Innovative optimization-based numerical methods for modeling the non-linear behavior of steel structures

Chadi El Boustani

UNIVERSITÉ PARIS-EST

Ecole doctorale SIE

THÈSE

présentée pour l'obtention du diplôme de

DOCTEUR

de l'Université Paris-Est

Spécialité : Structures et Matériaux

par

Chadi EL BOUSTANI

intitulée

Innovative optimization-based numerical methods for modeling the non-linear behavior of steel structures

soutenue le 2 novembre 2020, devant le jury composé de :

Habibou MAITOURNAM	Président du jury	ENSTA Paris
Mohammed HJIAJ	Rapporteur	INSA Rennes
Nicolas MOËS	Rapporteur	École Centrale de Nantes
Mickaël ABBAS	Examinateur	EDF
Jeremy BLEYER	Examinateur	École des Ponts ParisTech
Xavier CESPEDES	Examinateur	Strains
Laurence DAVAINE	Exminatrice	Ingérop
Karam SAB	Directeur de thèse	École des Ponts ParisTech
Mathieu ARQUIER	Invité	Strains

Remerciements

J'ai pu réaliser ce travail de thèse dans d'excellentes conditions, très bien entouré, et ceci grâce à la contribution de nombreuses personnes que j'aimerais ici remercier.

Je souhaite tout d'abord remercier les membres du jury qui ont accepté d'évaluer ma thèse, en particulier Habibou Maitournam qui m'a fait l'honneur de présider ce jury et également Mohammed Hjiaj et Nicolas Moës qui ont accepté d'être rapporteurs. Je tiens aussi à remercier Mickaël Abbas et Laurence Davaine qui ont examiné avec attention ce travail et ont apporté leur vision plus pragmatique.

Un grand merci à mon directeur de thèse Karam Sab et à mon encadrant Jeremy Bleyer. Outre la confiance qu'ils m'ont accordée en me proposant ce sujet, ils m'ont transmis leur méthodologie, leur vision de la recherche et ont toujours été disponibles. Tout en me laissant une grande autonomie, ils ont su me guider sur les pistes prometteuses et capitaliser sur mes connaissances transverses afin d'aboutir à ce travail accompli.

Coté entreprise, je tiens à remercier Xavier Cepesdes qui a apporté sa vision à la fois pragmatique et scientifique et m'a toujours poussé à garder à l'esprit les applications au travail quotidien de l'ingénieur, et Mathieu Arquier qui a suivi de près la mise-en-œuvre numérique et m'a permis de formater son code à ma sauce. Au reste de mes collègues de Strains qui ont toujours été très enthousiastes à chaque nouveau développement et qui m'ont fait confiance pour l'utiliser dans leurs missions d'ingénieur. À Rafal et Thomas pour la bonne humeur qui régnait et aux innombrables « push », « pull », « blame » et toutes nos investigations pour dévoiler les coupables d'une ligne de code ratée ou tout autre crime.

Ensuite, j'aimerai remercier tous mes collègues de l'équipe MSA pour l'ambiance conviviale dont j'ai profité pendant ces trois années malgré le fait que je ne sois pas souvent au laboratoire. Merci à Filippo, Leyla, Charlotte, Florian, Baptiste, Eki et tous les autres, pour leur bonne humeur, l'entraide, les gâteaux, et surtout les bons apéros passés ensemble.

À tous mes amis dont le hasard a fait qu'on se retrouve souvent à Paris ou ailleurs, parfois loin de nos repères, mais qui m'ont accompagné dans tout les moments. Merci à Sophie, Fabienne, Charbel, Rawad, Hadi, Rita, Rita, Gregoire, Renée, Antoine, Khalil, Christian, Joelle, Emilio, Imad, Dory, Yang, Pilar, Arthur, Ignasi et tant d'autres, qu'il serait trop long de citer, qui n'ont jamais manqué une occasion de me retrouver pour un diner, un voyage, un apéro ou une randonnée. À Joe qui me supporte depuis d'innombrables années. À Léonore qui m'épaule constamment et m'a permis d'avancer au quotidien dans la joie et la sérénité malgré cette année très mouvementée. Tous les deux ont même bravé le confinement pour assister à ma soutenance.

Enfin, j'aimerais remercier du fond du cœur mes parents ainsi que mon frère et ma sœur qui, même habitant loin, m'ont toujours supporté et sur qui je peux compter à tout moment. Sans eux, je ne serais certainement pas arrivé jusqu'ici.

Abstract

This work is the product of a three-year collaboration between *Laboratoire Navier* and the civil engineering firm *Strains*. It has been shaped by following the needs of the construction industry for a more reliable, efficient, robust and easy-to-use structural analysis software. Departing from traditional numerical methods used in non-linear finite-element analyses, this work aims at adapting and extending the use of optimization-based algorithms to solve a wide range of non-linear mechanical problems.

Based on a solid convex-optimization mathematical theory, the primal-dual interior point method is nowadays becoming a reliable technology capable of handling various non-linear and non-smooth problems. Various mechanical behavior such as plasticity or contact conditions can be written using second-order cone complementarity problems which perfectly fits in the conic optimization framework.

Under the small-strain assumption, the elastoplastic contact problem can be cast as a pair of dual optimization problems. These problems can also be extended to the upper and lower bound theorems of yield design/limit analysis theory. Appropriate displacement and stress-based finite-element discretization strategies are chosen and the corresponding minimization problems are then solved using a state-of-the art primal-dual interior-point solver coded from scratch, yielding respectively an upper and a lower bound estimate of the exact solution. The proposed framework is illustrated and validated through various steel structure examples and the results are compared to other finite-element commercial software and Eurocode design recommendations. Its efficiency compared to a standard step-by-step Newton procedure, is proven via important savings in computational time, mainly due to its remarkable robustness with respect to large load steps.

The framework has also been extended to a non-convex setting involving finite-strain plasticity using a total Lagrangian formulation based on a logarithmic strain measure. The proposed extension of the interior-point algorithm is implemented and tested on 3D examples involving plastic collapse and geometrical changes. Comparison with classical Newton-Raphson/return mapping methods shows that the interior-point method still exhibits good computational performance, especially in terms of convergence robustness. Similarly to what is observed for convex small-strain plasticity, the interior-point method is able to converge for much larger load steps than classical methods.

Finally, the potentialities of the proposed framework is illustrated on various complex engineering problems taken from various design studies such as 3D steel assemblies or the second-order analysis of a steel bridge section.

Keywords : Conic programming ; Interior-point method ; Elastoplasticity ; Limit Analysis ; Non-linear finite-elements.

Résumé

Ce travail est l'aboutissement de trois ans de collaboration entre le *Laboratoire Navier* et la société *Strains*. Les axes de recherche traités ont été motivés par le besoin de l'industrie d'un renouvellement des méthodes actuelles en calcul non-linéaire des structures. Cette thèse tente de développer des outils numériques robustes, efficaces et faciles d'usage, basés sur des algorithmes d'optimisation en guise d'alternative aux méthodes non-linéaires plus traditionnelles.

Reposant sur les fondements de la théorie d'optimisation convexe, les méthodes de point-intérieur sont aujourd'hui la technique incontournable pour la résolution de divers problèmes d'optimisation, en particulier dans le cas non-différentiable. À travers le formalisme des conditions de complémentarité sur des cônes du second-ordre, plusieurs comportements mécaniques comme la plasticité ou les conditions de contact peuvent s'exprimer sous la forme de problèmes coniques pouvant être résolus efficacement grâce à la méthode de point intérieur.

Sous l'hypothèse des déformations infinitésimales, le problème élastoplastique avec contact peut être écrit dans sa forme faible comme une paire de problèmes d'optimisation duaux. Ces deux problèmes se rapprochent des théorèmes cinématique et statique du calcul à la rupture/analyse limite et fournissent des bornes supérieure et inférieure de la solution exacte. En utilisant des discrétisations par des éléments finis cinématique et statique convenables, les problèmes de minimisation sont résolus avec l'algorithme de point intérieur présenté et spécifiquement développé par l'entreprise. L'outil numérique est illustré à travers divers exemples de structures métalliques et les résultats sont comparés à ceux obtenus par des logiciels d'éléments finis commerciaux et aux recommandations de l'Eurocode. L'efficacité de l'algorithme est mise en avant grâce à une réduction significative des temps de calculs surtout due à sa remarquable robustesse vis-à-vis de grands pas de chargements, contrairement aux méthodes classiques.

L'approche est étendue au cas non-convexe de la plasticité en grandes transformations en utilisant une approche par Lagrangien total et une mesure logarithmique des déformations. L'algorithme de point intérieur est alors adapté et testé sur des exemples 3D illustrant les effets de changement de géométrie et les concentrations des déformations plastiques. Comparé à la méthode de Newton-Raphson avec un algorithme de retour radial, la méthode de point intérieur garde plusieurs de ses avantages numériques surtout en terme de robustesse. Comme observé pour les problèmes précédents, il est possible de converger vers une solution en utilisant des pas de chargement assez larges.

Finalement, les potentialités de ce nouvel outil numérique et ses avantages industriels sont illustrés dans divers exemples issus d'études d'ingénierie comme le calcul d'assemblages 3D complexes et l'étude au second-ordre d'un caisson de pont métallique.

Mots-clés : Programmation conique ; Méthode de point intérieur ; Elastoplasticité ; Analyse Limite ; Éléments finis non-linéaires.

Contents

Chapter 1

Introduction

Abstract: *This Ph.D. has been carried out in collaboration with* Strains*, a company of civil engineers aiming to develop modern numerical tools for engineering practice. In this first chapter, we set the industrial context of the work presented in this manuscript, along with the scientific context which is also presented with a main focus on the current numerical methods. The aim of this thesis and the manuscript layout are also included.*

Contents

1.1 General scope

1.1.1 Historical, economic and industrial context

Structural engineers have historically been concerned by two main issues: the mechanical expertise of existing structures and the design of new constructions. After the Second World War, the need to reconstruct entire cities and infrastructures pushed engineers to explore new horizons. Encouraged by scientific advances, economical growth, the low cost of materials and the absence of environmental constraints, the construction industry thrived. Today, almost two-thirds of buildings are post 1948, reaching their predicted decommissioning unless refurbished.

Builders and designers were seeking reliability and efficiency for their new complex structures without saving materials or optimizing the structural behavior. This period of growth was long enough to strongly influence the structural engineering methods. It has led to the development of fast computational methods yielding buildings with safety margins that were not fully assessed. Building codes were adapted to include simple design rules and software engineers have designed time-efficient software able to provide quick studies for complex structures in the framework of elasticity. The need for « one-click » computer software has quickly become a clear evidence.

Today, the construction industry is at the verge of a new horizon. Structural paradigms are shifting, but also economic and, most importantly, environmental ones. On the one hand, new projects are becoming even more complex, coupling 3D geometrical effects and new innovative materials, while budgets are shrinking whereas security norms are becoming increasingly restrictive. In this respect, recent disasters had a profound impact on safety considerations e.g. the failure of Terminal 2E at Charles de Gaulle airport in 2004, the Fukushima nuclear incident in 2011, the Morandi bridge collapse in 2018, the "Pont de Mirepoix" in 2019 and much more. A ongoing review of all bridges in France showed that many of them, including late-80s iconic bridges, face increasing failure risks and must urgently be repaired.

On the other hand, design methods lag behind and often do not provide satisfying results within a reasonable amount of time and effort. Many easy-to-use « profession » software are enough to handle the majority of problems. However, general finite element software, such as Abaqus or ANSYS, powerful enough to handle all non-linearities and complexities, require great expertise to correctly model the structure, let alone to pilot convergence and obtain quality results. The need for powerful computer software coupling state-of-the-art numerical methods, performance, efficiency and ease-of-use is today greater than ever.

1.1.2 Structural analysis for steel structures

Steel is a material that offers several advantages: it is structurally sound, has a high strength-to-weight ratio and is quite sustainable. These characteristics make it ideal for constructing buildings and bridges of all sizes.

1.1. General scope

In France, a country well known for its concrete affinity, almost 48.5% [1] of its non-residential buildings are made of a steel-based structure. Their safety is ensured via the application of strict building codes such as the Eurocode, especially its third part [EN1993–1-1, 2005] which gives global recommendations and design methods.

The Eurocode ensures the structural integrity of a project throughout a set of local and global checks. In general, the various structural design requirements should comply with the notion of *limit states design*. More precisely, the structure must satisfy two different criteria: the *ultimate limit state* (ULS), i.e. the structure shall not fail, or the *serviceability limit state* (SLS), for instance its deflections shall not exceed prescribed limits.

These requirements can be simplified using a deterministic manner as presented in Fig. 1.1 where for a given geometry (a_d), action set (F_d) and material choice (X_d), the overall effects (E_d) should be less than an estimated resistance (R_d). In this manuscript, we focus on the methods used to evaluate the effects and the resistance of steel structures with a fixed geometry and material parameters and under a specific load case.

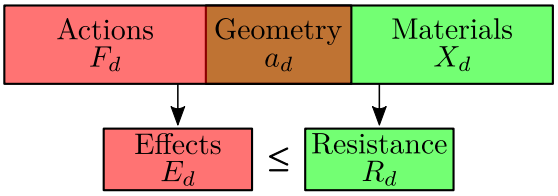


Figure 1.1 – Basic deterministic Eurocode check

For a ULS check, the engineer's task can therefore be divided into two parts:

1. the evaluation of the resistance, and
2. the calculation of the effects or the structural analysis method.

However, these two tasks are strongly related, since a more complex and complete structural analysis method greatly simplifies the remaining ULS check and vice versa as can be seen in Fig. 1.2. For instance, including all material and geometrical non-linearities in the structural analysis phase, thus implicitly accounting for the ultimate strength of each member, renders the final check rather straightforward since the only thing that remains is to check that the structure can withstand the current load with a specific safety factor. In contrast, if a simple linear elastic structural analysis is made, the ultimate resistance must then be explicitly calculated during the ULS check using various methods presented by the Eurocode.

In both cases, the resistance of the structure relies on each of its individual member and also on the steel assemblies connecting all the elements. Structural steel members are required to transmit axial and transverse forces, moments and torques and its resistance is determined both by its material strength and geometry. The Eurocode

1. Source: SCMF via Schilling Communication

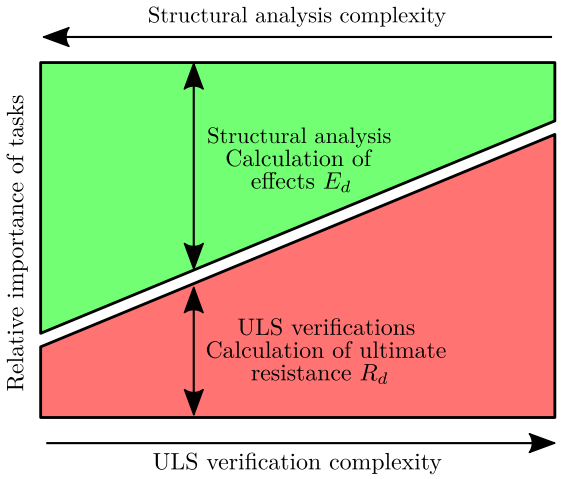


Figure 1.2 – Philosophy of a Eurocode design scheme

provides various approaches to estimate the resistance of each member depending on the slenderness of the structure and the interaction between the different members. A large set of coefficients, which take into account the local buckling of plates or yielding of bolts, etc., greatly complicate this step.

In theory, a design is considered economical when the utility ratio of all its member is maximized, i.e. for a given ultimate load we have full yielding of the members. Therefore, the main method used by the Eurocode to estimate the ultimate resistance is yield design. In simple cases such as unidirectional forces, flexural checks or simple assembly typologies, these calculations can easily be made by hand. But eventually, they become extremely long and even impossible in the case of complex assemblies, leaving the engineer the choice to either overestimate the required sections and thus leading to an uneconomical design, or to rethink his analysis procedure. Note that in some cases, such as class 4 slender sections, the Eurocode will even require the member to stay in its elastic domain, with a possible reduction of the resisting section, in order to calculate the resistance. It is in this specific cases where a complete structural analysis will be most interesting.

In this thesis, we will try to overcome some of these problems by either using yield design theory to estimate the ultimate resistance or by improving the structural analysis.

As regards structural analysis, different methods might be used, each relying on a specific set of hypotheses. A schematic diagram comparing these analyses is given in Fig. 1.3.

For many structures, it is common to use *first-order elastic analysis* which is based on linear elastic constitutive relationships and which ignores any geometrical non-linearities, instabilities or plasticity. This method is the simplest and the fastest method

found in various software. The deformations determined by such an analysis are proportional to the applied loads, and so the principle of superposition can be used to simplify the analysis. However, a first-order elastic analysis will overestimate the forces and underestimate the deformations of a structure. Although comparing elastic-obtained forces to plastic-obtained resistances will generally yield uneconomical design, this method remains the most used one among engineers.

The importance of instabilities on the structural behavior can be accounted for through an *elastic stability analysis* where the global buckling modes and related buckling coefficients can be calculated. On the contrary, a *second-order elastic analysis* accounts for both elastic behavior and geometrical non-linearities. Although generic computer programs can run this type of analysis, they may be quite difficult to carry out in engineering practice. Note, that the Eurocode allows the use of elastic stability results to amplify first-order moments and use a first-order like design rules as an alternative to second-order analysis.

Conversely, some structures may not be sensitive to instabilities, in which case only material non-linearities (e.g. plasticity) should be accounted for to determine the failure mechanisms. This *first-order plastic analysis* is also available in various software. However, common practice shows again that these methods are not always easy to converge and a solid knowledge of the underlying theory is required in order to correctly interpret the results and provide quality designs.

Finally, *second-order plastic analysis* in which the actual behavior is closely analyzed by allowing for instability, yielding, residual stresses, and initial crookedness is the most complex type of analysis. This method is rarely used as it requires experienced engineers and complex finite-element implementations. Its use is limited to specific projects and is extremely time-consuming and again difficult to converge.

These different analyses can be realized at different levels of the structure by having a global model and/or a set of different local models. Typically, a quick and effective design scheme relies on reducing the 3D structures to a set of simple sub-structures interacting with each other and with a specific resistance goal as can be seen in Fig. 1.4. It is therefore common practice to subdivide the structure into its small parts, to chose a structural analysis method and to apply that to each and every model. Obviously, the connection between the different parts is crucial. The set of hypotheses taken for each connection should be coherent with their design, which brings us to the most problematic part of a steel structure: the steel assemblies.

Steel assemblies connect different parts of a structure using various typologies designed specifically to transmit (or not) a given type of force. In simple industrial buildings, connections can be modeled using simplified hypotheses such as "nominally-pinned" or "fully-rigid" connections between the elements. However, engineers are facing today new 3D connection typologies where a large number of members converge to the same point in different angles and sizes (see Fig. 1.5).

Moreover, the actual behavior of a steel connection is neither fully-rigid nor nominally-pinned. In reality, all steel assemblies are "semi-rigid", and their stiffness

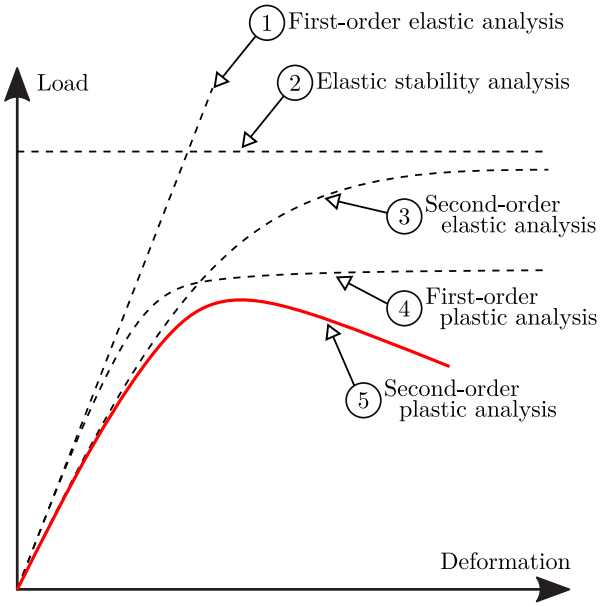


Figure 1.3 – Different structure behaviors (taken from Trahair et al. [2007])

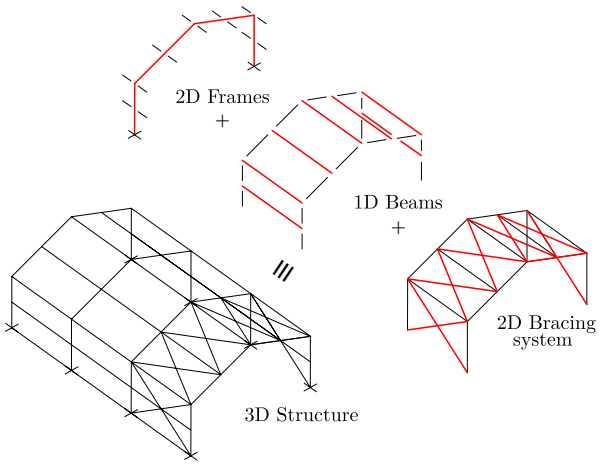


Figure 1.4 – Reduction of a 3D structure to simpler forms (taken from Trahair et al. [2007])

and ultimate resistance varies due to the contact between the steel plates, bolts, welds and the possible yielding or buckling of some parts under the different load cases.

As it will be seen in this manuscript, the Eurocode provides some methods to evaluate the mechanical characteristics of steel assemblies such as the sub-components method. However, these calculations are often extremely time-consuming and often rely on simplified models with large safety coefficients therefore providing uneconomical designs.

The design of steel assemblies is to be considered one of the biggest challenges of steel structure engineering. Few software can efficiently handle such intrinsically three-dimensional models using a first-order elastic analysis, let alone a full second-order

Figure 1.5 – A complex 3D steel structure - *Fondation Louis Vuitton*

plastic analysis coupling contact interfaces with material non-linearities and geometrical non-linearities. This problem is at the heart of this thesis.

1.1.3 Strains

Strains is a civil engineering design firm which was created in 2014 by a handful of structural engineers in order to contribute to answering the previously mentioned issues. The company aims at changing the current paradigm of structural analysis by developing innovative software dedicated to the structural analysis of existing and future buildings and infrastructures. In particular, the company tries to directly apply these new design methods throughout its various engineering missions.

A big part of the development is the numerical platform *Digital-Structure* dedicated to the 3D finite-element analysis of complex details. This platform groups different modules, each dedicated to a specific type of structures: *DS-Steel* for steel structures, *DS-Concrete* for reinforced concrete structures and *DS-Masonry* for masonry structures. However, all of these modules rely on the same innovative finite-element framework of which certain aspects will be explored in this work.

While this manuscript focuses mainly on the new tools dedicated specifically to steel structures, the developments made throughout this collaboration with *Strains*, mainly on the numerical methods and dual approaches, have been extended to all the different modules.

All the developments at *Strains* are based on modern computing techniques and embed cutting-edge numerical components. Programs are designed on the basis of a Software As A Service (SAAS) client-server model, meaning that the computations are processed on remote powerful servers. The client can carry out computations from any numerical device with a web connection without any software installation. Results are stored on a cloud and can be accessed online and/or downloaded locally by the client.

1.2 Aim of this work

Having provided an overview of structural analysis in civil engineering, it appears clear that performing a non-linear finite-element analysis involving plasticity, contact or geometrical changes on large-scale 3D structures is a daunting taks for engineers. If for simple projects such a complex analysis may not be required, these difficulties are clear obstacles to safe and economical designs for more complex projects. We can identify various blocking points which must be addressed in order to make non-linear analysis used more often in engineering practice:

- Discretization issues: choosing a reliable finite-element discretization, free from numerical locking problems, and providing reliable estimates of quantities of interest for the engineer is not a trivial task. As regards the latter point, accurate stress evaluations are usually the most important goal of a structural analysis in engineering practice. Unfortunately, standard displacement-type finite element discretizations do not necessarily provide safe approximations of stresses.

- Numerical resolution issues: standard non-linear procedures may often exhibit non-convergence for difficult problems, especially when the underlying material behavior is non-smooth. Such lack of robustness may be extremely detrimental for an engineer which has to assess the safety of a structure.

- "Experienced" user-parameters: some techniques alleviate the issue of convergence robustness by tweaking some mechanical (e.g. penalty stiffness for contact) or purely algorithmic user-parameters. Usually, such parameters have to be found by trial-and-error for a specific problem which is again undesirable in the context of safety assessment.

- Modeling and computing time: time schedules for civil engineering studies are usually extremely tight. The geometric modeling process must therefore be simplified and computation times must be reasonable.

This work aims at handling these various aspects by providing engineers with a simple to use, yet efficient numerical method to calculate the behavior of steel structures while relying on the latest state-of-the art numerical methods capable of handling large-scale problems with various non-linearities. These methods can either be used to estimate the ultimate resistance of steel members or assemblies or as a structural analysis method in order to simplify final ULS checks as explained early on.

We now give a brief review of existing numerical methods for contact, elastoplastic and yield analysis before positioning our work with respect to optimization-based approaches. We will also briefly mention discretization techniques involving equilibrium finite elements.

1.2.1 Numerical methods for contact problems

Contact problems are often solved using the penalty approach in which penetration between the two bodies, i.e. negative gap and tensile surface tractions are permitted, but penalized using a gap-element characterized by a large stiffness value [Kikuchi and Oden, 1988]. This method is easy to implement and transforms the problem into a regularized unconstrained minimization problem, the drawback being that the system becomes ill-conditioned for increasingly larger penalty factors. Besides, contact constraints violation generally produces poor results, underestimating the system's stresses, and thus preventing engineers from performing a safe and optimized design. Another method consists in enforcing the contact conditions using Lagrange multipliers thus transforming the problem into a saddle point problem with simple inequality constraints [Kikuchi and Oden, 1988]. Although contact conditions will be strictly verified, this method proved slow convergence rates. Augmented Lagrangian (AL) approaches emerged as an alternative [Hestenes, 1969; Powell, 1969]: it couples the Lagrange multipliers approach with the penalty approach while limiting the need to drive the stiffness coefficient to infinity, thereby providing better results [Fortin and Glowinski, 2000; Kikuchi and Oden, 1988; Wriggers and Zavarise, 1993].

Strategies for solving the corresponding non-linear problems are numerous, including fixed-point scheme or Uzawa algorithm [Laursen and Simo, 1993; Simo and Laursen, 1992], generalized/semi-smooth Newton methods [Alart, 1997; Alart and Curnier, 1991], primal-dual active sets [Hüeber and Wohlmuth, 2005], etc. Some of these methods have been shown to be equivalent in some specific cases [Hintermüller et al., 2002; Popp and Wall, 2014]. Frictional contact can be written as second-order cone complementarity conditions [Kanno, 2011; Kanno et al., 2006; Zhang et al., 2011] but the non-associated character of the Coulomb friction law makes the problem even more challenging, the problem becoming a non-monotone complementarity problem. Linear complementarity solvers have been proposed when adopting a facetized representation of the Coulomb cone [Klarbring, 1986; Stewart and Trinkle, 1996]. Bi-potential approaches have also been studied along with dedicated numerical procedures [Hjiaj et al., 2004; Joli and Feng, 2008].

In the above, we focussed on contact resolution strategies but contact detection and discretization issues are also particularly important. We will however not review these different approaches here since our work will only consider the most simple setting of node-to-node contact in a small strain setting.

1.2.2 Numerical methods for elastoplastic analysis

The analysis of stress-strain states or the stability of components or structures subjected to various complex loading and contact conditions beyond limits of elasticity requires a plasticity theory which can describe complex elastoplastic processes of deformation. Finding solution to those elastoplastic boundary value problems has always been one of the main concerns of the engineering community.

From a general point of view, the elastoplastic boundary value problem can be

divided into two parts: an equilibrium check and a local constitutive model check. For these two checks, one of the most common methods consists in using a Newton-type procedure that alternates between iterative global equilibrium corrections and local constitutive state updates at specific points, generally chosen as Gauss points. One may typically proceed by determining a displacement increment based on the current state of the system, then a stress increment is calculated at the specific chosen points by solving the local constitutive relations, and thus determining an updated estimate of the tangent elastoplastic stiffness modulus and values of the new state variables (e.g. cumulated plastic strain). The change of stiffness yields a residual expressing a global out-of-balance force which is driven to zero using the Newton-Raphson algorithm.

State-of-the art algorithms mainly differ in how the local stress update is performed and how the resulting stiffness is computed. The method used to integrate local constitutive equations composed by a set of non-linear differential equations can be classified as either implicit or explicit. Implicit methods which rely on a backward Euler scheme are the most commonly used, the most important example being the return mapping algorithm. An extensive study of these algorithms can be found in Simo and Hughes [2006]. Explicit methods such as the sub-stepping schemes developed by Sloan and his co-workers [Sloan, 1987] are less common.

Another method for integrating those equations relies Hills maximum dissipation theorem thus formulating the local constitutive equations as mathematical programs than can be solved using optimization algorithms [Maier, 1970; Martin et al., 1987].

A extensive presentation of computational methods for elastoplasticity can be found in the seminal work of de Souza Neto et al. [2008].

1.2.3 Numerical methods for yield analysis and extensions

A fundamentally different approach will be explored in this manuscript which is inspired by the numerical methods used for limit analysis/yield design and the development of a dedicated software at *Strains* [Vincent, 2018; Vincent et al., 2018].

Limit analysis theory [Hill, 1950], or more generally yield design theory [Salençon, 1983; Salençon, 2013], has been widely used as the basis of many design methods in civil engineering e.g. slope stability factors and footing bearing capacities [Chen, 2013], strut-and-tie models [Schlaich et al., 1987], yield line methods for slabs and metal plates [Johansen, 1962; Save, 1995; Save et al., 1997], etc.

Due to the degenerated nature of the limit analysis problem, it cannot be solved using traditional Newton-Raphson techniques. Recent works proposed to use a fictitious loading path using a specific arc-length procedure to reach convergence to the ultimate load using a standard Newton approach [Magisano and Garcea, 2020]. However, the current state-of-the-art approach for solving limit analysis problem is through the use of mathematical programming tools, in particular interior-point methods [Andersen et al., 2003; Boyd and Vandenberghe, 2004; Lobo et al., 1998; Wright, 2005]. With the advent of such efficient conic programming solvers and dedicated finite-element solution strategies, limit analysis problems can now be solved numerically at a moderate cost [Makrodimopoulos and Martin, 2006, 2007; Martin and Makrodimopoulos, 2008;

Vincent et al., 2018]. The robustness of such solvers in the context of estimating a structure ultimate load is a decisive advantage over traditional Newton-Raphson based techniques which generally fail to compute an acceptable solution within a reasonable amount of time [Lyamin and Sloan, 2002a,b; Makrodimopoulos and Martin, 2006, 2007].

Some contributions also explored the use of this framework for other applications such as granular media [Krabbenhoft et al., 2012b; Zhang et al., 2014], contact [Kanno, 2011; Kanno et al., 2006], viscoplastic flows [Bleyer, 2017; Bleyer et al., 2015] or elastoplastic computations [Kanno, 2011; Krabbenhøft et al., 2007; Krabbenhoft et al., 2007; Yonekura and Kanno, 2012]. The latter case is of interest since limit analysis does not guarantee a unique solution as regards to failure mechanisms, nor does it enable to compute the structure's displacements and total strains. Including elastic strains in a limit analysis model i.e. aiming at computing the ultimate load with an elastoplastic computation circumvents the previously mentioned limitations. Beyond assessing the structure's load-bearing capacity, it also enables to check for ultimate strain limits in yielded regions. The approaches bridging the numerical methods of limit analysis with contact or elastoplastic constitutive models seem particularly relevant for our objectives. It consists of casting the complete problem as a non-linear mathematical program which is subsequently solved using general optimization methods.

This approach was initiated by Maier [1968, 1969] and Capurso and Maier [1970] where the incremental problem is reduced to a convex quadratic problem. Since then, various optimization-based algorithms have been developed, each one adapted to a specific set of material behaviors, and they include: the parametric quadratic programming method [Maier, 1968], a general interior-point method [Krabbenhoft et al., 2007], methods based on the linear complementarity problem to couple frictional contact with elastoplastic media [Zhu, 1995] and the bi-potential method which deals with non-associative behavior [de Saxcé and Feng, 1998; Hjiaj et al., 2003].

Recently, many contributions were made to formulate the classical complementarity conditions of associated plastic flow, also known as the Karun-Kush-Tucker (KKT) optimality conditions as a set of self dual convex cones [Krabbenhøft et al., 2007]. It has been shown that the Mohr-Coulomb yield criterion, which is conventionally used in the limit analysis of geomaterials, can be represented as a linear matrix inequality that can be reduced to a semi-definite program (SDP) [Martin and Makrodimopoulos, 2008] and that the von Mises yield criterion with or without strain hardening, which is conventionally used in the analysis of metals, can be represented as a second order cone condition over a pair of dual kinematic and static variables [Kanno, 2011; Yonekura and Kanno, 2012]. With that, the incremental elastoplastic problem for the quasistatic analysis can be formulated as a pair of dual minimization problems under second-order cone constraints or more commonly know as second-order cone programs (SOCP). Although these algorithms present a guaranteed convergence for convex problems and a polynomial time complexity [Alizadeh and Goldfarb, 2003], few attempts were made to generalize their use and assess their performance in the case of large-scale 3D elastoplastic problems.

1.2.4 Displacement and equilibrium finite elements

Standard approaches of computational limit analysis make use of traditional displacement-based finite elements for implementing the upper bound kinematic approach but also of less standard equilibrium (or stress-based) discretizations for implementing the lower bound static approach. The latter choice is especially interesting for engineering application since it produces a safe (lower bound) estimate of the structure ultimate load, contrary to conventional approaches. Let us note that fully discontinuous displacement interpolations can also be used in the upper-bound kinematic approach (see Bleyer and Hassen [2020] for a presentation of various discretizations in limit analysis).

Initially developed after the works of Fraeijs de Veubeke [1963, 1965], these elements were used to bracket the energy of the exact solution, however their use was not very common due to some drawbacks and restrictions.

Alternatively, hybrid formulation were presented by de Almeida [de Almeida and de Freitas, 1991; de Almeida and Pereira, 1996] and by de Freitas [de Freitas and Ji, 1996] following the works of Pian [Pian, 1964]. These elements were called hybrid because of the discretization of the displacement field on their sides. Upon proper choice of the discretization spaces, such hybrid approaches can lead to completely statically admissible solutions. However, the corresponding discrete displacement fields are not kinematically admissible in such approaches. Recently, a new family of 3D tetrahedral elements based on the model developed in Fraeijs de Veubeke [1965] was introduced in Kempeneers et al. [2009]. We also refer to a recent work published by de Almeida and Maunder [2017] for an extensive presentation of equilibrium elements.

Building upon the common usage in limit analysis, our work will therefore investigate the simultaneous use of displacement and equilibrium finite-element discretizations in order to bracket the results of our computations. This choice will be referred to as a *dual approach* which we will exploit to devise a mesh adaptation scheme.

1.3 Contributions and manuscript organization

Inspired by the existing theory in yield design, in this work, we aim at providing a complete large-scale 3D dual approach based on the IPM which will be adapted in order to handle a large variety of problems. Technically, a general finite-element software should be able to handle elasticity, material non-linearities, boundary non-linearities and geometrical non-linearities. Relying on an optimization software instead of more standard non-linear equation solvers requires many changes compared to a more standard finite-element software. This includes in particular the possibility to include both displacement and equilibrium-based finite element discretizations.

The scientific contribution is not only limited to the algorithmic and implementation part or the mechanical reformulation of classical problems but it also includes various

improvements in equilibrium formulations, in dual error estimators and remesh scheme, along with an extension to logarithmic elastoplasticity and also to various mechanical behaviors not included in this work but already found in *Strains*'s software.

1.3.1 Manuscript organization

Chapter 2 gives an overview of non-linear optimization which is the starting point of all our developments. Since a primal-dual interior point algorithm will be used in all subsequent computations, it is presented in its most general form. Although the IPM is normally used to solve convex optimization problems, our presentation includes the case of non-convex problems with non-smooth constraints. The algorithm is presented in detail along with the advanced implementation aspects such as the symmetrical scaling and the predictor-corrector scheme.

The subsequent chapters constitute the different technical developments made in *Strains*' finite-element software:

Chapter 3 handles the adaptation of the dual approach to elastic problems with contact conditions. First, it includes the reformulation of associated frictional contact conditions as a pair of complementary dual cones to be used within the framework of second-order cone programming. Then, the minimization problems and error estimator including a surface contact term are reformulated relying on displacement and stress-based variational bounding principles. The error estimator will be used in a global 3D remeshing scheme based on the dual approach. Finite elements discretization for each of the two approaches based on classic kinematic elements and on improved equilibrium elements are introduced and illustrated by some basic and engineering examples.

Chapter 4 provides a general framework for efficiently computing the ultimate state of complex steel assemblies including plasticity. Following on the ideas introduced in chapter 3, we will resort to a dual finite-element discretization involving kinematic displacement-based elements as well as static equilibrium-based elements, providing respectively an upper and a lower bounds estimates to the exact solution and, in particular, to the exact structure ultimate load. Both computations can be also be used to compute a constitutive error indicator which can be used in a remesh algorithm. Therefore the incremental variational formulations of elastoplastic boundary value problems and their extension to limit analysis/yield design are presented along with the particularities of the finite-element discretizations used in both approaches. The whole process is illustrated using examples in order to assess its performance and accuracy with respect to classical elastoplastic solution strategies and to validate the method against different Eurocode design norms.

Chapter 5 extends the concepts to include geometrical non-linearities. Various civil engineering applications require to take into account large-strain effects. In this chapter, we apply the proposed framework to the specific case of logarithmic strain elastoplasticity which is well adapted to the study of steel structures. As it will be discussed later, the use of the logarithmic strain setting enables a simple extension of classical small-strain elastoplastic constitutive laws to the finite-strain setting. In particular, the additive decomposition between elastic and plastic strain is preserved and the elastic

energy densities and plastic dissipation potentials are still convex with respect to the corresponding strains. Non-convexity only arises due to the non-linear relation between displacement and total strain. Using such a framework, we indeed obtain a problem in which non-smoothness (due to the plastic dissipation) can still be expressed using convex constraints whereas non-convexity involves smooth terms (strain/displacement relation). A key assumption in our method relies on the fact that the additional non-convex terms are smooth, inducing additional contributions to the KKT system tangent matrix and residuals without changing the convex conic structure. Benchmark 3D examples will be used to validate our implementation and a comparison is made against standard Newton-Raphson methods with a return mapping inner procedure enabling us to assess the computational cost and convergence robustness of the IPM solver.

Chapter 6 exposes various engineering application that were made possible using these numerical developments and clearly showing the direct industrial reach of this thesis.

We finally give some conclusions and perspectives for future works.

1.3.2 Publications associated with the present work

The present work resulted in three peer-reviewed papers in international journals:

El Boustani, C., Bleyer, J., Arquier, M., Ferradi, M.-K., and Sab, K. (2019). Dual finite-element analysis using second-order cone programming for structures including contact. *Engineering Structures*, 10.1016/j.engstruct.2019.109892 [2]

El Boustani, C., Bleyer, J., Arquier, M., Ferradi, M.-K., and Sab, K. (2020). Elastoplastic and limit analysis of 3D steel assemblies using second-order cone programming and dual finite-elements. *Engineering Structures*, 10.1016/j.engstruct.2020.111041 [3]

El Boustani, C., Bleyer, J., Arquier, M., and Sab, K. . Extending interior-point methods to non-linear second-order cone programming: application to finite-strain elastoplasticity. *International Journal for Numerical Methods in Engineering*, 10.1002/nme.6537 [4]

one book chapter:

El Boustani, C., Bleyer, J., and Sab, K. (2020). Limit Analysis of Complex 3D Steel Structures Using Second-Order Cone Programming. *Direct Methods: Methodological Progress and Engineering Applications*, Chapter 3, 10.1007/978-3-030-48834-5_3 [5]

and two contributions in international or national conferences:

2. accessible at https://hal-enpc.archives-ouvertes.fr/hal-02428540
3. accessible at https://hal-enpc.archives-ouvertes.fr/hal-02884021
4. accessible at https://hal-enpc.archives-ouvertes.fr/hal-02930390
5. accessible at https://hal-enpc.archives-ouvertes.fr/hal-02906864

El Boustani, C., Bleyer, J., Arquier, M., Ferradi, M.-K., and Sab, K. (2019). An efficient and reliable steel assembly modeling scheme using second-order cone programming and dual error estimator. *Proceedings of the 9th International Conference on Steel and Aluminium Structures ICSAS19*, ISBN: 978-1-78972-197-3 [6]

El Boustani, C., Bleyer, J., Arquier, M., Ferradi, M.-K., and Sab, K. (2019). Analyse des structures par optimization convexe. Application aux charpentes métalliques. $14^{ème}$ *Colloque National en Calcul des Structures (CSMA2019)*

As a result, this report is made of parts entirely taken from these different articles and proceedings.

* *
*

6. accessible at https://hal-enpc.archives-ouvertes.fr/hal-02289785

Chapter 2

Non-linear optimization and the primal-dual interior point method

Abstract: *Throughout this manuscript, various mechanical problems will be introduced using the framework of minimization problems under a certain number of smooth and non-smooth constraints such as conic complementarity conditions. The aim of this chapter is to review these mathematical programming concepts along with the primal-dual interior point method which is particularly adapted to solving these type of minimization problems.*

Contents

2.1 Introduction

2.1.1 The need for a in-house optimization solver

The aim of this thesis, as explained in chapter 1 Sec. 1.2, is to provide engineers an innovative approach for solving non-linear mechanical problems using convex optimization tools. Upon reviewing the available software, there exists a large number of state-of-the-art ready-to-use optimization algorithms which can handle a wide variety of problems. Software like `Mosek`, `SeDuMi`, `CVXOPT`, `CPLEX`, `Knitro` and much more provide engineers and decision makers many efficient tools which rely on strong mathematical concepts and convergence proofs.

However, in this work, we decided to implement our own optimization software based on the same algorithm used in the most efficient software, namely the primal-dual interior point method (IPM). As it will be seen in this chapter, this main contribution of the thesis has been motivated by several drawbacks which are inherited from using black-box solvers relying on a very specific format which does not take into account the specific structure of the mechanical problems we will encounter. More importantly, experience has also showed that open-source alternatives, in particular first-order solvers like `SCS` [1], `COSMO` [2] or `CDCS` [3], often presented in the literature as more efficient and more scalable solvers than IPM do not perform well on our types of problems, most likely because their specific structure is not taken into account by such general purpose solvers. These disadvantages, along with *Strain*'s will to exploit at most the mechanical insight in order to provide quality results, has lead us to undertake the laborsome task of coding a new algorithm, well adapted to the needs of the industry.

Throughout this chapter, we will give a brief introduction on various mathematical programming concepts along with possible solving strategies and existing algorithms. General mathematical definitions used in this section can be found in appendix A. We will also focus on the interior-point method, its main ideas and the various implementation details required to produce an efficient solver. More importantly, contrary to state-of-the-art solvers on which we based our implementation and which are dedicated to a convex setting, we extend the ideas of the primal-dual IPM to a more general setting involving non-convex objectives/constraints.

2.1.2 A simple mechanical example

In order to better understand the meaning of each concept used in the algorithm through mechanical insight, we will illustrate its key ingredients on a simple example inspired from the mechanical problem that will be introduced later in chapter 4.

Let Ω be an elastoplastic body with Young modulus E, Poisson ratio ν and obeying a von Mises perfectly plastic criterion with an elastic limit σ_0. Let $\boldsymbol{\varepsilon}$ be the total small-

1. https://github.com/cvxgrp/scs
2. https://github.com/oxfordcontrol/COSMO.jl
3. https://github.com/oxfordcontrol/CDCS

strain tensor and $\boldsymbol{\varepsilon}^p$ be the plastic strain tensor. The body is subjected to an imposed deformation state: $\varepsilon_{11} = \varepsilon_0$, $\varepsilon_{22} = 0$ and ε_{33} is free. This corresponds to a uniaxial strain state of a body in plane-stress conditions in the third direction. At a specific material point of Ω, we now consider the following minimization problem:

$$\min_{\boldsymbol{\varepsilon},\, \boldsymbol{\varepsilon}^p} \quad \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p)\boldsymbol{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \sqrt{\frac{2}{3}}\sigma_0\|\boldsymbol{\varepsilon}^p\| \tag{2.1a}$$

$$\text{s.t.} \quad \varepsilon_{11} = \varepsilon_0 \tag{2.1b}$$

$$\varepsilon_{22} = 0 \tag{2.1c}$$

$$\operatorname{tr}(\boldsymbol{\varepsilon}^p) = 0 \tag{2.1d}$$

where $\boldsymbol{D}$ is the elasticity tensor and $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p$ are the state variables, or in mathematical terms, the *primal variables* represented as the vector of its 3 principal values $\boldsymbol{\varepsilon} = \{\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}\}$. Eq. (2.1a) is the objective function which contains a *smooth* quadratic form, the strain energy density $\frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p)\boldsymbol{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p)$, and a *non-smooth* term, the plastic dissipation $\frac{2}{3}\sigma_0\|\boldsymbol{\varepsilon}^p\|$, composed of the Frœbenius norm of the plastic strain $\|\boldsymbol{\varepsilon}^p\| = \sqrt{\varepsilon^p_{ij}\varepsilon^p_{ij}}$. Equality constraints given by equations (2.1b), (2.1c) and (2.1d) account for the imposed strain state and plastic incompressibility condition. Let us mention that this minimum problem can be seen as a simplified version of the incremental potential energy minimum principle in an elastoplastic setting and which will be detailed and presented more rigorously in chapter 4. Despite its simplicity, this problem is a representative example of the problems we will encounter later and will serve as a mechanical illustration of some notions on convex and conic optimization that we will now review.

2.2 Non-linear optimization and non-smooth optimization

A general way of formulating an optimization problem over a specific set $\mathcal{X} \subseteq \mathbb{R}^n$ is to consider non-linear (equality or inequality) constraints and a non-linear objective function, yielding so-called *non-linear programs* (NLP) which can be written as follows:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) = \boldsymbol{0} \quad \forall i \in \{1, \dots, m\} \\ & h_j(\boldsymbol{x}) \leq \boldsymbol{0} \quad \forall j \in \{1, \dots, p\} \\ & \boldsymbol{x} \in \mathcal{X} \end{array} \tag{2.2}$$

where f, g_i, and h_j be real-valued functions on $\mathcal{X}$ with at least one being non-linear.

Various methods exist for solving generic optimization problems. They can however be categorized depending on whether the problem is convex or non-convex and whether objective or constraints are smooth or non-smooth.

2.3 Convex and conic optimization

Convex optimization is one of the most important of the mathematical problems classes which finds applications in a wide range of fields from image processing, finance and economics, data analysis and numerous physical principles based on minimization of a convex potential.

Being able to formulate a problem as a convex optimization one offers great theoretical and conceptual advantages such as duality, where the associated dual problem often has a significant interpretation, and uniqueness of the solution (strictly convex case) or at least the fact that local minimizers are also global minimizers.

A *convex optimization* problem consists in minimizing a convex function over a convex set. More explicitly, a convex problem is of the form:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \mathcal{C} \end{aligned} \tag{2.3}$$

where $\mathcal{C}$ is a convex set and f is a convex function over $\mathcal{C}$. Note that any convex optimization problem can be reformulated into an equivalent problem involving only a linear objective function. Indeed, one can equivalently replace the objective $f(x)$ in (2.3) using its epigraph (see Def. A.9) i.e. by introducing a variable t and the convex constraint $f(\boldsymbol{x}) \leq t$ which will be attained at the optimum.

Problem (2.1) is indeed a convex problem due to the convexity of the quadratic strain energy ($\mathbb{D}$ being positive) and of the plastic dissipation potential.

It turns out that any convex optimization problem can be formulated as a *conic optimization* problem [Boyd and Vandenberghe, 2004; Glineur, 2001] which consists in minimizing a convex function [4] over a conic section such as:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \in \mathcal{K} \end{aligned} \tag{2.4}$$

where $\mathcal{K}$ is a convex cone and f is a convex function over $\mathcal{K}$. Quite often, it is assumed that the cone $\mathcal{K}$ entails a specific block-like structure i.e. it consists of a Cartesian product of smaller (and simpler) cones $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_q$ (see Def. A.7 for examples of different types of cones). In the above, the conic section is also characterized by linear equality constraints using matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ of rank $m \leq n$.

To illustrate problem (2.4), we can for instance reformulate (2.1) by introducing an

4. Following the previous remark, *conic optimization* usually refers to problem involving a linear objective and conic constraints.

additional slack variable γ as follows:

$$\min_{\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, \gamma} \quad \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) : \mathbb{D} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \sqrt{\frac{2}{3}}\sigma_0\gamma \tag{2.5a}$$

$$\text{s.t.} \quad \varepsilon_{11} = \varepsilon_0 \tag{2.5b}$$

$$\varepsilon_{22} = 0 \tag{2.5c}$$

$$\operatorname{tr}(\boldsymbol{\varepsilon}^p) = 0 \tag{2.5d}$$

$$\gamma \geq \|\boldsymbol{\varepsilon}^p\| \tag{2.5e}$$

where we can clearly see that Eq. (2.5e) constitutes a conic condition of the form $(\gamma, \boldsymbol{\varepsilon}^p) \in \mathcal{L}^4$ where $\mathcal{L}^4$ is the 4-dimensional quadratic Lorentz cone (see again Def. A.7). The total strain $\boldsymbol{\varepsilon}$ remains free of any conic constraint, Eq. (2.5e) can further be expressed using the following product $\boldsymbol{x} \in \mathcal{K} = \mathbb{R}^3 \times \mathcal{L}^4$ with $\boldsymbol{x} = (\boldsymbol{\varepsilon}, \gamma, \boldsymbol{\varepsilon}^p)$. In the above formulation, an important step has been taken. We replaced an initially non-smooth objective due to the presence of $\|\boldsymbol{\varepsilon}^p\|$ by a smooth quadratic + linear objective. This has been done at the expense of including an additional non-smooth conic constraint. This conic reformulation trick through the use of additional slack variables and conic constraints is a powerful method to treat non-smoothness and cast various optimization programs into the conic optimization format (see for instance [Ben-Tal and Nemirovski, 2001; Boyd and Vandenberghe, 2004; Lobo et al., 1998] and the Mosek Modeling Cookbook [Mosek, 2019] for various examples).

Let us remark that the reformulation of problem (2.1) has been done at the material point level. In the subsequent chapters, when solving a finite-element problem, plastic strain variables will be expressed at each quadrature point and numerical quadrature will be used to compute the total dissipation as a sum of local contributions. Similar reformulations will then be performed at each quadrature point [5] yielding a block-like structure of the total cone (ignoring for now total strains) $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_q$ where q is the total number of quadrature points and each $\mathcal{K}_i = \mathcal{L}^4$ is a quadratic Lorentz cone.

2.3.1 Linear programs and quadratic programs

While convex and smooth problems are usually simple to solve (see section 2.4.1), non-smooth convex problems are more difficult to handle. In particular, *Linear Programming* (LP) problems are the most simple form of non-smooth convex problems. They consist of minimizing a linear objective function with linear equality and inequality constraints, i.e. $\mathcal{K} \equiv \mathbb{R}^{n+}$.

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned} \tag{2.6}$$

5. One can in fact even reformulate directly the von Mises plastic dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}^p) = \begin{cases} \sqrt{\frac{2}{3}}\sigma_0\|\dot{\boldsymbol{\varepsilon}}^p\| & \text{if } \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^p) = 0 \\ +\infty & \text{otherwise} \end{cases}$ in a conic form as $\phi(\dot{\boldsymbol{\varepsilon}}^p, \dot{\gamma}) = \sqrt{\frac{2}{3}}\sigma_0\dot{\gamma} + \mathbb{1}_{\operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^p)=0} + \mathbb{1}_{\dot{\gamma} \geq \|\dot{\boldsymbol{\varepsilon}}^p\|}$ while retaining the framework of generalized standard materials.

LP problems can nowadays be efficiently solved and are well understood from a theoretical point of view. However their use in mechanics is somewhat limited due to the simplicity of the constraints and objective functions.

Quadratic Programs (QP) are a first extension of LP which includes a positive quadratic form in its objective function:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}_0\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{array} \tag{2.7}$$

Such problems are often found in basic mechanical and thermal problems where $\boldsymbol{Q}_0$ can be seen as the system stiffness matrix or conductivity matrix and $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ are a set of simple boundary conditions. Inequality constraints can be used to model simple unilateral conditions for instance. Removing the plastic part in (2.5), i.e. considering only a purely elastic material, directly yields a typical QP program, without inequality constraints.

2.3.2 Second order cone programs

Second-Order Cone Programs (SOCP) are a further generalization of LP which consist of minimizing a linear objective function over a set of equality constraints and generalized inequality constraints defined over Lorentz second-order cones, i.e. $\mathcal{K} \equiv \mathcal{L}$ (see Def. A.7), or a product of such cones:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \in \mathcal{L} \end{array} \tag{2.8}$$

It turns out that a particular Euclidean Jordan algebra underlies the analysis of SOCP and allowed the extension of LP solvers to the case of SOCP problems. SOCP is a particularly expressive class of conic optimization which enables to model many interesting situations. A comprehensible study of the relationship between SOCP and QP/LP can be found in Alizadeh and Goldfarb [2003].

One particular problem of interest is a further generalization of QP programs when introducing a quadratic term in the inequality constraints. The resulting problem called *quadratically constrained quadratic program* (QCQP) can be written as follows:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}_0\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}_i\boldsymbol{x} + \boldsymbol{d}_i^T\boldsymbol{x} + \boldsymbol{e}_i \leq 0 \quad \forall i \in \{1, \ldots, m\} \end{array} \tag{2.9}$$

In fact, QCQP can be systematically transformed into standard SOCP. Supposing that the matrix $\boldsymbol{Q}_i$ is positive semi-definite, the passage from a quadratic to a second-

order cone constraint is straightforward using this equivalence:

$$\begin{aligned} \tfrac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}_i\boldsymbol{x} + \boldsymbol{d}_i^T\boldsymbol{x} + \boldsymbol{e}_i \leq 0 \quad &\Leftrightarrow \quad \begin{cases} t + \boldsymbol{d}_i^T\boldsymbol{x} + \boldsymbol{e}_i \leq 0 \\ \tfrac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}_i\boldsymbol{x} \leq t \end{cases} \\ &\Leftrightarrow \quad \begin{cases} t + \boldsymbol{d}_i^T\boldsymbol{x} + \boldsymbol{e}_i \leq 0 \\ \|\boldsymbol{F}_i\boldsymbol{x}\|^2 \leq 2t \end{cases} \\ &\Leftrightarrow \quad \begin{cases} t + \boldsymbol{d}_i^T\boldsymbol{x} + \boldsymbol{e}_i \leq 0 \\ (t, 1, \boldsymbol{F}_i\boldsymbol{x}) \in \mathcal{L}_r^{n+2} \end{cases} \end{aligned} \tag{2.10}$$

where $\boldsymbol{Q}_i = \boldsymbol{F}_i^T\boldsymbol{F}_i$ is the Cholesky factorization of $\boldsymbol{Q}_i$ and $\mathcal{L}_r^{n+2}$ is the rotated second-order Lorentz cone (see again Def. A.7). The same thing can be done for the quadratic form $\boldsymbol{Q}_0$ appearing in the objective function so that the resulting problem can indeed be expressed in SOCP form.

Our illustrative problem (2.5) entails both a quadratic objective and a conic constraint. It is therefore more general than a QCQP program. It can be written in the following simple matrix format:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \dfrac{1}{2}\boldsymbol{x}^T\boldsymbol{K}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \in \mathbb{R}^3 \times \mathcal{L}^4 \end{array} \tag{2.11}$$

where:

$$\boldsymbol{x} = \{\boldsymbol{\varepsilon}, \gamma, \boldsymbol{\varepsilon}^p\}^T = \{\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, \gamma, \varepsilon_{11}^p, \varepsilon_{22}^p, \varepsilon_{33}^p\}^T \tag{2.12}$$

$$\boldsymbol{K} = \left[\begin{array}{c|c|c} \boldsymbol{D} & \boldsymbol{0} & -\boldsymbol{D} \\ \hline \boldsymbol{0} & 0 & \boldsymbol{0} \\ \hline -\boldsymbol{D} & \boldsymbol{0} & \boldsymbol{D} \end{array}\right] \tag{2.13}$$

$$\boldsymbol{A} = \left[\begin{array}{ccc|c|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right] \tag{2.14}$$

$$\boldsymbol{c} = \{0, 0, 0, \sqrt{2/3}\sigma_0, 0, 0, 0\}^T \tag{2.15}$$

$$\boldsymbol{b} = \{\varepsilon_0, 0, 0\}^T \tag{2.16}$$

$$\mathcal{L}^4 = \{\boldsymbol{z} = (z_0, \overline{\boldsymbol{z}}) = (\gamma, \boldsymbol{\varepsilon}^p) \mid \gamma \geq \|\boldsymbol{\varepsilon}^p\|\} \tag{2.17}$$

Unfortunately, only few solvers can handle both a quadratic objective and a conic constraint. Most solvers may handle either only QCQP programs or standard SOCP programs which must be transformed in the standard SOCP format (2.8). The quadratic term given by the stiffness matrix $\boldsymbol{K}$ must therefore be removed from the objective function and reformulated as a conic constraint in order to fit the SOCP standard format (2.8). This requires, as seen in equations (2.10), a Cholesky factorization of the elementary stiffness matrix. While this operation is technically possible, several drawbacks motivated us to avoid this method and to decide to adapt the solving algorithms in order to directly handle the quadratic elastic stiffness term:

1. A Cholesky factorization of an elasticity matrix does not entail a strong physical meaning. The elastic behavior, normally translated throughout this quadratic form, will be transformed into a different, not directly understandable, *elastic conic condition* for which we lack a clear physical interpretation.
2. When transforming the quadratic function into a conic constraint, additional slack variables are introduced. As it will later on be seen in Sec. 2.3.5, for each set of variables constrained in a cone, an equal number of *dual* variables will also be included, doubling the required number of variables necessary to solve the optimization program, impeding the final computational efficiency.
3. The need for conic reformulation is a way to handle non-smooth objective terms. Such a quadratic term is inherently smooth and it feels therefore unnecessary to treat it on the same level as the plastic non-smooth term.

Based on these different drawbacks, our work will extend regular SOCP solvers to take into account the additional quadratic (or smooth in general) terms in the objective function.

2.3.3 Semi-definite programs

Although not investigated in this manuscript, let us mention that *Semi-Definite Programs* (SDP) is a further generalization of SOCP (LP$\subset$ SOCP $\subset$ SDP) in which optimization variables may also include symmetric matrices $\boldsymbol{X} \in \mathcal{S}^n$ which are restricted to be positive semi-definite (PSD) $\boldsymbol{X} \succeq 0$, the cone of PSD matrices being denoted by $\mathcal{S}^n_+$ (see Def. A.7):

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{X}_i} \quad & \boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^{p} \boldsymbol{C}_i : \boldsymbol{X}_i & & \\ \text{s.t.} \quad & \sum_{j=1}^{n} \boldsymbol{A}_{ij}\boldsymbol{x}_j + \sum_{j=1}^{p} \boldsymbol{B}_{ij} : \boldsymbol{X}_i & \forall i \in \{1,\dots,m\} & \\ & \boldsymbol{x} \in \mathcal{K} & & \\ & \boldsymbol{X}_i \in \mathcal{S}^n_+ & \forall i \in \{1,\dots,p\} & \end{aligned} \tag{2.18}$$

where $\boldsymbol{A} : \boldsymbol{B}$ is the inner product of matrices.

Second-order cone conditions can easily be expressed as semi-definite cones, making every SOCP or QCQP problem solvable using an SDP algorithm. Indeed, for $\boldsymbol{x} = (x^0, \overline{\boldsymbol{x}}) \in \mathbb{R} \times \mathbb{R}^n$ we have that:

$$x^0 \geq \|\overline{\boldsymbol{x}}\| \quad \Leftrightarrow \quad \boldsymbol{X} = \boldsymbol{mat}(\boldsymbol{x}) = \begin{bmatrix} x^0 & \overline{\boldsymbol{x}}^T \\ \overline{\boldsymbol{x}} & x^0\boldsymbol{I}_n \end{bmatrix} \in \mathcal{S}^{n+1}_+ \tag{2.19}$$

SDP allows us to express an even wider variety of problems and mechanical behaviors. For instance, 3D limit analysis problems involving a Mohr-Coulomb, Tresca or Rankine strength criteria can only be formulated as SDP programs [Krabbenhøft et al., 2007; Makrodimopoulos, 2010]. However, the computational effort per iteration required to solve SDP problems of similar size and structure is greater than the one required to solve SOCP problems of similar size.

2.3.4 Duality in conic programming

Let us now briefly describe the dual problem associated with conic programs of the form (2.4) which will be important in the derivation of our algorithm. Similarly to classical Lagrangian duality (see Def. A.13), the Lagrangian associated with the conic program (2.4) reads as:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = f(\boldsymbol{x}) + \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{x}^T\boldsymbol{s} \tag{2.20}$$

where $\boldsymbol{y}$ is the linear constraint Lagrange multiplier vector and $\boldsymbol{s}$ the dual conic variables associated with $\boldsymbol{x}$.

The dual problem is then obtained from the Lagrangian as a max / min problem with the dual variables belonging to the dual cone of $\mathcal{K}$ (see Def. A.6) i.e. $\boldsymbol{s} \in \mathcal{K}^*$:

$$\max_{\boldsymbol{y}, \boldsymbol{s} \in \mathcal{K}^*} \min_{\boldsymbol{x}} \quad \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = f(\boldsymbol{x}) + \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{x}^T\boldsymbol{s} \tag{2.21}$$

Isolating the minimization over $\boldsymbol{x}$ and replacing it with a maximization yields:

$$\max_{\boldsymbol{y}, \boldsymbol{s} \in \mathcal{K}^*} \left\{ -\boldsymbol{b}^T\boldsymbol{y} - \max_{\boldsymbol{x}} \{ \boldsymbol{x}^T \left(-\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} \right) - f(\boldsymbol{x}) \} \right\} \tag{2.22}$$

Finally, the maximization problem over $\boldsymbol{x}$ can be simplified when recognizing it as a Legendre-Fenchel transformation, yielding finally the dual problem:

$$\begin{array}{ll} \max\limits_{\boldsymbol{y}, \boldsymbol{s}} & -\boldsymbol{b}^T\boldsymbol{y} - f^*(-\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s}) \\ \text{s.t.} & \boldsymbol{s} \in \mathcal{K}^* \end{array} \tag{2.23}$$

where f^* is the convex conjugate function of f (see Def. A.12).

In the important case where f is linear e.g. $f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x}$, $f^*(\boldsymbol{z}) = \mathbb{1}_{\boldsymbol{z}=\boldsymbol{c}}(\boldsymbol{z})$ and the dual problem reduces to:

$$\begin{array}{ll} \max\limits_{\boldsymbol{y}, \boldsymbol{s}} & -\boldsymbol{b}^T\boldsymbol{y} \\ \text{s.t.} & -\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c} \\ & \boldsymbol{s} \in \mathcal{K}^* \end{array} \tag{2.24}$$

Going back to our illustrative example (2.11), the dual cone $\mathcal{K}^* = 0^3 \times \mathcal{L}^4$ (the Lorentz cone $\mathcal{L}$ being self-dual) so that we have $\boldsymbol{s} \in \mathcal{K}^* = \{0, 0, 0, s_0, \bar{\boldsymbol{s}}\}$ with $\|\bar{\boldsymbol{s}}\| \leq s_0$:

$$\boldsymbol{z} = \begin{Bmatrix} \boldsymbol{z}_\epsilon \\ z_\gamma \\ \boldsymbol{z}_{\epsilon^p} \end{Bmatrix} = -\boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{s} = \begin{Bmatrix} -y_1 \\ -y_2 \\ 0 \\ s_0 \\ -y_3 + s_1 \\ -y_3 + s_2 \\ -y_3 + s_3 \end{Bmatrix} \tag{2.25}$$

When computing the conjugate function $f^*(\boldsymbol{z})$ of $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{K}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x}$, one obtains that $\boldsymbol{z}_{\boldsymbol{\varepsilon}} = -\boldsymbol{z}_{\boldsymbol{\varepsilon}^p}$, $z_\gamma = \sqrt{\frac{2}{3}}\sigma_0$ and $f^*(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{z}_{\boldsymbol{\varepsilon}}^T\boldsymbol{D}^{-1}\boldsymbol{z}_{\boldsymbol{\varepsilon}}$, so that the dual variables $\boldsymbol{y}$ and $\boldsymbol{s}$ are linked as follows:

$$\boldsymbol{z}_{\boldsymbol{\varepsilon}} = \begin{Bmatrix} -y_1 \\ -y_2 \\ 0 \end{Bmatrix} = \begin{Bmatrix} y_3 - s_1 \\ y_3 - s_2 \\ y_3 - s_3 \end{Bmatrix} \tag{2.26}$$

$$s_0 = z_\gamma = \sqrt{\frac{2}{3}}\sigma_0 \tag{2.27}$$

Introducing the following new notations: $\boldsymbol{\sigma} = \{\sigma_{11}, \sigma_{22}, 0\} = \boldsymbol{z}_{\boldsymbol{\varepsilon}}$, $y_3 = -p$ and $\widetilde{\boldsymbol{s}} = -\bar{\boldsymbol{s}}$, we finally obtain the dual problem to (2.1):

$$\max_{\boldsymbol{\sigma}, p, \boldsymbol{s}} \quad \varepsilon_0\sigma_{11} - \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{D}^{-1}\boldsymbol{\sigma} \tag{2.28a}$$

$$\text{s.t.} \quad \boldsymbol{\sigma} = -p\mathbf{1} + \widetilde{\boldsymbol{s}} \tag{2.28b}$$

$$\|\widetilde{\boldsymbol{s}}\| \leq \sqrt{\frac{2}{3}}\sigma_0 \tag{2.28c}$$

where we can see that the dual variable $\boldsymbol{z}_{\boldsymbol{\varepsilon}}$ can be reinterpreted as the stress state $\boldsymbol{\sigma}$ which is indeed in plane-stress conditions in the third direction, the Lagrange multiplier associated with the incompressibility condition can be interpreted as the hydrostatic stress p and the dual conic variables $(s_0, \bar{\boldsymbol{s}})$ respectively correspond to the yield limit $\sqrt{\frac{2}{3}}\sigma_0$ and to the deviatoric stress (see the hydrostatic/deviatoric decomposition of (2.28b)) up to a sign. The dual conic constraint $(s_0, \bar{\boldsymbol{s}}) \in \mathcal{K}^*$ rewritten as (2.28c) clearly expresses the von Mises yield criterion. Finally, up to changing its sign and transforming the maximization into a minimization, the objective can be recognized as the complementary energy minimum principle (expressed at the material point level) with the elastic energy density $\frac{1}{2}\boldsymbol{\sigma}\boldsymbol{D}^{-1}\boldsymbol{\sigma}$ expressed in terms of stresses, the term $\varepsilon_0\sigma_{11}$ being the work of the imposed displacements (strains here). Again, a more rigorous expression of the stress-based variational principle in an elastoplastic setting will be given later in chapter 4.

2.3.5 Non-linear and non-convex optimization problems over conic constraints

Although non-convex optimization problems are usually treated with radically different optimization techniques, due to the existence of local minima, we will in chapter 5 treat a non-convex conic problem which only slightly deviates from the convex conic programming framework (2.4). Indeed, we will assume that non-smooth objectives are always convex and can be reformulated as conic constraints. The remaining smooth objectives or constraints may however be non-convex. This leads us to the following Second-Order cone Non-Linear Programming (SONLP) problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{x} \in \mathcal{K} \end{array} \tag{2.29}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a scalar-valued and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$ a vector-valued function, both functions are assumed to be sufficiently smooth and possibly non-convex functions. Note that standard SOCP problems correspond to the particular case of (2.29) in which f is linear and $\boldsymbol{g}$ is absent (or linear, in such a case it can be appended to matrix $\boldsymbol{A}$). Refering again to our illustrative example, one example of such a situation would be the case of a non-convex strain energy density instead of the classical quadratic convex energy.

One important feature of problem (2.29) is that non-linear constraints appear in two very different forms: general but smooth non-linear constraints via $\boldsymbol{g}$ and non-smooth but specific SOC constraints via $\mathcal{K}$. In particular, both types of constraints will be treated differently in the solution algorithm described in section 2.5 by extending classical primal-dual interior point methods dedicated to tackling non-smooth conic constraints to the case of additional smooth non-linear constraints. Similarly, the objective function is assumed to be smooth. Modeling the presence of non-smooth terms such as $\|\boldsymbol{x}\|_2$ would require appropriate conic reformulation to treat it using SOC constraints as discussed before.

2.3.6 The optimality conditions

Let us now give the optimality conditions characterizing the solutions of generic problems of the form (2.29). Let $\boldsymbol{w} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{s})^T \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n$ be the concatenation of all variables and the constraints dual variables. The Lagrangian function of problem (2.29) is defined as follows:

$$\mathcal{L}(\boldsymbol{w}) = f(\boldsymbol{x}) + \boldsymbol{y}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) + \boldsymbol{z}^T\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{s}^T\boldsymbol{x} \tag{2.30}$$

Then Karush-Kuhn-Tucker (KKT) first-order conditions for optimality of problem (2.29) are given by the following system of non-linear equations:

$$\boldsymbol{r}(\boldsymbol{w}) = \left\{ \begin{array}{c} \nabla_x \mathcal{L} \\ \nabla_y \mathcal{L} \\ \nabla_z \mathcal{L} \\ \boldsymbol{x}^T\boldsymbol{s} \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{\nabla}_x f + \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{G}^T\boldsymbol{z} - \boldsymbol{s} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \boldsymbol{g}(\boldsymbol{x}) \\ \boldsymbol{x}^T\boldsymbol{s} \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{array} \right\} \quad \text{and } \boldsymbol{x} \in \mathcal{K},\ \boldsymbol{s} \in \mathcal{K}^* \tag{2.31}$$

with $\boldsymbol{G}(\boldsymbol{x}) \in \mathbb{R}^{p \times n}$ being the collection of the gradients of the non-linear constraints:

$$\boldsymbol{G}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{\nabla}_x g_1 \\ \vdots \\ \boldsymbol{\nabla}_x g_p \end{bmatrix} \tag{2.32}$$

and where $\boldsymbol{x}^T\boldsymbol{s} = 0$ along with $\boldsymbol{x} \in \mathcal{K},\ \boldsymbol{s} \in \mathcal{K}^*$ form the so-called *conic complementarity conditions* (see Def. A.11).

As regards our illustrative problem (2.5), we already discussed the interpretation of the dual variables. In this case, $\boldsymbol{g}$ and its associated Lagrange multiplier is absent. The KKT conditions can therefore be rewritten as:

$$\nabla_x \mathcal{L} = \boldsymbol{K}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{A}^T\boldsymbol{y} - \boldsymbol{s} = \boldsymbol{0} \Rightarrow \begin{cases} \boldsymbol{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) - \boldsymbol{\sigma} = \boldsymbol{0} \\ \sqrt{\frac{2}{3}}\sigma_0 - s_0 = 0 \\ -\boldsymbol{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) - p\boldsymbol{1} + \widetilde{\boldsymbol{s}} = \boldsymbol{0} \end{cases} \tag{2.33}$$

$$\nabla_y \mathcal{L} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} = 0 \Rightarrow \begin{cases} \varepsilon_{11} = \varepsilon_0 \\ \varepsilon_{22} = 0 \\ \mathrm{tr}(\boldsymbol{\varepsilon}^p) = 0 \end{cases} \tag{2.34}$$

$$\boldsymbol{x}^T\boldsymbol{s} = \gamma s_0 + (\boldsymbol{\varepsilon}^p)^T\bar{\boldsymbol{s}} = 0 \Rightarrow \sqrt{\frac{2}{3}}\sigma_0\gamma - \boldsymbol{\varepsilon}^p : \widetilde{\boldsymbol{s}} = 0 \tag{2.35}$$

$$\boldsymbol{x} \in \mathcal{K} \Rightarrow \|\boldsymbol{\varepsilon}^p\| \leq \gamma \tag{2.36}$$

$$\boldsymbol{s} \in \mathcal{K}^* \Rightarrow \|\widetilde{\boldsymbol{s}}\| \leq \sqrt{\frac{2}{3}}\sigma_0 \tag{2.37}$$

The first two sets of conditions express the elastoplastic constitutive relation and the imposed strain-state. The conic complementarity conditions (2.35) to (2.37) can be further combined with the Cauchy-Schwartz inequality yielding [6]

$$\sqrt{\frac{2}{3}}\sigma_0\boldsymbol{\varepsilon}^p - \gamma\widetilde{\boldsymbol{s}} = \boldsymbol{0} \tag{2.38}$$

Combining again with (2.35), one obtains that:

$$\boldsymbol{\varepsilon}^p = \gamma\frac{\widetilde{\boldsymbol{s}}}{\sqrt{\frac{2}{3}}\sigma_0} \tag{2.39}$$

$$\gamma\left(1 - \left(\frac{\|\widetilde{\boldsymbol{s}}\|}{\sqrt{\frac{2}{3}}\sigma_0}\right)^2\right) = 0 \tag{2.40}$$

which respectively express the plastic flow rule and plasticity consistency condition.

2.4 An overview of the solving algorithms

All of the presented problems can be qualified as a (generalized) inequality constrained optimization problems. Solving these problems rely on various algorithms each

6. Indeed, for a Lorentz cone and $\boldsymbol{x}, \boldsymbol{s} \in \mathcal{L}$ we have that $\boldsymbol{x}^T\boldsymbol{s} = 0 \Longleftrightarrow \boldsymbol{x} \circ \boldsymbol{s} = (x_0 s_0 + \bar{\boldsymbol{x}}^T\bar{\boldsymbol{s}}, x_0\bar{\boldsymbol{s}} + s_0\bar{\boldsymbol{x}}) = \boldsymbol{0}$, see Def. A.11 and Eq. (B.7).

of them treating the constraints differently. In addition to that, non-smoothness requires specific attention in order to adapt the algorithms. In the following, we do not discuss algorithms dedicated to finding a global minimum for non-convex problems nor do we give an extensive literature review and refer only to the monographs of Boyd and Vandenberghe [2004] and Nocedal et al. [2009].

2.4.1 Unconstrained optimization

The simplest optimization algorithms for smooth unconstrained problems are descent methods. For instance, the gradient descent method is a *first-order method* as it does not require to invert a linear system to find the optimal descent direction (it relies only on first order derivatives, the gradient of the objective function). Newton's method is its most natural extension, being a *second-order method* which requires solving a linear system involving the objective function Hessian matrix. Due to its simplicity, the gradient descent converges only linearly and cannot be applied to non-smooth functions. Proximal algorithms can be seen as a generalization in the non-smooth case.

Newton's method is an extremely attractive and widely used method due to its theoretical quadratic convergence near the optimal point. Computational cost caused by the Hessian system inversion can be overcome using approximation techniques such as quasi-Newton methods (Broyden, BFGS, Levenberg–Marquardt...).

2.4.2 Constrained optimization

Solving constrained optimization problems is much more difficult than unconstrained ones. A first set of algorithms rely on eliminating the constraints and solving an equivalent unconstrained optimization problem. Penalty methods are simple way of including the cost of violating the constraints in the objective function. The penalty parameter will set the trade-off cost of this violation compared to the initial objective function. Penalty methods are therefore *exterior-point methods* since iterates can fall outside the initial feasible region due to the relaxed constraints. Obviously an important difficult in practice is the selection of the penalty parameter and the associated problem ill-conditioning when it becomes too large. Augmented Lagrangian methods combine the idea of using Lagrange multipliers (it is a *primal-dual* method) with the penalty method and alleviate the need for using a large enough parameter. In particular, a variant of the standard augmented Lagrangian method that uses partial updates known as the alternating direction method of multipliers (ADMM) recently gained some attention in the context of conic optimization and has been implemented in open-source solvers such as `SCS`, `COSMO` or `CDCS`. Unfortunately, such first-order methods converge quite slowly and can be used in practice only at the expense of a reduced accuracy in the computed solution. Such kinds of algorithms are nowadays quite popular in the image and signal processing community which are mostly motivated by obtaining a sufficient decrease of some loss function (e.g. noise level) rather than compute the exact solution of the underlying optimization problem. As a result, these solutions do not seem mature enough yet for our structural engineering applications.

Following the same idea of removing the constraints, active-set methods identify,

within an iterate, the set of active and inactive constraints in order to solve unconstrained problems until final identification of the active set of constraints. Successive linear programming or sequential quadratic programming can be interpreted as active set methods and consist in replacing the objective and the constraints by a linear or quadratic approximation around the current iterate. Unfortunately, such techniques rely on some smoothness assumption of the objective or constraints.

Barrier methods can be seen as a counter-part of penalty methods in which constraints are replaced by a barrier function i.e. a function which diverges as the point approaches the feasible region boundary. These methods are therefore *interior-point methods* since each iterate will be feasible. In particular, this requires the initial point to be feasible which is not always easy to find without any knowledge of the underlying problem. Although not being extremely efficient since they suffer from similar drawbacks as penalty methods, barrier methods found new interests due to their connection with the much more efficient *primal-dual interior point methods*.

2.4.3 Non-smooth problems

As we have seen, many methods are dedicated to smooth problems. Reformulation and smoothing methods are a first class of methods which consist in using a specific regularization technique to eliminate non-smoothness, thus reformulating the problem as a smooth optimization program. These techniques suffer from a similar drawback as penalty methods, namely the choice of the regularization parameter, the ill-conditioning when approaching the original behavior and, even more importantly, a change in the underlying physical model when applied to mechanical problems. For instance, biviscous regularization is a widely used technique to replace intrinsic viscoplastic models of yield-stress fluids by a surrogate non-linear viscous behavior. Unfortunately, abandoning the original non-smooth behavior results in a loss on the existence of real solid (rigid) regions in viscoplastic flows, return to rest in finite time, etc.

Linear Programming is an important class of non-smooth optimization problems which was first efficiently solved using the *simplex method* developed by Dantzig et al. [1955]. The simplex method relies on finding the minimum of the objective function by exploring the vertices of the polytope defined by the constraints in the opposite direction of the gradient. The starting point for the simplex method must be a vertex. Thereafter, every iteration moves to an adjacent vertex, decreasing the objective as it goes, until an optimal vertex is found. Simplex methods are a well established technology for LP however their extension to non-linear constraints is not straightforward. Another drawback is the exponential theoretical complexity of the algorithm which makes it impractical for large-scale LP problems. This drawback has been resolved with the advent of interior-point algorithms which exhibit polynomial complexity and which we will discuss in the next section.

To conclude and further motivate the choices made in the present work, let us quote T. Rockafellar from [Rockafellar, 1994] on non-smooth optimization which perfectly summarizes our approach:

A function is smooth if it is differentiable and the derivatives are continuous. [...] From this perspective a nonsmooth function only has a negative description – it lacks some degree of properties traditionally relied upon in analysis. One could get the impression that "nonsmooth optimization" is a subject dedicated to overcoming handicaps which have to be faced in miscellaneous circumstances where mathematical structure might be poorer than what one would like. But this is far from right. Instead, nonsmooth optimization typically deals with highly structured problems, but problems which arise differently, or are modeled or cast differently, from ones for which many of the mainline numerical methods, involving gradient vectors and Hessian matrices, have been designed. [...]

A strong argument can be made for the notion that nonsmoothness in optimization is very often a question of modeling, and due to the prevalence of inequality constraints, is present anyway in almost all problems of importance, at least in the background. The issue from that angle is simply how to make use of available structure in the best possible way. Nonsmooth optimization gives the message that many effective approaches are possible, and one need not be confined to a classical view of how functions are to be approximated and evaluated. Because nonsmoothness has different manifestations and treatments, one shouldn't imagine that numerical techniques in nonsmooth optimization can act as "black boxes." Techniques are developed for the particular structures that compensate for the absence of differentiability. It's important therefore to understand the source of any nonsmoothness, before deciding how it might be handled.

2.5 Primal-dual interior point method for problems with conic constraints

2.5.1 Historical background

The 1960s were the heyday of unconstrained optimization, and, as a result, it was common practice to convert constrained problems into unconstrained subproblems or sequences of unconstrained subproblems using for instance penalty or barrier methods. Interior-point methods, which originated from barrier methods, mainly emerged in the late 1970s and 1980s, as a response to the concerns regarding the exponential complexity of simplex methods. In 1984, Karkamar was the first to report a much faster method to solve large-scale LP problems giving birth to interior-point methods with polynomial complexity [Karmarkar, 1984].

Since then, a lot of research activity produced a wide range of theoretical results and implementations. One decade after Karmarkar's publication a subclass of interior-point methods, namely *primal-dual methods*, arose as the most efficient version of interior-point methods. These methods roughly consist in applying Newton's method to a set of perturbed non-linear equations stated in terms of the original or « primal » problem

variables, along with « dual » variables representing the Lagrange multipliers. Authors focused on the path-following algorithms, asymptotic convergence properties, infeasible IPM and numerical implementations [Hertog, 1994; Jansen, 1997; Wright, 1997].

It wasn't until Nesterov's introduced the concept of « self-concordant barrier functions » that the IPM scope of polynomial-time complexity was extended to a wide family of convex optimization problems including SOCP and SDP [Nesterov and Nemirovskii, 1994; Nesterov and Todd, 1997, 1998]. The advent of interior-point method is often considered as a revolution in optimization theory which now views linear and non-linear programming in a unified perspective.

Since then, various improvements were implemented in what is today a reliable technology known as the path-following primal-dual interior point method [Alizadeh and Goldfarb, 2003; Andersen et al., 2003; Boyd and Vandenberghe, 2004; Mehrotra, 1992]. A condensed yet complete look at the evolution of IPM can be found in Wright [2005].

Finally, although theoretical complexity of IPM is polynomial, state-of-the-art implementations usually exhibit in practice much less iterations than what would be predicted by the theoretical bounds. As a result, IPM are usually viewed as really robust algorithms, scaling very well with the problem complexity. However, the exact origin of why IPM are much more efficient in practice than what theory would suggest is not fully understood yet.

2.5.2 General primal-dual interior point framework

The main idea of the IPM consists in finding a solution to the KKT conditions given by (2.31) by following the neighborhood of a curve called the *central path* which consists of a sequence of iterates $\boldsymbol{w}(\eta) = (\boldsymbol{x}(\eta), \boldsymbol{y}(\eta), \boldsymbol{z}(\eta), \boldsymbol{s}(\eta))$ parameterized by a *barrier parameter* $\eta \geq 0$. These iterates are *interior points* i.e. they satisfy the conic constraints $\boldsymbol{x} \in \operatorname{int} \mathcal{K}$ and $\boldsymbol{s} \in \operatorname{int} \mathcal{K}^*$.

The main property of the central path is that it defines a continuous set of strictly feasible points which are well-centered, i.e. far from the boundary of the feasible region except when reaching the optimum for $\eta \to 0$ (see Fig. 2.1). This will allow for large descent steps to be taken when minimizing the objective function from points lying in the neighborhood of the central path.

The central path is defined as the solution to the perturbed KKT system or *barrier* KKT system (BKKT) defined as follows:

$$\boldsymbol{r}(\boldsymbol{w}, \eta) = \begin{Bmatrix} \nabla_x \mathcal{L} \\ \nabla_y \mathcal{L} \\ \nabla_z \mathcal{L} \\ \boldsymbol{x} \circ \boldsymbol{s} - \eta \boldsymbol{e} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{\nabla}_x f + \boldsymbol{A}^T \boldsymbol{y} + \boldsymbol{G}^T \boldsymbol{z} - \boldsymbol{s} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \boldsymbol{g}(\boldsymbol{x}) \\ \boldsymbol{x} \circ \boldsymbol{s} - \eta \boldsymbol{e} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{Bmatrix} \quad \text{and } \boldsymbol{x} \in \operatorname{int} \mathcal{K},\ \boldsymbol{s} \in \operatorname{int} \mathcal{K}^* \tag{2.41}$$

in which operator $\circ$ and $\boldsymbol{e}$ depend on the types of the cones defining $\mathcal{K}$ (see for instance appendix B where $\boldsymbol{e} = (1, \boldsymbol{0})$ of dimension $n+1$ is defined for a second order cone $\mathcal{L}^{n+1}$).

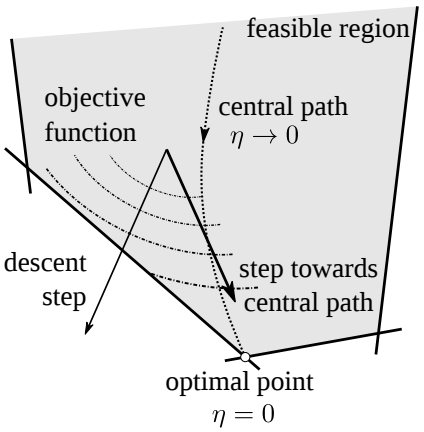


Figure 2.1 – General idea of a IP algorithm (adapted from [Bleyer, 2017])

The introduction of the perturbed KKT system, the main idea of the IPM, is designed in order to tackle the complementarity constraint $\boldsymbol{x}^T\boldsymbol{s} = 0$. All other constraints, especially those involving $\boldsymbol{g}$, remain unchanged. As it will be discussed later, the proposed algorithm will therefore be similar to a Newton-method on the smooth non-linear residuals and a modified Newton-method on the conic residuals, exactly as in the standard SOCP case.

2.5.3 Solving the perturbed KKT system

As mentioned earlier, the aim of the IPM is to find a series of iterates $\boldsymbol{w}(\eta) = (\boldsymbol{x}(\eta), \boldsymbol{y}(\eta), \boldsymbol{z}(\eta), \boldsymbol{s}(\eta))$ while driving η to 0, yielding, at convergence, a solution to the original unperturbed KKT system (2.31). For that, a *primal-dual* method will be used in which both primal and dual variables remain unknown.

At each iteration (k), a Newton step on the perturbed KKT system (2.41) is computed towards the central path for a fixed value of $\eta^{(k)}$. The solution is updated after a step-length calculation and the barrier parameter is reduced by some amount. The process is repeated until the residuals fall under a certain tolerance. This will result in a series of iterates $\boldsymbol{w}^{(k)} = (\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}, \boldsymbol{z}^{(k)}, \boldsymbol{s}^{(k)})$ which remain feasible with respect to the conic constraints.

Given an iteration (k) at which a point $\boldsymbol{w}^{(k)}$ satisfying the conic constraints is known and a value for the barrier parameter $\eta^{(k)}$ has been chosen, the next iterate is calculated by computing a new point near the central path. This is obtained by performing one iteration of the Newton method when linearizing the residual equations such as:

$$\boldsymbol{r}^{(k+1)}(\boldsymbol{w}, \eta) = \boldsymbol{r}^{(k)}(\boldsymbol{w}) + \boldsymbol{r'}^{(k)}(\boldsymbol{w}, \eta).\boldsymbol{\Delta w} = 0 \tag{2.42}$$

thus solving the system $\boldsymbol{J}^{(k)}.\Delta\boldsymbol{w} = -\boldsymbol{r}^{(k)}(\boldsymbol{w}, \eta)$ with the jacobian matrix $\boldsymbol{J}^{(k)} = \boldsymbol{r'}^{(k)}(\boldsymbol{w})$

and where $\Delta \boldsymbol{w} = (\Delta \boldsymbol{x}, \Delta \boldsymbol{y}, \Delta \boldsymbol{z}, \Delta \boldsymbol{s})$ is a descent direction. More precisely, one has:

$$\begin{bmatrix} \boldsymbol{H}^{(k)} & \boldsymbol{A}^T & (\boldsymbol{G}^{(k)})^T & -\boldsymbol{I} \\ \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{G}^{(k)} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{S}^{(k)} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{X}^{(k)} \end{bmatrix} \cdot \begin{Bmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \\ \Delta \boldsymbol{z} \\ \Delta \boldsymbol{s} \end{Bmatrix} = \begin{Bmatrix} -\boldsymbol{r}_d^{(k)} \\ -\boldsymbol{r}_p^{(k)} \\ -\boldsymbol{r'}_p^{(k)} \\ -\boldsymbol{r}_c^{(k)} \end{Bmatrix} \tag{2.43}$$

in which $\boldsymbol{H}^{(k)}$ is the Hessian of the Lagrangian with respect to $\boldsymbol{x}$ given by:

$$\boldsymbol{H}^{(k)} = \boldsymbol{\nabla}^2_{xx} \mathcal{L}^{(k)} = \boldsymbol{\nabla}^2_{xx} f^{(k)} - \sum_{i=1}^{p} z_i^{(k)} \boldsymbol{\nabla}^2_{xx} g_i^{(k)} \tag{2.44}$$

and $\boldsymbol{X}^{(k)} = \boldsymbol{diag}(\ldots, \boldsymbol{mat}(\boldsymbol{x}_i^{(k)}), \ldots)$ and $\boldsymbol{S}^{(k)} = \boldsymbol{diag}(\ldots, \boldsymbol{mat}(\boldsymbol{s}_i^{(k)}), \ldots)$ (see Eq. (B.2)) and the vector $\boldsymbol{r}^{(k)}$ corresponding to the vector of residuals for the linearized equations of the KKT system:

$$\boldsymbol{r}_d^{(k)} = \boldsymbol{\nabla}_x f^{(k)} + \boldsymbol{A}^T \boldsymbol{y}^{(k)} + (\boldsymbol{G}^{(k)})^T \boldsymbol{z}^{(k)} - \boldsymbol{s}^{(k)} \tag{2.45a}$$

$$\boldsymbol{r}_p^{(k)} = \boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} \tag{2.45b}$$

$$\boldsymbol{r'}_p^{(k)} = \boldsymbol{g}(\boldsymbol{x}^{(k)}) \tag{2.45c}$$

$$\boldsymbol{r}_c^{(k)} = \eta^{(k)} \boldsymbol{e} - \boldsymbol{X}^{(k)} \boldsymbol{S}^{(k)} \boldsymbol{e} \tag{2.45d}$$

2.5.4 The basic primal-dual interior point algorithm

2.5.4.1 Step-length calculation

The next iterate $\boldsymbol{w}^{(k+1)}$ is obtained by performing a line-search on the maximum allowable step length α in direction $\Delta \boldsymbol{w}$, i.e. $\boldsymbol{w}^{(k+1)} = \boldsymbol{w}^{(k)} + \alpha \Delta \boldsymbol{w}$ with $\alpha \in [0, 1]$.

This step should be chosen such that the iterate still satisfies the conic constraints. A fraction-to-boundary rule using a parameter θ_{fbr} is also used to limit its value to avoid that the next iterate will lie exactly on the boundary of the feasible domain and therefore causing numerical troubles when performing the Cholesky factorization. The parameter θ_{fbr} is typically taken equal to 0.95–0.99. This yields in the following step-length:

$$\alpha = \theta_{fbr}. \max \left\{ \alpha \in]0, 1] \mid (\boldsymbol{x}^{(k)} + \alpha \Delta \boldsymbol{x}) \in \mathcal{K} \text{ and } (\boldsymbol{s}^{(k)} + \alpha \Delta \boldsymbol{s}) \in \mathcal{K}^* \right\} \tag{2.46}$$

In certain IPM implementations, a more restrictive condition is put on α requiring that in addition to verifying conic constraints, the next iterate should lie in some neighborhood of the central path defined using a certain *centrality measure* [Alizadeh and Goldfarb, 2003].

2.5.4.2 Initialization and convergence criteria

All primal non-conic variables i.e. $\boldsymbol{x}$ are initially taken equal to 0, whereas the conic variables and their dual are taken as $\boldsymbol{x}_i = (x_i^0, \overline{\boldsymbol{x}}_i) = (1, \boldsymbol{0}) = \boldsymbol{e}$. This choice ensures that we have a starting point that is feasible with the conic constraints. The initial

complementarity gap is then equal to 1 everywhere. As the residual vectors and the complementarity gap are both reduced at the same rate, the same stopping criterion is used. The iterative process is stopped when the complementarity gap and the residual norms fall bellow a given tolerance, here fixed to 10^{-8}. This point will be considered as an optimal solution for the non-perturbed KKT system.

2.5.4.3 Choice of the barrier parameter

The choice of the barrier parameter η at the beginning of each step influences the whole performance of the IP algorithm. This choice will define the type of each step, whether a « centering » step is made in which the next iterate will be closer to the central path, or a « descent » step in which the complementarity gap is reduced but maybe to the detriment of going away from the central path. We recall that in order to solve the perturbed KKT system and converge to the optimal solution, the barrier parameter η should be driven to 0.

Various strategies can be found in the literature such as the *Fiaco-McCormick monotone approach* where the barrier parameter is held fixed for a series of iterations until the KKT conditions are satisfied to some accuracy and then it is decreased.

Another possibility is to use an *adaptive approach* in which the barrier parameter is updated at the end of each iteration $(k+1)$. Since a new feasible point $\boldsymbol{w}^{(k+1)}$ has been calculated, a typical choice for the new barrier parameter $\eta^{(k+1)}$ makes use of the value of the complementarity gap $\mathsf{g}_i^{(k+1)} = (\boldsymbol{x}_i^{(k+1)})^T \cdot \boldsymbol{s}_i^{(k+1)}$ for each cone:

$$\eta^{(k+1)} = \gamma \left(\frac{1}{m} \sum_{i=1}^{m} \mathsf{g}_i^{(k+1)} \right) = \gamma \overline{\mathsf{g}}^{(k+1)} \tag{2.47}$$

m being the total number of cones. The barrier parameter $\eta^{(k+1)}$ is therefore proportional to the average complementarity gap via a scalar $\gamma \in [0, 1]$.

Other adaptive strategies can be based on the deviation of the minimum complementarity pair from the average and calculated the centering parameter γ as follows:

$$\gamma = 0.1 \min \left(0.05 \frac{1-\xi}{\xi}, 2 \right)^3 \quad \text{, where } \xi = \frac{\min\{\boldsymbol{x}_i^T \boldsymbol{s}_i,\ i = 1, \cdots, m\}}{\overline{\mathsf{g}}^{(k+1)}} \tag{2.48}$$

Note that ξ measures the deviation of the smallest complementarity product $\boldsymbol{x}_i^T \boldsymbol{s}_i$ from the average. When $\xi = 1$ (all individual products are equal to their average) we have that $\gamma = 0$ and the algorithm takes an aggressive step. Note that the rule always chooses $\gamma \leq 0.8$, so that even though the value of η may increase from one iteration to the next, it will never be chosen to be larger than the current complementarity value.

Another strategy, explained in Sec. 2.5.5.1, is known as the Mehrotra's predictor-corrector method. This method determines the value of γ using a preliminary step computation called an affine scaling step.

Other strategies rely on an approach that selects η by approximately minimizing a quality function $q_N(\gamma)$. These methods are quite expensive depending on the chosen quality function (e.g. quadratic or linear) since it requires the evaluation of problem functions, derivatives and step-lengths for every value of γ.

2.5.5 Advanced aspects

2.5.5.1 The Mehrotra predictor-corrector scheme and the adaptive choice of the barrier parameter

The IPM implementation used in this manuscript adopts the predictor-corrector scheme proposed by Mehrotra [1992] which greatly improves the IPM convergence. The predictor-corrector scheme amounts to computing two different directions, taking advantage of having to factorize the KKT matrix $\boldsymbol{J}$ only once. The Mehrotra predictor-corrector relies on the expression of the complementarity conditions:

$$\begin{aligned}(\boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}) \circ (\boldsymbol{s}^{(k)} + \Delta\boldsymbol{s}) &= (\boldsymbol{X}^{(k)} + \Delta\boldsymbol{X})(\boldsymbol{S}^{(k)} + \Delta\boldsymbol{S})\boldsymbol{e} \\ &= \left(\boldsymbol{X}^{(k)}\boldsymbol{S}^{(k)} + \Delta\boldsymbol{X}\boldsymbol{S}^{(k)} + \boldsymbol{X}^{(k)}\Delta\boldsymbol{S} + \Delta\boldsymbol{X}\Delta\boldsymbol{S}\right)\boldsymbol{e} \end{aligned} \tag{2.49}$$

When this quantity is linearized to form the Newton step of the perturbed KKT system, the quadratic term $(\Delta\boldsymbol{X}\Delta\boldsymbol{S})$ is neglected. Mehrotra suggested estimating this term from a first pure Newton step on the unperturbed KKT system. The predictor-corrector scheme is therefore the following:

1. Tthe first step, called *affine step*, corresponds to solving (2.41) with $\eta = 0$ (unperturbed KKT system). The right hand side is given by:

$$\boldsymbol{r}_d^a = \boldsymbol{\nabla}_x f^{(k)} + \boldsymbol{A}^T\boldsymbol{y}^{(k)} + (\boldsymbol{G}^{(k)})^T\boldsymbol{z}^{(k)} - \boldsymbol{s}^{(k)} \tag{2.50a}$$
$$\boldsymbol{r}_p^a = \boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} \tag{2.50b}$$
$$\boldsymbol{r'}_p^a = \boldsymbol{g}(\boldsymbol{x}^{(k)}) \tag{2.50c}$$
$$\boldsymbol{r}_c^a = \boldsymbol{X}^{(k)}\boldsymbol{S}^{(k)}\boldsymbol{e} \tag{2.50d}$$

   Its solution is denoted $\boldsymbol{\Delta w}^a$ and a line-search is realized to determine the maximum step length for this step which will be noted α_{max}^a.

2. The centering parameter is then chosen using the following heuristic [Salahi et al., 2008]:

$$\gamma = (1 - \alpha_{max}^a)^3 \tag{2.51}$$

   therefore allowing full centering steps in order to quickly attain the central path if the affine step is small.

3. The final step direction $\boldsymbol{\Delta w}$ is then computed using the value of $\eta^{(k)}$ calculated using the Eq. (2.47) and a better estimate of the complementarity gap by including the quadratic term $(\Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a)$. The residual vector is also reduced depending on γ, ensuring that residuals and complementarity gaps evolve at the same rate. The residuals are given by:

$$\boldsymbol{r}_d^{(k)} = (1-\gamma)(\boldsymbol{\nabla}_x f^{(k)} + \boldsymbol{A}^T\boldsymbol{y}^{(k)} + (\boldsymbol{G}^{(k)})^T\boldsymbol{z}^{(k)} - \boldsymbol{s}^{(k)}) \tag{2.52a}$$
$$\boldsymbol{r}_p^{(k)} = (1-\gamma)(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}) \tag{2.52b}$$
$$\boldsymbol{r'}_p^{(k)} = (1-\gamma)\boldsymbol{g}(\boldsymbol{x}^{(k)}) \tag{2.52c}$$
$$\boldsymbol{r}_c^{(k)} = \eta^{(k)}\boldsymbol{e} - \boldsymbol{X}^{(k)}\boldsymbol{S}^{(k)}\boldsymbol{e} - \Delta\boldsymbol{X}^a\Delta\boldsymbol{S}^a\boldsymbol{e} \tag{2.52d}$$

The factor $(1-\gamma)$ has been added to the primal and dual residuals to ensure that each quantity is decreased at the same rate which is beneficial for the algorithm's efficiency [Bleyer, 2017]. We observe that if $\gamma = 1$, the primal and dual residual are not reduced and only the complementarity gap is improved such that the next iterate will be closer to the central path. Otherwise if $\gamma = 0$, the solution is almost the same as the affine step differing only by the quadratic correction term.

Further details about the choice of the parameters and the reduced residual expressions can be found in Andersen [2013] and Mehrotra [1992]. Note that the efficiency of these advanced aspects (adaptive barrier parameter, predictor-corrector schemes, etc.) and corresponding heuristics still remain open questions, especially in the SONLP case (see a discussion on that matter in Nocedal et al. [2009]). Fine tuning exists and should be tested to determine the best combination when devising the algorithm. However, in our work, such choices have been made once and for all, providing good results on our test cases, and should not be changed from the user's perspective.

2.5.5.2 Symmetric scaling and reduced linear system

Most IPM implementations further reduce the system size by eliminating the conic variables. A specific scaling procedure on the primal-dual pair $(\boldsymbol{x}, \boldsymbol{s})$ is usually adopted to yield a symmetric reduced system and to improve the algorithm convergence. In the following, we use the Nesterov-Todd scaling and search direction which has been described at length in Nesterov and Todd [1998] and in Andersen et al. [2003] which allows us to rewrite the linearized complementarity as follows:

$$(\boldsymbol{x} \circ \boldsymbol{s})^{(k+1)} \simeq (\boldsymbol{x} \circ \boldsymbol{s})^{(k)} + \boldsymbol{V}\boldsymbol{F}^{-T}\Delta\boldsymbol{s} + \boldsymbol{V}\boldsymbol{F}\Delta\boldsymbol{x} \tag{2.53}$$

The expressions of the scaling matrices and details of their calculations can be found in appendix B. The linearized KKT system (2.43) can then be efficiently reduced using (2.53):

$$\begin{bmatrix} \boldsymbol{H}^{(k)} & \boldsymbol{A}^T & \boldsymbol{G}^{T(k)} & -\boldsymbol{I} \\ \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{G}^{(k)} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ (\boldsymbol{V}\boldsymbol{F})^{(k)} & \boldsymbol{0} & \boldsymbol{0} & (\boldsymbol{V}\boldsymbol{F}^{-T})^{(k)} \end{bmatrix} \begin{Bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{y} \\ \Delta\boldsymbol{z} \\ \Delta\boldsymbol{s} \end{Bmatrix} = \begin{Bmatrix} -r_d^{(k)} \\ -r_p^{(k)} \\ -r_p'^{(k)} \\ -r_c^{(k)} \end{Bmatrix} \tag{2.54}$$

The next step is to eliminate all conic slack variables $\Delta\boldsymbol{s}$ using static condensation:

$$\Delta\boldsymbol{s} = -(\boldsymbol{F}\boldsymbol{V}^{-1})^{(k)} r_c^{(k)} - (\boldsymbol{F}^T\boldsymbol{F})^{(k)}\Delta\boldsymbol{x} \tag{2.55}$$

one finally obtains the following reduced KKT system combining only the primal variables and the Lagrange multipliers for the equality constraints:

$$\boldsymbol{J}^{(k)}\Delta\boldsymbol{w} = \begin{bmatrix} \boldsymbol{H}^{(k)} + (\boldsymbol{F}^T\boldsymbol{F})^{(k)} & \boldsymbol{A}^T & \boldsymbol{G}^{T(k)} \\ \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{G}^{(k)} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{Bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{y} \\ \Delta\boldsymbol{z} \end{Bmatrix} = \begin{Bmatrix} -r_d'^{(k)} \\ -r_p^{(k)} \\ -r_p'^{(k)} \end{Bmatrix} \tag{2.56}$$

where:

$$r'^{(k)}_d = \nabla_x f^{(k)} + \boldsymbol{A}^T \boldsymbol{y}^{(k)} + \boldsymbol{G}^{T(k)} \boldsymbol{z}^{(k)} - \boldsymbol{s}^{(k)} - (\boldsymbol{F}\boldsymbol{V}^{-1})^{(k)} r_c^{(k)} \tag{2.57a}$$

$$r_p^{(k)} = \boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} \tag{2.57b}$$

$$r'^{(k)}_p = \boldsymbol{g}(\boldsymbol{x}^{(k)}) \tag{2.57c}$$

In practice, additional static condensations of some groups of primal or dual variables can be done in order to further reduce the size of the linear system. However, these static condensations are of algorithmic and implementation nature and do not have any physical of mathematical meaning.

In the case of problem (2.5), this system is very similar the tangent elastoplastic stiffness matrix obtained during a classical Newton-Raphson algorithm. In fact, after the return mapping is performed to determine the local stress state and verify the plasticity criterion, the elastoplastic tangent modulus is calculated and then used to reassemble the new tangent stiffness matrix. In the case of the IPM, the stiffness matrix (or the Hessian matrix $\boldsymbol{H}^{(k)}$) stays constant in this case, and represents the *elastic* behavior. The local *plastic* corrections to the stiffness matrix are introduced through the scaling matrix $\boldsymbol{F}^{(k)}$ which depends on the conic condition, i.e. the plastic state.

2.5.6 Summary of the final algorithm

As explained earlier, most industrial IP software like `Mosek` or `SeDuMi` are very well adapted to solving regular SOCP in a black-box manner. These software use a specific standard SOCP formalism (such as problem (2.8)) to cover a wider set of problems. Also, they require that the quadratic form in the objective function be formulated as SOCP constraints, which can be obtained by introducing additional optimization variables [Andersen, 2013; Mosek, 2013].

Since the knowledge of the mechanical problem and the specific significance of each variable can be exploited in order to maximize the computational effort, the use of black-box solvers did not seem adapted in our case. Moreover NLP algorithms are not necessarily adapted to our case of studies and coding all the interfaces between the software was proven to be extremely time consuming. Therefore a new algorithm was coded from scratch in order to have full control over the solving method: a full primal-dual interior point algorithm adapted to handling both classical SOCP and more complex SONLP, has been implemented within the software developed at *Strains* and is summarized in Alg. 1.

For instance, in the case of regular SOCP, our implementation enables to keep the quadratic form, changing only slightly some aspects of the Newton-like system to be solved in a standard IPM implementation. In this SOCP case, the Hessian matrix $\boldsymbol{H}^{(k)}$ is constant throughout the solving process and is equal to the quadratic $\boldsymbol{Q}_0$ in the objective function. It enables to reduce the number of auxiliary variables needed to formulate the problem in a standard form. It also paves the way to tackling more complex models by taking directly into account the smooth non-linear part in the objective problem, which was not possible in the commercial software used previously.

Algorithm 1 Predictor-corrector primal-dual interior point algorithm

initialization ($k = 0$) ▷ Sec. 2.5.4.2
while $\|\boldsymbol{r'}_d^{(k)}, \boldsymbol{r}_p^{(k)}, \boldsymbol{r'}_p^{(k)}\| \geq \epsilon$ and $\bar{\mathrm{g}}^{(k)} \geq \epsilon$ **do**
  compute scaling matrices $\boldsymbol{F}$ and $\boldsymbol{V}$ ▷ Appendix B
  form reduced Newton system $\boldsymbol{J}^{(k)}$ ▷ Sec. 2.5.5.2
  form affine residual vector $\boldsymbol{r}^a = \boldsymbol{r}^{(k)}(\boldsymbol{w}, \eta = 0)$ ▷ Sec. 2.5.5.1
  solve affine step $\boldsymbol{J}^{(k)}.\Delta\boldsymbol{w}^a = -\boldsymbol{r}^a$ and calculate the affine solution $\Delta\boldsymbol{w}^a$
  compute maximum affine step length α_{max}^a ▷ Sec. 2.5.4.1
  estimate centering parameter γ using Eq. (2.51)
  set new barrier parameter $\eta^{(k)} = \gamma\bar{\mathrm{g}}^{(k)}$ ▷ Sec. 2.5.4.3
  form corrector residual vector $\boldsymbol{r}_c^{(k)}$ using Eq. (2.45d) and reduced residuals
  solve corrector step $\boldsymbol{J}^{(k)}.\Delta\boldsymbol{w} = -\boldsymbol{r}^{(k)}$ by reusing the factorization of $\boldsymbol{J}^{(k)}$
  compute maximum step length α ▷ Sec. 2.5.4.1
  update variables $\boldsymbol{w}^{(k+1)} = \boldsymbol{w}^{(k)} + \alpha\Delta\boldsymbol{w}$
  compute new complementarity gap $\bar{\mathrm{g}}^{(k+1)}$
  if $k > N_{iter,max}$ or $\alpha \leq \alpha_{min}$ **then**
    break ▷ Algorithm failed
  end if
  $k \leftarrow k + 1$
end while

2.6 Solving the illustrative example

In order to illustrate how the algorithm works, we will now solve problem (2.1). We will make different initialization choices to illustrate different behaviors and assess its sensitivity to warm-start, i.e. being close to the problem solution. Note however that, due to the problem simplicity and low number of variables, the obtained performances might not necessarily reflect what will be observed in practice on real test cases.

The chosen different initialization values are reported in Tab. 2.1, some of them are arbitrary and other use mechanical intuition to improve the convergence speed. All the remaining variables not included in Tab. 2.1 are set to zero. When running the algorithm, all the tests converge to the same values while respecting a relative tolerance of 10^{-8} on all residuals, however, tests 1 and 6 take 12 iterations while tests 2,3,5 and 7 take 4 iterations and test 4 takes 5 iterations as seen in Tab. 2.1. This is mainly due the conic variables arbitrary initialization in tests 1 and 6. Mathematically speaking, the value $\gamma^{(0)} = s_0^{(0)} = 1.0$ ensures a feasible conic starting point, however it is clear that, from a mechanical point of view, the equivalent plastic strain will be far less than the initial value and same goes for the equivalent von Mises Stress. The simple improvement in the coherence of the starting values of the cones allows us to reduce the number of iterations needed by a factor of 2.

Note that the initial value for the primal variables and Lagrange multipliers have

Table 2.1 – Initialization values and number of iterations for convergence

Test	$\varepsilon_{11}^{(0)}$	$\varepsilon_{33}^{(0)}$	$\sigma_{11}^{(0)}$ [MPa]	$\sigma_{22}^{(0)}$ [MPa]	$\gamma^{(0)}$	$s_0^{(0)}$ [MPa]	Remarks	N_{iter}
1	0	0	0	0	1.0	1.0	Similar to Sec. 2.5.4.2	12
2	0	0	0	0	0.001	1000	Complementarity gap equal to 1	4
3	0	0	0	0	0.001	355	Mechanical insight	4
4	0	0	-355	-355	0.001	355	Robustness test	5
5	-0.0025	0.0025	0	0	0.001	355	Robustness test	4
6	-0.0025	0.0025	-355	-355	1.0	1.0	Robustness test	5
7	0.0025	-0.00145	405	147	0.001	355	Close to exact solution guess	4

a small influence of the algorithm's number of iterations as can be seen in test 2 to 5. This is one of the main advantages of the primal-dual interior point method where all residuals are reduced at the same rate at each iteration. Therefore, initializing the algorithm at the exact primal and dual solution (if known) will not drastically change the number of iterations since the complementarity gap will still need to be reduced to reach the needed tolerance. It is clear that any warm-start strategy should take into account the mechanical aspect of the problem which brings light to yet another advantage of coding a full algorithm from scratch. Most commercial IP solvers do not allow warm strategies as this subject is still a open research question, however, using the mechanical insight, we can derive simple yet very effective strategies in order to cut computational costs.

Fig. 2.2 to 2.8 show an illustration of different iterates of the primal variables and Lagrange multipliers for the 7 different tests. One can see that sometimes large steps are taken. These are possible when we are far from the conic boundaries which is one of the main advantages of the IPM. The influence of the choice of the initial value of the conic variables can be seen in Fig. 2.2 and 2.7 where the algorithms uses the first several iterations to « find the central path » (see the go back and forth in the representations of $\boldsymbol{\varepsilon}^p$) and then converges in 5 iterations to the optimal value.

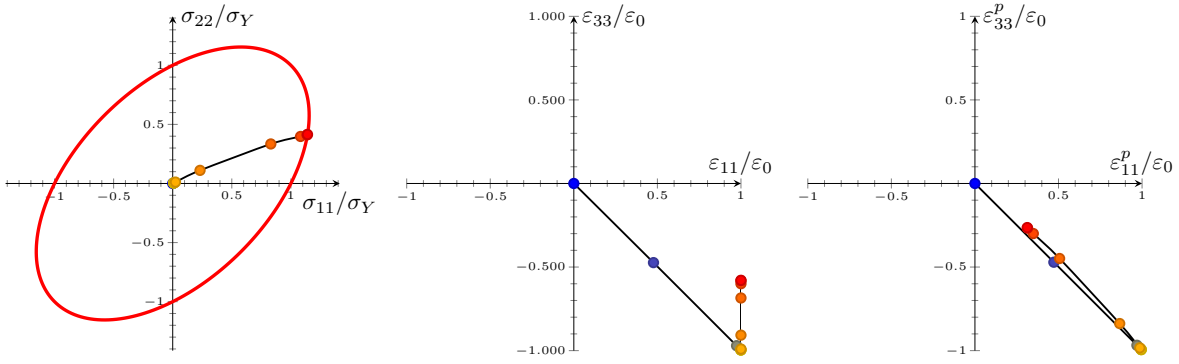


Figure 2.2 – Test 1 - Visualization of the different iterates

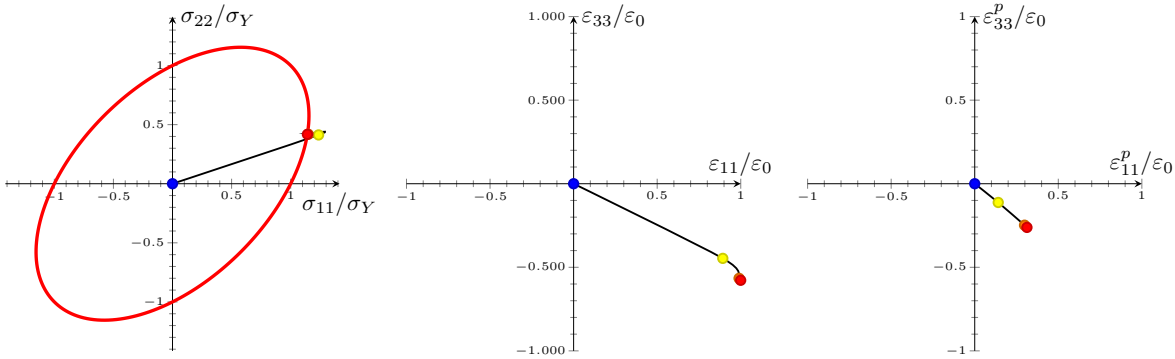


Figure 2.3 – Test 2 - Visualization of the different iterates

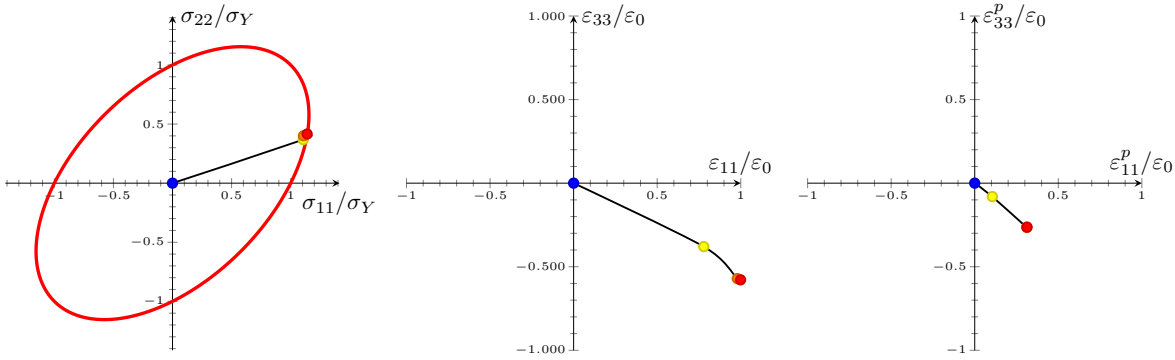


Figure 2.4 – Test 3 - Visualization of the different iterates

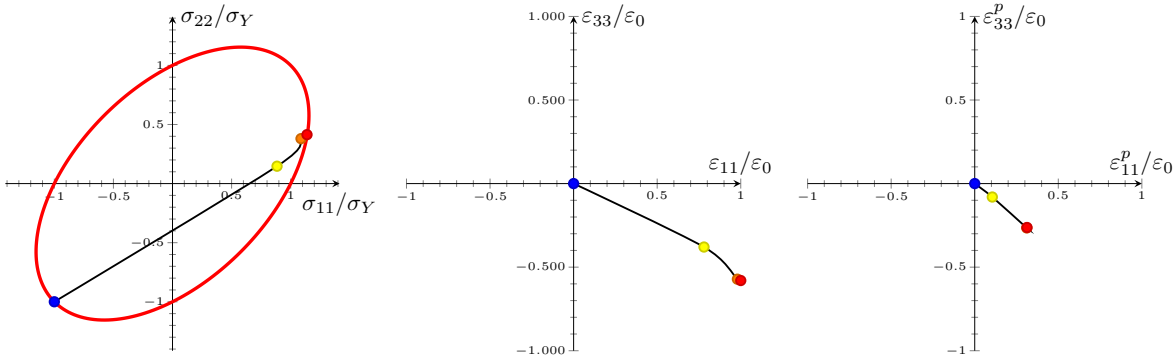


Figure 2.5 – Test 4 - Visualization of the different iterates

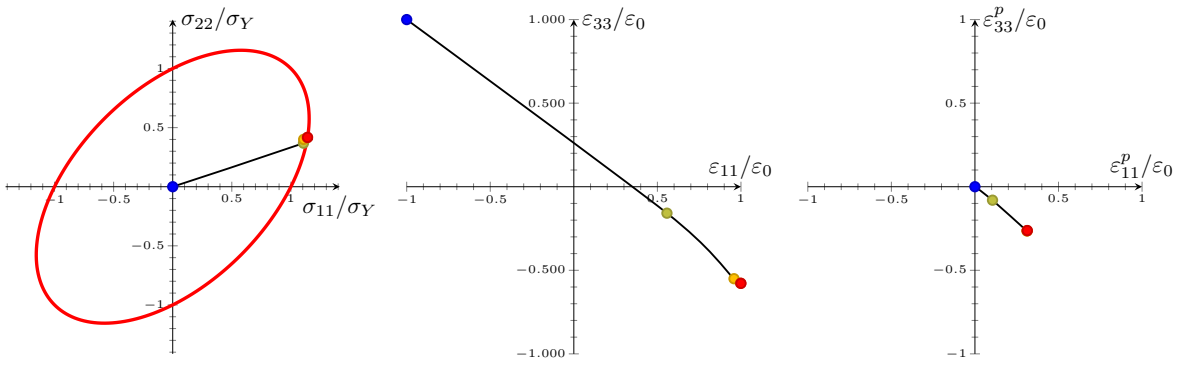


Figure 2.6 – Test 5 - Visualization of the different iterates

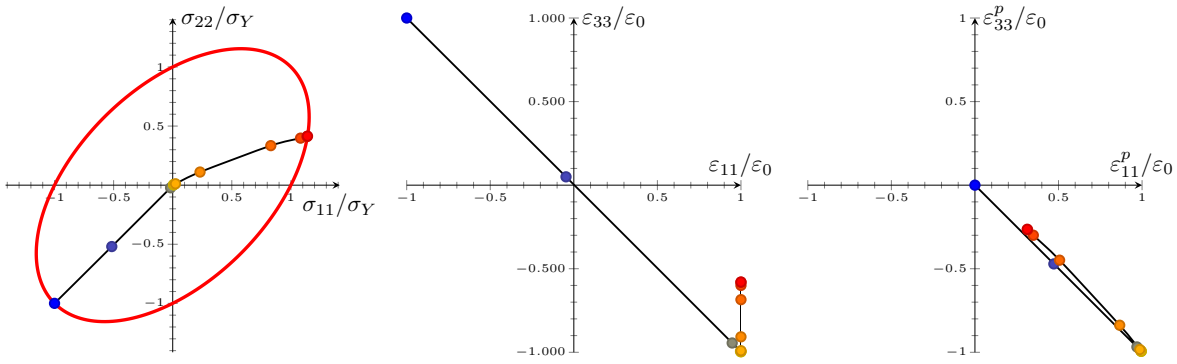


Figure 2.7 – Test 6 - Visualization of the different iterates

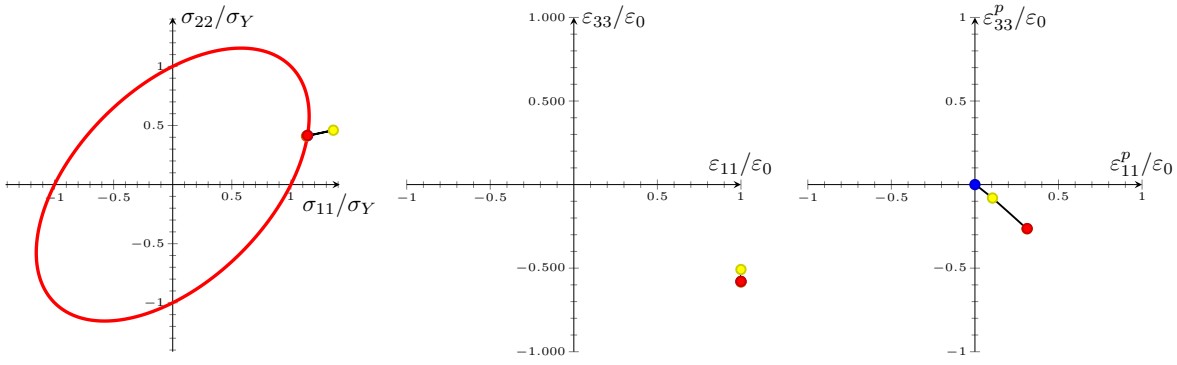


Figure 2.8 – Test 7 - Visualization of the different iterates

2.7 Conclusions

IPM are nowadays becoming an efficient and reliable technology for solving a wide range of optimization programs and specifically convex ones with a specific conic structure. As it will be seen the following chapters, this algorithm will provide a robust solving scheme for mechanical problems compared to other traditional approaches such as the Newton-Raphson or the augmented Lagrangian method. Its framework allows us to easily include various mechanical phenomena such as contact conditions, plasticity or finite-strain effects using the same formalism given by second-order cones.

Many commercial and open software are available and are capable of handling large-scale problems. However, as already expressed in this chapter and the subsequent ones, the main advantage out of coding from scratch an IP algorithm is an increased flexibility and the possibility to account for the known mechanical problem structure. This allows us to surpass many disadvantages of black-box solvers such as the reformulation of quadratic forms into conic constraints which proves to be unnatural in our class of problems and the use of warm start strategies in order to increase the convergence speed. It also allows us to simply include new mechanical behaviors using the same mathematical formalism without having to re-adapt the solving algorithm.

* *
*

Chapter 3

Dual finite-element analysis using second-order cone programming for structures including contact

Abstract: *In this chapter, the computation of elastic structures with contact conditions is performed by means of a dual analysis combining displacement-based and equilibrium-based finite elements and using the SOCP framework. The dual approach allows the user to assess the quality of convergence and to calculate a discretization error estimator which includes a contact error term. An efficient remeshing scheme, based on the local contributions of the elements to the global error, can then be used to efficiently improve the accuracy of the solution.*

Contents

3.1 Introduction

Non-linear finite element calculations are far from being used on a daily basis by structural engineers, especially when the non-linear behavior is non-smooth such as the case of plasticity or contact as seen in chapter 1. Though the finite-element method is a well established technology for simple 3D elastic continuum, the treatment of non-smooth conditions remains a problematic matter for inexperienced users

Contact between mechanical parts exists in almost all engineering fields, specifically in steel structures analysis where the verification of steel assemblies relies on determining the forces and displacements in the assembly which, by definition, joins distinct parts of the structure, transferring forces through plates, welds and bolts that mainly rely on contact. The non-smoothness of contact conditions is characterized by a pair of kinematic and static variables, the displacement gap and the reaction forces over the contact surface, which are linked together by a set of complementarity conditions. Such finite element implementations are quite difficult to solve and the introduction of additional non-linearities such as material plasticity for instance will only worsen the convergence of these models, these aspects being at the origin of the difficulties for engineers to check complex assemblies.

Throughout various project experiences at *Strains*, we were able to asses that such situations present a major difficulty for structural engineers as the modeling of a complex steel assembly requires a great amount of computer-aided design time, whereas convergence of numerical solvers may often be challenging to achieve for inexperienced users. We propose an approach for computing structures with contact, such as 3D steel assemblies, by using a novel solution algorithm which exhibits very robust convergence properties and does not require any fine tuning algorithmic parameters while still being very competitive in terms of computational costs compared to classical approaches implemented in commercial software.

In the present chapter, we will consider the case of elastic materials and an associated Coulomb friction law for modeling the contact behavior. In such a situation, the complementarity problem becomes a convex second-order cone problem for which very efficient interior-point algorithms have been developed by the mathematical programming community [Andersen et al., 2003; Boyd and Vandenberghe, 2004; Lobo et al., 1998; Wright, 2005]. It is interesting to note that interior-point methods (IPM) are becoming a state-of-the art numerical procedure for other non-smooth mechanical problems such as limit load computations in limit analysis [Lyamin and Sloan, 2002a,b; Makrodimopoulos and Martin, 2006, 2007] or computation of viscoplastic fluid flows for which conventional methods such as augmented Lagrangian (AL) approaches have quite slow convergence rates [Bleyer, 2017; Bleyer et al., 2015]. Only few attempts were made to solve contact problems using IPMs such as Krabbenhoft et al. [2012b] or Kleinert et al. [2014] for modeling granular contact dynamics. While earlier attempts did not use the most efficient forms of IPM methods [Christensen et al., 1998; Tanoh et al., 2004], more recent implementations showed promising results in non-smooth contact dynamics [Mangoni et al., 2018] or isogeometric contexts [Temizer et al., 2014].

As regards our choice of considering only associated friction for the present work, we point out that IPM rely on a convex conic optimization setting which does not hold in the presence of classical Coulomb friction. However, since quasi-static computations are considered here, we expect that the obtained solutions will be quite similar to a classical Coulomb friction computation, as it will be later shown in the example section. Moreover, some iterative strategies have already been proposed to tackle the case of non-associative plasticity or friction using a series of associated computations [Gilbert et al., 2006; Krabbenhoft et al., 2012a]. The associated formulation of the IPM problems could serve as a basis of a future extension to non-associated behaviors.

We aim at further establishing the potential of IPM for contact of complex 3D assemblies, especially concerning its convergence robustness and low iteration count in regards of its use by inexperienced engineer. The proposed methodology can handle large-scale problems and complex assemblies in an efficient manner with a minimum user intervention. We will also combine it with dual stress and displacement-based computations to assess the solution quality with respect to discretization errors.

3.2 Reference problem and governing equations

3.2.1 Contact constitutive equations

Let us consider in this section the formulation of contact laws on the contact boundary denoted by Γ_c described by a unit normal vector $\boldsymbol{n}$ pointing from the solid Ω_1 to the solid Ω_2 and two tangential vectors $\boldsymbol{t}_1, \boldsymbol{t}_2$ (see Fig. 3.1). We will also use the following notation $\boldsymbol{x} = x_N \boldsymbol{n} + \boldsymbol{x}_T$ with $\boldsymbol{x}_T$ the tangential component of any vector $\boldsymbol{x}$.

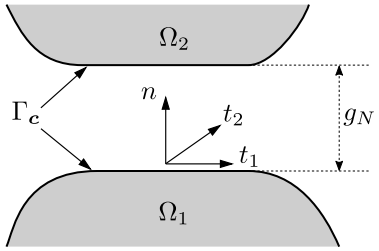


Figure 3.1 – Local coordinate system

In the case where the two deformable bodies Ω_1 and Ω_2 come in contact, kinematic variables will involve the displacement discontinuity vector $[\![\boldsymbol{u}]\!] = \boldsymbol{u}_2 - \boldsymbol{u}_1$, while the stress vector $\boldsymbol{T} = \sigma_N \boldsymbol{n} + \boldsymbol{\sigma}_T$ is expressed on one of the two faces since equilibrium conditions ensures equality between the two surface efforts on both sides of Γ_c. These local variables can be found using the following relations:

$$\boldsymbol{u} = \begin{pmatrix} u_N \\ u_{T_1} \\ u_{T_2} \end{pmatrix} = \begin{pmatrix} u_N \\ \boldsymbol{u}_T \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}.\boldsymbol{n} \\ \boldsymbol{u}.\boldsymbol{t} \end{pmatrix} \tag{3.1}$$

$$\boldsymbol{T} = \boldsymbol{\sigma}.\boldsymbol{n} = \begin{pmatrix} \sigma_N \\ \sigma_{T_1} \\ \sigma_{T_2} \end{pmatrix} = \begin{pmatrix} \sigma_N \\ \boldsymbol{\sigma}_T \end{pmatrix} = \begin{pmatrix} \sigma_{ij} n_i n_j \\ (\sigma_T)_i = \sigma_{ij} n_j - \sigma_N n_i \end{pmatrix} \tag{3.2}$$

Since the faces are not necessarily initially in contact, an initial gap vector $\boldsymbol{g}_0 = g_0 \boldsymbol{n}$ is introduced. We suppose that the initial gap have no tangential component while its normal component is positive $g_0 \geq 0$. This choice is appropriate considering that we will restrict to infinitesimal transformations and since contact will be implemented in a node-to-node fashion. As a consequence, meshes of both contacting surfaces will be conforming. The current gap vector is then defined by:

$$\boldsymbol{g}_{(\boldsymbol{u})} = \begin{pmatrix} g_N \\ \boldsymbol{g}_T \end{pmatrix} = \begin{pmatrix} g_0 + [\![u_N]\!] \\ [\![\boldsymbol{u}_T]\!] \end{pmatrix} \tag{3.3}$$

3.2.1.1 Unilateral contact

Unilateral contact is enforced using Signorini's conditions which express the non-penetration and the non-apparition of tensile surface tractions between contact surfaces

$$g_N \geq 0 \tag{3.4a}$$

$$\begin{cases} \sigma_N \leq 0 & \text{if} \quad g = 0 \\ \sigma_N = 0 & \text{if} \quad g \geq 0 \end{cases} \tag{3.4b}$$

Unilateral contact conditions are therefore one of the simplest and most common non-smooth constraints encountered in mechanics. One of the way of treating these constrains is to write it as a set of equations including complementarity condition thus involving no « if-clause ». The contact conditions are then equivalent to:

$$\begin{cases} g_N \geq 0 & \text{(3.5a)} \\ \sigma_N \leq 0 & \text{(3.5b)} \\ g_N.\sigma_N = 0 & \text{(3.5c)} \end{cases}$$

$g_N.\sigma_N = 0$ being the complementarity between normal stresses and the normal gap.

3.2.1.2 Frictional behavior

Coulomb frictional model is used to model tangential behavior of contact surfaces by means of the friction coefficient μ. One distinguishes between stick and slip conditions (expressed here in terms of a quasistatic displacement increment):

$$\begin{cases} \text{if } \left\|\boldsymbol{\sigma}_T\right\| < -\mu\sigma_N \text{ then } \boldsymbol{g}_T = 0 & \text{(3.6a)} \\ \text{if } \left\|\boldsymbol{\sigma}_T\right\| = -\mu\sigma_N \text{ then } \boldsymbol{g}_T = \lambda \dfrac{\boldsymbol{\sigma}_T}{\left\|\boldsymbol{\sigma}_T\right\|} \text{ with } \lambda \geq 0 & \text{(3.6b)} \end{cases}$$

This law shows the dependency over a frictional yield value $\mu\sigma_N$ and expresses the tangential slip as proportional to frictional stresses.

3.2.1.3 Contact complementarity conditions

When associated contact is assumed, the unilateral and frictional laws are both replaced by the following combined law:

$$\begin{cases} \text{if } \left\|\boldsymbol{\sigma}_T\right\| < -\mu\sigma_N \text{ then } \boldsymbol{g} = \boldsymbol{0} & \text{(3.7a)} \\ \text{if } \left\|\boldsymbol{\sigma}_T\right\| = -\mu\sigma_N \text{ then } \boldsymbol{g}_T = \lambda \dfrac{\boldsymbol{\sigma}_T}{\left\|\boldsymbol{\sigma}_T\right\|} \text{ with } \lambda = \dfrac{g_N}{\mu} \geq 0 & \text{(3.7b)} \end{cases}$$

which ensures that $\sigma_N \leq 0$, $g_N \geq 0$, but also couples the normal and tangential gaps. The previous condition can be expressed as the following two second-order cones (see Fig. 3.2):

$$(\sigma_N, \boldsymbol{\sigma}_T) \in \mathcal{K}_\mu \tag{3.8}$$

$$(g_N, \boldsymbol{g}_T) \in \mathcal{K}_{-1/\mu} \tag{3.9}$$

$$\text{where } \mathcal{K}_\alpha = \left\{(x, \boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^2 \;\middle|\; \left\|\boldsymbol{y}\right\| + \alpha x \leq 0\right\} \tag{3.10}$$

along with the following complementarity condition:

$$\sigma_N g_N + \boldsymbol{\sigma}_T \cdot \boldsymbol{g}_T = 0 \tag{3.11}$$

In the above, it can be observed that the cone $\mathcal{K}_{-1/\mu} = (\mathcal{K}_\mu)^\circ = \mathcal{K}_\mu^\circ$ where $\star^\circ$ refers to the polar cone (cf. definition A.5). The pair $(\sigma_N, \boldsymbol{\sigma}_T)$ and $(g_N, \boldsymbol{g}_T)$ therefore satisfies a complementarity condition over dual cones (more details can be found in appendix C). This remark is at the origin of the formulation of contact problems as complementarity problems in the general (non-associated) case and second-order cone programming in the present associated case. The use of associated friction signifies that slip only occurs when there is a normal opening. This aspect is indeed less realistic than a non-associated framework. However, the use of the dual approach along with the numerical robustness of this convex alternative makes the use of associated friction very interesting. More details on the non-associated case can be found in de Saxcé and Feng [1998] and in Kanno et al. [2006].

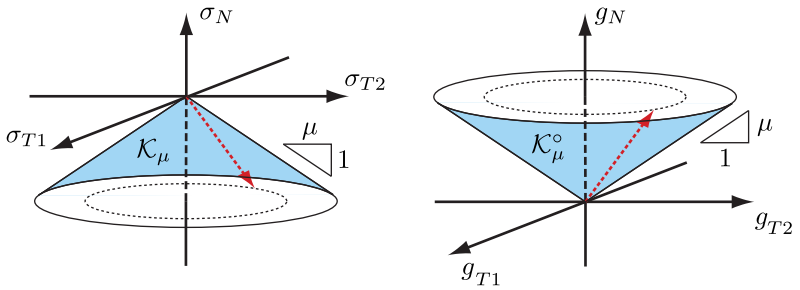


Figure 3.2 – Geometry of the friction cones

3.2.2 Local governing equations

In the following, $\boldsymbol{u}$ denotes the displacement field, $\boldsymbol{\varepsilon}_{(\boldsymbol{u})} = \boldsymbol{\nabla}^S \boldsymbol{u}$ the linearized strain tensor, $\boldsymbol{\sigma}$ the Cauchy stress tensor and $\mathbb{D}$ the elastic stiffness tensor. Imposed displacements and surface tractions will be respectively denoted by $\boldsymbol{u}_d$ and $\boldsymbol{t}_d$, and the body forces by $\boldsymbol{b}$. The normal vector $\boldsymbol{n}$ pointing outwards will be used to orient free faces and surfaces. For contact conditions, the direction of the normal vector will be specified when needed.

Let $\Omega = \Omega_1 \cup \Omega_2$ be the total space occupied by the solids sharing a potential contact surface $\Gamma_{\boldsymbol{c}}$ and let be $\Sigma_{\boldsymbol{\sigma}}$ an internal surface of potential stress discontinuities (see Fig. 3.3). The external boundary will be split into three distinct parts such as: $\Gamma = \Gamma_{\boldsymbol{u}} \cup \Gamma_{\boldsymbol{t}} \cup \Gamma_{\boldsymbol{c}}$ and $\Gamma_{\boldsymbol{u}} \cap \Gamma_{\boldsymbol{t}} \cap \Gamma_{\boldsymbol{c}} = \emptyset$ where:

- displacements $\boldsymbol{u}$ are imposed to $\boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}} = \Gamma_{1,\boldsymbol{u}} \cup \Gamma_{2,\boldsymbol{u}}$,
- surface tractions $\boldsymbol{t} = \boldsymbol{\sigma} \cdot \boldsymbol{n}$ are imposed to $\boldsymbol{t}_d$ on $\Gamma_{\boldsymbol{t}} = \Gamma_{1,\boldsymbol{t}} \cup \Gamma_{2,\boldsymbol{t}}$, and
- contact conditions between Ω_1 and Ω_2 are imposed on $\Gamma_{\boldsymbol{c}}$. Generally the contact surface is unknown but in the scope of this chapter and the small displacements hypothesis, it is considered that the contact surface is known, predefined and does not change.

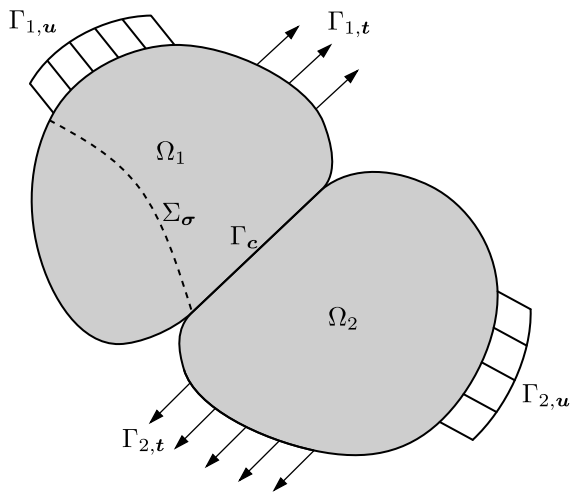


Figure 3.3 – Reference model

The general governing equations for an elastic body are the following:

Equilibrium and static boundary conditions

$$\mathrm{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 \quad \text{in } \Omega \tag{3.12}$$

$$[\![\boldsymbol{\sigma}]\!].\boldsymbol{n} = 0 \quad \text{on } \Sigma_{\boldsymbol{\sigma}} \tag{3.13}$$

$$\boldsymbol{\sigma}.\boldsymbol{n} = \boldsymbol{t}_d \quad \text{on } \Gamma_{\boldsymbol{t}} \tag{3.14}$$

Strain-displacement compatibility and kinematic boundary conditions

$$\boldsymbol{\varepsilon}_{(\boldsymbol{u})} = \boldsymbol{\nabla}^S \boldsymbol{u} \quad \text{in } \Omega \tag{3.15}$$

$$\boldsymbol{u} = \boldsymbol{u}_d \quad \text{on } \Gamma_{\boldsymbol{u}} \tag{3.16}$$

Elastic constitutive law

$$\boldsymbol{\sigma} = \mathbb{D} : \boldsymbol{\varepsilon}_{(\boldsymbol{u})} \quad \text{in } \Omega \tag{3.17}$$

Contact constitutive laws between $\boldsymbol{\sigma}.\boldsymbol{n}$ and $\boldsymbol{g}_{(\boldsymbol{u})}$ on $\Gamma_{\boldsymbol{c}}$ as discussed in Sec. 3.2.1.

3.2.2.1 Statically admissible stress field

A stress field $\boldsymbol{\sigma} \in \mathbb{S}$ belonging to the set $\mathbb{S}$ of symmetric second-rank tensors is said to be statically admissible (SA) with the body forces $\boldsymbol{b}$ in Ω and the imposed surfaces tractions $\boldsymbol{t}_d$ on $\Gamma_{\boldsymbol{t}}$ if it satisfies the equilibrium equations with these data:

$$SA = \left\{ \boldsymbol{\sigma} \in \mathbb{S} \;\middle|\; \begin{array}{ll} \mathrm{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 & \text{in } \Omega \\ [\![\boldsymbol{\sigma}]\!].\boldsymbol{n} = 0 & \text{through } \Sigma_{\boldsymbol{\sigma}} \\ \boldsymbol{\sigma}.\boldsymbol{n} = \boldsymbol{t}_d & \text{on } \Gamma_{\boldsymbol{t}} \end{array} \right\} \tag{3.18}$$

3.2.2.2 Kinematically admissible displacement field

A displacement field $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ is said to be kinematically admissible (KA) with the imposed displacements $\boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}}$ if it is piece-wise continuous and continuously differentiable and such that $\boldsymbol{u} = \boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}}$:

$$KA = \left\{ \boldsymbol{u} \in \mathbb{R}^3 \;\middle|\; \boldsymbol{u} = \boldsymbol{u}_d \quad \text{on } \Gamma_{\boldsymbol{u}} \right\} \tag{3.19}$$

3.2.2.3 Classical variational principles of elastostatics

Finite-element methods in structural and solid mechanics can be formulated using variational principles. This means that there exist a scalar integral functional of the chosen variable fields, such that finding a solution that verifies the strong equation forms given previously, implies that the functional is stationary. In this section, we introduce the classical variational principles of 3D elastic media in order to extend it hereafter to include contact conditions. The general 3 fields variational principle was presented simultaneously and independently by Fraeijs de Veubeke [1951, 1965], by Hai-Chang Hu [Hai-Chang, 1954] and by Washizu [Washizu, 1968]. The functional Π_{HW} takes the displacements, the strains and the stresses as 3 independent fields with no imposed constraints.

$$\begin{aligned} \Pi_{HW}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega (\boldsymbol{\sigma} : (\boldsymbol{\nabla}^S \boldsymbol{u} - \boldsymbol{\varepsilon}) + \psi(\boldsymbol{\varepsilon}) - \boldsymbol{b}\boldsymbol{u}) d\Omega \\ - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{u} - \boldsymbol{u}_d)(\boldsymbol{\sigma}\boldsymbol{n}) d\Gamma - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d \boldsymbol{u} d\Gamma \end{aligned} \tag{3.20}$$

with $\psi(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{D} : \boldsymbol{\varepsilon}$ being the energy density of elastic deformation. This principle takes into account all the conditions of Sec. 3.2.2. Some parts may be seen as Lagrange

multipliers for these conditions which link this general principles to the 3 other most known principles: the Hellinger-Reissner (HR) two-field principle $(\boldsymbol{u}, \boldsymbol{\sigma})$, and the total potential (TPE) and total complementary (TCPE) energies principles respectively involving the displacement field $(\boldsymbol{u})$ and the stress field $(\boldsymbol{\sigma})$. The two fields HR principle is obtained from the latter by imposing the and constitutive relation (3.17) as an essential condition thus giving the following functional

$$\Pi_{HR}(\boldsymbol{u}, \boldsymbol{\sigma}) = \int_{\Omega} (\boldsymbol{\sigma} : \boldsymbol{\nabla}^S \boldsymbol{u} + \psi^*(\boldsymbol{\sigma}) - \boldsymbol{b}\boldsymbol{u}) d\Omega - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{u} - \boldsymbol{u}_d)(\boldsymbol{\sigma}\boldsymbol{n}) d\Gamma - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d \boldsymbol{u} d\Gamma \quad (3.21)$$

with $\psi^*(\boldsymbol{\sigma})$ being the Fenchel transform of $\psi(\boldsymbol{\varepsilon})$ called the complementary energy density (see Def. A.12):

$$\psi^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\varepsilon}} \{\boldsymbol{\sigma} : \boldsymbol{\varepsilon} - \psi(\boldsymbol{\varepsilon})\} = \frac{1}{2} \boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma} \quad (3.22)$$

The TPE functional is obtained by then imposing the kinematic compatibility equations (3.15) as essential conditions, thus the single field displacement minimization problem is as follows:

$$\begin{array}{ll} \min\limits_{\boldsymbol{u}} & \Pi_p(\boldsymbol{u}) \\ \text{s.t.} & \boldsymbol{u} \in KA(\Omega) \\ & \boldsymbol{\varepsilon} = \boldsymbol{\nabla}^S \boldsymbol{u} \end{array} \quad (3.23)$$

with

$$\Pi_p(\boldsymbol{u}) = \int_{\Omega} \psi(\boldsymbol{\varepsilon})\, d\Omega - \int_{\Omega} \boldsymbol{b}.\boldsymbol{u}\, d\Omega - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d.\boldsymbol{u}\, d\Gamma \quad (3.24)$$

The TCPE functional is then obtained by imposing the equilibrium equations (3.12) and (3.13) as essential conditions, thus the single field stress minimization problem is as follows:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\sigma}} & \Pi_c(\boldsymbol{\sigma}) \\ \text{s.t.} & \boldsymbol{\sigma} \in SA(\Omega) \end{array} \quad (3.25)$$

with

$$\Pi_c(\boldsymbol{\sigma}) = \int_{\Omega} \psi^*(\boldsymbol{\sigma})\, d\Omega - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{u}_d\, d\Gamma \quad (3.26)$$

A comprehensible reference for the link between the different principles by using Lagrange multipliers can be found in Pian [1976].

3.3 Constitutive error estimator

As stated in Ladevèze and Pelle [2005], many constitutive relations can be defined using a pair of convex dual functions ψ and ψ^* such that:

$$\psi_{(\boldsymbol{\varepsilon})} + \psi^*_{(\boldsymbol{\sigma})} - \boldsymbol{\sigma} : \boldsymbol{\varepsilon} \geq 0 \quad (3.27)$$

This is no other than the Fenchel inequality for a pair of Fenchel conjugate functions (see again Def. A.12). At any point in Ω, the equality in the above equation is strictly equivalent to enforcing the constitutive relation at that point:

$$\psi_{(\boldsymbol{\varepsilon})} + \psi^*_{(\boldsymbol{\sigma})} - \boldsymbol{\sigma} : \boldsymbol{\varepsilon} = 0 \Leftrightarrow \boldsymbol{\sigma} = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} \text{ and } \boldsymbol{\varepsilon} = \frac{\partial \psi^*}{\partial \boldsymbol{\sigma}} \tag{3.28}$$

Using this concept, we define a constitutive relation error including the same volume term defined in Ladevèze and Pelle [2005] and adding a surface term expressed over the contact boundary $\Gamma_{\boldsymbol{c}}$ as in Debongnie [1983], Debongnie et al. [1995] and Coorevits et al. [2001]:

$$e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} = e^2_{\Omega(\boldsymbol{u},\boldsymbol{\sigma})} + e^2_{\Gamma_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{\sigma})} \tag{3.29}$$

$$\begin{aligned} e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} = &\int_\Omega \left(\psi_{(\boldsymbol{\varepsilon}_{(\boldsymbol{u})})} + \psi^*_{(\boldsymbol{\sigma})} - \boldsymbol{\sigma} : \boldsymbol{\varepsilon}_{(\boldsymbol{u})} \right) d\Omega \\ &+ \int_{\Gamma_{\boldsymbol{c}}} \left(\mathbb{1}_{\mathcal{K}^\circ_\mu(\boldsymbol{g}_{(\boldsymbol{u})})} + \mathbb{1}_{\mathcal{K}_\mu(\boldsymbol{\sigma}.\boldsymbol{n})} - (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{g}_{(\boldsymbol{u})} \right) d\Gamma \end{aligned} \tag{3.30}$$

with $\mathbb{1}_{\mathcal{K}^\circ_\mu(\boldsymbol{g})}$ and $\mathbb{1}_{\mathcal{K}_\mu(\boldsymbol{\sigma}.\boldsymbol{n})}$ being the indicator functions of the contact cones defined in Sec. 3.2.1.3 such as:

$$\mathbb{1}_{\mathcal{K}(\boldsymbol{x})} = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \mathcal{K} \\ +\infty & \text{otherwise} \end{cases} \tag{3.31}$$

Indeed, for the convex cone $\mathcal{K}_\mu$, we have that the conjugate of the indicator function of the cone if no other than the indicator function of its polar cone i.e. $(\mathbb{1}_{\mathcal{K}_\mu})^* = \mathbb{1}_{\mathcal{K}^\circ_\mu}$ and it can be easily shown that:

$$\mathbb{1}_{\mathcal{K}^\circ_\mu(\boldsymbol{g})} + \mathbb{1}_{\mathcal{K}_\mu(\boldsymbol{\sigma}.\boldsymbol{n})} - (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{g} = 0 \tag{3.32}$$

is equivalent to conditions (3.8),(3.9) and (3.11). The important property of this error functional is that it is always positive:

$$e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} \geq 0 \quad \forall (\boldsymbol{u}, \boldsymbol{\sigma}) \tag{3.33}$$

In particular, it will be infinite if the contact variables do not satisfy the contact constitutive equations.

For the purpose of the present chapter, we consider only elastic behavior in the bulk domain, $\psi_{(\boldsymbol{\varepsilon})}$ and $\psi^*_{(\boldsymbol{\sigma})}$ therefore coincide with Helmholtz's free energy and Gibbs's complementary energy in the case of linear elasticity:

$$\psi_{(\boldsymbol{\varepsilon})} = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{D} : \boldsymbol{\varepsilon} \quad \text{and} \quad \psi^*_{(\boldsymbol{\sigma})} = \frac{1}{2}\boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma} \tag{3.34}$$

It can be easily shown that the error can be expressed as follows for any $\boldsymbol{\sigma} \in SA$ and $\boldsymbol{u} \in KA$:

$$e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} = \Pi_{p(\boldsymbol{u})} + \Pi_{c(\boldsymbol{\sigma})} \tag{3.35}$$

with $\Pi_{p(\boldsymbol{u})}$ being the total elastic potential energy (TPE) functional including a surface contact term:

$$\Pi_{p(\boldsymbol{u})} = \int_{\Omega} \frac{1}{2}\boldsymbol{\varepsilon}_{(\boldsymbol{u})} : \mathbb{D} : \boldsymbol{\varepsilon}_{(\boldsymbol{u})}\, d\Omega + \int_{\Gamma_{\boldsymbol{c}}} \mathbb{1}_{\mathcal{K}_{\mu}^{\circ}(\boldsymbol{g}_{(\boldsymbol{u})})}\, d\Gamma - \int_{\Omega} \boldsymbol{b}.\boldsymbol{u}\, d\Omega - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d.\boldsymbol{u}\, d\Gamma \tag{3.36}$$

and $\Pi_{c(\boldsymbol{\sigma})}$ being the total elastic complementary energy (TCPE) functional also including a surface contact term:

$$\Pi_{c(\boldsymbol{\sigma})} = \int_{\Omega} \frac{1}{2}\boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma}\, d\Omega + \int_{\Gamma_{\boldsymbol{c}}} \mathbb{1}_{\mathcal{K}_{\mu}(\boldsymbol{\sigma}.\boldsymbol{n})}\, d\Gamma - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{u}_d\, d\Gamma - \int_{\Gamma_{\boldsymbol{c}}} (\boldsymbol{\sigma}.\boldsymbol{n}).(g_0\boldsymbol{n})\, d\Gamma \tag{3.37}$$

The introduced error also possesses the important property to be zero if and only if the constitutive equations are satisfied. As a result, the solution $(\boldsymbol{u}^*, \boldsymbol{\sigma}^*)$ can be characterized by:

$$(\boldsymbol{u}^*, \boldsymbol{\sigma}^*) \in KA \times SA \quad \left\{ \begin{array}{c} \boldsymbol{\sigma}^* = \mathbb{D} : \boldsymbol{\varepsilon}_{(\boldsymbol{u}^*)} \\ (\boldsymbol{g}_{(\boldsymbol{u}^*)}, \boldsymbol{\sigma}^*.\boldsymbol{n}) \in (\mathcal{K}_{\mu} \times \mathcal{K}_{\mu}^{\circ}) \end{array} \right. \quad \Leftrightarrow \; e^2_{(\boldsymbol{u}^*, \boldsymbol{\sigma}^*)} = 0 \tag{3.38}$$

Owing to the introduced assumptions, a solution to the previous problem always exists. It has been demonstrated in Debongnie [1983], Debongnie et al. [1995] and in Coorevits et al. [2001] that the error estimator described above always gives an upper bound for the real error i.e. let $(\boldsymbol{u}^*, \boldsymbol{\sigma}^*)$ be a solution to the contact problem, and $(\boldsymbol{u}_h, \boldsymbol{\sigma}_h) \in KA \times SA$ be an admissible approximation, then:

$$\begin{aligned} \frac{1}{2}\left\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}^*\right\|^2_{\mathbb{D}^{-1},\Omega} + \frac{1}{2}\left\|\boldsymbol{\varepsilon}_{(u_h)} - \boldsymbol{\varepsilon}_{(u^*)}\right\|^2_{\mathbb{D},\Omega} \\ - \mu \int_{\Gamma_{\boldsymbol{c}}} (\boldsymbol{\sigma}_h.\boldsymbol{n} - \boldsymbol{\sigma}^*.\boldsymbol{n})(\boldsymbol{g}_{(\boldsymbol{u}_h)} - \boldsymbol{g}_{(\boldsymbol{u}^*)}) d\Gamma \leq e^2_{(\boldsymbol{u}_h, \boldsymbol{\sigma}_h)} \end{aligned} \tag{3.39}$$

with $\left\|\boldsymbol{\varepsilon}\right\|_{\mathbb{D},\Omega}$ and $\left\|\boldsymbol{\sigma}\right\|_{\mathbb{D}^{-1},\Omega}$ being the traditional energy norms. Further information about posteriori error estimation for contact problems can be found in Louf et al. [2003], Ben Belgacem et al. [2012] and Ladevèze and Chamoin [2016].

3.4 Primal and dual variational principles

The following variational principles can be derived using the constitutive error estimator presented in Sec. 3.3:

$$e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} = e^2_{\Omega(\boldsymbol{u},\boldsymbol{\sigma})} + e^2_{\Gamma_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{\sigma})} = \Pi_{p(\boldsymbol{u})} + \Pi_{c(\boldsymbol{\sigma})} \tag{3.40}$$

Finding an approximate solution consists of computing an optimal pair $(\boldsymbol{u}_h, \boldsymbol{\sigma}_h)$ of displacements and stress fields such that the estimated error, and thus the real error, are minimal over subsets of the space of K.A. displacements and S.A. stress fields, respectively noted $KA_h \subseteq KA$ and $SA_h \subseteq SA$:

$$\begin{array}{rll} (\boldsymbol{u}_h, \boldsymbol{\sigma}_h) = & \underset{\boldsymbol{u},\,\boldsymbol{\sigma}}{\arg\min} & e^2_{(\boldsymbol{u},\boldsymbol{\sigma})} \\ & \text{s.t.} & (\boldsymbol{u}, \boldsymbol{\sigma}) \in KA_h \times SA_h \end{array} \tag{3.41}$$

Both functionals Π_p and Π_c are a sum of a positive definite quadratic function, some linear terms and indicator functions of convex sets thus they are strictly convex. Since the error functional is the sum of two functions of $\boldsymbol{u}$ and $\boldsymbol{\sigma}$ respectively, the above problem can be transformed into two variational forms, namely a kinematic approach in which a kinematically admissible field is obtained by minimizing the total potential energy Π_p:

$$\begin{aligned} \min_{\boldsymbol{u}} \quad \Pi_{p(\boldsymbol{u})} = & \int_{\Omega} \frac{1}{2}\boldsymbol{\varepsilon}_{(\boldsymbol{u})} : \mathbb{D} : \boldsymbol{\varepsilon}_{(\boldsymbol{u})}\, d\Omega + \int_{\Gamma_{\boldsymbol{c}}} \mathbb{1}_{\mathcal{K}^{\circ}_{\mu}(\boldsymbol{g}_{(\boldsymbol{u})})}\, d\Gamma \\ & - \int_{\Omega} \boldsymbol{b}.\boldsymbol{u}\, d\Omega - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d.\boldsymbol{u}\, d\Gamma \\ \text{s.t.} \quad & \boldsymbol{u} \in KA_h \end{aligned} \tag{3.42}$$

and a static approach in which a statically admissible field is obtained by minimizing the total complementary energy Π_c:

$$\begin{aligned} \min_{\boldsymbol{\sigma}} \quad \Pi_{c(\boldsymbol{\sigma})} = & \int_{\Omega} \frac{1}{2}\boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma}\, d\Omega + \int_{\Gamma_{\boldsymbol{c}}} \mathbb{1}_{\mathcal{K}_{\mu}(\boldsymbol{\sigma}.\boldsymbol{n})}\, d\Gamma \\ & - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{u}_d\, d\Gamma - \int_{\Gamma_{\boldsymbol{c}}} (\boldsymbol{\sigma}.\boldsymbol{n}).(g_0\boldsymbol{n})\, d\Gamma \\ \text{s.t.} \quad & \boldsymbol{\sigma} \in SA_h \end{aligned} \tag{3.43}$$

Since the error becomes zero for the optimal solution, both energies are then opposite: $\Pi_{p(\boldsymbol{u}^*)} = -\Pi_{c(\boldsymbol{\sigma}^*)}$. As a result both minimization problems (3.42) and (3.43) provide respectively an upper bound and a lower bound of the real potential energy of the system (after inverting the sign of the complementary energy):

$$-\Pi_{c(\boldsymbol{\sigma}_h)} \leq -\Pi_{c(\boldsymbol{\sigma}^*)} = \Pi_{p(\boldsymbol{u}^*)} \leq \Pi_{p(\boldsymbol{u}_h)} \tag{3.44}$$

These bounds provide the engineer the capacity to judge the quality of convergence regarding the mesh since one can evaluate the proximity of approximate solutions to the exact solution. The relative difference between the upper and lower bound therefore represents the above-introduced error estimator and serves as an excellent indicator of convergence which can be defined as:

$$\Delta E = \frac{\Pi_{p(\boldsymbol{u}_h)} + \Pi_{c(\boldsymbol{\sigma}_h)}}{\Pi_{c(\boldsymbol{\sigma}_h)}} = \frac{e^2_{(\boldsymbol{u}_h,\boldsymbol{\sigma}_h)}}{\Pi_{c(\boldsymbol{\sigma}_h)}} \tag{3.45}$$

where the value of the complementary energy has been arbitrarily chosen for normalization.

3.5 Finite elements discretization

3.5.1 Kinematic approach

Displacement-type elements are used in the finite element discretization of problem (3.42). For the present case, 10-node quadratic tetrahedra will be used. More details

about the kinematic finite element are given in Sec. 4.5.1. Problem (3.42) can then be expressed after assembling the global stiffness matrix $\boldsymbol{K}$ and the nodal force vector $\boldsymbol{f}$ as a conic minimization problem using the same formalism as in Sec. 2.3.2:

$$\begin{array}{lll} (TPE) \quad \min\limits_{\hat{\boldsymbol{u}},\,\hat{\boldsymbol{g}}} & \frac{1}{2}\hat{\boldsymbol{u}}^T\boldsymbol{K}\hat{\boldsymbol{u}} - \boldsymbol{f}^T\hat{\boldsymbol{u}} & \\ \text{s.t.} & \hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}_d & \text{on } \Gamma_{\boldsymbol{u}} \\ & \hat{\boldsymbol{g}} = \hat{\boldsymbol{g}}_0 + [\![\hat{\boldsymbol{u}}]\!] & \text{on } \Gamma_{\boldsymbol{c}} \\ & \hat{\boldsymbol{g}} \in \mathcal{K}_\mu^\circ & \text{on } \Gamma_{\boldsymbol{c}} \end{array} \tag{3.46}$$

in which the primal optimization variables are $\boldsymbol{x} = (\hat{\boldsymbol{u}},\ \hat{\boldsymbol{g}})$, with $\hat{\boldsymbol{u}}$ a vector collecting all nodal degrees of freedom, and $\hat{\boldsymbol{g}}$ a vector of gaps at each node of the finite element mesh faces belonging to a contact surface $\Gamma_{\boldsymbol{c}}$. Contact constraints will be enforced at all 6 nodes of each face with the possibility of adding additional contact points within the face in order to prevent all penetration possible due to quadratic elements. The node-to-node constraints will therefore tend to mimic a complete face-to-face constraint. After introducing a proper change of variable, the contact condition $\hat{\boldsymbol{g}} \in \mathcal{K}_\mu^\circ$ can be easily transformed into a standard second-order cone constraint using a Lorentz cone.

3.5.2 Static approach

3.5.2.1 Regular equilibrium elements

Finite element discretization of problem (3.43) can be done using simple equilibrium tetrahedra. The simplest element can be obtained by considering a linear interpolation of the stress tensor components σ_{ij} over each element. 4-node linear tetrahedra with discontinuous interpolation can be used for which continuity of the stress vector over the element facets should be explicitly written (these faces will be denoted as the set $\Sigma_{\boldsymbol{\sigma}}$) and included as linear constraints of the corresponding optimization problem.

Let $i = 1, \ldots, 4$, be the node order index. The stress state of a point is defined by six stress components, ${\boldsymbol{\sigma}^i}^T = \{\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{yz}, \sigma_{xz}\}$. Let $P_i(\boldsymbol{x})$ be the linear shape function associated with each node. The element stress state is defined by the 24 stress components of the nodes as ${\boldsymbol{\sigma}^e}^T = \{{\boldsymbol{\sigma}^1}^T, \ldots, {\boldsymbol{\sigma}^4}^T\}$. We can write the stresses of an arbitrary point as:

$$\boldsymbol{\sigma} = [\boldsymbol{I}P_1, \ldots, \boldsymbol{I}P_4]\boldsymbol{\sigma}^e = \boldsymbol{P}\boldsymbol{\sigma}^e \tag{3.47}$$

The element compliance matrix can then be calculated as follows:

$$\boldsymbol{F}^e = \int_{\Omega^e} \boldsymbol{P}^T\mathbb{D}^{-1}\boldsymbol{P}\, d\Omega \tag{3.48}$$

using a 4 Gauss-points quadrature rule.

Using this equilibrium element and the typical definition for the normal projection matrix operator $\boldsymbol{N}$ on a facet of normal $\boldsymbol{n}$ and the divergence matrix operator $\boldsymbol{D}$, problem (3.43) can be expressed after assembling the global matrices as a conic minimization problem using the same formalism as in Sec. 2.3.2:

$$
\begin{aligned}
(TCPE) \quad \min_{\hat{\boldsymbol{\sigma}}, \boldsymbol{t}_c} \quad & \frac{1}{2}\hat{\boldsymbol{\sigma}}^T \boldsymbol{F}\hat{\boldsymbol{\sigma}} - (\boldsymbol{N}\hat{\boldsymbol{\sigma}})^T \boldsymbol{u}_d - \boldsymbol{t}_c^T \hat{\boldsymbol{g}}_0 \\
\text{s.t.} \quad & \boldsymbol{D}\hat{\boldsymbol{\sigma}} + \boldsymbol{b} = \boldsymbol{0} && \text{in } \Omega \\
& [\![\boldsymbol{N}\hat{\boldsymbol{\sigma}}]\!] = \boldsymbol{0} && \text{on } \Sigma_{\boldsymbol{\sigma}} \\
& \boldsymbol{N}\hat{\boldsymbol{\sigma}} = \boldsymbol{t}_d && \text{on } \Gamma_{\boldsymbol{t}} \\
& \boldsymbol{N}\hat{\boldsymbol{\sigma}} = \boldsymbol{t}_c && \text{on } \Gamma_{\boldsymbol{c}} \\
& \boldsymbol{t}_c \in \mathcal{K}_\mu && \text{on } \Gamma_{\boldsymbol{c}}
\end{aligned}
\tag{3.49}
$$

in which the primal optimization variables are $\boldsymbol{x} = (\hat{\boldsymbol{\sigma}},\ \boldsymbol{t}_c)$, with $\hat{\boldsymbol{\sigma}}$ a vector collecting all nodal stresses, and $\boldsymbol{t}_c$ a vector of contact surface tractions at each node of the finite element mesh faces belonging to a contact surface $\Gamma_{\boldsymbol{c}}$. Contact constraints will be enforced at all 3 nodes of each face. Similarly, after introducing a proper change of variable, the contact condition $\boldsymbol{t}_c \in \mathcal{K}_\mu$ can be easily transformed into a standard second-order cone constraint using a Lorentz cone.

After implementing this method, poor performance results were obtained mainly due to the size of the problem and to linearly dependent conditions in the linear constraint matrix collecting equilibrium and traction continuity conditions. This made the problem ill-posed and caused troubles when computing the solution of the linear system equations using a direct solver. The dependencies come from the divergence condition which can be easily eliminated by integrating this condition in the definition of the finite element thus using a 21 components equivalent stress vector rather than 24, but more importantly they come from the traction continuity condition over the tetrahedra faces $\Sigma_{\boldsymbol{\sigma}}$. In fact, it has been shown that there exists some null eigenvalues in the compliance matrix in addition to those due to the natural rigid body modes. They correspond to zero energy modes called spurious kinematic modes (SKM) [de Almeida and de Freitas, 1991; de Almeida and Maunder, 2017; Fraeijs de Veubeke, 1951; Kempeneers et al., 2009; Pian, 1964]. A comprehensible analysis of the spurious modes and their illustration can be found in the seminal work of de Almeida and Maunder [2017].

3.5.2.2 Improved equilibrium elements

Following the works of Pian [1964], Kempeneers et al. [2009] and de Almeida and Maunder [2017], improved equilibrium finite elements can be derived by choosing an appropriate interpolation for $\boldsymbol{\sigma}$ which will already include the local equilibrium condition $\mathrm{Div}(\boldsymbol{\sigma}) = 0$. To do so, the stress distribution inside an element is chosen in order to satisfy local equilibrium equations with body forces such as:

$$
\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{q} + \boldsymbol{\sigma}_{\boldsymbol{b}}(\boldsymbol{x}) \tag{3.50}
$$

where $\boldsymbol{S}$ is a matrix of independent shape functions of degree p, of which the columns $\boldsymbol{S}_k$ all verify $\mathrm{Div}(\boldsymbol{S}_k) = 0$. They can be obtained either from Morera [1892] or Maxwell [1866] stress potentials. In order to verify the equilibrium conditions on all the internal facets of the mesh, Lagrange multipliers which can be interpreted as the displacements of each facet are also interpolated. Contact conditions can easily be written using these new variables. The details of this implementation are presented in Sec. 4.5.2.

3.6 Mesh adaptivity

In order to limit the number of elements used in a 3D analysis, a general adaptive remeshing scheme is implemented using the error estimator defined in Eq. (3.30). First, the volume contribution of the error term is computed for each element of the mesh using the dual approximate solutions $(\boldsymbol{u}_h, \boldsymbol{\sigma}_h) \in KA_h \times SA_h$ obtained from the resolution of problems (3.46) and (3.49).

Remeshing algorithms usually require a node-based metric, for that purpose the previously computed element-wise error contribution is converted to a nodal field based on error density by performing an average on all elements sharing a common node. A scalar error ratio field is therefore obtained over all nodes of the model. For that, an elementary error value e_e^2 for each element e of the mesh should be calculated using the dual optimal solution $(\boldsymbol{u}^*, \boldsymbol{\sigma}^*) \in KA(\Omega) \times SA(\Omega)$. Thus, we have:

$$e^2(\boldsymbol{u}, \boldsymbol{\sigma}) = \sum_{e=1}^{N_e} e_e^2(\boldsymbol{u}, \boldsymbol{\sigma}) \tag{3.51}$$

with N_e being the total number of elements in the mesh. Finally, since the error estimator also contains surface terms over the contact area, the elementary surface term of a given contact facet is split evenly between all nodes of the contact facet and added to the previous contribution. A scalar nodal error estimator is therefore obtained.

Using an h-version remeshing scheme [Ladevèze and Pelle, 2005; Zienkiewicz and Taylor, 2000], one can use this nodal scalar field to change the mesh size accordingly, thus improving the quality of the results while limiting the number of elements compared to a global mesh-size reduction. For this purpose, we used a general anisotropic remesh computed on the basis of an anisotropic 3×3 metric tensor. More details about determining such an anisotropic mesh size map can be found for instance in Alauzet and Frey [2006]. Distene's MeshGems [1] surface (MeshGems-CADSurf) and volume (MeshGems-Tetra) are used to produce the required meshes for the finite elements analyses.

The complete procedure therefore involves the following steps:

1. Over an automatically software-generated coarse mesh $\Omega_h^{(0)}$, an initial static and kinematic calculation are performed.
2. On this mesh $\Omega_h^{(0)}$, an error calculation is made, and global, elementary and relative errors are computed along with the relative difference between objectives functions as given by Eq. (3.45).
3. Nodal relative errors are then used to compute a new anisotropic mesh size map.
4. A new mesh $\Omega_h^{(1)}$ is generated and new static and kinematic analyses are performed.

The process is repeated until a target objective difference is reached. However, this does not prevent bad quality meshes with depends on the software used to produce that mesh. A post-treatment is realized over the newly-obtained mesh where quality measures are checked and the mesh is cleaned.

1. http://www.meshgems.com/

3.7 Illustrative examples

3.7.1 A simple example for remesh scheme with contact error

This example is a simple test for the error and remesh scheme using a square beam with contact conditions. The dimensions of the beam are chosen to be in agreement with the Euler-Bernoulli aspect ratio hypothesis: the main beam has a section of 0.02×0.02 m^2 and a length of 0.8 m and underneath is a similar section beam but with a length of 0.4 m. The main beam is fully fixed on one end while the second one is fully fixed at its base and right end. Unilateral contact conditions are imposed between the two beams. A downwards 2 kN force is applied on the edge of the main beam, thus simulating a fixed beam with a pinned support at its middle. The steel material for the beam is modeled as linear elastic with Young modulus of 210 GPa and a Poisson's ratio of 0.3. Fig. 3.4 gives a quick description of the model. This 3D problem can be well approximated by a simple 1D Euler-Bernoulli beam model with intermediate support as sketched at the bottom of Fig. 3.4.

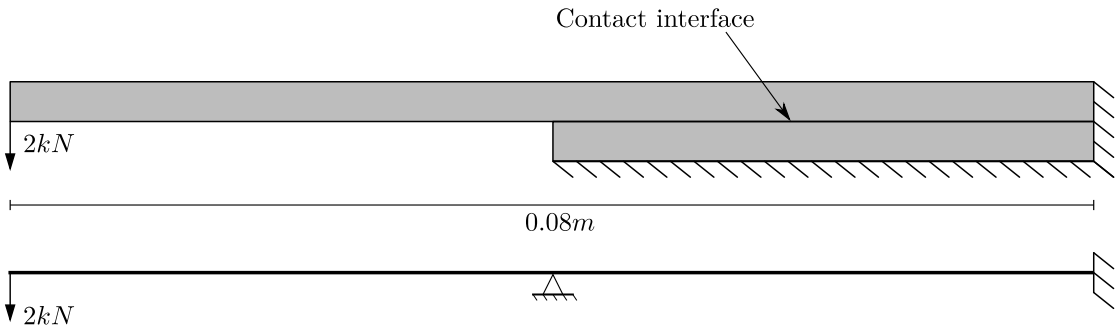
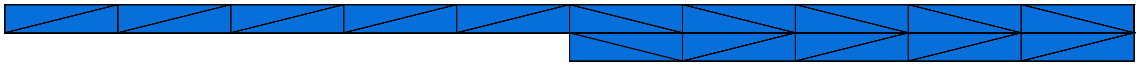


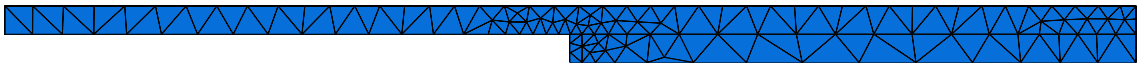
Figure 3.4 – Description of the 3D model and equivalent Euler-Bernoulli beam

Six remesh iterations are made and the difference between the two objective functions of the approaches is measured to assess the convergence quality. Fig. 3.5 shows three of the seven different meshes used. One can clearly see the effectiveness of the remesh procedure which concentrates meshes over the contact surface between the beams. Since the relative difference computed from (3.45) gives an upper bound of the real error, the dual approach allows us to assess the global convergence quality and that by just calculating the relative difference between the two objective functions of the minimization problems.

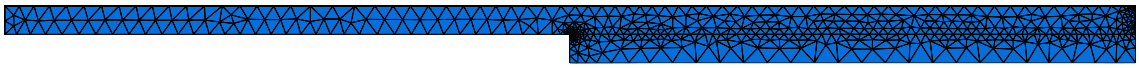
Fig. 3.6a shows the evolution of this difference over the mesh iterations. We can clearly see that after the second iteration, the difference is already less than 2% which can be considered as a good convergence level for engineering design applications. Fig. 3.6b plots the different total errors calculated over the different meshes which shows that the remesh scheme is very efficient in reducing the volume error, but more importantly the surface term also which constitutes the greater contribution for this specific case.



(a) Initial mesh



(b) Third remesh iteration



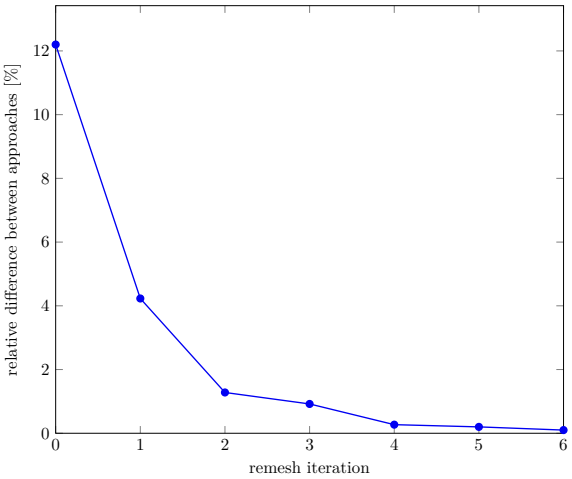
(c) Fifth remesh iteration

Figure 3.5 – View of some of the different meshes

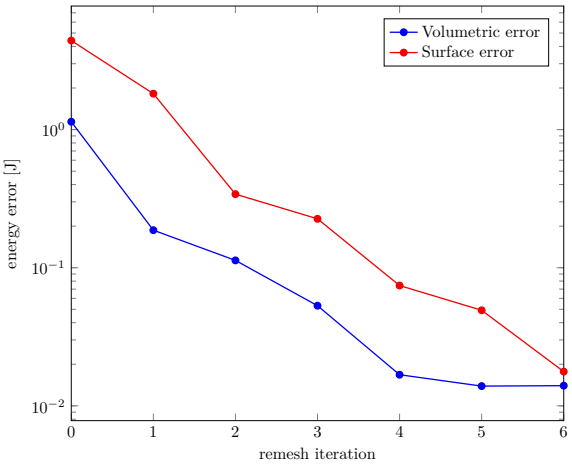
Local quantities can also be compared such as the displacement obtained by the two approaches. Its evolution over the remeshes follows the same tendency as the errors. Indeed, the discrepancy between displacement obtained from the kinematic approach and the one post-processed from the static approach tend to close and eventually become indistinguishable at the fourth iteration as shown in Fig. 3.6c. The maximum displacement at the tip of the beam is also in accordance with the analytical value of 2.66 cm obtained from the simple 1D beam model of Fig. 3.4. Plots of the deformed mesh and the normal gap are also given in Fig. 3.7 and 3.8.

3.7.2 Performance and results comparison with Abaqus

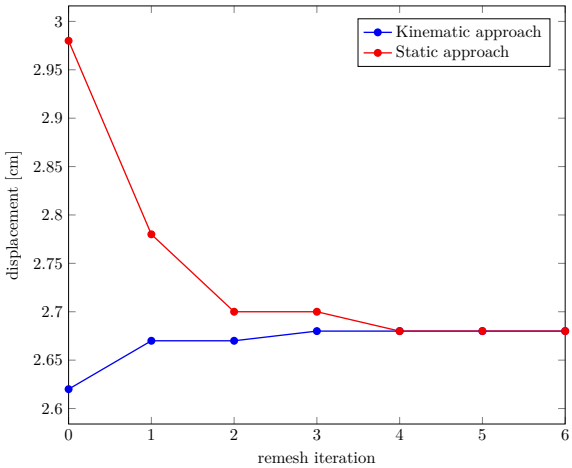
The next example aims at assessing the procedure's computational performance with respect to computations made using Abaqus. The considered problem is a beam-column continuous connection frequently found in steel-structures buildings, the constitutive material being the same as before. To eliminate all forms of variations and to focus on the performance of the solvers only, the same series of 5 iteratively refined meshes is used in all the studies, both for our implementation and for Abaqus computations. This model is calculated using the IP algorithm and various configurations in Abaqus, each corresponding to a different modeling choice of contact enforcement method. These are summarized in Tab. 3.1. Note that these choices have been made in order to position the IPM with respect to a very simple and general-purpose enforcement of contact constraints (penalty approach) as well as a more advanced one (Augmented Lagrangian). For the latter case, additional contact points other than the nodes are used by Abaqus in the surface-to-surface computation. For the comparison to be fair, we also added 4 additional points inside each triangular facets for contact enforcement in our implementation, thereby reducing the potential violation of contact conditions in the node-to-node formulation inside the facet. With such additional contact points, similar system sizes are obtained. Default parameters have been retained for assessing the robustness of our



(a) Relative difference between energies



(b) Total volume and surface energy errors



(c) Maximum vertical displacement at the tip of the first beam

Figure 3.6 – Convergence plots for the first example

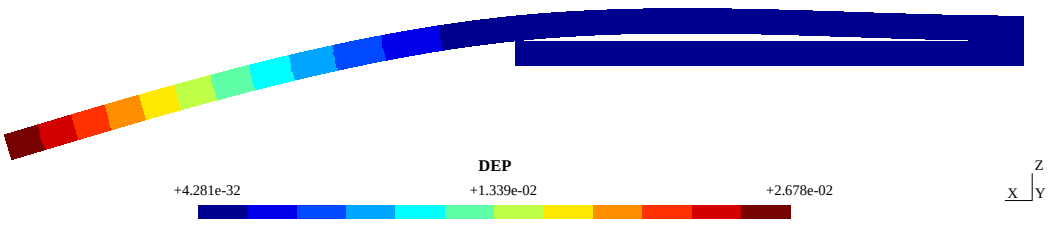


Figure 3.7 – Deformed shape and displacement magnitude (in meters) for the fourth remesh iteration

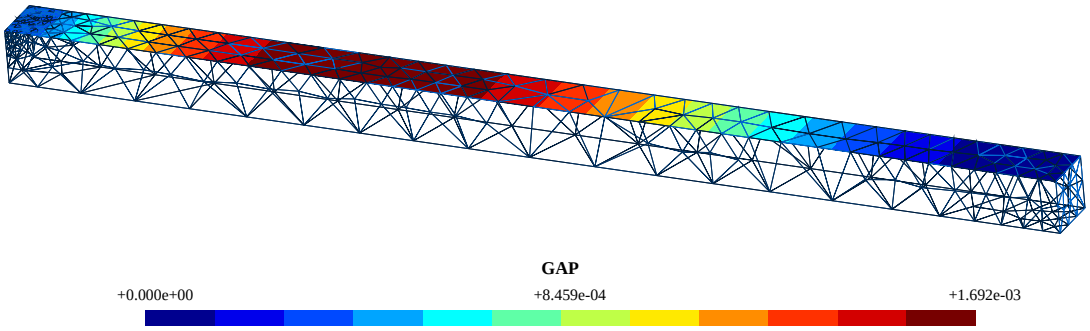


Figure 3.8 – Normal gap (in meters) on the contact surface between the two beams

approach with respect to its use by an inexperienced user.

Two variants are defined by changing the friction coefficient between the beam's end-plates and the column's flanges:

- a frictionless contact is used for model 1 (only unilateral contact conditions are enforced so that the comparison is most relevant), and
- a $\mu = 0.3$ friction coefficient is used for model 2 (note that Abaqus solves standard non-associated Coulomb friction)

Fig. 3.9 gives a general description of the model while three mesh iterations are represented on Fig. 3.10. It consists of a HEB200 central column with two IPE360 beams attached over the flanges using welded end plates and bolts. The end plates have a 15 mm thickness and 6 M18 bolts are used to attach each beam. Web stiffeners are used to prevent web buckling: for the HEB column, their thickness is 12.7 mm to coincide with the HEB's flange thickness, whereas for the IPE beams, 15 mm thick stiffeners are used over the supports to correctly channel shear forces. The stiffeners are glued to the beams and column using tie constraints in Abaqus. We suppose that the bolt hole is equal to its diameter and, to prevent rigid body motions, one of the bolt heads is glued to the plate. A 45 MPa normal pressure is applied over the top section of the HEB column and the displacement of the nodes belonging to the surface defined by the intersection of the IPE's web stiffeners are blocked in all the 3 directions, the HEB sections remaining free.

Since Abaqus offers displacement-based elements only, the calculation times are compared with respect to the kinematic approach. The same meshes are used in Abaqus therefore the solved system size is nearly the same and the system's sparsity patterns and conditioning are also comparable. The main difference lies in contact modeling and resolution strategy. In order to closely compare the numerical performance of the algorithm, all the calculations were made using the same machine. OpenMP technology was used to parallelize over 8 threads. Tab. 3.2 summarizes the machine's specifications.

Tab. 3.3 and 3.4 show the CPU times for the models described in Tab. 3.1. The

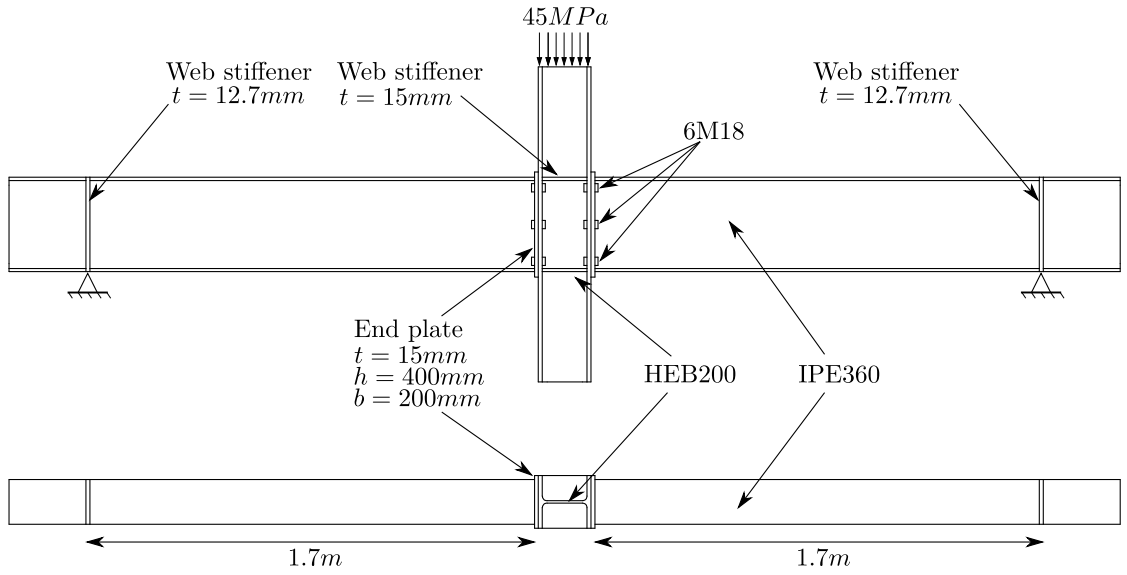
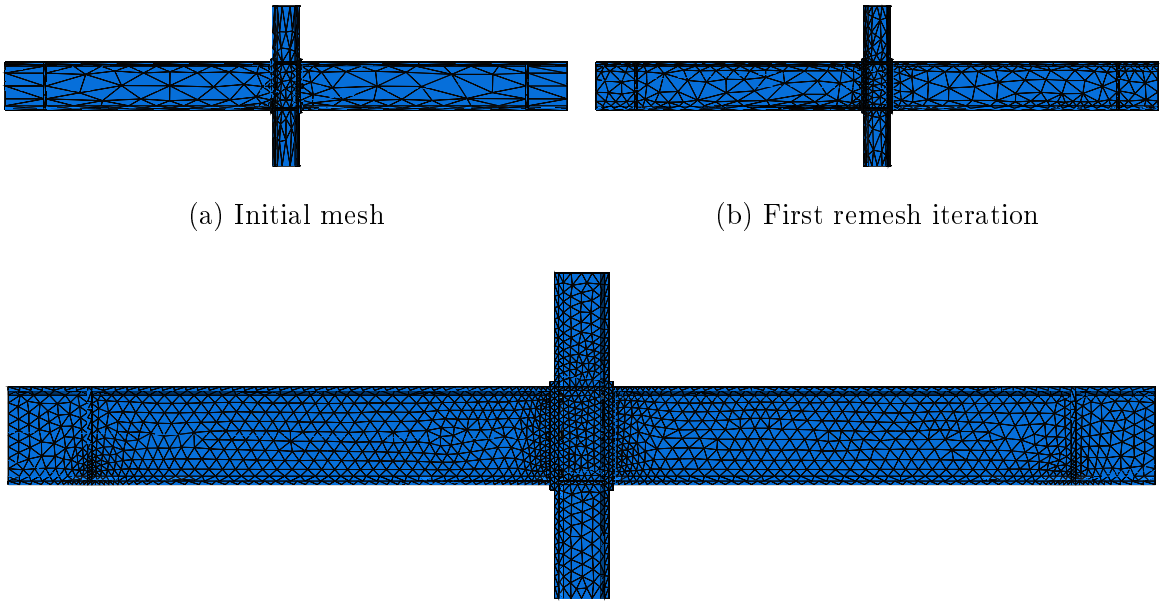


Figure 3.9 – Description of the used model



(a) Initial mesh

(b) First remesh iteration

(c) Third remesh iteration

Figure 3.10 – View of some of the different meshes

Table 3.1 – Definition of contact modeling choices in Abaqus

	Model 1-1	**Model 1-2**	**Model 2-1**	**Model 2-2**
Mesh type	C3D10	C3D10	C3D10	C3D10
Contact pairs type	node to surface	surface to surface	node to surface	surface to surface
Use supplementary contact points	No	Yes	No	Yes
Normal behavior	Hard contact	Hard contact	Hard contact	Hard contact
Enforcement method	Penalty	Augmented Lagrangian	Penalty	Augmented Lagrangian
Stiffness reduction	0.01	No reduction	0.01	No reduction
Tangential behavior	Frictionless	Frictionless	Frictional	Frictional
Enforcement method			Penalty	Penalty
Friction coefficient			0.30	0.30

Table 3.2 – Computer specifications

Processor	Intel core i7-4700MQ
Number of cores	4
Number de threads	8
Base frequency	2.40 GHz
Maximum frequency	3.40 GHz
Cache	6 MB –SmartCache
RAM	16.0 GB (15.7 GB usable)

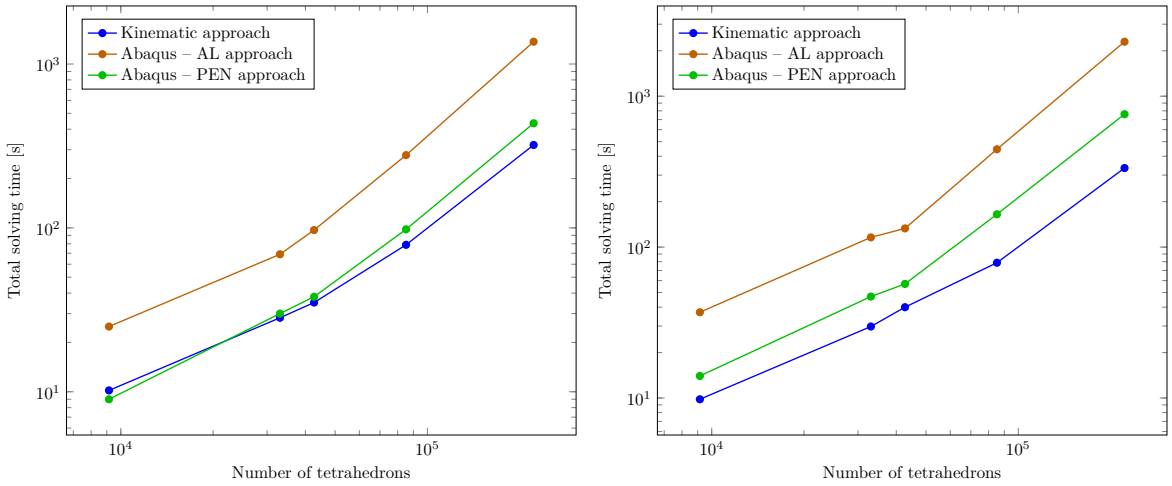
IPM is largely comparable to Abaqus penalty approach (PEN) in terms of CPU times, with even a small speed-up factor when compared to the augmented Lagrangian (AL) approach. Compared to the latter, which generally yields more accurate results, the IPM shows a speed-up factor of 2 to 4 in frictionless contact and up to 6 in frictional contact, the speed-up factor increasing with the mesh size. This is an attractive feature of the IPM since it can handle large-scale sparse problems efficiently. While Abaqus number of iterations required for convergence tends to increase (from 20 to 30 for the AL approach), the IPM number of iteration remains fairly constant varying between 18

Table 3.3 – First model (frictionless contact) CPU times and speed-up factors of the IPM over Abaqus augmented Lagrangian approach and penalty approach

Remesh iteration	**Mesh Size**	**IPM kinematic approach** (s)(N_{iter})	**Abaqus AL approach** (s)(N_{iter})	**Speed-up factor**	**Abaqus PEN approach** (s)(N_{iter})	**Speed-up factor**
0	$N_e = 9150$	10.2 (17)	25.0 (21)	2.5	9.0 (8)	0.9
1	$N_e = 33051$	28.3 (18)	69.0 (22)	2.4	30.0 (9)	1.1
2	$N_e = 42696$	35.0 (17)	97.0 (24)	2.8	38.0 (9)	1.1
3	$N_e = 85187$	78.8 (17)	278.0 (25)	3.5	98.0 (9)	1.2
4	$N_e = 222082$	320.6 (18)	1368.0 (30)	4.3	435.0 (10)	1.4

Table 3.4 – Second model (frictional contact) CPU times and speed-up factors of the IPM over Abaqus augmented Lagrangian approach and penalty approach

Remesh iteration	Mesh Size	IPM kinematic approach (s)(N_{iter})	Abaqus AL approach (s)(N_{iter})	Speed-up factor	Abaqus PEN approach (s)(N_{iter})	Speed-up factor
0	$N_e = 9150$	9.8 (18)	37.0 (23)	3.8	14.0 (8)	1.4
1	$N_e = 33051$	29.8 (18)	116.0 (24)	3.9	47.0 (10)	1.6
2	$N_e = 42696$	40.0 (18)	133.0 (21)	3.3	57.0 (9)	1.4
3	$N_e = 85187$	78.8 (18)	445.0 (25)	5.6	165.0 (9)	2.1
4	$N_e = 222082$	334.1 (19)	2296.0 (29)	6.9	760.0 (10)	2.3



(a) Model 1 – Frictionless contact

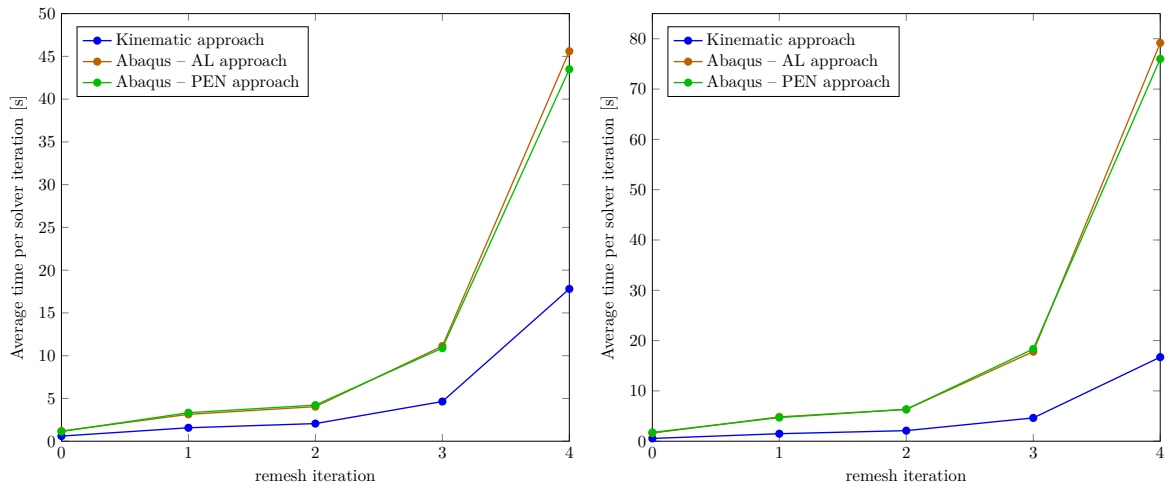
(b) Model 2 – Frictional contact

Figure 3.11 – Total solving time for each mesh size

and 20 iterations for all the models and remeshes, this almost independence of number of iteration with respect to problem size is a known attractive feature of IPM methods. Fig. 3.11 and 3.12 respectively illustrate the total CPU cost and the average CPU cost per inner iteration of the non-linear solvers for the different configurations.

One should note that the dual solution requires as much as time to calculate as the primal one since the linear system is roughly the same size. Since the problems are independent, they could be easily parallelized. The cost of the error calculation is a simple post-processing step which takes less than 1% of the time needed for the whole resolution which is negligible.

To assess the global convergence level with mesh refinement, total quantities are considered. The influence of mesh size over the relative difference between the static and kinematic IP approaches has been represented in Fig. 3.13 and 3.14. The quality of convergence of the finite element analysis can therefore be directly evaluated by checking the relative difference between the approaches which reaches acceptable values (less than 5% directly after the first remesh iteration) and 1.68% for the frictionless model



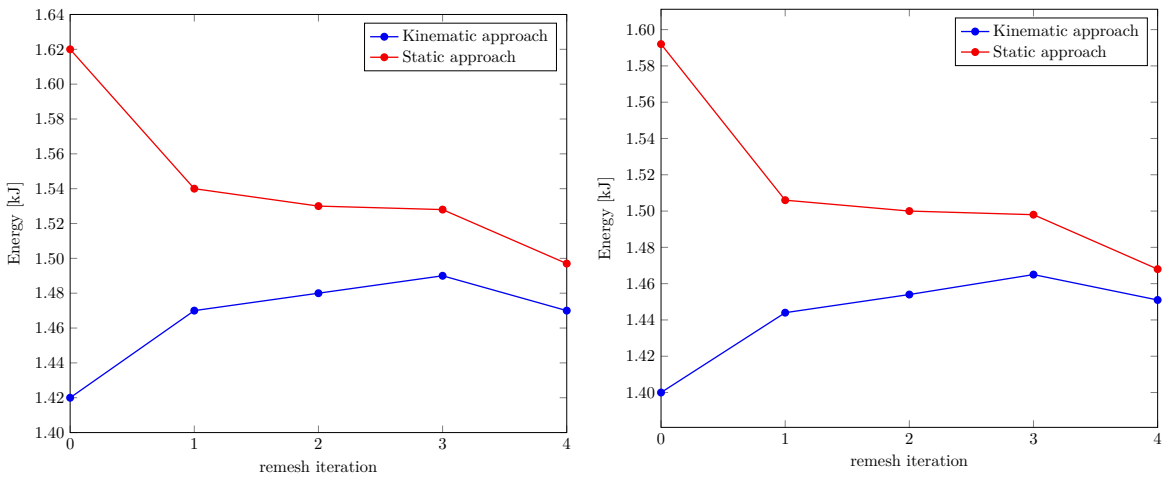
(a) Model 1 – Frictionless contact

(b) Model 2 – Frictional contact

Figure 3.12 – Average solving time per internal algorithm iteration for each mesh iteration

and 1.2% for the frictional model on the fourth remesh iteration. When inspecting the convergence rate of the global error when refining the mesh, we observe that it is roughly linear with respect to the number of degrees of freedom.

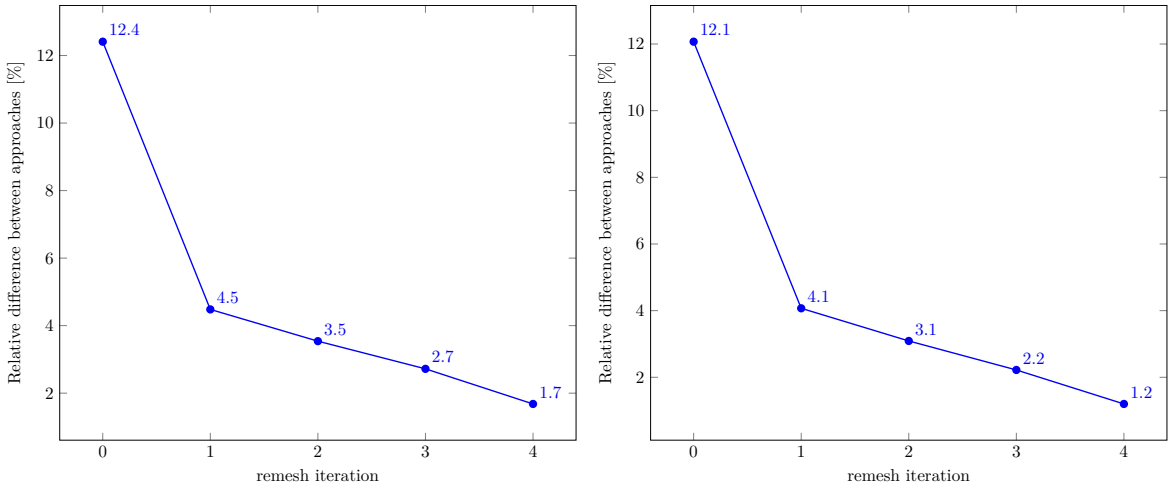
Let us now consider some local quantities. Fig. 3.15 shows the evolution of the total vertical displacement of the column over the six meshes. We can clearly see the upper and lower bound offered by the dual approach that we used. Abaqus AL approach shows nearly the same results as the kinematic approach whereas with the penalty contact constraints, one can easily over-estimate displacements thus leading to misinterpretations. These aspects are generally unknown to engineers who do not have means to estimate an upper bound for displacements, contrary to what offers equilibrium-based computations. The maximum gap between the end plate and the column's flange presented in Fig. 3.16 shows the consequences of the contact constraints enforcement method. It is clear that with the penalty, the maximum gap is overestimated. An important point is that with the IPM, no contact constraint violation is permitted, thus we have zero penetration, whereas in Abaqus, a penetration tolerance is allowed. The results show that a « negative » gap, equivalent to up to 8% of the maximum gap, is possible with the penalty approach ($-0.055\,\text{mm}$ compared to $0.707\,\text{mm}$ for the first mesh in model 1 as seen in Fig. 3.17c and $-0.058\,\text{mm}$ compared to $0.697\,\text{mm}$ for the first mesh in model 2 as seen in Fig. 3.19c). However, with the AL approach, negative gaps are limited to 0.0001% of the maximum gap, thus yielding better results that are comparable to the IP kinematic approach. With the penalty approach in Abaqus and the removal of supplementary contact constraints, one can manage to obtain results in a more reasonable time. Nonetheless, the quality of the results is poor and contact conditions are generally over-evaluated or violated as shown in Fig. 3.16 and in the gap iso-values shown in Fig. 3.17 to 3.20 (COPEN being the gap keyword in Abaqus).



(a) Model 1 – Frictionless contact

(b) Model 2 – Frictional contact

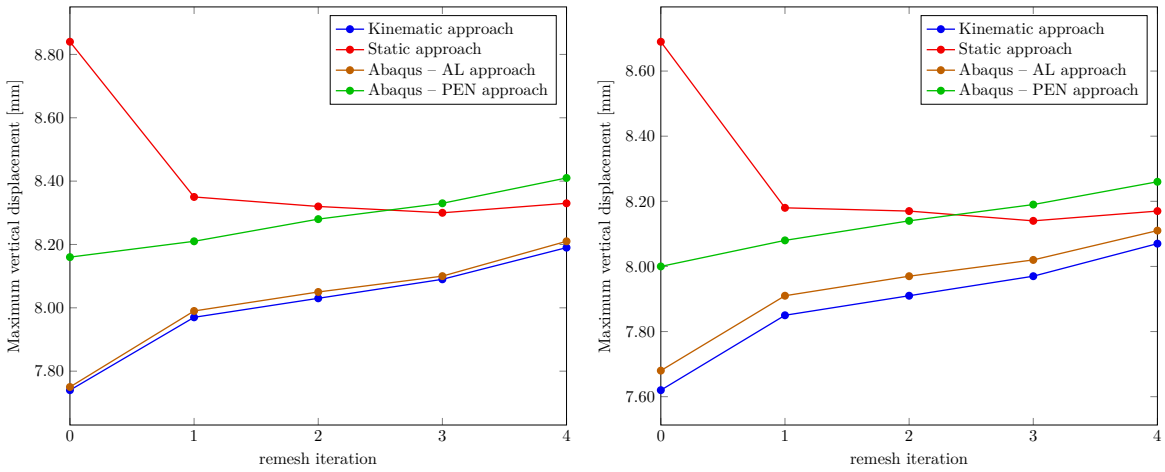
Figure 3.13 – Objective function values for the static and the kinematic IP approach



(a) Model 1 – Frictionless contact

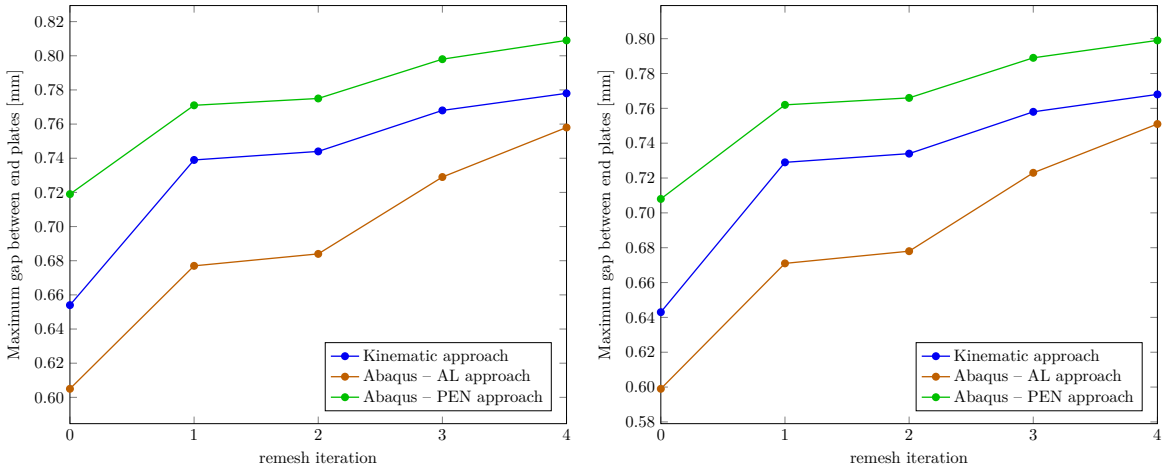
(b) Model 2 – Frictional contact

Figure 3.14 – Relative difference between static and kinematic IP approach's objective functions



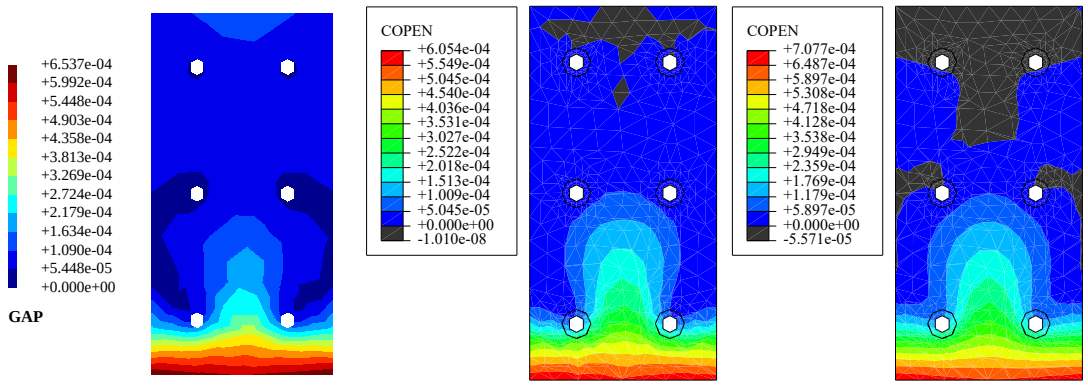
(a) Model 1 – Frictionless contact (b) Model 2 – Frictional contact

Figure 3.15 – Total vertical displacement of the column



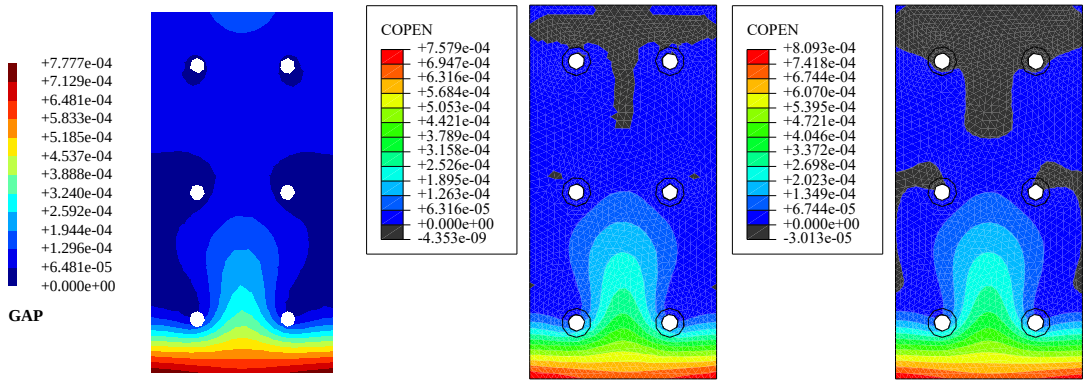
(a) Model 1 – Frictionless contact (b) Model 2 – Frictional contact

Figure 3.16 – Maximum gap between the end plate and the column's flange



(a) IPM kinematic approach (b) Abaqus AL approach (c) Abaqus PEN approach

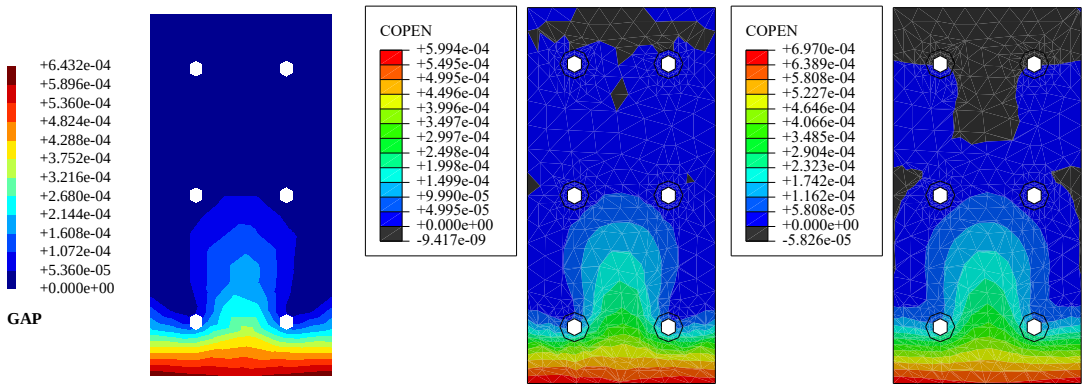
Figure 3.17 – Model 1 – frictionless contact, initial coarse mesh, gap iso-values in (m) over one of the end-plates



(a) IPM kinematic approach (b) Abaqus AL approach (c) Abaqus PEN approach

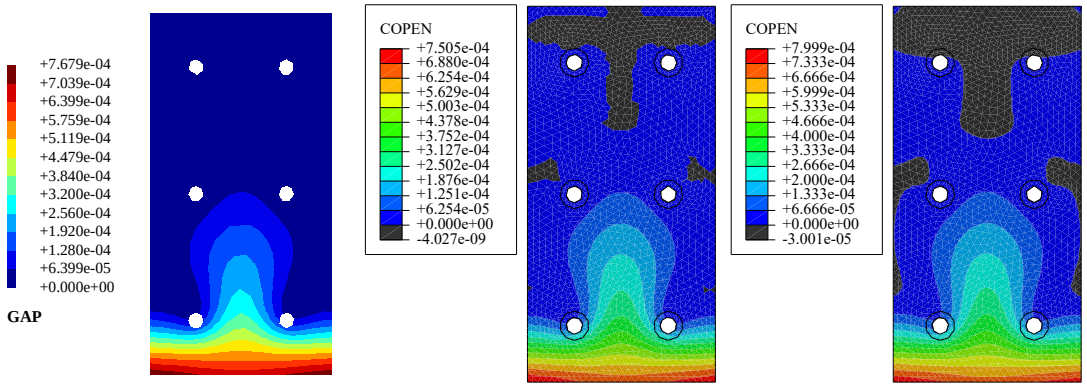
Figure 3.18 – Model 1 – frictionless contact, fourth remesh iteration, gap iso-values in (m) over one of the end-plates

It must be noted that in the case of frictionless contact, both our model and Abaqus impose the same Signorini unilateral conditions so that the comparison is relevant. However, the validity of the comparison between the models including friction can be questioned because of our choice of associated friction. In particular, it can be expected that both computations will differ, especially regarding the normal and tangential components of the gap vector since associated friction predicts normal opening during sliding. Nevertheless, due to the end-plates flexural behavior considered in this example, surfaces in sliding are quite limited in size (they approximately correspond to the black region of Fig. 3.20b) and the maximum tangential gap amplitude is less than 5% of the maximum normal gap. The impact of associated friction is therefore quite limited for this specific situation. Fig. 3.21 and 3.22 show the local contact pressure over the end plates, one obtained using the IP static approach and the two other from Abaqus



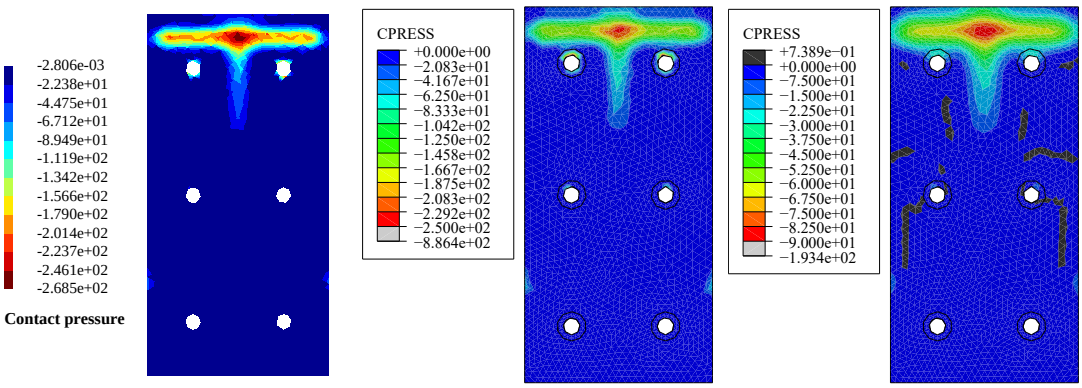
(a) IPM kinematic approach (b) Abaqus AL approach (c) Abaqus PEN approach

Figure 3.19 – Model 2 – frictional contact, initial coarse mesh, gap iso-values in (m) over one of the end-plates



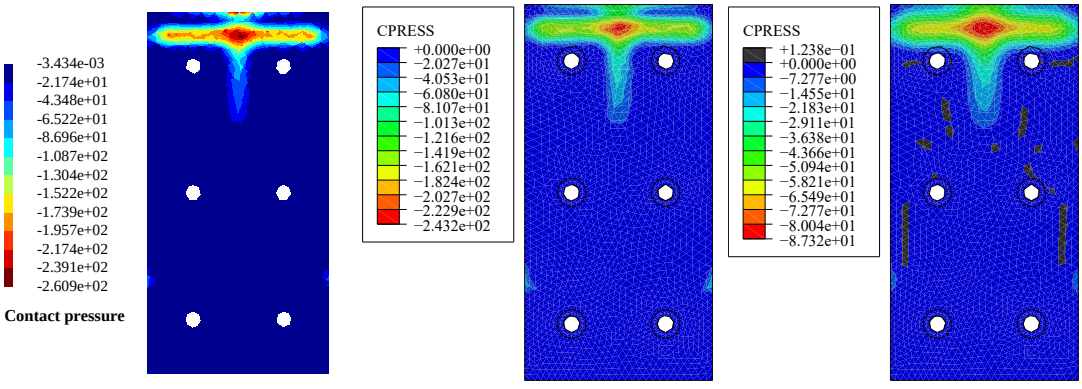
(a) IPM kinematic approach (b) Abaqus AL approach (c) Abaqus PEN approach

Figure 3.20 – Model 2 – frictional contact, fourth remesh iteration, gap iso-values in (m) over one of the end-plates



(a) IPM static approach (b) Abaqus AL approach (c) Abaqus PEN approach

Figure 3.21 – Model 1 – frictionless contact, fourth remesh iteration, normal pressure iso-values in (MPa) over one of the end-plates



(a) IPM static approach (b) Abaqus AL approach (c) Abaqus PEN approach

Figure 3.22 – Model 2 – frictional contact, fourth remesh iteration, normal pressure iso-values in (MPa) over one of the end-plates

(CPRESS being the normal pressure keyword in Abaqus). We can clearly see that the stresses post-processed by Abaqus from a displacement field are underestimated. While the AL approach gives comparable results to the IP static approach (247 MPa compared to 268 MPa for model 1 and 243 MPa compared to 261 MPa for model 2, again, despite the modeling difference of frictional behavior), the penalty approach largely underestimates these values (89.2 MPa for model 1 and 87.3 MPa for model 2). Again, close agreement between our computations and Abaqus AL computations in terms of normal pressure values is found due to a limited impact of associated friction in this bending-dominated case. However, the comparison should be made with greater care in the case of compressed end plates with an important torsional loading for instance, where the tangential behavior would be dominant. These aspects will be investigated in further studies in which non-associated frictional contact will be considered. In any case, relying on a dual approach enables to produce results of good quality in a robust and computationally efficient manner which is very attractive in the context of design engineering.

3.8 Conclusions and orientation

Modeling three-dimensional elastic structures in contact under static loads using SOCP and IPM has been investigated in this chapter. Using a dual principle approach, a kinematically admissible displacement field and statically admissible stress field can be obtained by solving a pair of minimization problems. The obtained second-order cone optimization problems are solved using a primal-dual interior point method for which convergence properties are ensured. The optimal pair of variables allows for computing an error estimator based on dual analysis which is then used to produce a new anisotropic mesh size map. The whole process of solving the minimization problems, calculating errors and remeshing, has been automatized. The process is repeated until a target relative difference between objective functions is reached. Since this difference measures the global quality of convergence of the two approaches and provides an upper bound estimate of the global error, it is an excellent indicator of the convergence of the finite elements computations.

The two examples presented in this chapter show the numerical and technical efficiency of this process. The error estimator which consists of an elementary volume term and a local surface contribution due to contact complementarity condition violation enables to remesh areas of interest while limiting the number of elements. Few remesh iterations are generally enough to ensure a difference under 2%, this number can be reduced if the coefficient of error reduction used in the remesh algorithm is increased and if the initial mesh is less coarse. However, the mesh size map will be more difficult to process by automatic mesh generators which could produce bad quality meshes.

The IPM's capacity to take into account a great number of conic constraints shows all its advantages over traditional penalty or augmented Lagrangian approaches. Excellent results were obtained for large-scale problems, with a relatively constant number of iterations ranging between 17 and 20 while the number of iterations in Abaqus AL

approach tended to increase with the problem size. At the expense of quality, contact problems can be solved with nearly the same time as the IPM using Abaqus penalty approach. However, it is known that enforcing contact conditions with such a general-purpose strategy can yield poor estimates of stress levels and larger displacements in structures. The IPM dual approach appears therefore as an interesting alternative, offering high-quality displacement and stress estimates while relying on a robust solver with good scaling properties for large models. This aspect may be very appealing for safety considerations in a structural design context.

Another main advantage is the absence of user intervention while using the IPM. While contact solution algorithms and/or parameters can greatly influence calculation times or sometimes cause divergence, the IPM offers robust and efficient convergence behavior with no user intervention and no algorithmic input parameter.

In the next chapter, we will extend the previous methodology to the case of an elastoplastic behavior in order to consider the problematic of the ultimate limit state design.

* *
*

Chapter 4

Elastoplastic analysis and yield design using second-order cone programming and dual finite-elements

Abstract: *The proposed dual analysis coupled with the SOCP framework is extended to include material elastoplasticity. Displacement and stress-based variational formulations are considered and appropriate finite-element discretization strategies are chosen, yielding respectively an upper and lower bound estimate of the exact solution. The link with classical upper and lower bound theorems of limit analysis is also established. The proposed concepts are illustrated with various applications taken from the Eurocodes and throughout the study of steel assemblies.*

Contents

4.1 Motivation

The design of steel assemblies poses formidable challenges to the designer when the steel joint geometry becomes complex. On the one hand, normative design methods involve lots of manual verification and, on the other hand, the use of numerical simulation by civil engineers is often limited to crude elastic computations and elastic limit checks, largely underestimating the plastic capacity of the steel assembly which is taken into account by the design norms. Resorting to a fully non-linear elastoplastic computation on a large-scale 3D model is still extremely rare in today's engineering practices, mostly due to the difficulty of running such computations which require both time and expertise and often suffer from robustness issues which may be detrimental to the structure's safety.

Building up on the concepts introduced in chapter 3, in this chapter we consider the formulation of elastoplastic problems in a convex optimization setting, sharing therefore close similarities with the formulation of limit analysis problems. This approach was initiated by Maier [Capurso and Maier, 1970; Maier, 1968, 1969] where the incremental problem is reduced to a convex quadratic problem in the case of a piecewise-linear yield criterion. The standard elastic predictor/plastic corrector approach is obtained when the local minimization problem over plastic strain variables is solved exactly (return mapping step) and the stress field balance equations (optimality condition on the displacement field) is expressed as a non-linear function of the total strain. Other approaches may include non-smooth Newton methods [Christensen, 2002; Sander and Jaap, 2019], general interior-point methods [Krabbenhoft et al., 2007], sequential quadratic programming [Bilotta et al., 2012], accelerated proximal gradient methods [Kanno, 2016], etc. One can also mention the use of a bi-potential method for non-associative behaviors [de Saxcé and Feng, 1998; Hjiaj et al., 2003].

When generalizing the results of limit analysis problems and their conic representation depending on the chosen plasticity yield criterion [Krabbenhøft et al., 2007; Makrodimopoulos, 2010], the obtained elastoplastic optimization problem can belong to the class of quadratic programming (QP), second-order cone programming (SOCP) or semi-definite programming (SDP) problems. For instance, one obtains QP problems for piecewise-linear yield surfaces, SOCP for 2D or 3D von Mises yield surface [Kanno, 2011; Yonekura and Kanno, 2012] or SDP for 3D Mohr-Coulomb yield criterion [Martin and Makrodimopoulos, 2008]. When applying the IPM to an elastoplastic computation, this technique exhibits a fundamental difference with classical elastic predictor/plastic corrector schemes coupled with a Newton-Raphson procedure. Indeed, contrary to the latter which alternates between satisfying global equilibrium and verifying the plasticity conditions, the IPM method will produce a sequence of iterates which satisfy none of the two conditions except at convergence. This specific feature is at the origin of the good robustness of the method in this context [Bilotta et al., 2012].

When discretizing in time the rate equations of plasticity, it is assumed that the evolution is monotonous during the time increment (no elastic unloading occurs) so that

the obtained formulation is that of *holonomic* or *finite-deformation* plasticity. In many cases of interest, one looks for an ultimate state solution under proportional loading for which such formulations are appropriate. Otherwise, one needs to discretize the entire load path into smaller increments over which holonomic plasticity formulations can be adopted. In the former case, we advocate for the use of a single large load increment for computing the elastoplastic ultimate state of the structure if one is not interested in details by the elastic phase. The robustness of the IPM enables to solve this large load step efficiently, contrary to more classical procedures.

The purpose of the present chapter is to provide a general framework for computing efficiently the ultimate state of complex steel assemblies including plasticity. Similarly to the precedent chapter, we will resort to a dual finite-element discretization involving kinematic displacement-based elements as well as static equilibrium-based elements providing, respectively, an upper and lower bounds estimate to the exact solution and, in particular, to the exact structure ultimate load. Both computations can also be used to compute a constitutive error indicator which can be used in a remeshing algorithm. The extension to limit analysis problems is straightforward and various examples, taken from real engineering situations, are presented.

4.2 Variational formulation for elastoplasticity

Since contact conditions can be handled independently from elastoplasticity, in the following, we will only consider one elastoplastic body Ω as seen in Fig. 4.1. In addition to all the notations already introduced in Sec. 3.2.2, the plastic strain tensor will be denoted $\boldsymbol{\varepsilon}^p$.

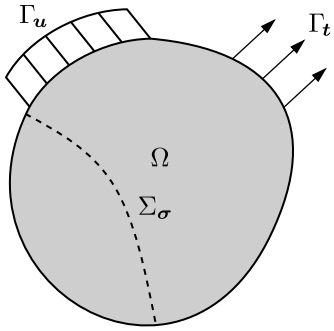


Figure 4.1 – Reference model

In this chapter, we will consider an elastoplastic von Mises material with linear isotropic hardening, although more general hardening models could well be considered. Its free energy density is given by:

$$\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) : \mathbb{D} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \frac{1}{2} E_h p^2 \tag{4.1}$$

where $\mathbb{D}$ is the elasticity stiffness tensor, E_h the hardening modulus and $p(t) = \int_0^t \sqrt{\frac{2}{3}}\|\dot{\boldsymbol{\varepsilon}}^p\|dt$ the accumulated plastic strain where $\|\boldsymbol{a}\| = \sqrt{a_{ij}a_{ij}}$. The plastic dissipation potential is given by:

$$\phi(\dot{\boldsymbol{\varepsilon}}^p) = \begin{cases} \sqrt{\frac{2}{3}}\sigma_0\|\dot{\boldsymbol{\varepsilon}}^p\| & \text{if } \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^p) = 0 \\ +\infty & \text{otherwise} \end{cases} \tag{4.2}$$

corresponding to the von Mises plastic yield criterion:

$$f(\boldsymbol{\sigma}, p) = \sqrt{\frac{3}{2}}\|\operatorname{dev}(\boldsymbol{\sigma})\| - (\sigma_0 + E_h.p) \tag{4.3}$$

Note that one could include a large variety of yield criteria using the framework of generalized standard materials (Some details can be found in appendix D).

We now focus on an incremental formulation between times $[t_n, t_{n+1}]$ among the total time interval $[0, T]$. Knowing all mechanical fields at time t_n, the unknown fields at time t_{n+1} can be obtained from the solution of the following incremental variational formulation [Miehe, 2002; Ortiz and Stainier, 1999]:

$$\underset{\boldsymbol{u},\boldsymbol{\varepsilon}^p,p}{\arg\min} \int_\Omega \int_{t_n}^{t_{n+1}} (\dot{\psi}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p) + \phi(\dot{\boldsymbol{\varepsilon}}^p))dtd\Omega - \int_{t_n}^{t_{n+1}} P_{ext}(\dot{\boldsymbol{u}})dt \tag{4.4}$$

where P_{ext} is the power of external loads.

An approximate solution to (4.4) can be obtained by restricting the above variational formulation to evolutions on $[t_n, t_{n+1}]$ in which external forces and plastic strain rates are assumed to be constant over the time interval:

$$\dot{\boldsymbol{\varepsilon}}^p(t) = \frac{\boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n}{t_{n+1} - t_n} \tag{4.5}$$

where we explicitly wrote the time dependence but not the spatial one ($\dot{\boldsymbol{\varepsilon}}^p, \boldsymbol{\varepsilon}^p_{n+1}, \boldsymbol{\varepsilon}^p_n$ all being tensor fields).

More generally, we get the same final result if restricting to the case of plastic strain rates following radial evolution over $[t_n, t_{n+1}]$ i.e.:

$$\dot{\boldsymbol{\varepsilon}}^p(t) = \dot{\lambda}(t)\frac{\boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n}{\lambda(t_{n+1}) - \lambda(t_n)} \tag{4.6}$$

so that $\int_{t_n}^{t_{n+1}} \dot{\boldsymbol{\varepsilon}}^p(t)dt = \boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n$. In the above, $\lambda(t)$ is any increasing scalar field i.e. $\dot{\lambda}(t) \geq 0$ and obviously encompasses (4.6) for $\lambda(t) = t$. With these assumptions and owing to the fact that ϕ is positively-homogeneous, we get that:

$$\int_{t_n}^{t_{n+1}} \phi(\dot{\boldsymbol{\varepsilon}}^p)dt = \int_{t_n}^{t_{n+1}} \frac{\dot{\lambda}(t)}{\lambda(t_{n+1}) - \lambda(t_n)}\phi(\boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n)dt = \phi(\boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n) \tag{4.7}$$

As a result, an approximate solution to (4.4) is given by:

$$(\boldsymbol{u}_{n+1}, \boldsymbol{\varepsilon}^p_{n+1}, p_{n+1}) = \underset{\boldsymbol{u},\boldsymbol{\varepsilon}^p,p}{\arg\min} \int_\Omega \left(\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p) - \psi(\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}^p_n, p_n) + \phi\left(\boldsymbol{\varepsilon}^p - \boldsymbol{\varepsilon}^p_n\right)\right) d\Omega - P_{ext}(\boldsymbol{u} - \boldsymbol{u}_n) \tag{4.8}$$

The optimality conditions of this incremental problem yield the classical relations of holonomic plasticity. The solution also coincides with that obtained from a classical return mapping procedure using a backward Euler discretization in time. Let us recall again that the obtained solution is only approximate but usually of good quality, especially for a proportional loading.
Denoting by $\Delta\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}_n$ the strain increment and injecting (4.1) and (4.2), the minimization problem becomes:

$$\begin{aligned} \min_{\Delta\boldsymbol{u},\Delta\boldsymbol{\varepsilon}^p,\Delta p} \quad & \int_\Omega \left(\frac{1}{2}(\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) : \mathbb{D} : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) + \boldsymbol{\sigma}_n : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) \right. \\ & \left. + \frac{1}{2}E_h\Delta p^2 + E_h p_n \Delta p + \sqrt{\frac{2}{3}}\sigma_0 \|\Delta\boldsymbol{\varepsilon}^p\| \right) d\Omega - P_{ext}(\Delta\boldsymbol{u}) \\ \text{s.t.} \quad & \operatorname{tr}(\Delta\boldsymbol{\varepsilon}^p) = 0 \end{aligned} \tag{4.9}$$

Since $\dot{p} = \sqrt{\frac{2}{3}}\|\dot{\boldsymbol{\varepsilon}}^p\|$, we also have $\Delta p = \sqrt{\frac{2}{3}}\|\Delta\boldsymbol{\varepsilon}^p\|$ therefore

$$E_h p_n \Delta p + \sqrt{\frac{2}{3}}\sigma_0\|\Delta\boldsymbol{\varepsilon}^p\| = \sqrt{\frac{2}{3}}\sigma_{Y,n}\|\Delta\boldsymbol{\varepsilon}^p\|$$

and

$$\frac{1}{2}E_h(\Delta p)^2 = \frac{1}{3}E_h\|\Delta\boldsymbol{\varepsilon}^p\|^2$$

with $\sigma_{Y,n} = \sigma_0 + E_h.p_n$.

Writing explicitly the link between the displacement increments, the total strain increments and the Dirichlet boundary conditions, we have:

$$\begin{aligned} \min_{\Delta\boldsymbol{u},\,\Delta\boldsymbol{\varepsilon}^p} \quad & \int_\Omega \left[\frac{1}{2}(\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) : \mathbb{D} : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) + (\boldsymbol{\varepsilon}_n - \boldsymbol{\varepsilon}^p_n) : \mathbb{D} : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) \right. \\ & \left. + \frac{1}{3}E_h\|\Delta\boldsymbol{\varepsilon}^p\|^2 + \sqrt{\frac{2}{3}}\sigma_{Y,n}\|\Delta\boldsymbol{\varepsilon}^p\| \right] d\Omega - P_{ext}(\Delta\boldsymbol{u}) && (4.10a) \\ \text{subject to} \quad & \Delta\boldsymbol{\varepsilon} = \boldsymbol{\nabla}^S(\Delta\boldsymbol{u}) && \text{in } \Omega && (4.10b) \\ & \Delta\boldsymbol{u} + \boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} && \text{on } \Gamma_{\boldsymbol{u}} && (4.10c) \\ & \operatorname{tr}(\Delta\boldsymbol{\varepsilon}^p) = 0 && \text{in } \Omega && (4.10d) \end{aligned}$$

This optimization problem fits into the conic programming framework for which interior point methods are well suited. Indeed, one can introduce an auxiliary scalar variable $\Delta\gamma$

such that $\Delta\gamma \geq \|\Delta\boldsymbol{\varepsilon}^p\|$ in order to eliminate the non-smooth objective part as already explained in Sec. 2.3. The previous problem can then be equivalently reformulated as:

$$\min_{\Delta\boldsymbol{u},\Delta\boldsymbol{\varepsilon}^p,\Delta\gamma} \quad \int_\Omega \left[\frac{1}{2}(\Delta\boldsymbol{\varepsilon}-\Delta\boldsymbol{\varepsilon}^p):\mathbb{D}:(\Delta\boldsymbol{\varepsilon}-\Delta\boldsymbol{\varepsilon}^p)+(\boldsymbol{\varepsilon}_n-\boldsymbol{\varepsilon}_n^p):\mathbb{D}:(\Delta\boldsymbol{\varepsilon}-\Delta\boldsymbol{\varepsilon}^p)\right.$$
$$\left.+\frac{1}{3}E_h(\Delta\gamma^2)+\sqrt{\frac{2}{3}}\sigma_{Y,n}\Delta\gamma\right]d\Omega - P_{ext}(\Delta\boldsymbol{u}) \tag{4.11a}$$

$$\text{subject to} \quad \Delta\boldsymbol{\varepsilon} = \boldsymbol{\nabla}^S(\Delta\boldsymbol{u}) \qquad \text{in } \Omega \tag{4.11b}$$
$$\Delta\boldsymbol{u}+\boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} \qquad \text{on } \Gamma_{\boldsymbol{u}} \tag{4.11c}$$
$$\mathrm{tr}(\Delta\boldsymbol{\varepsilon}^p) = 0 \qquad \text{in } \Omega \tag{4.11d}$$
$$\|\Delta\boldsymbol{\varepsilon}^p\| \leq \Delta\gamma \qquad \text{in } \Omega \tag{4.11e}$$

The traceless constraint on $\Delta\boldsymbol{\varepsilon}^p$ can be removed by introducing directly the deviatoric operator dev() in the elastic constitutive relation as shown in Yonekura and Kanno [2012]:

$$\boldsymbol{\sigma}_{n+1} = \boldsymbol{\sigma}_n + \mathbb{D}:(\Delta\boldsymbol{\varepsilon}-\mathrm{dev}(\Delta\boldsymbol{\varepsilon}^p)) \tag{4.12}$$

The previous optimization problem is then changed to:

$$J_{\text{kin}} = \min_{\Delta\boldsymbol{u},\,\Delta\boldsymbol{\varepsilon}^p,\,\Delta\gamma} \quad \int_\Omega \left[\frac{1}{2}(\Delta\boldsymbol{\varepsilon}-\mathrm{dev}(\Delta\boldsymbol{\varepsilon}^p)):\mathbb{D}:(\Delta\boldsymbol{\varepsilon}-\mathrm{dev}(\Delta\boldsymbol{\varepsilon}^p))\right.$$
$$+\boldsymbol{\sigma}_n:(\Delta\boldsymbol{\varepsilon}-\mathrm{dev}(\Delta\boldsymbol{\varepsilon}^p))$$
$$\left.+\frac{1}{3}E_h(\Delta\gamma)^2+\sqrt{\frac{2}{3}}\sigma_{Y,n}\Delta\gamma\right]d\Omega$$
$$-P_{ext}(\Delta\boldsymbol{u}) \tag{4.13a}$$

$$\text{subject to} \quad \Delta\boldsymbol{\varepsilon} = \boldsymbol{\nabla}^S(\Delta\boldsymbol{u}) \qquad \text{in } \Omega \tag{4.13b}$$
$$\Delta\boldsymbol{u}+\boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} \qquad \text{on } \Gamma_{\boldsymbol{u}} \tag{4.13c}$$
$$\|\Delta\boldsymbol{\varepsilon}^p\| \leq \Delta\gamma \qquad \text{in } \Omega \tag{4.13d}$$

Similarly, the dual minimization problem acting on statically admissible stress fields

$\boldsymbol{\sigma} \in SA(\Omega)$ is expressed as follows [Kanno, 2011]:

$$J_{\text{stat}} = \min_{\boldsymbol{\sigma}_{n+1},\, \sigma_{Y,n+1}} \int_{\Omega} \left[\frac{1}{2}(\boldsymbol{\sigma}_{n+1} - \boldsymbol{\sigma}_n) : \mathbb{D}^{-1} : (\boldsymbol{\sigma}_{n+1} - \boldsymbol{\sigma}_n) + \frac{1}{2E_h}(\sigma_{Y,n+1} - \sigma_{Y,n})^2 \right] d\Omega - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}_{n+1}.\boldsymbol{n}).(\boldsymbol{u}_{d,n+1} - \boldsymbol{u}_{d,n}) d\Gamma \tag{4.14a}$$

$$\text{subject to} \quad \text{Div}(\boldsymbol{\sigma}_{n+1}) + \boldsymbol{b}_{n+1} = 0 \quad \text{in } \Omega \tag{4.14b}$$

$$[\![\boldsymbol{\sigma}_{n+1}]\!].\boldsymbol{n} = 0 \quad \text{on } \Sigma_\sigma \tag{4.14c}$$

$$\boldsymbol{\sigma}_{n+1}.\boldsymbol{n} = \boldsymbol{t}_{d,n+1} \quad \text{on } \Gamma_{\boldsymbol{t}} \tag{4.14d}$$

$$\sqrt{\frac{3}{2}} \|\text{dev}(\boldsymbol{\sigma}_{n+1})\| \leq \sigma_{Y,n+1} \quad \text{in } \Omega \tag{4.14e}$$

As a result, both objective functions are opposite $-J_{\text{stat}} = J_{\text{kin}}$ for the true elastoplastic solution. When restricting both minimum problems to finite-dimensional subspace obtained from a finite-element discretization for instance, both objective functions will differ and under careful choice on how discretization is performed, the kinematic approach will yield an upper bound to the true solution, whereas the static approach will yield a lower bound: $-J_{\text{stat},h} \leq -J_{\text{stat}} = J_{\text{kin}} \leq J_{\text{kin},h}$. It must be kept in mind that the discrete optimization problems associated with $J_{\text{stat},h}$ and $J_{\text{kin},h}$ are not dual to each other in the convex optimization sense since they are associated with different discretization strategies. They however offer a bracketing of the true solution since the accuracy of the discretized solution can be compared by computing $J_{\text{kin},h} + J_{\text{stat},h}$ (see chapter 3). We later refer to this quantity as the primal-dual gap i.e. the discretization gap between the primal (kinematic) discretization and the dual (static) discretization.

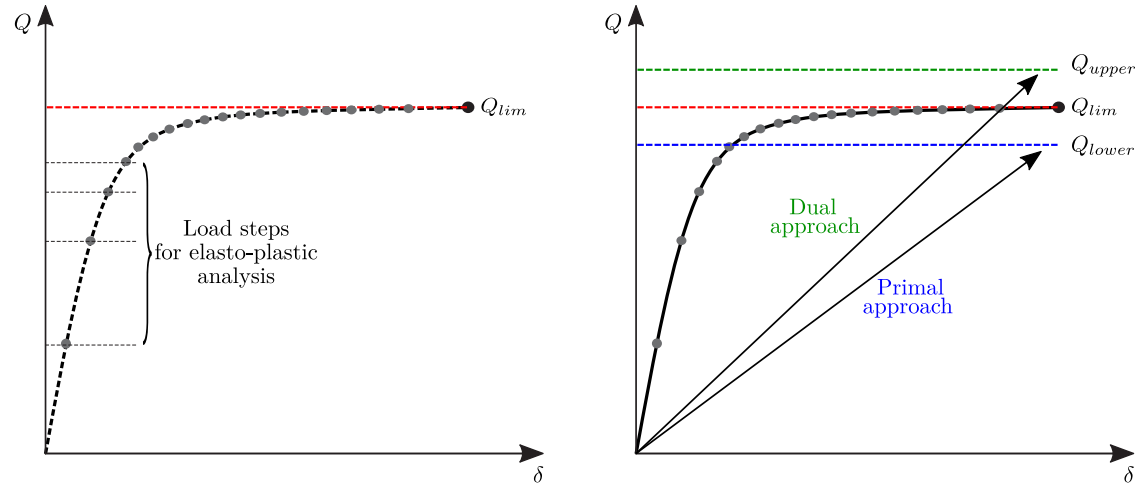
The increment solution of an elastoplastic problem can then be obtained from the resolution of a convex optimization problem in terms of the displacement and plastic strain increments for the kinematic formulation or in terms of the new stress and yield limit at time t_{n+1} for the dual static formulation. With a von Mises material, both problems fall into the class of second-order cone programming problems for which interior-point algorithms are well suited. Let us mention that we can also add unilateral or associated frictional contact conditions to both formulations following the lines of chapter 3 without changing the second-order cone nature of the problem.

4.3 Limit analysis and yield design

4.3.1 Motivation

In various engineering applications, one of the main concerns is to find the ultimate supportable load of a structure under a given set of loading parameters and a local material yield criterion.

Considering an elastoplastic material, determining the ultimate load can be achieved through a series of incremental analyses where the loading parameters are increased until divergence of the solving algorithm as can be seen in Fig. 4.2a. While this approach is general and applicable in most cases, the process can be incredibly time-consuming and difficult to monitor. Full elastoplastic analyses generally require a certain level of mastery of the mechanical theory involved in modeling non-elastic behavior and convergence problems are often faced when trying to push the system to its maximum resisting load. This is mainly due to the solving algorithm, commonly the Newton-Raphson method, which requires the computation of a tangent stiffness matrix at each internal iteration: this matrix tends to become singular when plasticity occurs which in consequence causes the algorithm to diverge.



(a) Elastoplastic load-displacement curve (b) Limit load computed using limit analysis

Figure 4.2 – General idea of limit analysis

Aiming at computing this limit load in a direct fashion, i.e. without having to run a full elastoplastic incremental analysis, is the purpose of limit analysis theory (in the case of perfectly plastic materials) [Hill, 1950] or yield design theory [Salençon, 2013] in a more general framework. In the following sections, the general approaches of the theory of yield design will be presented. It consists of two main theorems or *approaches* as seen in Fig. 4.2b:

1. A primal (static-based) approach which provides a lower bound estimate for the limit load;
2. A dual (kinematic-based) approach which provides an upper bound estimate for the limit load.

4.3.2 Static lower bound approach of yield design

Let $\hat{\boldsymbol{u}}$ be a kinematically admissible (KA) virtual velocity field with the imposed displacements $\boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}}$ i.e. is piecewise continuous and continuously differentiable and

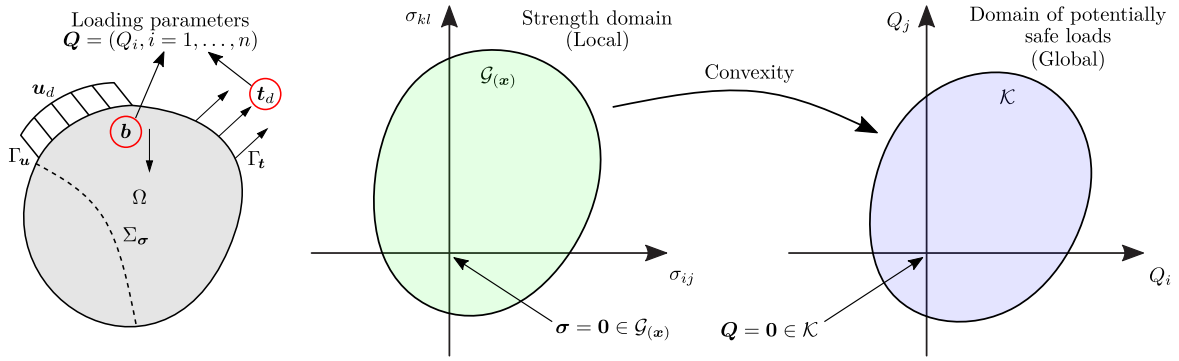


Figure 4.3 – Reference model and definition of the domain of potentially safe loads

such that $\hat{\boldsymbol{u}} = \boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}}$:

$$KA = \left\{ \boldsymbol{u} \in \mathbb{R}^3 \;\middle|\; \boldsymbol{u} = \boldsymbol{u}_d \quad \text{on } \Gamma_{\boldsymbol{u}} \right\} \tag{4.15}$$

$\hat{\boldsymbol{d}} = \boldsymbol{\nabla}^S \hat{\boldsymbol{u}}$ then denotes the linearized virtual strain rate tensor.

Let $\boldsymbol{\sigma}$ be a statically admissible (SA) stress field i.e. it satisfies the local equilibrium equations, traction continuity and traction boundary conditions:

$$SA = \left\{ \boldsymbol{\sigma} \in \mathbb{S} \;\middle|\; \begin{array}{ll} \mathrm{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 & \text{in } \Omega \\ [\![\boldsymbol{\sigma}.\boldsymbol{n}]\!] = 0 & \text{through } \Sigma_{\boldsymbol{\sigma}} \\ \boldsymbol{\sigma}.\boldsymbol{n} = \boldsymbol{t}_d & \text{on } \Gamma_{\boldsymbol{t}} \end{array} \right\} \tag{4.16}$$

Along with the geometrical and loading data, the third set of data needed for yield design is the constitutive material strength properties. At any point $M(\boldsymbol{x})$ of Ω, a convex strength domain $\mathcal{G}(\boldsymbol{x})$ is defined in the 6-dimensional space of the stress tensor $\boldsymbol{\sigma}(\boldsymbol{x})$ as follows:

$$\mathcal{G}(\boldsymbol{x}) = \left\{ \; \boldsymbol{\sigma}(\boldsymbol{x}) \;\middle|\; f(\boldsymbol{\sigma}_{(\boldsymbol{x})}) \leq 0 \; \right\} \tag{4.17}$$

where f corresponds to the classical yield function in the context of plasticity. In general, we will refer to $\mathcal{G}(\boldsymbol{x})$ as the material strength criterion (see Fig. 4.3).

The domain of potentially safe loads $\mathcal{K}$ therefore consists of the set of loads $\boldsymbol{Q}$ such that there exists a stress field satisfying the equilibrium equations and the material strength conditions [Salençon, 2013]:

$$\boldsymbol{Q} \in \mathcal{K} \Longleftrightarrow \left\{ \begin{array}{ll} \exists \boldsymbol{\sigma}_{(\boldsymbol{x})} \text{ S.A. with } \boldsymbol{Q} & \text{(Equilibrium condition)} \\ \boldsymbol{\sigma}_{(\boldsymbol{x})} \in \mathcal{G}_{(\boldsymbol{x})} \; \forall \boldsymbol{x} \in \Omega & \text{(Strength condition)} \end{array} \right. \tag{4.18}$$

The lower bound approach of yield design therefore consists in approximating $\mathcal{K}$ by determining a stress field verifying the equations given in (4.18) for a fixed reference loading direction $\boldsymbol{Q}_0$. One then looks for the maximum value of a load factor α such that $\alpha \boldsymbol{Q}_0 \in \mathcal{K}$. The maximum load factor α^* and can then be obtained by solving a maximization problem given by:

$$\alpha^* = \max\{\alpha\} \text{ such that } \exists \boldsymbol{\sigma}_{(\boldsymbol{x})} \left\{ \begin{array}{l} \boldsymbol{\sigma}_{(\boldsymbol{x})} \text{ S.A. with } \alpha \boldsymbol{Q}_0 \\ \boldsymbol{\sigma}_{(\boldsymbol{x})} \in \mathcal{G}_{(\boldsymbol{x})} \; \forall \boldsymbol{x} \in \Omega \end{array} \right. \tag{4.19}$$

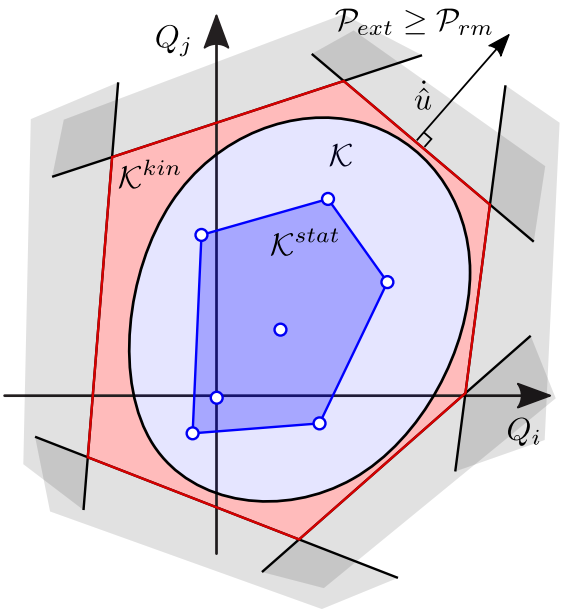


Figure 4.4 – Schematic representation of the interior and exterior approximation of the domain $\mathcal{K}$

If restricting the above maximization to a subspace of the statically admissible stress fields, e.g. one obtained from a stress-based finite-element discretization, this approach will therefore provide a lower bound $\alpha_{lower} \leq \alpha^*$ to the true limit load factor α^*. The set of these lower bound approximations for varying loading directions provide an interior approximation $\mathcal{K}_{stat} \subseteq \mathcal{K}$ to the safe load domain $\mathcal{K}$ (see Fig. 4.4).

4.3.3 Kinematic upper bound approach of yield design

Deriving the dual approach of the theory of yield design starts by writing the virtual work principle, corresponding to the dualization of the equilibrium equations given in (4.16), as follows:

$$\begin{aligned} \mathcal{P}_{int(\hat{\boldsymbol{u}})} &= \int_\Omega \boldsymbol{\sigma} : \hat{\boldsymbol{d}} d\Omega + \int_\Sigma (\boldsymbol{\sigma}.\boldsymbol{n})[\![\hat{\boldsymbol{u}}]\!] d\Gamma \\ &= \int_\Omega \boldsymbol{b}\hat{\boldsymbol{u}} d\Omega + \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}.\boldsymbol{n})\boldsymbol{u}_d d\Gamma + \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d \hat{\boldsymbol{u}} d\Gamma = \mathcal{P}_{ext(\hat{\boldsymbol{u}})} \end{aligned} \tag{4.20}$$

Using the properties of convex sets, one can rewrite the convex set $\mathcal{G}$ using support functions:

$$\boldsymbol{\sigma} \in \mathcal{G} \Rightarrow \left\{ \begin{array}{lll} \boldsymbol{\sigma} : \hat{\boldsymbol{d}} & \leq \sup\{\boldsymbol{\sigma} : \hat{\boldsymbol{d}} \mid \boldsymbol{\sigma} \in \mathcal{G}\} & =: \Pi(\hat{\boldsymbol{d}}) \quad \forall \hat{\boldsymbol{d}} \\ (\boldsymbol{\sigma}.\boldsymbol{n})[\![\hat{\boldsymbol{u}}]\!] & \leq \sup\{(\boldsymbol{\sigma}.\boldsymbol{n})[\![\hat{\boldsymbol{u}}]\!] \mid \boldsymbol{\sigma} \in \mathcal{G}\} & =: \Pi(\boldsymbol{n}, [\![\hat{\boldsymbol{u}}]\!]) \quad \forall [\![\hat{\boldsymbol{u}}]\!] \end{array} \right. \tag{4.21}$$

The work of internal forces $\mathcal{P}_{int(\hat{\boldsymbol{u}})}$ can then be bounded from above using these support functions as follows:

$$\begin{aligned} \Rightarrow \mathcal{P}_{int(\hat{\boldsymbol{u}})} &= \int_\Omega (\boldsymbol{\sigma} : \hat{\boldsymbol{d}}) d\Omega + \int_\Sigma (\boldsymbol{\sigma}.\boldsymbol{n}) [\![\hat{\boldsymbol{u}}]\!] d\Sigma \\ &\leq \int_\Omega \Pi(\hat{\boldsymbol{d}}) d\Omega + \int_\Sigma \Pi(\boldsymbol{n}, [\![\hat{\boldsymbol{u}}]\!]) d\Sigma = \mathcal{P}_{rm(\hat{\boldsymbol{u}})} \end{aligned} \tag{4.22}$$

where the last term is corresponds to the maximum resisting work (or the plastic dissipation in a limit analysis context) of the structure for a given virtual field $\hat{\boldsymbol{u}}$.
One can then show that we have:

$$\boldsymbol{Q} \in \mathcal{K} \quad \Leftrightarrow \forall \hat{\boldsymbol{u}} \text{ K.A.} \quad \mathcal{P}_{ext(\hat{\boldsymbol{u}})} \leq \mathcal{P}_{rm(\hat{\boldsymbol{u}})} \tag{4.23}$$

Note that for a fixed reference load $\boldsymbol{Q}_0$, we can therefore characterize the maximum load factor α associated with $\boldsymbol{Q}_0$ by:

$$\alpha \leq \alpha^* \quad \Leftrightarrow \forall \hat{\boldsymbol{u}} \text{ K.A.} \quad \alpha \mathcal{P}_{ext,0(\hat{\boldsymbol{u}})} \leq \mathcal{P}_{rm(\hat{\boldsymbol{u}})} \tag{4.24}$$

i.e.

$$\alpha^* = \min_{\hat{\boldsymbol{u}} \text{ K.A.}} \frac{\mathcal{P}_{rm(\hat{\boldsymbol{u}})}}{\mathcal{P}_{ext,0(\hat{\boldsymbol{u}})}} \tag{4.25}$$

The equation $\mathcal{P}_{ext(\hat{\boldsymbol{u}})} \leq \mathcal{P}_{rm(\hat{\boldsymbol{u}})}$ defines a hyperplane subdividing the space into two half-spaces, one of them containing the strength domain $\mathcal{K}$ and the origin as shown in Fig. 4.3. Therefore the use of a virtual velocity field approximates the limit load from the exterior.

The upper bound approach of yield design therefore consists in approximating the convex domain $\mathcal{K}$ by minimizing the maximum resisting work for various virtual velocity fields (which can be interpreted as the structure collapse mechanisms). When adding a normalization condition $\mathcal{P}_{ext,0(\hat{\boldsymbol{u}})} = 1$, the maximum load factor α^* can be obtained by solving the following minimization problem:

$$\begin{aligned} \alpha^* = \quad & \min_{\hat{\boldsymbol{u}}} \quad \mathcal{P}_{rm(\hat{\boldsymbol{u}})} \\ & \text{s.t.} \quad \hat{\boldsymbol{u}} \in \text{ K.A.} \\ & \qquad \mathcal{P}_{ext,0(\hat{\boldsymbol{u}})} = 1 \end{aligned} \tag{4.26}$$

Again, if restricting the above minimization to a subspace of the kinematically admissible fields, e.g. one obtained from a displacement-based finite-element discretization, this approach will therefore provide an upper bound $\alpha_{upper} \geq \alpha^*$ to the true limit load factor α^*. The set of these upper bound approximations for varying loading directions provide an exterior approximation $\mathcal{K}_{kin} \supseteq \mathcal{K}$ to the safe load domain $\mathcal{K}$ (see Fig. 4.4).

The numerical resolution of finite-element discretized yield analysis can also be efficiently achieved using interior point methods for conic programming [Lyamin and Sloan, 2002a,b; Makrodimopoulos and Martin, 2006, 2007].

4.4 Modification for yield analysis optimization problems

One can easily obtain these two lower and upper bound problems from formulations (4.13a) and (4.14a) by considering a rigid perfectly plastic material i.e. $\mathbb{D} \to \infty$ hence $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^p$ and $E_h = 0$. Besides, in order to obtain the maximum load factor, we consider that the loading consist of a fixed part (e.g. the body forces $\boldsymbol{b}$) and a reference loading scaled by a load multiplier α (e.g. for the surface tractions $\alpha \boldsymbol{t}$) which we want to maximize. In this case, the work of external loads take the following form:

$$\alpha P_{ext}(\boldsymbol{u}) + P_{ext,0}(\boldsymbol{u}) = \alpha \int_{\Gamma_t} \boldsymbol{t} \cdot \boldsymbol{u} d\Gamma + \int_{\Omega} \boldsymbol{b} \cdot \boldsymbol{u} d\Omega$$

Assuming that the the deformations are small at incipient collapse and that the material can be modeled with sufficient accuracy using perfect plasticity and an associated flow rule, the upper is obtained by simply scaling the plastic dissipation, also called in the context of yield analysis the maximum resisting energy, with respect to the rate of work done by the reference tractions $\boldsymbol{t}_d$ i.e. $P_{ext}(\boldsymbol{u}) = 1$.

Considering only homogeneous Dirichlet boundary conditions for simplicity, problems (4.13a) and (4.14a) respectively become:

$$\begin{aligned} \alpha_{upper} = \min_{\boldsymbol{u},\, \boldsymbol{\varepsilon}^p,\, \gamma} \quad & \int_{\Omega} \sqrt{\frac{2}{3}} \sigma_0 \gamma d\Omega - P_{ext,0}(\boldsymbol{u}) & \\ \text{s.t.} \quad & \boldsymbol{\varepsilon}^p = \boldsymbol{\nabla}^S \boldsymbol{u} & \text{in } \Omega \\ & \boldsymbol{u} = 0 & \text{on } \Gamma_{\boldsymbol{u}} \\ & P_{ext}(\boldsymbol{u}) = 1 & \\ & \mathrm{tr}(\boldsymbol{\varepsilon}^p) = 0 & \text{in } \Omega \\ & \|\boldsymbol{\varepsilon}^p\| \leq \Delta\gamma & \text{in } \Omega \end{aligned} \tag{4.27}$$

$$\begin{aligned} \alpha_{lower} = \max_{\alpha,\, \boldsymbol{\sigma}} \quad & \alpha & \\ \text{s.t.} \quad & \mathrm{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 & \text{in } \Omega \\ & [\![\boldsymbol{\sigma}]\!].\boldsymbol{n} = 0 & \text{on } \Sigma_\sigma \\ & \boldsymbol{\sigma}.\boldsymbol{n} = \alpha \boldsymbol{t}_d & \text{on } \Gamma_{\boldsymbol{t}} \\ & \sqrt{\frac{3}{2}} \|\mathrm{dev}(\boldsymbol{\sigma})\| \leq \sigma_0 & \text{in } \Omega \end{aligned} \tag{4.28}$$

in which the optimal objective values respectively produce an upper and lower bound estimate of the true collapse load multiplier α^+: $\alpha_{\text{lower}} \leq \alpha^+ \leq \alpha_{\text{upper}}$.

4.5 Finite-element discretizations and solution procedure

In this section, we briefly describe the different finite-element discretizations that we can use for solving problem (4.13a) or (4.14a) and (4.27) or (4.28).

4.5.1 Kinematic approach

4.5.1.1 Continuous discretization

Standard continuous displacement-type discretization with 10-node quadratic tetrahedra are used to solve the upper bound problem (4.13a). The element degrees of freedom (DoF) as presented in Fig. 4.5 are the following:

- $3 \times 10 = 30$ nodal displacement increments $\Delta \boldsymbol{u}$;
- $6 \times n_G$ plastic strain increments $\Delta \boldsymbol{\varepsilon}^p$ at each integration point;
- n_G scalar auxiliary variables $\Delta \gamma$ at each integration point;

where n_G is the number of quadrature points per tetrahedron for numerical integration (minimum 4).

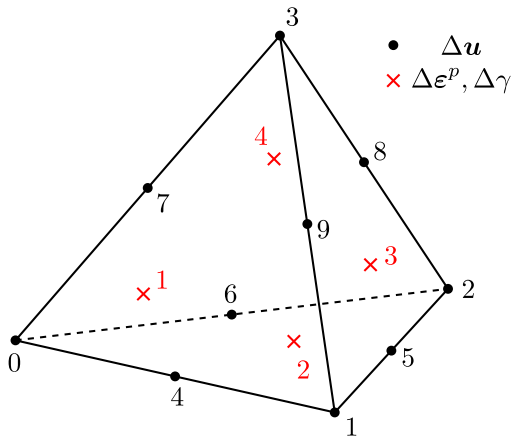


Figure 4.5 – Kinematic finite-element: 10-node quadratic tetrahedron

Let $\boldsymbol{x}$ be the aggregation of of a single tetrahedron DoF i.e. for $n_G = 4$:

$$\boldsymbol{x} = (30\,\Delta \boldsymbol{u}, 24\,\Delta \boldsymbol{\varepsilon}^p, 4\Delta\gamma)$$

Let $\boldsymbol{B}$ be the strain-displacement matrix operator and $\mathbb{K}$ the projector over deviatoric tensors such that $\mathrm{dev}(\Delta \boldsymbol{\varepsilon}^p) = \mathbb{K} : \Delta \boldsymbol{\varepsilon}^p$. The elementary stiffness matrix for each tetrahedron can then be calculated taking into account the traceless condition for $\Delta \boldsymbol{\varepsilon}^p$

as follows:

$$\boldsymbol{K}_e = \left[\begin{array}{c|c|c} \boldsymbol{B}^T\mathbb{D}\boldsymbol{B} & -\boldsymbol{B}^T\mathbb{D}\mathbb{K} & 0 \\ \hline -\mathbb{K}^T\mathbb{D}\boldsymbol{B} & \mathbb{K}^T\mathbb{D}\mathbb{K} & 0 \\ \hline 0 & 0 & \frac{2}{3}E_h \end{array}\right] \tag{4.29}$$

Problem (4.13a) can then be expressed after assembling the global stiffness matrix $\hat{\boldsymbol{K}}$, the nodal plastic dissipation vector $\hat{\boldsymbol{d}}$ and the nodal force vector $\hat{\boldsymbol{f}}$ as a conic minimization problem using the SOCP formalism as follows:

$$\begin{array}{lll} \min\limits_{\hat{\boldsymbol{x}}} & \frac{1}{2}\Delta\hat{\boldsymbol{x}}^T\hat{\boldsymbol{K}}\Delta\hat{\boldsymbol{x}}^T + \hat{\boldsymbol{K}}\hat{\boldsymbol{x}}_n\Delta\hat{\boldsymbol{x}}^T + \hat{\boldsymbol{d}}^T\Delta\hat{\boldsymbol{x}}^T - \hat{\boldsymbol{f}}_{n+1}^T\Delta\hat{\boldsymbol{x}}^T & \\ \text{s.t.} & \Delta\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}_n = \hat{\boldsymbol{u}}_{d,n+1} & \text{on } \Gamma_{\boldsymbol{u}} \\ & \|\Delta\boldsymbol{\varepsilon}^p\| \leq \Delta\gamma & \text{at each integration point} \end{array} \tag{4.30}$$

in which the primal optimization variables are $\hat{\boldsymbol{x}} = (\Delta\hat{\boldsymbol{u}}, \Delta\hat{\boldsymbol{\varepsilon}}^p, \Delta\gamma)$. Let us mention that, for the above formulation to yield a true upper bound of the exact response, quadrature points (i.e. the location at which plastic flow rule will be enforced) have to be located at the tetrahedra vertices (see Makrodimopoulos and Martin [2007]). However, in practice, choosing standard Gauss points yields a very close response.

When implementing the algorithm, the local nature of $\Delta\hat{\boldsymbol{\varepsilon}}^p$ and $\Delta\gamma$ (i.e. they are expressed at the quadrature nodes) enables us to perform a static condensation on the element level by eliminating these variables thus further reducing the system's size which will finally be of the same size of standard Newton-Raphson stiffness matrix and greatly improving the computational performance.

As mentioned previously, contact constraints can be added and enforced at all 6 nodes of each face following chapter 3. After introducing a proper change of variable, the contact conditions can be easily transformed into a standard second-order cone constraint using a Lorentz cone. Contact condition can therefore be considered in incremental form by introducing the following constraints:

$$\Delta\hat{\boldsymbol{g}} = \hat{\boldsymbol{g}}_n + [\![\Delta\hat{\boldsymbol{u}}]\!] \quad \text{on } \Gamma_{\boldsymbol{c}} \tag{4.31a}$$

$$\hat{\boldsymbol{g}}_{n+1} \in \mathcal{K}_\mu^\circ \quad \text{on } \Gamma_{\boldsymbol{c}} \tag{4.31b}$$

4.5.1.2 Discontinuous mesh

The plastic condition is known for inducing volumetric locking problems, even for quadratic tetrahedra. One possible remedy consists in considering a fully discontinuous displacement interpolation and accounting for the displacement jump contribution to the plastic dissipation potential:

$$\phi([\![\Delta \boldsymbol{u}]\!]) = \begin{cases} \sqrt{\frac{1}{3}}\sigma_0 \|[\![\Delta \boldsymbol{u}.\boldsymbol{t}]\!]\| & \text{if } [\![\Delta \boldsymbol{u}.\boldsymbol{n}]\!] = 0 \\ +\infty & \text{otherwise} \end{cases} \tag{4.32}$$

Note that only tangential discontinuities are allowed and that the full displacement jump contributes to the plastic potential, meaning that no discontinuity will occur during the purely elastic phase. Similarly to the bulk plastic dissipation, an auxiliary scalar variable $\Delta\zeta$ must be introduced such that $\Delta\zeta \geq \|[\![\Delta \boldsymbol{u}.\boldsymbol{t}]\!]\|$ in order to fit into the conic formalism.

The introduction of discontinuities prevents volumetric locking and also improves the discretization accuracy. It can be compared to a *discontinuous Galerkin* approach where the interpolated displacement field is piecewise-continuous and interface interaction laws are included in the plastic domain. However, it introduces additional degrees of freedom since each tetrahedron displacements are now independent. This implies higher computational times which can reach 4 to 10 times those of the continuous kinematic approach.

4.5.1.3 Hybrid formulation

One other remedy for mitigating locking effects is to resort to a so-called mixed, or hybrid, discretization in which the traceless condition (4.11d) is enforced only on average on each element Ω^e, instead of being enforced at each quadrature point as it should be. Note that, in this case, the bounding character of the kinematic approach is inevitably lost.

By including such a local condition as (4.33d), a piecewise-constant Lagrange multiplier is associated with the constraint. This Lagrange multiplier bears striking similarities with an additional degree of freedom corresponding to the P_0 interpolation of the pressure $p = \frac{1}{3}\operatorname{tr}(\boldsymbol{\sigma})$ in u-p mixed finite elements when treating quasi-incompressible materials.

$$\min_{\Delta \boldsymbol{u}, \Delta \boldsymbol{\varepsilon}^p, \Delta\gamma} \int_\Omega \left[\frac{1}{2}(\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) : \mathbb{D} : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) + (\boldsymbol{\varepsilon}_n - \boldsymbol{\varepsilon}_n^p) : \mathbb{D} : (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p) \right.$$
$$\left. + \frac{1}{3}E_h(\Delta\gamma^2) + \sqrt{\frac{2}{3}}\sigma_{Y,n}\Delta\gamma \right] d\Omega - P_{ext}(\Delta \boldsymbol{u}) \tag{4.33a}$$

$$\text{subject to} \quad \Delta\boldsymbol{\varepsilon} = \boldsymbol{\nabla}^S(\Delta \boldsymbol{u}) \qquad \text{in } \Omega \tag{4.33b}$$
$$\Delta \boldsymbol{u} + \boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} \qquad \text{on } \Gamma_{\boldsymbol{u}} \tag{4.33c}$$
$$\textstyle\int_{\Omega^e} \operatorname{tr}(\Delta\boldsymbol{\varepsilon}^p) = 0 \qquad \forall \Omega^e \subset \Omega \tag{4.33d}$$
$$\|\Delta\boldsymbol{\varepsilon}^p\| \leq \Delta\gamma \qquad \text{in } \Omega \tag{4.33e}$$

4.5.2 Static approach

For the static approach and as explained previously in chapter 3, we use improved equilibrium elements based on the formulations given in Pian [1964], Kempeneers et al. [2009] and de Almeida and Maunder [2017] but accounting for some changes to fit into the convex optimization framework. To simplify notations isotropic hardening is not considered and the time increments are dropped so that problem (4.14a) becomes:

$$\min_{\boldsymbol{\sigma}} \quad \int_{\Omega} \frac{1}{2}\boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma} d\Omega - \int_{\Gamma_{\boldsymbol{u}}} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{u}_d d\Gamma \tag{4.34a}$$

$$\text{s.t.} \quad \operatorname{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 \qquad \text{in } \Omega \tag{4.34b}$$

$$[\![\boldsymbol{\sigma}]\!].\boldsymbol{n} = 0 \qquad \text{on } \Sigma_{\boldsymbol{\sigma}} \tag{4.34c}$$

$$\boldsymbol{\sigma}.\boldsymbol{n} = \boldsymbol{t}_d \qquad \text{on } \Gamma_{\boldsymbol{t}} \tag{4.34d}$$

$$\sqrt{\frac{3}{2}} \| \operatorname{dev}(\boldsymbol{\sigma}) \| \leq \sigma_0 \qquad \text{in } \Omega \tag{4.34e}$$

When introducing Lagrange multipliers associated with the linear constraint (4.34c), these can be interpreted as displacements of each internal facet of the mesh. Similarly, Lagrange multipliers associated with constraint (4.34d) can be interpreted as displacement of faces lying on the boundary $\Gamma_{\boldsymbol{t}}$. Denoting by $\boldsymbol{u}$ both internal and external face displacements and prescribing explicitly $\boldsymbol{u} = \boldsymbol{u}_d$ on $\Gamma_{\boldsymbol{u}}$, the previous minimization problem can be rewritten as follows:

$$\min_{\boldsymbol{\sigma}, \boldsymbol{u}} \quad \int_{\Omega} \frac{1}{2}\boldsymbol{\sigma} : \mathbb{D}^{-1} : \boldsymbol{\sigma} d\Omega - \int_{\Sigma_{\boldsymbol{\sigma}}} [\![\boldsymbol{\sigma}]\!].\boldsymbol{n}.\boldsymbol{u} d\Gamma - \int_{\Gamma_{\boldsymbol{t}}} \boldsymbol{t}_d.\boldsymbol{u} d\Gamma \tag{4.35a}$$

$$\text{s.t.} \quad \operatorname{Div}(\boldsymbol{\sigma}) + \boldsymbol{b} = 0 \qquad \text{in } \Omega \tag{4.35b}$$

$$\boldsymbol{u} = \boldsymbol{u}_d \qquad \text{on } \Gamma_{\boldsymbol{u}} \tag{4.35c}$$

$$\sqrt{\frac{3}{2}} \| \operatorname{dev}(\boldsymbol{\sigma}) \| \leq \sigma_0 \qquad \text{in } \Omega \tag{4.35d}$$

Next, appropriate interpolations are chosen for $\boldsymbol{\sigma}$ and $\boldsymbol{u}$. To do so, the stress distribution inside an element is chosen in order to satisfy local equilibrium equations with body forces [de Almeida and Maunder, 2017; Kempeneers et al., 2009; Pian, 1964] such as:

$$\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{q} + \boldsymbol{\sigma}_{\boldsymbol{b}}(\boldsymbol{x}) \tag{4.36}$$

where $\boldsymbol{S}$ is a matrix of independent shape functions of degree p, the columns of which $\boldsymbol{S}_k$ all verify $\operatorname{Div}(\boldsymbol{S}_k) = 0$. They can be obtained either from Morera [1892] or Maxwell [1866] stress potentials. In the following, we consider only linear stress shape functions ($p = 1$) i.e. $\boldsymbol{S}$ is of dimension (6×21). The 21 variables $\boldsymbol{q}$ are the generalized stresses and

are not related to any nodes but to the entire finite-element volume. $\boldsymbol{\sigma_b}$ is a particular solution for the equilibrium equations with $\boldsymbol{b} \neq 0$ and taken as equal to zero in absence of body forces. For simplicity, we will consider the latter case in the following. The expressions for the polynomial functions $\boldsymbol{S}$ can be found in de Almeida and Pereira [1996].

Since, stress fields vary in a linear fashion, equations (4.34c) and (4.34d) can be enforced exactly using a linear interpolation for the face displacement Lagrange multipliers $\boldsymbol{u}$. For face j, the nodal displacement vector is $\boldsymbol{u}^{(j)} = \{u_{x1}, u_{y1}, \ldots, u_{z3}\}$ so that the displacement of an arbitrary point of face j is:

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{P}_j(\boldsymbol{x})\boldsymbol{u}^{(j)} \tag{4.37}$$

with $\boldsymbol{P}_j$ being linear shape functions. We later denote by $\boldsymbol{u}_e = \{\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(4)}\}$ the nodal displacements of all faces and by $\boldsymbol{P}$ the corresponding aggregated shape function matrix. The final DoF vector $\boldsymbol{x}$ of the tetrahedral element as seen in Fig. 4.6 is the following:

$$\boldsymbol{x} = (21\,\boldsymbol{q}, 36\,\boldsymbol{u})$$

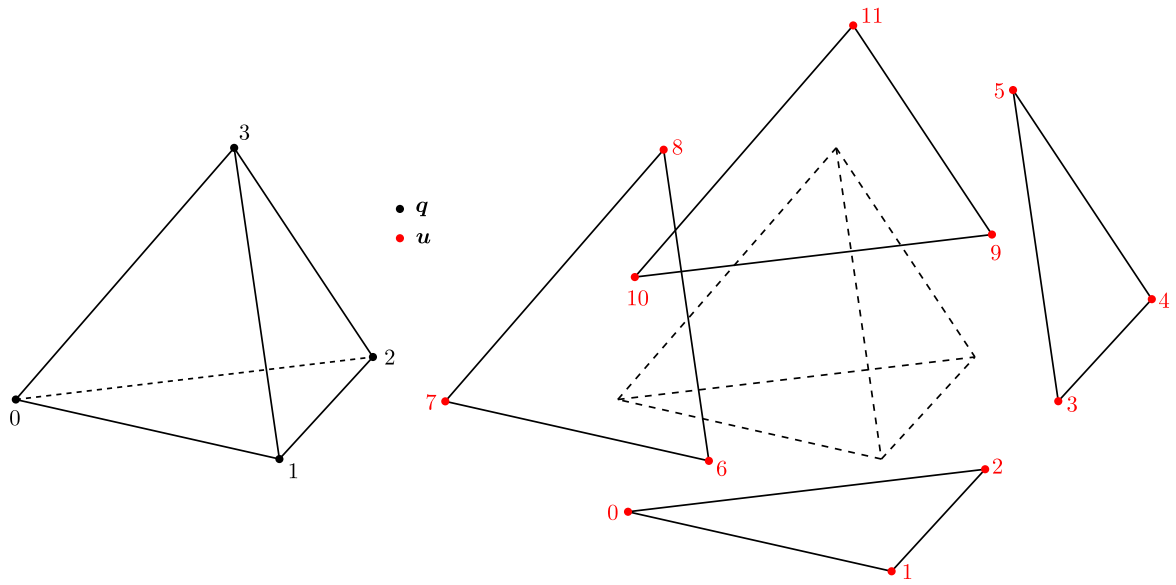


Figure 4.6 – 4-node linear static tetrahedron with additional DoF on the faces

Using all of these, the elementary matrix for each tetrahedron can be calculated using the first variation of the objective function (4.35a):

$$\delta\boldsymbol{q}^T\boldsymbol{F}\boldsymbol{q} + \delta\boldsymbol{q}^T\boldsymbol{H}^T\boldsymbol{u}_e + \delta\boldsymbol{u}_e^T\boldsymbol{H}\boldsymbol{q} = \delta\boldsymbol{u}_e^T\boldsymbol{f} \qquad \forall(\delta\boldsymbol{q}, \delta\boldsymbol{u}_e) \tag{4.38}$$

where

$$\boldsymbol{F} = \int_{\Omega_e} \boldsymbol{S}^T \mathbb{D}^{-1} \boldsymbol{S} d\Omega \tag{4.39}$$

$$\boldsymbol{H} = \int_{\partial\Omega_e} \boldsymbol{P}^T \boldsymbol{N} \boldsymbol{S} d\Gamma \tag{4.40}$$

$$\boldsymbol{f} = \int_{\partial\Omega_e \cap \Gamma_{\boldsymbol{t}}} \boldsymbol{P}^T \boldsymbol{t}_d d\Gamma \tag{4.41}$$

with $\boldsymbol{N}$ being the normal projection matrix operator on a facet of normal $\boldsymbol{n}$.
Eq. (4.38) can be rewritten in the following matrix form:

$$\left[\begin{array}{c|c} \boldsymbol{F} & \boldsymbol{H}^T \\ \hline \boldsymbol{H} & \boldsymbol{0} \end{array}\right] \left\{\begin{array}{c} \boldsymbol{q} \\ \boldsymbol{u}_e \end{array}\right\} = \left\{\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{f} \end{array}\right\} \tag{4.42}$$

The obtained matrix is singular. As shown in Kempeneers et al. [2009], it has 6 zero eigenvalues corresponding to the rigid body movements, and 9 more corresponding to spurious kinematic modes (SKM).

These spurious modes can be eliminated using the *super-element* method, consisting in considering one tetrahedral element as an assembly of four sub-tetrahedra based on each four faces and sharing a common inner vertex lying at the center of the macro-tetrahedron as can be seen in Fig. 4.7. This combination results in 4 tetrahedrons ($4 \times 21 = 84$ generalized stresses $\hat{\boldsymbol{q}}$), 4 external faces ($4 \times 9 = 36$ external face displacements $\hat{\boldsymbol{u}}_e$) and 6 internal faces ($6 \times 9 = 54$ internal faces displacements $\hat{\boldsymbol{u}}_i$) yielding the corresponding elementary matrix:

$$\left[\begin{array}{c|c|c} \hat{\boldsymbol{F}} & \hat{\boldsymbol{H}}_i^T & \hat{\boldsymbol{H}}_e^T \\ \hline \hat{\boldsymbol{H}}_i & \boldsymbol{0} & \boldsymbol{0} \\ \hline \hat{\boldsymbol{H}}_e & \boldsymbol{0} & \boldsymbol{0} \end{array}\right] \left\{\begin{array}{c} \hat{\boldsymbol{q}} \\ \hat{\boldsymbol{u}}_i \\ \hat{\boldsymbol{u}}_e \end{array}\right\} = \left\{\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{0} \\ \hat{\boldsymbol{f}}_e \end{array}\right\} \tag{4.43}$$

in which the $\hat{\star}$ symbol is used to denote aggregated matrix/vector over the corresponding degrees of freedom.

The final element matrix is obtained by performing static condensation over the internal degrees of freedom $\hat{\boldsymbol{u}}_i$ by taking advantage that $\boldsymbol{F}$ is invertible:

$$\left[\begin{array}{c|c} \hat{\boldsymbol{F}} & \hat{\boldsymbol{H}}_c \\ \hline \hat{\boldsymbol{H}}_e & \boldsymbol{0} \end{array}\right] \left\{\begin{array}{c} \hat{\boldsymbol{q}} \\ \hat{\boldsymbol{u}}_e \end{array}\right\} = \left\{\begin{array}{c} \boldsymbol{0} \\ \hat{\boldsymbol{f}}_e \end{array}\right\} \tag{4.44}$$

where

$$\hat{\boldsymbol{H}}_c = \hat{\boldsymbol{H}}_e^T - \hat{\boldsymbol{H}}_i^T . (\hat{\boldsymbol{H}}_i \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_i^T)^{-1} . (\hat{\boldsymbol{H}}_i \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_e^T) \tag{4.45}$$

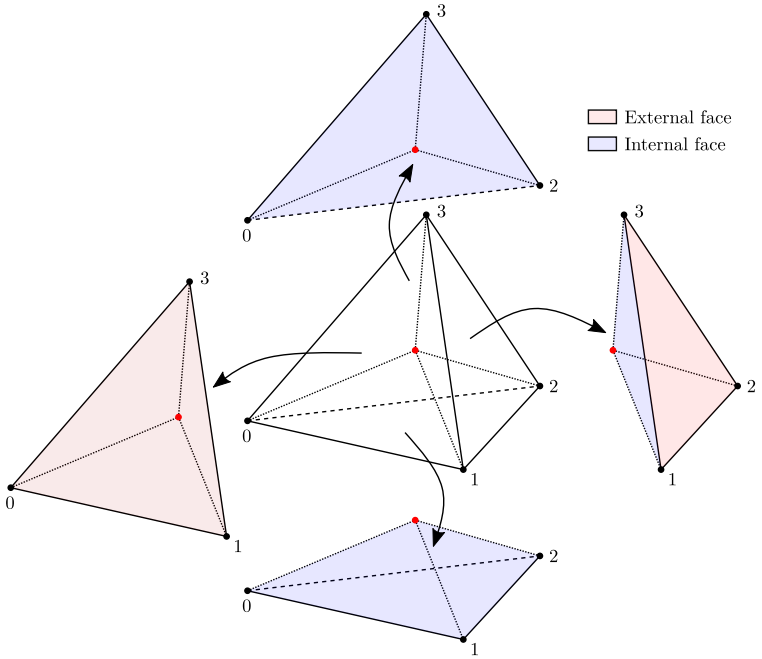


Figure 4.7 – Subdivision on the original tetrahedron into 4 sub-elements

Assembling the elementary matrices to a global matrix $\hat{\boldsymbol{A}}$ leads to a minimization problem of the following form:

$$\min_{\hat{\boldsymbol{x}}} \quad \frac{1}{2}\hat{\boldsymbol{x}}^T\hat{\boldsymbol{A}}\hat{\boldsymbol{x}} - \hat{\boldsymbol{f}}^T\hat{\boldsymbol{x}} \tag{4.46a}$$

$$\text{s.t.} \quad \hat{\boldsymbol{u}}_e = \boldsymbol{u}_d \qquad \text{on } \Gamma_{\boldsymbol{u}} \tag{4.46b}$$

$$\sqrt{\frac{3}{2}}\|\mathbb{K}\boldsymbol{S}\hat{\boldsymbol{q}}\| \leq \sigma_0 \qquad \text{at each vertex} \tag{4.46c}$$

where $\hat{\boldsymbol{x}}$ is the global vector of degrees of freedom containing both $\hat{\boldsymbol{q}}$ and $\hat{\boldsymbol{u}}_e$ for all elements. Note that since the stress field is piecewise linear inside each sub-tetrahedra, it is enough to check the plasticity criterion at each vertex of each sub-tetrahedra to guarantee the lower bound status of the static approach. Including back time increments and body forces will only change expressions of right-hand side terms whereas including hardening will have to include the new yield stress as an additional optimization variable with its corresponding quadratic energy term.

Similarly, contact constraint can be easily added to the above conic optimization problem. Using the Lagrange multipliers interpreted as displacements, one can easily include the contact conditions by introducing the following additional constraints:

$$\hat{\boldsymbol{g}} = [\![\hat{\boldsymbol{u}}_e]\!] \qquad \text{on } \Gamma_{\boldsymbol{c}} \tag{4.47a}$$

$$\hat{\boldsymbol{g}} \in \mathcal{K}_\mu^\circ \qquad \text{on } \Gamma_{\boldsymbol{c}} \tag{4.47b}$$

Note that in practice another static condensation is made at the element level before assembling the global matrix $\hat{\boldsymbol{A}}$ in order to further reduce the size of the system (4.44) and to restore the symmetry of the matrix and that is by eliminating $\boldsymbol{q}$ which yields the following:

$$\hat{\boldsymbol{K}}_c \hat{\boldsymbol{u}}_e = \hat{\boldsymbol{f}}_e \tag{4.48}$$

where $\hat{\boldsymbol{K}}_c$ is equivalent to a stiffness matrix:

$$\hat{\boldsymbol{K}}_c = \hat{\boldsymbol{H}}_e \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_c = \hat{\boldsymbol{H}}_e \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_e^T - (\hat{\boldsymbol{H}}_i \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_e^T)^T . (\hat{\boldsymbol{H}}_i \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_i^T)^{-1} . (\hat{\boldsymbol{H}}_i \hat{\boldsymbol{F}}^{-1} \hat{\boldsymbol{H}}_e^T) \tag{4.49}$$

4.6 Illustrative examples

4.6.1 Elastoplastic analysis of a cylinder subject to torsion

This first example consists in a theoretical validation of the algorithm using a comparison of the obtained numerical results and some analytical results. The example considers a simple steel cylinder of radius $R = 5\,\text{cm}$ and of height $H = 20\,\text{cm}$ subject to torsion. The material is supposed elastic perfectly plastic with a Young modulus $E = 210\,000\,\text{MPa}$, a Poisson ratio of $\nu = 0.3$, and a elastic yield limit $\sigma_0 = 275\,\text{MPa}$. To simplify notations, let $G = E/2(1+\nu)$ be the shear modulus and $k = \sigma_0/\sqrt{3}$ the shear yield limit. The boundary conditions described in the cylindrical coordinate system

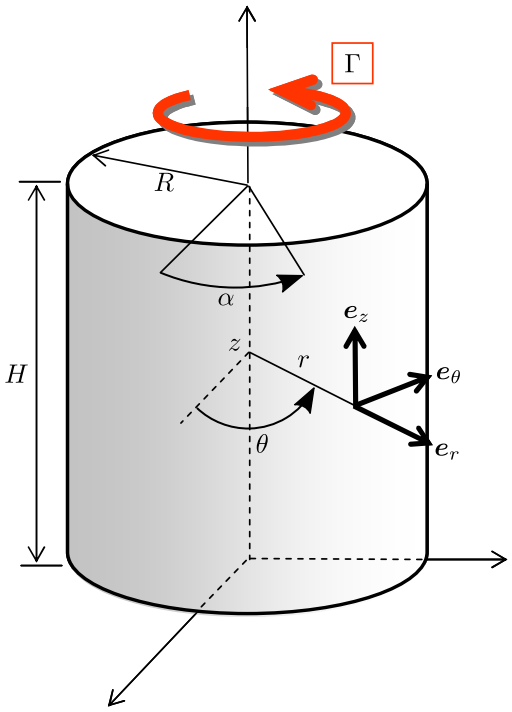


Figure 4.8 – Description of the steel cylinder

$(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z)$ are the following:

1. the bottom face is fully clamped i.e.

$$\boldsymbol{u}(r,\theta,z=0,t)=\boldsymbol{0}$$

2. the top face is subject to an imposed torsion around the z axis parameterized using the angle $\alpha(t)$ such that:

$$\boldsymbol{u}(r,\theta,z=H,t)=\alpha(t)r\boldsymbol{e}_\theta$$

3. the lateral surface has zero tractions i.e.

$$\boldsymbol{t}_d(r=R,\theta,z,t)=\boldsymbol{0}$$

We are interested in computing the evolution of the resulting torque $\Gamma(t)$ as a function of the torsion angle of the top surface $\alpha(t)$ where:

$$\Gamma(t)=\int\limits_{S(z=H)}\left[(r\boldsymbol{e}_r)\wedge(\boldsymbol{\sigma}.\boldsymbol{e}_z)\right].\boldsymbol{e}_z\,d\Gamma \tag{4.50}$$

The closed-form analytical solution (in a small strain setting) for an imposed rotation evolving from $\alpha=0$ to $\alpha=\alpha_u$ and them from $\alpha=\alpha_u$ to $\alpha=\alpha_r$ with α_r corresponding to full unloading can be described by the following phases (see Fig. 4.9):

1. A fully elastic phase:

$$\begin{cases}\forall\alpha(t)\leq\alpha^e=\frac{kH}{GR}\\ \Gamma(\alpha)=\frac{G\alpha(t)\pi R^4}{2H}\leq\Gamma^e=\Gamma(\alpha=\alpha^e)=\frac{k\pi R^3}{2}\end{cases} \tag{4.51}$$

2. An elastoplastic phase where a part of the cylinder from $r=0$ to $r=\rho(t)$ is still elastic and the rest has fully yielded:

$$\begin{cases}\forall\alpha\geq\alpha^e=\frac{kH}{GR}\\ \rho(t)=\frac{kH}{G\alpha(t)}=R\frac{\alpha^e}{\alpha(t)}\\ \Gamma(\alpha)=\Gamma^l\left[1-\frac{1}{4}\left(\frac{\alpha^e}{\alpha(t)}\right)^3\right]\leq\Gamma^l=\Gamma(\rho(t)\to 0)=\frac{2k\pi R^3}{3}\end{cases} \tag{4.52}$$

3. A limit state when all the cylinder has yielded i.e. $\rho(t)\to 0$. In this state, and using a perfectly plastic material, $\alpha(t)$ is not bounded however the torsion moment reaches its limit value:

$$\Gamma^l=\Gamma(\rho(t)\to 0)=\frac{2k\pi R^3}{3} \tag{4.53}$$

4. A residual phase which can be calculated using elastic unloading with a slope parallel to the elastic part from 0 to Γ^e. If the torsion couple reached a value Γ^u corresponding to an angle α^u before complete unloading, the residual torsion couple will be 0 and the residual angle is given by:

$$\alpha^r=\alpha^u-\frac{2H}{G\pi R^4}\Gamma^u \tag{4.54}$$

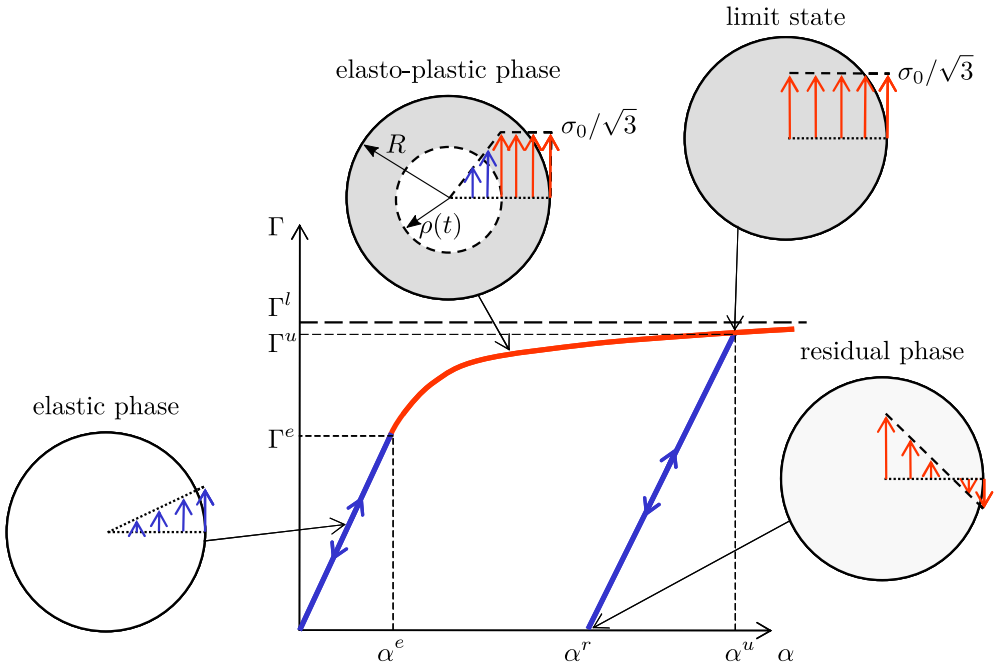


Figure 4.9 – Description of the different phases of the elastoplastic solution

Using the upper and lower bound approaches, the algorithm is able to attain very satisfying bounds for the limit value Γ^l:

$$\begin{cases} \Gamma^l_{analytical} = 0.0355\,\text{MN m} \\ \Gamma^l_{lower} = 0.0344\,\text{MN m} \\ \Gamma^l_{upper} = 0.0350\,\text{MN m} \end{cases} \tag{4.55}$$

The small difference is actually due to the mesh resolution since the steel cylinder is discretized using 9974 tetrahedrons and the circular section is approximated using an equivalent inscribed polygon. The discretized circular section using this mesh constitutes only 99.35% of the total section while the gap between the upper bound and the analytical value is 99.24%.

As for the elastoplastic phase, the interior point method allows us to determine the complete moment–torsion angle curve as seen in Fig. 4.10. The curve is drawn using 10 load steps. However, it would have been possible to attain the fully yielded state using one large step as will be seen in the following example which is a great advantage of using the IPM method. The results show a good match with the analytical results with less than 2% gap between the results. Again, the difference between the upper bound curve and the analytical one is mainly due to the discretization error as explained earlier. Fig. 4.11 to 4.15 show the different phases identified in Fig. 4.9 and numbered from 1 to 5 in Fig. 4.10

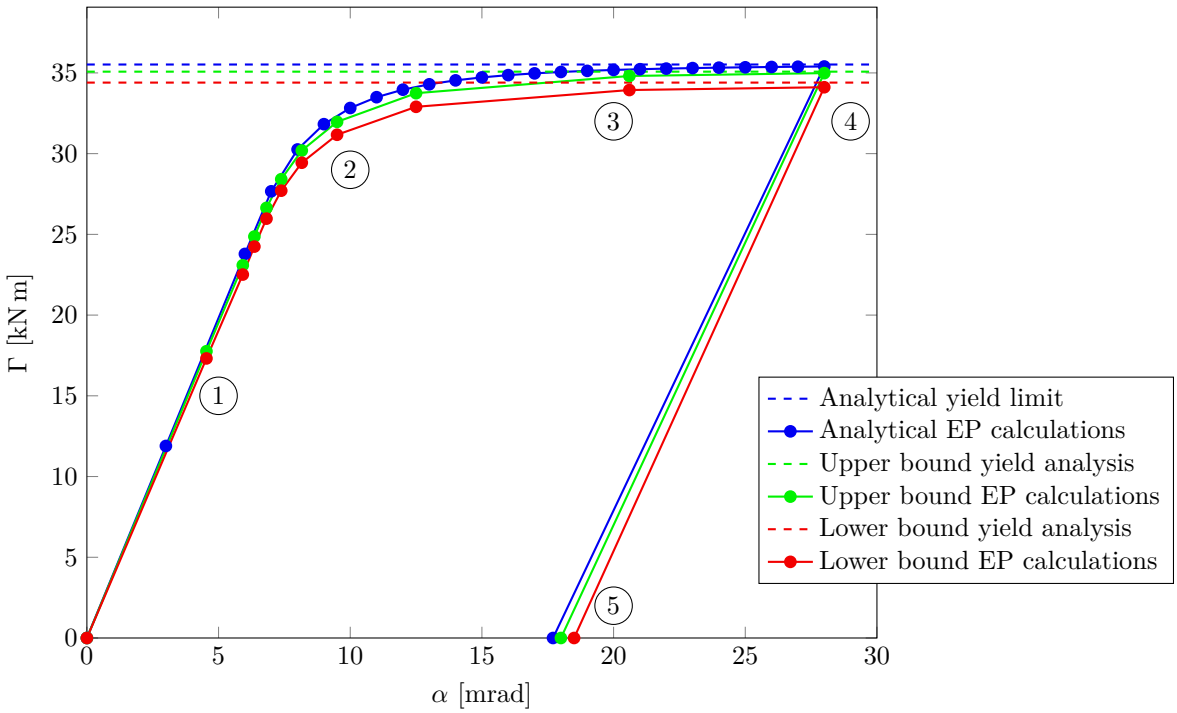


Figure 4.10 – Moment–torsion angle curve

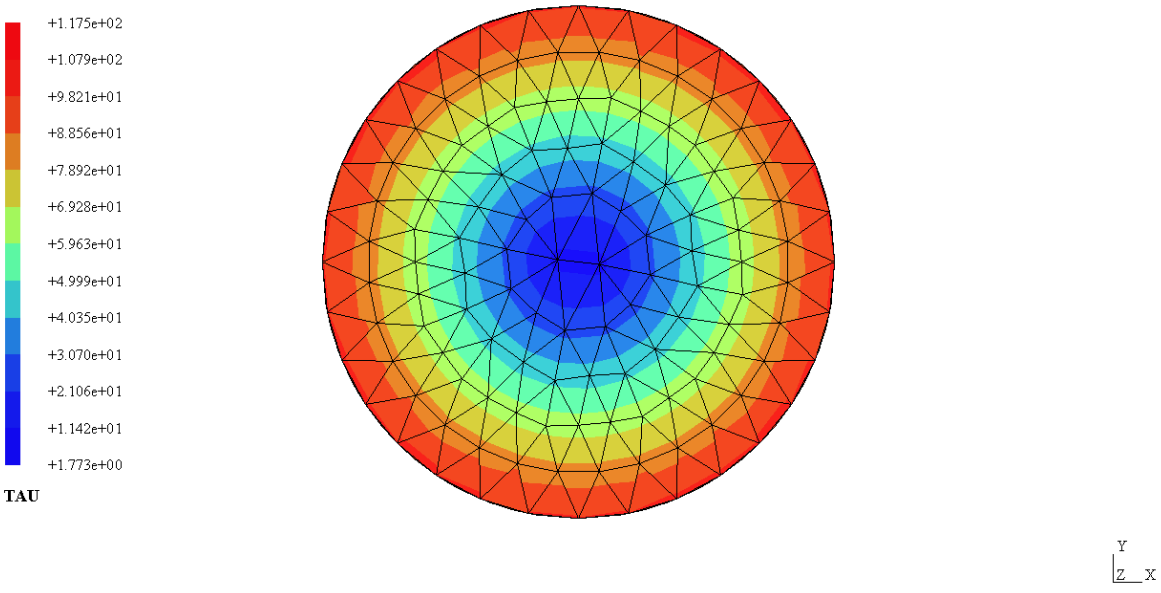


Figure 4.11 – Shear stress isovalues in elastic phase (state 1 in Fig. 4.10)

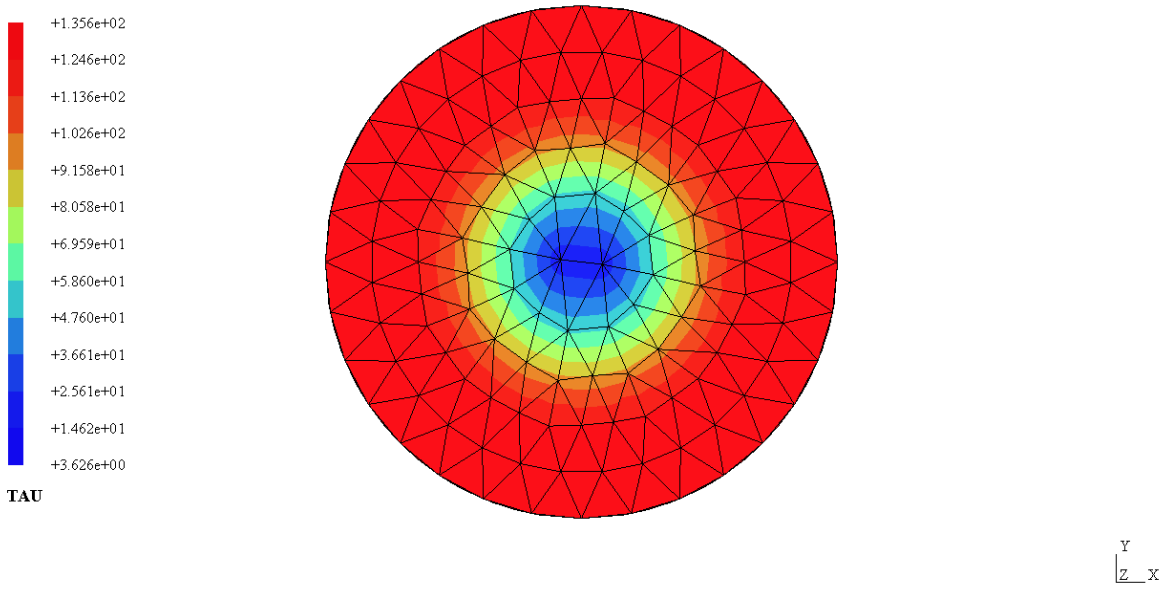


Figure 4.12 – Shear stress isovalues in elastoplastic phase (state 2 in Fig. 4.10)

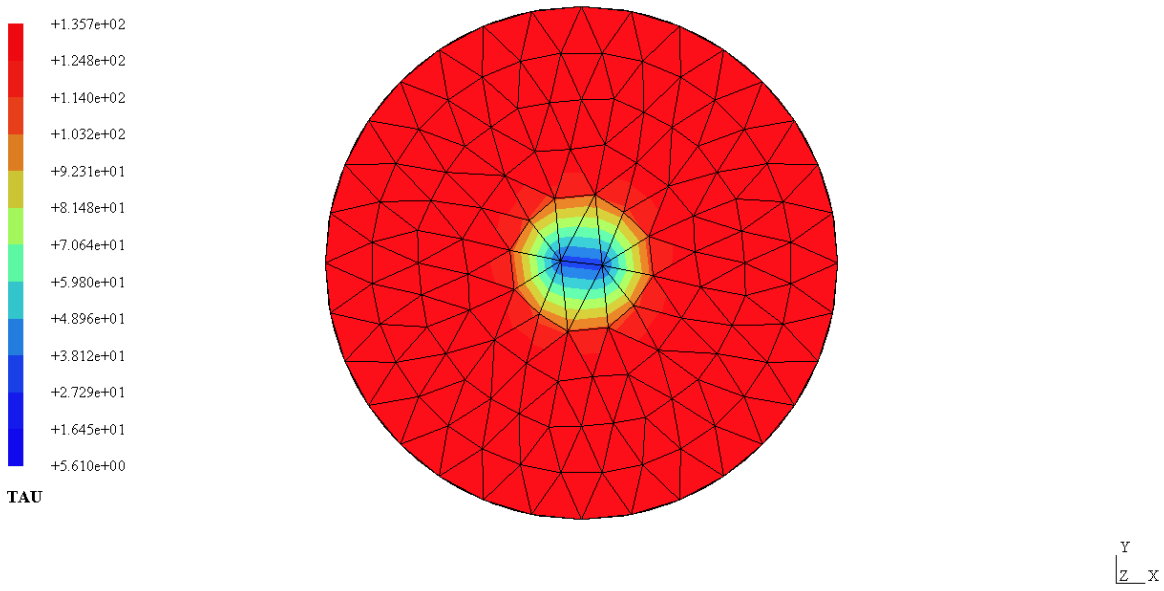


Figure 4.13 – Shear stress isovalues in elastoplastic phase (state 3 in Fig. 4.10)

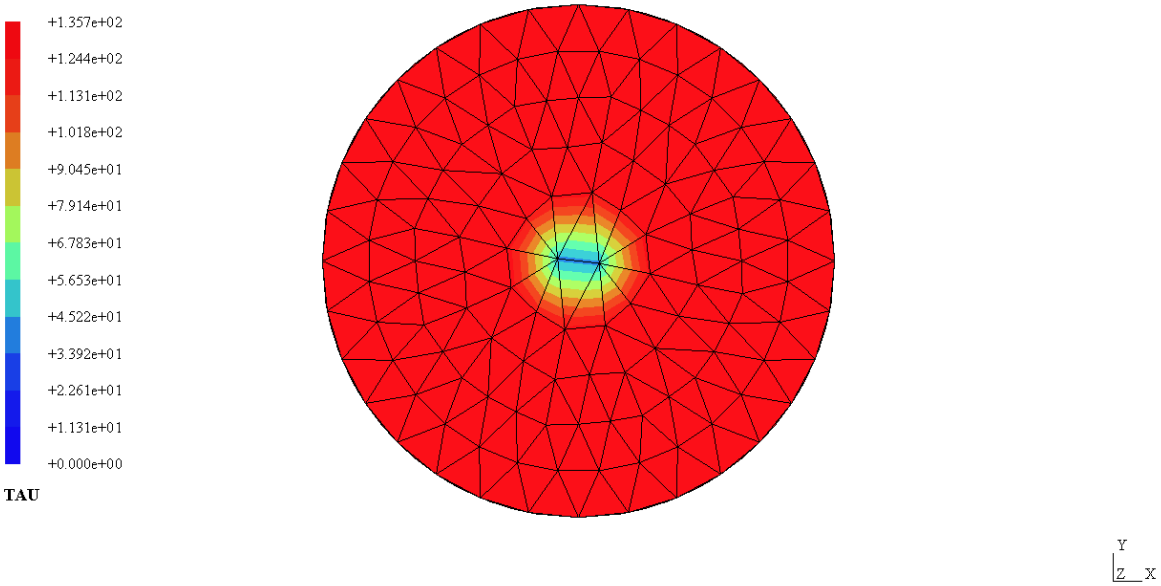


Figure 4.14 – Shear stress isovalues in limit phase (state 4 in Fig. 4.10)

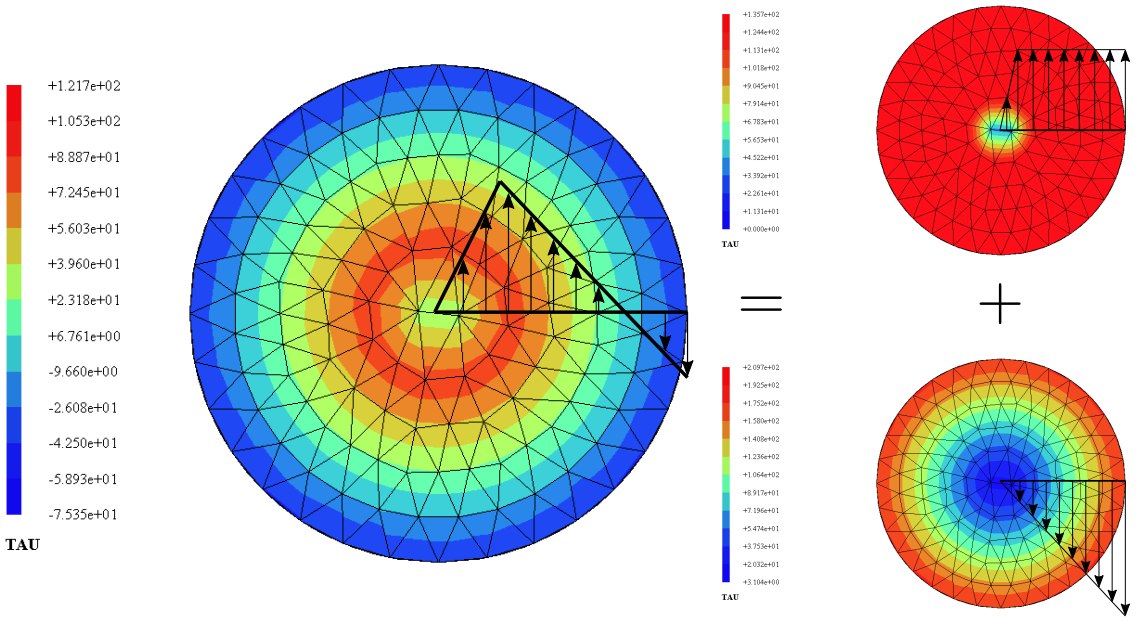


Figure 4.15 – Shear stress isovalues in residual phase (state 5 in Fig. 4.10)

4.6.2 Performance and result comparison with Abaqus

4.6.2.1 Convergence analysis

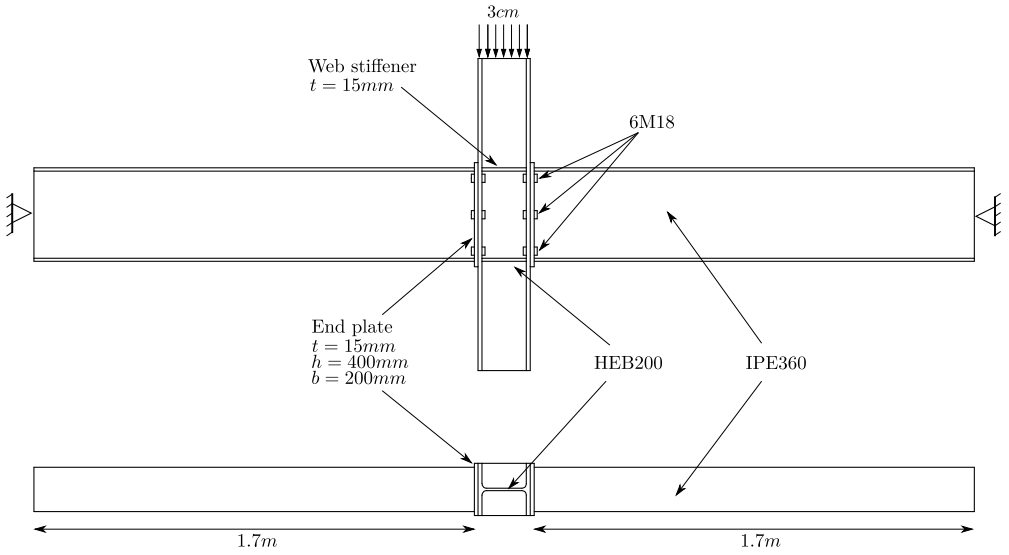
This second example is the continuation of the one found in chapter 3 with some minor changes in boundary conditions. The aim of this example is to assess the proposed procedure computational cost and quality of the obtained results by comparing it with computations made using Abaqus. Abaqus uses classical approaches such as a Newton-Raphson algorithm to solve the global non-linear balance equations along with a return mapping algorithm to correct local plasticity and either a penalty or an augmented Lagragian method to tackle contact conditions. Conversely, our proposed method based on convex optimization and the interior point algorithm treats all three aspects simultaneously, with plasticity and contact conditions expressed using conic constraints. Aiming at a fair comparison, the same series of 3 iteratively refined meshes is used in all the studies, both for our implementation and for Abaqus computations. Since Abaqus offers continuous displacement-based elements only, the calculation times are compared with respect to our kinematic approach using continuous interpolation. In order to closely compare the numerical performance of both algorithms, all the calculations were made using the same computer (see Tab. 4.1). OpenMP technology was used to parallelize over 8 threads.

Fig. 4.16a gives a general description of the model and the three mesh iterations are represented on Fig. 4.16b to 4.16d. The considered example consists of a HEB200 central column with two IPE360 beams attached over the flanges using welded end-plates and bolts. The end-plates have a 15 mm thickness and 6 M18 bolts are used to connect each beam. We suppose that the bolt hole is equal to its diameter and, to prevent rigid body motions, one of the bolts heads is glued to the plate. For the HEB column, the thickness is 12.7 mm, the same as the IPE flange thickness. The steel grade for all beams and plates is S275 ($f_y = 275\,\text{MPa}$), the bolts are of grade 8.8 ($f_y = 800\,\text{MPa}$) with no initial stress and the yield limit of the welds is taken equal to $f_y = 400\,\text{MPa}$, no hardening being considered here. In this example, we will only consider unilateral contact conditions.

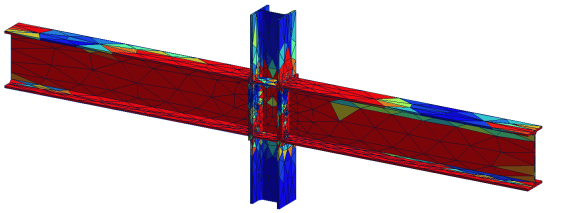
A 3 cm vertical displacement is prescribed over the top section of the HEB column and the average displacements of the two end sections of the IPE are blocked.

Table 4.1 – Computer specifications

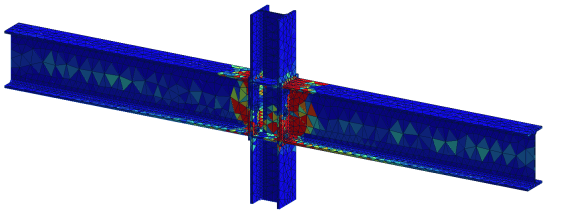
Processor	Intel core i7-4700MQ
Number of cores	4
Number of threads	8
Base frequency	2.40 GHz
Maximum frequency	3.40 GHz
Cache	6 MB –SmartCache
RAM	16.0 GB (15.7 GB usable)



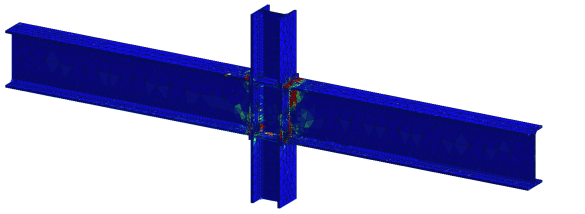
(a) Description of the used model



(b) Initial mesh



(c) First remesh iteration



(d) Second remesh iteration

Figure 4.16 – Description of the model, initial and adapted meshes (isovalues represent the local computed error)

Tab. 4.2 shows the CPU times for the different analyses using the 3 different meshes: the kinematic IPM method and Abaqus penalty and augmented Lagrangian methods for treating contact. In this first case, we aim at testing the robustness of the algorithms regarding large load steps. For this purpose, one displacement increment of 3 cm of vertical displacement has been applied to the column. Each solver, using its default options, is given enough time to converge. While with the interior point method, convergence was always ensured with a steady number of iterations $(18-21)$, Abaqus could not find a solution in a single load-step and relied on its automatic load incrementation to find a sequence of converged increments until reaching the final prescribed displacement. The number of increments needed depends on the type of approach to solve contact: the penalty approach (PEN) needed 4 to 10 load subdivisions to converge whereas the augmented Lagrangian (AL) required 10 to 13. As a result, although for each increment, the number of iterations to reach convergence is usually smaller than the IPM, in total, the final number of iterations is much higher than the IPM. This obviously has an impact on the total computation cost since one iteration has a similar cost for all methods. One can also note that the IPM method scales well with the system size since the number of iterations only weakly increases. The IPM is largely comparable to Abaqus PEN approach in terms of CPU times with even a small speed-up factor, and compared to the AL approach, a speed-up factor of up to 4.1 can be reached for the last mesh. We therefore obtain similar conclusions in presence of plasticity as those obtained previously in chapter 3 when considering contact as the only non-linearity.

One should note that the dual solution requires as much time to calculate as the primal one since the linear system is roughly the same size. The cost of the error calculation is a simple post-processing step which takes less than 1% of the time needed for the whole resolution which is negligible. As regards the kinematic approach with internal discontinuities, the number of DoF is greater than its continuous counterpart. The solving time is therefore much larger but the results are of better quality. In particular, it is known that volumetric locking appears when dealing with incompressible plasticity. Ten-noded tetrahedra suffer from such a numerical difficulty and hybrid formulations must therefore be used to mitigate the locking effect. On the contrary, discontinuous interpolations are free of any locking problem.

Table 4.2 – CPU times and speed-up factors of the IPM over Abaqus augmented Lagrangian approach and penalty approach for example 2

Remesh iteration	**Mesh Size**	**IPM kinematic approach** (s)[N_{inc}](N_{iter})	**Abaqus AL approach** (s)[N_{inc}](N_{iter})	**Speed-up factor**	**Abaqus Penalty approach** (s)[N_{inc}](N_{iter})	**Speed-up factor**
0	10045	37.5 [1] (18)	137.0 [10] (75)	3.6	45.0 [4] (45)	1.2
1	34017	140.3 [1] (20)	517.0 [13] (93)	3.7	319.0 [9] (55)	2.3
2	63137	243.8 [1] (21)	992.0 [13] (90)	4.1	536.0 [10] (52)	2.2

Global convergence levels with mesh refinement can be assessed using figures 4.17 and 4.18. The influence of mesh size over the relative difference between the static and

kinematic IP approaches has been represented in Fig. 4.17. We can notice that the discontinuous kinematic approach offers better convergence gaps. A relative difference of nearly less than 10% can be reached with only 2 remesh steps. Regarding the energy values and reaction forces shown in Fig. 4.18, we can see that the IPM results coincide with Abaqus AL approach which is known to provide quality results. As for the PEN approach, the values are quite different and that can be explained by some violations of contact constraints as shown in chapter 3

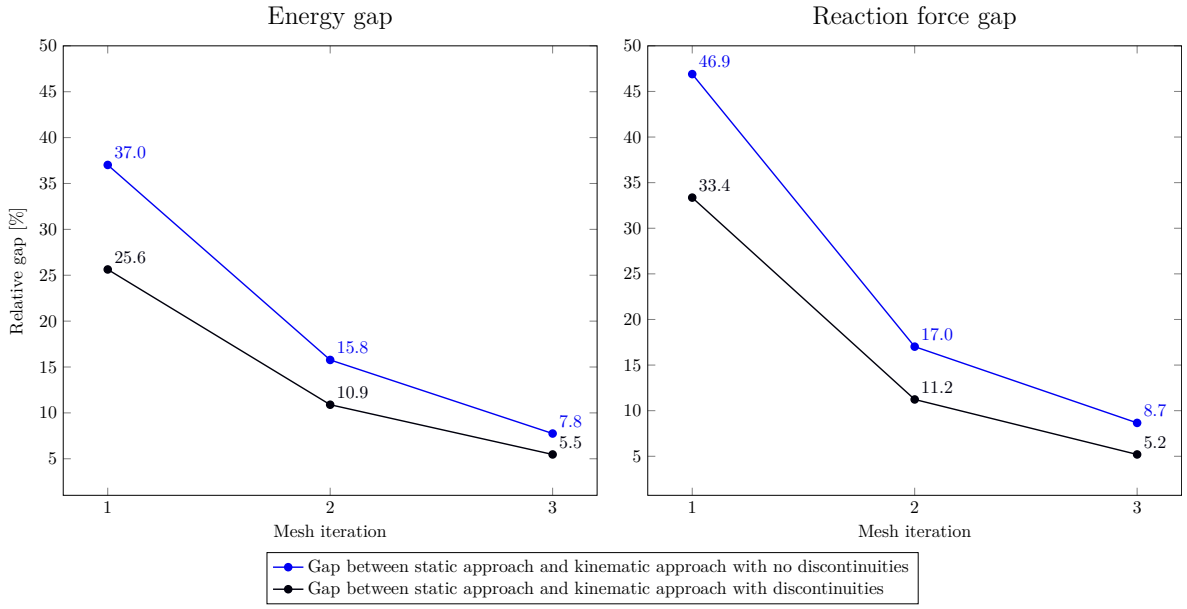


Figure 4.17 – Convergence analysis

4.6.2.2 Force to displacements curve

The final mesh obtained from the convergence analysis is used to draw the typical force–displacement curves for the structural system. The prescribed displacement varies from 0 to 3 cm and the total vertical reaction force is then calculated. Fig. 4.19 shows the different curves obtained using the developed approaches and using Abaqus. The adequacy between solutions is clear: Abaqus C3D10I elements provide the same solution as the continuous kinematic approach, the C3D10H hybrid elements provide the same solution as our own implementation of the continuous hybrid approach; all of these solutions are located over the discontinuous kinematic solution which has an upper bound status, itself being larger than the static solution which has a lower bound status.

Besides, assuming a radial loading path, any load state can be computed from the elastoplastic problems using different load steps strategies. Using the IPM, one can have different incrementation strategies where the engineer can reached the desired stress state using 3, 2 or even one load step as can be seen in Fig. 4.20. It is remarkable that using only one load step does not induce any time discretization error when comparing to the computed reaction force with that obtained when using 13 load steps.

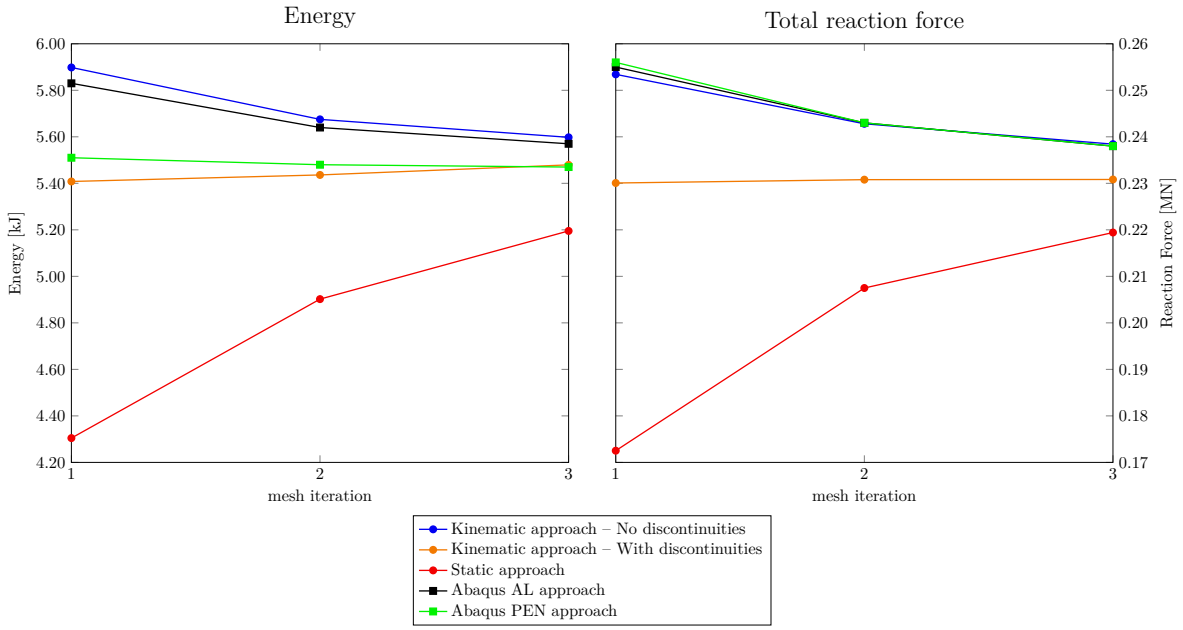


Figure 4.18 – Values comparison with Abaqus

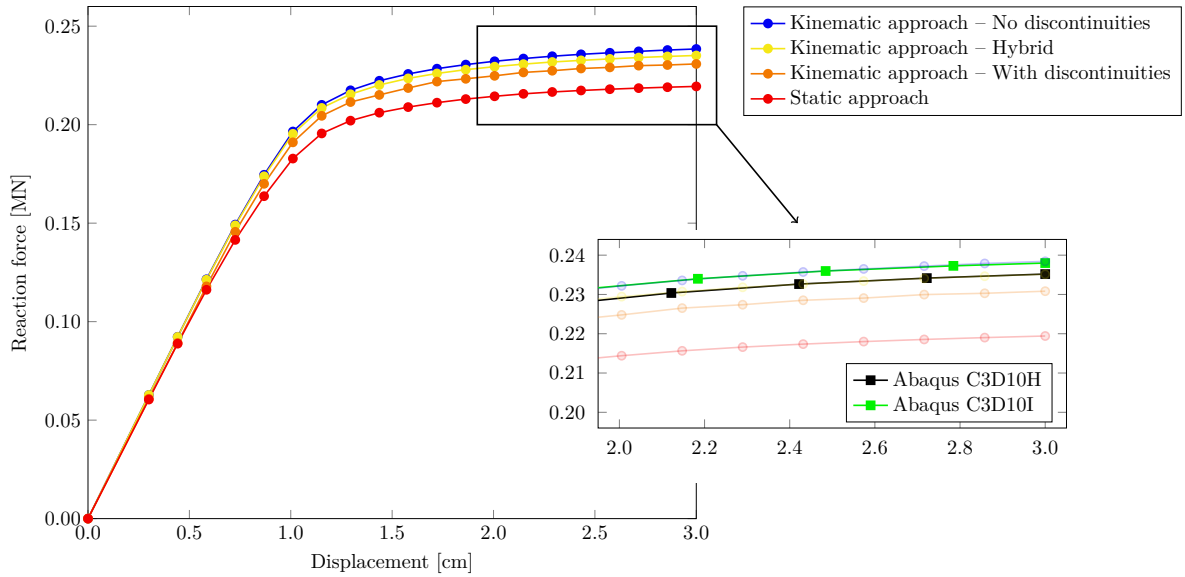
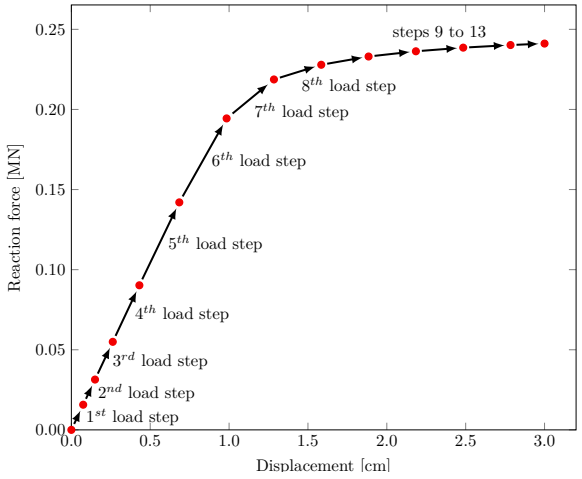
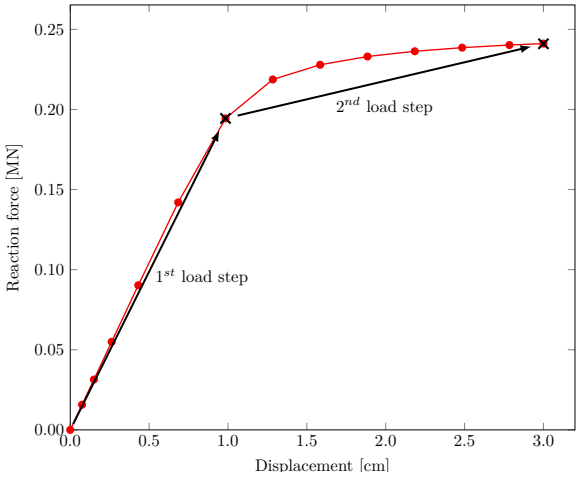


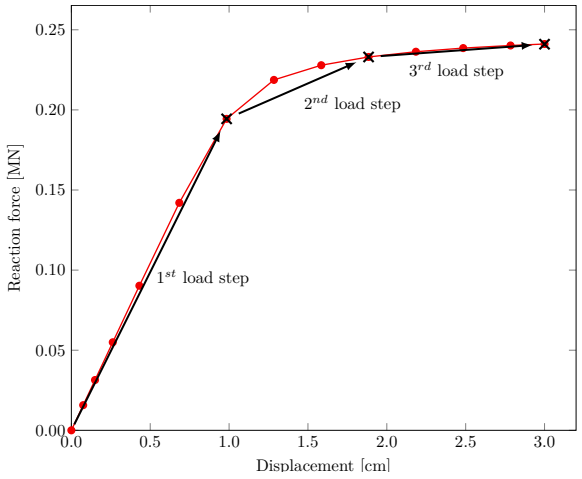
Figure 4.19 – Force–Displacement curves



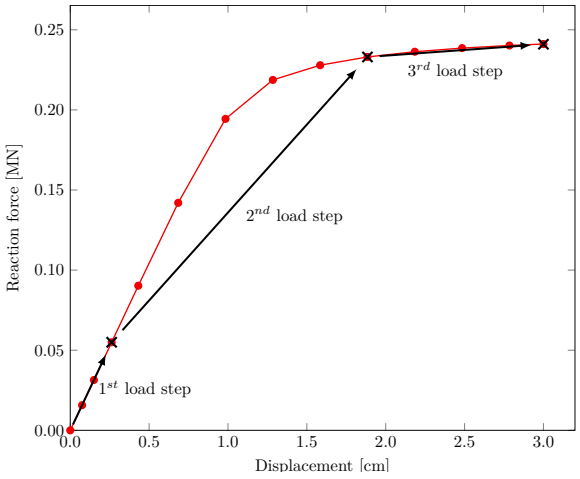
(a) Abaqus step-by-step strategy



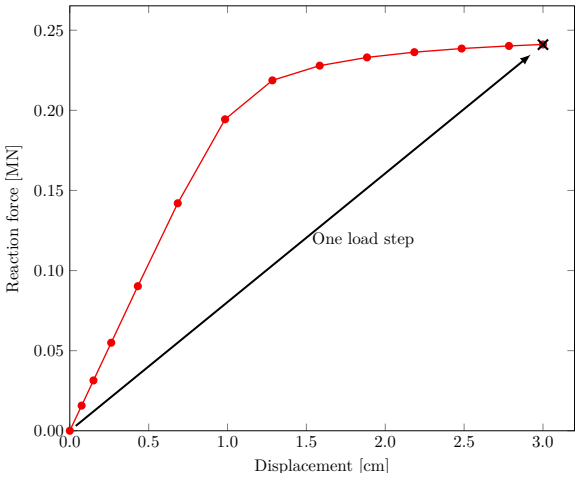
(b) IPM using 2 load steps



(c) IPM using 3 load steps



(d) IPM using 3 different load steps



(e) IPM using one load step

Figure 4.20 – Different load incrementation strategies

4.6.3 Basic steel checks according to the Eurocodes

A typical steel design check according to the Eurocodes [EN1993–1-1, 2005; EN1993–1-8, 2005] relies on a series of basic components checks which are then assembled to derive an estimate of the ultimate resistance of the structure or the joint. In order to show the validity of the yield design SOCP framework, elementary examples such as basic weld checks, bolt shear and plate bearing, biaxial bending interaction diagram and T-stub resistance are presented in the following paragraphs.

4.6.3.1 Some fillet weld checks

According to EN1993–1-8 [2005] paragraph 4.5.3 , the ultimate resistance of fillet welds can be determined using the *directional method* which takes into account a sort of dependence of the weld strength on the direction of loading by assuming that the normal stress parallel to the axis of the weld throat does not influence the design resistance. The ultimate resistance is given by Eq. (4.1) of the EN1993–1-8 [2005]:

$$\sqrt{\sigma_\perp^2 + 3(\tau_\perp^2 + \tau_\parallel^2)} \leq \frac{f_u}{\beta_w \gamma_{M2}} \quad \text{and} \quad \sigma_\perp \leq \frac{0.9 f_u}{\gamma_{M2}} \tag{4.56}$$

Using these two inequalities, the ultimate resistance of basic fillet welds can be determined. In the following examples, a fillet weld of throat thickness $a \geq 3\,\text{mm}$, of total length L and of effective length $L_{eff} = L - 2a \geq L_{min} = \max(30\,\text{mm}, 6a)$ is presented in 3 different configurations:

- a frontal configuration as seen in Fig. 4.21,
- a lateral configuration as seen in Fig. 4.22, and
- an inclined configuration parameterized with an angle α as seen in Fig. 4.23.

The assembled pieces are supposed to have the same steel grade where $f_u(\text{MPa})$ is used to denote the ultimate tensile strength. The assembled pieces are subjected to an axial force N_{Ed} as seen in Fig. 4.21 to 4.23. The ultimate resistance $N_{Rd} = \min(N_{Rd,1}; N_{Rd,2})$ can be calculated from the resistance $N_{Rd,i}$ obtained from the two conditions of (4.56). The analytical ultimate resistance for the three configurations are summarized in Tab. 4.3.

Table 4.3 – Ultimate resistance calculated using equation (4.1) of EN1993–1-8 [2005]

Configuration	$\sigma_\perp = \tau_\perp$	$\tau_\parallel$	1^{st} criterion $N_{Ed} \leq N_{Rd,1}$	2^{nd} criterion $N_{Ed} \leq N_{Rd,2}$
Frontal configuration	$\frac{N_{Ed}}{2aL_{eff}}\frac{\sqrt{2}}{2}$	0	$\frac{0.9f_u}{\gamma_{M2}} 2\sqrt{2} aL_{eff}$	$\frac{f_u}{\beta_w \gamma_{M2}} \sqrt{2} aL_{eff}$
Lateral configuration	0	$\frac{N_{Ed}}{2aL_{eff}}$	verified	$\frac{f_u}{\beta_w \gamma_{M2}} \frac{2aL_{eff}}{\sqrt{3}}$
Inclined configuration	$\frac{N_{Ed}}{aL_{eff}} \sin(\alpha) \frac{\sqrt{2}}{2}$	$\frac{N_{Ed}}{aL_{eff}} \cos(\alpha)$	$\frac{0.9f_u}{\gamma_{M2}} \frac{\sqrt{2} aL_{eff}}{\sin(\alpha)}$	$\frac{f_u}{\beta_w \gamma_{M2}} \frac{aL_{eff}}{\sqrt{3 - \sin^2 \alpha}}$

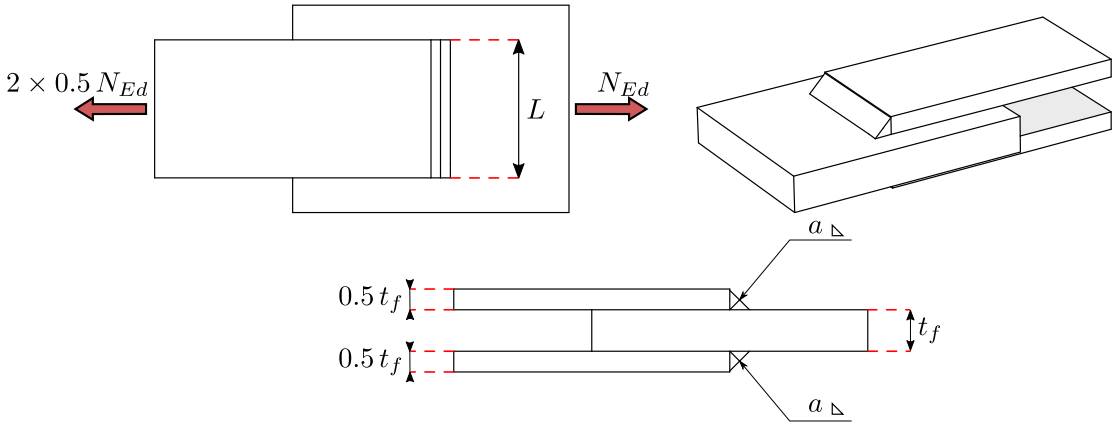


Figure 4.21 – Frontal fillet weld configuration

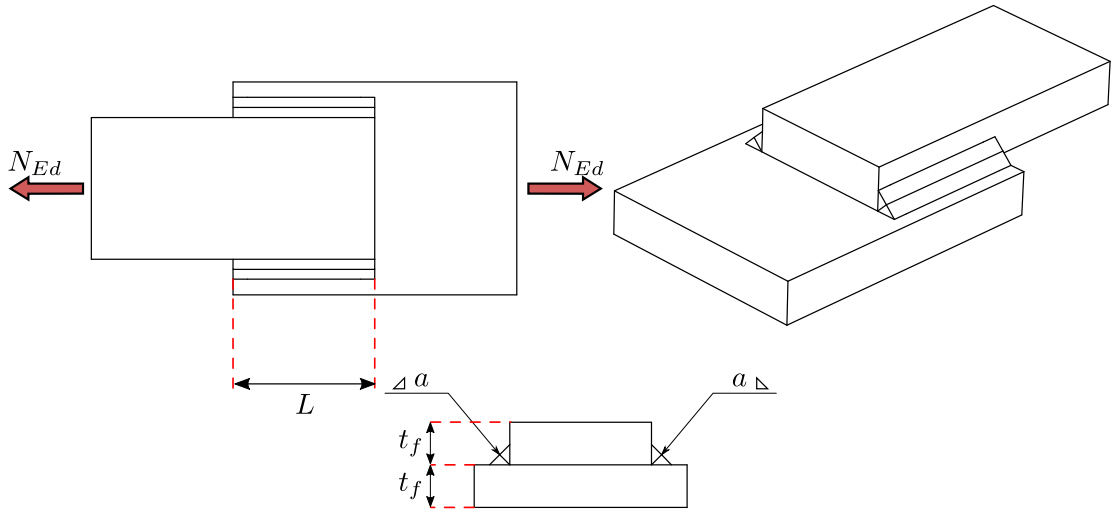


Figure 4.22 – Lateral fillet weld configuration

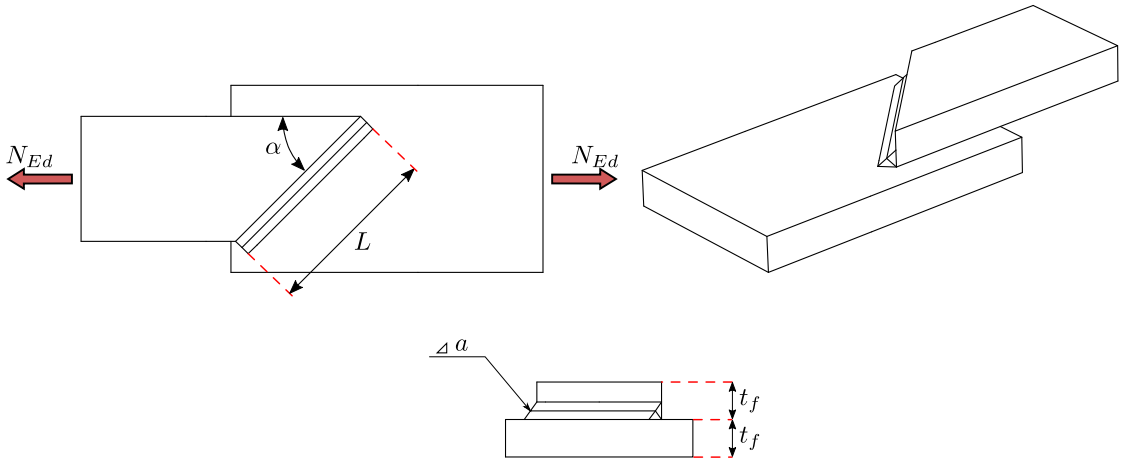


Figure 4.23 – Inclined fillet weld configuration

Table 4.4 – Ultimate resistance of the different weld configurations calculated using equation (4.1) of EN1993–1-8 [2005] and the dual SOCP method

Configuration	$N_{Rd,EC}$	$N_{Rd,upper}$	$N_{Rd,lower}$
Frontal configuration	492.7 kN	568.2 kN	531.3 kN
Lateral configuration	402.4 kN	401.3 kN	388.3 kN
Inclined configuration	334.5 kN	385.8 kN	358.1 kN

Considering $L = 100\,\text{mm}$, $a = 10\,\text{mm}$, $\alpha = 45°$ and a steel grade S355N/NL such as $f_u = 490\,\text{MPa}$, $\beta_w = 0.9$ and $\gamma_{M2} = 1.25$, the ultimate strengths of the different configurations are calculated using the EC equations and using the finite-element SOCP limit analysis method using both lower and upper bound approaches. Results are summarized in Tab. 4.4. The gap between the upper and lower bound can be reduced when using a finer mesh, although it is already less than 5%. The difference with the EN1993–1-8 [2005] value is mainly due to the hypothesis taken in equation (4.1) where the normal stress parallel to the axis of th weld does not influence the design resistance of the weld whereas in finite-element limit analysis computations, a full 3D yield criterion is considered. However, despite these intrinsic differences, the results differ by 7% at most for all cases when comparing the EC prediction with the safe lower bound approximation. We can observe that the Eurocode prediction tends to be conservative for the frontal and inclined configuration and slightly non-conservative for the lateral configuration.

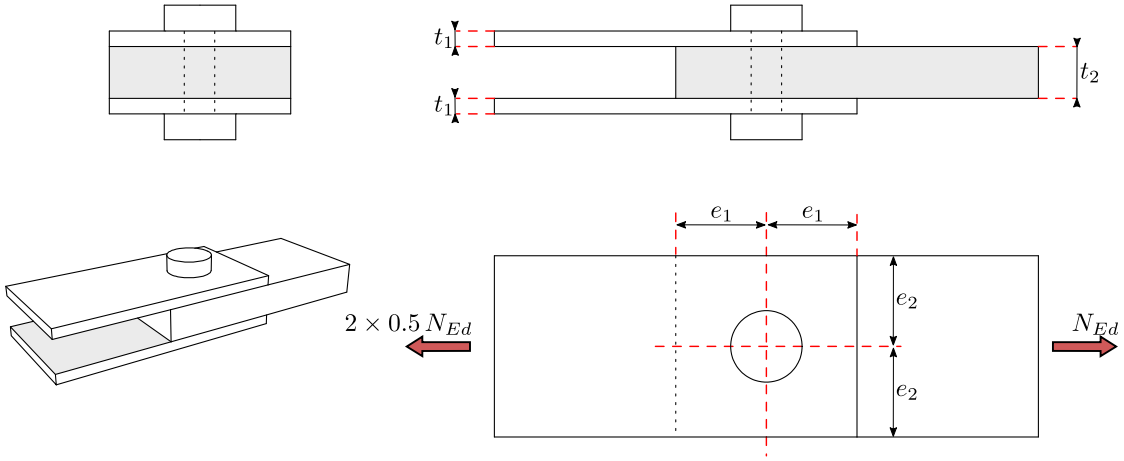


Figure 4.24 – Bearing pressure and bold shear model configuration

4.6.3.2 Bolt shear and plate bearing

According to EN1993–1-8 [2005] paragraph 3.6.1, the ultimate resistance of individual fasteners can be determined by considering 4 different failure modes: shear resistance of the bolt itself, bearing resistance of the plates surrounding the bolts, tension resistance of the bold and punching shear resistance. All the formulas and explanations can be found in Table 3.4 of the EN1993–1-8 [2005].

We consider the following model presented in Fig. 4.24. where traction forces acting perpendicularly to the bolt's axis are studied, thus modeling the shear failure of the bolts and the plate bearing resistance. The assembled pieces are supposed to have the same steel grade f_u(MPa). The ultimate tensile strength of the bolt is denoted f_{ub}(MPa).
In the case of regular not preloaded bolts, the resistance $F_{v,Rd}$ in shear depends on the shear strength of the bolt and the area A of the bolt in a particular shear plane and is given by:

$$F_{v,Rd} = \frac{\alpha_v f_{ub} A}{\gamma_{M2}} \tag{4.57}$$

where α_v takes into account the bolt's steel grade in order to evaluate its shear ultimate strength (in theory $\alpha_v = 1/\sqrt{3}$) and A is the tensile stress area of the bolt.
The bearing resistance of each individual plate can be calculated using:

$$F_{b,Rd} = \frac{k_1 \alpha_b f_u d t}{\gamma_{M2}} \tag{4.58}$$

where k_1 is a function of the edge distance e_1 which takes into account different shear or traction failure modes of the piece due to parallel forces, while α_b is a function of the edge distance e_2 which takes into account the perpendicular effects, d is the bolt's diameter and t the plate's thickness.

Table 4.5 – Ultimate resistance calculated the EN1993–1-8 [2005] and the dual SOCP method

Configuration	Failure Mode	F_{Rd}	$F_{Rd,upper}$	$F_{Rd,lower}$
$t_1 = 6\,\text{mm}$	Plate bearing	55.2 kN	62.3 kN	55.1 kN
$t_1 = 10\,\text{mm}$	Bolt shear	72.4 kN	83.9 kN	72.0 kN

For this example, the plates are of grade S355N/NL modelled as a perfectly plastic material with $f_y = f_u = 355\,\text{MPa}$, the bolt is of grade 8.8 also modelled as a perfectly plastic material with $f_y = f_u = 800\,\text{MPa}$. The edge distances are all equal to $e_1 = e_2 = 35\,\text{mm}$ and the center plate has a thickness of $t_2 = 20\,\text{mm}$. Two configurations for t_1 are studied:

1. $t_1 = 6\,\text{mm}$ where a failure mode with plate bearing is expected;
2. $t_1 = 10\,\text{mm}$ where a failure mode involving bolt shearing is expected.

The value for the safety coefficient γ_{M2} is taken equal to 1.0. For these values the EN1993–1-8 [2005] equations gives us the following resistances:

$$F_{v,Rd} = \frac{0.6 \times 800 \times 157 \times 10^{-6}}{1.0} \times 1000 = 75.4\,\text{kN} \tag{4.59}$$

$$F_{b,Rd,6\,\text{mm}} = \frac{2.5 \times 0.65 \times 355 \times 0.016 \times 0.006}{1.0} \times 1000 = 55.2\,\text{kN} \tag{4.60}$$

$$F_{b,Rd,10\,\text{mm}} = \frac{2.5 \times 0.65 \times 355 \times 0.016 \times 0.01}{1.0} \times 1000 = 92.0\,\text{kN} \tag{4.61}$$

The total resistance is obtained by taking the minimum value:

$$F_{Rd} = \min\{F_{v,Rd}, F_{b,Rd}\} \tag{4.62}$$

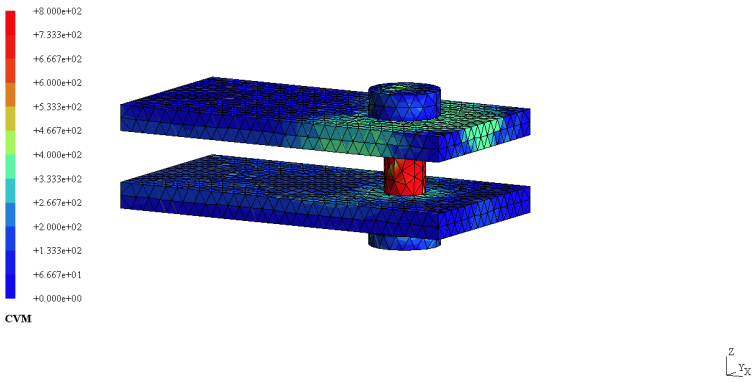
Results are summarized in Tab. 4.5. The gap between the upper and lower bound can be reduced when using a finer mesh. The differences with the results of the EN1993–1-8 [2005] are mainly due to the discretization of the bolt's section and hole using a polygon and to various 3D effects that aren't taken into account in the analytical formula. However, despite these intrinsic differences, the gap between the lower bound and the predicted value is negligible. We can observe in Fig. 4.25 and 4.26 that our approach correctly reproduces the expected failure mechanisms.

4.6.3.3 Biaxial bending of a beam

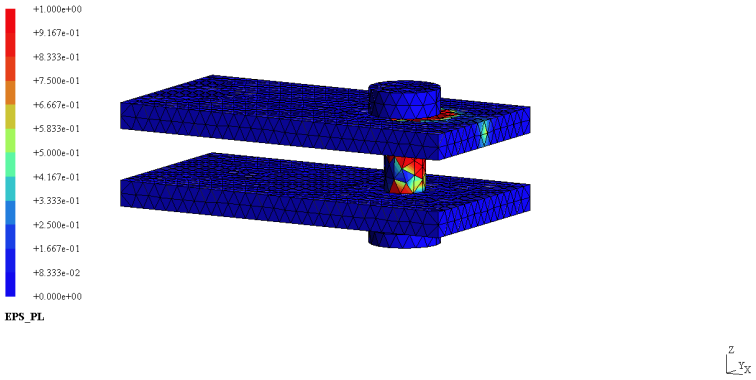
The ultimate state check of a steel class 1 or 2 section under biaxial bending action, excluding buckling instability, can be verified using two methods according to the EN1993–1-1 [2005]:

– A conservative plastic check using equation (6.2):

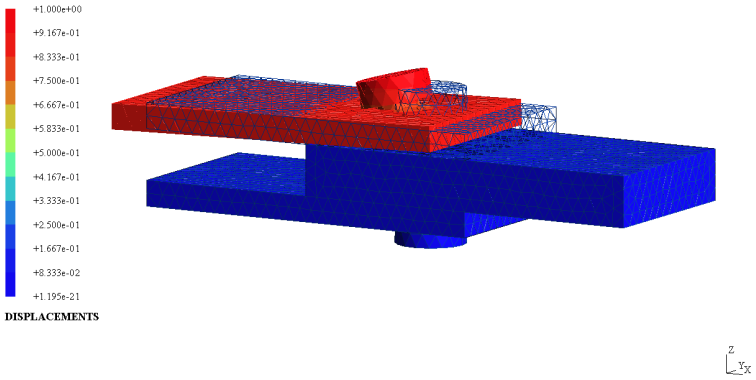
$$\frac{N_{Ed}}{N_{Rd}} + \frac{M_{y,Ed}}{M_{y,Rd}} + \frac{M_{z,Ed}}{M_{z,Rd}} \leq 1.0 \tag{4.63}$$



(a) Lower bound approach, von Mises stress isovalues

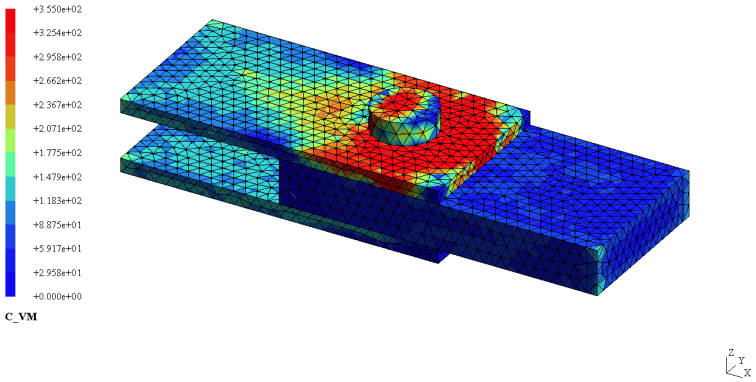


(b) Upper bound approach, equivalent plastic strain isovalues

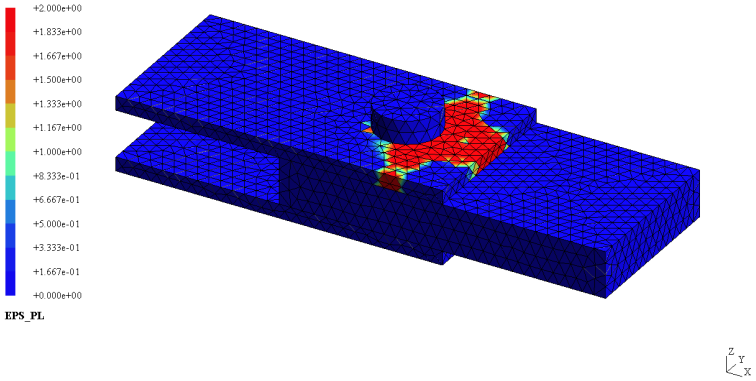


(c) Upper bound approach, obtained failure mechanism

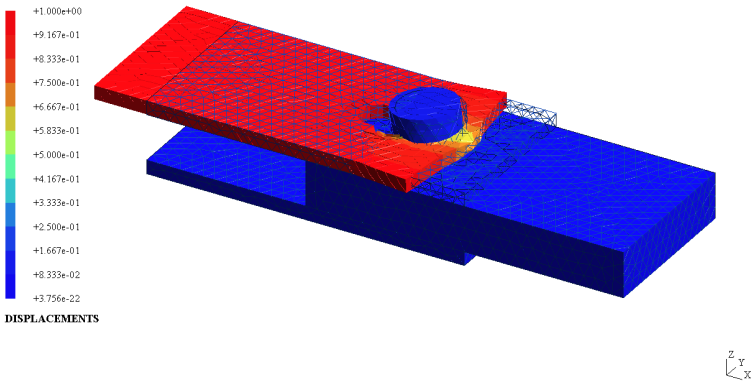
Figure 4.25 – Bolt shearing results using the dual approach



(a) Lower bound approach, von Mises stress isovalues



(b) Upper bound approach, equivalent plastic strain isovalues



(c) Upper bound approach, obtained failure mechanism

Figure 4.26 – Plate bearing results using the dual approach

– A more refined plastic check using equation (6.41):

$$\left[\frac{M_{y,Ed}}{M_{y,Rd}}\right]^{\alpha} + \left[\frac{M_{z,Ed}}{M_{z,Rd}}\right]^{\beta} \leq 1.0 \tag{4.64}$$

Eq. (4.63) corresponds to a lower-bound approximation of the real section interaction diagram which relies on a linear elastic normal stress distribution across the cross section. The second, Eq. (4.64), assuming a non-linear interaction cases, is more involved as it includes the coefficients α and β which depend on the section geometrical characteristics and on the normal force level. The different steps to calculate these coefficients can be found in detail in paragraph 6.2 of the EN1993–1-1 [2005].

This example is limited to the evaluation of the interaction diagram of HEB300 of steel grade 355N/NL under biaxial bending with no normal force. For this section, the mechanical characteristics are summarised in Tab. 4.6.

Table 4.6 – HEB300 section characteristics

Section	$A_s = 14\,900\,\text{mm}^2$
Plastic section modulus along the y axis	$W_{pl,y} = 1\,869\,000\,\text{mm}^3$
Plastic section modulus along the z axis	$W_{pl,z} = 870\,100\,\text{mm}^3$
Eq. (6.41) coefficient	$\alpha = 2$
Eq. (6.41) coefficient	$\beta = 1$

Using the different equations of the Eurocode, the interaction diagram for this section is determined and is shown in Fig. 4.27. The normalisation in equations (4.63) and (4.64) is done by calculating $M_{y,Rd}$ and $M_{z,Rd}$ using equation (6.13) of the EN1993–1-1 [2005]. Similarly, we compute the interaction diagram obtained from the numerical results by integration the stress diagram for the lower bound approach as the one seen Fig. 4.28, and by retrieving the correct Lagrange multipliers corresponding to the reaction moments in the case of the upper bound approach. Eq. (4.63) is clearly conservative as the full plastic capacity of the section is not mobilised where as Eq. (4.64) allows for a more economical design which coincides well enough with the upper and lower bound approach where a full 3D yield criterion is used. Fig. 4.28 shows the normal stress distributions obtained in our approach for various imposed bending moments. In particular, the pure bending strengths $M_{y,Rd}$ and $M_{z,Rd}$ are correctly retrieved.

4.6.3.4 T-stub resistance

A typical steel-assembly verification according to EN1993–1-8 [2005] relies on the components method where the complex assembly is divided into basic components. We limit ourselves for more simplicity to the tensile case where the equivalent T-stub method should be used to determine the ultimate resistance. The first step consists in determining the set of components subject to tensile loading under different configurations

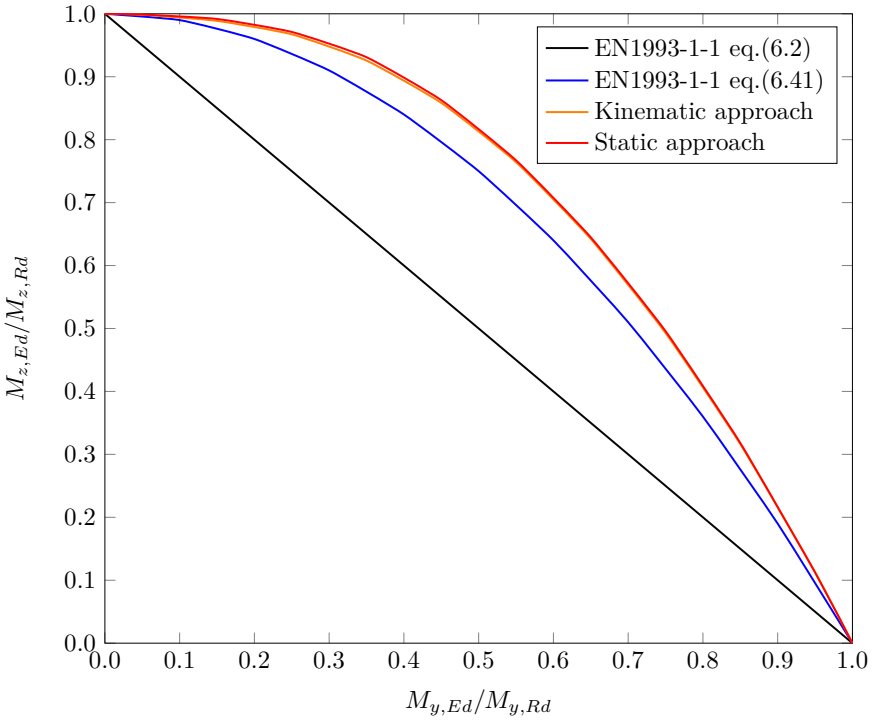


Figure 4.27 – Interaction diagram under biaxial bending action

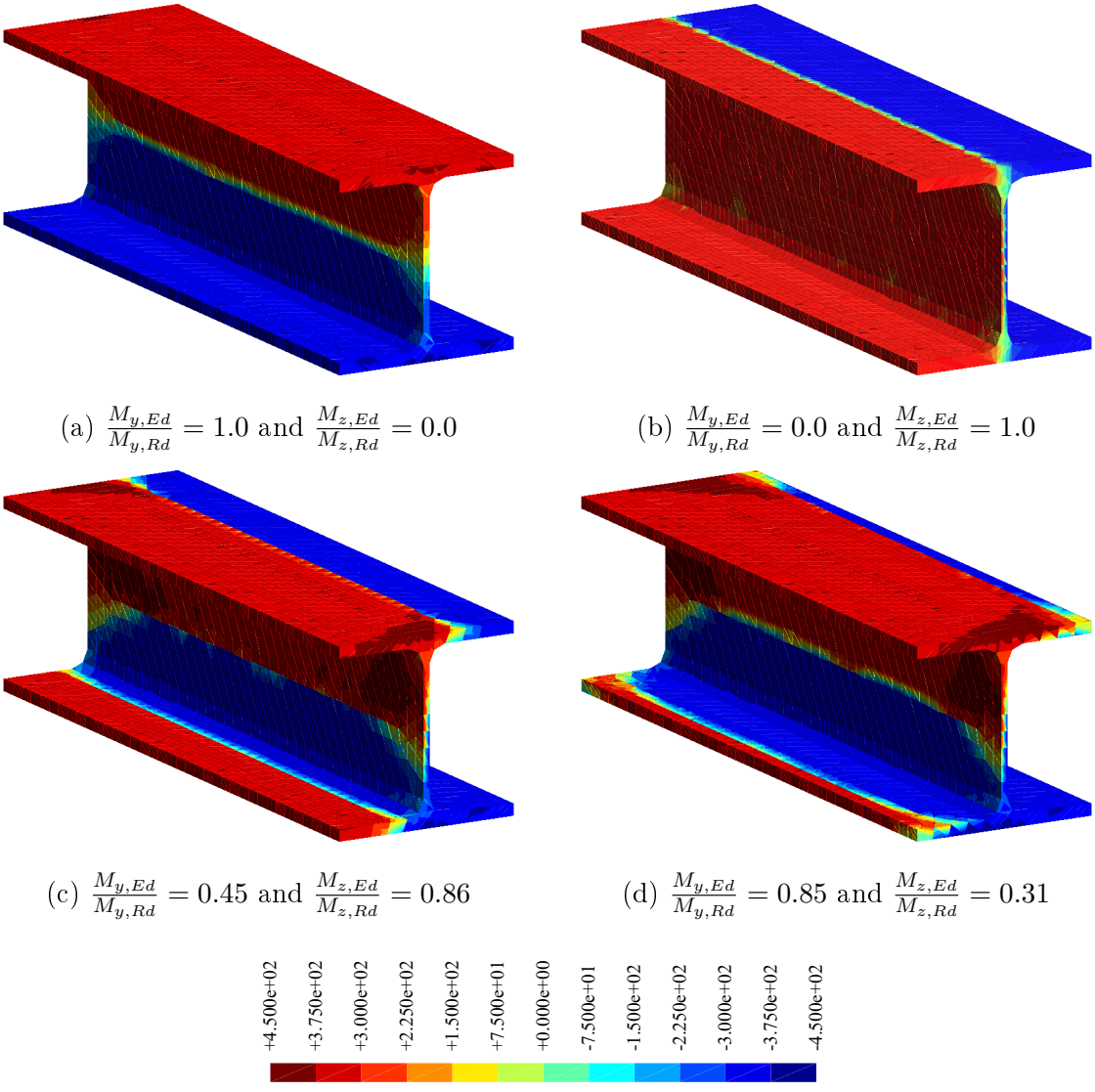
and subsequently calculating their equivalent T-stub lengths called L_{eff} (see Fig. 4.29). The assembly strength will be obtained as the lowest tensile strength of its different components. In order to validate our proposed yield analysis approach, we compute the strength of a single T-stub and compare it with the semi-analytical formulas proposed by the Eurocode. The description of the various variables in the considered model are represented in Fig. 4.30 which is adapted based on Fig. 6.2 of the EN1993–1-8 [2005]. More precisely, we consider three different variants described in Tab. 4.7 and, for each of them, the base flange thickness t_f will be varied. The thickness variation will allow us to test the different failure modes proposed by the Eurocode (cf. EN1993–1-8 [2005] Table 6.2) and described in Fig. 4.31:

– Mode 1: failure by complete flexural yielding of the flange characterized by the apparition of 4 plastic hinges

$$F_{T,1,Rd} = \frac{4.M_{pl,1,Rd}}{m} \tag{4.65}$$

– Mode 2: failure by partial flexural yielding of the flange characterized by the apparition of 2 plastic hinges, and partial tensile yielding of the bolts

$$F_{T,2,Rd} = \frac{2.M_{pl,2,Rd} + n\sum F_{t,Rd}}{n+m} \tag{4.66}$$



(a) $\frac{M_{y,Ed}}{M_{y,Rd}} = 1.0$ and $\frac{M_{z,Ed}}{M_{z,Rd}} = 0.0$

(b) $\frac{M_{y,Ed}}{M_{y,Rd}} = 0.0$ and $\frac{M_{z,Ed}}{M_{z,Rd}} = 1.0$

(c) $\frac{M_{y,Ed}}{M_{y,Rd}} = 0.45$ and $\frac{M_{z,Ed}}{M_{z,Rd}} = 0.86$

(d) $\frac{M_{y,Ed}}{M_{y,Rd}} = 0.85$ and $\frac{M_{z,Ed}}{M_{z,Rd}} = 0.31$

Figure 4.28 – Normal stresses isovalues for different loading states

– Mode 3: failure by complete tensile yielding of the bolts

$$F_{T,3,Rd} = \sum_{n_{bolts}} F_{t,Rd} \tag{4.67}$$

– Mode 4: failure by complete tensile yielding of the web

$$F_{T,4,Rd} = L_{eff}.t_w.f_{y,d} \tag{4.68}$$

– Mode 5: failure by complete shear yielding of the welds

$$F_{T,5,Rd} = L_{eff}.(2a_{w,T}).f_{vw,d} \tag{4.69}$$

The detailed description and formulas can be found in Table 6.2 of the EN1993-1-8 and specifically for:

– $M_{pl,1,Rd}$ which is the plastic resisting moment of the flange under failure mode 1;

– $M_{pl,2,Rd}$ which is the plastic resisting moment of the flange under failure mode 2;

– $n = \min(e, 1.25m)$ which is a geometrical parameter which depends on the position of the bolts.

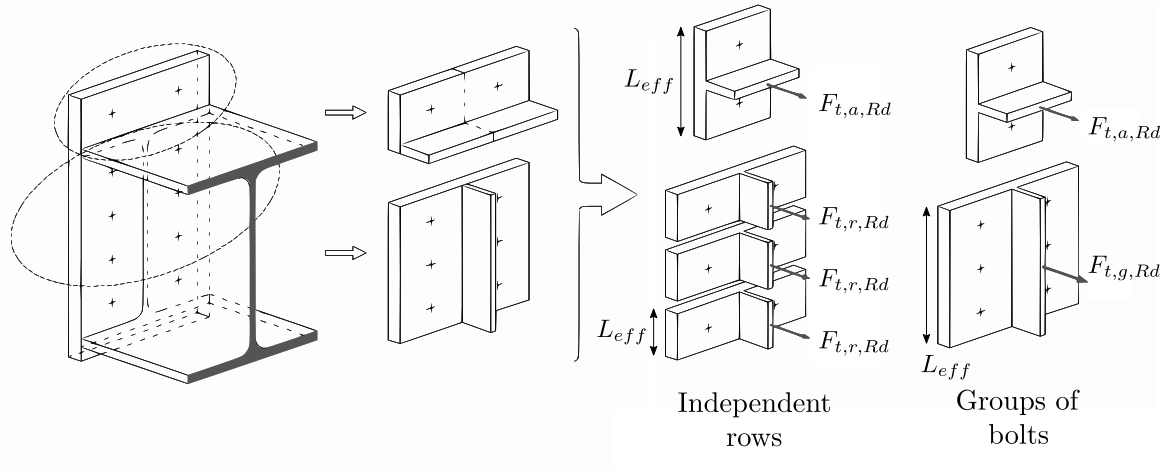


Figure 4.29 – A descriptive explanation of the components method

Stiff bolts are used in model 1, therefore no yielding is expected in these. Since $2a \leq t_w$, we expect that the welds will yield before the web since their yield strength under shear loading is less than the one for the web under tensile loading. We then expect a failure of type 1 for thin flanges and of type 5 for thick flanges. In models 2 and 3, smaller bolts are used, therefore modes 2 and 3 are expected to appear depending of the thickness of the flanges. In model 3, we chose $2a > t_w$ so that the yielding will occur in the web.

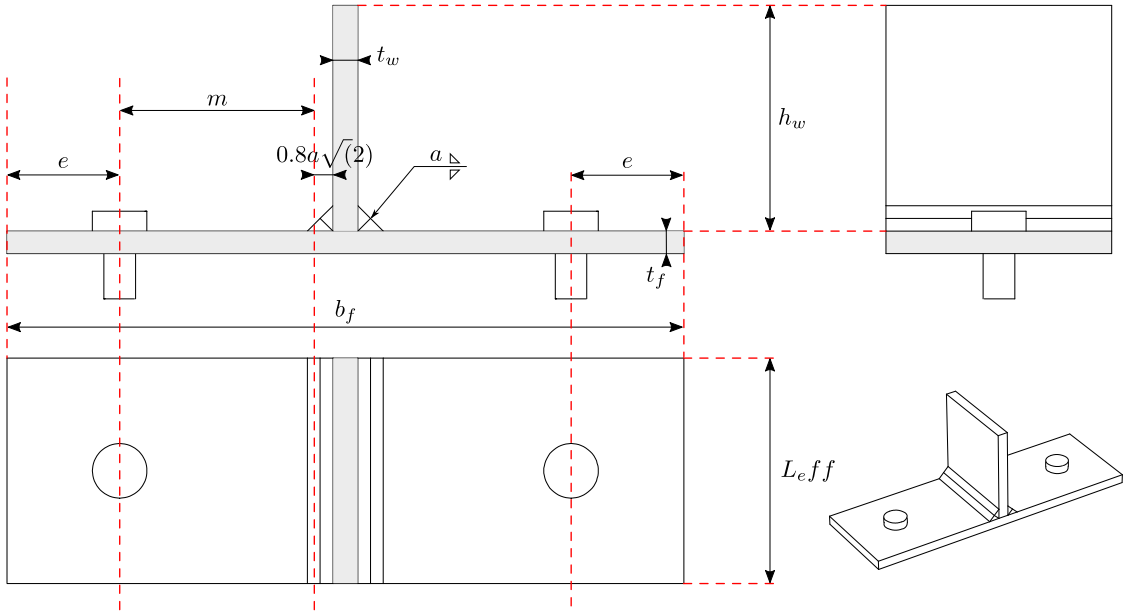
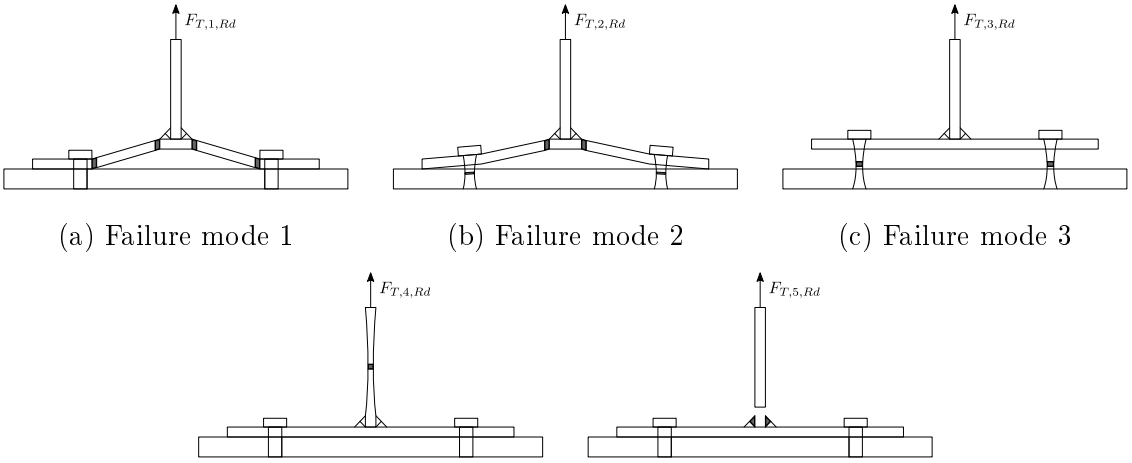


Figure 4.30 – Description of the used model

Table 4.7 – Description of the various models

Model name	Length L (mm)	Web			Flange			Welds		Bolts			Expected failure modes
		t_w (mm)	h_w (mm)	f_{uw} (MPa)	t_f (mm)	b_f (mm)	f_{uf} (MPa)	a (mm)	f_{vw} (MPa)	Type	f_{ub} (MPa)	e (mm)	
Model 1	100	10	100	355	Variable	300	355	5	205	M20	1000	50	1 – 5
Model 2	100	10	100	355	Variable	300	355	5	205	M14	1000	50	1 – 2 – 5
Model 3	100	11	100	355	Variable	300	355	8	205	M14	1000	50	1 – 2 – 3

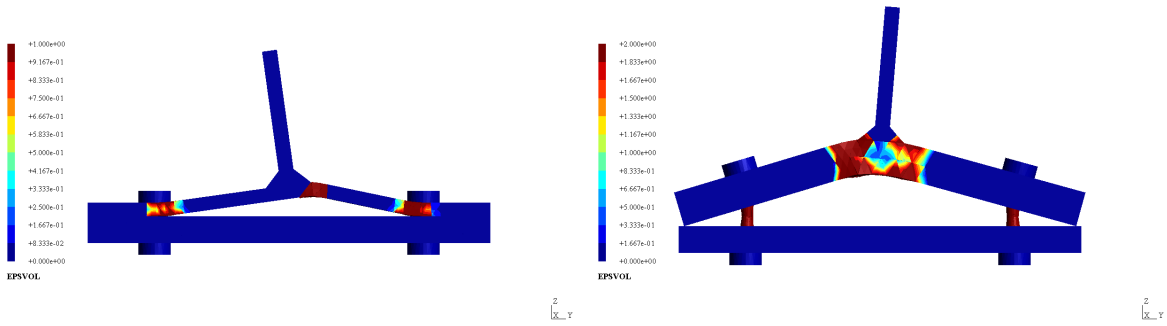
The obtained failure mechanisms are in accordance with the expected modes. For example, Fig. 4.32 illustrates the three different failure modes obtained with the third model when varying the flange thickness: complete flexural yielding of the flange (Mode 1) for a thin flange ($t_f = 10\,\text{mm}$), partial flexural yielding of the flange and tensile yielding of the bolts (Mode 2) for a moderately thick flange ($t_f = 25\,\text{mm}$) and a complete yielding of the bolts (Mode 3) with a thick flange ($t_f = 40\,\text{mm}$). The results obtained for all the considered configurations are presented in Fig. 4.33. Using the upper and lower bound yield analysis, we obtain a very satisfying estimate of the T-stub strength complying with that proposed by the Eurocode. The small differences observed in Fig. 4.33 with respect to the Eurocode semi-analytical formulae are due to the fact that the latter are obtained from 2D failure modes while our model is fully 3D. For instance, yield lines of modes 1 and 2 are considered straight of length L_{eff}, while our computations produce curved yield lines depending on the depth of the analysed model.



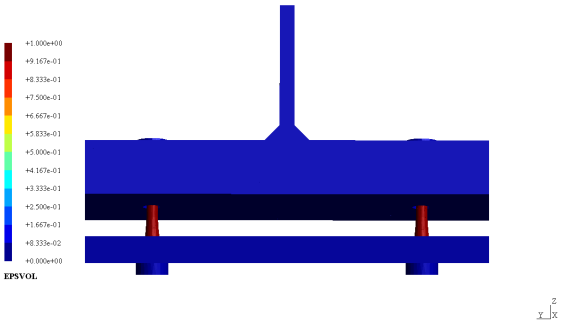
(a) Failure mode 1 (b) Failure mode 2 (c) Failure mode 3

(d) Failure mode 4 (e) Failure mode 5

Figure 4.31 – Different failure modes proposed by the Eurocode



(a) Model 3 – $t_f = 10\,\text{mm}$ – Failure mode 1 (b) Model 3 – $t_f = 25\,\text{mm}$ – Failure mode 2



(c) Model 3 – $t_f = 40\,\text{mm}$ – Failure mode 3

Figure 4.32 – Different failure modes obtained for 3 configurations of the third model (isocolours correspond to the equivalent von Mises plastic strain)

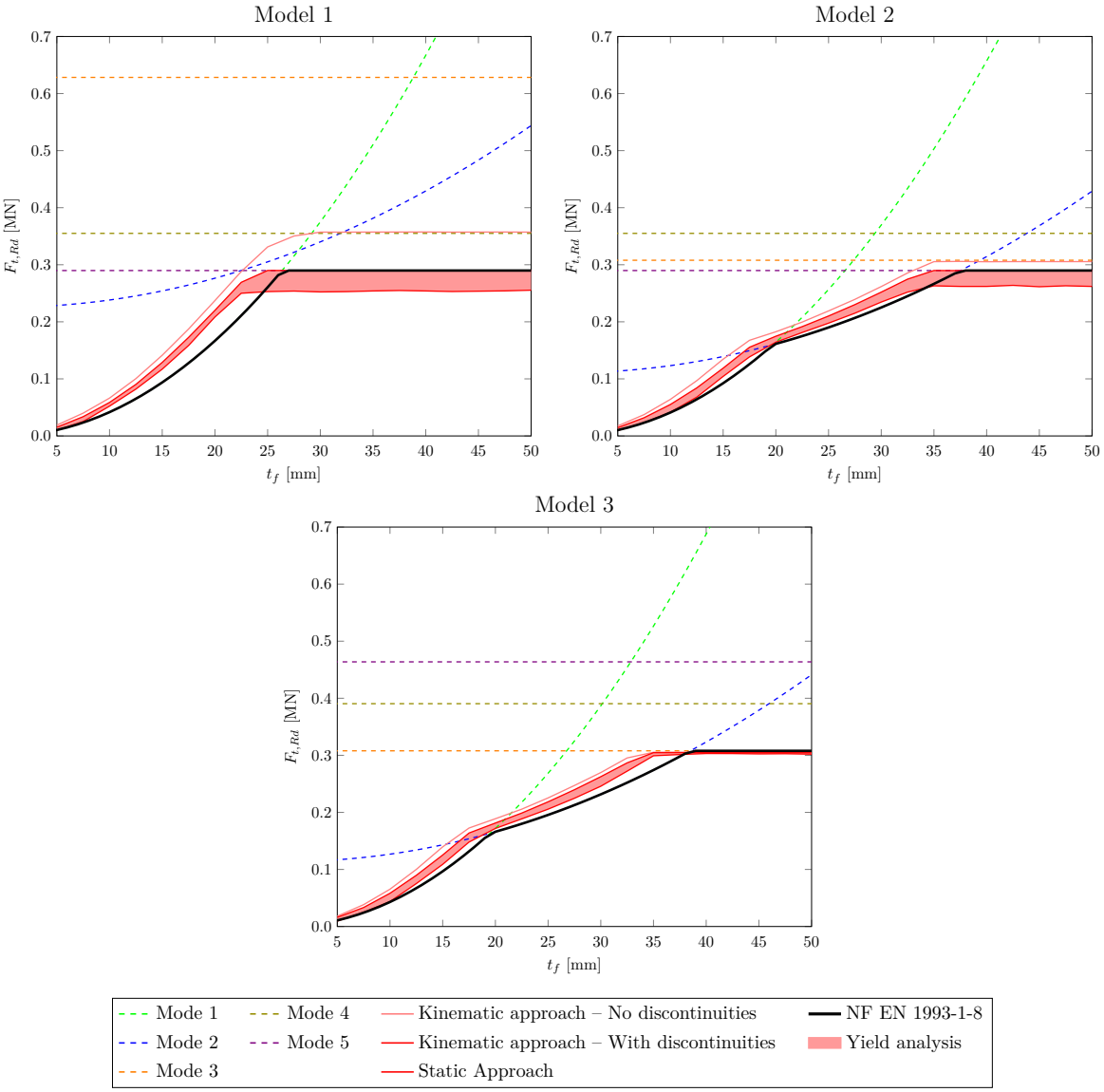


Figure 4.33 – Results comparisons with the EN1993–1-8 [2005]

4.7 Conclusions and orientation

Modeling three-dimensional elastoplastic steel structures (including also contact) using SOCP and IPM has been investigated in this chapter. Two kinematic and static variational principles are considered, leading to the formulation of convex minimization problems. Computation of the structure ultimate load using limit analysis concepts can be obtained as a particular case of the elastoplastic variational problems with only minor modifications. Besides, assuming a radial loading path, limit loads can also be computed from the elastoplastic problems using a single large load step, without the need to perform a step-by-step load subdivision. Robustness of the solution procedure to large load steps is ensured by resorting to a primal-dual interior point method which is really efficient for SOCP problems.

The use of a dual analysis based on both a kinematic and static approach enables to assess the solution convergence by comparing their associated energies. The comparison of local fields also enables to compute an error indicator used for adaptive remeshing. The static-based solution has also the interesting property of being a lower-bound to the true solution. It is therefore, a safer solution in terms of stress quantities than a displacement-based solution, which is appealing in the context of safety verification including elastoplastic behavior and limit load computations.

The robustness and efficiency of the solution procedure has been compared against Abaqus elastoplastic computations. When aiming at computing a fully yielded solution, the computational benefit of performing only one large load step compared to a standard step-by-step Newton procedure is quite important. Limit load computations can therefore be obtained at reasonable computational cost and offer a very useful design aid for the engineer as shown by the considered example. In particular, we have demonstrated that our numerical tool produces excellent estimates of the ultimate capacity of welds, bolts, plate bearings and cross-section flexural capacities We also showed that it enables to retrieve the different failure modes of a T-stub considered by the Eurocodes, paving the way to using this tool for the ultimate state design of more complex assemblies as it will be shown later in chapter 6.

Before tackling such complex engineering problems, we investigate in chapter 5 the possibility of taking into account geometrical non-linearities in a 3D elastoplastic setting using our optimization tools.

* *
*

Chapter 5

Finite-strain elastoplastic analysis using non-linear second-order cone programming

Abstract: *In this chapter, the possibility of extending the SOCP framework to include geometrical non-linearities is explored. The IPM dedicated to optimization problems involving non-smooth but convex objectives or constraints is adapted to the case of a non-convex objective. A typical application for such kind of problems is finite-strain elastoplasticity which we address using a total Lagrangian formulation based on a logarithmic strain measure. The proposed interior-point algorithm is implemented and tested on 3D benchmark examples involving plastic collapse and geometrical changes.*

Contents

5.1 Motivation

Motivated by the promising results obtained in chapters 3 and 4, and various engineering projects requiring further developments at *Strains*, we aim at extending the presented framework to include yet another type of non-linearities. Often in complex projects, the engineer is obliged to consider geometrical non-linearities and large-strain effects in addition to plasticity. As seen in chapter 1, such analysis called *second-order non-linear analysis*, is required in building codes whenever the small strain hypothesis is not sufficient to model the real behavior of the mechanical system. A typical example is the calculation of a cable-stayed bridge where the cable system does not hold without the geometrical effects, thus the engineer cannot correctly asses the structural integrity. Another common problem is plate buckling often found in thin-walled steel structures, and specifically the local buckling of web plates in complex 3D assemblies. Modeling these phenomena is not an easy task, and providing an efficient tool similar to the one presented in the precedent chapters will greatly improve the efficiency and the security of designs. With that in mind, a first but major step is to investigate the possibility to apply the IPM framework to this type of problem.

The present chapter therefore aims at exploring one step further in the direction of extending IPM to the case of finite-strain plasticity. Building upon the previous chapters and works on a custom IPM solver including smooth convex terms [Bleyer, 2017], we investigate the case of problems containing smooth but non-convex terms. Obviously, proofs of convergence of the IPM algorithm will necessarily be lost in the non-convex case. However, our heuristic reasoning is that we will restrict to a case in which difficulties will be decoupled. On the one hand, non-smoothness is present only in conic constraints which we still consider to be convex, while, on the other hand, non-convexity of some objective terms or constraints will concern only terms which we assume to be smooth. Since the IPM can be seen as a Newton method with continuation along the so-called *central path*, we hope that smooth terms will be properly handled by the Newton method and that continuation along the central path will sill handle properly the non-smooth but convex conic constraints. Our proposed algorithm is therefore a simple extension of a classical IPM to the previously mentioned non-convex case.

We apply the proposed framework to the specific case of logarithmic strain elastoplasticity [Miehe et al., 2002]. As it will be discussed later, the use of the logarithmic strain setting enables a simple extension of classical small-strain elastoplastic constitutive laws to the finite-strain setting. In particular, the additive decomposition between elastic and plastic strain is preserved and elastic energy densities and plastic dissipation potentials are still convex with respect to the corresponding strains. Non-convexity only arises due to the non-linear relation between displacement and total strain. Using such a framework, we indeed obtain a problem in which non-smoothness (due to the plastic dissipation) can still be expressed using convex constraints whereas non-convexity involves smooth terms (strain/displacement relation). Benchmark 3D examples will validate our implementation. Comparison is made against standard Newton-Raphson methods based on a return-mapping inner procedure enabling to assess the computational cost

and convergence robustness of the IPM solver.

5.2 Large strain elastoplastic problems

5.2.1 Kinematics

Let $\mathcal{B}^0$ be an elastoplastic body that occupies the space Ω^0 in the reference configuration. Its external boundary $\Gamma^0 := \partial\mathcal{B}^0$ is split into two distinct parts $\Gamma^0 = \Gamma^0_{\boldsymbol{u}} \cup \Gamma^0_{\boldsymbol{t}}$ and $\Gamma^0_{\boldsymbol{u}} \cap \Gamma^0_{\boldsymbol{t}} = \emptyset$ where:

- displacements $\boldsymbol{u}$ are imposed to $\boldsymbol{u}_d$ on $\Gamma^0_{\boldsymbol{u}}$;
- surface tractions are imposed to $\boldsymbol{t}_d$ on $\Gamma^0_{\boldsymbol{t}}$.

The deformation process maps every point $\boldsymbol{X} \in \Omega^0$ in the reference configuration to a point $\boldsymbol{x}_{(t)} = \boldsymbol{X} + \boldsymbol{u}_{(\boldsymbol{X},t)}$ in the equilibrium configuration. The deformation gradient is then given by $\boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{\nabla}_{\mathrm{X}}\,\boldsymbol{x} = \boldsymbol{I} + \boldsymbol{\nabla}_{\mathrm{X}}\,\boldsymbol{u}$ with $J = \det(\boldsymbol{F}) > 0$. The right polar decomposition of the deformation gradient $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$ allows the definition of two fundamental tensors: the material stretch tensor $\boldsymbol{U}$ and the rotation tensor $\boldsymbol{R}$ with $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^T = \boldsymbol{I}$.

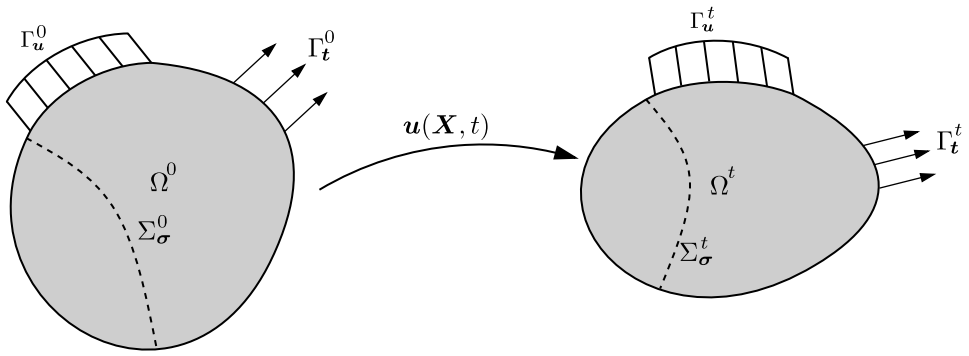


Figure 5.1 – Reference model

5.2.2 The logarithmic strain measure and its work-conjugate stress

In this work, we will adopt the logarithmic strain framework proposed in Miehe et al. [2002] which has been shown to be well suited for describing finite-strain metal plasticity. In this framework, the chosen total strain measure is the material Hencky logarithmic strain measure:

$$\boldsymbol{E} = \frac{1}{2}\ln(\boldsymbol{F}^T\boldsymbol{F}) = \ln(\boldsymbol{U}) \tag{5.1}$$

An attractive feature of using logarithmic strain measures is that classical small strain constitutive relations can be naturally extended to a finite-strain setting. In

particular, the total (Hencky) strain can be split additively into many contributions (elastic, plastic, thermal, swelling, etc.). Its trace is also linked with the volume change $J = \exp(\mathrm{tr}(\boldsymbol{E}))$. Classical von Mises plasticity can therefore be used in the space of logarithmic strains, in particular the total strain will still be assumed to consist of the sum of an elastic and a plastic part:

$$\boldsymbol{E} = \boldsymbol{E}^e + \boldsymbol{E}^p \tag{5.2}$$

We will denote by $\boldsymbol{T}$ the stress measure associated with $\boldsymbol{E}$ with respect to the power density of internal forces i.e. $p_{\text{int}} = \boldsymbol{T} : \dot{\boldsymbol{E}}$. This pair of variables can be related to other classical stress/strain measures such as the Green-Lagrange strain measure $\boldsymbol{E}_{GL} = \frac{1}{2}(\boldsymbol{U}^2 - \boldsymbol{I})$ and its work-conjugate stress, the second Piola-Kirchhoff stress tensor $\boldsymbol{S}$. One has for instance:

$$p_{\text{int}} = \boldsymbol{S} : \dot{\boldsymbol{E}}_{GL} = \boldsymbol{T} : \dot{\boldsymbol{E}} = \boldsymbol{T} : \left(\mathbb{M} : \dot{\boldsymbol{E}}_{GL}\right) \tag{5.3}$$

where

$$\mathbb{M} = \frac{\partial \boldsymbol{E}}{\partial \boldsymbol{E}_{GL}} \tag{5.4}$$

is a fourth-order geometrical tensor mapping both strain rate measures. We, therefore, also have:

$$\boldsymbol{S} = \boldsymbol{T} : \mathbb{M} \tag{5.5}$$

or equivalently:

$$\boldsymbol{T} = \boldsymbol{S} : \mathbb{M}^{-1} \tag{5.6}$$

Constitutive tangent stiffness operators can also be expressed using the two different stress/strain pairs and are related as follows:

$$\mathbb{D}_{GL} = \mathbb{M} : \mathbb{D} : \mathbb{M} + \boldsymbol{T} : \mathfrak{L} \tag{5.7}$$

where $\mathbb{D}_{GL}$ is the tangent operator in the Green-Lagrange setting i.e. such that $\dot{\boldsymbol{S}} = \mathbb{D}_{GL} : \dot{\boldsymbol{E}}_{GL}$, $\mathbb{D}$ is the tangent operator in the logarithmic setting i.e. such that $\dot{\boldsymbol{T}} = \mathbb{D} : \dot{\boldsymbol{E}}$ and finally $\mathfrak{L}$ is a sixth-order mapping tensor defined as:

$$\mathfrak{L} = \frac{\partial^2 \boldsymbol{E}}{\partial \boldsymbol{E}_{GL} \partial \boldsymbol{E}_{GL}} \tag{5.8}$$

In fact both Green-Lagrange and Hencky strain measures belong to the more general family of Seth-Hill strain measures. General details concerning these strain measures and the geometric mappings existing between them, especially the general expression for $\mathbb{M}$ and $\mathfrak{L}$, can be found in appendix E and in Miehe et al. [2002], Caminero et al. [2011] and Latorre and Montáns [2016]. Although our focus is mostly on the use of the Hencky strain measure, our implementation has been based on the Seth-Hill framework, allowing to change easily the chosen strain measure (keeping in mind that choosing another strain measure than the Hencky strain for describing finite-strain plasticity may not be appropriate).

5.2.3 The global incremental variational problem for elastoplastic media

We place ourselves in the framework of standard generalized materials [Halphen and Son Nguyen, 1975; Lemaitre and Chaboche, 1994], namely we postulate the existence of a Helmholtz free energy function of the following form:

$$\psi(\boldsymbol{E}, \boldsymbol{E}^p, p) = \psi^e(\boldsymbol{E}^e) + \psi^p(p) \tag{5.9}$$

where $\psi^e(\boldsymbol{E}^e)$ is the stored elastic energy density and $\psi^p(p)$ the hardening energy density with p being an internal state variable. In this decomposition, ψ^p is assumed to be convex and ψ^e polyconvex [Abbas et al., 2018]. Besides, we postulate the existence of a convex positively-homogeneous plastic dissipation potential $\phi(\dot{\boldsymbol{E}}^p)$.

We now focus on an incremental formulation over $[t_n, t_{n+1}]$ among the total time interval $[0, T]$, assuming that all mechanical fields are known at time t_n. Following Miehe et al. [2002], and in the same spirit as in chapter 4, the unknown fields at time t_{n+1} can be obtained from the solution to the following incremental variational formulation:

$$\begin{aligned}(\boldsymbol{u}_{n+1}, \boldsymbol{E}^p_{n+1}, p_{n+1}) = \underset{\boldsymbol{u}, \boldsymbol{E}^p, p}{\arg\min} \int_{\Omega^0} \int_{t_n}^{t_{n+1}} (\dot{\psi}(\boldsymbol{E}, \boldsymbol{E}^p, p) + \phi(\dot{\boldsymbol{E}}^p)) dt d\Omega \\ - \int_{t_n}^{t_{n+1}} P_{ext}(\dot{\boldsymbol{u}}) dt\end{aligned} \tag{5.10}$$

where P_{ext} is the power of external loads.

Restricting the above minimization to radial evolutions (see chapter 4) of $\dot{\boldsymbol{E}}^p(t)$ over the time interval, we obtain the following incremental minimization principle:

$$(\boldsymbol{u}_{n+1}, \boldsymbol{E}^p_{n+1}, p_{n+1}) = \underset{\boldsymbol{u}, \boldsymbol{E}^p, p}{\arg\min} \int_{\Omega^0} \Psi_n^{n+1}(\boldsymbol{E}, \boldsymbol{E}^p, p) d\Omega - P_{ext}(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) \tag{5.11}$$

in which we assumed that external forces are constant and where the incremental pseudo-energy density is given by:

$$\Psi_n^{n+1}(\boldsymbol{E}, \boldsymbol{E}^p, p) = \psi^e(\boldsymbol{E}, \boldsymbol{E}^p) + \psi^p(p) - \psi^e(\boldsymbol{E}_n, \boldsymbol{E}^p_n) - \psi^p(p_n) + \phi(\boldsymbol{E}^p - \boldsymbol{E}^p_n) \tag{5.12}$$

5.2.4 The case of von Mises plastic with linear isotropic hardening

We now particularize the problem to von Mises plasticity with linear isotropic hardening. In this case:

$$\psi(\boldsymbol{E}, \boldsymbol{E}^p, p) = \psi^e(\boldsymbol{E}, \boldsymbol{E}^p) + \psi^p(p) = \frac{1}{2}(\boldsymbol{E} - \boldsymbol{E}^p) : \mathbb{D} : (\boldsymbol{E} - \boldsymbol{E}^p) + \frac{1}{2} E_h p^2 \tag{5.13}$$

where $\mathbb{D}$ is the elastic modulus tensor of small-strain isotropic linear elasticity:

$$\mathbb{D} = \lambda \boldsymbol{I} \otimes \boldsymbol{I} + 2\mu \mathbb{I} \tag{5.14}$$

E_h is the hardening modulus and $p(t) = \int_0^t \sqrt{\frac{2}{3}}\|\dot{\boldsymbol{E}}^p\|dt$ is the accumulated plastic strain where $\|\boldsymbol{a}\| = \sqrt{a_{ij}a_{ij}}$.
The plastic dissipation potential is here:

$$\phi(\dot{\boldsymbol{E}}^p) = \begin{cases} \sqrt{\frac{2}{3}}\sigma_0\|\dot{\boldsymbol{E}}^p\| & \text{if } \operatorname{tr}(\dot{\boldsymbol{E}}^p) = 0 \\ +\infty & \text{otherwise} \end{cases} \tag{5.15}$$

Introducing the notation $\bar{\star} = \star_{n+1} - \star_n$ to denote the variable increment over the current time step, (5.12) becomes:

$$\begin{aligned} \Psi_n^{n+1}(\boldsymbol{E}, \boldsymbol{E}^p, p) = \frac{1}{2}(\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) : \mathbb{D} : (\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) + \boldsymbol{T}_n : (\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) \\ + \frac{1}{2}E_h(\bar{p})^2 + E_h p_n \bar{p} + \sqrt{\frac{2}{3}}\sigma_0\|\bar{\boldsymbol{E}}^p\| \end{aligned} \tag{5.16}$$

with $\operatorname{tr}(\bar{\boldsymbol{E}}^p) = \operatorname{tr}(\boldsymbol{E}^p) = 0$.
Since $\dot{p} = \sqrt{\frac{2}{3}}\|\dot{\boldsymbol{E}}^p\|$, we also have $\bar{p} = \sqrt{\frac{2}{3}}\|\bar{\boldsymbol{E}}^p\|$ therefore

$$E_h p_n \bar{p} + \sqrt{\frac{2}{3}}\sigma_0\|\bar{\boldsymbol{E}}^p\| = \sqrt{\frac{2}{3}}\sigma_{Y,n}\|\bar{\boldsymbol{E}}^p\|$$

with $\sigma_{Y,n} = \sigma_0 + E_h.p_n$ being the yield stress at the previous time step. The minimization problem (5.11) therefore reads as:

$$\begin{aligned} \min_{\boldsymbol{u}, \bar{\boldsymbol{E}}^p, p} \int_{\Omega^0} \Bigg(\frac{1}{2}(\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) : \mathbb{D} : (\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) + \boldsymbol{T}_n : (\bar{\boldsymbol{E}} - \bar{\boldsymbol{E}}^p) \\ + \frac{1}{2}E_h(\bar{p})^2 + \sqrt{\frac{2}{3}}\sigma_{Y,n}\|\bar{\boldsymbol{E}}^p\| \Bigg) d\Omega - P_{ext}(\bar{\boldsymbol{u}}) \end{aligned} \tag{5.17}$$

under the following constraints:

$$\bar{\boldsymbol{E}} + \boldsymbol{E}_n = \frac{1}{2}\ln\left(\boldsymbol{F}^T \cdot \boldsymbol{F}\right) \tag{5.18a}$$

$$\boldsymbol{F} = \boldsymbol{F}_n + \nabla_X \bar{\boldsymbol{u}} \tag{5.18b}$$

$$\bar{\boldsymbol{u}} + \boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} \quad \text{on } \Gamma_{\boldsymbol{u}}^0 \tag{5.18c}$$

$$\operatorname{tr}(\bar{\boldsymbol{E}}^p) = 0 \tag{5.18d}$$

Following Yonekura and Kanno [2012], the traceless constraint $\operatorname{tr}(\bar{\boldsymbol{E}}^p) = 0$ can be removed by introducing directly in the elastic energy every occurrences of $\bar{\boldsymbol{E}}^p$ by $\operatorname{dev}(\bar{\boldsymbol{E}}^p) = \mathbb{K} : \bar{\boldsymbol{E}}^p$ where $\mathbb{K}$ is the projector over deviatoric tensors.

Moreover, in the objective function of (5.17), the only non-smooth term is $\sqrt{\frac{2}{3}}\sigma_{Y,n}\|\bar{\boldsymbol{E}}^p\|$ which can be transformed into a second-order cone constraint by introducing an additional variable $\bar{\gamma}$ such that $\|\bar{\boldsymbol{E}}^p\| \leq \bar{\gamma}$. Replacing also the quadratic term $\frac{1}{2}E_h(\bar{p})^2$ with this new variable, one finally obtains:

$$\underset{\bar{\boldsymbol{u}},\bar{\boldsymbol{E}}^p,\bar{\gamma}}{\text{minimize}} \quad \int_{\Omega^0} \left[\frac{1}{2}(\bar{\boldsymbol{E}} - \mathbb{K} : \bar{\boldsymbol{E}}^p) : \mathbb{D} : (\bar{\boldsymbol{E}} - \mathbb{K} : \bar{\boldsymbol{E}}^p) + \boldsymbol{T}_n : (\bar{\boldsymbol{E}} - \mathbb{K} : \bar{\boldsymbol{E}}^p) \right. \tag{5.19a}$$
$$\left. + \frac{1}{3}E_h(\bar{\gamma})^2 + \sqrt{\frac{2}{3}}\sigma_{Y,n}\bar{\gamma} \right] d\Omega - P_{ext}(\bar{\boldsymbol{u}})$$

$$\text{subject to} \quad \bar{\boldsymbol{E}} + \boldsymbol{E}_n = \frac{1}{2}\ln\left(\boldsymbol{F}^T \cdot \boldsymbol{F}\right) \quad \text{in } \Omega^0 \tag{5.19b}$$
$$\boldsymbol{F} = \boldsymbol{F}_n + \nabla_X \bar{\boldsymbol{u}} \quad \text{in } \Omega^0 \tag{5.19c}$$
$$\bar{\boldsymbol{u}} + \boldsymbol{u}_n = \boldsymbol{u}_{d,n+1} \quad \text{on } \Gamma^0_{\boldsymbol{u}} \tag{5.19d}$$
$$\|\bar{\boldsymbol{E}}^p\| \leq \bar{\gamma} \quad \text{in } \Omega^0 \tag{5.19e}$$

which fits into the non-linear second-order cone programming format of (2.29) where the last constraint is expressed as a second-order Lorentz cone constraint $(\bar{\gamma}, \bar{\boldsymbol{E}}^p) \in \mathcal{Q}^7$. Let us finally remark that constraints (5.19b) and (5.19c) can be eliminated and $\bar{\boldsymbol{E}}$ can be replaced by its non-linear expression as a function $\bar{\boldsymbol{u}}$:

$$\bar{\boldsymbol{E}} = \bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) = \frac{1}{2}\ln\left((\boldsymbol{F}_n + \nabla_X \bar{\boldsymbol{u}})^T \cdot (\boldsymbol{F}_n + \nabla_X \bar{\boldsymbol{u}})\right) - \boldsymbol{E}_n \tag{5.20}$$

so that the final problem expressed in the format of (2.29) involves the unknowns $\boldsymbol{x} = (\bar{\boldsymbol{u}}, \bar{\boldsymbol{E}}^p, \bar{\gamma})$, linear constraints associated with the kinematic boundary condition (5.19d) later expressed as $\boldsymbol{A}\bar{\boldsymbol{u}} = \boldsymbol{b}$, no non-linear constraints $\boldsymbol{g}(\boldsymbol{x})$ and the non-linear smooth objective function:

$$f(\boldsymbol{x}) = \frac{1}{2}(\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) : \mathbb{D} : (\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) + \boldsymbol{T}_n : (\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) \tag{5.21}$$
$$+ \frac{1}{3}E_h(\bar{\gamma})^2 + \sqrt{\frac{2}{3}}\sigma_{Y,n}\bar{\gamma} - P_{ext}(\bar{\boldsymbol{u}})$$

5.2.5 Residuals and KKT system

Let us now detail the expression of the residuals and the associated KKT conditions around a current iterate k. We will denote by $\star^{(k)}$ the value of a quantity $\star$ at this iterate e.g. $\boldsymbol{F}^{(k)} = \boldsymbol{F}_n + \boldsymbol{\nabla}_{\mathrm{X}}\,\bar{\boldsymbol{u}}^{(k)}$. First, the variation of the total logarithmic strain in direction $\delta\bar{\boldsymbol{u}}$ is given by:

$$\delta\bar{\boldsymbol{E}} = \mathbb{M}^{(k)} : \delta\bar{\boldsymbol{E}}_{GL} = \mathbb{M}^{(k)} : \mathrm{sym}((\boldsymbol{F}^{(k)})^T \cdot \boldsymbol{\nabla}_{\mathrm{X}}\,\delta\bar{\boldsymbol{u}}) \tag{5.22}$$

with $\mathrm{sym}(\boldsymbol{A}) = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)$
The variation of the objective function $f(\boldsymbol{x})$ in direction $\delta\boldsymbol{x}$ is given by:

$$(\nabla_{\boldsymbol{x}} f, \delta\boldsymbol{x}) = (\nabla_{\bar{\boldsymbol{u}}} f, \delta\bar{\boldsymbol{u}}) + (\nabla_{\bar{\boldsymbol{E}}^p} f, \delta\bar{\boldsymbol{E}}^p) + (\nabla_{\bar{\gamma}} f, \delta\bar{\gamma}) \tag{5.23}$$

where:

$$
\begin{aligned}
(\nabla_{\bar{\boldsymbol{u}}} f, \delta\bar{\boldsymbol{u}}) &= \int_{\Omega^0} \left(\mathbb{D} : (\bar{\boldsymbol{E}}^{(k)} - \mathbb{K} : \bar{\boldsymbol{E}}^{p,(k)}) + \boldsymbol{T}_n\right) : \delta\bar{\boldsymbol{E}} d\Omega - P_{ext}(\delta\bar{\boldsymbol{u}}) \\
&= \int_{\Omega^0} \boldsymbol{T}^{(k)} : \mathbb{M}^{(k)} : \mathrm{sym}((\boldsymbol{F}^{(k)})^T \cdot \boldsymbol{\nabla}_{\mathrm{X}}\, \delta\bar{\boldsymbol{u}}) d\Omega - P_{ext}(\delta\bar{\boldsymbol{u}}) \qquad (5.24) \\
(\nabla_{\bar{\boldsymbol{E}}^p} f, \delta\bar{\boldsymbol{E}}^p) &= \int_{\Omega^0} \left(-\mathbb{D} : (\bar{\boldsymbol{E}}^{(k)} - \mathbb{K} : \bar{\boldsymbol{E}}^{p,(k)}) - \boldsymbol{T}_n\right) : \mathbb{K} : \delta\bar{\boldsymbol{E}}^p d\Omega \\
&= \int_{\Omega^0} -\,\mathrm{dev}(\boldsymbol{T}^{(k)}) : \delta\bar{\boldsymbol{E}}^p d\Omega \qquad (5.25) \\
(\nabla_{\bar{\gamma}} f, \delta\bar{\gamma}) &= \int_{\Omega^0} \left(\frac{2}{3} E_h \bar{\gamma}^{(k)} + \sqrt{\frac{2}{3}}\sigma_{Y,n}\right) \delta\bar{\gamma} d\Omega \qquad (5.26)
\end{aligned}
$$

where we introduced $\boldsymbol{T}^{(k)}$ the value of the stress tensor at iterate k given by:

$$
\boldsymbol{T}^{(k)} = \boldsymbol{T}_n + \mathbb{D} : (\bar{\boldsymbol{E}}^{(k)} - \mathbb{K} : \bar{\boldsymbol{E}}^{p,(k)})
$$

The global conic constraints $\boldsymbol{x} \in \mathcal{K}$ correspond here to $\bar{\boldsymbol{u}} \in \mathbb{R}^n$ (free variable) and $(\bar{\gamma}, \bar{\boldsymbol{E}}^p) \in \mathcal{Q}^7$ so that the dual conic variable $\boldsymbol{s} = (\boldsymbol{s}_{\bar{\boldsymbol{u}}}, \boldsymbol{s}_{\bar{\boldsymbol{E}}^p}, \boldsymbol{s}_{\bar{\gamma}}) \in \mathcal{K}^*$ are such that $\boldsymbol{s}_{\bar{\boldsymbol{u}}} = 0$ and $(\boldsymbol{s}_{\bar{\gamma}}, \boldsymbol{s}_{\bar{\boldsymbol{E}}^p}) \in \mathcal{Q}^7$. As a result, one obtains the following expressions for the first two residuals of (2.31):

$$
\boldsymbol{r}_d^{(k)} = \begin{Bmatrix} \boldsymbol{r}_{d,\bar{\boldsymbol{u}}}^{(k)} \\ \boldsymbol{r}_{d,\bar{\boldsymbol{E}}^p}^{(k)} \\ \boldsymbol{r}_{d,\bar{\gamma}}^{(k)} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{\nabla}_{\bar{\boldsymbol{u}}} f^{(k)} + \boldsymbol{A}^T \boldsymbol{y}^{(k)} - \boldsymbol{s}_{\bar{\boldsymbol{u}}}^{(k)} \\ \boldsymbol{\nabla}_{\bar{\boldsymbol{E}}^p} f^{(k)} - \boldsymbol{s}_{\bar{\boldsymbol{E}}^p}^{(k)} \\ \boldsymbol{\nabla}_{\bar{\gamma}} f^{(k)} - \boldsymbol{s}_{\bar{\gamma}}^{(k)} \end{Bmatrix} \qquad (5.27)
$$

$$
\boldsymbol{r}_p^{(k)} = \boldsymbol{A}\bar{\boldsymbol{u}}^{(k)} - \boldsymbol{b} \qquad (5.28)
$$

Expressing that the residuals should be zero for the solution, the three blocks of $\boldsymbol{r}_d^{(k)}$ respectively yield:

$$
\begin{aligned}
&\int_{\Omega^0} \boldsymbol{T}^{(k)} : \mathbb{M}^{(k)} : \mathrm{sym}((\boldsymbol{F}^{(k)})^T \cdot \boldsymbol{\nabla}_{\mathrm{X}}\, \delta\bar{\boldsymbol{u}}) d\Omega \\
&\qquad - P_{ext}(\delta\bar{\boldsymbol{u}}) + \int_{\Gamma_{\boldsymbol{u}}^0} \boldsymbol{y}^{(k)} \cdot \delta\bar{\boldsymbol{u}} dS = 0 \quad \forall \delta\bar{\boldsymbol{u}}
\end{aligned} \qquad (5.29)
$$

$$
\boldsymbol{s}_{\bar{\boldsymbol{E}}^p}^{(k)} = -\,\mathrm{dev}(\boldsymbol{T}^{(k)}) \qquad (5.30)
$$

$$
\boldsymbol{s}_{\bar{\gamma}}^{(k)} = \sqrt{\frac{2}{3}}\sigma_{Y,n} + \frac{2}{3} E_h \bar{\gamma}^{(k)} \qquad (5.31)
$$

The first block expresses the virtual work principle whereas the last two, combined with the conic constraint $\|\boldsymbol{s}_{\bar{\boldsymbol{E}}^p}\| \leq \boldsymbol{s}_{\bar{\gamma}}$ express the plastic yield criterion:

$$\sqrt{\frac{3}{2}}\|\mathrm{dev}(\boldsymbol{T}^{(k)})\| \leq \sigma_{Y,n} + \sqrt{\frac{2}{3}}E_h\bar{\gamma}^{(k)} \tag{5.32}$$

One can also easily check that, combined with the complementarity condition $\boldsymbol{x}^T\boldsymbol{s} = 0$ of the Lorentz cone (B.7), these relations are equivalent to the plastic consistency condition and plastic flow rule:

$$\bar{\gamma}^{(k)}\left(\left(\sigma_{Y,n} + \sqrt{\frac{2}{3}}E_h\bar{\gamma}^{(k)}\right)^2 - \frac{3}{2}\|\mathrm{dev}(\boldsymbol{T}^{(k)})\|^2\right) = 0 \tag{5.33}$$

$$\bar{\boldsymbol{E}}^{p,(k)} = \frac{\sqrt{\frac{3}{2}}\bar{\gamma}^{(k)}}{\sigma_{Y,n} + \sqrt{\frac{2}{3}}E_h\bar{\gamma}^{(k)}}\,\mathrm{dev}(\boldsymbol{T}) \tag{5.34}$$

where it is clear that $\bar{\gamma}^{(k)} = \|\bar{\boldsymbol{E}}^{p,(k)}\| = \sqrt{\frac{2}{3}}\bar{p}^{(k)}$ during a plastic evolution.
Finally, the Jacobian system (2.43) reads in this case as:

$$\boldsymbol{J}^{(k)} = \begin{bmatrix} \boldsymbol{H}^{(k)} & \boldsymbol{A}^T & -\boldsymbol{I} \\ \boldsymbol{A} & 0 & 0 \\ \boldsymbol{S}^{(k)} & 0 & \boldsymbol{X}^{(k)} \end{bmatrix} \tag{5.35}$$

in which the Hessian $\boldsymbol{H}^{(k)}$ is obtained using the chain rule resulting in a sum of six different contributions, the first five corresponding to the material contribution to the stiffness matrix and the sixth being the geometrical stiffness matrix:

$$\begin{aligned}
\Delta\bar{\boldsymbol{x}}\cdot\boldsymbol{H}^{(k)}\cdot\delta\bar{\boldsymbol{x}} = &\int_{\Omega^0}\Big[\mathrm{sym}((\boldsymbol{F}^{(k)})^T\cdot\boldsymbol{\nabla}_{\mathrm{X}}\,\Delta\bar{\boldsymbol{u}}) : \mathbb{D}^{(k)}_{GL} : \mathrm{sym}((\boldsymbol{F}^{(k)})^T\cdot\boldsymbol{\nabla}_{\mathrm{X}}\,\delta\bar{\boldsymbol{u}})\Big]d\Omega \\
&-\int_{\Omega^0}\Big[\Delta\bar{\boldsymbol{E}}^p : (\mathbb{K}:\mathbb{D}:\mathbb{M}^{(k)}) : \mathrm{sym}((\boldsymbol{F}^{(k)})^T\cdot\boldsymbol{\nabla}_{\mathrm{X}}\,\delta\bar{\boldsymbol{u}})\Big]d\Omega \\
&-\int_{\Omega^0}\Big[\mathrm{sym}((\boldsymbol{F}^{(k)})^T\cdot\boldsymbol{\nabla}_{\mathrm{X}}\,\Delta\bar{\boldsymbol{u}}) : (\mathbb{M}^{(k)}:\mathbb{D}:\mathbb{K}) : \delta\bar{\boldsymbol{E}}^p\Big]d\Omega \\
&+\int_{\Omega^0}\Big[\Delta\bar{\boldsymbol{E}}^p : (\mathbb{K}:\mathbb{D}:\mathbb{K}) : \delta\bar{\boldsymbol{E}}^p\Big]d\Omega \\
&+\int_{\Omega^0}\Big[\Delta\bar{\gamma}.\frac{2}{3}E_h.\delta\bar{\gamma}\Big]d\Omega \\
&+\int_{\Omega^0}\Big[\boldsymbol{T}^{(k)} : \mathbb{M}^{(k)} : (\boldsymbol{\nabla}_{\mathrm{X}}\,\delta\bar{\boldsymbol{u}})^T\cdot\boldsymbol{\nabla}_{\mathrm{X}}\,\Delta\bar{\boldsymbol{u}})\Big]d\Omega
\end{aligned} \tag{5.36}$$

with the tangent Green-Lagrange stiffness $\mathbb{D}_{GL}^{(k)}$ being given by (5.7):

$$\mathbb{D}_{GL}^{(k)} = \mathbb{M}^{(k)} : \mathbb{D} : \mathbb{M}^{(k)} + \boldsymbol{T}^{(k)} : \mathfrak{L}^{(k)} \tag{5.37}$$

5.2.6 Finite-element implementation

The finite-element discretization of (5.19) is quite standard and very close to the kinematic approach discussed in Sec. 4.5.1. In particular, we used 10-noded quadratic tetrahedra for the displacement field interpolation. The logarithmic plastic strain unknowns correspond to its six components expressed at all quadrature points (4 Gauss points per tetrahedron in the present case) and similarly to the additional scalar variable $\Delta\bar{\gamma}$. As discussed in chapter 4, the equations involving $\Delta\bar{\boldsymbol{E}}^p$ and $\Delta\bar{\gamma}$ in the KKT system are all of local nature i.e. they are expressed at the quadrature point level and are all uncoupled. They can therefore be easily condensed when forming system (2.56) yielding a further reduced system involving only the displacement variables and the boundary condition Lagrange multipliers as final unknowns. As a consequence, the proposed primal-dual IPM exhibits the same computational cost as regards linear system resolutions as a standard Newton-Raphson method.
Let $\boldsymbol{x}$ be the aggregation of DoF for a tetrahedron i.e. for $n_G = 4$:

$$\boldsymbol{x} = (30\Delta\bar{\boldsymbol{u}}\ , 24\Delta\bar{\boldsymbol{E}}^p\ , 4\Delta\bar{\gamma})$$

The different differential operators in matrix notation are then given by:

$$\mathrm{sym}((\boldsymbol{F}^{(k)})^T \cdot \boldsymbol{\nabla}_{\mathrm{X}}\, \Delta\bar{\boldsymbol{u}}) = \boldsymbol{B}^{(k)}\Delta\bar{\boldsymbol{u}}^e \tag{5.38}$$

$$\boldsymbol{\nabla}_{\mathrm{X}}\, \Delta\bar{\boldsymbol{u}} = \boldsymbol{G}^{(k)}\Delta\bar{\boldsymbol{u}}^e \tag{5.39}$$

The elementary material stiffness matrix for each tetrahedron can then be calculated taking into account the traceless condition for $\Delta\bar{\boldsymbol{E}}^p$:

$$(\boldsymbol{H}^{(k)})^e = \left[\begin{array}{c|c|c} (\boldsymbol{B}^{(k)})^T\mathbb{D}_{GL}\boldsymbol{B}^{(k)} & -(\boldsymbol{B}^{(k)})^T\mathbb{M}^{(k)}\mathbb{D}\mathbb{K} & \boldsymbol{0} \\ +(\boldsymbol{G}^{(k)})^T(\boldsymbol{T}^{(k)}\mathbb{M}^{(k)})\boldsymbol{G}^{(k)} & & \\ \hline -\Big((\boldsymbol{B}^{(k)})^T\mathbb{M}^{(k)}\mathbb{D}\mathbb{K}\Big)^T & \mathbb{K}\mathbb{D}\mathbb{K} & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{0} & \dfrac{2}{3}E_h \end{array}\right] \tag{5.40}$$

Considering that the external work is expressed as follows:

$$P_{ext}(\bar{\boldsymbol{u}}) = \int_{\Omega^0} \boldsymbol{b}_{d,n+1}.\bar{\boldsymbol{u}}d\Omega + \int_{\Gamma_{\boldsymbol{t}}^0} \boldsymbol{t}_{d,n+1}.\bar{\boldsymbol{u}}d\Gamma \tag{5.41}$$

On can derive the nodal force vector $\boldsymbol{f}_{n+1}$ such that:

$$P_{ext}(\delta\bar{\boldsymbol{u}}) = \boldsymbol{f}_{n+1}\delta\bar{\boldsymbol{u}} \tag{5.42}$$

The variation of the objective function $f(\boldsymbol{x})$ in direction $\delta\boldsymbol{x}$ is given then by:

$$\nabla_{\boldsymbol{x}} f^{(k)} = \begin{Bmatrix} \boldsymbol{\nabla}_{\bar{\boldsymbol{u}}} f^{(k)} \\ \boldsymbol{\nabla}_{\bar{\boldsymbol{E}}^p} f^{(k)} \\ \boldsymbol{\nabla}_{\bar{\gamma}} f^{(k)} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{T}^{(k)} \mathbb{M}^{(k)} \boldsymbol{B}^{(k)} - \boldsymbol{f}_{n+1} \\ \mathbb{K} \boldsymbol{T}^{(k)} \\ \frac{2}{3} E_h \bar{\gamma}^{(k)} \end{Bmatrix} \tag{5.43}$$

5.2.7 Continuation methods

Often in non-linear analysis a *continuation method* is required to obtain solution paths. The easiest consists in a simple radial incrementation of the load factor $\lambda_{n+1} = \lambda_n + \bar{\lambda}$, hence $\boldsymbol{t}_{d,n+1} = \lambda_{n+1} \boldsymbol{t}_d$ is the external load vector in the case of a force-driven continuation method and $\boldsymbol{u}_{d,n+1} = \lambda_{n+1} \boldsymbol{u}_d$ in a displacement-driven one.

More complex approaches such as the arc-length method consider the load factor $\bar{\lambda}$ as a variable. Then, an extra new constraint equation, known as the arc-length constraint, is added to the principle of virtual work for defining the next equilibrium point solution as an intersection between the solution path and the constraint equation of the form:

$$g(\bar{\lambda}, \bar{\boldsymbol{u}}) = \theta^2 (\lambda_n + \bar{\lambda})^2 \boldsymbol{t}_d^T \boldsymbol{t}_d + (\boldsymbol{u}_n + \bar{\boldsymbol{u}})^T (\boldsymbol{u}_n + \bar{\boldsymbol{u}}) - l^2 = 0 \tag{5.44}$$

where θ is a scaling parameter and l is the prescribed arc-length. When $\theta = 0$ the constraint is known as the fixed normal plane constraint or the cylindrical arc-length method.

The non-linear smooth objective function $f(\boldsymbol{x})$ is then changed to include the additional variable $\bar{\lambda}$:

$$\begin{aligned} f(\boldsymbol{x}) = {} & \frac{1}{2} (\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) : \mathbb{D} : (\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) + \boldsymbol{T}_n : (\bar{\boldsymbol{E}}(\bar{\boldsymbol{u}}) - \mathbb{K} : \bar{\boldsymbol{E}}^p) \\ & + \frac{1}{3} E_h (\bar{\gamma})^2 + \sqrt{\frac{2}{3}} \sigma_{Y,n} \bar{\gamma} - P_{ext}(\bar{\lambda}, \bar{\boldsymbol{u}}) \end{aligned} \tag{5.45}$$

with the external work being expressed as follows:

$$P_{ext}(\bar{\lambda}, \bar{\boldsymbol{u}}) = \int_{\Omega^0} (\lambda_n + \bar{\lambda}) \boldsymbol{b}.\bar{\boldsymbol{u}} d\Omega + \int_{\Gamma_{\boldsymbol{t}}^0} (\lambda_n + \bar{\lambda}) \boldsymbol{t}_d.\bar{\boldsymbol{u}} d\Gamma = (\lambda_n + \bar{\lambda}) \boldsymbol{f}_0 \bar{u} \tag{5.46}$$

$\boldsymbol{f}_0$ being the initial nodal force vector.

The variation of the objective function $f(\boldsymbol{x})$ in direction $\delta\boldsymbol{x}$ given by Eq. (5.23) is adjusted to include the new direction $\delta\bar{\lambda}$:

$$(\nabla_{\boldsymbol{x}} f, \delta\boldsymbol{x}) = (\nabla_{\bar{\lambda}} f, \delta\bar{\lambda}) + (\nabla_{\bar{\boldsymbol{u}}} f, \delta\bar{\boldsymbol{u}}) + (\nabla_{\bar{\boldsymbol{E}}^p} f, \delta\bar{\boldsymbol{E}}^p) + (\nabla_{\bar{\gamma}} f, \delta\bar{\gamma}) \tag{5.47}$$

The residuals and Hessian matrix are then adjusted to include the gradient and the Hessian of the non-linear constraint added as seen in Sec. 2.5.3. Thus the variation of the objective function $f(\boldsymbol{x})$ in direction $\delta\boldsymbol{x}$ is given then by:

$$\nabla_{\boldsymbol{x}} f^{(k)} = \begin{Bmatrix} \boldsymbol{\nabla}_{\bar{\lambda}} f^{(k)} \\ \boldsymbol{\nabla}_{\bar{\boldsymbol{u}}} f^{(k)} \\ \boldsymbol{\nabla}_{\bar{\boldsymbol{E}}^p} f^{(k)} \\ \boldsymbol{\nabla}_{\bar{\gamma}} f^{(k)} \end{Bmatrix} = \begin{Bmatrix} -\boldsymbol{f}_0^T \bar{\boldsymbol{u}}^{(k)} \\ \boldsymbol{T}^{(k)} \mathbb{M}^{(k)} \boldsymbol{B}^{(k)} - \bar{\lambda}^{(k)} \boldsymbol{f}_0 \\ \mathbb{K} \boldsymbol{T}^{(k)} \\ \frac{2}{3} E_h \bar{\gamma}^{(k)} \end{Bmatrix} \tag{5.48}$$

The Hessian of the objective function is given by:

$$\boldsymbol{\nabla}^2_{xx} f^{(k)} = \left[\begin{array}{c|c|c|c} 0 & -\boldsymbol{f}_0^T & \mathbf{0} & \mathbf{0} \\ \hline -\boldsymbol{f}_0 & \begin{array}{c}(\boldsymbol{B}^{(k)})^T \mathbb{D}_{GL} \boldsymbol{B}^{(k)} \\ +(\boldsymbol{G}^{(k)})^T (\boldsymbol{T}^{(k)} \mathbb{M}^{(k)}) \boldsymbol{G}^{(k)}\end{array} & -(\boldsymbol{B}^{(k)})^T \mathbb{M}^{(k)} \mathbb{D}\mathbb{K} & \mathbf{0} \\ \hline \mathbf{0} & -\Big((\boldsymbol{B}^{(k)})^T \mathbb{M}^{(k)} \mathbb{D}\mathbb{K}\Big)^T & \mathbb{K}\mathbb{D}\mathbb{K} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \dfrac{2}{3} E_h \end{array}\right] \tag{5.49}$$

As for the non-linear equality constraint, the associated gradient vector is given by:

$$\boldsymbol{G}^{(k)} = \left[\begin{array}{c|c|c|c} 2\theta^2 \bar{\lambda}^{(k)} \boldsymbol{f}_0^T \boldsymbol{f}_0 & 2\bar{\boldsymbol{u}}^{(k)} & \mathbf{0} & \mathbf{0} \end{array}\right] \tag{5.50}$$

its Hessian matrix being:

$$\boldsymbol{\nabla}^2_{xx} g^{(k)} = \left[\begin{array}{c|c|c|c} 2\theta^2 \boldsymbol{f}_0^T \boldsymbol{f}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & 2\boldsymbol{I} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \tag{5.51}$$

The Hessian of the Lagrangian can be found using (2.44) and the dual residual using (2.45a) where the contribution of the non-linear constraint should be taken into account. The primal residual vector associated to the non-linear constraint and given by Eq. (2.45c) is the following:

$$r'^{(k)}_p = g(\boldsymbol{x}^{(k)}) = \theta^2 (\bar{\lambda}^{(k)})^2 \boldsymbol{f}_0^T \boldsymbol{f}_0 + (\bar{\boldsymbol{u}}^{(k)})^T \bar{\boldsymbol{u}}^{(k)} - l^2 \tag{5.52}$$

5.2.8 Initialization points of the IPM

Because of the incremental aspect of the mechanical problem, once the solution of the incremental problem for a specific load step, say, t_n, is obtained, loading conditions as well as state variables, e.g. $\sigma_{Y,n}$, are updated before solving the new incremental problem for the next step t_{n+1}. It is usual that the change in data of these problems is small so that solutions are often close from each other between consecutive time steps. Exploiting a good initial guess of the solution, also known as *warm-start*, is a current challenge for interior point methods due to the fact that conic variables should initially be feasible and ideally far from the feasible region boundary.

In our implementation, all variables are initialized with the converged values of the previous load step, except the conic variables $(\boldsymbol{x}, \boldsymbol{s})$. Indeed, in order to have a starting

point located far from the feasible region boundary, the initial value for the next load step is taken as follows:

$$\left.\begin{array}{l}(x_0, \bar{\mathbf{x}})_{n+1}^{(0)} = (x_0, \beta\bar{\mathbf{x}})_n \\ (s_0, \bar{\mathbf{s}})_{n+1}^{(0)} = (s_0, \beta\bar{\mathbf{s}})_n\end{array}\right\} \text{ with } \beta = 0.7 \text{ typically} \tag{5.53}$$

5.3 Illustrative examples

5.3.1 Membrane effect in a fully-clamped elastoplastic beam

In this example, we consider a fully clamped rectangular beam of length $L = 2.0\,\mathrm{m}$ oriented in direction x, of height $h = 0.1\,\mathrm{m}$ and width $b = 0.04\,\mathrm{m}$. The beam consists of a von Mises plastic material with no hardening and is subject to a uniformly distributed body force $\boldsymbol{b} = -f\boldsymbol{e}_z$. The material parameter values are summarized in Tab. 5.1.

Table 5.1 – Beam material properties

Young modulus	$E = 210\,\mathrm{GPa}$
Poisson's ratio	$\nu = 0.30$
Initial yield stress	$\sigma_0 = 250\,\mathrm{MPa}$
Hardening modulus	$h = 0.00\,\mathrm{MPa}$

The present implementation results have been compared with computations using the commercial finite-element software Abaqus and also using the open-source finite-element platform FEniCS [Alnæs et al., 2015; Logg et al., 2012; Logg and Wells, 2010] coupled with MFront [Helfer et al., 2020] for the constitutive behavior integration.

The FEniCS/MFront implementation consists in a total-Lagrangian implementation of logarithmic plasticity using a standard Newton-Raphson/return mapping procedure. The coupling between both libraries relies on the `MFrontGenericInterfaceSupport` project [Helfer et al., 2020].

The Abaqus implementation relies on an updated-Lagrangian formulation using the Cauchy stress tensor and its work-conjugate rate of deformation. The integration technique for the total deformation gives the logarithmic strains (`LE` in Abaqus notation) which is used in the case of metal plasticity. For more details, one can refer to sections 1.4 and 1.5 in the Abaqus theory manual [Smith, 2009].

We monitor the evolution of the mid-span deflection u and the horizontal support reaction force H when increasing the body force up to $f = 50\,\mathrm{MN/m^3}$. The evolutions of u and H have been represented in Fig. 5.2. It can first be observed that all three different implementations yield very similar results, the slight difference observed with respect to the FEniCS/MFront computations can be attributed to the fact that a different mesh, although of similar element size, was used. The obtained results clearly exhibit a first elastic then plastic stage (for a load factor below 0.5) when geometrical non-linear effects do not play an important role. A secondary stiffening stage (for load factor larger than

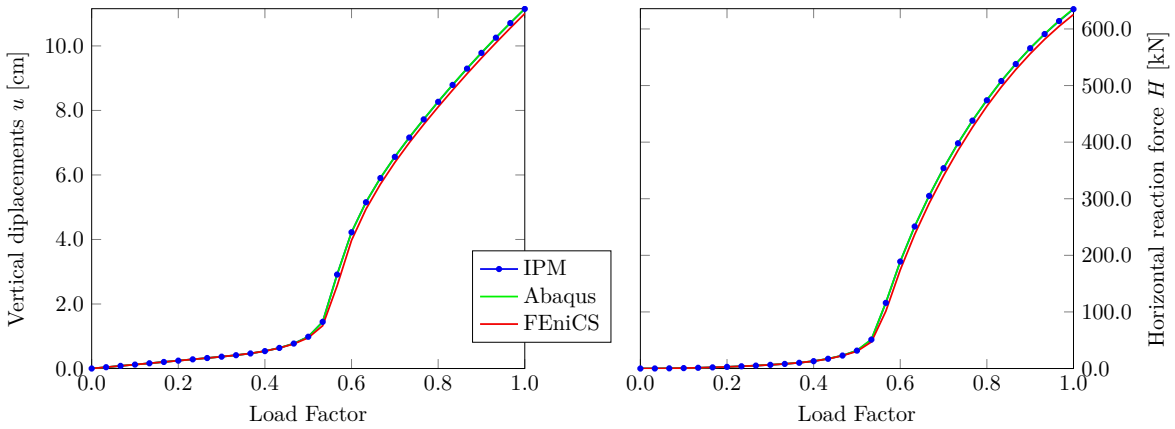


Figure 5.2 – Evolution of the mid-span deflection u and the horizontal reaction force H

Table 5.2 – Total number of iterations for the different methods and load-stepping

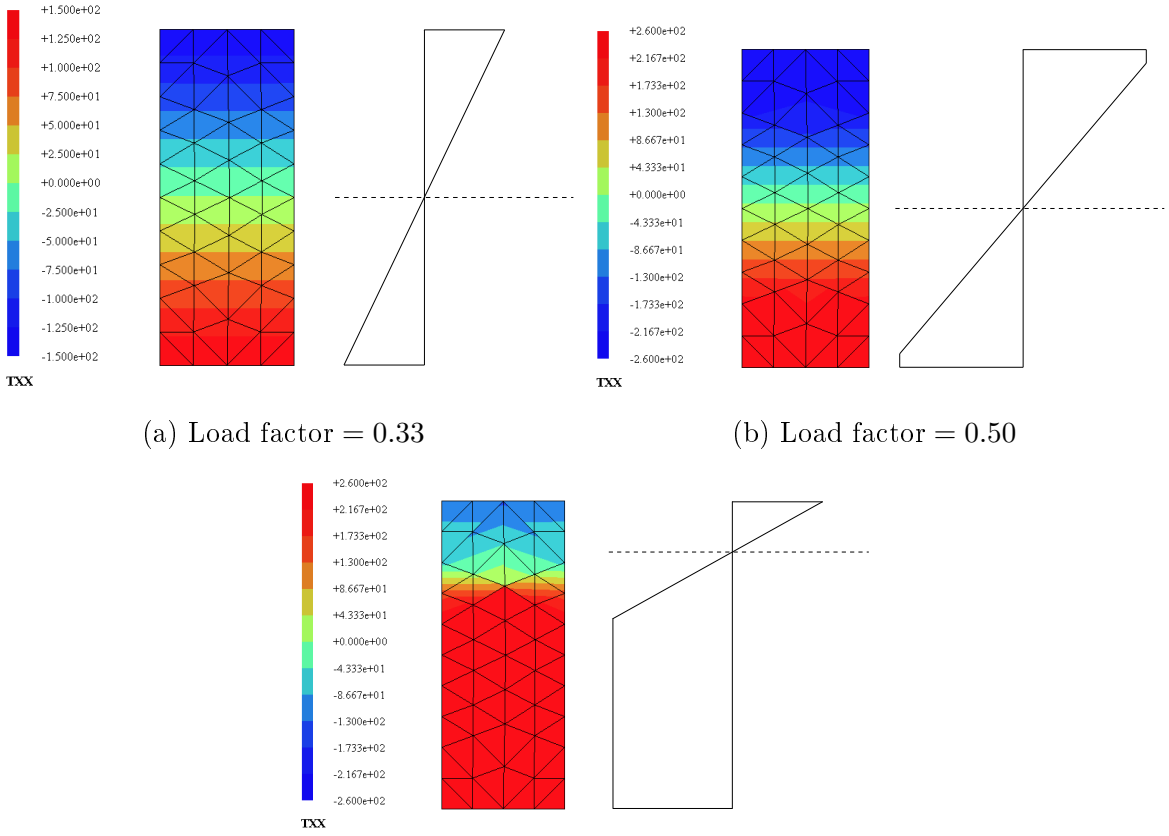
Method	Total iterations
FEniCS (30 steps)	128
IPM (30 steps)	385
IPM (15 steps)	192
IPM (5 steps)	70

0.5) is then observed due to membrane catenary effect (see the increase of the horizontal reaction force) when geometrical non-linear effects become more and more important. This is further confirmed when inspecting the normal stress diagram along the mid-span cross-section (Fig. 5.3) showing an elastic stage solution (Fig. 5.3a), the onset of cross-section yielding in pure bending (Fig. 5.3b) and finally a membrane-dominated plastic stage (Fig. 5.3c). The final deformed configuration has also been represented in Fig. 5.4.

In order to assess the numerical solution procedure, we will compare the FEniCS solution with the proposed IPM solution since both approaches rely on a total-Lagrangian formulation. In particular, we compared the number of iterations per load step to reach convergence using the same relative residual tolerance. It must be recalled that, apart from the way boundary conditions are handled, the linear system size, and hence the cost per iteration, is similar for both methods.

Results are reported in Fig. 5.5 where it can be observed that the required number of iterations is much larger for the IPM than for the Newton method used in FEniCS for 30 load steps. This is by no means surprising due to the quadratic convergence of the Newton method close to a solution. It can also be observed that the required number of iterations increases in the second stage of the problem where plasticity and geometrically non-linear effects become much more dominant.

However, an extremely interesting feature of the IPM is its robustness over large



(a) Load factor = 0.33

(b) Load factor = 0.50

(c) Load factor = 1.0

Figure 5.3 – generalized normal stress component T_{XX} in the middle-section for different load factors

load steps. Indeed, the Newton method was unable to converge with less than 20 load steps whereas the IPM method could converge using only 5 load steps. Besides, this robustness does not seem to deteriorate the convergence quality since roughly the same number of iterations is required for the same load level when using smaller load steps (see again Fig. 5.5). Overall, the total number of iterations (Tab. 5.2) using 5 load steps becomes competitive compared with the Newton method, whereas the Newton method is more efficient than the IPM with similar load-stepping. Moreover, Fig. 5.6 clearly shows that using fewer load steps yields similar values for the displacement and reaction forces.

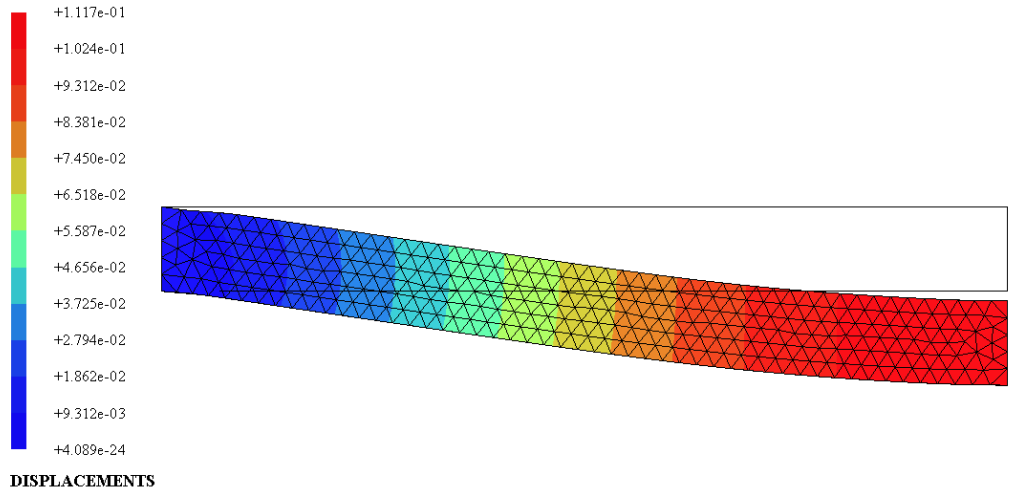


Figure 5.4 – Deformed configuration and displacement isovalues in m (Load factor = 1.0)

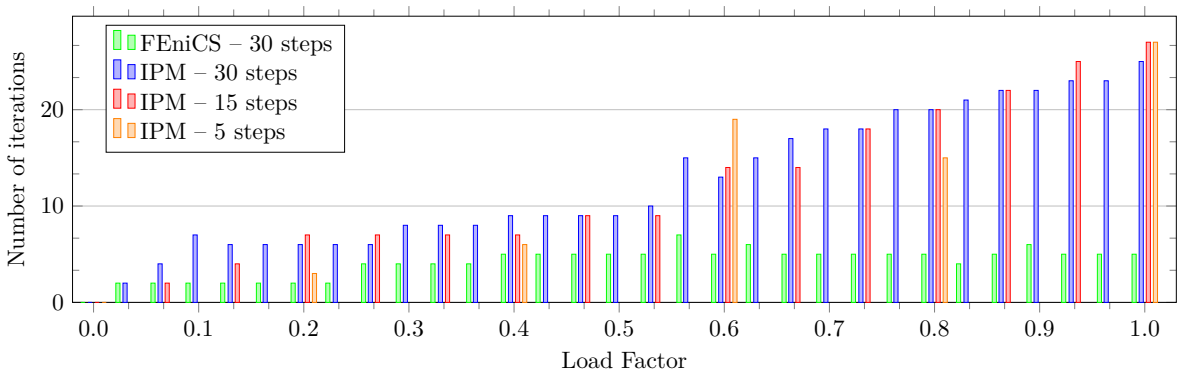


Figure 5.5 – Number of iterations per load step

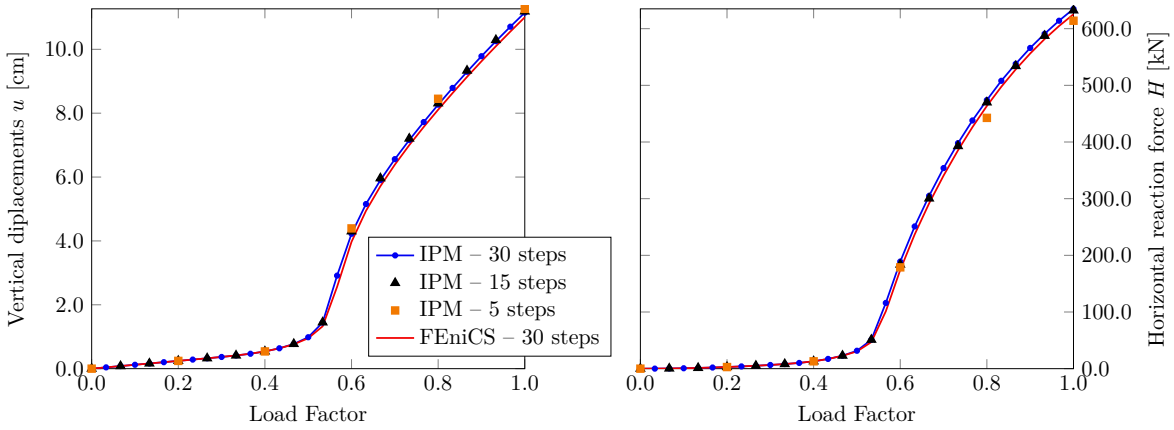


Figure 5.6 – Comparison of the evolution of the mid-span deflection u and the horizontal reaction force H using 30, 15 and 5 load steps

5.3.2 Necking of a rod

We consider the rod-necking problem, a standard benchmark problem of finite plasticity that have been used by various authors [Miehe et al., 2002; Papadopoulos and Lu, 1998; Simo, 1992]. The goal of this example is to compare the results given by the IPM algorithm with reference results found in Miehe et al. [2002] and using an Abaqus implementation. The initial length of the rod is $l = 53.34\,\text{mm}$, the radius $r_0 = 6.4135\,\text{mm}$. The necking is triggered by an initial imperfection of the rod in the middle section represented by a continuous decrease of the radius to $r = 0.982 r_0$ over a length of 8.98 mm.

The constitutive response of the material is characterized by the logarithmic strain plasticity formulation using a von Mises yield criterion and isotropic hardening. The following saturation-type non-linear isotropic hardening is considered:

$$\sigma_Y(p) = hp + (\sigma_\infty - \sigma_0)(1 - \exp[-\omega p]) \tag{5.54}$$

As a first approach, this non-linear hardening model has been represented in our linear hardening model implementation by a piecewise-linear model. Namely, at each time step, the new value of the elastic yield limit is calculated, and the linear hardening modulus E_h is set to:

$$E_h = \left.\frac{d\sigma_Y(p)}{dp}\right|_{p=p_n} = h + (\sigma_\infty - \sigma_0)\omega \exp[-\omega p_n] \tag{5.55}$$

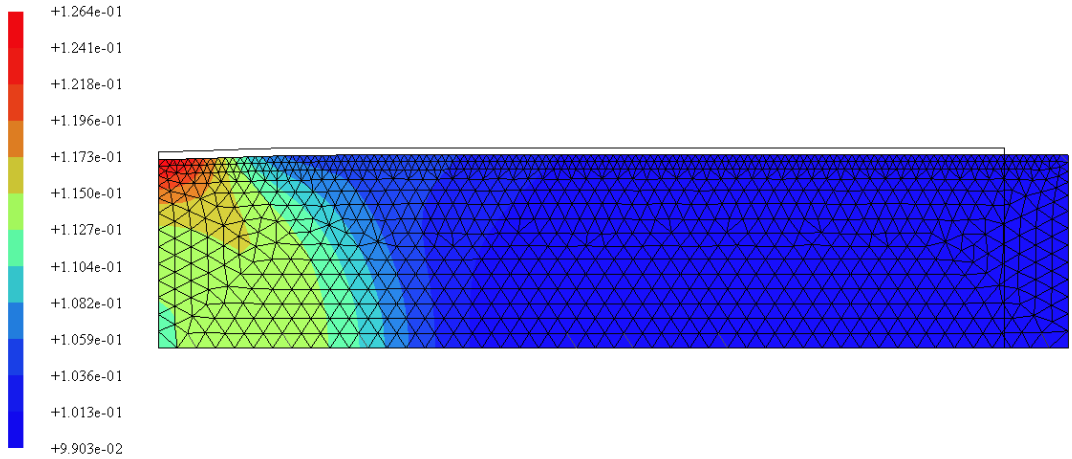
Note, however, that it would have been possible to model directly the non-linear hardening law in format (2.29). The material parameter values are summarized in Tab. 5.3.

Table 5.3 – Rod material properties

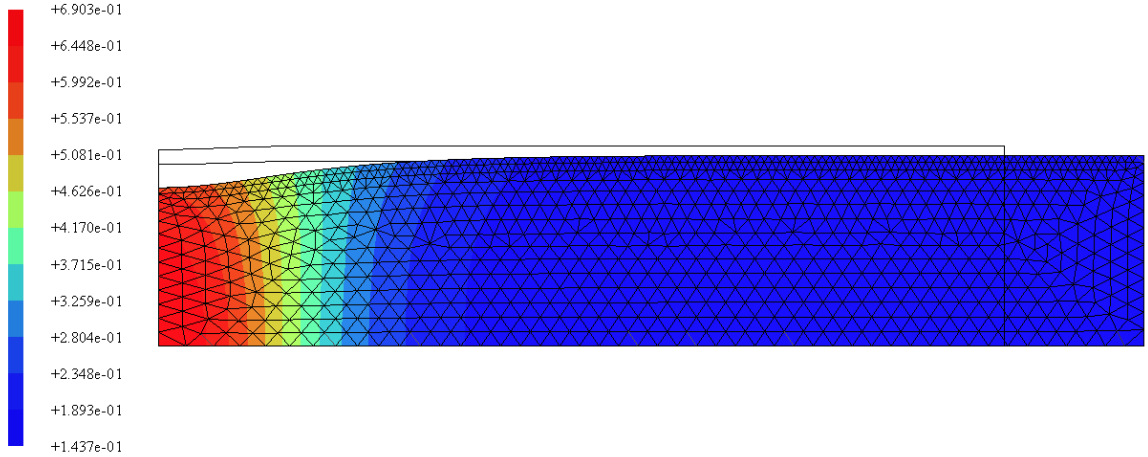
Young modulus	$E = 207.9\,\text{GPa}$
Poisson ratio	$\nu = 0.29$
Initial yield stress	$\sigma_0 = 450\,\text{MPa}$
Infinite yield stress	$\sigma_\infty = 715\,\text{MPa}$
Hardening modulus	$h = 129.24\,\text{MPa}$
Saturation parameter	$\omega = 16.93$

Due to the problem symmetry, we consider only one sixteenth of the specimen, discretized with 44 000 quadratic tetrahedra. Boundary conditions consist of symmetry conditions and a displacement-driven condition $u_x = u(t)$ where a total imposed displacement of $u = 4.5\,\text{mm}$ applied in 65 uniform load increments (approximately 0.07 mm per increment).

Fig. 5.7a and 5.7b depict two deformed meshes for $u = 2.0\,\text{mm}$ and $u = 4.5\,\text{mm}$ with the equivalent plastic strain isovalues. One can clearly see in the first figure the



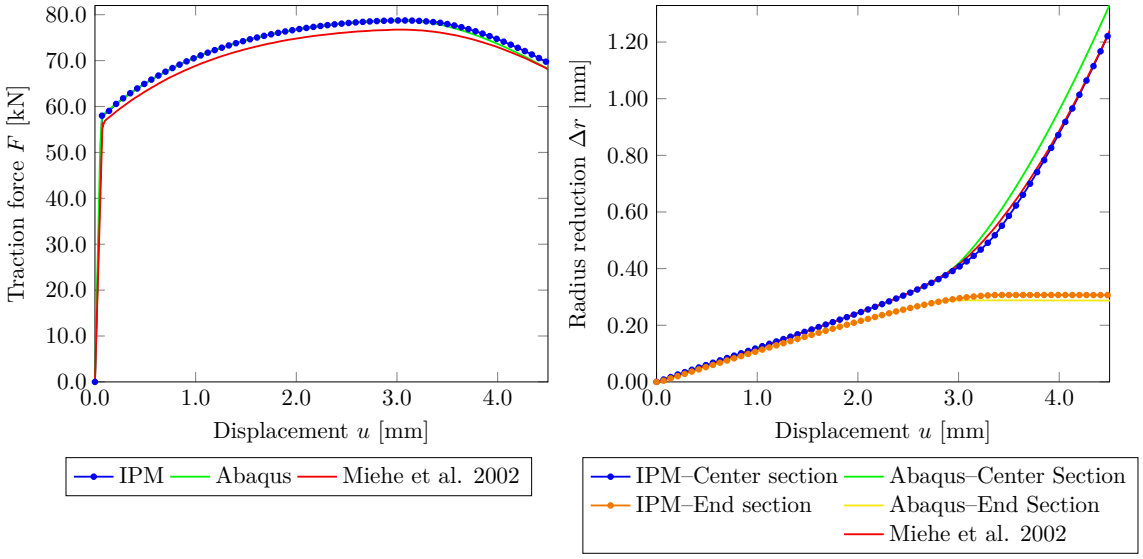
(a) $u = 2.0\,\text{mm}$



(b) $u = 4.5\,\text{mm}$

Figure 5.7 – Deformed geometry and equivalent plastic strain isovalues for different elongation values u

accumulation of plastic strain on the top of the reduced section where the necking will onset. Fig. 5.8a represent the load-deflection curve where we can clearly see that our results correspond exactly with Abaqus where as the small difference with the reference results (less than 5%) can be explained by the calculation method of the traction force F. Fig. 5.8b represents the load-deflection curve as well as the radial contraction at the center section and the end sections of the rod. The reference model, the Abaqus model and the IPM solution all provide very similar results, especially for $u \leq 3.0\,\text{mm}$, i.e. before necking occurs. In the necking stage, small differences are observed which may be attributed to different algorithm tolerances or the incrementation process. However, the difference remains less than 1% which is clearly very satisfying. Fig. 5.9a shows a



(a) Load displacement curve

(b) Radial reduction Δr of the mid and end section

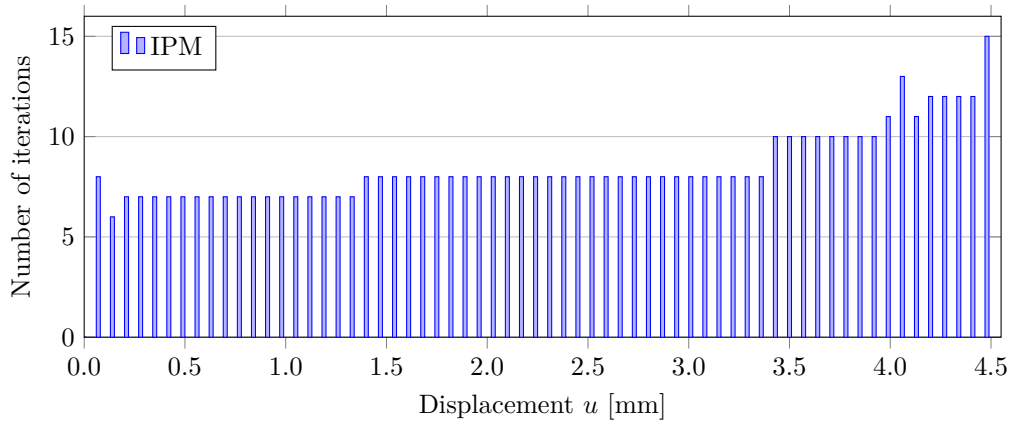
Figure 5.8 – Comparison of the results

number of iterations for the IPM ranging from 7 to 15 per increment. Iteration number slightly increases with increasing geometrical non-linearities but remains at a reasonable level. This result further confirms the robustness of the IPM with problem complexity. This aspect is one of the main advantages of the classical IPM when applied to convex problems which seems to be conserved in the present non-convex case.

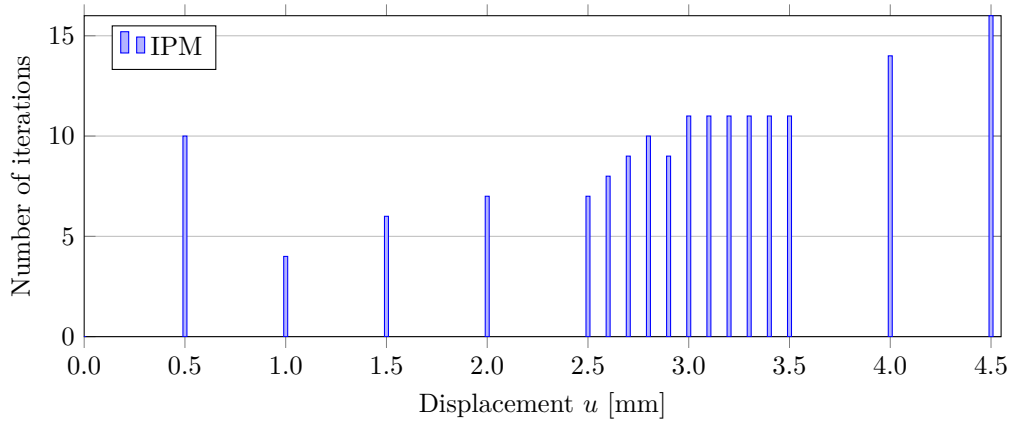
Finally, we further tested the robustness of the IPM with respect to large load steps. Obviously, one can expect that this problem is challenging for any solver due to the striction regime occurring around 3 mm. We were able to obtain a converged solution using only 5 load steps for $u \leq 2.5\,\text{mm}$, 10 load steps for u between 2.5 mm and 3.5 mm and 2 load steps between 3.5 mm and 4.5 mm. Fig. 5.9b shows the number of iterations for each of these 17 load steps. Interestingly, the number of iteration is quite similar to those of Fig. 5.9a with smaller load steps. The total number of iterations in this case is 166 against 543 for the initial load-stepping using 65 load steps. This further confirms that the IPM is particularly robust in terms of convergence properties to large load steps.

5.4 Conclusions and orientation

In this chapter, we investigated the use of a primal-dual interior-point algorithm for solving second-order cone programming problems involving non-convex objectives or constraints. A key assumption in our method relies on the fact that the additional non-convex terms are smooth, inducing additional contributions to the KKT system tangent matrix and residuals. In particular, the non-smooth terms, i.e. the conic constraints, are still assumed to be convex as for standard convex SOCP problems. Our proposed



(a) with 65 increments



(b) with 17 increments

Figure 5.9 – Number of iterations per displacement increment

algorithm is therefore a straightforward extension to the standard primal-dual IPM.

The resolution of such non-convex optimization problems has been illustrated in the case of finite-strain elastoplasticity problems relying on a logarithmic strain framework. Indeed, in such models the elastoplastic constitutive law exhibits the same expression as in the small strain-case whereas only the total strain/displacement relationship is modified. The non-smooth characteristic of the plastic law can therefore still be reformulated as conic constraints whereas the non-convex strain/displacement expression is smooth. As a result, the corresponding incremental variational problem of logarithmic strain plasticity fits into the considered optimization problem format.

Implementation has been validated on numerical benchmarks and compared to standard Newton-type procedures. Since both approaches require the resolution of a Newton system of similar size, computational efficiency can be assessed by comparing the required number of iterations. Our results show that the IPM exhibits a good convergence behavior with respect to the problem complexity and the load step size. Indeed, convergence robustness and relative stability of the number of iterations is one of the key

interesting features which are classically observed with IPM in the convex setting. Our results seem to extend this observation in the present non-convex setting. As a result, it was possible to compute elastoplastic solutions with only a few large load steps when Newton-type methods failed to converge in such situations. The total number of iterations of the IPM became competitive with respect to Newton methods when increasing the load step size, without impairing too much the solution quality.

Nevertheless further work is needed to improve the efficiency of the solution procedure. First, it is known that IPM cannot fully exploit the knowledge of points close to the solution. More efficient warm-start strategies than the simple strategy taken here could therefore improve the method convergence, especially when considering small load steps. Second, our results indicate that Newton-type method become more interesting with small load-stepping discretization but exhibit less robustness than IPM for larger load steps. A potentially efficient strategy would then to use an IPM algorithm in the first iterations and switch to a Newton-type method for the final iterations, thereby benefiting from the Newton method quadratic convergence near the solution. Some implementation details may also be worth investigating for improving the solver efficiency. For instance, some IPM implementations use merit functions to measure the quality of the next iterate. Finally, in Mehrotra's predictor-corrector scheme the complementarity gap linearization accuracy is improved by taking advantage of the affine step solution. Investigating if a similar approach can be used for improving the other non-linear terms would also be interesting to explore.

* *
*

Chapter 6

Complex engineering examples

Abstract: *Throughout this thesis, various engineering missions at Strains were directly solved using the proposed methods. The software developed at Strains was adapted to include all the precedent numerical developments, enabling us to solve complex engineering applications. In this chapter, a selection of concrete projects and the main results are briefly presented in order to prove the industrial reach of this thesis.*

Contents

6.1 Various 3D assemblies studies

6.1.1 Ultimate resistance of a base column under tensile loading

Following on the T-stub resistance example of Sec. 4.6.3.4, we apply our to more complex assemblies, such as a classical base column under tensile loading as shown in Fig. 6.1. The HEB column is welded to a base plate which is bolted to its final support. In order to determine the resistance of this assembly, more than 20 sub-components should be analyzed. For each of them, all failure modes of the corresponding equivalent T-stub should be checked as can be seen in Fig. 6.2. A complete Eurocode check identifies the yield line mechanism of Fig. 6.3 as the most critical one. The manual computation of this simplified mechanism gives a yield strength of $F_{T,rd} = 312\,\text{kN}$

The yield design approach can provide valuable insight on the failure mode and the shape of the yield lines as seen in Fig. 6.4, therefore reducing considerably the computational effort required to verify this assembly. The proposed approaches give $F_{T,upper} = 354\,\text{kN}$ using the upper bound approach and $F_{T,lower} = 335\,\text{kN}$ using the lower bound approach. The small difference is again due to the fact that the Eurocode does not take into consideration 3D effects, nor the complete curvature seen in the yield line, nor the contribution of the bolt heads to the assembly strength. The resistance value given by the Eurocode is therefore slightly more conservative.

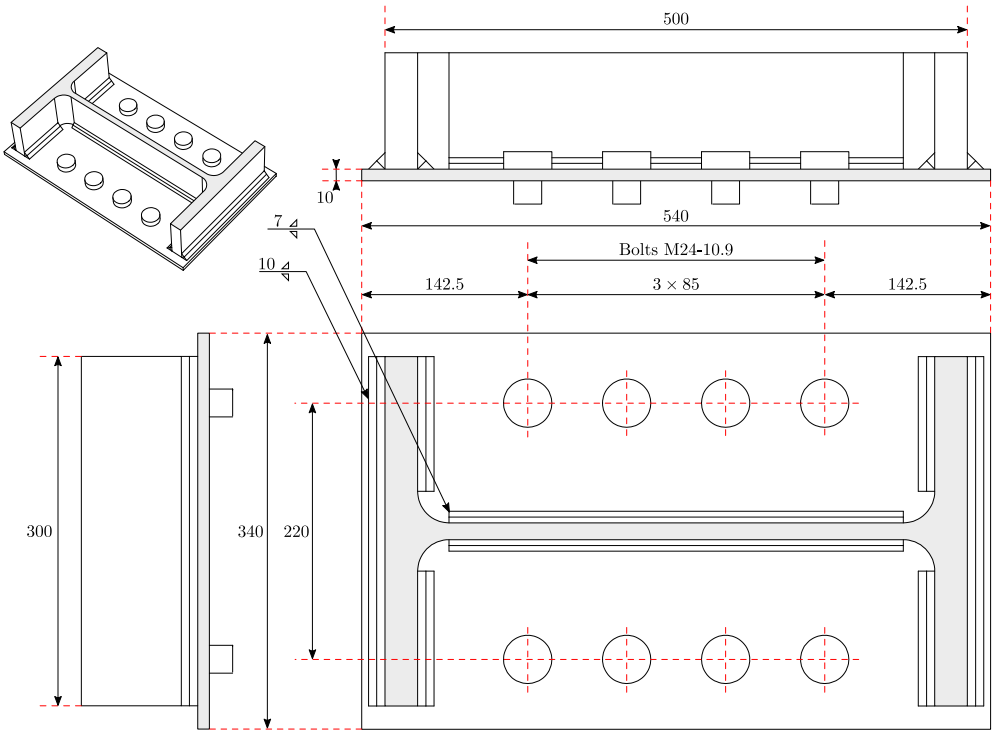


Figure 6.1 – Description of the used model

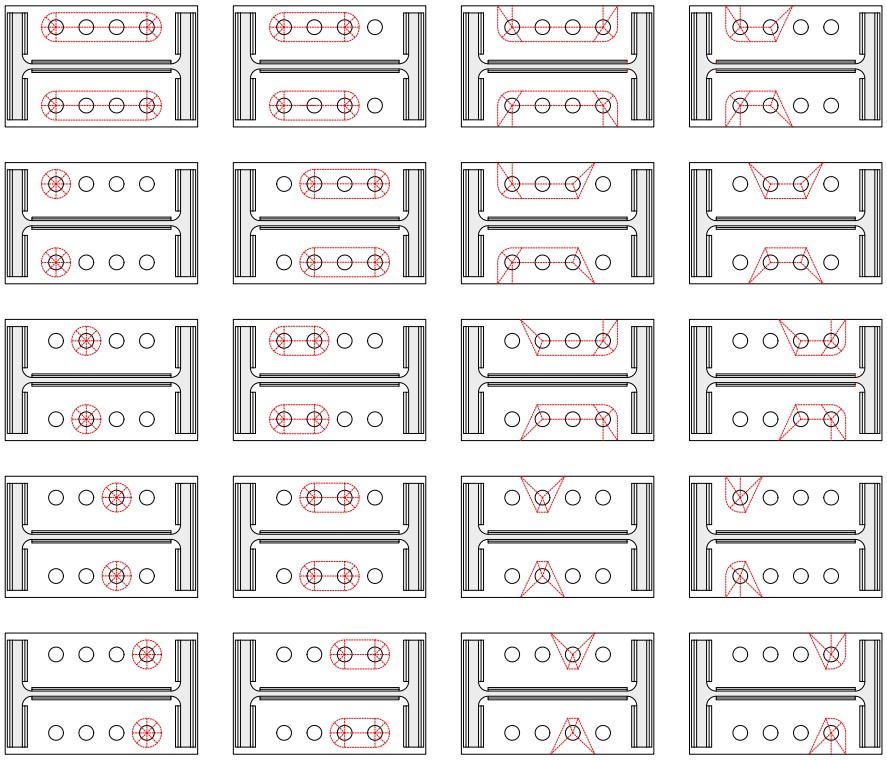


Figure 6.2 – Illustration of all failure mechanisms to check (the dashed red lines represent yield lines)

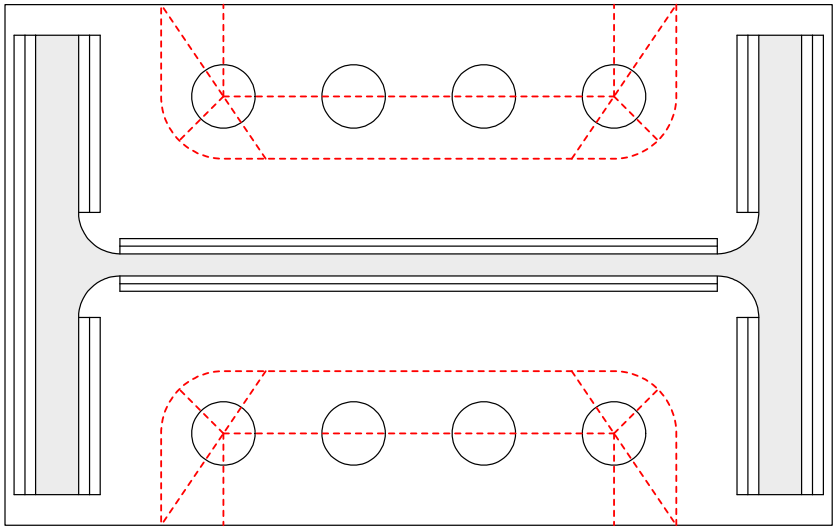
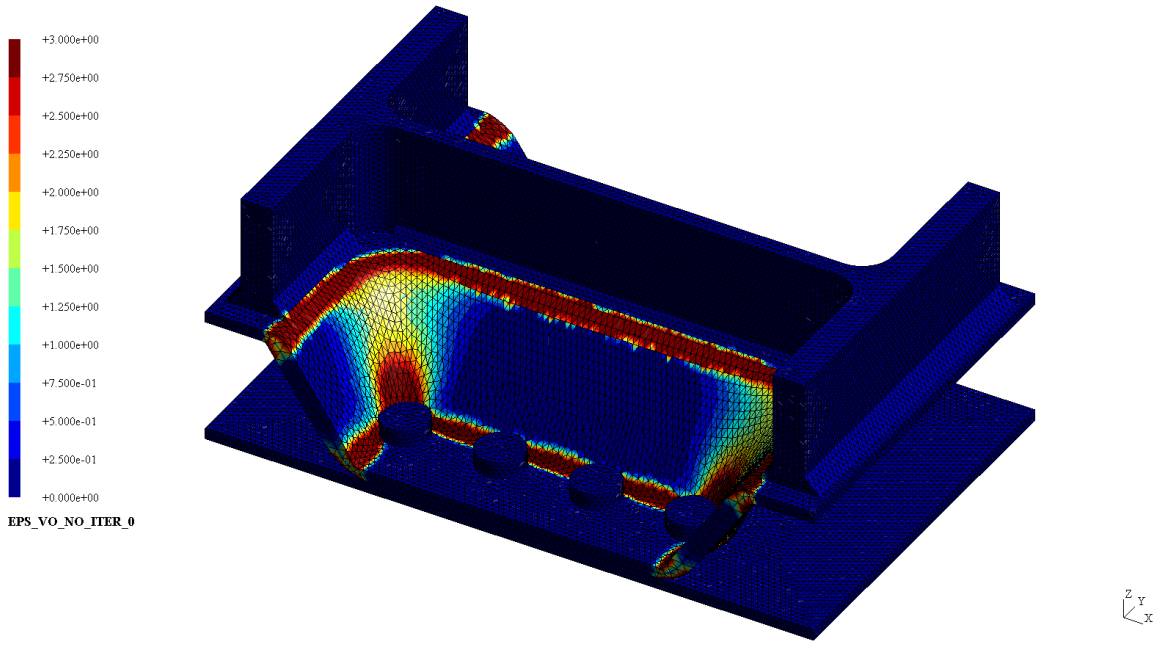
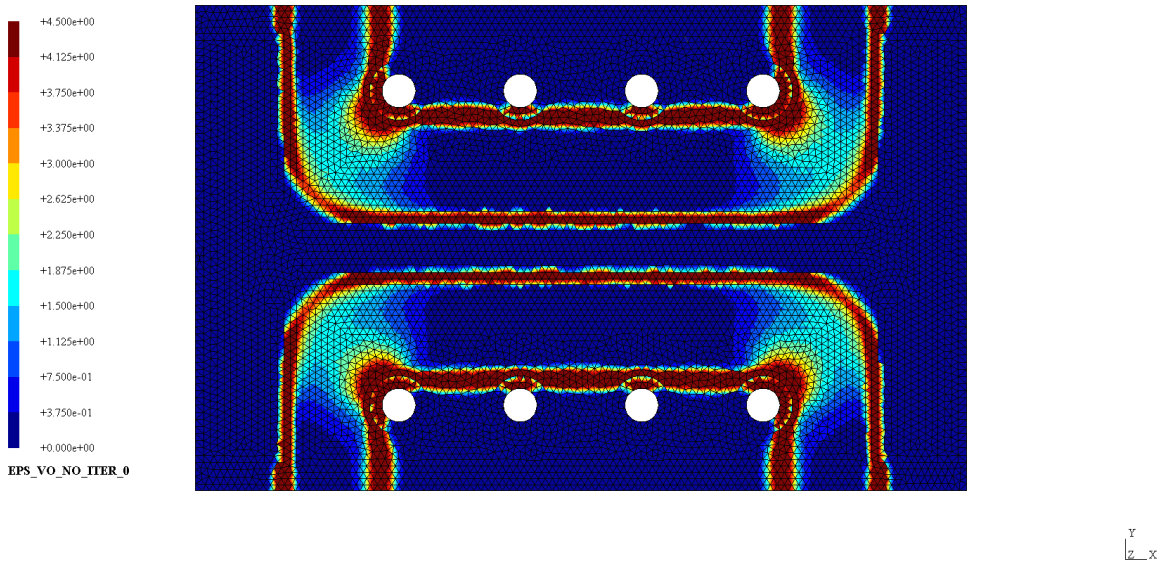


Figure 6.3 – Expected yield mechanism (the dashed red lines represent yield lines)



(a) Failure mode obtained by the upper bound yield analysis



(b) Plastic dissipation concentration matches the expected shape of the yield lines in Fig. 6.3

Figure 6.4 – Solution of the upper bound yield analysis approach

6.1.2 Study of a moment-transmitting assembly

This example aims at illustrating the efficiency of the solving and remesh procedures when applied to a real steel assembly consisting of a particular joint designed for transmitting axial forces and bending moments. This specific joint is rather hard to verify according to the Eurocode design rules and must therefore be assessed numerically. Fig. 6.5 gives a general description of the steel assembly and the bolts disposition. It consists of two HEB500 beams of grade S275 ($f_y = 275\,\text{MPa}$) attached to one another using welded 40 mm end-plates of grade S355 ($f_y = 355\,\text{MPa}$) and 20 M27 bolts of grade 8.8 ($f_y = 800\,\text{MPa}$) with no initial stress. The depth of the weld throat is 15 mm thus verifying $2 \times 15 = 30\,\text{mm} \geq t_w = 28\,\text{mm}$ and its yield limit is taken as $f_y = 400\,\text{MPa}$. Steel is modeled as an elastic perfectly plastic material with Young modulus $E = 210\,\text{GPa}$, Poisson ratio $\nu = 0.3$ and yield stress $\sigma_0 = f_y$. Frictionless contact conditions are imposed over the end-plates forbidding penetration. Moreover, the beams are supposed disconnected from the plates so that forces will be transmitted through the welds, thus simulating a small construction gap between the beam ends and the plates.

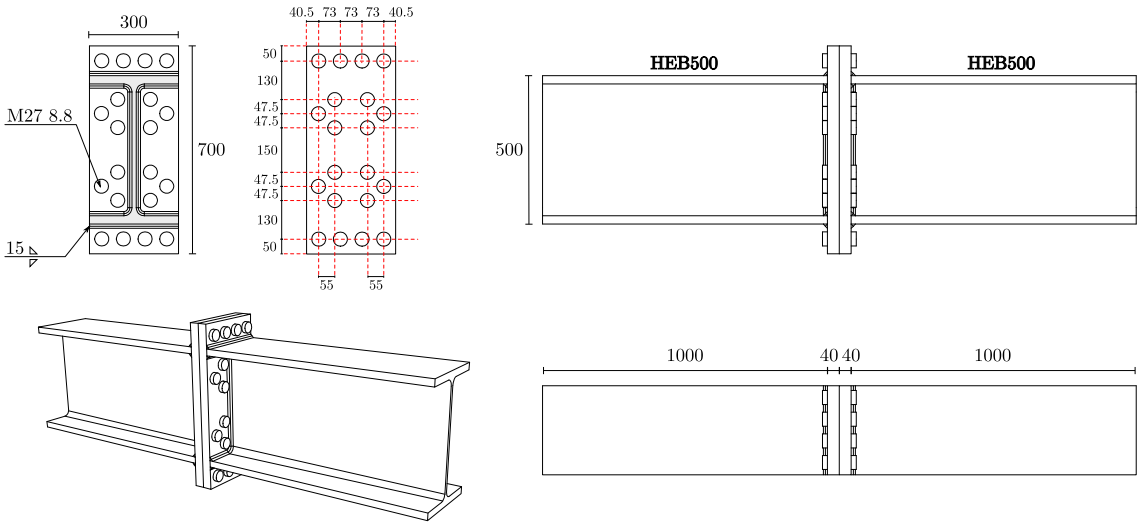
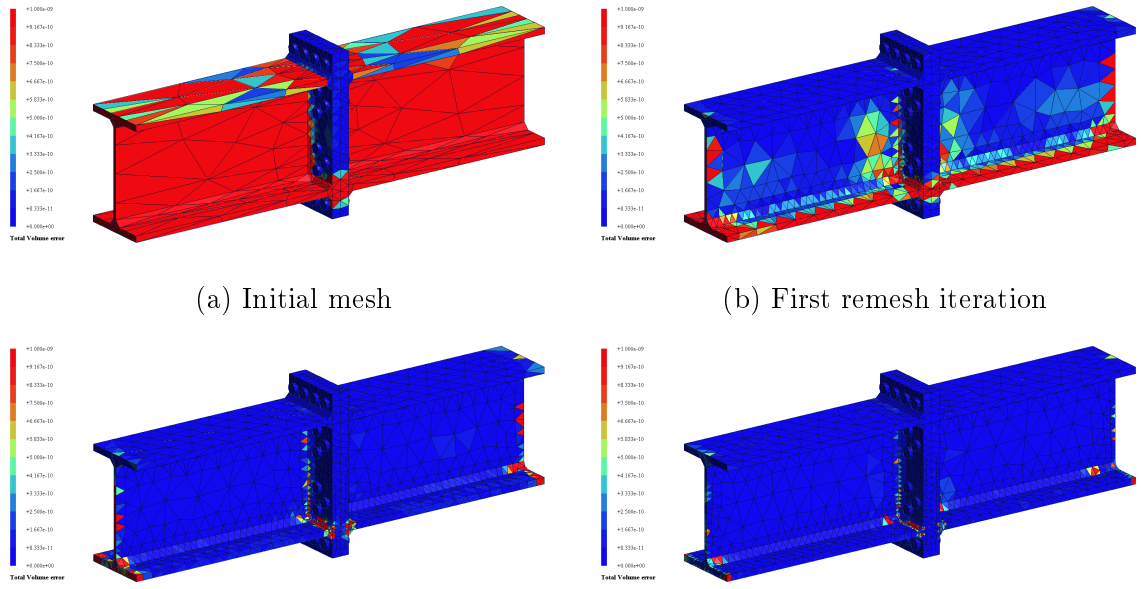


Figure 6.5 – Description of the studied steel assembly

The displacements of all the nodes belonging to the foremost left IPE surface are blocked, and a prescribed displacement composed of a compressive axial displacement (2 mm), a downward vertical displacement (8 mm) and a rotation is imposed (8 mrad), thus simulating an assembly transmitting simultaneously an axial force and a bending moment typically located at mid-span in a structure where shear forces are negligible. Reaction forces are then calculated from the finite element solution.

6.1.2.1 Convergence analysis

We first perform a convergence analysis of the quantities of interest by solving both kinematic and static approaches in an elastoplastic setting with a single load increment step. A 7.5% objective gap between $-J_{\text{stat},h}$ and $J_{\text{kin},h}$ is requested at the beginning of the study and the whole calculation–error maps–remesh scheme is repeated until the desired value is obtained. A total of 4 meshes (one initial mesh and 3 remeshes, see Fig. 6.6) were necessary for the gap between the static approach and the kinematic approach with no discontinuities to reach 7.1%, as shown in Fig. 6.7. The gap between the static and the kinematic approach with discontinuous elements reaches 3% on the second remesh and 1.8% on the third remesh. As regards the differences on reaction forces, they follow the same tendency: 9.5% between the static approach and the kinematic approach with no discontinuities and 2.0% between the static and the discontinuous kinematic approach. Fig. 6.8 and 6.9 show the evolution of the reaction forces with the 4 different meshes.



(a) Initial mesh

(b) First remesh iteration

(c) Second remesh iteration

(d) Third remesh iteration

Figure 6.6 – Initial and adapted meshes for the first example (isovalues represent the local computed error)

6.1.2.2 Determining the interaction diagram

The previous assembly is now compared to a complete HEB500 beam and a hollow beam designed for optimizing material use. Fig. 6.10 describes the hollow beam geometry, typical of those found in practice. The chosen diameter-to-height ratio is taken here equal to 1.45 with a steel grade S275 ($f_y = 275\,\text{MPa}$). As regards to the assemblies, we consider in fact two different cases:

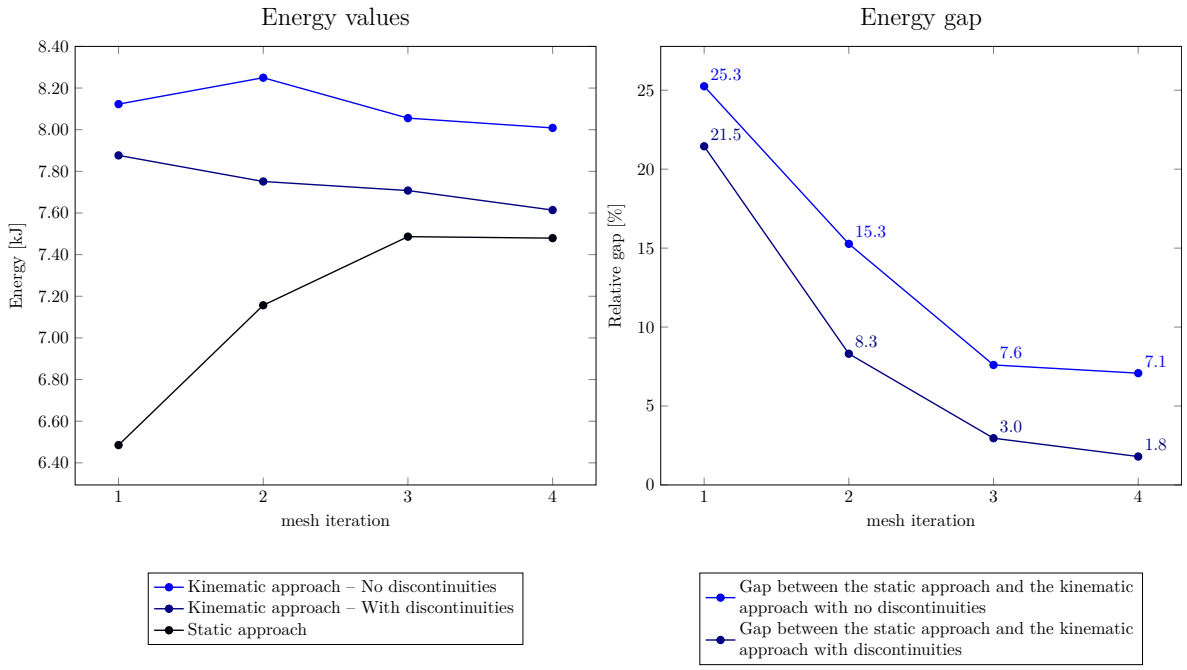


Figure 6.7 – Convergence analysis with respect to objective functions: $J_{\text{kin},h}$ for the kinematic approach and $-J_{\text{stat},h}$ for the static approach

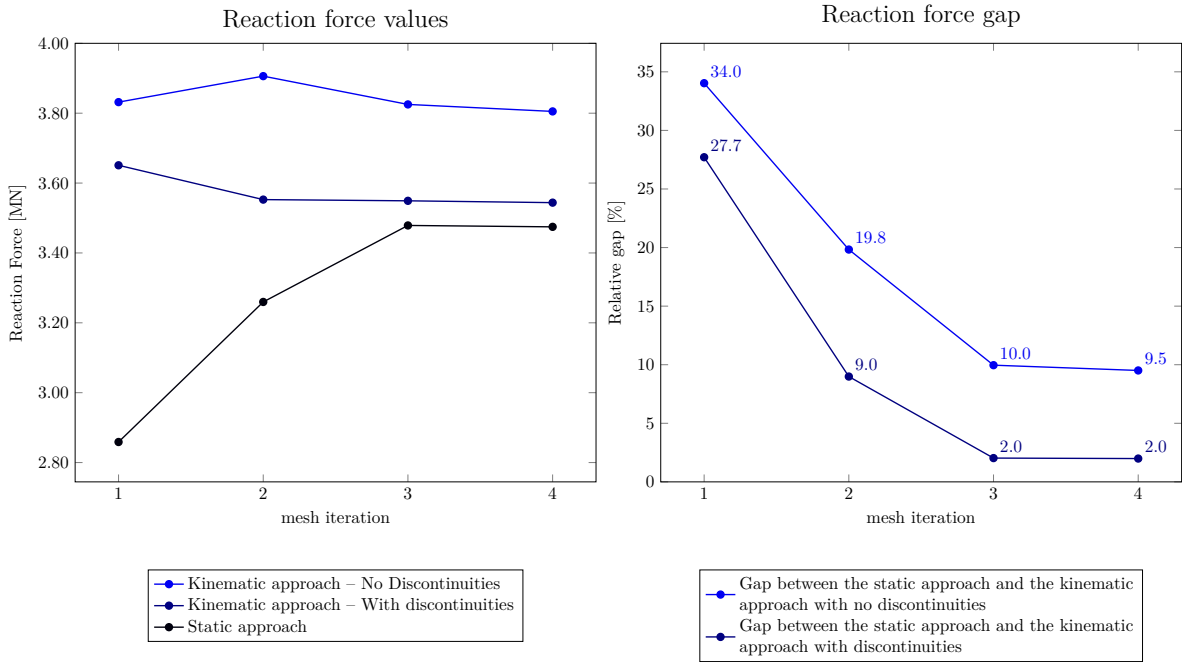


Figure 6.8 – Convergence analysis with respect to the reaction force

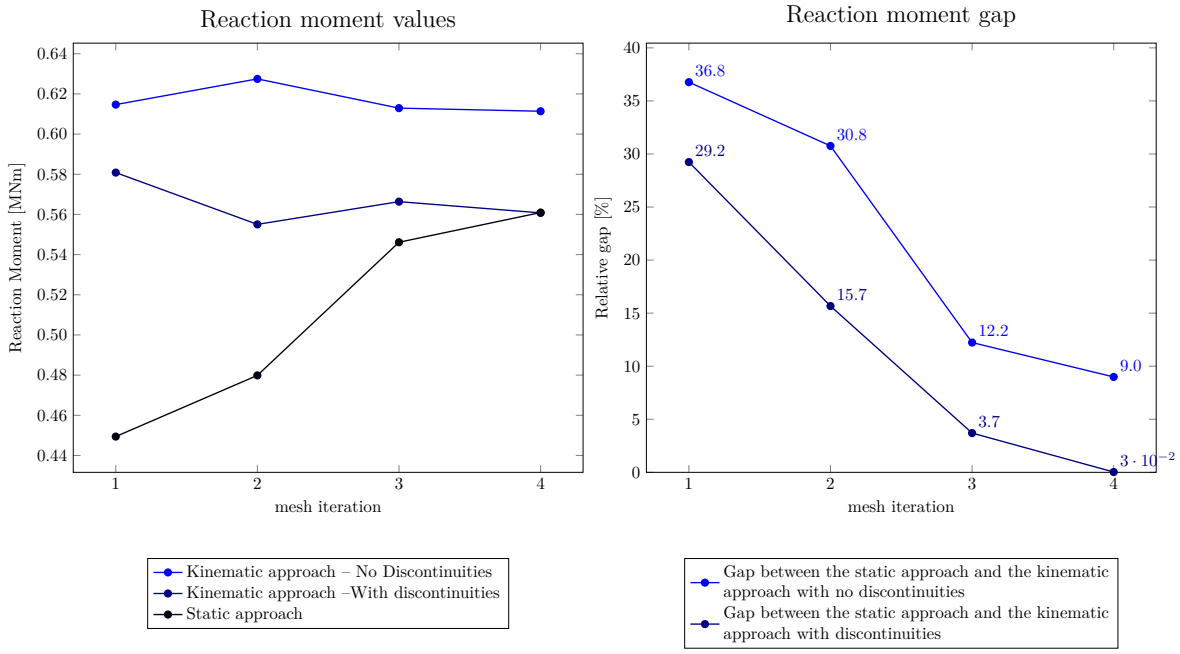


Figure 6.9 – Convergence analysis with respect to the reaction moment

– Assembly 1: a weaker, badly-designed assembly, where the depth of the weld throat is 10 mm thus being less than the thickness of the flanges;

– Assembly 2: the well-designed assembly introduced before (see Fig. 6.5).

In this example, we consider upper and lower bound yield design computations as described in Sec. 4.3. The reference loading consists of a prescribed bending moment M around the strong axis of the beam and a normal force N. For a fixed value of (N, M), we maximize the load multiplier factor α such that $(\alpha N, \alpha M)$ corresponds to the maximum normal force and bending moment for the chosen reference values. By varying the reference load direction (N, M) in the normal force-bending moment space, we compute different values of the load multiplier α, corresponding to different points on the failure domain, describing, in the end, the beam section interaction diagram. We restrict here the diagram computation to only one quarter, corresponding to tensile force $N \geq 0$ and positive moment $M \geq 0$.

Fig. 6.11 shows the interaction diagram obtained by following the previous procedure using either upper or lower bound approaches. For each configuration, the corresponding exact interaction diagram boundary lies inside the shaded area delimited by the lower and upper bound calculations. As expected, the hollow beam presents a smaller strength than the entire beam. As regards to the assemblies, the first topology (assembly 1) with weaker welds exhibits a significantly reduced strength whereas the second one exhibits a strength comparable to that of the hollow beam. In practice, the second assembly presents a better design since the engineer would be able to completely utilize the beam and the assembly strengths. With the first assembly design, the beam would not be

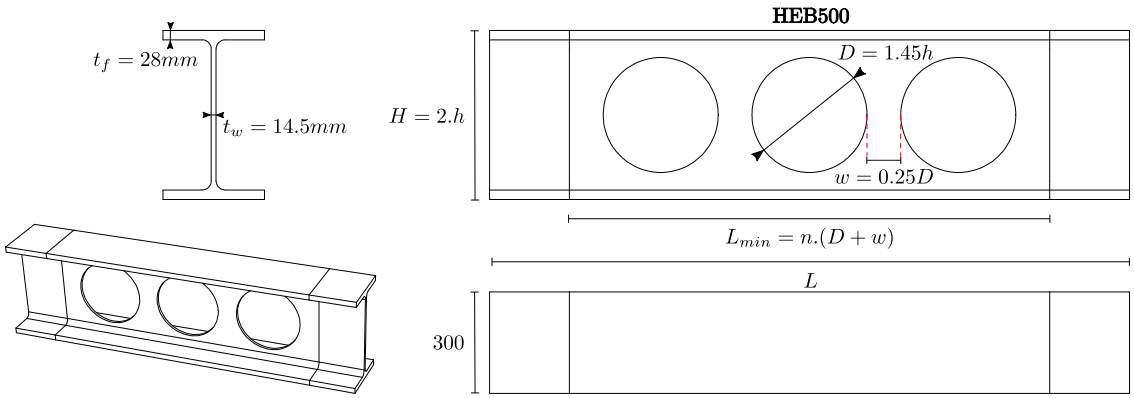


Figure 6.10 – Description of the hollow beam

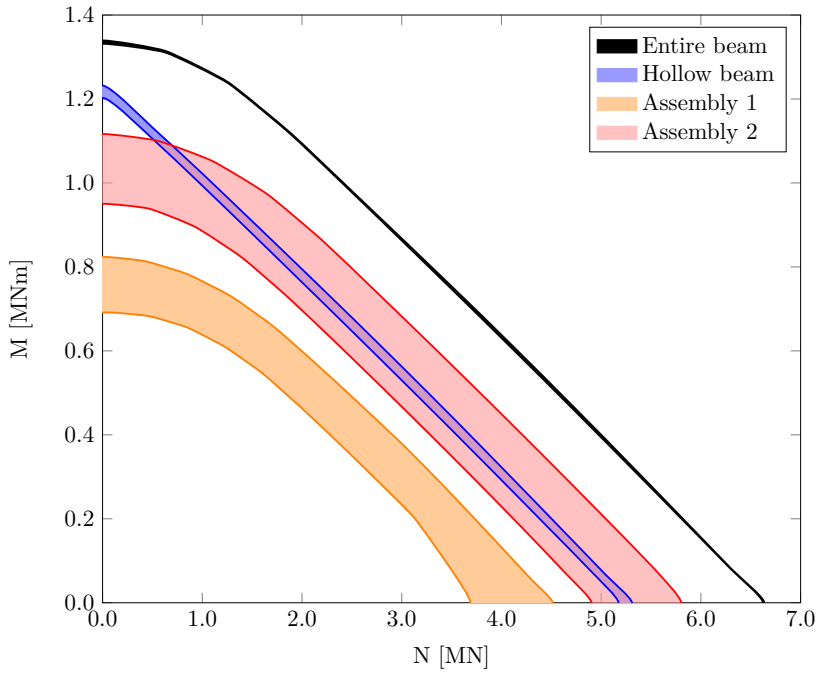


Figure 6.11 – Axial force-bending moment interaction diagram

fully utilized since failure would be dictated by the weaker assembly. The proposed method provides the engineer the ability to estimate the yielding domain of the studied structure and therefore can be used to give insight on bad conceptions and the possible failure modes which can then be improved. This will ensure a safer and more economical design.

6.1.3 A safety analysis of more complex assemblies

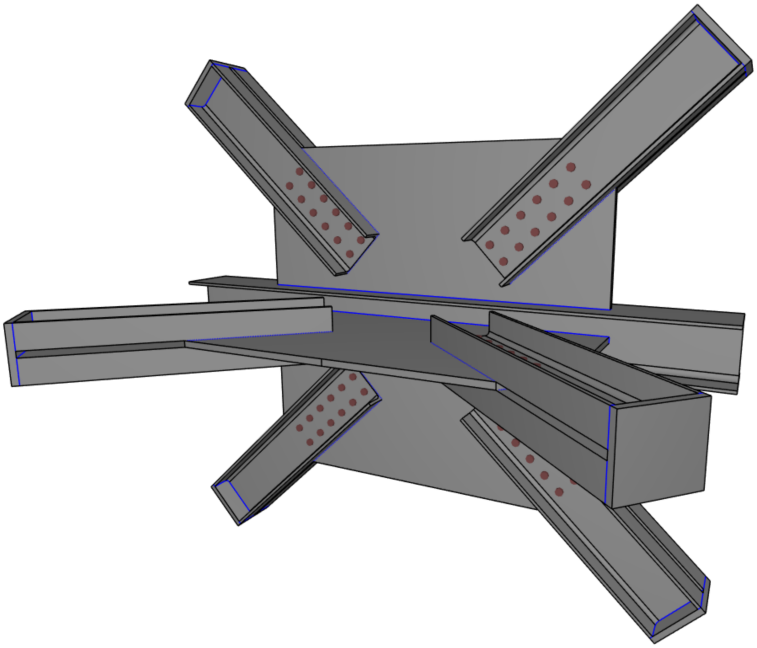
In this paragraph, we briefly describe two other complex assemblies that were analyzed using the software *DS-Steel* developed by *Strains*. These assemblies are part of a large set of checks that were made for a structural engineering firm. The geometry details and load cases are omitted in order to simplify the presentation and preserve the confidentiality of the study. Two complex assemblies are briefly presented:

- A 3D bracing assembly as shown in Fig. 6.12a in which the applied loads are mainly tension or compression forces in the converging bracing members obtained from a global 3D model of the whole structure;

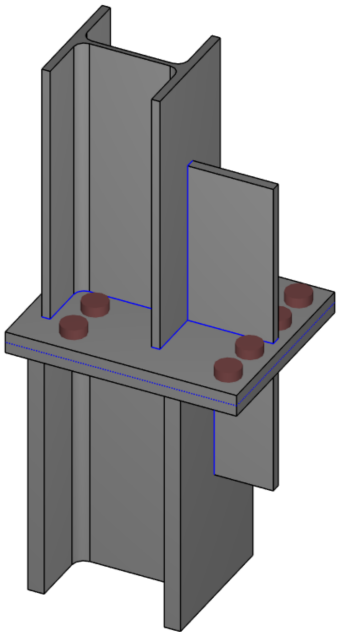
- A moment transmitting assembly as shown in Fig. 6.12b in which the applied load is a combination of a uniaxial bending moment, along with a normal and a shear force.

In these two studies, only unilateral contact constraints were imposed.

Using the upper bound approach, one can determine the most critical failure mechanism. In the case of the first assembly, failure occurs in one of the bracing members where the web plate fails in bearing due to a critical bearing pressure as shown in Fig. 6.13a. One can also see in Fig. 6.13b that the elastic limit, in this case $f_y = 355\,\text{MPa}$ is reached in the same diagonal member. The upper bound provides a load factor of 2.2 for the corresponding load case while the lower bound approach yields a load factor of 1.9. The gap between the two results can be reduced by reducing the mesh size in the failure area. For the second assembly, failure occurs within the end-plate where clear yield lines can be seen when mapping the plastic strains isovalues as shown in Fig. 6.14. In fact, in order to obtain those clear yield lines, the remesh scheme was used which allowed us to obtain, from a first coarse mesh as seen in Fig. 6.14a, a much better result presented in Fig. 6.14b. The mesh used for each of the two examples consisted of 200 000 quadratic tetrahedrons for the upper bound problem and 800 000 linear tetrahedrons for the lower bound problem, yielding roughly 5.5 million degrees of freedom for each study. The interior point method shows a very efficient behavior with a number of iteration remaining stable compared to smaller problems (22 to 24 iterations) and a CPU time of almost 315 s per iteration with an OpenMP parallelization over 8 cores.

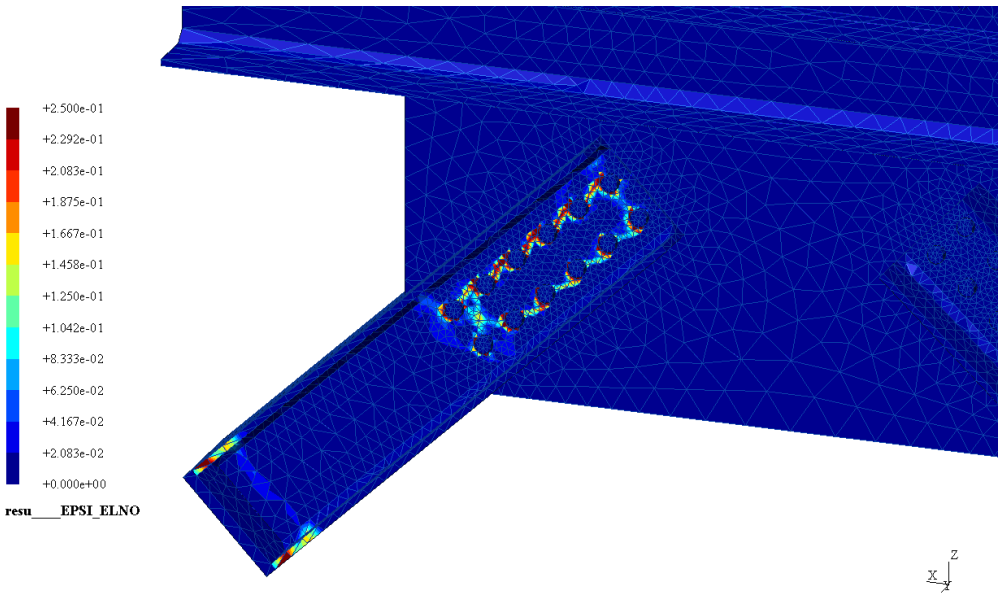


(a) A wind bracing assembly

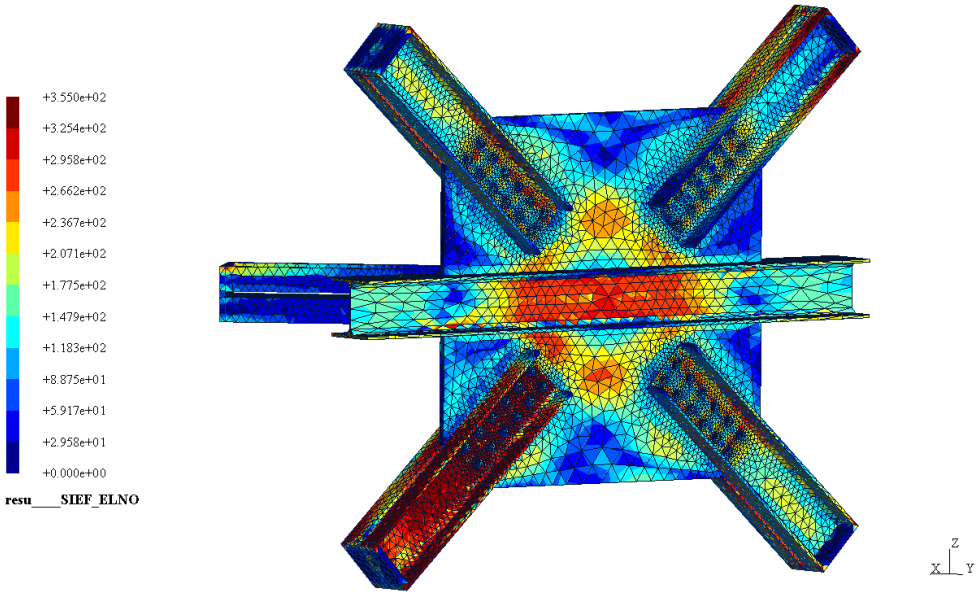


(b) A continuity assembly

Figure 6.12 – Complex 3D assemblies

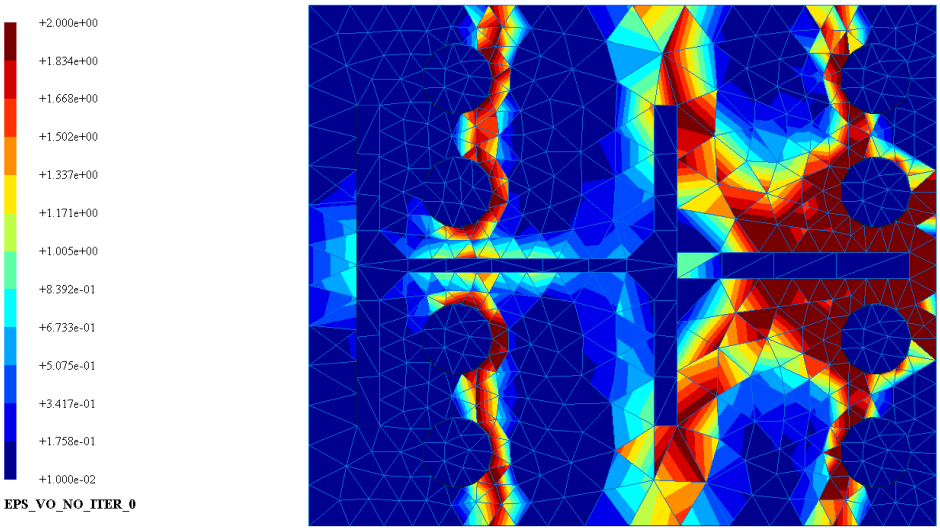


(a) Equivalent von Mises plastic strain isovalues in the critical diagonal

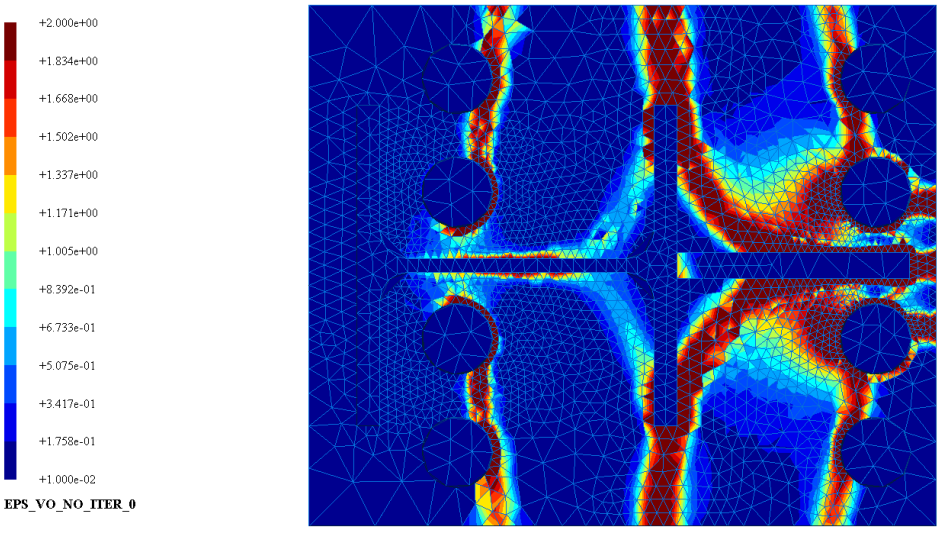


(b) Equivalent von Mises plastic stress isovalues

Figure 6.13 – Equivalent von Mises plastic strain and stress isovalues for the wind bracing



(a) Equivalent von Mises plastic strain isovalues – first mesh iteration



(b) Equivalent von Mises plastic strain isovalues – second mesh iteration

Figure 6.14 – Kinematic results for the continuity assembly

6.2 Second order non-linear analysis of a steel bridge section

This last example is dedicated to a real-case civil engineering application on a mixed steel–concrete bridge. Assessing the ultimate state of a 3D steel structure including second-order geometrical effets and material non-linearities is a current challenge in the engineering practice. Following Eurocode design rules, a second-order non-linear structural analysis is conducted based on the introduction of an initial geometrical imperfection whereas steel is modeled using a finite-strain von Mises linear isotropic hardening behavior following the developments of chapter 5.

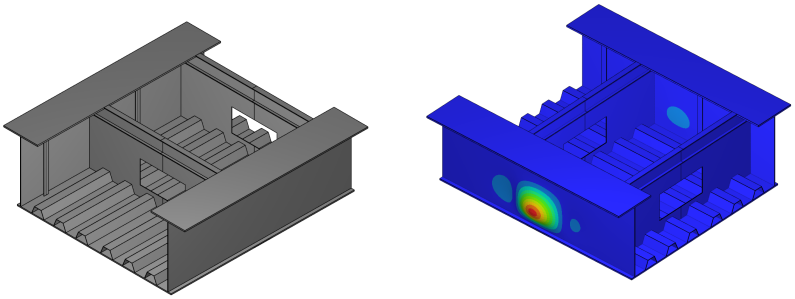
In the following, we will analyze on of the construction phases of the steel–concrete bridge by considering a 3D local model of the steel girder without the top concrete slab. In the construction planning, the pre-assembed steel girder is launched from one abutment of the bridge until it reaches the second. During the incremental launching phase, each section of the bridge will experience different loading combinations. This is especially critical when a supposed middle-span section of the final configuration passes over a pile, experiencing bending moment shifts and tensile forces. The configuration id considered as critical since this section is not necessarily optimized to withstand negative bending moment coupled with a shear force due to the pile reaction. In such a configuration, the thin web plate may be unstable and fail by buckling.

To assess this risk of failure, a specific load case for the critical mid-span section is obtained from a global bridge model from which all launching configurations are calculated. The load case is then applied to the local 3D model containing an initial geometrical imperfection. This load case is then incremented until a certain *failure criterion* (taking into account all normative and technical aspects) such as a excessive out-of-the-plane lateral displacement of the web or an excessive plastic deformation are reached. The ultimate load multiplier is finally compared with load safety factors proposed by the Eurocode.

The 3D model is represented in Fig. 6.15a where one can clearly see the girder cross-section as well as equally-spaced diaphragms and web reinforcements. Initial imperfections have been obtained from the first elastic buckling mode and have been represented in Fig. 6.15b. The load case seen in Fig. 6.15c consists of a bending moment and shear force along with the pile reaction force over the bearing device surfaces.

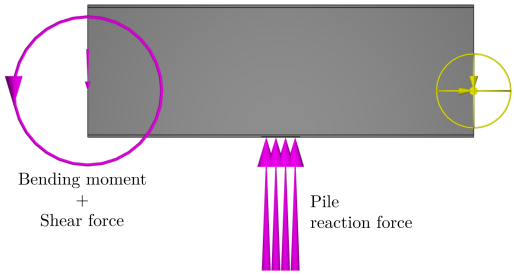
The numerical simulation enabled to verify that the steel section as well as its transverse stiffeners can withstand buckling and will most likely fail due to excessive plastic yielding of the top flanges as seen in Fig. 6.16a. The out-of-the-plane displacements of the thin web remained within a 1 cm limit in the middle of the pre-deformed web as seen in Fig. 6.16b.

Finally, even for this complex engineering example consisting of a mesh of roughly 150 000 tetrahedra, the IPM solver exhibited good convergence properties, the required number of iterations ranged from 6 to 21 for the first six load steps, the last converged load step required 42 iterations.



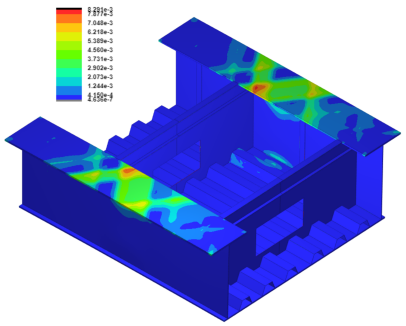
(a) Local 3D model

(b) Initial geometrical imperfection isovalues

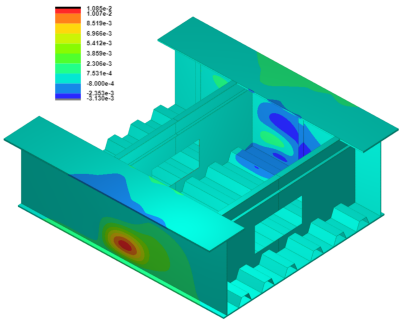


(c) Illustrative representation of the load case

Figure 6.15 – Geometry of the steel bridge, initial geometrical imperfection and the considered load case



(a) Equivalent plastic strain isovalues at failure state



(b) Out-of-plane displacements isovalues at failure state

Figure 6.16 – Some results at failure states

* *
*

Chapter 7

Conclusions and outlook

Conclusions

In order to answer the industry's need for powerful and robust numerical methods for modeling the non-linear behavior of steel structures, we explored and adapted in this work the use of one the most efficient optimization tool, the primal-dual interior point method, to solve large-scale non-linear finite elements problems. Relying on a proven efficiency in the mathematical programming community, we solved elastoplastic problems with contact conditions using the IPM and we extended its use to include large-strain effects. The results are very promising and show great advantages compared to classical approaches such as the Newton-Raphson/Return mapping algorithms.

Chapter 1 has set the general context on this thesis. Due to increasing economical, environmental and technical challenges, engineers are nowadays required to greatly optimize their designs while still ensuring the safety of the users. In order to achieve their goals, they increasingly rely on more sophisticated and complex structural analysis software which are hard to use and not necessarily robust. We insisted on the fact that modeling a complex 3D steel assembly using non-linear finite elements is an extremely time-consuming task and is often impossible to realize in the prescribed projects deadlines. The need for new and powerful numerical approaches which can provide quality results in a reasonable amount of time of effort is as great as ever. The proposed solution relies on powerful optimization software which require the formulation of classical mechanical problems in the framework of non-linear mathematical programs.

In chapter 2, we introduced the mathematical concepts on which we rely along with a complete presentation of the primal-dual interior point method. One of the main advantages of this method is its capacity to handle large-scale problems with various constraints and specifically second-order cone constraints which allows us to solve a wide variety of non-smooth problems via a simple reformulation using additional slack variables. Contrary to most implementation available in sate-of-the-art solvers, we extended the primal-dual IPM scope to the case involving smooth non-linear objectives (instead of usual linear objectives) and also possibly non-convex, convergence being not necessarily guaranteed in the latter case. These extensions require only minor modifications of traditional implementation, the specific convex conic constraints structure being unchanged.

Chapter 3 can be seen as a proof of concept in which the IPM advantages over a classical solving methods and the choice of an in-house algorithm are illustrated. In this chapter, we explored the use of SOCP and IPM for modeling three-dimensional elastic structures in contact. For that, the mechanical problem is cast in the form of two dual energy principles which are then solved using adequate kinematic and static finite-elements discretization. The advantages of a dual approach is not only limited to its bounding properties but it also yields great insight to build a dedicated error estimator. This estimator can then be used in an adaptive remeshing scheme. The IPM is shown to be very robust and efficient and is capable to take into account a great number of conic constraints corresponding to the contact conditions. This shows all its advantages over traditional penalty or augmented Lagrangian approaches in terms of results quality and computational efficiency. The approach could perfectly be extended to non-matching meshes with a node-to-surface method. Curved surfaces such as the bolt's cores or plates' holes can also be finely modeled using curved or isogeometric elements. These changes are independent of the contact resolution method which remains the same.

In chapter 4 we presented a complete calculation scheme for elastoplastic bodies with contact, which we applied to the study of 3D steel assemblies. This chapter directly represents the main goal of this thesis where we devised a robust and efficient solution procedure. The dual principles introduced in the previous chapter are extended to include elastoplasticity and the link with the classical theorems of yield design are established. Specific finite-elements discretization are used to bypass some problems such as volumetric locking which is solved using discontinuous displacements fields or hybrid formulations. Improved equilibrium elements removing linear dependencies and spurious modes have also been used for the static formulation. The behavior of the IPM is shown to be extremely robust, the number of iterations for convergence remaining reasonable, ranging between 15 and 30 iterations, even for very large problems. However, the most interesting feature is the capacity of the IPM to handle large load steps. In fact, for each monotonous loading path, one large step is sufficient in the case of perfect plasticity to compute the structure ultimate state. This greatly reduces computational times and the need of an experienced engineer to pilot the algorithm convergence.

Chapter 5 was a development made for *Strains* in order to answer a direct need for one specific engineering study. We explored the possibility to include large-strain effects using the same framework, which, to the author's knowledge, has never been implemented using mathematical programming tools. The formulation relies on logarithmic strains which allow us to model finite-strain metal plasticity using the same concepts as in the small strain hypothesis. The IPM method which we extended to SONLP is then used to solve the obtained minimization problem and showed several of the benefits found in the convex, small-strain, setting. Its robustness to large load steps remains the key feature as it becomes competitive regarding the classical methods

without impairing too much on the solution quality.

The industrial reach of this Ph.D. is illustrated in chapter 6 where various complex engineering examples are presented. It was in fact the needs defined through these studies at *Strains* and the team's will to provide a quality service that pushed us to explore and implement all of the proposed framework introduced here and many more.

The *optimization-based* approach shows all of its benefits: using the same mathematical formalism and one algorithm, a solution is obtained in a reasonable amount of time and effort without relying to the back and forth methods such as the Newton-Raphson and radial return schemes for plasticity, of the penalty or augmented Lagrangian approaches for modeling contact conditions. The IPM appears to be very scalable, it allows us to handle large-scale problems with no to little effects on convergence time and the number of iterations.

The proposed method has been proven to be in accordance with the Eurocode's design requirements through the various examples presented in this thesis. The results shows that the failure mechanisms covered by the Eurocode can be correctly predicted with a impressive accuracy. However, one should keep in mind that for security and legal reasons, the Eurocode checks are still required until a full normative setting is provided by the regulatory authorities. Surely, the non-linear finite-element method is the go-to whenever an engineering project cannot be modeled by hand, however, there is still some debate around the validity of the models and the hypothesis taken, therefore finite-element models are not yet ready to overcome all normative hand-checks. In contrast, the optimization framework explored in this thesis could be of great importance in engineering practices. Not only does it allows the engineer to efficiently obtain quality results, but it also gives him additional mechanical insight such as the upper and lower bounds or the failure mechanism which allows him to correctly assess the risks and to better choose the critical Eurocode check. The proposed method paves the way to more a generalized and automatic use of non-linear finite-element models in a complete project design process.

Perspectives

Improvement of the IPM solver

Although the efficiency of IPM has been be illustrated in this manuscript, various subjects remain to be explored. Improved path-following strategies, arc-length strategies, adapted merit functions for assessing the quality of the iterates, filter methods, active set methods, etc. are still active research subjects and receive a lot of attention in the mathematical programming community. Another main numerical bottleneck that needs to be tackled is the linear resolution of the Newton system. The use of direct solvers for really large-scale problems becomes prohibitive and iterative solvers would certainly be more appropriate. Development of good preconditioners of IPM in order to use iterative conjugate-gradient methods is however still a subject of current research

due to the strong ill-conditioning of the system when approaching the optimal solution.

Extension to other mechanical behaviors

The convex optimization setting allows us to take into account various mechanical behaviors using the generalized standard material framework and to write the dual principles in the same form. In fact, the possibility of extending yield design to 3D concrete structures and masonry structures, using a Mohr-Coulomb or Rankine strength criterion, has already been explored at *Strains* in previous works [Vincent, 2018; Vincent et al., 2018]. The implementation took into account the existence of steel reinforcement using either a homogenized material or a real 3D modeling of rebars. These strength criteria require an SDP solver which turns out less efficient than the one developed for SOCP. The in-house solver available at *Strains* is still under constant development in order to reach an efficiency comparable to commercial SDP solvers such as `Mosek`. The use of these type of mechanical behaviors in elastoplastic analysis with or without large-strain effects is still an active research subject.

For masonry structures, it is clear that in order to better represent the real material behavior, a non-associative Coulomb friction law should be considered. While an elasto-plastic load-path following analysis with a non-associated law is interesting to perform, in yield design, a non-associated solution raises many theoretical questions. Moreover, due to the formulation relying on a convex optimization problem,the mathematical formalism used here is only limited to associated laws. The proposed IPM approach cannot take into account the real non-associative behavior of Coulomb's frictional contact as such. Some strategies have nonetheless been already proposed to circumvent this aspect [Gilbert et al., 2006; Kleinert et al., 2014; Krabbenhoft et al., 2012a; Kučera et al., 2013] and this problem will definitely deserve further consideration.

Local-global models and model reduction

The ultimate objective is to be able to describe a complete steel structure using a combination of local 3D models for steel joints and a global model using beam elements. As already explained in chapter 1, a steel assembly is neither fully rigid nor nominally pinned. Its structural behavior is rather *semi-rigid* and the evaluation of its stiffness and displacement tolerances poses immense challenges when modeling full 3D structures. Using the reliable tool presented in this thesis, one is able to better estimate these parameters.

This is an ongoing project at *Strains* and it has already benefited form a European grant through the project *Fortissimo 2* which was led in collaboration with *Egis Industries* and the *Edinburgh Parallel Computing Center*. An important objective of this collaborative project is to process complex computations with an High Performance Computing (HPC) infrastructure. A proof of concept was realized where 3D models of assemblies are connected to 1D classical Euler-Bernoulli beams via rigid body movements of the interface sections (see Fig. 7.1-7.2). This has led to a new global *online* calculation scheme:

1. A first estimate of the *condensed* stiffness matrix for each joint is calculated using

the IPM;

2. The global stiffness matrix of the whole structure is then assembled and a global Newton-Raphson scheme is launched;
3. At each iteration of the Newton-Raphson algorithm, the stiffness matrix and nodal forces vector for each assembly is updated via a full 3D analysis using the IPM;
4. The process is repeated until a specified tolerance is reached.

This algorithm is said to be *embarrassingly parallel* i.e. it can easily be subdivided into independent calculations using a *master-slave* parallel computing design pattern (see Fig. 7.3). This is where the powerful HPC infrastructure comes into action. The *master node* holds the global Newton scheme and sends out the required information for each set of *slave nodes* responsible for calculating the assemblies. These nodes would simultaneously use the IPM to calculate the strain and stress state and derive the required stiffness and nodal vector update which is then sent to the *master node* where the reduced system is solved and so on.

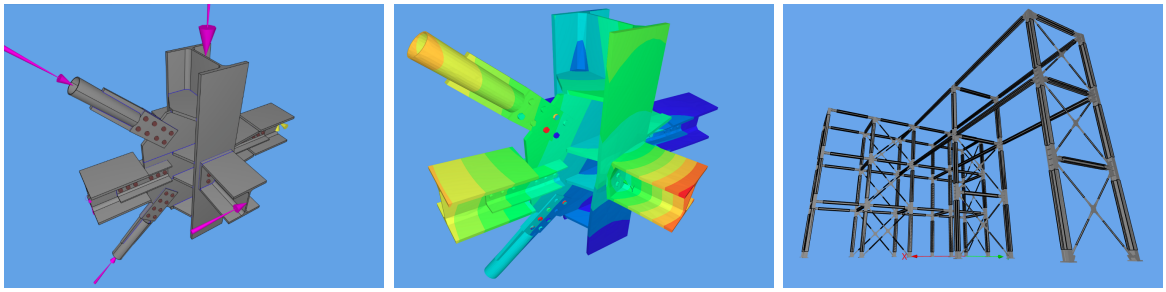


Figure 7.1 – Examples of more complex local and global models

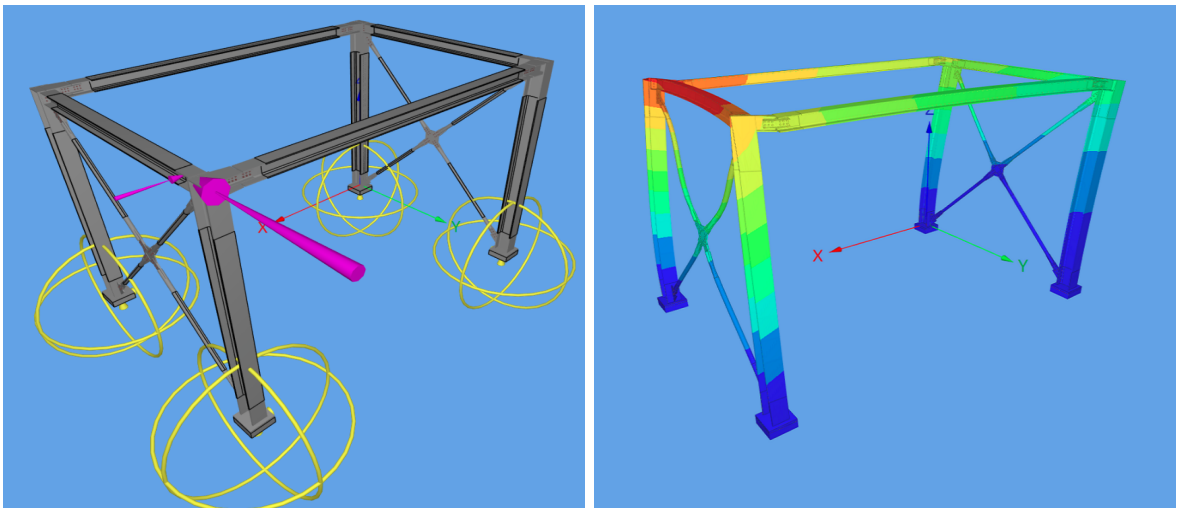


Figure 7.2 – Finite elements analysis of a simple steel structure

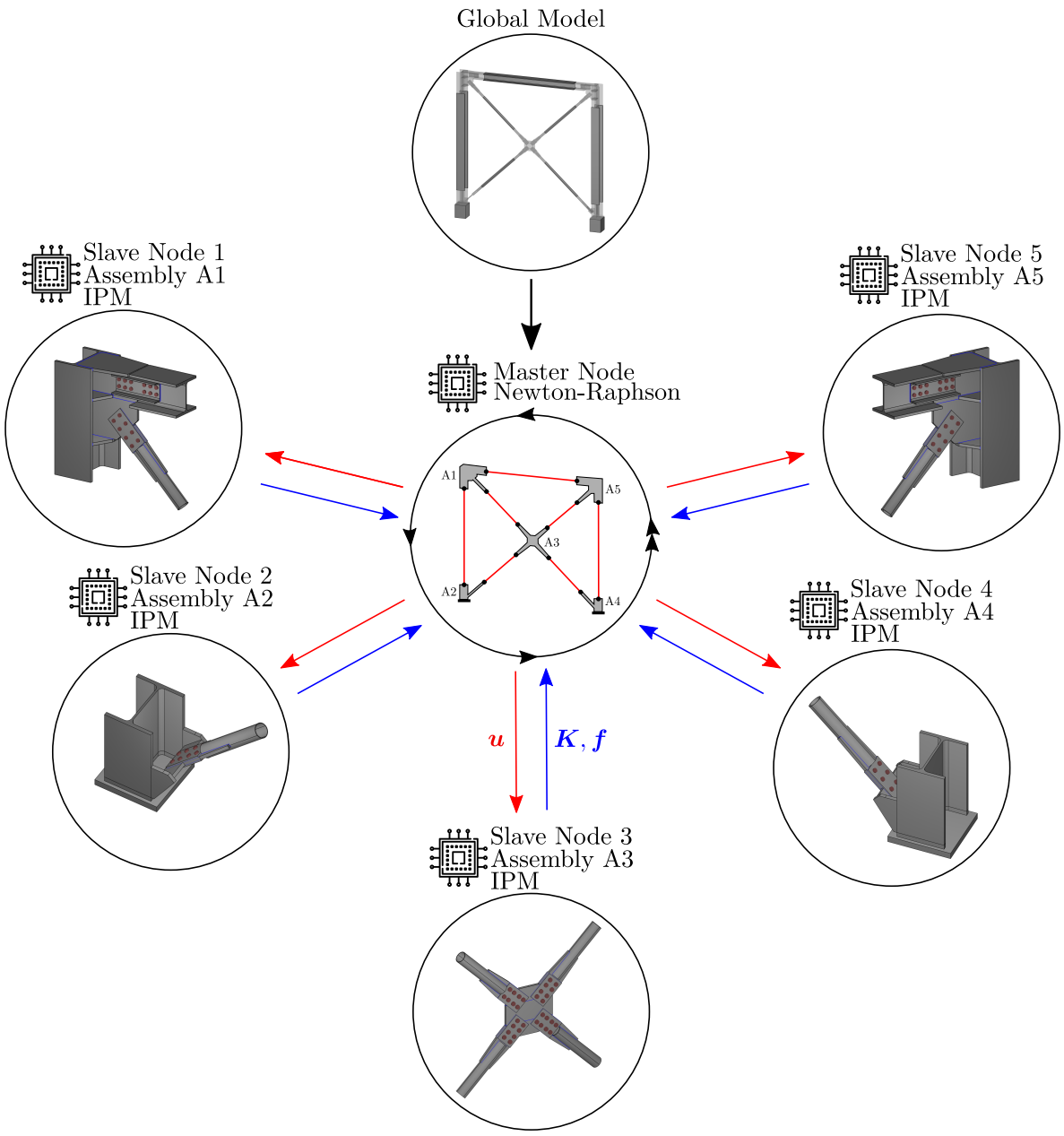


Figure 7.3 – Schematic representation of the master-slave design pattern for a simple steel structure

The key problem remains the 3D-1D connection via a rigid section interface. When assuming a Euler-Bernoulli section kinematics at the 1D/3D junction, transverse 3D information is inevitably lost. In order to better estimate these effects, the aim would be to connect the 3D assembly to enriched beam models already developed at *Strains* through various previous collaborations [Corre et al., 2018, 2020; Ferradi et al., 2013, 2016]. This remains an active research project at *Strains*.

Non-deterministic design

To conclude, let us go back to the basic deterministic Eurocode check given in Fig. 1.1 where for a given geometry, action set and material choice, the overall effects should be less than an estimated resistance. In this manuscript, we focused on the methods used to evaluate the effects and the resistance of steel structures with a fixed geometry and material parameters and under a specific load case.

But in reality, neither the geometry, nor the materials, nor the action set can be accurately determined. The deterministic approach is only a mere idealization of reality where various uncertainties exist, yet it still is the most used among engineers.

One way to surpass this is by relying on Monte-Carlo methods which consist in repeated random sampling to obtain approximate probability estimates. However, such analyses are extremely time-consuming and catching tail-risks (i.e. exceptional events), which are technically the ones we are looking for in a structural analysis, is very hard. Note that if this method is to be adopted, the one large step analysis scheme presented here will show all its promise over more traditional Newton-Raphson approaches since we are mostly interested by the final limit state.

Another possibility is to extend the presented framework, and specifically limit analysis optimization problems, to include uncertainties in the mechanical problem formulation. There exists a large literature on optimization under uncertainties including *stochastic optimization*, *chance-constrained optimization* or *robust optimization*. These methods rely on objective and constraint functions $f_i(\boldsymbol{x}, \zeta)$ which depend on optimization variables $\boldsymbol{x}$ and a some random or uncertain variable ζ.

In its most general form, the goal of *stochastic optimization* amounts to finding $\boldsymbol{x}$ so that constraints are satisfied almost surely and the objective expectation is minimized for a given probability distribution $\mathcal{P}$ of the uncertain variables:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & F_0(\boldsymbol{x}) = \mathbb{E}[f_0(\boldsymbol{x}, \zeta)] \\ \text{s.t.} & f_i(\boldsymbol{x}, \zeta) \leq 0, \qquad\qquad \text{a.s. for } \zeta \in \mathcal{P} \end{array} \tag{7.1}$$

To remove the potentially strongly conservative constraint satisfaction, *chance constraints* are often easier to handle. In a mechanical problem setting, this would amount, for instance, to satisfy the yield function $f(\boldsymbol{\sigma})$ up to a certain η-confidence level:

$$\mathbf{Prob}(f_i(\boldsymbol{\sigma}, \zeta) \leq 0) \geq \eta \tag{7.2}$$

However, these problems are often difficult to solve.

Robust optimization (RO) offers a simpler mathematical framework when assuming that the uncertain parameters belong to a known *uncertainty set* $\mathcal{U}$. Such sets are usually assumed to be convex and simple (L^2-ball for instance). A robust optimization constraint therefore reads as $f_i(\boldsymbol{x}, \zeta) \leq 0 \quad \forall \zeta \in \mathcal{U}$. A main advantage of RO problems is that the uncertainty parameters ζ can often be eliminated by reformulating the problem

e.g. the previous constraint is satisfied if and only if $\max_{\zeta \in \mathcal{U}} f_i(\boldsymbol{x}, \zeta) = f_i^{RO}(\boldsymbol{x}) \leq 0$. Under some specific conditions, $f_i^{RO}(\boldsymbol{x})$ can be computed analytically and retain a simple form. For instance, depending on the shape of $\mathcal{U}$, a robust LP problem can be reformulated as a deterministic LP or SOCP problem, possibly of larger size. Such results therefore call for the possibility of reusing computational approaches for deterministic problems (e.g. our IPM solver) to the context of uncertain problems. The use of robust optimization theory in mechanical engineering is however quite scarce and many remains to be explored.

* *
*

Appendix A

Some mathematical aspects

A.1 Convex sets and cones

Definition A.1 (Convex sets)**.** A subset $\mathcal{C}$ of $\mathbb{R}^n$ is called convex if $(1-\lambda)x+\lambda y \in \mathcal{C}$ whenever $x \in \mathcal{C}$, $y \in \mathcal{C}$ and $\lambda \in [0,1]$.

Definition A.2 (Cones)**.** A subset $\mathcal{K}$ of $\mathbb{R}^n$ is called a cone if it is closed under positive scalar multiplication, i.e. $\lambda x \in \mathcal{K}$ when $x \in \mathcal{K}$, and $\lambda \geq 0$.

Definition A.3 (Proper cone)**.** A cone $\mathcal{K} \subseteq \mathbb{R}^n$ is called a proper cone if it satisfies the following:

(a) $\mathcal{K}$ is convex;

(b) $\mathcal{K}$ is closed;

(c) $\mathcal{K}$ is solid, which means it has non-empty interior;

(d) $\mathcal{K}$ is pointed, which means it contains no line, or equivalently $\boldsymbol{x} \in \mathcal{K}$, $-\boldsymbol{x} \in \mathcal{K} \Rightarrow \boldsymbol{x} = \boldsymbol{0}$.

Definition A.4 (Generalized inequalities)**.** A generalized inequality is a partial ordering on $\mathbb{R}^n$ defined using a proper cone $\mathcal{K}$ defined as follows:

$$\boldsymbol{x} \preceq_{\mathcal{K}} \boldsymbol{y} \Leftrightarrow \boldsymbol{y} - \boldsymbol{x} \in \mathcal{K} \tag{A.1}$$

Similarly, a strict partial ordering can be defined as such:

$$\boldsymbol{x} \prec_{\mathcal{K}} \boldsymbol{y} \Leftrightarrow \boldsymbol{y} - \boldsymbol{x} \in \operatorname{int} \mathcal{K} \tag{A.2}$$

When $\mathcal{K} = \mathbb{R}^{n+}$, the partial ordering $\preceq_{\mathcal{K}}$ is nothing else than the usual ordering $\leq$ on $\mathcal{R}$.

Definition A.5 (Polar cones)**.** Let $\mathcal{K} \subset \mathbb{R}^n$ be a non empty convex cone. The polar of $\mathcal{K}$ is the set

$$\mathcal{K}^\circ = \{\boldsymbol{s} \in \mathbb{R}^n \mid \langle \boldsymbol{s}, \boldsymbol{x} \rangle \leq 0 \ \forall \boldsymbol{x} \in \mathcal{K}\} \tag{A.3}$$

Polarity may be seen as a generalization, in an unilateral way, of orthogonality. Hence, if $\mathcal{K}$ is a subspace then $\mathcal{K}^\circ$ is its orthogonal subspace. The polar cone obtained from $\mathcal{K}$ depends on the scalar product that is used in the definition. When $\mathcal{K}$ is a non empty closed convex cone, then $\mathcal{K}^\circ$ is also a non empty closed convex cone, and $\mathcal{K}^{\circ\circ} = \mathcal{K}$.

Definition A.6 (Conjugate or dual cone)**.** Let $\mathcal{K} \subset \mathbb{R}^n$ be a non empty convex cone. The dual cone of $\mathcal{K}$ is the set

$$\mathcal{K}^* = \{\boldsymbol{s} \in \mathbb{R}^n \mid \langle \boldsymbol{s}, \boldsymbol{x} \rangle \geq 0 \; \forall \boldsymbol{x} \in \mathcal{K}\} \tag{A.4}$$

Therefore, $\mathcal{K}^\circ = -\mathcal{K}^*$ (see Fig. A.1).

Remark. Given a non empty set $\mathcal{K}$, not necessarily convex, one may define also its dual cone using the same definition.

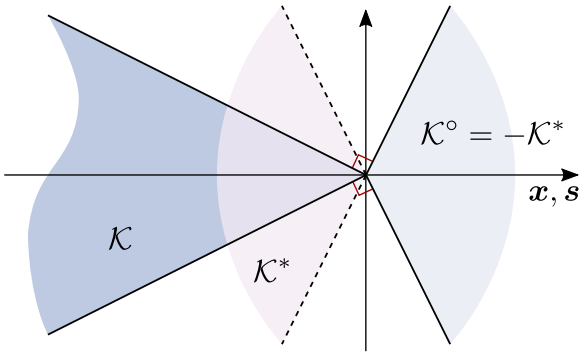


Figure A.1 – Primal, polar and dual cones

Definition A.7 (Self-dual cones)**.** Let $\mathcal{K} \subset \mathbb{R}^n$ be a non empty convex cone. A cone is said to be self-dual if and only if $\mathcal{K}^* = \mathcal{K}$.

The most important self-dual cones used in convex optimization are the following:

– the nonngative orthant

$$\mathcal{K} = \mathbb{R}^{n+} = \{\boldsymbol{x} \in \mathbb{R}^n \mid x_i \geq 0 \; \forall i = 1, \ldots, n\} = \mathcal{K}^* \tag{A.5}$$

– the second-order cone or Lorentz cone

$$\mathcal{K} = \mathcal{L}^{n+1} = \left\{\boldsymbol{x} = (x^0, \overline{\boldsymbol{x}}) \in \mathbb{R} \times \mathbb{R}^n \;\middle|\; x^0 \geq \|\overline{\boldsymbol{x}}\|_2\right\} = \mathcal{K}^* \tag{A.6}$$

– the rotated second-order cone

$$\mathcal{K} = \mathcal{L}_r^{n+2} = \left\{\boldsymbol{x} = (x^0, x^1, \overline{\boldsymbol{x}}) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^n \;\middle|\; 2x^0 x^1 \geq \|\overline{\boldsymbol{x}}\|_2^2\right\} = \mathcal{K}^* \tag{A.7}$$

As the name indicates, there is a simple relationship between quadratic and rotated quadratic cones (see Fig. A.2). Using the orthogonal transformation:

$$\boldsymbol{T}_n := \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & \boldsymbol{I}_{n-2} \end{bmatrix}$$

It is easy to verify that:

$$\boldsymbol{x} \in \mathcal{L} \quad \Leftrightarrow \quad \boldsymbol{T}_n \boldsymbol{x} \in \mathcal{L}_r$$

– the positive-semidefinite cone

$$\mathcal{K} = \mathcal{S}_+^n = \left\{ \boldsymbol{M} \in \mathcal{M}_n(\mathbb{R}) \;\middle|\; \boldsymbol{M}^T = \boldsymbol{M}, \boldsymbol{M} \succeq 0 \right\} = \mathcal{K}^* \tag{A.8}$$

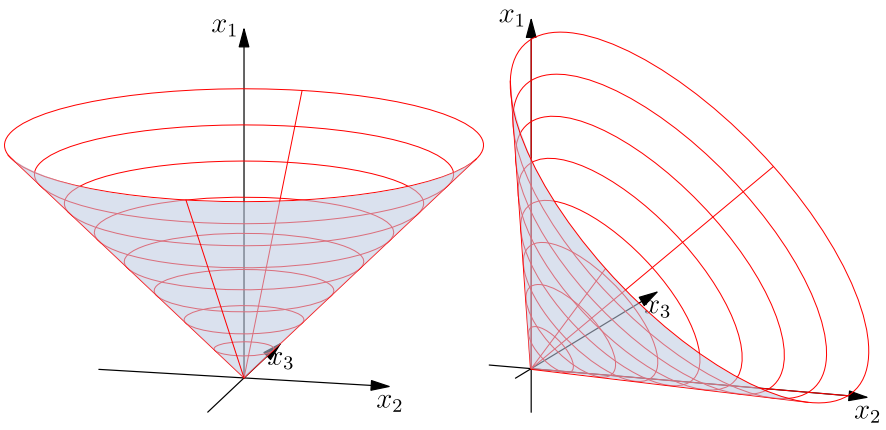


Figure A.2 – Boundary of the quadratic cone $x_1 \geq \sqrt{x_2^2 + x_3^2}$ and the rotated quadratic cone $2x_1x_2 \geq x_3^2,\ x_1 \geq 0,\ x_2 \geq 0$

Definition A.8 (Normal cone)**.** Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set. Often an indicator function, $\mathbb{1}_C : \mathbb{R}^n \longmapsto \mathbb{R}$ for the set C, is employed to remove the constraints of an optimization problem:

$$\min_{\boldsymbol{x} \in \mathcal{C}} f(\boldsymbol{x}) \;\Leftrightarrow\; \min_{\boldsymbol{x}} f(\boldsymbol{x}) + \mathbb{1}_{\mathcal{C}}(\boldsymbol{x}) \;\text{ where } \mathbb{1}_{\mathcal{C}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in \mathcal{C} \\ \infty & \text{if } \boldsymbol{x} \notin \mathcal{C} \end{cases} \tag{A.9}$$

The sub-differential of the indicator function x is known as the normal cone, $N_{\mathcal{C}}(x)$, of $\mathcal{C}$:

$$N_{\mathcal{C}}(\boldsymbol{x}) = \partial \mathbb{1}_{\mathcal{C}}(\boldsymbol{x}) = \left\{ \boldsymbol{g} \in \mathbb{R}^n : \boldsymbol{g}^T \boldsymbol{x} \geq \boldsymbol{g}^T \boldsymbol{y}\ ; \quad \forall \boldsymbol{y} \in \mathcal{C} \right\} \tag{A.10}$$

Therefore, the sub-differential of the indicator of a cone is its polar cone.

Definition A.9 (Epigraph)**.** The epigraph of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set

$$\mathbf{epi}(f) = \{(\boldsymbol{x}, t) \mid \boldsymbol{x} \in \mathbf{dom}(f),\ f(\boldsymbol{x}) \leq t\} \tag{A.11}$$

f is convex if and only if $\mathbf{epi}(f)$ is a convex set

Definition A.10 (Support function)**.** The support function of a set $\mathcal{C} \in \mathbb{R}^n$ describes the distances of the supporting hyperplanes of $\mathcal{C}$ from the origin:

$$\pi_{\mathcal{C}}(\boldsymbol{y}) = \sup \left(\boldsymbol{y}^T \boldsymbol{x} \,|\, \boldsymbol{x} \in \mathcal{C} \right) \tag{A.12}$$

The support function is convex and positively homogeneous of degree 1.

Definition A.11 (Second-order cone complementarity condition)**.** Let $\boldsymbol{x} = (x^0, \overline{\boldsymbol{x}}) \in \mathbb{R} \times \mathbb{R}^n$ and $\boldsymbol{s} = (s^0, \overline{\boldsymbol{s}}) \in \mathbb{R} \times \mathbb{R}^n$. $\boldsymbol{x}$ and $\boldsymbol{s}$ satisfy a second-order cone complementarity condition if:

$$\boldsymbol{x} \in \mathcal{K}, \quad \boldsymbol{s} \in \mathcal{K}^*, \quad \langle \boldsymbol{x}, \boldsymbol{s} \rangle = \boldsymbol{x}^T \boldsymbol{s} = x^0 s^0 + \overline{\boldsymbol{x}}^T \overline{\boldsymbol{s}} = 0 \tag{A.13}$$

Conventionally, the second-order cone constraint is defined as an inequality between a scalar x_0 and the Euclidean norm $\|\overline{\boldsymbol{x}}\|$ of the tail vector $\overline{\boldsymbol{x}}$.
In this manuscript, we slightly generalize the notion of second-order cone constraint for a tensor $\boldsymbol{\xi}$ instead of a vector $\overline{\boldsymbol{x}}$ in order to treat mechanical behavior defined using second-order symmetrical tensors. Specifically, for $(x^0, \boldsymbol{\xi}) \in \mathbb{R} \times \mathbb{S}^3$ and $(z^0, \boldsymbol{\zeta}) \in \mathbb{R} \times \mathbb{S}^3$, we consider:

$$x^0 \geq \|\boldsymbol{\xi}\|, \quad z^0 \geq \|\boldsymbol{\zeta}\|, \quad x^0 z^0 + \boldsymbol{\xi} : \boldsymbol{\zeta} = 0 \tag{A.14}$$

where $\|\boldsymbol{\xi}\| = \sqrt{\xi_{ij}\xi_{ij}}$ is the Frœbenius norm of the tensor.
The solution set can be explicitly described as follows. $(x^0, \boldsymbol{\xi}) \in \mathbb{R} \times \mathbb{S}^3$ and $(z^0, \boldsymbol{\zeta}) \in \mathbb{R} \times \mathbb{S}^3$ satisfy the second-order cone complementarity condition if any one of the following six cases is true:

(i) $x^0 > \|\boldsymbol{\xi}\|, \ (z^0, \boldsymbol{\zeta}) = (0, \mathbf{0})$
(ii) $(x^0, \boldsymbol{\xi}) = (0, \mathbf{0}), \ z^0 > \|\boldsymbol{\zeta}\|$
(iii) $x^0 = \|\boldsymbol{\xi}\| \neq 0, \ z^0 = \|\boldsymbol{\zeta}\| \neq 0, \ \exists \alpha > 0$ such that $\boldsymbol{\xi} = -\alpha \boldsymbol{\zeta}$
(iv) $x^0 = \|\boldsymbol{\xi}\| \neq 0, \ (z^0, \boldsymbol{\zeta}) = (0, \mathbf{0})$
(v) $(x^0, \boldsymbol{\xi}) = (0, \mathbf{0}), \ z^0 = \|\boldsymbol{\zeta}\| \neq 0$
(vi) $(x^0, \boldsymbol{\xi}) = (0, \mathbf{0}), \ (z^0, \boldsymbol{\zeta}) = (0, \mathbf{0})$

Another method also used in this manuscript is to identify second-order symmetrical tensors with their canonical vector representation thus using the initial definition of the complementarity condition.

A.2 Duality

Definition A.12 (Conjugate function)**.** Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ (not necessarily convex), the conjugate function of f is a function $f^* : (\mathbb{R}^n)^* = \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ defined by:

$$f^*(\boldsymbol{s}) = \sup_{\boldsymbol{x}} \{\langle \boldsymbol{s}, \boldsymbol{x} \rangle - f(\boldsymbol{x}) | \boldsymbol{x} \in \mathbb{R}^n\} \tag{A.15}$$

The mapping f^* is called the *Legendre-Fenchel* transformation. Moreover f^* is convex and we have:

$$f(\boldsymbol{x}) + f^*(\boldsymbol{s}) \geq \langle \boldsymbol{x}, \boldsymbol{s} \rangle \quad \forall \boldsymbol{x}, \boldsymbol{s} \in \mathbb{R}^n \times (\mathbb{R}^n)^* \tag{A.16}$$

The *Fenchel inequality* holds for all functions. In the case where f is closed proper convex function, the following three statements are equivalent:

(a) $\boldsymbol{s} \in \partial f(\boldsymbol{x})$
(b) $f(\boldsymbol{x}) + f^*(\boldsymbol{s}) = \langle \boldsymbol{x}, \boldsymbol{s} \rangle$

(c) $\boldsymbol{x} \in \partial f^*(\boldsymbol{s})$

A particularly interesting pair of conjugate functions are the indicator function of a convex set and the support function of the same set:

$$\mathbb{1}^*_{\mathcal{C}}(\boldsymbol{y}) = \sup_{\boldsymbol{x}}\{\langle \boldsymbol{y}, \boldsymbol{x}\rangle - \mathbb{1}_{\mathcal{C}}(\boldsymbol{y}) | \boldsymbol{x} \in \mathcal{C}\} = \pi_{\mathcal{C}}(\boldsymbol{y}) \tag{A.17}$$

Definition A.13 (Lagrangian duality)**.** Considering the following optimization problem in the standard form with $\boldsymbol{x} \in \mathbb{R}^n$:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) = \boldsymbol{0} \quad \forall i \in \{1, \dots, m\} \\ & h_j(\boldsymbol{x}) \leq \boldsymbol{0} \quad \forall j \in \{1, \dots, p\} \end{array} \tag{A.18}$$

We assume that the feasibility domain is not empty and that f is not necessarily convex. The basic idea in Lagrangian duality is to take the constraints in (A.18) into account by augmenting the objective function with a weighted sum of the constraint functions. Thus the Lagrangian $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with problem (A.18) is defined as:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} \nu_j h_j(\boldsymbol{x}) \tag{A.19}$$

We refer to $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ as the *Lagrange multipliers* or *dual variables* associated with the problem. The Lagrange dual function $l : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined by:

$$l(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\boldsymbol{x}} \quad \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \tag{A.20}$$

If a solution exist, the dual function yields a lower bound to the optimal value $\boldsymbol{x}^*$ of problem (A.18) which is called *weak duality* i.e.:

$$\forall \boldsymbol{\lambda}, \boldsymbol{\nu} \geq \boldsymbol{0} : l(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq f(\boldsymbol{x}^*) \tag{A.21}$$

In certain cases *strong duality* holds i.e. the value of the dual function coincide with the optimal value. A necessary condition is the *Slater constraint qualification* i.e.: if there exist a vector $\boldsymbol{x}$ satisfying

$$\begin{array}{ll} g_i(\boldsymbol{x}) = \boldsymbol{0} & \forall i \in \{1, \dots, m\} \\ h_j(\boldsymbol{x}) < \boldsymbol{0} & \forall j \in \{1, \dots, p\} \end{array}$$

* *
*

Appendix B

Nesterov-Todd symmetric scaling for second-order cones

B.1 Jordan algebra over the Lorentz cone

General results concerning Jordan algebra over the Lorentz second-order cone following Alizadeh and Goldfarb [2003] are given here. For a vector $\boldsymbol{v} = (v^0, \overline{\boldsymbol{v}}) \in \mathcal{K}^{n+1}$, we define:

$$\boldsymbol{e} = (1, \boldsymbol{0}) \in \mathbb{R}^{n+1} \tag{B.1}$$

$$\boldsymbol{mat}(\boldsymbol{v}) = \begin{bmatrix} v^0 & \overline{\boldsymbol{v}}^T \\ \overline{\boldsymbol{v}} & v^0 \boldsymbol{I}_n \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \tag{B.2}$$

$$\det(\boldsymbol{v}) = \det(\boldsymbol{mat}(\boldsymbol{v})) = (v^0)^2 - \|\overline{\boldsymbol{v}}\|^2 \tag{B.3}$$

$$\boldsymbol{Q} = \boldsymbol{diag}(1, -\boldsymbol{I}_n) \tag{B.4}$$

$$\hat{\boldsymbol{v}} = \boldsymbol{Q}\boldsymbol{v} = (v^0, -\overline{\boldsymbol{v}}) \tag{B.5}$$

$$(\boldsymbol{mat}(\boldsymbol{v}))^{-1} = \frac{1}{\det(\boldsymbol{v})} \begin{bmatrix} v^0 & -\overline{\boldsymbol{v}}^T \\ -\overline{\boldsymbol{v}} & \dfrac{\det(\boldsymbol{v})\boldsymbol{I}_n + \overline{\boldsymbol{v}}\overline{\boldsymbol{v}}^T}{v^0} \end{bmatrix} \tag{B.6}$$

The last equation is valid only if $\boldsymbol{v}$ is located strictly inside $\mathcal{K}^{n+1}$, i.e. $\boldsymbol{v} \in int(\mathcal{K}^{n+1})$, which is equivalent to $\det(\boldsymbol{v}) > 0$.

B.2 The symmetric scaling matrix

Let $\boldsymbol{X} = \boldsymbol{mat}(\boldsymbol{x})$ and $\boldsymbol{S} = \boldsymbol{mat}(\boldsymbol{s})$, the complementarity condition $\boldsymbol{x}^T\boldsymbol{s} = 0$ for the second-order Lorentz cone can be rewritten as:

$$\boldsymbol{x} \circ \boldsymbol{s} = \begin{Bmatrix} x^0 s^0 + \overline{\boldsymbol{x}}^T\overline{\boldsymbol{s}} \\ x^0\overline{\boldsymbol{s}} + s^0\overline{\boldsymbol{x}} \end{Bmatrix} = \boldsymbol{X}\boldsymbol{s} = \boldsymbol{S}\boldsymbol{x} = \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} \tag{B.7}$$

Starting from the linearized form of the complementarity condition written as:

$$(\boldsymbol{x} \circ \boldsymbol{s})^{(k+1)} \simeq \boldsymbol{X}^{(k)}\boldsymbol{S}^{(k)}\boldsymbol{e} + \boldsymbol{X}^{(k)}\Delta\boldsymbol{s} + \boldsymbol{S}^{(k)}\Delta\boldsymbol{x} \tag{B.8}$$

It can be shown that for $\boldsymbol{x}, \boldsymbol{s} \in \text{int}(\mathcal{K}^{m+1})$, there exists a unique matrix $\boldsymbol{F}$, depending on $\boldsymbol{x}$ and $\boldsymbol{s}$, such that

$$\boldsymbol{F}\boldsymbol{x} = \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{s}} = \boldsymbol{F}^{-T}\boldsymbol{S} \quad \text{and} \quad \boldsymbol{x} \circ \boldsymbol{s} = \tilde{\boldsymbol{x}} \circ \tilde{\boldsymbol{s}} \tag{B.9}$$

Let $\boldsymbol{V} = \boldsymbol{mat}(\tilde{\boldsymbol{x}}) = \boldsymbol{mat}(\tilde{\boldsymbol{s}})$ the associated matrix for the scaled point $\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{s}}$. The scaled point satisfies the following properties:

$$\boldsymbol{x} \in \mathcal{K}^{n+1} \Leftrightarrow \boldsymbol{v} \in \mathcal{K}^{n+1} \tag{B.10}$$

$$\boldsymbol{x} \in \operatorname{int}(\mathcal{K}^{n+1}) \Leftrightarrow \boldsymbol{v} \in \operatorname{int}(\mathcal{K}^{n+1}) \tag{B.11}$$

$$\boldsymbol{x} \circ \boldsymbol{s} = \|\boldsymbol{v}\|^2 \tag{B.12}$$

Using this symmetrical Nesterov-Todd scaling, the linearized complementarity condition can be rewritten as

$$(\boldsymbol{x} \circ \boldsymbol{s})^{(k+1)} \simeq (\boldsymbol{x} \circ \boldsymbol{s})^{(k)} + \boldsymbol{V}\boldsymbol{F}^{-T}\Delta\boldsymbol{s} + \boldsymbol{V}\boldsymbol{F}\Delta\boldsymbol{x} \tag{B.13}$$

In the specific case of the second-order Lorentz cone, $\boldsymbol{F}^{-T} = \boldsymbol{F}^{-1}$ and is explicitly given by the following symmetric matrix:

$$\boldsymbol{F} = \theta \begin{bmatrix} w^0 & \overline{\boldsymbol{w}}^T \\ \overline{\boldsymbol{w}} & \boldsymbol{I}_n + \frac{\overline{\boldsymbol{w}}^T\overline{\boldsymbol{w}}}{1+w^0} \end{bmatrix} = \theta\left(-\boldsymbol{Q} + \frac{(\boldsymbol{e}+\boldsymbol{w})(\boldsymbol{e}+\boldsymbol{w})^T}{1+w^0}\right) \tag{B.14}$$

$$\boldsymbol{F}^{-1} = \theta^{-2}\boldsymbol{Q}\boldsymbol{F}\boldsymbol{Q} \tag{B.15}$$

$$\boldsymbol{w} = \frac{\theta^{-1}\boldsymbol{s} + \theta\hat{\boldsymbol{x}}}{\sqrt{2}\sqrt{\boldsymbol{x}^T\boldsymbol{s} + \sqrt{\det(\boldsymbol{x})\det(\boldsymbol{s})}}} \tag{B.16}$$

$$\theta = \left(\frac{\det(\boldsymbol{s})}{\det(\boldsymbol{x})}\right)^{1/4} \tag{B.17}$$

Further details concerning the scaling matrix $\boldsymbol{F}$ can be found in Nesterov and Todd [1998], Alizadeh and Goldfarb [2003], Andersen et al. [2003] and the seminal works of Wright [1997] and Boyd and Vandenberghe [2004].

* *
*

Appendix C

Second-order cone complementarity formulation for the associated friction cones

Starting from the support function for the static cone $\Pi(\boldsymbol{n}, [\![\boldsymbol{u}]\!])$ we can demonstrate than the Lagrange multipliers for the static variables are no other than the kinematic variables. The proof of the duality of both cones is as follows:

$$\Pi(\boldsymbol{n}, \boldsymbol{g}) = \sup \left\{ (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{g} \mid \left\| \boldsymbol{\sigma}_T \right\| \leq -\mu\sigma_N \right\} \tag{C.1}$$

This upper bound problem can be resolved by forming the Lagrangian:

$$\mathcal{L}(\sigma_N, \boldsymbol{\sigma}_T, \lambda_0, \boldsymbol{\lambda}) = \sigma_N.g_N + \boldsymbol{\sigma}_T.\boldsymbol{g_T} + \begin{Bmatrix} -\mu\sigma_N \\ \boldsymbol{\sigma}_T \end{Bmatrix} . \begin{Bmatrix} \lambda_0 \\ \boldsymbol{\lambda} \end{Bmatrix} \tag{C.2}$$

where g_N and $\boldsymbol{g_T}$ are fixed parameters and $(\lambda_0, \boldsymbol{\lambda}) \in \mathcal{L}$ being the Lagrange multiplier associated with the conic constraint $(-\mu\sigma_N, \boldsymbol{\sigma}_T) \in \mathcal{L}$ verifying the complementarity conditions:

$$\begin{Bmatrix} -\mu\sigma_N \\ \boldsymbol{\sigma}_T \end{Bmatrix} \in \mathcal{L} \; ; \quad \begin{Bmatrix} \lambda_0 \\ \boldsymbol{\lambda} \end{Bmatrix} \in \mathcal{L} \; ; \quad \begin{Bmatrix} -\mu\sigma_N \\ \boldsymbol{\sigma}_T \end{Bmatrix} . \begin{Bmatrix} \lambda_0 \\ \boldsymbol{\lambda} \end{Bmatrix} = 0 \tag{C.3}$$

The Lagrangian can also be rewritten as follows:

$$\mathcal{L}(\sigma_N, \boldsymbol{\sigma}_T, \lambda_0, \boldsymbol{\lambda}) = \sigma_N(g_N - \mu.\lambda_0) + \boldsymbol{\sigma}_T(\boldsymbol{g_T} + \boldsymbol{\lambda}) \tag{C.4}$$

First order optimality conditions yield:

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \sigma_N} = 0 \;\Rightarrow\; g_N - \mu.\lambda_0 = 0 \;\Rightarrow\; \lambda_0 = \dfrac{g_N}{\mu} & \text{(C.5a)} \\ \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{\sigma}_T} = 0 \;\Rightarrow\; \boldsymbol{g}_T + \boldsymbol{\lambda} = 0 \;\Rightarrow\; \boldsymbol{\lambda} = -\boldsymbol{g}_T & \text{(C.5b)} \end{cases}$$

The conic constraint on the Lagrange multiplier therefore reads as:

$$\begin{Bmatrix} g_N/\mu \\ -\boldsymbol{g}_T \end{Bmatrix} \in \mathcal{L} \Leftrightarrow \|\boldsymbol{g}_T\| \leq \frac{g_N}{\mu} \Leftrightarrow \begin{Bmatrix} g_N \\ \boldsymbol{g}_T \end{Bmatrix} \in \mathcal{K}^*_{-1/\mu} \tag{C.6}$$

* *
*

Appendix D

Thermodynamical aspects for elasto-plastic media

We briefly recall here some basic notions of building constitutive behaviors through the framework of generalized standard materials [Halphen and Son Nguyen, 1975].

D.1 General aspects

Let us introduce the free energy density for an elastoplastic continuum in the case of isothermal evolution:

$$\rho\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, \boldsymbol{\alpha}) = \rho\psi^e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \rho\psi^p(\boldsymbol{\alpha}) \tag{D.1}$$

where $\rho\psi^e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p)$ is the Helmholtz elastic free energy and $\rho\psi^p(\boldsymbol{\alpha})$ is the stored hardening energy. We also introduce the plastic dissipation pseudo-potential such as $\mathcal{D}_1 = \phi(\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}})$ which is assumed to be a positive function and convex and positively homogeneous with respect to the state variable rates $\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}$.

The first law of thermodynamics can be expressed as follows:

$$\int_\Omega \int_{\tau_n}^{\tau_{n+1}} \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} d\Omega dt = \int_\Omega \Big[\rho\psi^e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \rho\psi^p(\boldsymbol{\alpha})\Big]_{\tau_n}^{\tau_{n+1}} d\Omega + \int_\Omega \int_{\tau_n}^{\tau_{n+1}} \phi(\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}}) d\Omega dt \tag{D.2}$$

along with the following state laws:

$$(\boldsymbol{\sigma}^{rev}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\varepsilon}^p}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}) \in \partial_{(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, \boldsymbol{\alpha})} \rho\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, \boldsymbol{\alpha}) \tag{D.3}$$

and complementary laws:

$$(\boldsymbol{\sigma}^{irr}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}) \in \partial_{(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}})} \phi(\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}}) \tag{D.4}$$

Let $\phi^*(\boldsymbol{\sigma}^{irr}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ be the Legendre-Fenchel transform of $\phi(\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}})$ with respect to $\dot{\boldsymbol{\varepsilon}}$ and $\dot{\boldsymbol{\alpha}}$. ϕ^* is no other than the indicator function for the convex domain of thermodynamical forces $\boldsymbol{\mathcal{A}}$, with a yield function $f(\boldsymbol{\sigma}^{irr}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \leq 0$. The complementary law are then equivalent to the flux evolution laws given by:

$$(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}) \in \partial_{(\boldsymbol{\sigma}^{irr}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}})} \phi^*(\boldsymbol{\sigma}^{irr}, \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{D.5}$$

In the case of time-independent plasticity, and considering that $\boldsymbol{\varepsilon}$ is not a dissipative variable, we have:

$$\boldsymbol{\sigma}^{irr} = 0 \tag{D.6a}$$
$$\boldsymbol{\sigma}^{rev} = \boldsymbol{\sigma} \tag{D.6b}$$

The plastic dissipation pseudo-potential is then independent from $\boldsymbol{\varepsilon}$ and $\dot{\boldsymbol{\varepsilon}}$: $\mathcal{D}_1 = \phi(\boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}})$ and the evolution law are as follows:

$$\dot{\boldsymbol{\alpha}} = \frac{\partial \phi^*(\boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha})}{\partial \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}} = \lambda \frac{\partial f(\boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}})}{\partial \boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}} \tag{D.7}$$

with λ being a plastic multiplier such as:

$$\lambda \geq 0, \quad f(\boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}) \leq 0, \quad \lambda . f(\boldsymbol{\mathcal{A}}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}) = 0 \tag{D.8}$$

This is classically known as the normality law.

D.2 Von Mises plasticity with isotropic hardening

In the case of von Mises plasticity with isotropic hardening, the two variables $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}^p$ are not sufficient to determine the mechanical state. It is necessary to introduce an additional internal variable such as the accumulated plastic strain or commonly known as the equivalent von Mises strain:

$$p(t) = \int_0^t \sqrt{\frac{2}{3}} \|\dot{\boldsymbol{\varepsilon}}^p\| dt \tag{D.9}$$

The internal variables are then $\boldsymbol{\alpha} = (\boldsymbol{\varepsilon}^p, p)$.

In the case of linear isotropic hardening, the free energy and dissipation pseudo-potential can be expressed as follows:

$$\rho\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) : \mathbb{D} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \frac{1}{2} E_h p^2 \tag{D.10}$$

$$\phi(\dot{\boldsymbol{\varepsilon}}^p) = \sqrt{\frac{2}{3}} \sigma_0 \|\dot{\boldsymbol{\varepsilon}}^p\| + \mathbb{1}_{\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^p)=0} \tag{D.11}$$

where E_h is the hardening modulus.
The state law are therefore as follows:

$$\boldsymbol{\sigma}^{rev} = +\rho \frac{\partial \psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p)}{\partial \boldsymbol{\varepsilon}} = \mathbb{D} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) = \boldsymbol{\sigma} \tag{D.12a}$$

$$\boldsymbol{\mathcal{A}}_{\boldsymbol{\varepsilon}^p} = -\rho \frac{\partial \psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p)}{\partial \boldsymbol{\varepsilon}^p} = \mathrm{dev}(\boldsymbol{\sigma}) = \boldsymbol{s} \tag{D.12b}$$

$$\mathcal{A}_p = -\rho \frac{\partial \psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, p)}{\partial p} = E_h . p \tag{D.12c}$$

Thus the yield function is then given by:

$$f(\boldsymbol{\sigma}, p) = \|\boldsymbol{s}\| - (\sigma_0 + E_h.p) \tag{D.13}$$

and the evolution laws:

$$\dot{\boldsymbol{\varepsilon}}^p = \lambda \frac{\partial f(\boldsymbol{\sigma}, p)}{\partial \boldsymbol{\sigma}} = \lambda \sqrt{\frac{3}{2}} \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} \tag{D.14a}$$

$$\dot{p} = \lambda \tag{D.14b}$$

The distinction between purely elastic behavior and plastic loading can be expressed as follows:

$$\begin{cases} f(\boldsymbol{\sigma}) < 0 & \Rightarrow \quad \dot{\boldsymbol{\varepsilon}}^p = \boldsymbol{o} \\ \dot{\boldsymbol{\varepsilon}}^p \neq \boldsymbol{o} & \Rightarrow \quad f(\boldsymbol{\sigma}) = 0 \end{cases} \tag{D.15}$$

* *
*

Appendix E

Strain measures, stress measures and mapping tensors

E.1 The generalized strain measure

Within an infinitesimal neighbourhood of a generic material particle, pure rotations can be distinguished from pure stretching by means of the polar decomposition of the deformation gradient $\boldsymbol{F} = \boldsymbol{RU}$. Under the action of pure rotations ($\boldsymbol{F} = \boldsymbol{R}$), the distances between particles within this neighbourhood remain fixed. Under stretching, we say that the region surrounding the material particle is strained. To quantify straining, i.e. to evaluate how much $\boldsymbol{U}$ departs from $\boldsymbol{I}$ being a rigid deformation, some kind of strain measure needs to be defined.

Since $\boldsymbol{U}$ is symmetrical, it follows that it admits a unique spectral decomposition

$$\boldsymbol{U} = \sum_{i=1}^{3} \lambda_i \boldsymbol{N}_i \otimes \boldsymbol{N}_i = \sum_{i=1}^{3} \lambda_i \boldsymbol{M}_{ii} \tag{E.1}$$

where $\{\lambda_1, \lambda_2, \lambda_3\}$ are the eigenvalues and the triad $\{\boldsymbol{N}_1, \boldsymbol{N}_2, \boldsymbol{N}_3\}$ the Lagrangian triad or Lagrangian principle directions. The same triad can be expressed using $\boldsymbol{M}_{ij}$ which defines the full-symmetric spectral basis tensors:

$$\boldsymbol{M}_{ij} = \begin{cases} \boldsymbol{N}_i \otimes \boldsymbol{N}_i & \text{if } i = j \\ \frac{1}{2}(\boldsymbol{N}_i \otimes \boldsymbol{N}_j + \boldsymbol{N}_j \otimes \boldsymbol{N}_i) & \text{if } i \neq j \end{cases} \tag{E.2}$$

The generalized Lagrangian strain measure is defined through an isotropic function of the pure stretching tensor $\boldsymbol{U}$ defined in its spectral space along the principal Lagrangian directions:

$$\boldsymbol{E}^* = f^*(\boldsymbol{U}) = \sum_{i=1}^{3} f^*(\lambda_i) \boldsymbol{N}_i \otimes \boldsymbol{N}_i = \sum_{i=1}^{3} f^*(\lambda_i) \boldsymbol{M}_{ii} \tag{E.3}$$

One of the most common families of strain measures is the Seth-Hill family defined by $f^*(x) = \frac{1}{m}(x^m - 1)$, that is:

$$\boldsymbol{E}^{(m)} = \sum_{i=1}^{3} \frac{1}{m}(\lambda_i^m - 1) \boldsymbol{M}_{ii} \tag{E.4}$$

This family contains the most common strain measures used in literature:

$$\boldsymbol{E}_{GL} = \boldsymbol{E}^{(2)} = \sum_{i=1}^{3} \frac{1}{2}(\lambda_i^2 - 1)\boldsymbol{M}_{ii} \qquad \text{Green-Lagrange deformation tensor} \tag{E.5}$$

$$\boldsymbol{E}^{(1)} = \sum_{i=1}^{3} (\lambda_i - 1)\boldsymbol{M}_{ii} \qquad \text{Biot deformation tensor} \tag{E.6}$$

$$\boldsymbol{E}^{(0)} = \sum_{i=1}^{3} \frac{1}{2}\ln(\lambda_i)\boldsymbol{M}_{ii} \qquad \text{Hencky deformation tensor} \tag{E.7}$$

where $\boldsymbol{E}^{(0)} = \lim_{m \to 0} \boldsymbol{E}^{(m)}$.

E.2 Fourth order mapping tensors

Since all strain measures given by the generalized definition, or more specifically by the Seth-Hill family, are unique for a given deformation gradient, there is a one-to-one mappings that allows the passage between the different measures [Caminero et al., 2011; Latorre and Montáns, 2016; Miehe et al., 2002]. Let $\boldsymbol{E}^*$ and $\boldsymbol{E}^\dagger$ be two different generalized material strain measures as functions of the material stretch tensor $\boldsymbol{U}$. The fourth order geometrical mapping tensor $\mathbb{M}_{\boldsymbol{E}^\dagger}^{\boldsymbol{E}^*}$ such as $\boldsymbol{E}^* = \mathbb{M}_{\boldsymbol{E}^\dagger}^{\boldsymbol{E}^*} : \boldsymbol{E}^\dagger$ is given by:

$$\mathbb{M}_{\boldsymbol{E}^\dagger}^{\boldsymbol{E}^*} = \sum_{i=1}^{3} \frac{f^*(\lambda_i)}{f^\dagger(\lambda_i)} \boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ii} \tag{E.8}$$

one can easily verify that $\boldsymbol{E}^\dagger = \mathbb{M}_{\boldsymbol{E}^*}^{\boldsymbol{E}^\dagger} : \boldsymbol{E}^* = (\mathbb{M}_{\boldsymbol{E}^\dagger}^{\boldsymbol{E}^*})^{-1} : \boldsymbol{E}^*$ with:

$$\mathbb{M}_{\boldsymbol{E}^*}^{\boldsymbol{E}^\dagger} = (\mathbb{M}_{\boldsymbol{E}^\dagger}^{\boldsymbol{E}^*})^{-1} = \sum_{i=1}^{3} \frac{f^\dagger(\lambda_i)}{f^*(\lambda_i)} \boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ii} \tag{E.9}$$

In a similar way, there is a one-to-one mapping between the rates of each measure such that:

$$\dot{\boldsymbol{E}}^* = \mathbb{M}_{\dot{\boldsymbol{E}}^\dagger}^{\dot{\boldsymbol{E}}^*} : \dot{\boldsymbol{E}}^\dagger \tag{E.10}$$

$$\mathbb{M}_{\dot{\boldsymbol{E}}^\dagger}^{\dot{\boldsymbol{E}}^*} = \frac{\partial \boldsymbol{E}^*}{\partial \boldsymbol{E}^\dagger} = \sum_{i=1}^{3} \frac{df^*(\lambda_i)/d\lambda_i}{df^\dagger(\lambda_i)/d\lambda_i} \boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ii} + \sum_{i=1}^{3}\sum_{j \neq i}^{3} \frac{f^*(\lambda_j) - f^*(\lambda_i)}{f^\dagger(\lambda_j) - f^\dagger(\lambda_i)} \boldsymbol{M}_{ij} \otimes \boldsymbol{M}_{ij} \tag{E.11}$$

In most cases, the reference strain measure used is the Green-Lagrange measure $\dot{\boldsymbol{E}}^\dagger = \dot{\boldsymbol{E}}_{GL}$, and the geometrical mapping tensor taken according to this measure:

$$\mathbb{M}_{\dot{\boldsymbol{E}}_{GL}}^{\dot{\boldsymbol{E}}^*} = \sum_{i=1}^{3} \frac{df^*(\lambda_i)/d\lambda_i}{\lambda_i} \boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ii} + \sum_{i=1}^{3}\sum_{j \neq i}^{3} \frac{2(f^*(\lambda_j) - f^*(\lambda_i))}{\lambda_j^2 - \lambda_i^2} \boldsymbol{M}_{ij} \otimes \boldsymbol{M}_{ij} \tag{E.12}$$

Note that all the defined geometrical tensors verify both the major and the minor symmetry conditions:

$$\mathbb{M}_{ijkl} = \mathbb{M}_{klij} \quad \text{and} \quad \mathbb{M}_{ijkl} = \mathbb{M}_{jikl} = \mathbb{M}_{ijlk} \tag{E.13}$$

and that:

$$\mathbb{M}^{\dot{\boldsymbol{E}}_{GL}}_{\dot{\boldsymbol{E}}^*} = \left(\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}\right)^{-1} \tag{E.14}$$

E.3 Classical stress measures

Using Cauchy's theorem, the Cauchy stress tensor $\boldsymbol{\sigma}$ is defined as the linear application that corresponds for each normal vector $\boldsymbol{n}$ of the current configuration $\Omega \equiv \Omega^t$, the current surface traction vector $\boldsymbol{t}$:

$$\boldsymbol{t} = \boldsymbol{\sigma}.\boldsymbol{n} \tag{E.15}$$

Local equilibrium in the actual configuration is written as follows:

$$\operatorname{div} \boldsymbol{\sigma} + \rho \boldsymbol{b}_d = 0 \tag{E.16}$$

The equilibrium of the resulting moments implies the symmetry of the Cauchy stress tensor, i.e.: $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$.

Using Nanson's formula, the traction forces can be written in terms of the first Piola–Kirchhoff stress tensor and the surface area in the reference configuration

$$\int_{\Omega=\varphi(\partial\Omega^0)} \boldsymbol{\sigma}.\boldsymbol{n} ds = \int_{\partial\Omega^0} J\boldsymbol{\sigma}\boldsymbol{F}^{-T}\boldsymbol{N} dS = \int_{\partial\Omega^0} \boldsymbol{P}\boldsymbol{N} dS \tag{E.17}$$

where $J = \det \boldsymbol{F}$ and $\boldsymbol{P} = J\boldsymbol{\sigma}\boldsymbol{F}^{-T}$ is the first Piola-Kirchhoff stress tensor. This tensor is not symmetrical, thus the local equilibrium over the reference configuration is written as follows:

$$\operatorname{Div} \boldsymbol{P} + \rho_0 \boldsymbol{b}_d = 0 \tag{E.18}$$

and the resulting moment equilibrium as:

$$\boldsymbol{P}.\boldsymbol{F}^T = \boldsymbol{F}.\boldsymbol{P}^T \tag{E.19}$$

To ensure symmetry, the second Piola-Kirchhoff stress tensor is introduced:

$$\boldsymbol{S} = \boldsymbol{F}^{-1}\boldsymbol{P} = J\boldsymbol{F}^{-1}\boldsymbol{\sigma}\boldsymbol{F}^{-T} \tag{E.20}$$

such as $\boldsymbol{S} = \boldsymbol{S}^T$ and:

$$\operatorname{Div}(\boldsymbol{F}\boldsymbol{S}) + \rho_0 \boldsymbol{b}_d = 0 \tag{E.21}$$

E.4 Work-conjugacy and the generalized stress measure

The principle of virtual work can be found by introducing a virtual velocity field $\boldsymbol{v}$ and dualising the equilibrium Eq. (E.16) as follows:

$$\int_{\Omega} \operatorname{div} \boldsymbol{\sigma}.\boldsymbol{v} d\Omega + \int_{\Omega} \rho \boldsymbol{b}_d.\boldsymbol{v} d\Omega = 0 \qquad \forall \boldsymbol{v} \tag{E.22}$$

Using the divergence theorem, Eq. (E.22) can be written as such:

$$-\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\nabla}_x \boldsymbol{v} d\Omega + \int_{\partial\Omega} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{v} ds + \int_{\Omega} \rho \boldsymbol{b}_d.\boldsymbol{v} d\Omega = 0 \qquad \forall \boldsymbol{v} \tag{E.23}$$

Since $\boldsymbol{\sigma}$ is symmetrical, $\boldsymbol{\sigma} : \boldsymbol{\nabla}_x \boldsymbol{v} = \boldsymbol{\sigma} : \operatorname{sym}(\boldsymbol{\nabla}_x \boldsymbol{v}) = \boldsymbol{\sigma} : \boldsymbol{d}$, with $\boldsymbol{d}$ being the classical spatial deformation rate, Eq. (E.23) becomes:

$$\underbrace{-\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{d} d\Omega}_{\substack{\text{Intenal} \\ \text{mechanical power}}} + \underbrace{\int_{\partial\Omega} (\boldsymbol{\sigma}.\boldsymbol{n}).\boldsymbol{v} ds + \int_{\Omega} \rho \boldsymbol{b}_d.\boldsymbol{v} d\Omega}_{\substack{\text{External} \\ \text{mechanical power}}} = 0 \qquad \forall \boldsymbol{v} \tag{E.24}$$

known as the principle of virtual work in the current configuration.

The *internal mechanical power* is the rate of change of the work done by internal forces. Let p_m be the internal stress power with respect to the unit volume such as in the current configuration we have $p_m = \boldsymbol{\sigma} : \boldsymbol{d}$. Starting from the stress power in the current configuration, various alternative representation can be expressed as a function of the other stress tensors:

$$\begin{aligned} \int_{\Omega} p_m d\Omega = \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{d} d\Omega = \int_{\Omega^0} J\boldsymbol{\sigma} : \boldsymbol{d} d\Omega \\ = \int_{\Omega^0} \boldsymbol{P} : \dot{\boldsymbol{F}} d\Omega = \int_{\Omega^0} \boldsymbol{S} : \dot{\boldsymbol{E}}_{GL} d\Omega = \int_{\Omega^0} \boldsymbol{T}^* : \dot{\boldsymbol{E}}^* d\Omega \end{aligned} \tag{E.25}$$

Each pair of strain measure and stress tensor constitutes a *work-conjugate* pair. Using this equivalence, any stress tensor $\boldsymbol{T}^*$ associated to the generalized strain measure $\dot{\boldsymbol{E}}^*$ can be written in terms of the other, more classical, measures. In most cases, the Green-Lagrange measure is considered as the reference measure since its derivatives are the easiest to calculate and it is found in almost every finite elements code. Thus we have:

$$p_m = \boldsymbol{S} : \dot{\boldsymbol{E}}_{GL} = \boldsymbol{T}^* : \dot{\boldsymbol{E}}^* = \boldsymbol{T}^* : \left(\frac{\partial \boldsymbol{E}^*}{\partial \boldsymbol{E}_{GL}} : \dot{\boldsymbol{E}}_{GL} \right) = \boldsymbol{T}^* : \left(\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} : \dot{\boldsymbol{E}}_{GL} \right) \tag{E.26}$$

The generalized stress measure is then defined in reference to the second Piola-Kirchhoff stress tensor by the following purely geometric relation:

$$\boldsymbol{S} = \boldsymbol{T}^* : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} \tag{E.27}$$

or equivalently:

$$\boldsymbol{T}^* = \boldsymbol{S} : \mathbb{M}^{\dot{\boldsymbol{E}}_{GL}}_{\dot{\boldsymbol{E}}^*} \tag{E.28}$$

E.5 Computing constitutive tangent modulus

In a general constitutive equation, the constitutive tangent modulus $\mathbb{D}^*$ relating the generalized strain increments and the generalized stress increments is required. Note that time derivatives for Lagrangian measures are always objective. This tangent modulus can also be written in terms of the classical tangent modulus $\mathbb{D}_{GL}$ relating the Green-Lagrange strain increments and the second Piola-Kirchhoff stress increments using the geometric mapping tensors. Assuming that we have derived the constitutive tangent modulus $\mathbb{D}^*$ for a generalized strain measure $\boldsymbol{E}^*$ and its work conjugate stress tensor $\boldsymbol{T}^*$ such that:

$$\dot{\boldsymbol{T}}^* = \mathbb{D}^* : \dot{\boldsymbol{E}}^* \tag{E.29}$$

the equivalent tangent modulus $\mathbb{D}_{GL}$ such that $\dot{\boldsymbol{S}} = \mathbb{D}_{GL} : \dot{\boldsymbol{E}}_{GL}$ can be calculated as follows:

$$\dot{\boldsymbol{S}} = \dot{\boldsymbol{T}}^* : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} + \boldsymbol{T}^* : \dot{\mathbb{M}}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}$$

$$\dot{\boldsymbol{S}} = \left(\frac{\partial \boldsymbol{T}^*}{\partial \boldsymbol{E}^*} : \frac{\partial \boldsymbol{E}^*}{\partial \boldsymbol{E}_{GL}} : \dot{\boldsymbol{E}}_{GL}\right) : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} + \boldsymbol{T}^* : \left(\frac{\partial \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}}{\partial \boldsymbol{E}_{GL}} : \dot{\boldsymbol{E}}_{GL}\right)$$

$$\dot{\boldsymbol{S}} = \mathbb{D}^* : \left(\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} : \dot{\boldsymbol{E}}_{GL}\right) : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} + \boldsymbol{T}^* : \left(\mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} : \dot{\boldsymbol{E}}_{GL}\right)$$

$$\dot{\boldsymbol{S}} = \left(\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} : \mathbb{D}^* : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} + \boldsymbol{T}^* : \mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}\right) : \dot{\boldsymbol{E}}_{GL} \tag{E.30}$$

Thus we have

$$\mathbb{D}_{GL} = \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} : \mathbb{D}^* : \mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} + \boldsymbol{T}^* : \mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} \tag{E.31}$$

The previous relation is found using the major symmetry of $\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}$ and the introduction of the sixth-order geometric tensor $\mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}$ relating the rate of $\mathbb{M}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} = \partial \boldsymbol{E}^* / \partial \boldsymbol{E}_{GL}$ and the rate of $\boldsymbol{E}_{GL}$:

$$\mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} = \frac{\partial^2 \boldsymbol{E}^*}{\partial \boldsymbol{E}_{GL} \partial \boldsymbol{E}_{GL}} \tag{E.32}$$

Both geometrical tensors are required in order to formally map the tangent modulus associated to one strain measure to the tangent modulus associated to the other strain measure. In practice, computing the six-order tensor $\mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}$ then doing the double contraction is not efficient therefore formulas for calculating directly the double contraction $\boldsymbol{T}^* : \mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}}$ have been derived:

$$
\begin{aligned}
\boldsymbol{T}^* : \mathfrak{L}^{\dot{\boldsymbol{E}}^*}_{\dot{\boldsymbol{E}}_{GL}} &= \sum_{i=1}^{3} F(\lambda_i) T^*_{ii} \boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ii} \\
&+ \sum_{i=1}^{3} \sum_{j \neq i} G(\lambda_i, \lambda_j) T^*_{ii} \boldsymbol{M}_{ij} \otimes \boldsymbol{M}_{ij} \\
&+ \sum_{i=1}^{3} \sum_{j \neq i} G(\lambda_i, \lambda_j) T^*_{ij} \left(\boldsymbol{M}_{ii} \otimes \boldsymbol{M}_{ij} + \boldsymbol{M}_{ij} \otimes \boldsymbol{M}_{ii} \right) \\
&+ \sum_{i=1}^{3} \sum_{j \neq i} \sum_{k \neq j \neq i} H(\lambda_i, \lambda_j, \lambda_k) T^*_{ik} \left(\boldsymbol{M}_{ij} \otimes \boldsymbol{M}_{jk} + \boldsymbol{M}_{jk} \otimes \boldsymbol{M}_{ij} \right)
\end{aligned}
\tag{E.33}
$$

where

$$T^*_{ij} = \boldsymbol{T}^* : \boldsymbol{M}_{ij} \tag{E.34}$$

$$F(\lambda_i) = -\frac{2}{\lambda_i^4} \tag{E.35}$$

$$G(\lambda_i, \lambda_j) = \frac{8(f^*(\lambda_j) - f^*(\lambda_i)) - 4\Lambda_{ij}/\lambda_i}{\Lambda_{ij}^2} \tag{E.36}$$

$$H(\lambda_i, \lambda_j, \lambda_k) = 8 \frac{-\Lambda_{jk} f^*(\lambda_i) - \Lambda_{ki} f^*(\lambda_j) - \Lambda_{ij} f^*(\lambda_k)}{\Lambda_{jk} \Lambda_{jk} \Lambda_{ij}} \tag{E.37}$$

$$\Lambda_{ij} = \lambda_j^2 - \lambda_i^2 \tag{E.38}$$

Note that $H(\lambda_i, \lambda_j, \lambda_k) = H(\lambda_i, \lambda_k, \lambda_j) = H(\lambda_k, \lambda_i, \lambda_j)$ but $G(\lambda_i, \lambda_j) \neq G(\lambda_j, \lambda_i)$. Also, when two or three principal stretches converge to the same value we have:

$$H(\lambda_i, \lambda_j, \lambda_k \to \lambda_j) = G(\lambda_i, \lambda_j) \tag{E.39}$$

$$H(\lambda_i, \lambda_j \to \lambda_i, \lambda_k \to \lambda_i) = G(\lambda_i, \lambda_j \to \lambda_i) = F(\lambda_i) \tag{E.40}$$

Note that all the corresponding formulas for the logarithmic strain measure and their geometrical mapping tensors can be found by replacing $f^*(\lambda_i)$ with $\ln(\lambda_i)$.

* *
*

Bibliography

Abbas, M., A. Ern, and N. Pignet (2018). Hybrid high-order methods for finite deformations of hyperelastic materials. *Computational Mechanics*, 62(4):909–928. 125

Alart, P. (1997). Méthode de Newton généralisée en mécanique du contact. *Journal de Mathématiques Pures et Appliquées*, 76(1):83–108. 9

Alart, P. and A. Curnier (1991). A mixed formulation for frictional contact problems prone to newton like solution methods. *Computer methods in applied mechanics and engineering*, 92(3):353–375. 9

Alauzet, F. and P. Frey (2006). Estimateur d'erreur géométrique et métriques anisotropes pour l'adaptation de maillage. Partie I : aspects théoriques. *Institut national de recherche en informatique et en automatique*, 37(1):55–57. 58

Alizadeh, F. and D. Goldfarb (2003). Second-order cone programming. *Mathematical Programming*, 95(1):3–51. 11, 23, 33, 35, 173, 174

Alnæs, M. S., J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells (2015). The fenics project version 1.5. *Archive of Numerical Software*, 3(100). 133

Andersen, E. (2013). On formulating quadratic functions in optimization problems. Technical Report TR-1-2013, MOSEK. 38, 39

Andersen, E., C. Roos, and T. Terlaky (2003). On implementing a primal-dual interior-point method for conic quadratic optimization. *Mathematical Programming*, 95(2):249–277. 10, 33, 38, 46, 174

Ben Belgacem, F., C. Bernardi, A. Blouza, and M. Vohralik (2012). On the unilateral contact between membranes. Part 2: a posteriori analysis and numerical experiments. *IMA Journal of Numerical Analysis*, 32(3):1147–1172. 54

Ben-Tal, A. and A. Nemirovski (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Siam. 22

Bilotta, A., L. Leonetti, and G. Garcea (2012). An algorithm for incremental elastoplastic analysis using equality constrained sequential quadratic programming. *Computers & Structures*, 102:97–107. 76

Bleyer, J. (2017). Advances in the simulation of viscoplastic fluid flows using interior-point methods. *Computer Methods in Applied Mechanics and Engineering*. 11, 34, 38, 46, 122

Bleyer, J. and G. Hassen (2020). Automated formulation and resolution of limit analysis problems. *arXiv:2005.04779*. 12

Bleyer, J., M. Maillard, P. De Buhan, and P. Coussot (2015). Efficient numerical computations of yield stress fluid flows using second-order cone programming. *Computer Methods in Applied Mechanics and Engineering*, 283:599–614. 11, 46

Boyd, S. P. and L. Vandenberghe (2004). *Convex optimization*. Cambridge University Press, Cambridge, UK ; New York. 10, 21, 22, 30, 33, 46, 174

Caminero, M. n., F. J. Montáns, and K.-J. Bathe (2011). Modeling large strain anisotropic elasto-plasticity with logarithmic strain and stress measures. *Computers & Structures*, 89(11-12):826–843. 124, 182

Capurso, M. and G. Maier (1970). Incremental elastoplastic analysis and quadratic optimization. *Meccanica*, 5(2):107–116. 11, 76

Chen, W.-F. (2013). *Limit analysis and soil plasticity*. Elsevier. 10

Christensen, P. W. (2002). A nonsmooth newton method for elastoplastic problems. *Computer Methods in Applied Mechanics and Engineering*, 191(11-12):1189–1219. 76

Christensen, P. W., A. Klarbring, J. S. Pang, and N. Strömberg (1998). Formulation and comparison of algorithms for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 42(1):145–173. 46

Coorevits, P., P. Hild, and M. Hjiaj (2001). A posteriori error control of finite element approximations for Coulomb's frictional contact. *SIAM Journal on Scientific Computing*, 23(3):976–999. 53, 54

Corre, G., A. Lebée, K. Sab, M. K. Ferradi, and X. Cespedes (2018). Higher-order beam model with eigenstrains: theory and illustrations. *Journal of Applied Mathematics and Mechanics*, 98(7):1040–1065. 164

Corre, G., A. Lebée, K. Sab, M. K. Ferradi, and X. Cespedes (2020). A new higher-order elastoplastic beam model for reinforced concrete. *Meccanica*, 55(4):791–813. 164

Dantzig, G., A. Orden, and P. Wolfe (1955). The generalized simplex method for minimizing a linear form under linear inequality restraints. *Pacific Journal of Mathematics*, 5(2):183–195. 31

de Almeida, J. and J. de Freitas (1991). Alternative approach to the formulation of hybrid equilibrium finite elements. *Computers & Structures*, 40(4):1043–1047. 12, 57

de Almeida, J. P. M. and E. A. Maunder (2017). *Equilibrium finite element formulations*. John Wiley & Sons, Ltd, Chichester, UK. 12, 57, 90

de Almeida, J. P. M. and O. J. B. A. Pereira (1996). A set of hybrid equilibrium finite elements models for the analysis of three dimensional solids. *International Journal for Numerical Methods in Engineering*, 39(16):2789–2802. 12, 91

de Freitas, J. and Z. Ji (1996). Hybrid-trefftz boundary integral formulation for simulation of singular stress fields. *International Journal for Numerical Methods in Engineering*, page 28. 12

de Saxcé, G. and Z. Q. Feng (1998). The bipotential method: a constructive approach to design the complete contact law with friction and improved numerical algorithms. *Mathematical and Computer Modelling*, 28(4):225–245. 11, 49, 76

de Souza Neto, E. A., D. Peri, and D. R. J. Owen (2008). *Computational Methods for Plasticity*. John Wiley & Sons, Ltd, Chichester, UK. 10

Debongnie, J., H. Zhong, and P. Beckers (1995). Dual analysis with general boundary conditions. *Computer Methods in Applied Mechanics and Engineering*, 122(1-2):183–192. 53, 54

Debongnie, J.-F. (1983). A general theory of dual error bounds by finite elements. Technical report, Université de Liège. 53, 54

EN1993–1-1 (2005). *Eurocode 3: Design of Steel Structures, Part 1.1: General rules and rules for buildings*. AFNOR. 3, 106, 110, 113

EN1993–1-8 (2005). *Eurocode 3: Design of Steel Structures, Part 1.8: Design of Joints*. AFNOR. 106, 108, 109, 110, 113, 114, 119

Ferradi, M. K., X. Cespedes, and M. Arquier (2013). A higher order beam finite element with warping eigenmodes. *Engineering Structures*, 46:748–762. 164

Ferradi, M. K., A. Lebée, A. Fliscounakis, X. Cespedes, and K. Sab (2016). A model reduction technique for beam analysis with the asymptotic expansion method. *Computers & Structures*, 172:11–28. 164

Fortin, M. and R. Glowinski (2000). *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*. Elsevier. 9

Fraeijs de Veubeke, B. (1951). Diffusion des inconnues hyperstatiques dans les voilures à longerons couplés. *Bulletin du Service Technique de l'Aéronautique*, 24. 51, 57

Fraeijs de Veubeke, B. (1963). Upper and lower bounds in matrix structural analysis. Technical Report P-24, LTAS. 12

Fraeijs de Veubeke, B. (1965). Displacement and equilibrium models in the finite element method. *Stress analysis*, page 145–197. 12, 51

Gilbert, M., C. Casapulla, and H. Ahmed (2006). Limit analysis of masonry block structures with non-associative frictional joints using linear programming. *Computers & structures*, 84(13-14):873–887. 47, 162

Glineur, F. (2001). Conic optimization: an elegant framework for convex optimization. *Belgian Journal of Operations Research, Statistics and Computer Science*, 41(1-2):5–28. 21

Hai-Chang, H. (1954). On some variational principles in the theory of elasticity and the theory of plasticity. *Acta Physica Sinica*, 10(3):259. 51

Halphen, B. and Q. Son Nguyen (1975). Sur les matériaux standard généralisés. *Journal de Mécanique*, 14:39–63. 125, 177

Helfer, T., J. Bleyer, T. Frondelius, I. Yashchuk, T. Nagel, and D. Naumov (2020). The 'mfrontgenericinterfacesupport' project. *Journal of Open Source Software*, 5(48):2003. 133

Hertog, D. d. (1994). *Interior Point Approach to Linear, Quadratic and Convex Programming: Algorithms and Complexity*. Mathematics and Its Applications. Springer Netherlands. 33

Hestenes, M. R. (1969). Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5):303–320. 9

Hill, R. (1950). *The mathematical theory of plasticity*. Clarendon Press, Oxford. 10, 82

Hintermüller, M., K. Ito, and K. Kunisch (2002). The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888. 9

Hjiaj, M., Z.-Q. Feng, G. De Saxcé, and Z. Mróz (2004). Three-dimensional finite element computations for frictional contact problems with non-associated sliding rule. *International Journal for Numerical Methods in Engineering*, 60(12):2045–2076. 9

Hjiaj, M., J. Fortin, and G. de Saxcé (2003). A complete stress update algorithm for the non-associated Drucker–Prager model including treatment of the apex. *International Journal of Engineering Science*, 41(10):1109–1143. 11, 76

Hüeber, S. and B. I. Wohlmuth (2005). A primal–dual active set strategy for non-linear multibody contact problems. *Computer Methods in Applied Mechanics and Engineering*, 194(27-29):3147–3166. 9

Jansen, B. (1997). *Interior Point Techniques in Optimization: Complementarity, Sensitivity and Algorithms*. Applied Optimization. Springer US. 33

Johansen, K. (1962). *Yield-line theory*. Cement and Concrete Association London. 10

Joli, P. and Z.-Q. Feng (2008). Uzawa and Newton algorithms to solve frictional contact problems within the bi-potential framework. *International Journal for Numerical Methods in Engineering*, 73(3):317–330. 9

Kanno, Y. (2011). *Nonsmooth mechanics and convex optimization*. CRC Press, Boca Raton, FL. 9, 11, 76, 81

Kanno, Y. (2016). A fast first-order optimization approach to elastoplastic analysis of skeletal structures. *Optimization and Engineering*, 17(4):861–896. 76

Kanno, Y., J. A. C. Martins, and A. Pinto da Costa (2006). Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem. *International Journal for Numerical Methods in Engineering*, 65(1):62–83. 9, 11, 49

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395. 32

Kempeneers, M., J.-F. Debongnie, and P. Beckers (2009). Pure equilibrium tetrahedral finite elements for global error estimation by dual analysis. *International Journal for Numerical Methods in Engineering*, 81(4):513–536. 12, 57, 90, 92

Kikuchi, N. and J. Oden (1988). *Contact problems in elasticity : a study of variational inequalities and finite element methods*. Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics. 9

Klarbring, A. (1986). A mathematical programming approach to three-dimensional contact problems with friction. *Computer Methods in Applied Mechanics and Engineering*, 58(2):175–200. 9

Kleinert, J., B. Simeon, and M. Obermayr (2014). An inexact interior point method for the large-scale simulation of granular material. *Computer Methods in Applied Mechanics and Engineering*, 278:567–598. 46, 162

Krabbenhøft, K., A. Lyamin, and S. Sloan (2007). Formulation and solution of some plasticity problems as conic programs. *International Journal of Solids and Structures*, 44(5):1533–1549. 11, 25, 76

Krabbenhoft, K., M. Karim, A. Lyamin, and S. Sloan (2012a). Associated computational plasticity schemes for nonassociated frictional materials. *International Journal for Numerical Methods in Engineering*, 90(9):1089–1117. 47, 162

Krabbenhoft, K., A. Lyamin, J. Huang, and M. Vicente da Silva (2012b). Granular contact dynamics using mathematical programming methods. *Computers and Geotechnics*, 43:165–176. 11, 46

Krabbenhoft, K., A. V. Lyamin, S. W. Sloan, and P. Wriggers (2007). An interior-point algorithm for elastoplasticity. *International Journal for Numerical Methods in Engineering*, 69(3):592–626. 11, 76

Kučera, R., J. Machalová, H. Netuka, and P. Ženčák (2013). An interior-point algorithm for the minimization arising from 3d contact problems with friction. *Optimization methods and software*, 28(6):1195–1217. 162

Ladevèze, P. and L. Chamoin (2016). The Constitutive Relation Error Method: A General Verification Tool. In Chamoin, L. and P. Díez, editors, *Verifying Calculations - Forty Years On: An Overview of Classical Verification Techniques for FEM Simulations*, pages 59–94. Springer International Publishing, Cham. 54

Ladevèze, P. and J. P. Pelle (2005). *Mastering calculations in linear and nonlinear mechanics*. Mechanical engineering series. Springer Science, New York. 52, 53, 58

Latorre, M. and F. J. Montáns (2016). Stress and strain mapping tensors and general work-conjugacy in large strain continuum mechanics. *Applied Mathematical Modelling*, 40(5-6):3938–3950. 124, 182

Laursen, T. and J. Simo (1993). Algorithmic symmetrization of coulomb frictional problems using augmented lagrangians. *Computer methods in applied mechanics and engineering*, 108(1-2):133–146. 9

Lemaitre, J. and J.-L. Chaboche (1994). *Mechanics of solid materials*. Cambridge university press. 125

Lobo, M. S., L. Vandenberghe, S. Boyd, and H. Lebret (1998). Applications of second-order cone programming. *Linear Algebra and its Applications*, 284(1-3):193–228. 10, 22, 46

Logg, A., K.-A. Mardal, and G. Wells (2012). *Automated solution of differential equations by the finite element method: The FEniCS book*, volume 84. Springer Science & Business Media. 133

Logg, A. and G. N. Wells (2010). Dolfin: Automated finite element computing. *ACM Transactions on Mathematical Software*, 37(2). 133

Louf, F., J.-P. Combe, and J.-P. Pelle (2003). Constitutive error estimator for the control of contact problems involving friction. *Computers & Structures*, 81(18-19):1759–1772. 54

Lyamin, A. V. and S. W. Sloan (2002a). Lower bound limit analysis using non-linear programming. *International Journal for Numerical Methods in Engineering*, 55(5):573–611. 11, 46, 85

Lyamin, A. V. and S. W. Sloan (2002b). Upper bound limit analysis using linear finite elements and non-linear programming. *International Journal for Numerical and Analytical Methods in Geomechanics*, 26(2):181–216. 11, 46, 85

Magisano, D. and G. Garcea (2020). Fiber-based shakedown analysis of three-dimensional frames under multiple load combinations: Mixed finite elements and incremental-iterative solution. *International Journal for Numerical Methods in Engineering*, n/a(n/a). 10

Maier, G. (1968). A quadratic programming approach for certain classes of non linear structural problems. *Meccanica*, 3(2):121–130. 11, 76

Maier, G. (1969). Complementary plastic work theorems in piecewise-linear elastoplasticity. *International Journal of Solids and Structures*, 5(3):261–270. 11, 76

Maier, G. (1970). A matrix structural theory of piecewise linear elastoplasticity with interacting yield planes. *Meccanica*, 5(1):54–66. 10

Makrodimopoulos, A. (2010). Remarks on some properties of conic yield restrictions in limit analysis. *International Journal for Numerical Methods in Biomedical Engineering*, 26(11):1449–1461. 25, 76

Makrodimopoulos, A. and C. M. Martin (2006). Lower bound limit analysis of cohesive-frictional materials using second-order cone programming. *International Journal for Numerical Methods in Engineering*, 66(4):604–634. 10, 11, 46, 85

Makrodimopoulos, A. and C. M. Martin (2007). Upper bound limit analysis using simplex strain elements and second-order cone programming. *International Journal for Numerical and Analytical Methods in Geomechanics*, 31(6):835–865. 10, 11, 46, 85, 88

Mangoni, D., A. Tasora, and R. Garziera (2018). A primal–dual predictor–corrector interior point method for non-smooth contact dynamics. *Computer Methods in Applied Mechanics and Engineering*, 330:351–367. 46

Martin, C. M. and A. Makrodimopoulos (2008). Finite-element limit analysis of Mohr–Coulomb materials in 3d using semidefinite programming. *Journal of Engineering Mechanics*, 134(4):339–347. 10, 11, 76

Martin, J. B., B. D. Reddy, T. B. Griffin, and W. W. Bird (1987). Applications of mathematical programming concepts to incremental elastic-plastic analysis. *Engineering Structures*, 9(3):171–176. 10

Maxwell, J. C. (1866). On reciprocal diagrams in space, and their relation to Airy's function of stress. *Proceedings of the London Mathematical Society*, 1(1):58–63. 57, 90

Mehrotra, S. (1992). On the Implementation of a Primal-Dual Interior Point Method. *SIAM Journal on Optimization*, 2(4):575–601. 33, 37, 38

Miehe, C. (2002). Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation. *International Journal for numerical methods in engineering*, 55(11):1285–1322. 78

Miehe, C., N. Apel, and M. Lambrecht (2002). Anisotropic additive plasticity in the logarithmic strain space: modular kinematic formulation and implementation based on incremental minimization principles for standard materials. *Computer Methods in Applied Mechanics and Engineering*, 191(47-48):5383–5425. 122, 123, 124, 125, 137, 182

Morera, G. (1892). Soluzione generale delle equazioni indefinite dell'equilibrio di un corpo continuo. *Atti Accad. Naz. Lincei, Rend. Cl. Fis. Mat. Natur., V. Ser*, 1(1):137–141. 57, 90

Mosek (2013). Mosek modeling manual. User manual, Mosek. 39

Mosek (2019). Mosek Modeling Cookbook. 22

Nesterov, Y. E. and A. Nemirovskii (1994). *Interior-Point Polynomial Algorithms in Convex Programming*. Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics. 33

Nesterov, Y. E. and M. J. Todd (1997). Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations research*, 22(1):1–42. 33

Nesterov, Y. E. and M. J. Todd (1998). Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on optimization*, 8(2):324–364. 33, 38, 174

Nocedal, J., A. Wächter, and R. A. Waltz (2009). Adaptive barrier update strategies for nonlinear interior methods. *SIAM Journal on Optimization*, 19(4):1674–1693. 30, 38

Ortiz, M. and L. Stainier (1999). The variational formulation of viscoplastic constitutive updates. *Computer Methods in Applied Mechanics and Engineering*, 171(3-4):419–444. 78

Papadopoulos, P. and J. Lu (1998). A general framework for the numerical solution of problems in finite elasto-plasticity. *Computer Methods in Applied Mechanics and Engineering*, 159(1-2):1–18. 137

Pian, T. H. H. (1964). Derivation of element stiffness matrices by assumed stress distributions. *AIAA Journal*, 2(7):1333–1336. 12, 57, 90

Pian, T. H. H. (1976). Variational principles for incremental finite element methods. *Journal of the Franklin Institute*, 302(5):473–488. 52

Popp, A. and W. Wall (2014). Dual mortar methods for computational contact mechanics–overview and recent developments. *GAMM-Mitteilungen*, 37(1):66–84. 9

Powell, M. J. (1969). *A method for nonlinear constraints in minimization problems*. Academic Press. 9

Rockafellar, R. (1994). Nonsmooth optimization. *Mathematical Programming: State of the Art*, pages 248–258. 31

Salahi, M., J. Peng, and T. Terlaky (2008). On Mehrotra-Type Predictor-Corrector Algorithms. *SIAM Journal on Optimization*, 18(4):1377–1397. 37

Salençon, J. (1983). *Calcul à la rupture et analyse limite*. Presses de l'Ecole Nationale des Ponts et Chaussées. 10

Salençon, J. (2013). *Yield design*. Mechanical engineering and solid mechanics series. ISTE ; Wiley, London : Hoboken, NJ. 10, 82, 83

Sander, O. and P. Jaap (2019). Solving primal plasticity increment problems in the time of a single predictor–corrector iteration. *Computational Mechanics*, pages 1–23. 76

Save, M. (1995). *Atlas of limit loads of metal plates, shells and disks*. North-Holland Series in Applied Mathematics & Mechanics. Elsevier Science Ltd. 10

Save, M. A., C. E. Massonnet, and G. de Saxce (1997). *Plastic limit analysis of plates, shells, and disks*. North Holland. 10

Schlaich, J., K. Schäfer, and M. Jennewein (1987). Toward a consistent design of structural concrete. *PCI journal*, 32(3):74–150. 10

Simo, J. (1992). Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory. *Computer Methods in Applied Mechanics and Engineering*, 99(1):61–112. 137

Simo, J. C. and T. J. Hughes (2006). *Computational inelasticity*. Springer. 10

Simo, J. C. and T. Laursen (1992). An augmented lagrangian treatment of contact problems involving friction. *Computers & Structures*, 42(1):97–116. 9

Sloan, S. W. (1987). Substepping schemes for the numerical integration of elastoplastic stress-strain relations. *International Journal for Numerical Methods in Engineering*, 24(5):893–911. 10

Smith, M. (2009). *ABAQUS/Standard User's Manual, Version 6.9*. Simulia. 133

Stewart, D. E. and J. C. Trinkle (1996). An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and Coulomb friction. *International Journal for Numerical Methods in Engineering*, 39(15):2673–2691. 9

Tanoh, G., Y. Renard, and D. Noll (2004). Computational experience with an interior point algorithm for large scale contact probles. *Optimization Online*, 10(2):18. 46

Temizer, I., M. Abdalla, and Z. Gürdal (2014). An interior point method for isogeometric contact. *Computer Methods in Applied Mechanics and Engineering*, 276:589–611. 46

Trahair, N. S., M. A. Bradford, D. Nethercot, and L. Gardner (2007). *The Behaviour and Design of Steel Structures to EC3, Fourth Edition*. CRC Press. 6

Vincent, H. (2018). *Développement d'un modèle de calcul de la capacité ultime d'éléments de structure (3D) en béton armé, basé sur la théorie du calcul à la rupture*. phdthesis, Université Paris-Est. 10, 162

Vincent, H., M. Arquier, J. Bleyer, and P. de Buhan (2018). Yield design-based numerical analysis of three-dimensional reinforced concrete structures. *International Journal for Numerical and Analytical Methods in Geomechanics*, 42(18):2177–2192. 10, 11, 162

Washizu, K. (1968). *Variational methods in elasticity and plasticity*. Pergamon Press. 51

Wriggers, P. and G. Zavarise (1993). Application of augmented Lagrangian techniques for non-linear constitutive laws in contact interfaces. *Communications in Numerical Methods in Engineering*, 9(10):815–824. 9

Wright, M. (2005). The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bulletin of the American mathematical society*, 42(1):39–56. 10, 33, 46

Wright, S. J. (1997). *Primal-dual Interior-Point Methods*. SIAM. 33, 174

Yonekura, K. and Y. Kanno (2012). Second-order cone programming with warm start for elastoplastic analysis with von Mises yield criterion. *Optimization and Engineering*, 13(2):181–218. 11, 76, 80, 126

Zhang, H., J. Li, and S. Pan (2011). New second-order cone linear complementarity formulation and semi-smooth newton algorithm for finite element analysis of 3d frictional contact problem. *Computer Methods in Applied Mechanics and Engineering*, 200(1-4):77–88. 9

Zhang, X., K. Krabbenhoft, and D. Sheng (2014). Particle finite element analysis of the granular column collapse problem. *Granular Matter*, 16(4):609–619. 11

Zhu, C. (1995). A finite element-mathematical programming method for elastoplastic contact problems with friction. *Finite Elements in Analysis and Design*, 20(4):273–282. 11

Zienkiewicz, O. C. and R. L. Taylor (2000). *The finite element method*. Butterworth-Heinemann, Oxford ; Boston, 5th ed edition. 58