear systems of governing equations in structural optimisation, *i.e.* reduction of

1. the number of design iterations,
2. the computational effort per solve, and
3. the number of solves per design iteration.

The first technique has shown great potential to reduce computational effort, for instance using advanced sequential approximate optimisation schemes (*e.g.* see (Bruyneel *et al.* 2002; Li *et al.* 2015)). However, these approaches are out of scope for this discussion, independent of the presented methodology.

A common approach to reduce computation *time* per linear solve is to employ parallel computing (Aage *et al.* 2017; Borrvall *et al.* 2001), a technique which *distributes* the computational effort. However, to *reduce* this effort, approximation techniques should be considered, such as approximated reanalysis (Amir 2015; Kirsch 1991), iterative solution techniques (Amir *et al.* 2014, 2010b; Borrvall *et al.* 2001), and approximated model order reduction (Choi *et al.* 2019; Ma *et al.* 1993). Alternatively, static condensation (Guyan 1965; Irons 1965) allows for exact model order reduction, decreasing the system dimensionality without loss of information (*e.g.* see (Yang *et al.* 1996)). For a comprehensive review of techniques aiming to decrease the computational effort per solve in the context of topology optimisation, the reader is referred to the recent work by Mukherjee *et al.* (2021).

The third category—approaches to reduce the number of solves per design iterations—includes the adjoint sensitivity analysis method itself, for instance, when applied to most self-adjoint problems (Arora *et al.* 1979; Belegundu 1986; Vanderplaats 1980). For problems considering many physical static loads, Zhang *et al.* (2020) reduce the number of deterministic loads to a single approximated load using sampling schemes. Recent study shows that static condensation allows for a reduction of the number of factorisations/preconditioning steps and the number of solves in multi-partition problems; which are problems that, as a result of changing boundary conditions, require multiple different partitions of the stiffness matrix (Koppen *et al.* 2022c).

In contrast to that study, in this paper, we focus on compound problems with a *single* partitioning of the system matrix. We introduce another method of the third category that reduces the number of solves per design iteration design problems with equivalent

partitioning of degrees of freedom. Different boundary condition values can be handled as long as the partition remains the same. We herein assume linear state-based optimisation problems under (quasi-)static loading, which constitutes a significant fraction of all problems studied in the topology optimisation community (Bendsøe *et al.* 2004). By automatically detecting linear dependencies between physical and adjoint loads, unnecessary solves in compound problems involving the same partition of system matrix can be avoided entirely while maintaining equal accuracy of the solution of the states. To help the reader recognize linear dependencies that may arise in common design optimization problems, we distinguish three cases of such linear dependency:

1. Linearly Dependent Physical-Physical (LDPP) loads. Such cases are common in design problems involving multiple loading conditions with applied loads of varying magnitudes, for example, present in the case study of Section 6.4. Optimisation problems with LDPP loads are relatively easily detected manually and regularly avoided by the user.

2. Linearly Dependent Adjoint-Physical (LDAP) load pairs. Typical problems include cases where the adjoint load depends linearly on the *corresponding* physical load, as common in conventional self-adjoint[1] problems (Belegundu 1986; Rozvany *et al.* 1993). The most well-known design problem in the topology optimisation community involving such load pairs is the classical compliance minimisation problem. Such cases are typically detected by academics in this field but may be overlooked otherwise.

3. Mixed Linear Dependencies (MLD), *i.e.* cases where physical loads or adjoint loads can be written as a linear combination of previously considered physical or adjoint loads. MLDs also include linear dependencies between adjoint loads and between non-corresponding adjoint and physical loads (as well as any linear combination). These MLDs are the most general situation and the most difficult to foresee and consider by hand. Such cases are common in problems with multiple response functions depending on multiple states. More specifically, such cases often occur when the locations where the loads are applied and the locations of the performance measures coincide, such as typical in the design of compliant mechanisms. These MLDs will be elaborately clarified in all numerical examples.

A user will typically be unaware of the presence and type of most of such linear dependencies. A Linear Dependency Aware Solver (LDAS) can be employed to detect and exploit any linear dependency, including any of the three aforementioned types, automatically. In this work, we demonstrate the need and benefits of an LDAS in the context of gradient-based, structural optimisation for compound problems and provide one such solver in the form of a simple algorithm to automatically detect and exploit any linear dependence in a (possibly large) set of loads. The focus is on MLDS since these linear dependencies

[1] It is a common misconception that self-adjoint problems *always* exhibit an LDAP pair, as such problems can (and originally were) often of analytical nature and do not require a solve to obtain sensitivities (*e.g.* design for a single natural frequency) (Rozvany *et al.* 1993; Shield *et al.* 1970a). Also, problems that exhibit an LDAP pair are by no definition *per se* self-adjoint (*e.g.* the optimisation for deflection constraints constitutes a *non-self-adjoint problem*, although exhibiting an LDAP pair (Rozvany *et al.* 1993)).

are typically the hardest to detect. However, due to the generality of the method, it also automatically resolves unnecessary solves in LDPP and LDAP pairs. Thus, it is ensured that only the minimum number of linear solves is performed in each iteration. This advantage makes the approach suitable for general-purpose structural and topology optimisation implementations.

Note that the presented algorithm does *not* exclude other additional techniques to reduce the computational effort and time, such as parallel computing, approximate modelling, or reduced order techniques, which can be implemented alongside the presented methodology.

6.2 Method

Consider a general inequality-constrained nonlinear structural optimisation problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] \\ \text{s.t.} \quad & \mathbf{g}[\mathbf{x}] \leq \mathbf{0} \end{aligned} \tag{6.1}$$

with objective $f \in \mathbb{R}$, m inequality constraints $\mathbf{g} \in \mathbb{R}^m$ and N design variables $\mathbf{x} \in \mathbb{X}^N \subseteq \mathbb{R}^N$.

6.2.1 Response and sensitivity analysis

The responses (objective and constraint functions) commonly depend on physical states $\mathbf{U} := [\mathbf{u}_1, \dots, \mathbf{u}_a] \in \mathbb{R}^{n\times a}$, where n is the dimensionality of the discretised governing equations and a the number of states. These states implicitly depend on the design variables, *i.e.* $\mathbf{U} = \mathbf{U}[\mathbf{x}]$. We consider a setting in which these physical states are obtained by solving a linear system of discretised governing equations, *i.e.*

$$\mathbf{K}[\mathbf{x}]\mathbf{U} = \mathbf{F}[\mathbf{x}], \tag{6.2}$$

with $\mathbf{F}[\mathbf{x}] := [\mathbf{f}_1[\mathbf{x}], \dots, \mathbf{f}_a[\mathbf{x}]] \in \mathbb{R}^{n\times a}$ the physical loads and $\mathbf{K}[\mathbf{x}] \in \mathbb{R}^{n\times n}$ a design dependent, symmetric, and non-singular system matrix. In the following we assume the system in Equation (6.2) constitutes a single partition, thus the physical loads are applied on the system under the same boundary conditions.

In gradient-based optimisation, the sensitivities of the responses to the design variables are required to update the design variables. For structural optimisation problems with a large ratio of the number of design variables to the number of state-based response functions, commonly, the adjoint method is applied to efficiently obtain this sensitivity information (Arora *et al.* 1979; Vanderplaats 1980). To this end, consider the augmented response

$$\mathcal{L}_j[\mathbf{x}, \mathbf{U}[\mathbf{x}]] = g_j[\mathbf{x}, \mathbf{U}[\mathbf{x}]] - \mathbf{\Lambda}_j : (\mathbf{K}[\mathbf{x}]\mathbf{U} - \mathbf{F}[\mathbf{x}]). \tag{6.3}$$

with $\mathbf{\Lambda}_j := \left[\lambda_{j,1}, \dots, \lambda_{j,a}\right] \in \mathbb{R}^{n\times a}$. Here, a suitable choice of the adjoint states $\mathbf{\Lambda}_j$ can circumvent calculation of the computationally expensive derivative $\frac{\partial \mathbf{U}}{\partial x_k}$ (Vanderplaats 1980). Doing so, full differentiation of Equation (6.3) yields

$$\frac{\mathrm{d}\mathcal{L}_j}{\mathrm{d}x_k} = \frac{\partial g_j}{\partial x_k} - \mathbf{\Lambda}_j : \left(\frac{\partial \mathbf{K}}{\partial x_k}\mathbf{U}\right), \tag{6.4}$$

with

$$\mathbf{K}[\mathbf{x}]\boldsymbol{\Lambda}_j = \frac{\partial g_j}{\partial \mathbf{U}}, \tag{6.5}$$

where $\frac{\partial g_j}{\partial \mathbf{U}}$ is referred to as the adjoint loads of response g_j.

Each of the physical and adjoint loads can be linearly dependent on any combination of previously considered loads and thus can be reconstructed as their linear combination. Exploiting possible linear dependence can significantly reduce the costs required to find all states. Consider a set of a loads, of which b are linearly independent, then the computational effort scales roughly with $\frac{b}{a}$, as only b solves are required to reconstruct all states. To avoid unnecessarily solving Equations (6.2) and (6.5) for linear dependent loads we propose

1. to compute each load's dependency on previous loads, and
2. to keep track of the states corresponding to linearly independent loads.

Various possible methods exist to check for linear dependency and necessary bookkeeping. We consider one such algorithm that detects linear dependencies and builds orthogonal bases of linear independent loads and their corresponding states.

6.2.2 Orthogonalisation and reconstruction

Consider the non-empty orthogonal bases of loads $\mathcal{F}$ and states $\mathcal{U}$ of length c. One can investigate the linear dependency of a load $\mathbf{f}$ (*e.g.* a physical load $\mathbf{f}$ or adjoint load $\frac{\partial g}{\partial \mathbf{u}}$) with respect to $\mathcal{F}$ by applying the last step of the well known Gram-Schmidt orthogonalisation procedure[2] (Gram 1883; Laplace 1820; Schmidt 1907). The residual $\mathbf{r}$ is obtained via

$$\mathbf{r} := \mathbf{f} - \sum_{i=1}^{c} \alpha_i \mathcal{F}_i, \quad \text{with} \quad \alpha_i = \frac{\mathcal{F}_i \cdot \mathbf{f}}{\mathcal{F}_i \cdot \mathcal{F}_i}, \tag{6.6}$$

with $\mathcal{F}_i$ the ith load in $\mathcal{F}$. A possible implementation is given by the pseudo-code Algorithm 5.

Algorithm 5 Gram-Schmidt orthogonalisation

```
α = []
r = copy(f)
for f in F do
    α = (r · f)/(f · f)
    r −= αf
    α.append(α)
end for
return (α, r)
```

If the norm of the residual $\mathbf{r}$ is zero, then $\mathbf{f}$ is linearly dependent to basis $\mathcal{F}$. As a result, the corresponding state $\mathbf{u}$ (or adjoint state $\boldsymbol{\lambda}$) is linearly dependent on basis $\mathcal{U}$. Thus, the

[2] Although the method is named after Jørgen Pedersen Gram and Erhard Schmidt, Pierre-Simon Laplace had been familiar with it before, see (Leon *et al.* 2013).

state $\mathbf{u}$ may be reconstructed via

$$\mathbf{u} = \sum_{i=1}^{c} \alpha_i \mathcal{U}_i. \tag{6.7}$$

As such, one can obtain the *exact* numerical solution of state $\mathbf{u}$, while avoiding solving the governing equations for loads $\mathbf{f}$. However, if the norm of the residual vector $\mathbf{r}$ is nonzero (or bigger than a relatively small value ε), $\mathbf{f}$ is linearly *independent* with respect to basis $\mathcal{F}$ and the expensive solve cannot be avoided.

We solve for the state $\mathbf{v}$ corresponding to residual load $\mathbf{r}$ defined by

$$\mathbf{K}[\mathbf{x}]\mathbf{v} = \mathbf{r}. \tag{6.8}$$

Subsequently load $\mathbf{r}$ and state $\mathbf{v}$ are added to bases $\mathcal{F}$ and $\mathcal{U}$, respectively. Since $\mathbf{r}$ is orthogonal with respect to basis $\mathcal{F}$, so is $\mathbf{v}$ to $\mathcal{U}$. As a result, both enriched bases $\mathcal{F}$ and $\mathcal{U}$ remain orthogonal. The state $\mathbf{u}$ is then reconstructed from Equation (6.6) and Equation (6.7). The above procedure can be repeated using the enriched bases, as defined in Algorithm 6. Due to the general nature of the algorithm, the proposed procedure is independent of the type of dependencies as defined in Section 6.1. The equivalence of solutions is extensively verified for many test problems.

Algorithm 6 Linear Dependency Aware Solver

```
for (i, f) in enumerate(F) do
    (α, r) = GSO(f, F)
    if ‖r > ε‖ then
        F.append(r)
        U.append(solve(K, r))
        α.append(1)
    end if
    U[i] += α · U
end for
return U
```

Although Algorithm 6 introduces additional computational operations, *i.e.* computing vector norms and orthogonality coefficients, their computational cost is typically negligible compared to the costs of solving a system of equations, as illustrated in Section 6.5. The computational effort increases with the number of loads to consider, however, remains negligible as long as the number of loads (both physical and adjoint) is smaller than the dimensionality of the load vectors.Furthermore, these operations do not change when considering distributed-memory parallelism. Alternatively, for loads that do not depend on the states, it is possible to rearrange Algorithm 6 to determine all the independent loads first and evaluate their solutions in parallel afterwards.

6.3 Analytical example

Compound problems may appear in any real-world problem, modelled by (a sequence of) linear governing equations. Typical examples of compound problems are formulations

Figure 6.1: One-dimensional two degrees of freedom compliant mechanism model.

with multiple loading conditions and multiple response functions in which the degrees of freedom of (some of) the loads coincide with (some of) the degrees of freedom that define the response functions. For example, one may think of the design of a structure with multiple critical loading conditions, where the displacements of a loading condition are measured at the same degrees of freedom where the loads are applied at another loading condition. A direct example of this are multi-input-multi-output compliant mechanisms, see *e.g.* (Frecker *et al.* 1999) or (Liu *et al.* 2009). The problem formulation of such mechanisms includes multiple physical loads *and* responses, all applied to, or dependent on, the input and output degrees of freedom of the mechanism. As a result, MLD is commonly present. However, it generally remains unnoticed. To clarify the cases in which one might encounter linear dependency, we here exemplify the three different types of unnecessary solves, as introduced in Section 6.1.

6.3.1 Problem formulation

Consider the two degrees of freedom spring model as depicted in Figure 6.1. Note that this example—after applying static condensation—can *exactly* represent any single-input-single-output compliant mechanism, see *e.g.* (Hasse *et al.* 2017; Wang 2009b). Therefore, this two degrees of freedom example is fully representative of large-scale linear problems considering multiple physical loads and responses while better suited to illustrate the proposed method.

6.3.2 Problem analysis

Next, we analyse the properties of this optimisation problem in light of the proposed method, with a specific emphasis on the required number of systems of equations that are to be solved.

Forward analysis

The physical and adjoint states can be obtained by solving the design-dependent discretised governing equations following Equations (6.2) and (6.5). A set of the following three physical loads is considered:

$$\mathbf{F} = \left[\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} \right]. \tag{6.9}$$

The first residual *by definition* equals the first load, that is $\mathbf{r}_1 = \mathbf{f}_1$. As a result, the state $\mathbf{v}_1 = \mathbf{u}_1$. Since the basis is initially empty when this load is considered, the resulting load and state are directly added to corresponding bases. The second residual is calculated via Equation (6.6), that is

$$\mathbf{r}_2 = \mathbf{f}_2 - \alpha_1 \mathcal{F}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}. \tag{6.10}$$

$\mathcal{F}$		Loads					
$\mathbf{r}_1 = \mathbf{f}_1$	$\mathbf{r}_2$	$\mathbf{f}_1$	$\mathbf{f}_2$	$\mathbf{f}_3$	$\frac{\partial g_1}{\partial \mathbf{u}_2}$	$\frac{\partial g_2}{\partial \mathbf{u}_1}$	$\frac{\partial g_2}{\partial \mathbf{u}_3}$
$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\2\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\2\end{bmatrix}$	$\begin{bmatrix}4\\4\end{bmatrix}$	$\begin{bmatrix}\frac{1}{2}\\1\end{bmatrix}$	$\begin{bmatrix}2\\1\end{bmatrix}$	$\begin{bmatrix}1\\3\end{bmatrix}$
		$\mathbf{r}_1$	$\mathbf{r}_1 + \mathbf{r}_2$	$4\mathbf{r}_1 + 2\mathbf{r}_2$	$\frac{1}{2}\mathbf{r}_1 + \frac{1}{2}\mathbf{r}_2$	$2\mathbf{r}_1 + \frac{1}{2}\mathbf{r}_2$	$\mathbf{r}_1 + \frac{3}{2}\mathbf{r}_2$
$\mathcal{U}$		States					
$\mathbf{v}_1 = \mathbf{u}_1$	$\mathbf{v}_2$	$\mathbf{u}_1$	$\mathbf{u}_2$	$\mathbf{u}_3$	$\boldsymbol{\lambda}_{1,2}$	$\boldsymbol{\lambda}_{2,1}$	$\boldsymbol{\lambda}_{2,3}$
		$\mathbf{v}_1$	$\mathbf{v}_1 + \mathbf{v}_2$	$4\mathbf{v}_1 + 2\mathbf{v}_2$	$\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2$	$2\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2$	$\mathbf{v}_1 + \frac{3}{2}\mathbf{v}_2$

Table 6.1: Overview of both physical and adjoint loads and states, as well as the orthogonal bases encountered in the illustrative example presented in Figure 6.1. The right-hand side displays the load and states vectors expressed as linear combinations of the corresponding bases given on the left-hand side.

Since $\mathbf{r}_2$ is non-zero, the first and second physical loads are linearly independent. The corresponding physical state $\mathbf{v}_2$ is obtained by solving for the non-zero load $\mathbf{r}_2$ via Equation (6.8). As a result the following bases, consisting of orthogonal vectors, are obtained after solving for the first two loads:

$$\mathcal{F} = [\mathbf{f}_1, \mathbf{r}_2] \quad \text{and} \quad \mathcal{U} = [\mathbf{u}_1, \mathbf{v}_2]. \tag{6.11}$$

The second physical state is now reconstructed following Equation (6.7) and reads

$$\mathbf{u}_2 = \alpha_1 \mathcal{U}_1 + \mathbf{v}_2 = \mathbf{u}_1 + \mathbf{v}_2. \tag{6.12}$$

The third physical load can be written as a linear combination of the current orthogonal basis $\mathcal{F}$, resulting in a zero residual load $\mathbf{r}_3 = \mathbf{0}$. These are thus LDPP loads. Thus the basis $\mathcal{U}$ can be used to reconstruct the third physical state without an additional solve as in Equation (6.7), *i.e.*

$$\mathbf{u}_3 = \alpha_1 \mathcal{U}_1 + \alpha_2 \mathcal{U}_2 = 4\mathbf{u}_1 + 2\mathbf{v}_2. \tag{6.13}$$

Sensitivity analysis

Now consider a response function $g_1[\mathbf{u}_2]$ that is a measure for the strain energy due to load $\mathbf{f}_2$, *i.e.*

$$g_1[\mathbf{u}_2] = \frac{1}{2}\mathbf{f}_2 \cdot \mathbf{u}_2. \tag{6.14}$$

The second adjoint load for this response is linearly dependent on the corresponding physical load $\mathbf{f}_2$ as

$$\frac{\partial g_1}{\partial \mathbf{u}_2} = \frac{1}{2}\mathbf{f}_2, \tag{6.15}$$

thus this is an LDAP pair, and consequently $\mathbf{r}_4 = \mathbf{0}$. As a result, one can use the basis $\mathcal{U}$ to reconstruct the second adjoint state, which yields

$$\boldsymbol{\lambda}_{1,2} = \frac{1}{2}\mathbf{u}_2 = \alpha_1 \mathcal{U}_1 + \alpha_2 \mathcal{U}_2 = \frac{1}{2}\mathbf{u}_1 + \frac{1}{2}\mathbf{v}_2, \tag{6.16}$$

with $\boldsymbol{\lambda}_{j,i}$ the adjoint state of response j with respect to state i. Note that both the first and third adjoint loads of this response, that is $\frac{\partial g_1}{\partial \mathbf{u}_1}$ and $\frac{\partial g_1}{\partial \mathbf{u}_3}$ are zero, and thus so are $\boldsymbol{\lambda}_{1,1}$ and $\boldsymbol{\lambda}_{1,3}$.

Finally consider a (fictitious) response function $g_2[\mathbf{u}_1, \mathbf{u}_3]$ that depends on both degrees of freedom of the first state *and* third state via

$$g_2[\mathbf{u}_1, \mathbf{u}_3] = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \mathbf{u}_1 + \begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot \mathbf{u}_3. \tag{6.17}$$

The adjoint loads for this response function can be written as

$$\begin{aligned} \frac{\partial g_2}{\partial \mathbf{u}_1} &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\mathbf{f}_1 + \frac{1}{2}\mathbf{r}_2 \quad \text{and} \\ \frac{\partial g_2}{\partial \mathbf{u}_3} &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \mathbf{f}_1 + \frac{3}{2}\mathbf{r}_2. \end{aligned} \tag{6.18}$$

Note that both adjoint loads are linearly dependent on *a combination of* previously considered loads, *i.e.* an MLD. In this case, the adjoint loads are both linearly dependent on both loads in basis $\mathcal{F}$. As a result, one may again use the states in $\mathcal{U}$ to reconstruct the adjoint states via

$$\boldsymbol{\lambda}_{2,1} = 2\mathbf{u}_1 + \frac{1}{2}\mathbf{v}_2 \quad \text{and} \quad \boldsymbol{\lambda}_{2,3} = \mathbf{u}_1 + \frac{3}{2}\mathbf{v}_2. \tag{6.19}$$

The loads, states, and bases of this example are summarised in Table 6.1.

Concluding remarks

Six solves are required when all loads (physical and adjoint) are considered. If both LDPPs and LDAP pairs are taken into account, only three solves are needed. Finally, considering MLDs (and thus also LDPPs and LDAP pairs), only two solves are required. Although the presented example is simplified, more complex MLDs do appear in large-scale compound problems, as will be demonstrated in Sections 6.4 and 6.5.

6.4 Numerical example 1: design of a bridge

In this section we demonstrate the use of an LDAS for a practically relevant numerical example. The emphasis will be on the potential gain, not on formulation, design or optimization convergence aspects.

6.4.1 Problem formulation

Consider the design of a simplified bridge-deck supporting structure. A schematic of the problem setting, together with an optimized design, is shown in Figure 6.2. The engineer has selected a set of crucial loading conditions and (derived) constraints based on an extensive set of requirements and loading conditions, as typical in the design of such a bridge.

The aim is to design a stiff bridge with limited material for a given set of four loading conditions considering three points of interest: one at a quarter, one at the middle and one at three-quarters of the bridge deck. The magnitudes of forces applied to the DOFs

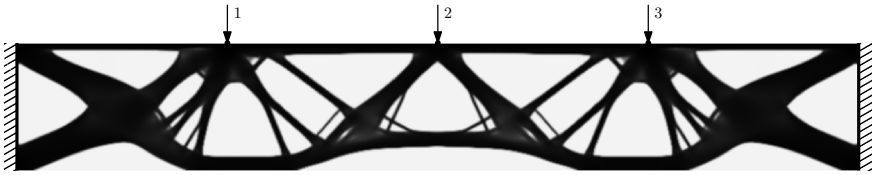


Figure 6.2: Optimized result of topology optimization problem Equation (6.20). The solution (800 × 120 finite elements and design variables) satisfies all constraints (all active) and the optimization process terminated in 59 design iterations. Corresponding displacements at the DOF of interest are listed in Table 6.3.

DOF \ LC	1	2	3	4
1	3	0	0	1
2	0	2	0	1
3	0	0	3	1

Table 6.2: Magnitude of forces applied at DOFs 1, 2 and 3 (numbered as assigned in Figure 6.2) for loading conditions (LC) 1, 2, 3 and 4.

of interest for the four loading conditions are as shown in Table 6.2. Furthermore, it is decided that the difference in deformations from loading conditions with concentrated loads and combined loads must be restricted. As such, the design has to satisfy several constraints on the deflection of the points of interest under the given loading conditions.

The topology optimization problem formulation reads

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{X}^N} \quad & f[\mathbf{x}] : \sum_j^4 \mathcal{E}_j\left[\mathbf{u}_j[\mathbf{x}]\right] \\ \text{s.t.} \quad & g^{\mathrm{v}}[\mathbf{x}] : v[\mathbf{x}] \le \overline{v} \\ & g_1^{\mathrm{u}}[\mathbf{x}] : u_{1,1}[\mathbf{x}] - u_{1,4}[\mathbf{x}] \le \overline{u} \\ & g_2^{\mathrm{u}}[\mathbf{x}] : u_{2,2}[\mathbf{x}] - u_{2,4}[\mathbf{x}] \le \overline{u} \\ & g_3^{\mathrm{u}}[\mathbf{x}] : u_{3,3}[\mathbf{x}] - u_{3,4}[\mathbf{x}] \le \overline{u} \end{aligned} \tag{6.20}$$

The objective is to minimize the strain energy $\mathcal{E}_j$, or equivalently maximize the stiffness, under the four loading conditions by finding design variables x_k that are bounded by $\mathbb{X} = \{x \in \mathbb{R} \mid 0 \le x \le 1\}$. Constraint $g^{\mathrm{v}}[\mathbf{x}]$ limits the maximum material usage by fraction $\overline{v} = 0.5$. Constraint g_i^{u} limits the difference in displacement of DOF i between loading conditions i and 4 to $\overline{u} = 20$. Herein $u_{i,j}\left[\mathbf{u}_j[\mathbf{x}]\right]$ is defined as the displacement at DOF of interest i for loading condition j.

An optimized solution is shown in Figure 6.2. The displacements at the DOFs of interest for the four loading conditions of this constrained optimised design are shown in Table 6.3. Note that the deformations at the points of interest now satisfy the imposed restrictions.

DOF \ LC	1	2	3	4
1	96	65	34	76
2	97	170	97	150
3	34	65	96	76

Table 6.3: Displacements at DOF 1 and 2 for loading conditions 1, 2 and 3 of the optimized design with deformation constraints.

6.4.2 Problem analysis

Now we analyse the potential gain of using an LDAS for solving this bridge design optimization problem.

Forward analysis

Let us first consider the objective of the problem formulation posed in Equation (6.20). The objective is a function of the states of four loading conditions, that is $f[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4]$. Straightforward analysis would thus require four solves. However, upon closer inspection, it can be observed that the fourth loading condition uniquely uses LDPP loads. The fourth state $\mathbf{u}_4$ can, thus, be written as a linear combination of states $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$. No solves are required for the forward analysis of the constraints since all states have previously been determined to calculate the objective. Thus, using an LDAS to solve Equation (6.20) can save the user one of the four solves required in the forward analysis, thus requiring three solves per design iteration.

Sensitivity analysis

Straightforward sensitivity analysis of the objective requires four more solves. However, the adjoint loads $\frac{\mathrm{d}f}{\mathrm{d}\mathbf{u}_i}$ for $i = 1, 2, 3, 4$ can all be written as a linear combination of $\mathbf{f}_1$ and $\mathbf{f}_2$ and $\mathbf{f}_3$, that is four LDAP pairs. Considering the additional constraint functions, the number of solves required for sensitivity analysis quickly increases. Each constraint depends on two states, thus requiring two adjoint solves per constraint. This sensitivity analysis thus requires a total of six additional solves. Closer inspection, similar to the preceding section, brings to light the MLDs in these constraints; all the adjoint loads can be written as a linear combination of physical loads $\mathbf{f}_1$, $\mathbf{f}_2$ and $\mathbf{f}_3$. Using an LDAS thus avoids all of the ten solves, and the sensitivity analysis would not require any solve.

Concluding remarks

A straightforward implementation to solve the bridge design problem would require a total of fourteen solves per design iteration, four for the forward analysis and ten for the sensitivity analysis. Using an LDAS one only requires three solves per design iteration. That is a decrease in the number of solves by almost 80%.

6.5 Numerical example 2: design of a multi-DOF compliant mechanism

To further demonstrate the benefits of the proposed method, we consider as illustrative case study the topology optimisation of a planar, multiple degree-of-freedom micro-mechanism for use, for example, as analogue gate in a mechanical computer (Larsen *et al.* 1997). Note that the focus here is not on the optimisation (problem formulation) of the micro-mechanism but on demonstrating the numerical benefits of an LDAS.

6.5.1 Problem formulation

Consider the design problem depicted in Figure 6.3a. The domain consists of four points of interest, each consisting of two Degrees Of Freedom (DOFs), u_x and u_y, respectively. The target is to design a monolithic compliant mechanism that *doubles* a unit input motion at DOF 6 to the output motion at DOF 4 *and* a unit input motion at DOF 8 to an equivalent magnified output motion at DOF 2. Thus we consider two independent *kinematic* DOFs. Furthermore, we also consider parasitic motion, input coupling and output coupling: all remaining DOFs—apart from the intended input and output—are restricted to displace a maximum of 0.1% of the input motion.

The force paths have to cross, making this a challenging problem that is not necessarily intuitive for engineers to solve. Therefore we solve this problem using topology optimisation (Bendsøe *et al.* 2004). We consider the following compound topology optimisation problem formulation[3]:

$$
\begin{aligned}
&\min_{\mathbf{x}\in\mathbb{X}^N} \\
f[\mathbf{x}]: \quad & \sum_j \mathcal{E}_j\left[\mathbf{u}_j[\mathbf{x}]\right] \qquad \forall\, j\in\{1,3,5,7\} \\
\text{s.t.} \\
g^{\mathrm{v}}[\mathbf{x}]: \quad & v[\mathbf{x}] \le \overline{v} \\
g^{\mathrm{in}}_{j,j}[\mathbf{x}]: \quad & u_{j,j}\left[\mathbf{u}_j[\mathbf{x}]\right] \ge u_{\mathrm{in}} \qquad \forall\, j\in\{6,8\} \\
g^{\mathrm{ct}}_{i,j}[\mathbf{x}]: \quad & u_{i,j}\left[\mathbf{u}_j[\mathbf{x}]\right] \le u_{\mathrm{ct}} \\
& -u_{i,j}\left[\mathbf{u}_j[\mathbf{x}]\right] \le u_{\mathrm{ct}} \\
& \qquad \forall\, i,j \in \begin{cases} \{1,2,3,5,7,8\},\{6\} \\ \{1,3,4,5,6,7\},\{8\} \end{cases} \\
g^{\mathrm{t}}_{i,j}[\mathbf{x}]: \quad & J_k u_{i,j}\left[\mathbf{u}_j[\mathbf{x}]\right] - u_{j,j}\left[\mathbf{u}_j[\mathbf{x}]\right] \le u_{\mathrm{t}} \\
& u_{j,j}\left[\mathbf{u}_j[\mathbf{x}]\right] - J_k u_{i,j}\left[\mathbf{u}_j[\mathbf{x}]\right] \le u_{\mathrm{t}} \\
& \qquad \forall\, i,j \in \begin{cases} \{4\},\{6\} \\ \{2\},\{8\} \end{cases}
\end{aligned}
\tag{6.21}
$$

[3] We do not claim this formulation is (best) suited for the considered problem, we merely employ this formulation for demonstration of the proposed method.

The objective is to minimize the strain energy $\mathcal{E}_j$, or equivalently maximize the stiffness, by finding design variables s_k that are bounded by $\mathbb{X} = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$. Constraint $g^{\mathrm{v}}[\mathbf{x}]$ limits the maximum material usage by fraction $\overline{v} = 0.25$. The other constraints enforce a minimum displacement at the input DOFs ($g^{\mathrm{in}}_{j,j}$), limit cross talk ($g^{\mathrm{ct}}_{i,j}$) to tolerance u_{ct}, and enforce the transmission between input and output displacements ($g^{\mathrm{t}}_{i,j}$) within a tolerance u_{t}. In the next subsection these constraints will be further explained.

This problem formulation consists of standard, well-documented response functions as well as corresponding sensitivity analysis. An extensive description is therefore omitted. For an in-depth discussion on the design of compliant mechanisms using topology optimisation, the reader is referred to earlier works, such as (Ananthasuresh *et al.* 1994; Frecker *et al.* 1997; Sigmund 2001; Sigmund 1997) and the review of Cao *et al.* (2013) and references therein. For works regarding multiple degrees of freedom systems, the works by (Alonso *et al.* 2014; Frecker *et al.* 1999; Koppen *et al.* 2021b; Zhan *et al.* 2010; Zhu *et al.* 2018) may be consulted.

The proposed compound topology optimisation problem Equation (6.21) was discretised using 200 by 200 finite elements (and design variables). The design variable field is blurred using a linear convolution operator with a filter radius of two elements to eliminate modelling artefacts(Bruns *et al.* 2001).

A post-processed (via design variable thresholding) version of a solution is shown in Figure 6.3b. This solution is obtained from a uniform initial guess in 58 design iterations using the method of moving asymptotes (Svanberg 1987). This solution adheres to the constraints imposed and, thus, satisfies the design requirements on displacement transmission and maximum parasitic motion. As expected, the solution to the topology optimisation problem using an LDAS is fully equivalent to the reference method.

Note the presence of rigid bodies and hinges and their location and connections. The resulting deformation and displacements of the DOFs of interest for one of the use-cases are displayed by the prototype in Figure 6.3c. A movie of the prototype—available as supplementary material and provided on Github by Koppen *et al.* (2022b)—demonstrates that the intended functionality has been achieved.

6.5.2 Problem analysis

Let us analyse the properties of this optimisation problem in light of the proposed method, with a specific emphasis on the required number of systems of equations to be solved.

Forward analysis

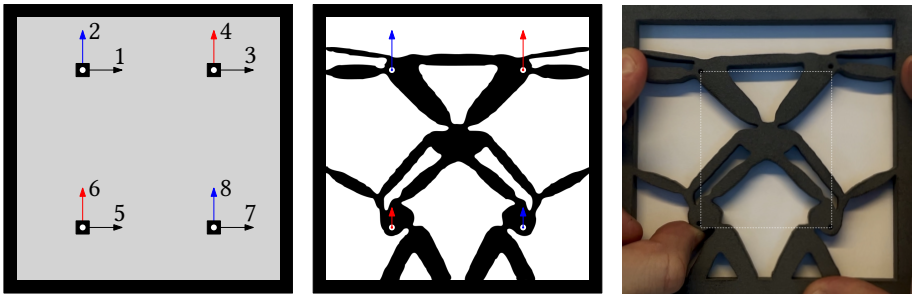
The objective function $f[\mathbf{x}]$ is a summation of strain energies, obtained by analysing the deformed structure under a unit load at DOFs $\{1,3,5,7\}$. The internal strain energy corresponding to each displacement field $\mathbf{u}_j$ reads as

$$\mathcal{E}_j = \frac{1}{2}\mathbf{u}_j \cdot \mathbf{K}[\mathbf{x}]\,\mathbf{u}_j, \tag{6.22}$$

where $\mathbf{u}_j$ is found by solving the system of equations

$$\mathbf{K}[\mathbf{x}]\,\mathbf{u}_j = \mathbf{f}_j, \tag{6.23}$$

with $\mathbf{f}_j$ the unit load vector that contains zeros at all entries except at DOF j of interest. To evaluate the objective function, the system of equations (Equation (6.23)) needs to be



(a) Initial design and degree of freedom numbering. The intended kinematic degrees of freedom are highlighted using colours; (i) motion from input DOF 6 to output DOF 4, and (ii) motion from input DOF 8 to output DOF 2.

(b) Final (post-processed) material distribution as obtained from the optimization. Arrow lengths indicate the transmissions between input and output.

(c) Prototype design in deformed configuration. The corners of the dashed line indicate the position of the DOFs of interest in undeformed configuration. A movie of the prototype is available on Github (Koppen *et al.* 2022b).

Figure 6.3: Design of a planar, decoupled multiple degrees of freedom compliant mechanism as described in Section 6.5.1. From left to right: (a) the initial design with the four points of interest each with two degrees of freedom (u_x, u_y), (b) the topology as obtained from the optimization, and (c) a prototype model in deformed configuration.

solved repeatedly, since the four physical loads are linearly independent. By minimising these strain energy terms, the motion corresponding to these DOFs is restricted in the resulting structure. None of the points of interest can significantly move in the x-direction.

Constraints $g^{\text{in}}_{j,j}[\mathbf{x}]$ are required to enforce a minimum displacement of u_{in} at $u_{j,j}$ with j the DOFs of interest 6 and 8, requiring two additional solves. Note, $u_{i,j}$ denotes the displacement at DOF i due to a unit load at DOF j. One may observe that the remaining displacement-based constraints are only dependent on $\mathbf{u}_6$ and $\mathbf{u}_8$. Since these were previously evaluated to determine $g^{\text{in}}_{j,j}[\mathbf{x}]$, inspection shows that *no* additional solves are required for the forward analysis.

Constraints $g^{\text{ct}}_{i,j}[\mathbf{x}]$ are imposed to limit the crosstalk (parasitic motion) $u_{i,j}$ of DOFs $\{1,2,3,5,7,8\}$ due to a unit load at DOF 6 *and* the motion of DOFs $\{1,3,4,5,6,7\}$ due to a unit load at DOF 8 from below by $-u_{\text{ct}}$ and from above by u_{ct}, with $u_{\text{ct}} = 0.001 u_{\text{in}}$. The number of crosstalk constraints is found by multiplying two kinematic DOF, six constraints per kinematic DOF, and two bounds per constraint, resulting in 24 constraint functions.

Constraints $g^{\text{t}}_{i,j}[\mathbf{x}]$ enforce a desired input-output transmission $J_k := \frac{u_{\text{out},k}}{u_{\text{in},k}}$ for kinematic DOF k with a maximum transmission deviation of $u_{\text{t}} = 0.1 u_{\text{in}}$. The input-output transmission for the first kinematic mode is defined as the motion transmission from DOF 2 to DOF 4 $J_1 := \frac{u_{4,6}}{u_{6,6}}$, and the second input-output transmission is defined as the motion transmission from DOF 8 to DOF 2 $J_2 := \frac{u_{2,8}}{u_{8,8}}$. This introduces four constraints, as each constraint is bound from below *and* above.

All response functions combined require 32 response functions to be evaluated for this optimisation problem, which are fully resolved by performing a total of *six* solves (four for the objective and two for $g^{\text{in}}_{j,j}[\mathbf{x}]$).

Sensitivity analysis

To obtain the sensitivities of the responses to the design variables, one generally loops over the responses, and *consecutively* calculates the corresponding sensitivities.

For the considered problem, the adjoint loads of the objective are linearly dependent on *corresponding* physical loads, i.e. they form four LDAP pairs. In this case $\frac{\partial \mathcal{E}_j}{\partial \mathbf{u}_j} = \mathbf{f}_j$, and thus $\boldsymbol{\lambda}_{j,j} = \mathbf{u}_j$. Thus, to obtain the sensitivities of the objective no additional solves are required.

The adjoint loads corresponding to $g^{\text{in}}_{j,j}[\mathbf{x}]$ read

$$\frac{\partial g^{\text{in}}_{j,j}[\mathbf{x}]}{\partial \mathbf{u}_j} = \frac{1}{u_{\text{in}}}\mathbf{l}_j, \tag{6.24}$$

which can be written as a linear combination of the physical loads $\mathbf{f}_6$ and $\mathbf{f}_8$ previously considered to evaluate $g^{\text{in}}_{j,j}[\mathbf{x}]$.

The sensitivities of the crosstalk constraints $g^{\text{ct}}_{i,j}[\mathbf{x}]$ exhibit MLDs. Furthermore, for $i = \{1,3,5,7\}$ and $j = \{6,8\}$ the following holds

$$\frac{\partial g^{\text{ct}}_{i,j}[\mathbf{x}]}{\partial \mathbf{u}_j} = \pm\frac{1}{u_{\text{ct}}}\mathbf{l}_j = \pm\frac{1}{u_{\text{ct}}}\mathbf{f}_j, \tag{6.25}$$

and the adjoint loads are therefore linearly dependent on *non-corresponding* physical loads. However, for $i,j = \{2,6\}$ and $i,j = \{4,8\}$ the adjoint load can *not* be written as (a combination) of previously evaluated physical and/or adjoint loads and the corresponding systems of equations (Equation (6.5)) need to be solved accordingly. Note, *only* two solves are required as the adjoint loads for the constraints related to lower and upper bounds are linear dependent (these only show a sign difference).

Lastly, the adjoint loads corresponding to transmission constraint $g^{\text{t}}_{i,j}[\mathbf{x}]$ are given by

$$\frac{\partial g^{\text{t}}_{i,j}[\mathbf{x}]}{\partial \mathbf{u}_j} = \pm\left(\frac{J_k}{u_{\text{t}}}\mathbf{l}_i - \frac{1}{u_{\text{t}}}\mathbf{l}_j\right), \tag{6.26}$$

which can all be written as a summation of the previous adjoint loads of $g^{\text{in}}_{i,j}[\mathbf{x}]$ (or physical loads $\mathbf{f}_6$ and $\mathbf{f}_8$ and $g^{\text{ct}}_{i,j}[\mathbf{x}]$. For such 'combined' loads it can be particularly obscure to manually express them as a linear combination of previous physical and/or adjoint loads.

Concluding remarks

The problem analysis reveals that if no linear dependencies are taken into account, 40 systems of equations need to be solved (of which 34 in the sensitivity analysis), as opposed to the minimum of 8 when considering all linear dependencies (MLDs). That is, one may expect a maximum decrease of computational effort by 80%. If only LDAP pairs are considered (this is generally the case), then 34 equations have to be solved. If, in addition to this, it is recognised that the adjoint loads of the constraints on lower and upper bounds only differ by a sign (and are thus linearly dependent), one still has to solve 20 systems of equations. The results of the foregoing problem analysis are summarised in Table 6.4, aiding in the detection of linear dependency between loads and calculation of states.

Although manually finding all linear dependencies and their corresponding coefficients is achievable and yields significant savings, it is time-consuming, cumbersome, and error-prone. Moreover, it does not readily permit implementation in commercial software. In the following, we demonstrate how an LDAS, such as Algorithm 6 provides the same result in an automated manner with negligible computational overhead.

6.5.3 Verification by run-time experiment

The following discusses a run-time measurement comparison between the LDAS and manual implementations considering LDAP pair and MLD detection. This comparison is based on the design problem as proposed and analysed in Sections 6.5.1 and 6.5.2. We aim to measure the run-time of a single design iteration using an automatic LDAS for solving the linear systems involved in a single design iteration of the problem proposed in Section 6.5.1, and compare this to the run-time required for manual implementations. In addition, we also focus on the attained performance improvements across a range of discretisations, indicated by the number of DOFs n, for a single design iteration. Assuming the physical and adjoint loads do not alter during the optimization process, the linear dependencies remain constant throughout the optimization process. Therefore, the computational effort of a complete optimization process simply scales with the number of design iterations. All presented run times are normalised to the implementation without exploiting linear dependencies. From the previous problem analysis, we found the number of solves required for each method: 40 for no detection, 34 considering LDAP, and 8 when including MLD, already hinting at potential performance improvements.

In order to consider the influence of different types of solution methods, we define the ratio χ as the ratio between the computational effort a solution method requires for preprocessing and the effort required for a solve. To capture a wide range of solution methods, we opt to compare two extremes:

- A high-χ solution method with predominant effort in the preprocessing; we opt here for a direct method, such as a Cholesky factorisation (Benoit 1924) with back-substitution, and
- A low-χ solution method with predominant effort in solving the equations. We opt here for an iterative solution process, such as Incomplete Cholesky preconditioning with Conjugate Gradient (Saad 2003).

The presented experiments consider a moderate number of DOFs: small enough to highlight the change in performance as the number of DOFs is increased while large enough to ensure the computational effort and run-time are dominated by preprocessing and solving. These aspects are therefore emphasised in the following analysis, and other computational overhead is assumed negligible[4]. In all cases, we reused the preprocessing information (factorisation/preconditioner) when possible. The results of this run-time experiment are shown in Figure 6.4. The figures show the normalised run-time $\hat{t}$, i.e. normalised to the run-time required without any linear dependency detection, of the solves required for a single design iteration, both for high and low-χ methods.

[4] Although very little computational overhead is present in the manual approaches, the required problem analysis (Section 6.5.2) is time-consuming and error-prone.

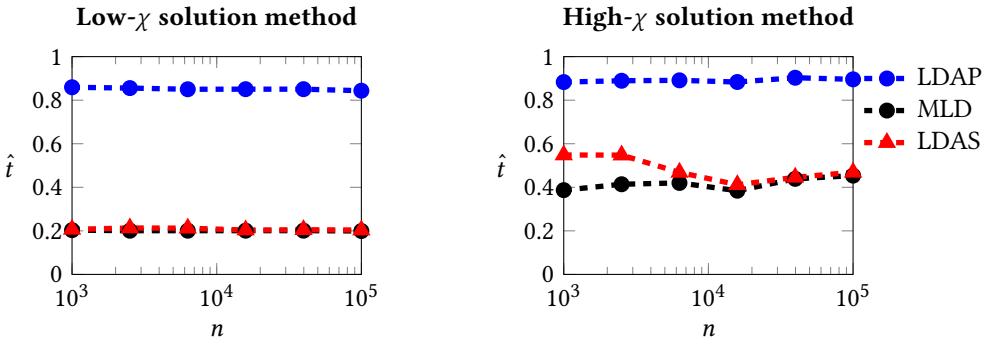


Figure 6.4: Normalized run-time $\hat{t}$ versus number of DOFs n of three implementations: LDAP (•), MLD (•) and LDAS (▲). Herein LDAP and MLD are implementations that manually detect linear dependencies. The LDAP implementation detects only adjoint-physical load pairs, whereas the MLD implementation detects all linear dependencies. The LDAS implementation uses automatic detection, with a slight overhead to the manual MLD implementation. The figures include both a high-χ and low-χ solution method to solve the system of equations related to the numerical example presented in Section 6.5. For each of the six data points, the measurements are averaged over respectively 1000, 250, 64, 16, 4 and 1 repeated experiments on a high performance computing cluster to obtain a stable time measurement.

For high-χ solution methods, the gains for LDAS and MLD converge toward each other, indicating the relative overhead of the LDAS decreases with problem size. It should be noted that the ideal normalised run time $\hat{t} = 0.2$ is not achieved for high-χ methods since the chosen preprocessing is relatively expensive (or vice versa, the solve is relatively cheap), thereby limiting the possible gains in run-time in this situation to $\hat{t} = 0.4$. Clearly, the maximum achievable gain is higher for low-χ solution methods (the difference is fully defined by the difference in χ). Counting the number of linearly independent solves of the different schemes gives an accurate estimate of relative computational efficiency. For the presented example, an 80% reduction may indeed be expected using an LDAS with a low-χ solution method.

Regardless of the solution method, taking into account only LDAP pairs is not computationally efficient compared to using an LDAS for this problem. For both high-χ and low-χ solution methods, the overhead of the LDAS is negligible for problems of moderate to large size.

Conclusions

The computational effort required to solve a gradient-based structural optimisation problem in a nested analysis and design setting is typically dominated by finding solutions to state equations. However, in real-world optimisation problems—that are typically *compound*, i.e. they consider multiple combinations of physical loading conditions and a wide variety of response functions—many avoidable linear system solves are executed regardless. This paper proposes the use of linear dependency aware solvers, complementary to methods aiming to reduce the total number of design iterations, or the cost per solve, by effectively reducing the *number* of solves per design iteration without compromising accuracy of the solution. The proposed concept leverages the linearity of the systems of equations—a trait

present in many commonly considered topology optimisation problems—to automatically omit expensive solves if the solutions can be expressed as a linear combination of previously evaluated solutions for a given design iteration.

We proposed one such algorithm that is simple, as illustrated by the provided supplementary Python and MATLAB implementations of Algorithm 6, and can be integrated non-intrusively into existing optimisation software. Although the potential benefits of the proposed method hinge on the presence of linear dependencies of the problem at hand, it has been illustrated that the accompanying overhead is negligible, allowing the method to be applied freely and achieving significant performance improvements when linear dependencies are abundant. Additionally, the concept does not restrict other methods to reduce the computational time per solve, such as parallel computing, approximation techniques, or model order reduction, which allows the user to focus on the design problem formulation and avoids laborious manual linearly dependency analysis altogether.

Loads	1	2	3	4	5	6	7	8
$\mathbf{f}_1$	1							
$\mathbf{f}_3$			1					
$\mathbf{f}_5$					1			
$\mathbf{f}_7$							1	
$\mathbf{f}_6$						1		
$\mathbf{f}_8$								1
$\frac{\partial f}{\partial \mathbf{u}_1}$	1							
$\frac{\partial f}{\partial \mathbf{u}_3}$			1					
$\frac{\partial f}{\partial \mathbf{u}_5}$					1			
$\frac{\partial f}{\partial \mathbf{u}_7}$							1	
$\frac{\partial g^{\text{in}}_{6,6}}{\partial \mathbf{u}_6}$						$\frac{1}{u_{\text{in}}}$		
$\frac{\partial g^{\text{in}}_{8,8}}{\partial \mathbf{u}_8}$								$\frac{1}{u_{\text{in}}}$
$\frac{\partial g^{\text{ct}}_{1,6}}{\partial \mathbf{u}_6}$	$\frac{1}{u_{\text{ct}}}$							
$\frac{\partial g^{\text{ct}}_{1,6}}{\partial \mathbf{u}_6}$	$-\frac{1}{u_{\text{ct}}}$							
$\frac{\partial g^{\text{ct}}_{2,6}}{\partial \mathbf{u}_6}$		$\frac{1}{u_{\text{ct}}}$						
$\frac{\partial g^{\text{ct}}_{2,6}}{\partial \mathbf{u}_6}$		$-\frac{1}{u_{\text{ct}}}$						
$\vdots$								
$\frac{\partial g^{\text{ct}}_{8,6}}{\partial \mathbf{u}_6}$								$\frac{1}{u_{\text{ct}}}$
$\frac{\partial g^{\text{ct}}_{8,6}}{\partial \mathbf{u}_6}$								$-\frac{1}{u_{\text{ct}}}$
$\frac{\partial g^{\text{ct}}_{1,8}}{\partial \mathbf{u}_8}$	$\frac{1}{u_{\text{ct}}}$							
$\frac{\partial g^{\text{ct}}_{1,8}}{\partial \mathbf{u}_8}$	$-\frac{1}{u_{\text{ct}}}$							
$\frac{\partial g^{\text{ct}}_{3,8}}{\partial \mathbf{u}_8}$			$\frac{1}{u_{\text{ct}}}$					
$\frac{\partial g^{\text{ct}}_{3,8}}{\partial \mathbf{u}_8}$			$-\frac{1}{u_{\text{ct}}}$					
$\vdots$								
$\frac{\partial g^{\text{ct}}_{7,8}}{\partial \mathbf{u}_8}$							$\frac{1}{u_{\text{ct}}}$	
$\frac{\partial g^{\text{ct}}_{7,8}}{\partial \mathbf{u}_8}$							$-\frac{1}{u_{\text{ct}}}$	
$\frac{\partial g^{\text{t}}_{4,6}}{\partial \mathbf{u}_6}$				$\frac{J_{4,6}}{u_{\text{t}}}$		$-\frac{1}{u_{\text{t}}}$		
$\frac{\partial g^{\text{t}}_{4,6}}{\partial \mathbf{u}_6}$				$-\frac{J_{4,6}}{u_{\text{t}}}$		$\frac{1}{u_{\text{t}}}$		
$\frac{\partial g^{\text{t}}_{2,8}}{\partial \mathbf{u}_8}$		$\frac{J_{2,8}}{u_{\text{t}}}$						$-\frac{1}{u_{\text{t}}}$
$\frac{\partial g^{\text{t}}_{2,8}}{\partial \mathbf{u}_8}$		$-\frac{J_{2,8}}{u_{\text{t}}}$						$\frac{1}{u_{\text{t}}}$

Table 6.4: Result of the problem analysis (Section 6.5.2); relation between loads and DOF of interest. The horizontal axis states the eight DOFs of interest, and the vertical axis the physical and adjoint loads, respectively.

7

Discussion and Recommendations

7.1 On energy-based formulations

Considering the posed limitations, we recommend the development of simple extensions, variations or modifications of the proposed energy-based formulations to gain (more accurate and preferably uncoupled) control over (i) motion or load transmission, (ii) a variety of specific stiffnesses (*e.g.* as initiated by the adaptive formulation), and (iii) strain-energy distribution and stress levels.

The performance of multi-DOF applications, such as optical mounts, are typically susceptible to thermo-mechanical and dynamic disturbances. We recommend the development of variations to the proposed formulations, including simplified or derived requirements on thermo-mechanical and dynamical stability of the points of interest. It would be interesting to see if the concept of 'motion patterns' and corresponding 'characteristic stiffness' measured via energy terms can be extended to the fields of thermo-mechanics and dynamics to minimize the sensitivity to user-defined thermo-mechanical and dynamical modes.

Stress constraints are of utmost importance when designing practically relevant compliant mechanisms. However, the impact on the mechanism performance and optimization process of such localized measures are typically high. We recommend the development of novel problem formulations that implicitly allow for controlling the stress levels with minimum burden on computational effort and optimization problem complexity.

What is more, we recommend bridging further the gap between the fields of traditional mechanism design and topology optimization. The use of static condensation has demonstrated useful in reduction of computational effort, but, possibly, even more so in understanding the compliant mechanism design problem as demonstrated throughout this dissertation, see *e.g.* Sections 3.B, 3.2.1, 5.5.2 and 6.3. To promote bridging this gap further, it is recommended that new methods explore including design requirements and synthesis tools from the field of (traditional) compliant mechanism design.

In contrast to design for stiffness, solutions to compliant mechanism design problems are particularly sensitive to the stiffness lower bound; the proposed formulations are no

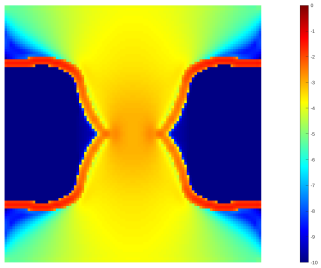
Figure 7.1: Various solutions to the problem formulation as proposed by Koppen *et al.* (2021b) (maximization of shear stiffness (ty), subjected to maximum axial stiffness (ty)) without volume constraint for various ratios of solid to void stiffness. From left to right these ratio is 1×10^3 , 1×10^6 and 1×10^9 , respectively. For each of those solutions the filtered design variable field is displayed on the right side; a post-processed design interpretation is displayed on the left.

exception, see Figure 7.1 for a range of solutions to the same non volume-constrained optimization problem with varying lower bound. Ideally, the lower bound is as low as possible, without leading to numerical issues due to poor conditioning. Proper argumentation of the magnitude of the lower bound, however, is typically absent (including this work).

This sensitivity to the lower bound can be emphasized by removal of the, commonly arbitrarily added (including this work), volume constraint. As per observation of Figure 7.1; solutions obtained with higher lower bounds tend to use less material. This is caused by the fact that solutions with higher lower bounds require more void elements in series between two solid members to achieve the same flexibility a solution with decreased lower bound may have. What is more, the resulting designs from this relaxed optimization problem typically consist of (possibly large) regions with variables that take intermediate values, see Figure 7.2 for a elucidating example based on preliminary results. Those regions are a result of the absence of sensitivity of any of the response functions with respect to the design variables, see Figure 7.2b. In a preliminary study it is found that the variables in these regions take on the value chosen as initial guess in solving the approximated sub-problems. This indicates the relative impact of lower bound magnitude and termination criteria in solving the approximated sub-problems has a large impact on the final solution of such problems.

These regions of variables that take intermediate values are generally undesired for manufacturability. A preliminary study indicates that, in contrast to design variables that take intermediate values due to optimality or filtering, the design variables in such regions may be arbitrarily chosen with limited influence on the performance, see Figure 7.2c. Such regions indicate where a tight manufacturing tolerance may be relaxed, and projection of those regions to void or solid can be beneficial for manufacturing, cost or any requirements not included in the problem formulation, see Figure 7.3. However, care should be taken and future studies are required to confirm this hypothesis. Although the sensitivities of these regions are close to zero, it is not guaranteed these sensitivities remain zero after projection.

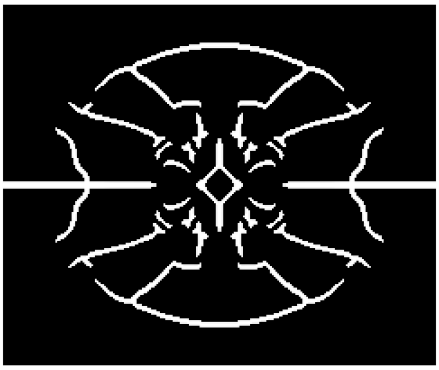
(a) Filtered design variable field of a solution that shows large regions with design variables that take intermediate values.



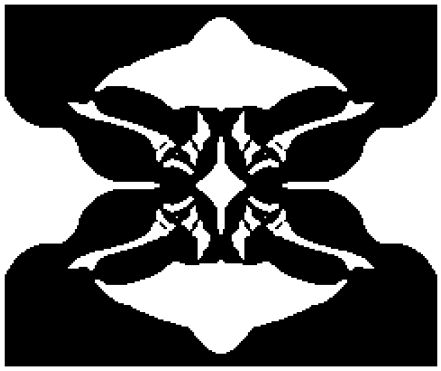
(b) Visualization of the elementwise-absolute maximum sensitivities $\max(\mathrm{d}f, \mathrm{d}g)$ at final design iteration (on $\log$ scale 1×10^{-10} blue to 1 red) of the solution posed in Figure 7.2a.

(c) Resulting solution after projection of ternary variables to solid such that for each design variable $x(\max(\mathrm{d}f, \mathrm{d}g) < 1 \times 10^{-6}) = 1$. This designs shows close performance to the design in Figure 7.2a, with slight infeasibility of the constraint function.

Figure 7.2: Preliminary results of a study on the non volume-constrained problem formulation proposed in Koppen *et al.* (2021b) (maximization of shear stiffness, subjected to maximum rotational stiffness) with high ratio of solid to void stiffness.



(a)



(b)

Figure 7.3: Example of two flexure designs based on the same solution of the non volume-constrained problem formulation proposed in Koppen *et al.* (2021b). The filtered design variable fields are projected to (a) solid and (b) void. These design have very close stiffness characteristics, yet differ substantially in terms of volume/mass and manufacturability.

7.2 On large(r) motion

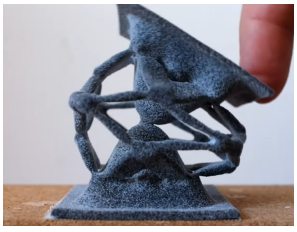
Throughout this dissertation we assumed that the compliant mechanisms under design are utilized in applications that require such short strokes that everywhere in the design deformations and rotations are so small that the use of infinitesimal strain theory is reasonably justified. This allows for major reduction of complexity and computational effort of analysis and optimization.

This implies that proper functioning of the designed compliant mechanism is solely guaranteed for, indeed, infinitesimally small displacements relative to the smallest relevant dimension of the body. To guarantee proper functioning for any finite stroke, it is therefore of utmost importance to post-analyse any of such designs using a geometrically (and possibly material) nonlinear analysis in the desired range of motion. For finite stroke mechanism design the proposed formulations should, at first, be considered as tools for the concept design phase.

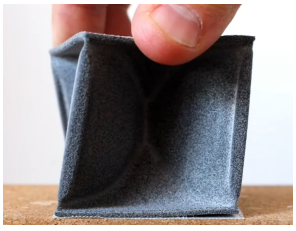
What is a reasonably accepted error due to nonlinear behavior in practice depends on the impact those non-linearities have on the function of the compliant mechanism, *e.g.* deviation of expected motion transmission or magnitude of cross-talk. The stroke is limited to the range of motion in which the performance error is sufficiently small. From our experience, applications in the high-tech sector typically justify an error up to 1 % to 10 % of parasitic motion relative to the magnitude of the stroke for most problems.

Inspection of additively manufactured prototypes, such as the planar designs presented in Figures 3.7 and 6.3b and three-dimensional prototypes of Figures 1.1 and 2.5, indicate that for some prototypes the useful stroke of the intended degrees of freedom far exceeds the useful stroke that is typically expected using linear theory, see *e.g.* Figure 7.4a. However, for other prototypes the stiffness of the intended degree of freedom rapidly increases, see Figure 7.4b. In addition, even for small prescribed motion of the degrees of constraint(s) (*e.g.* the twisting motion of designs presented in Figures 2.5a and 2.5d.) local buckling occurs and the support stiffness is fully lost, see Figure 7.4c. This observation is confirmed by the recent work of Duenser *et al.* (2021), who analyzed a subset of the designs presented in Figure 2.5 using an extension of linear eigenmodes for large-deformation analysis based on physical principles.

A substantial limitation of the infinitesimal strain theory in the context of design



(a) Example of a prototyped flexure of which the stiffness of the degree of freedom remains relatively constant over a larger range of motion.



(b) One of the prototyped flexures that shows a stiffening effect of the displayed twisting motion.

(c) A prototyped flexure that demonstrates to be particularly prone for loss of stiffness of the degree of constraint by buckling.

Figure 7.4: Prototyped designs generated using the problem formulation proposed in Koppen *et al.* (2021b). A video of these prototypes in motion is can be found in Koppen (2020).

optimization is that the set of possible response functions is limited to those expressing linear behaviour; one cannot design for stiffening or softening effects, let alone design compliant mechanisms whose functionality is based on zero-stiffness or bi-stability (Lange *et al.* 2008; Wallin *et al.* 2021). Given the complexity of analysis and optimization of compliant systems, especially in the combined form of topology optimization, the extension to nonlinear behaviour is on purpose largely avoided.

However, some attempts to bridge the gap have been made. In recent work by Hoevenaars (2021) we approximate the nonlinear behaviour around the undeformed configuration. Preliminary results indicate the promise of approximated description of mildly nonlinear effects (and thus the use of a larger set of response functions) within limited yet finite range using relatively low computational effort and without exposure to the issues originating from highly distorted finite elements. Another suggestion in Hoevenaars (2021), is to augment the, often highly complex, topology optimization problem formulation for large-deformation compliant mechanisms with lower bound constraints on the characteristic stiffness of desired constrained motion patterns. As such, this adds an intention of design for stiffness into the, otherwise, flexibility-based formulation, with marginal additional effort. Both approaches are actively researched by the author, with the aim to explore the potential use and extension of the presented energy-based formulations in the topology optimization of compliant mechanisms exhibiting larger displacements and rotations.

7.3 Further recommendations

We recommend the introduction of new benchmarks problems and/or design challenges (both single and multi-degree-of-freedom) that include at least the performance measures of motion *and* load transmission, maximum stress, input and output stiffnesses and the robustness of previous measures to (i) external stiffness variations, and (ii) manufacturing errors.

As observed in the demonstrations, the optimization process is typically characterized by distinct phases: a feasible solution is found in the first phase. This phase comes with significant design changes and variations of objective value. In the second phase, the performance is increased while satisfying the design constraints. We envision that (part of) the first phase can be skipped by a strategic choice of the initial design. Preliminary studies show that the optimization process can be sped up using initial designs inspired by rigid-body mechanisms with similar kinematics.

When size and complexity of the optimization problem increases inferior convergence properties are found to be increasingly common. We recommend adaptation of the solution method to fit the problem of interest; *i.e.* choice of approximation type and move limit strategy in a sequential approximate optimization context (Pappas 2020). This is particularly of interest when dealing with problems that contain multiple responses and/or variables of different behaviour as well as design-dependent and multi-physical problems. The author is involved in ongoing research on this topic.

8

Conclusions

Future high-tech instrumentation requires unprecedented and continuously increasing alignment precision and reliability of mirrors and lenses for customization, pre-use calibration and in-use adaptation. To meet such extreme requirements, engineers resort to the use of parallel multi-degree of freedom compliant mechanisms. The design of such monolithic elastic structures with multiple degrees of freedom is prone to introducing parasitic motions on input and output ports. Computational design approaches such as topology optimization demonstrate high potential to form the basis of a general tool for solving such complicated problems.

Multi-degree of freedom compliant mechanism optimization problems typically involve (i) multiple loading conditions, (ii) multiple loads (per loading condition), and (iii) a wide variety of stringent design constraints. As a result, straightforward optimization problem formulations are complex and limited in applicability, and related optimization problems are typically hard to solve and computationally expensive.

The stiffness at a limited number of locations and directions fully defines the behaviour of compliant mechanisms. Based on this, it is found that using static condensation, one can efficiently, and without compromising exactness, describe the mechanism behaviour and design sensitivities using low-dimensional stiffness/compliance matrices, as is common in traditional short-stroke compliant mechanism design. As such, exact reduced-order modelling provides a natural bridge between the fields of topology optimization and classical compliant mechanism design. Such reduced models allow for the development of a deeper understanding of response functions and aid in generating novel optimization problem formulations.

This dissertation contributes to the development of effective and useful problem formulations to design complex compliant structures. The proposed examples show that energy-based characteristic stiffness measures demonstrate a simple, versatile and computationally attractive basis for response functions in compliant mechanism optimization problem formulations. Such measures allow for (i) a drastic reduction of the number of stringent constraint functions and corresponding restrictions on the design space, (ii) a simple description of stiffness for arbitrarily complex motion patterns, and (iii) the use of

commercial software, typically restricted to solving ‘compliance minimization problems’, to solve such complex compliant mechanism design problems.

However, using such simple energy measures (i) comes with the loss of exact control of the stiffness and motion of independent points of interest and (ii) couples the characteristic motion stiffness to the error between intended and obtained kinematics. This study shows that additional measures can mitigate these disadvantages at the cost of increasing the complexity of the optimization formulation. However, exact transmission is not of the highest importance in many applications, such as for manually adjusted optical mounts.

Based on these conclusions, other researchers are recommended to develop modifications, variations and extensions of the proposed formulations to strengthen its use and broaden its scope of applicability, such as the use in design for mechanisms exhibiting large displacements and rotations. We foresee that this work boosts the interest and use of topology optimization to design a large variety of complex compliant mechanism design problems in academia and industry.

Bibliography

Aage, N., E. Andreassen, and B. Lazarov (2015). "Topology optimization using PETSc: An easy-to-use, fully parallel, open source topology optimization framework". In: *Struct. Multidiscip. Optim.* 51.3, pp. 565–572. DOI: 10.1007/s00158-014-1157-0.

Aage, N., E. Andreassen, B. S. Lazarov, and O. Sigmund (Oct. 2017). "Giga-voxel computational morphogenesis for structural design". In: *Nature* 550.7674, pp. 84–86. DOI: 10.1038/nature23911.

Abel, J. F. and M. S. Shephard (Jan. 1979). "An algorithm for multipoint constraints in finite element analysis". In: *International Journal for Numerical Methods in Engineering* 14.3, pp. 464–467. DOI: 10.1002/nme.1620140312.

Ahmad, A. (2017). "Adjustment Mechanisms". In: *Handbook of Optomechanical Engineering*, pp. 211–253. DOI: 10.4324/9781315153247.

Alonso, C., R. Ansola, and O. M. Querin (Oct. 2014). "Topology synthesis of multi-input-multi-output compliant mechanisms". In: *Advances in Engineering Software* 76, pp. 125–132. DOI: 10.1016/j.advengsoft.2014.05.008.

Amir, E. and O. Amir (Jan. 2019). "Topology optimization for the computationally poor: efficient high resolution procedures using beam modeling". In: *Struct. Multidiscip. Optim.* 59.1, pp. 165–184. DOI: 10.1007/s00158-018-2058-4.

Amir, O. (2015). "Revisiting approximate reanalysis in topology optimization: on the advantages of recycled preconditioning in a minimum weight procedure". In: *Struct. Multidiscip. Optim.* 51.1, pp. 41–57. DOI: 10.1007/s00158-014-1098-7.

Amir, O., N. Aage, and B. Lazarov (2014). "On multigrid-CG for efficient topology optimization". In: *Struct. Multidiscip. Optim.* 49.5, pp. 815–829. DOI: 10.1007/s00158-013-1015-5.

Amir, O., M. Bendsøe, and O. Sigmund (2009). "Approximate reanalysis in topology optimization". In: *Int. J. Numer. Methods Eng.* 78.12, pp. 1474–1491. DOI: 10.1002/nme.2536.

Amir, O. and B. S. Lazarov (Nov. 2018). "Achieving stress-constrained topological design via length scale control". In: *Struct. Multidiscip. Optim.* 58.5, pp. 2053–2071. DOI: 10.1007/s00158-018-2019-y.

Amir, O. and O. Sigmund (2010a). "On reducing computational effort in topology optimization: How far can we go?" In: *Struct. Multidiscip. Optim.* 44.1, pp. 25–29. DOI: 10.1007/s00158-010-0586-7.

Amir, O., M. Stolpe, and O. Sigmund (2010b). "Efficient use of iterative solvers in nested topology optimization". In: *Struct. Multidiscip. Optim.* 42.1, pp. 55–72. DOI: 10.1007/s00158-009-0463-4.

Ananthasuresh, G. K., S. Kota, and Y. Gianchandanif (1994). "A methodical approach to the design of compliant micromechanisms". In: *Technical Digest of the Solid-State Sensor and Actuator Workshop*. DOI: 10.31438/trf.hh1994.43.

Andreassen, E., A. Clausen, M. Schevenels, B. Lazarov, and O. Sigmund (2010). "Efficient topology optimization in MATLAB using 88 lines of code". In: *Struct. Multidiscip. Optim.* 43.1, pp. 1–16. DOI: 10.1007/s00158-010-0594-7.

Arora, J. and E. Haug (1979). "Methods of design sensitivity analysis in structural optimization". In: *AIAA J.* 17.9, pp. 970–974.

Assis Pereira, A. de and E. L. Cardoso (2018). "On the influence of local and global stress constraint and filtering radius on the design of hinge-free compliant mechanisms". In: *Struct. Multidiscip. Optim.* 58.2, pp. 641–655. DOI: 10.1007/s00158-018-1915-5.

Awtar, S., J. Ustick, and S. Sen (2012). "An XYZ Parallel-Kinematic Flexure Mechanism With Geometrically Decoupled Degrees of Freedom". In: *Journal of Mechanisms and Robotics* 5.1. DOI: 10.1115/1.4007768.

Belegundu, A. (Aug. 1986). "Interpreting adjoint equations in structural optimization". In: *J. Struct. Eng. (United States)* 112.8, pp. 1971–1976. DOI: 10.1061/(ASCE)0733-9445(1986)112:8(1971).

Bendsøe, M. P. and O. Sigmund (2004). *Topology Optimization*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-05086-6.

Bendsøe, M. P. (1989). "Optimal shape design as a material distribution problem". In: *Structural Optimization* 1.4, pp. 193–202. DOI: 10.1007/BF01650949.

Bendsøe, M. P. and N. Kikuchi (1988). "Generating optimal topologies in structural design using a homogenization method". In: *Computer Methods in Applied Mechanics and Engineering* 71, pp. 197–224. DOI: 10.1016/0045-7825(88)90086-2.

Benoit, C. (Apr. 1924). "Note Sur Une Méthode de Résolution des équations Normales Provenant de L'Application de la MéThode des Moindres Carrés a un Système D'équations Linéaires en Nombre Inférieur a Celui des Inconnues. — Application de la Méthode a la Résolution D'un Système Defini D'éQuations LinéAires". In: *Bulletin Géodésique* 2.1, pp. 67–77. DOI: 10.1007/bf03031308.

Benscoter, S. (1948). "The partitioning of matrices in structural analysis". In: *J. Appl. Mech.* 15.4, pp. 303–307.

Bensoussan, A., J. L. Lions, and G. Papanicolau (1978). *Asymptotic analysis for periodic structures*. Amsterdam: North-Holland Publishing Company.

Boer, A. de, A. H. van Zuijlen, and H. Bijl (2007). "Review of coupling methods for non-matching meshes". In: *Computer methods in applied mechanics and engineering* 196.8, pp. 1515–1525.

Borrvall, T. and J. Petersson (Sept. 2001). "Large-scale topology optimization in 3D using parallel computing". In: *Comput. Methods Appl. Mech. Eng.* 190.46-47, pp. 6201–6229. DOI: 10.1016/S0045-7825(01)00216-X.

Botkin, M. and R. Yang (1989). "Three-Dimensional Shape Optimization with Substructuring". In: *AIAA J.* 29.3, pp. 486–488. DOI: 10.2514/3.59924.

Boyd, S., S. P. Boyd, and L. Vandenberghe (2004). *Convex optimization*. Cambridge university press.

Bruns, T. and D. Tortorelli (2001). "Topology optimization of non-linear elastic structures and compliant mechanisms". In: *Comput. Methods Appl. Mech. Eng.* 190.26-27, pp. 3443–3459. DOI: 10.1016/S0045-7825(00)00278-4.

Bruyneel, M., P. Duysinx, and C. Fleury (2002). "A family of MMA approximations for structural optimization". In: *Struct. Multidiscip. Optim.* 24.4, pp. 263–276. DOI: 10.1007/s00158-002-0238-7.

Burns, R. and F. Crossley (1968). "Kinetostatic synthesis of flexible link mechanisms". In: *Mechanical Engineering*. Vol. 90. 11. ASME, p. 67.

Cao, L., A. T. Dolovich, and W. J. Zhang (2013). "On understanding of design problem formulation for compliant mechanisms through topology optimization". In: *Mechanical Sciences* 4.2, pp. 357–369. DOI: 10.5194/ms-4-357-2013.

Cavazzuti, M., A. Baldini, E. Bertocchi, D. Costi, E. Torricelli, and P. Moruzzi (July 2011). "High performance automotive chassis design: A topology optimization based approach". In: *Struct. Multidiscip. Optim.* 44.1, pp. 45–56. DOI: 10.1007/s00158-010-0578-7.

Chen, Q. and X. Zhang (2017). "The local optimum in topology optimization of compliant mechanisms". In: *Lecture Notes in Electrical Engineering*. Vol. 408, pp. 621–632. DOI: 10.1007/978-981-10-2875-5_51.

Choi, Y., G. Oxberry, D. White, and T. Kirchdoerfer (2019). "Accelerating design optimization using reduced order models". In: DOI: 10.13140/RG.2.2.16056.08965.

Cook, R., D. Malkus, M. Plesha, and R. Witt (2001). *Concepts and Applications of Finite Element Analysis*. 4th Editio. John Wiley & Sons, Ltd., pp. 40–42.

Cook, R. D. *et al.* (2007). *Concepts and applications of finite element analysis*. John wiley & sons.

Danun, A. N., P. D. Palma, C. Klahn, and M. Meboldt (2021). "Building Block Synthesis of Self-Supported Three-Dimensional Compliant Elements for Metallic Additive Manufacturing". In: *Journal of Mechanical Design* 143.5.

Davis, T. (2006). *Direct methods for sparse linear systems*. SIAM.

Deepak, S. R., M. Dinesh, D. K. Sahu, and G. K. Ananthasuresh (Feb. 2009). "A Comparative Study of the Formulations and Benchmark Problems for the Topology Optimization of Compliant Mechanisms". In: *Journal of Mechanisms and Robotics* 1.1, p. 011003. DOI: 10.1115/1.2959094.

Diaz, A. R. and M. P. Bendsøe (Mar. 1992). "Shape optimization of structures for multiple loading conditions using a homogenization method". In: *Struct. Optim.* 4.1, pp. 17–22. DOI: 10.1007/BF01894077.

Diegel, O., A. Nordin, and D. Motte (2019). *A Practical Guide to Design for Additive Manufacturing*. Springer.

Dinesh, M. and G. Ananthasuresh (2007). "A topology-optimized large-range compliant XY micro stage". In: *13th National Conference on Mechanisms and Machines (NaCoMM07), IISc, Bangalore, India*.

Duenser, S., B. Thomaszewski, R. Poranne, and S. Coros (2021). "Nonlinear Compliant Modes for Large-Deformation Analysis of Flexible Structures". In: *ACM Transactions on Graphics (Proc. ACM SIGGRAPH Asia 2021)*.

Evgrafov, A., C. Rupp, K. Maute, and M. Dunn (2008). "Large-scale parallel topology optimization using a dual-primal substructuring solver". In: *Struct. Multidiscip. Optim.* 36.4, pp. 329–345. DOI: 10.1007/s00158-007-0190-7.

Fernández, E., K.-k. Yang, S. Koppen, P. Alarcón, S. Bauduin, and P. Duysinx (2020). "Imposing minimum and maximum member size, minimum cavity size, and minimum separation distance between solid members in topology optimization". In: *Computer Methods in Applied Mechanics and Engineering* 368, p. 113157. DOI: 10.1016/j.cma.2020.113157.

Frecker, M. I., G. K. Ananthasuresh, S. Nishiwaki, N. Kikuchi, and S. Kota (June 1997). "Topological Synthesis of Compliant Mechanisms Using Multi-Criteria Optimization". In: *Journal of Mechanical Design, Transactions of the ASME* 119.2, p. 238. DOI: 10.1115/1.2826242.

Frecker, M. I., N. Kikuchi, and S. Kota (1999). "Topology optimization of compliant mechanisms with multiple outputs". In: *Structural Optimization* 17.4, pp. 269–278. DOI: 10.1007/BF01207003.

Gallego, J. A. and J. Herder (2009). "Synthesis methods in compliant mechanisms: An overview". In: *Proc. ASME Des. Eng. Tech. Conf.* 7.PARTS A AND B, pp. 193–214. DOI: 10.1115/DETC2009-86845.

Gangadharan, S., R. Haftka, and E. Nikolaidis (Apr. 1990). "Easily implemented static condensation method for structural sensitivity analysis". In: *Commun. Appl. Numer. Methods* 6.3, pp. 161–171. DOI: 10.1002/cnm.1630060302.

Gao, W., Y. Zhang, D. Ramanujan, K. Ramani, Y. Chen, C. B. Williams, C. C. Wang, Y. C. Shin, S. Zhang, and P. D. Zavattieri (2015). "The status, challenges, and future of additive manufacturing in engineering". In: *Computer-Aided Design* 69, pp. 65–89.

Garaigordobil, A., R. Ansola, E. Veguería, and I. Fernandez (Apr. 2019). "Overhang constraint for topology optimization of self-supported compliant mechanisms considering additive manufacturing". In: *CAD Computer Aided Design* 109, pp. 33–48. DOI: 10.1016/j.cad.2018.12.006.

Gaynor, A. T. and J. K. Guest (2016). "Topology optimization considering overhang constraints: Eliminating sacrificial support material in additive manufacturing through design". In: *Structural and Multidisciplinary Optimization* 54.5, pp. 1157–1172.

Gram, J. P. (1883). "Ueber die Entwickelung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate." In: *Journal für die reine und angewandte Mathematik* 1883.94, pp. 41–73.

Groen, J., M. Langelaar, O. Sigmund, and M. Ruess (June 2017). "Higher-order multi-resolution topology optimization using the finite cell method". In: *Int. J. Numer. Methods Eng.* 110.10, pp. 903–920. DOI: 10.1002/nme.5432.

Guyan, R. (1965). "Reduction of stiffness and mass matrices". In: *AIAA J.* 3.2, pp. 380–380. DOI: 10.2514/3.2874.

Hao, G. and H. Li (2015). "Conceptual designs of multi-degree of freedom compliant parallel manipulators composed of wire-beam based compliant mechanisms". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 229.3, pp. 538–555. DOI: 10.1177/0954406214535925.

Hasse, A. and L. F. Campanile (2008). "Synthesis of compliant mechanisms for shape adaptation by means of a modal procedure". In: *19th International Conference on Adaptive Structures and Technologies 2008, ICAST 2008.*

— (Sept. 2009). "Design of compliant mechanisms with selective compliance". In: *Smart Materials and Structures* 18.11, p. 115016. DOI: 10.1088/0964-1726/18/11/115016.

Hasse, A., M. Franz, and K. Mauser (2017). "Synthesis of compliant mechanisms with defined kinematics". In: *Mech. Mach. Sci.* Vol. 45. Springer, Cham, pp. 227–238. DOI: 10.1007/978-3-319-45387-3_20.

Haynsworth, E. V. (1968). *On the Schur complement.* Tech. rep. University of Basel, Institution of Mathematics.

Herder, J. (2017). "Flexure future". In: *Mikroniek: vakblad voor precisie-technologie* 57.3.

Hestenes, M. R., E. Stiefel, *et al.* (1952). *Methods of conjugate gradients for solving linear systems.* Vol. 49. 1. NBS Washington, DC.

Hoevenaars, J. (2021). "Approximate geometric non-linear analysis in topology optimization: A novel kind of structural analysis for topology optimization of finite range compliant mechanisms". MA thesis. Delft University of Technology.

Howell, L. L. (2013). "Compliant mechanisms". In: *21st century kinematics*. Springer, pp. 189–216.

— (2001). *Compliant Mechanisms*. New York: John Wiley & Sons, Ltd., pp. 189–216. DOI: 10.1007/978-1-4471-4510-3_7.

Hu, R., S. Liu, and Q. Li (May 2017). "Topology-optimization-based design method of flexures for mounting the primary mirror of a large-aperture space telescope". In: *Appl. Opt.* 56.15, p. 4551. DOI: 10.1364/ao.56.004551.

Irons, B. (May 1965). "Structural eigenvalue problems - elimination of unwanted variables". In: *AIAA J.* 3.5, pp. 961–962. DOI: 10.2514/3.3027.

Jin, M., X. Zhang, Z. Yang, and B. Zhu (Nov. 2017). "Jacobian-Based Topology Optimization Method Using an Improved Stiffness Evaluation". In: *J. Mech. Des.* 140.1, p. 011402. DOI: 10.1115/1.4038332.

— (2018). "Jacobian-Based Topology Optimization Method Using an Improved Stiffness Evaluation". In: *Journal of Mechanical Design, Transactions of the ASME* 140.1, pp. 1–11. DOI: 10.1115/1.4038332.

Jovanova, J., A. Nastevska, and M. Frecker (2019). "Tailoring energy absorption with functional grading of a contact-aided compliant mechanism". In: *Smart Materials and Structures* 28.8, p. 084003. DOI: 10.1088/1361-665x/ab281d.

Khurana, J., B. Hanks, and M. Frecker (2018). "Design for additive manufacturing of cellular compliant mechanism using thermal history feedback". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 51753. American Society of Mechanical Engineers, V02AT03A035.

Kirmse, S., L. F. Campanile, and A. Hasse (Mar. 2021a). "Synthesis of compliant mechanisms with selective compliance – An advanced procedure". In: *Mechanism and Machine Theory* 157, p. 104184. DOI: 10.1016/j.mechmachtheory.2020.104184.

— (2021b). "Synthesis of compliant mechanisms with selective compliance–An advanced procedure". In: *Mechanism and Machine Theory* 157, p. 104184.

— (2021c). *Topology-optimization based design of multi-degree-of-freedom compliant mechanisms (mechanisms with multiple pseudo-mobility)*.

Kirsch, U. (Sept. 1993). "Efficient reanalysis for topological optimization". In: *Struct. Optim.* 6.3, pp. 143–150. DOI: 10.1007/BF01743505.

Kirsch, U. (1991). "Reduced basis approximations of structural displacements for optimaldesign". In: *AIAA journal* 29.10, pp. 1751–1758.

Klarbring, A. and N. Strömberg (2012). "A note on the min-max formulation of stiffness optimization including non-zero prescribed displacements". In: *Structural and Multidisciplinary Optimization* 45.1, pp. 147–149. DOI: 10.1007/s00158-011-0674-3.

— (2013). "Topology optimization of hyperelastic bodies including non-zero prescribed displacements". In: *Structural and Multidisciplinary Optimization* 47.1, pp. 37–48. DOI: 10.1007/s00158-012-0819-z.

Koppen, S. (2020). *Topology optimization of 3D compliant mechanism flexures*. Youtube. URL: https://youtu.be/jHqkL6dLo2M.

Koppen, S., E. Hoes, M. Langelaar, and M. I. Frecker (Aug. 2021a). "Local Redesign for Additive Manufacturability of Compliant Mechanisms Using Topology Optimization". In: vol. Volume 8A: 45th Mechanisms and Robotics Conference (MR). International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. DOI: 10.1115/DETC2021-67642.

Koppen, S., M. van der Kolk, S. van den Boom, and M. Langelaar (2022a). "Efficient computation of states and sensitivities for compound structural optimisation problems using a Linear Dependency Aware Solver (LDAS)". In: *Structural and Multidisciplinary Optimization* 65.9, pp. 1–14. DOI: 10.1007/s00158-022-03378-8.

— (2022b). *LDAS*. https://github.com/artofscience/ldas.

Koppen, S., M. van der Kolk, F. C. M. van Kempen, J. de Vreugd, and M. Langelaar (Sept. 2018). "Topology optimization of multicomponent optomechanical systems for improved optical performance". In: *Structural and Multidisciplinary Optimization* 58.3, pp. 885–901. DOI: 10.1007/s00158-018-1932-4.

Koppen, S., M. Langelaar, and F. van Keulen (June 2021b). "A simple and versatile topology optimization formulation for flexure synthesis". In: *Mechanism and Machine Theory* 172, p. 104743. DOI: 10.1016/j.mechmachtheory.2022.104743.

— (2022c). "Efficient multi-partition topology optimization". In: *Computer Methods in Applied Mechanics and Engineering* 393, p. 114829. DOI: 10.1016/j.cma.2022.114829.

— (2022d). *Flexures*. https://github.com/artofscience/flexures.

— (2022e). "Synthesis of multi-degree of freedom compliant mechanisms via the kinetoelastic energy-based topology optimization method: principles and applications". In: *In review for publication in Structural and Multidisciplinary Optimization*.

Kota, S., J. Joo, Z. Li, S. M. Rodgers, and J. Sniegowski (2001). "Design of compliant mechanisms: applications to MEMS". In: *Analog integrated circuits and signal processing* 29.1, pp. 7–15. DOI: 10.1023/A:1011265810471.

Kron, G. (1955). "Solving highly complex elastic structures in easy stages". In: *J. Appl. Mech.* 22.2, pp. 235–244.

Kuo, Y.-H., C.-C. Cheng, Y.-S. Lin, and C.-H. San (2018). "Support structure design in additive manufacturing based on topology optimization". In: *Structural and Multidisciplinary Optimization* 57.1, pp. 183–195.

Lange, D. J. d., M. Langelaar, and J. L. Herder (2008). "Towards the design of a statically balanced compliant laparoscopic grasper using topology optimization". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 43260, pp. 293–305.

Langelaar, M. (2016). "Topology optimization of 3D self-supporting support structures for additive manufacturing". In: *Additive Manufacturing* 21, pp. 666–682. DOI: 10.1016/j.addma.2018.04.016.

— (2017). "An additive manufacturing filter for topology optimization of print-ready designs". In: *Structural and multidisciplinary optimization* 55.3, pp. 871–883.

— (2018). "Combined optimization of part topology, support structure layout and build orientation for additive manufacturing". In: *Structural and Multidisciplinary Optimization* 57.5, pp. 1985–2004.

Laplace, P. S. (1820). *Théorie analytique des probabilités*. Courcier.

Larsen, U., O. Sigmund, and S. Bouwstra (June 1997). "Design and fabrication of compliant micromechanisms and structures with negative Poisson's ratio". In: *Journal of Microelectromechanical Systems* 6.2, pp. 99–106. DOI: 10.1109/84.585787.

Lazarov, B., M. Schevenels, and O. Sigmund (Aug. 2011). "Robust design of large-displacement compliant mechanisms". In: *Mech. Sci.* 2.2, pp. 175–182. DOI: 10.5194/ms-2-175-2011.

Lazarov, B., F. Wang, and O. Sigmund (2016). "Length scale and manufacturability in density-based topology optimization". In: *Arch. Appl. Mech.* 86.1-2, pp. 189–218. DOI: 10.1007/s00419-015-1106-4.

Leon, D. M. de, J. Alexandersen, J. S. O. Fonseca, O. Sigmund, D. M. de Leon, J. Alexandersen, J. S. O. Fonseca, and O. Sigmund (Nov. 2015). "Stress-constrained topology optimization for compliant mechanism design". In: *Struct. Multidiscip. Optim.* 52.5, pp. 929–943. DOI: 10.1007/s00158-015-1279-z.

Leon, S. J., Å. Björck, and W. Gander (2013). "Gram-Schmidt orthogonalization: 100 years and more". In: *Numerical Linear Algebra with Applications* 20.3, pp. 492–532.

Li, L. and K. Khandelwal (2015). "An adaptive quadratic approximation for structural and topology optimization". In: *Computers & Structures* 151, pp. 130–147.

Li, L. and X. Zhu (Oct. 2019). "Design of compliant revolute joints based on mechanism stiffness matrix through topology optimization using a parameterization level set method". In: *Struct. Multidiscip. Optim.* 60.4, pp. 1475–1489. DOI: 10.1007/s00158-019-02278-8.

Linss, S., S. Henning, and L. Zentner (Sept. 2019). "Modeling and Design of Flexure Hinge-Based Compliant Mechanisms". In: *Kinemat. - Anal. Appl.* IntechOpen. DOI: 10.5772/intechopen.85224.

Liu, F., W. Li, W. Zhao, X. Wang, and X. Wang (Dec. 2021a). "Fast Optimization Design of the Flexure for a Space Mirror Based on Mesh Deformation". In: *Photonics* 8.12, p. 567. DOI: 10.3390/photonics8120567.

Liu, M., J. Zhan, and X. Zhang (May 2020). "Topology optimization of distributed flexure hinges with desired performance". In: *Eng. Optim.* 52.3, pp. 405–425. DOI: 10.1080/0305215X.2019.1595612.

Liu, M., J. Zhan, B. Zhu, and X. Zhang (Oct. 2018). "Topology Optimization of Flexure Hinges with Distributed Stress for Flexure-Based Mechanisms". In: *MARSS 2018 - International Conference on Manipulation, Automation and Robotics at Small Scales*. Institute of Electrical and Electronics Engineers Inc. DOI: 10.1109/MARSS.2018.8481150.

— (2021b). "Topology optimization of flexure hinges with a prescribed compliance matrix based on the adaptive spring model and stress constraint". In: *Precision Engineering*.

Liu, M., X. Zhang, and S. Fatikow (Jan. 2015). "Topology optimization of large-displacement flexure hinges". In: *Proc. ASME Des. Eng. Tech. Conf.* Vol. 5A-2015. American Society of Mechanical Engineers (ASME). DOI: 10.1115/DETC2015-46444.

— (May 2016). "Design and analysis of a high-accuracy flexure hinge". In: *Rev. Sci. Instrum.* 87.5, p. 055106. DOI: 10.1063/1.4948924.

— (2017a). "Design and analysis of a multi-notched flexure hinge for compliant mechanisms". In: *Precis. Eng.* 48, pp. 292–304. DOI: 10.1016/j.precisioneng.2016.12.012.

— (Dec. 2017b). "Design of flexure hinges based on stress-constrained topology optimization". In: *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* 231.24, pp. 4635–4645. DOI: 10.1177/0954406216671346.

Liu, Z. and J. Korvink (2009). "Using artificial reaction force to design compliant mechanism with multiple equality displacement constraints". In: *Finite Elements in Analysis and Design* 45.8-9, pp. 555–568. DOI: 10.1016/j.finel.2009.03.005.

Lu, K.-J. and S. Kota (2003). "Design of compliant mechanisms for morphing structural shapes". In: *Journal of intelligent material systems and structures* 14.6, pp. 379–391. DOI: 10.1177/1045389X03035563.

Ma, Z. .-., N. Kikuchi, and I. Hagiwara (Dec. 1993). "Structural topology and shape optimization for a frequency response problem". In: *Computational Mechanics* 13.3, pp. 157–174. DOI: 10.1007/bf00370133.

Ma, Z.-D., H.-C. Cheng, and N. Kikuchi (1994). "Structural design for obtaining desired eigenfrequencies by using the topology and shape optimization method". In: *Computing Systems in Engineering* 5.1, pp. 77–89. DOI: 10.1016/0956-0521(94)90039-6.

Machekposhti, D. F., N. Tolou, and J. L. Herder (2015). "A review on compliant joints and rigid-body constant velocity universal joints toward the design of compliant homokinetic couplings". In: *J. Mech. Des. Trans. ASME* 137.3. DOI: 10.1115/1.4029318.

Machekposhti, D. F., N. Tolou, and J. Herder (2018). "A statically balanced fully compliant power transmission mechanism between parallel rotational axes". In: *Mechanism and Machine Theory* 119, pp. 51–60. DOI: 10.1016/j.mechmachtheory.2017.08.018.

Midha, A., T. W. Norton, and L. L. Howell (Mar. 1994). "On the Nomenclature, Classification, and Abstractions of Compliant Mechanisms". In: *Journal of Mechanical Design* 116.1, pp. 270–279. DOI: 10.1115/1.2919358.

Mirth, J. A. (2016). "The Design and Prototyping of Complex Compliant Mechanisms via Multi-Material Additive Manufacturing Techniques". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 50152. American Society of Mechanical Engineers, V05AT07A003.

Mukherjee, S., D. Lu, B. Raghavan, P. Breitkopf, S. Dutta, M. Xiao, and W. Zhang (Jan. 2021). "Accelerating Large-scale Topology Optimization: State-of-the-Art and Challenges". In: *Archives of Computational Methods in Engineering* 1, p. 3. DOI: 10.1007/s11831-021-09544-3.

Noor, A., H. Kamel, and R. Fulton (1978). "Substructuring techniques-status and projections". In: *Comput. Struct.* 8.5, pp. 621–632. DOI: 10.1016/0045-7949(78)90100-1.

Ouden, M. van den, S. Koppen, and M. Langelaar (2021). "Topology Optimization for Thermoelastic Metamaterials". In: vol. IS24 - Optimal Design of Structures and Metamaterials: Innovative Techniques for Engineering Applications, 2021. 9th edition of the International Conference on Computational Methods for Coupled Problems in Science and Engineering. DOI: 10.23967/coupled.2021.058.

Pappas, I. (2020). "Adaptive Structural Optimization Methods". MA thesis. Delft University of Technology.

Patel, Y. D. and P. M. George (Aug. 2012). "Parallel Manipulators Applications—A Survey". In: *Modern Mechanical Engineering* 02.03, pp. 57–64. DOI: 10.4236/mme.2012.23008.

Pedersen, P. and N. L. Pedersen (2011). "Design objectives with non-zero prescribed support displacements". In: *Structural and Multidisciplinary Optimization* 43.2, pp. 205–214. DOI: 10.1007/s00158-010-0577-8.

Pellens, J., G. Lombaert, B. Lazarov, and M. Schevenels (2019). "Combined length scale and overhang angle control in minimum compliance topology optimization for additive manufacturing". In: *Structural and Multidisciplinary Optimization* 59.6, pp. 2005–2022.

Pinskier, J. and B. Shirinzadeh (2019). "Topology optimization of leaf flexures to maximize in-plane to out-of-plane compliance ratio". In: *Precis. Eng.* 55.July 2018, pp. 397–407. DOI: 10.1016/j.precisioneng.2018.10.008.

Pinskier, J., B. Shirinzadeh, M. Ghafarian, T. K. Das, A. Al-Jodah, and R. Nowell (2020). "Topology optimization of stiffness constrained flexure-hinges for precision and range maximization". In: *Mech. Mach. Theory* 150, p. 103874. DOI: 10.1016/j.mechmachtheory.2020.103874.

Prager, W. (1968). "Optimality criteria in structural design". In: *Proceedings of the National Academy of Sciences* 61.3, pp. 794–796. DOI: 10.1073/pnas.61.3.794.

Przemieniecki, J. (1963). "Matrix structural analysis of substructures". In: *AIAA J.* 1.1, pp. 138–147.

Qian, X. and O. Sigmund (2013). "Topological design of electromechanical actuators with robustness toward over- and under-etching". In: *Computer Methods in Applied Mechanics and Engineering* 253, pp. 237–251. DOI: 10.1016/j.cma.2012.08.020.

Qiu, L., X. Yue, and Z. Xie (2019). "Design and analysis of Multicavity Flexure Hinge (MCFH) based on three-dimensional continuum topology optimization". In: *Mechanism and Machine Theory* 139, pp. 21–33. DOI: 10.1016/j.mechmachtheory.2019.04.004.

Qiu, L., X. Yue, L. Zheng, and Y. Li (May 2020). "Design and analysis of porous flexure hinge based on dual-objective topology optimization of three-dimensional continuum". In: *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 42.5, p. 225. DOI: 10.1007/s40430-020-02312-7.

Ramesh, T., R. Bharanidaran, and V. Gopal (Sept. 2014). "Design and Development of XY Micro-Positioning Stage Using Modified Topology Optimization Technique". In: *Dynamics of Machines and Mechanisms, Industrial Research.* Vol. 592. Applied Mechanics and Materials. Trans Tech Publications Ltd, pp. 2220–2224. DOI: 10.4028/www.scientific.net/AMM.592-594.2220.

Rong, J., X. Rong, L. Peng, J. Yi, and Q. Zhou (2021). "A new method for optimizing the topology of hinge-free and fully decoupled compliant mechanisms with multiple inputs and multiple outputs". In: *International Journal for Numerical Methods in Engineering* 122.12, pp. 2863–2890. DOI: 10.1002/nme.6644.

Rosen, D. (2014). "Design for additive manufacturing: past, present, and future directions". In: *Journal of Mechanical Design* 136.9.

Rozvany, G., O. Sigmund, T. Lewiński, D. Gerdes, and T. Birker (Sept. 1993). "Exact optimal structural layouts for non-self-adjoint problems". In: *Structural Optimization* 5.3, pp. 204–206. DOI: 10.1007/bf01743359.

Rozvany, G. I., M. Zhou, M. Rotthaus, W. Gollub, and F. Spengemann (Mar. 1989). "Continuum-type optimality criteria methods for large finite element systems with a displacement constraint. Part I". In: *Structural Optimization* 1.1, pp. 47–72. DOI: 10.1007/BF01743809.

Saad, Y. (Jan. 2003). *Iterative Methods for Sparse Linear Systems.* Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9780898718003.

Sanchez-Palencia, E. (1980). "Non-Homogeneous Media and Vibration Theory". In: *Non-Homogeneous Media Vib. Theory*. Lecture Notes in Physics. Springer Berlin Heidelberg. DOI: 10.1007/3-540-10000-8.

Saxena, A. (2005). "Topology design of large displacement compliant mechanisms with multiple materials and multiple output ports". In: *Structural and Multidisciplinary Optimization* 30.6, pp. 477–490. DOI: 10.1007/s00158-005-0535-z.

Saxena, A. and G. K. Ananthasuresh (Aug. 2001). "Topology optimization of compliant mechanisms with strength considerations". In: *Mech. Struct. Mach.* 29.2, pp. 199–221. DOI: 10.1081/SME-100104480.

Schevenels, M., B. Lazarov, and O. Sigmund (Dec. 2011). "Robust topology optimization accounting for spatially varying manufacturing errors". In: *Comput. Methods Appl. Mech. Eng.* 200.49-52, pp. 3613–3627. DOI: 10.1016/j.cma.2011.08.006.

Schmidt, E. (1907). "Zur Theorie der linearen und nicht linearen Integralgleichungen Zweite Abhandlung". In: *Mathematische Annalen* 64.2, pp. 161–174.

Schmit, L. and M. Hirokazu (1976). *Approximation concepts for efficient structural synthesis*. Tech. rep. NASA, pp. 1–314.

Schur, I. (1917). "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind [I]". In: *J. für die Reine und Angew. Math.* 147, pp. 205–232.

Shield, R. T. and W. Prager (1970a). "Optimal structural design for given deflection". In: *Zeitschrift für angewandte Mathematik und Physik ZAMP* 21.4, pp. 513–523.

— (July 1970b). "Optimal structural design for given deflection". In: *Zeitschrift für angewandte Mathematik und Physik ZAMP* 21.4, pp. 513–523. DOI: 10.1007/BF01587681.

Shih, C. J., C. F. Lin, and H. Y. Chen (Jan. 2006). "An integrated design of flexure hinges and topology optimization for monolithic compliant mechanism". In: *Journal of Integrated Design and Process Science* 10.3, pp. 1–16.

Sigmund, O. (Oct. 2001). "Design of multiphysics actuators using topology optimization - Part I: One-material structures". In: *Computer Methods in Applied Mechanics and Engineering* 190.49-50, pp. 6577–6604. DOI: 10.1016/S0045-7825(01)00251-1.

Sigmund, O. (Sept. 1994). "Materials with prescribed constitutive parameters: An inverse homogenization problem". In: *Int. J. Solids Struct.* 31.17, pp. 2313–2329. DOI: 10.1016/0020-7683(94)90154-6.

— (June 1995). "Tailoring materials with prescribed elastic properties". In: *Mech. Mater.* 20.4, pp. 351–368. DOI: 10.1016/0167-6636(94)00069-7.

— (Jan. 1997). "On the design of compliant mechanisms using topology optimization". In: *Mechanics of Structures and Machines* 25.4, pp. 493–524. DOI: 10.1080/08905459708945415.

— (2007). "Morphology-based black and white filters for topology optimization". In: *Struct. Multidiscip. Optim.* 33.4-5, pp. 401–424. DOI: 10.1007/s00158-006-0087-x.

— (2009). "Manufacturing tolerant topology optimization". In: *Acta Mech. Sin. Xuebao* 25.2, pp. 227–239. DOI: 10.1007/s10409-009-0240-z.

Sigmund, O. and J. Petersson (1998). "Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima". In: *Structural Optimization* 16.1, pp. 68–75. DOI: 10.1007/BF01214002.

Sigmund, O. and S. Torquato (1997). "Design of materials with extreme thermal expansion using a three-phase topology optimization method". In: *Journal of the Mechanics and Physics of Solids* 45.6, pp. 1037–1067. DOI: 10.1016/S0022-5096(96)00114-7.

Silva, G. A. da, A. T. Beck, and O. Sigmund (June 2019). "Topology optimization of compliant mechanisms with stress constraints and manufacturing error robustness". In: *Computer Methods in Applied Mechanics and Engineering*. DOI: 10.1016/j.cma.2019.05.046.

Silva, G. A. da, E. L. Cardoso, and A. T. Beck (Feb. 2020). "Comparison of robust, reliability-based and non-probabilistic topology optimization under uncertain loads and stress constraints". In: *Probabilistic Engineering Mechanics*, p. 103039. DOI: 10.1016/j.probengmech.2020.103039.

Sung, M. K., S. Lee, and D. E. Burns (Jan. 2022). "Robust topology optimization of a flexural structure considering multi-stress performance for force sensing and structural safety". In: *Structural and Multidisciplinary Optimization* 65.1, p. 6. DOI: 10.1007/s00158-021-03088-7.

Svanberg, K. (Feb. 1987). "The method of moving asymptotes—a new method for structural optimization". In: *International Journal for Numerical Methods in Engineering* 24.2, pp. 359–373. DOI: 10.1002/nme.1620240207.

Vanderplaats, G. (1980). "Comment on "Methods of Design Sensitivity Analysis in Structural Optimization"". In: *AIAA J.* 18.11, pp. 1406–1407.

Ven, E. van de, C. Ayas, M. Langelaar, R. Maas, and F. van Keulen (2001). "Accessibility of support structures in topology optimization for additive manufacturing". In: *International Journal for Numerical Methods in Engineering*.

Ven, E. van de, R. Maas, C. Ayas, M. Langelaar, and F. van Keulen (2018). "Continuous front propagation-based overhang control for topology optimization with additive manufacturing". In: *Structural and Multidisciplinary Optimization* 57.5, pp. 2075–2091.

— (2020). "Overhang control based on front propagation in 3D topology optimization for additive manufacturing". In: *Computer Methods in Applied Mechanics and Engineering* 369, p. 113169.

Verbart, A., M. Langelaar, and F. van Keulen (Feb. 2017). "A unified aggregation and relaxation approach for stress-constrained topology optimization". In: *Struct. Multidiscip. Optim.* 55.2, pp. 663–679. DOI: 10.1007/s00158-016-1524-0.

Vukobratovich, D. and R. M. Richard (1988). "Flexure mounts for high-resolution optical elements". In: *Optomechanical and Electro-Optical Design of Industrial Systems*. Vol. 959. SPIE, pp. 18–36. DOI: 10.1117/12.947774.

Wallin, M., A. Dalklint, and D. Tortorelli (2021). "Topology optimization of bistable elastic structures–An application to logic gates". In: *Computer Methods in Applied Mechanics and Engineering* 383, p. 113912.

Wang, F., B. Lazarov, and O. Sigmund (2011). "On projection methods, convergence and robust formulations in topology optimization". In: *Struct. Multidiscip. Optim.* 43.6, pp. 767–784. DOI: 10.1007/s00158-010-0602-y.

Wang, M. Y. (2009a). "A kinetoelastic formulation of compliant mechanism optimization". In: *Journal of Mechanisms and Robotics* 1.2, pp. 1–10. DOI: 10.1115/1.3056476.

— (May 2009b). "Mechanical and geometric advantages in compliant mechanism optimization". In: *Front. Mech. Eng. China* 4.3, pp. 229–241. DOI: 10.1007/s11465-009-0066-1.

Wang, M. Y. and S. Chen (2009). "Compliant mechanism optimization: Analysis and design with intrinsic characteristic stiffness". In: *Mech. Based Des. Struct. Mach.* 37.2, pp. 183–200. DOI: 10.1080/15397730902761932.

Wilson, E. (Jan. 1974). "The static condensation algorithm". In: *Int. J. Numer. Methods Eng.* 8.1, pp. 198–203. DOI: 10.1002/nme.1620080115.

Wu, J., C. Dick, and R. Westermann (2016). "A System for High-Resolution Topology Optimization". In: *IEEE Trans. Vis. Comput. Graph.* 22.3, pp. 1195–1208. DOI: 10.1109/TVCG.2015.2502588.

Wu, Z. and Q. Xu (2018). "Survey on recent designs of compliant micro-/nano-positioning stages". In: *Actuators*. Vol. 7. 1. Multidisciplinary Digital Publishing Institute, p. 5. DOI: 10.3390/act7010005.

Wu, Z., L. Xia, S. Wang, and T. Shi (2019). "Topology optimization of hierarchical lattice structures with substructuring". In: *Comput. Methods Appl. Mech. Eng.* 345, pp. 602–617. DOI: 10.1016/j.cma.2018.11.003.

Xia, L., J.-H. Zhu, and W.-H. Zhang (May 2012). "A superelement formulation for the efficient layout design of complex multi-component system". In: *Struct. Multidiscip. Optim.* 45.5, pp. 643–655. DOI: 10.1007/s00158-011-0720-1.

Yang, R. and C. Lu (1996). "Topology Optimization with Superelements". In: *AIAA J.* 34.7, pp. 1533–1535. DOI: 10.2514/3.60028.

Yano, M., J. D. Penn, G. Konidaris, and A. T. Patera (2012). "Math, Numerics, & Programming (for Mechanical Engineers)". In: September. MIT.

Yin, L. and G. K. Ananthasuresh (2003). "Design of distributed compliant mechanisms". In: *Mechanics Based Design of Structures and Machines* 31.2, pp. 151–179. DOI: 10.1081/SME-120020289.

Yoon, G.-H. (2010). "Structural topology optimization for frequency response problem using model reduction schemes". In: *Comput. Methods Appl. Mech. Eng.* 199.25-28, pp. 1744–1763. DOI: 10.1016/j.cma.2010.02.002.

Yoon, G.-H., Y.-S. Joung, and Y.-Y. Kim (Feb. 2007). "Optimal layout design of three-dimensional geometrically non-linear structures using the element connectivity parameterization method". In: *Int. J. Numer. Methods Eng.* 69.6, pp. 1278–1304. DOI: 10.1002/nme.1808.

Zhan, J. and X. Zhang (2010). "Topology optimization of multiple inputs and multiple outputs compliant mechanisms using the ground structure". In: *ICIMA 2010 - 2010 2nd International Conference on Industrial Mechatronics and Automation* 1, pp. 20–24. DOI: 10.1109/ICINDMA.2010.5538111.

Zhang, W.-H. and S. Sun (Nov. 2006). "Scale-related topology optimization of cellular materials and structures". In: *Int. J. Numer. Methods Eng.* 68.9, pp. 993–1011. DOI: 10.1002/nme.1743.

Zhang, X. and B. Zhu (2018). "Topology Optimization of Flexure Hinges". In: *Topol. Optim. Compliant Mech.* Springer Singapore, pp. 25–80. DOI: 10.1007/978-981-13-0432-3_2.

Zhang, X. S., E. de Sturler, and G. H. Paulino (Oct. 2017). "Stochastic sampling for deterministic structural topology optimization with many load cases: Density-based and ground structure approaches". In: *Comput. Methods Appl. Mech. Eng.* 325, pp. 463–487. DOI: 10.1016/j.cma.2017.06.035.

Zhang, X. S., E. de Sturler, and A. Shapiro (May 2020). "Topology Optimization With Many Right-Hand Sides Using Mirror Descent Stochastic Approximation—Reduction From Many to a Single Sample". In: *J. Appl. Mech.* 87.5. DOI: 10.1115/1.4045902.

Zhu, B., Q. Chen, M. Jin, and X. Zhang (Aug. 2018). "Design of fully decoupled compliant mechanisms with multiple degrees of freedom using topology optimization". In: *Mechanism and Machine Theory* 126, pp. 413–428. DOI: 10.1016/j.mechmachtheory.2018.04.028.

Zhu, B., X. Zhang, and S. Fatikow (Mar. 2014). "Design of single-axis flexure hinges using continuum topology optimization method". In: *Sci. China Technol. Sci.* 57.3, pp. 560–567. DOI: 10.1007/s11431-013-5446-4.

Zhu, B., X. Zhang, H. Zhang, J. Liang, H. Zang, H. Li, and R. Wang (Jan. 2020). "Design of compliant mechanisms using continuum topology optimization: A review". In: *Mech. Mach. Theory* 143, p. 103622. DOI: 10.1016/J.MECHMACHTHEORY.2019.103622.

Zhu, D., Y. Feng, and W. Zhan (2019). "Topology optimization of three-translational degree-of-freedom spatial compliant mechanism". In: *Advances in Mechanical Engineering* 11.2, p. 1687814019828228. DOI: 10.1177/1687814019828228.

Curriculum Vitæ

Stijn Koppen

24-03-1993 Born in Haarlem, the Netherlands

2022–2024 Postdoctoral Researcher
at Delft University of Technology
Design optimization of compliant shape-morphing surgical tools
Supervision: dr. ir. L. Noël and dr. ir. A. Sakes

2018–2022 Doctor of Philosophy (Phd) in Structural Optimization
at Delft University of Technology
Topology optimization of compliant mechanisms with multiple degrees of freedom
Promotors: Prof. dr. ir. A. van Keulen and dr. ir. M. Langelaar

2015–2017 Master of Science (MSc) in Mechanical Engineering
at Delft University of Technology
Topology optimization of optomechanical systems
at TNO Optomechatronics, Delft

2014–2015 Pre-master in Mechanical Engineering
at Delft University of Technology

2011–2015 Bachelor of Engineering (BEng) in Mechanical Engineering
at The Hague University of Applied Sciences
Design of the overlay measurement tool coarse positioner
at TNO Optomechatronics, Delft

List of Publications

S. Koppen, M. Langelaar, and F. van Keulen (2022e). "Synthesis of multi-degree of freedom compliant mechanisms via the kinetoelastic energy-based topology optimization method: principles and applications". In: *In review for publication in Structural and Multidisciplinary Optimization*

S. Koppen, M. van der Kolk, S. van den Boom, and M. Langelaar (2022a). "Efficient computation of states and sensitivities for compound structural optimisation problems using a Linear Dependency Aware Solver (LDAS)". in: *Structural and Multidisciplinary Optimization* 65.9, pp. 1–14. DOI: 10.1007/s00158-022-03378-8

S. Koppen, M. Langelaar, and F. van Keulen (2022c). "Efficient multi-partition topology optimization". In: *Computer Methods in Applied Mechanics and Engineering* 393, p. 114829. DOI: 10.1016/j.cma.2022.114829

S. Koppen, E. Hoes, M. Langelaar, and M. I. Frecker (Aug. 2021a). "Local Redesign for Additive Manufacturability of Compliant Mechanisms Using Topology Optimization". In: vol. Volume 8A: 45th Mechanisms and Robotics Conference (MR). International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. DOI: 10.1115/DETC2021-67642

M. van den Ouden, S. Koppen, and M. Langelaar (2021). "Topology Optimization for Thermoelastic Metamaterials". In: vol. IS24 - Optimal Design of Structures and Metamaterials: Innovative Techniques for Engineering Applications, 2021. 9th edition of the International Conference on Computational Methods for Coupled Problems in Science and Engineering. DOI: 10.23967/coupled.2021.058

S. Koppen, M. Langelaar, and F. van Keulen (June 2021b). "A simple and versatile topology optimization formulation for flexure synthesis". In: *Mechanism and Machine Theory* 172, p. 104743. DOI: 10.1016/j.mechmachtheory.2022.104743

E. Fernández, K.-k. Yang, S. Koppen, P. Alarcón, S. Bauduin, and P. Duysinx (2020). "Imposing minimum and maximum member size, minimum cavity size, and minimum separation distance between solid members in topology optimization". In: *Computer Methods in Applied Mechanics and Engineering* 368, p. 113157. DOI: 10.1016/j.cma.2020.113157

S. Koppen, M. van der Kolk, F. C. M. van Kempen, J. de Vreugd, and M. Langelaar (Sept. 2018). "Topology optimization of multicomponent optomechanical systems for improved optical performance". In: *Structural and Multidisciplinary Optimization* 58.3, pp. 885–901. DOI: 10.1007/s00158-018-1932-4

Included in this thesis.

Conference talks

S. Koppen, M. Langelaar and F. van Keulen (2021). "Topology optimization of flexures: a simple and versatile formulation". At: *14th World Congress on Structural and Multidisciplinary Optimization*, May 2021. Colorado, Boulder (virtual).

S. Koppen, M. Langelaar and F. van Keulen (2019). "Synthesis of multi-input-multi-output compliant mechanisms with defined transmission ratios using density-based topology optimization". At: *13th World Congress on Structural and Multidisciplinary Optimization*, May 2019. Beijing, China.

S. Koppen, M. van der Kolk, F.C.M. van Kempen, J. de Vreugd and M. Langelaar (2018). "Integrated multicomponent topology optimization of optomechanical systems". At: *6th International Conference on Engineering Optimization (EngOpt)*, September 2018. Lisbon, Portugal.

S. Koppen, M. van der Kolk, F.C.M. van Kempen, J. de Vreugd and M. Langelaar (2018). "Improving optical performance of multicomponent optomechanical systems through topology optimization". At: *4th DSPE Conference on Precision Mechatronics*, September 2018. Sint Michielsgestel, The Netherlands.

Supervision MSc Theses

Schaab, H., Sonneveld, D., Koppen, S. & Ostayen, R. van (2022). *Design of Windmill Hydrostatic Bearing using Topology Optimization.*

Kerkhove, C., Koppen, S. & Langlaar, M. (2022). *Approximate Geometric Non-Linear Analysis in Topology Optimization (continued).*

Hoevenaars, J., Koppen, S. & Langlaar, M. (2021). *Approximate Geometric Non-Linear Analysis in Topology Optimization*

Maas, W., Koppen, S., Galffy, A. (TU Vienna), Schitter, G. (TU Vienna) & Keulen, F. van (2021). *The Topology Optimization of a Compliant Variable-camber Morphing Wing.*

Ouden, F. van den, Koppen, S. & Langelaar, M. (2021). *Topology Optimization for Manufacturable Thermoelastic Metamaterials.*

Hoes, E., Koppen, S., Frecker, M. (Penn State University) & Langelaar, M. (2020). *Additive Manufacturing of Locally Self-supporting Compliant Mechanisms.*

Pappas, I., Koppen, S. & Langelaar, M. (2020). *Adaptive Structural Optimization Methods.*

Blokland, G., Koppen, S. & Vreugd, J. de, Langelaar, M. (2019). *Mechanical Metamaterials by Topology Optimization.*

Acknowledgements

This endeavor would not have been possible without the funding and in-kind contributions of NWO and industrial partners involved in the Stable and Adjustable Mechanisms for Optical Instruments and Implants (SAMOII) project. I am deeply indebted to Fred, Just and Matthijs for the initiative in proposing this project and placing trust and confidence in my abilities for fulfillment. I would like to express my deepest gratitude to Fred for inspiring me to study the field of optimization and mechanics and pursue a career in research. I could not have undertaken this journey without the *begeleiding* of Matthijs. I particularly would like to emphasize my pleasure of collaborating in the supervision of numerous students. I would like to extend my sincere thanks to the external dissertation committee members for acceptance of critical assessment of my dissertation and defense, and, as such, maintain academic standards and scientific principles.

I am thankful for all those stimulating the growth of my intellectual integrity throughout this journey. Special thanks to Dirk for teaching me the art of critical thinking and scientific integrity whilst remaining *pretty chilled.* Max; you have been my continuous inspiration, critical mentor, advisor and collaborator from the beginning. I thank you. These years would not have been fantastic without the inspiring *and* hysterical discussions I had with Arnoud both during good and bad times. It stands to reason for us that without our close collaboration I would not have fulfilled this journey. I had the pleasure of working with a vast variety of inspiring, colleagues, friends and students, especially within the SOM and MSD research groups. I would be remiss in not mentioning the particular impact of Sanne, Floris, Emiel, Reinier, Abdullah, Casper, Jasper and Ioannis.

I am eternally grateful for the moral support and encouragement of my family, friends–both old and new–and *partner*; I have not been easy. Special mention goes to my trusted friends Niek, Bojan and Mounir and the countless evenings with fantastic docking of minds. I would like to extend my sincere thanks to Souad and Roland; you make me feel welcomed, accepted, and appreciated. Afaf; words cannot express my gratitude towards you. Your love, endless patience, and sixth sense to seamless understanding of my mind are invaluable to me. You are my rock. I am privileged by unconditional love and care from my parents and brother. Thank you for continuously reminding me that this milestone is valuable, yet not the most important thing in life.