Manuel Kutschka

Robustness Concepts for Knapsack and Network Design Problems under Data Uncertainty

Gamma-, Multi-band, Submodular, and Recoverable Robustness

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Robustness Concepts for Knapsack and Network Design Problems under Data Uncertainty

Robustness Concepts for Knapsack and Network Design Problems under Data Uncertainty

Gamma-, Multi-band, Submodular, and Recoverable Robustness

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Wirtschaftsmathematiker

Manuel Kutschka

aus Berlin

Berichter: Univ.-Prof. Dr. Ir. Arie M. C. A. Koster
Prof. Dr. Andreas Bley

Tag der mündlichen Prüfung: 2. Dezember 2013

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2013
Zugl.: (RWTH) Aachen, Univ., Diss., 2013

978-3-95404-593-8

D 82 (Diss. RWTH Aachen University, 2013)

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2013
Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft

978-3-95404-593-8

ABSTRACT

In this thesis, we consider mathematical optimization under data uncertainty using mixed integer linear programming (MILP) techniques. Our investigations follow the deterministic paradigm known as robust optimization. It allows to tackle an uncertain variant of a problem without increasing its complexity in theory or decreasing its computational tractability in practice.

We consider four robustness concepts for robust optimization and describe their parametrization, application, and evaluation. The concepts are Γ-robustness, its generalization multi-band robustness, the more general submodular robustness, and the two-stage approach called recoverable robustness.

For each concept, we investigate the corresponding robust generalization of the knapsack problem (KP), a fundamental combinatorial problem and subproblem of almost every integer linear programming (ILP) problem, and many other optimization problems. We present ILP formulations, detailed polyhedral investigations including new classes of valid inequalities, and algorithms for each robust KP. In particular, our results for the submodular and recoverable robust KP are novel. Additionally, the recoverable robust KP is experimentally evaluated in detail.

Further, we consider the Γ-robust generalization of the capacitated network design problem (NDP). For example, the NDP arises from many application areas such as telecommunications, transportation, or logistics. For the Γ-robust NDP, we present MILP formulations, detailed polyhedral insights with new classes of valid inequalities, and algorithms. Moreover, we consider the multi-band robust NDP, its MILP formulations, and generalized polyhedral results of the Γ-robust NDP.

Furthermore, we present computational results for the Γ-robust NDP using real-life measured uncertain data from telecommunication networks. These detailed representative studies are based on our work with the German ROBUKOM project in cooperation with Nokia Siemens Networks GmbH & Co. KG.

Finally, we give concluding remarks on the presented robustness concepts and discuss future research directions.

Keywords: Γ-Robustness, Multi-Band Robustness, Submodular Robustness, Recoverable Robustness, Robust Knapsack Problem, Robust Network Design Problem

Mathematics Subject Classification (2010): 90C27, 90C35, 90C57, 90C90, 90B18

ZUSAMMENFASSUNG

Die vorliegende Dissertation untersucht mathematische Optimierung unter Unsicherheiten mittels Methoden der gemischt-ganzzahligen linearen Programmierung (MILP). Dabei folgen wir dem deterministischen Paradigma der robusten Optimierung. Dieses ermöglicht die Lösung unsicherer Problemvarianten ohne Erhöhung der theoretischen Komplexität oder Verschlechterung der praktischen Lösbarkeit.

Wir untersuchen vier Robustheitskonzepte und beschreiben deren Parametrisierung, Anwendung, und Evaluierung. Die untersuchten Konzepte sind Γ-Robustheit [Γ-*robustness*], deren neue Verallgemeinerung Multi-Band-Robustheit [*multi-band robustness*], die neue allgemeinere submodulare Robustheit [*submodular robustness*], sowie der adaptive zweistufige Ansatz der wiederherstellbaren Robustheit [*recoverable robustness*].

Für jedes Konzept untersuchen wir die entsprechende robuste Verallgemeinerung des Rucksackproblems [*knapsack problem*] (KP), eines der fundamentalen kombinatorischen Probleme und Teilproblem fast jeden Problems der ganzzahligen linearen Programmierung (ILP) und vieler anderer Optimierungsprobleme. Wir präsentieren ILP-Formulierungen, detaillierte polyedrische Studien mit neuen Klassen gültiger Ungleichungen und Algorithmen für jedes robuste KP. Dabei sind insbesondere unsere Ergebnisse für das submodular- und wiederherstellbar-robuste KP neuartig. Zusätzlich evaluieren wir das wiederherstellbar-robuste KP experimentell in einer detaillierten Rechenstudie.

Außerdem betrachten wir die Γ-robuste Verallgemeinerung des kapazitierten Netzwerkplanungsproblems [*capacitated network design problem*] (NDP). Das NDP ist z. B. in Anwendungsproblemen aus den Bereichen Telekommunikation, Transport oder Logistik zu finden. Für das Γ-robuste NDP präsentieren wir MILP-Formulierungen, detaillierte polyedrische Ergebnisse, neue Klassen gültiger Ungleichungen und Algorithmen. Zusätzlich untersuchen wir das Multi-Band-robuste NDP, dessen MILP-Formulierungen, sowie dessen polyedrische Struktur als Verallgemeinerung des Γ-robusten NDP.

Im Weiteren beschreiben wir unsere detaillierten Rechenstudien zum Γ-robusten NDP mit real gemessenen unsicheren Daten verschiedener Telekommunikationsnetze. Diese repräsentativen Rechenergebnisse basieren auf unserer Arbeit im Projekt ROBUKOM in Kooperation mit der Nokia Siemens Networks GmbH & Co. KG.

Wir schließen diese Dissertation mit einigen zusammenfassenden Bemerkungen und einer Diskussion zukünftiger Forschungsrichtungen ab.

DANKSAGUNG

Meine Forschung, die schlussendlich zu dieser Dissertation geführt hat, begann 2008 in England. Seither haben viele meinen wissenschaftlichen Werdegang begleitet und beeinflusst. Ich möchte die Gelegenheit nutzen ihnen meinen Dank auszusprechen.

Zuerst gebührt natürlich Arie Koster Dank für die Betreuung meiner Arbeit. Seit meiner Zeit als Student am Zuse Institut Berlin hat er mir das wissenschaftliche Arbeiten nähergebracht, mir die Möglichkeit gegeben an spannenden Projekten mitzuwirken und mich dabei immer "gefordert und gefördert".

Weiter bedanke ich mich bei Andreas Bley für die gute Zusammenarbeit im ROBUKOM-Projekt und nicht zuletzt für seine Bereitschaft meine Dissertation als zweiter Berichter zu begutachten.

Dann möchte ich mich bei meinen beiden Bürokollegen Grit Claßen und Stephan Lemkens für die angenehme Arbeitsatmosphäre und die unzähligen Diskussionen bedanken, die immer hilfreich und manchmal auch fachlich waren.

Ebenso möchte ich Christina Büsing für die sehr gute lange Zusammenarbeit danken.

Eberhard Triesch und meinen Kollegen am Lehrstuhl II für Mathematik danke ich für etwas mehr als vier interessante Jahre gemeinsamen Arbeitens.

Außerdem danke ich allen meinen Mitautoren von Veröffentlichungen im Zusammenhang mit dieser Dissertation: Thomas Bauschert, Christina Büsing, Grit Claßen, Fabio D'Andreagiovanni, Svetoslav Duhovnikov, Sebastian Goderbauer, Arie Koster, Christian Raack, Franz Rambach, Dominic Schupke und Issam Tahiri.

Für das geduldige und gründliche Korrekturlesen früherer Versionen dieser Arbeit bedanke ich mich nochmals bei Grit Claßen und Arie Koster.

Schließlich möchte ich meiner wundervollen Frau Yvonne von ganzem Herzen dafür danken, dass sie mich stets unterstützt, mir den Rücken frei gehalten, und manchen Unmut ertragen hat. Danke!

Aachen, im Dezember 2013 Manuel Kutschka

CONTENTS

INTRODUCTION

Our life is affected by countless complex interdependent processes forming the backbone of economy including production, trading, logistics, distribution, and communication. Oftentimes, thousands, millions or even more decisions must be taken into account to plan and operate these processes.

Mathematical optimization strives for providing theory, models, and methods to tackle these problems and obtain relevant solutions in practice. As a sub-discipline of mathematics, mathematical optimization investigates the (hidden) problem structure to identify and exploit reoccurring (sub)structures and develop tailor-made strategies.

An example of reoccurring structures are networks. A network describes the dependencies between entities, e. g., logistics networks describe the places from which and where to commodities are sent, energy networks characterize the supply of energy, or telecommunication networks specify the possible ways of information exchange. The problem to plan a network, i. e., determine its layout and the rules how it can be used later on, is called the capacitated network design problem (NDP) in mathematical optimization.

Another crucial aspect are scarce resources and the resulting challenging question of prioritizing their usage. For example, the (scarce) loading capacity of a delivery truck must be utilized in the best way in logistics, the limited monetary budget must be managed in finance, or the available bandwidth of an optical fiber must be shared among several optical data signals in telecommunication. Mathematically, this leads to the so-called knapsack problem (KP).

The understanding of problems such as the NDP or the KP allows a more accurate mathematical modeling of the underlying real-world problem. However, the mathematical model is always a simplification; oftentimes a rather rough one. In particular, the temporal dynamics and uncertainties of real-life processes are hard to take into account, e. g., travel times are not constant but subject to delays like traffic jams in practice, the food production depends on future weather conditions, or telecommunication demands fluctuate significantly by the daytime with peaks during certain hours of the day and lows during the night.

Mathematical optimization offers several paradigms to incorporate uncertainty into the mathematical framework. Robust optimization is one of these. Here, the data uncertainty is modeled implicitly by a so-called uncertainty set. The robust optimization problem asks to find an optimal solution that is feasible for any possible data realization in this uncertainty set. In particular, robust linear optimization offers several advantages over other approaches. The definition of an uncertainty set does not rely on the knowledge of probability distributions and is thus often better suited to applied problems where only a finite discrete set of historical data is available, if any. In addition, robust solutions are feasible for all realizations in the uncertainty set by definition. Further, the complexity of robust linear programs does not increase compared to the original non-robust linear program under mild conditions. Instead, there often exist compact reformulations, i. e., formulations that are at most polynomially larger than the original non-robust formulations. Thus, robust linear optimization problems are more computationally tractable than other mathematical optimization problems applying different paradigms to handle uncertainties.

In this thesis, we consider robust integer linear optimization problems. In particular, we consider four different robustness concepts and the associated uncertainty sets. For each concept, we investigate the corresponding robust KP presenting integer linear programming formulations, results on the polyhedral structure of the solution sets, and algorithms to solve the occurring separation problems or the robust KP themselves. Moreover, we study the corresponding robust NDP problem for two of the concepts, also presenting several integer linear programming formulations, polyhedral insights, and (separation) algorithms to solve the (separation) problems.

Our theoretical investigations are completed by two extensive computational studies: one for the recoverable robust KP, the other for the Γ-robust NDP. The latter uses real-life uncertain data from an application in telecommunication and is based on our work with the German ROBUKOM project in cooperation with Nokia Siemens Networks GmbH & Co. KG.

Contributions. Some results are partially based on joint work as common in the area of applied mathematical optimization. Whenever this is the case, we state explicitly our coauthors and possible prior published publications of our joint work in the beginning of the corresponding chapters.

The main contributions of the thesis are the following.

- The introduction and study of the concept of submodular robustness.
- A detailed investigation of the recoverable robust KP. In particular with a Γ-robust scenario set and the k-removal recovery rule.
- A detailed investigation of the submodular robust KP introducing the classes of submodular robust $(1, k)$-configuration and weight inequalities.
- A study of the structure of covers and their extendability for each considered robust KP.
- A detailed investigation of the Γ-robust NDP including new classes of valid and facet-defining inequalities (e. g., Γ-robust cutset inequalities, Γ-robust envelope

inequalities, Γ-robust arc residual capacity inequalities, and Γ-robust metric inequalities) and algorithms solving the corresponding separation problems as well as the Γ-robust NDP problem itself.

- A first-time investigation of the multi-band robust NDP including mixed integer linear programming formulations, polyhedral studies yielding new classes of valid inequalities (multi-band robust cutset inequalities and multi-band robust metric inequalities), and algorithms to solve the corresponding separation problems. In particular, we point out by examples how results of the Γ-robust NDP can be generalized to the multi-band robust setting.

- Representative extensive computational studies for two recoverable robust knapsack variants and one robust network design problem (the latter with application to telecommunications).

Outline. This thesis is structured into three parts.

In *Part I - Concepts*, we introduce the relevant mathematical methodology, provide a survey on related work, and introduce the objects of research for this thesis. Therefore, we first recap mathematical requirements focusing on mathematical optimization and introducing the classic knapsack and capacitated network design problems in Chapter 1. A brief primer on relevant applications in telecommunications is given at the end of the same chapter. In Chapter 2, we present a detailed survey on literature related to mathematical optimization under data uncertainty and in particular robust optimization. Next in Chapter 3, we introduce the four robustness concepts which are our main focus of investigation in this thesis: Γ-robustness, multi-band robustness, submodular robustness, and recoverable robustness. Moreover, we address the evaluation of robustness discussing several alternative approaches.

In *Part II - Robust Knapsack Problems*, we consider the robust counterpart of the classic knapsack problem for each of the four robustness concepts. For each resulting robust knapsack problem, we present mathematical formulations, study the corresponding polyhedral solution sets identifying strong classes of valid inequalities, and develop algorithms solving the occurring separation problems as well as the robust knapsack problem itself. Following this structure of investigation, we consider the Γ-robust knapsack problem in Chapter 4, the more general multi-band robust knapsack problem in Chapter 5, and the submodular robust knapsack problem in Chapter 6 which generalizes the multi-band robust knapsack problem even further. In Chapter 7, we consider the recoverable robust knapsack problem which is an integrated two-stage problem. Two special cases are of particular interest for us whereof one generalizes the one-stage Γ-robust knapsack problem. We conclude this part of the thesis in Chapter 8 reporting on the results of extensive computational studies we carried out on recoverable robust knapsack problems. Therefore, we focus on the rather general recoverable robust knapsack problem evaluating the effect of the robustness parameters, the strength of the derived valid inequalities and finally the overall performance in a cut-and-branch approach to solve this problem.

Part III - Robust Network Design Problems is structured similarly to Part II. Here, we consider the robust counterpart of the classic (capacitated) network design problem for selected robustness concepts. We primarily focus on the Γ-robust network design problem and provide several mathematical formulations for this problem, investigate the corresponding polyhedral structure, derive several classes of valid inequalities, and algorithms. Our investigation is described in great detail in Chapter 9. Afterwards, we consider the more general multi-band robust network design problem in Chapter 10 pointing out how results for the Γ-robust network design problem are generalized to the multi-band robust setting. In Chapter 11, we describe the results of computational studies on robust network design problems and in particular the Γ-robust design of telecommunication networks. We experimentally compare its different formulations, the derived classes of valid inequalities, separation algorithms, and algorithms to solve the Γ-robust network design problem itself. For our experiments, we use historical real-life traffic measurements to define the data uncertainty.

Finally, we give concluding remarks to the contributions of this thesis and discuss potential future research directions.

PART ONE

CONCEPTS

CHAPTER ONE

MATHEMATICAL PRELIMINARIES

The first chapter of this thesis gives a brief survey of the mathematical prerequisites and thereby introduces the reader to the used notation. Furthermore, we present two important optimization problems: the knapsack problem and the capacitated network design problem. In later parts of this thesis we will consider variants of both problems where the input data is subject to some random uncertainty. In this chapter we introduce both problems in their classic deterministic settings, reporting on related work, important results, and polyhedral insights to their solution sets. The last section of this chapter is a primer to the telecommunications application area. There, we give a short introduction to the structure and operation of telecommunication networks and the related mathematical optimization challenges.

1.1 Basics

In the following, we introduce our notation while reminding the reader of some basics of (integer) linear optimization and polyhedral combinatorics. We assume that most results are well-known and therefore give only a brief introduction without making the claim to be complete. For further reading, we refer to some well-established monographs below.

An introduction to graphs, networks, flows, and related algorithms is given in the excellent book by Ahuja et al. [10]. The standard book about complexity theory is written by Garey and Johnson [72]. A well-written introduction to linear optimization with various examples is given by Chvatal [53]. A more formal and more recent survey on linear programming can be found in Dantzig [61].

Algorithmic combinatorial optimization is described by Grötschel et al. [76]. The books by Schrijver [143], Nemhauser and Wolsey [126], and Wolsey [159] consider the theory of integer programming, combinatorial optimization and the related polyhedral theory. The three-volume encyclopedic book by Schrijver [144] gives an excellent survey on state-of-the-art combinatorial optimization theory and techniques referencing thousands of original work. A good survey of the algorithmic aspects of solving mixed integer linear programs and experimental results are described in detail by Achterberg [6].

Linear algebra. We denote by $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, the sets of *integer*, *rational*, and *real numbers*, respectively. The set of positive natural numbers is denoted by $\mathbb{N}$. Let $K \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Then, $K_{>0}$, $K_{\geq 0}$, $K_{\leq 0}$, and $K_{<0}$ denotes the positive, nonnegative, nonpositive, and negative subset of K, respectively. For two arbitrary sets A and B, let $A \cup B$ denote the *union*, $A \cap B$ the *intersection*, and $A \triangle B := (A \cup B) \setminus (A \cap B)$ the *symmetric difference* of A and B. The *power set* of A, i. e., the set of all subsets of A, is denoted by 2^A. For a function $x : A \to \mathbb{R}$, we define the notation $x(A) := \sum_{a \in A} x(a)$. A function $f : A \to \mathbb{R}$ is called *submodular* if $f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$ holds for all $S, T \subseteq A$. A function $f : A \times A \to \mathbb{R}_{\geq 0}$ is called *metric* if the following three conditions hold for all $a, b, c \in A$: (i) $f(a, a) = 0$, (ii) $f(a, b) = f(b, a)$, and $f(a, b) \leq f(a, c) + f(c, b)$.

For $x \in \mathbb{R} \setminus \mathbb{Z}$, the largest integer number smaller than x is denoted by $\lfloor x \rfloor$. Analogously, the smallest integer number larger than x is denoted by $\lceil x \rceil$. For $x \in \mathbb{Z}$, we define $\lfloor x \rfloor = \lceil x \rceil = x$. Further, we define $\operatorname{frac}(x) := x - (\lceil x \rceil - 1)$ as the *fractional part* of $x \in \mathbb{R}$. Note, $\operatorname{frac}(x) = 1$ if $x \in \mathbb{Z}$.

Let $v \in \mathbb{R}^n$, $M \in \mathbb{R}^{m \times n}$ be a vector and a matrix, respectively. If not stated differently, all vectors are column vectors and $v^\top$ denotes the transposed vector of v. Further, v_i is the i-th component or entry of v. Analogously, $M_{i.}$ is the i-th row, $M_{.j}$ the j-th column, and M_{ij} the entry in row i and column j of M. Let e_i denote the i-th *unit vector*, i. e., the vector whose i-th entry is 1 and all others are 0. Let $x_1, x_2, \ldots, x_t \in \mathbb{R}^n$. A vector $x \in \mathbb{R}^n$ is a *linear (affine, conic, or convex) combination* of $x_1, \ldots, x_t$ if there exists a $\lambda \in \mathbb{R}^n$ so that $x = \sum_{i=1}^t \lambda_i x_i$ (and $\sum_{i=1}^t \lambda_i = 1$, $\lambda \geq 0$, or $\sum_{i=1}^t \lambda_i = 1$ and $\lambda \geq 0$, respectively). Considering one type of combination, if $\lambda = 0$ is the only solution, we call x *linearly, affinely, conic, or convex independent* of $x_1, \ldots, x_t$, respectively. Let $X \subseteq \mathbb{R}^n$. Then $\operatorname{lin}(X)$, $\operatorname{aff}(X)$, $\operatorname{cone}(X)$, or $\operatorname{conv}(X)$, denotes the *linear, affine, conic, or convex hull*, i. e., the set of all linear, affine, conic, or convex combinations of vectors in X, respectively. The *dimension* $\dim(X)$ is defined as the maximum number of affinely independent vectors in X minus 1.

Let E be a finite set, and $\mathcal{I} \subset 2^E$. Then, the pair $(E, \mathcal{I})$ is called an *independence system* if $\emptyset \in \mathcal{I}$ and $A \in \mathcal{I} \Rightarrow B \in \mathcal{I}$ for all $B \subseteq A \in \mathcal{I}$. The elements of $\mathcal{I}$ are called *independent sets*.

Complexity theory. Next, we provide a rather informal introduction to mathematical optimization and complexity theory. Based on Garey and Johnson [72] we define a *problem* as a question to be answered. Usually the answer depends on some input *parameters*. A *(problem) instance* of a problem is an assignment of values to all its input parameters. If only "yes" and "no" are feasible answers, we call the problem a *decision problem*. An *optimization problem* is a problem whose answer is the minimum or maximum value of a given *objective function*. For each optimization problem there exists a corresponding decision problem asking if the objective function is less than or greater than a given value (depending on if the optimization problem is a minimization or maximization problem).

An *algorithm* is a procedure which answers the problem for all problem instances. It *solves* the problem if the answers given are always feasible to the problem.

In this thesis, we assume all problem instances are coded binary and the model of computation is a deterministic one-tape Turing machine, see Garey and Johnson [72] for details. The *(time) complexity* of an algorithm is the number of elementary operations executed to solve a problem instance. The big-O notation is used to denote the complexity of an algorithm or problem depending on the size n of its input data. Then $\mathcal{O}(f(n))$ means there exist functions $f, g : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ with $\lim_{n\to\infty} \frac{f(n)}{g(n)} > 1$ and a scalar $a \in \mathbb{R}_{\geq 0}$ such that (there exists an algorithm such that) the number of elementary operations to solve the given instance of size n is bounded by the term $a \cdot f(n) + g(n)$. Notice that the scalar a and the dominated term $g(n)$ are usually dropped in the corresponding big-O notation. If $f(n)$ is a polynomial, we say the algorithm or problem has polynomial (time) complexity. In addition to the time complexity, we can also count the elementary read and write operations accessing the information storage. Analogously, this yields the so-called *memory* or *size complexity*.

We define $\mathcal{P}$ as the class of all decision problems for which a polynomial time algorithm exists. The class $\mathcal{NP}$ consists of all decision problems for which a "yes"-instance can be verified in polynomial time by another algorithm. It holds $\mathcal{P} \subseteq \mathcal{NP}$. The class *$\mathcal{NP}$-hard* consists of all problems as hard as the hardest problems in $\mathcal{NP}$. Although the misleading name, a $\mathcal{NP}$-hard problem may not be in $\mathcal{NP}$. If a problem is both in $\mathcal{NP}$ and $\mathcal{NP}$-hard, then it is called *$\mathcal{NP}$-complete*. We define the class *co-$\mathcal{NP}$* as all decision problems for which a "no"-instance can be verified with polynomial complexity. It holds $\mathcal{P} \subseteq$ co-$\mathcal{NP}$. Analogously, we define the complexity classes co-$\mathcal{NP}$-hard and co-$\mathcal{NP}$-complete. An algorithm has *pseudo-polynomial* complexity if its complexity is polynomial w.r.t. numeric value of the input and not its binary encoding. A $\mathcal{NP}$-hard problem with pseudo-polynomial complexity is called *weakly $\mathcal{NP}$-hard*. A problem in $\mathcal{NP}$-hard is called *strongly $\mathcal{NP}$-hard* if it is proven that no pseudo-polynomial algorithm solving this problem exists (unless $\mathcal{P} = \mathcal{NP}$). There exist polynomial time algorithms to $\mathcal{NP}$-hard optimization problems if and only if $\mathcal{P} = \mathcal{NP}$ holds which is one famous open question in complexity theory; cf. Cook [56]. An extensive list of classical combinatorial problems known to be $\mathcal{NP}$-complete is given by Garey and Johnson [72]. The definition of (fully) polynomial approximation schemes can also be found therein.

Graph theory. An (undirected) *graph* is defined by a set of *nodes* V, a set of *edges* $E \subset V \times V$ and an incidence function $\psi : E \to V^2$ relating each edge $e = \{i, j\} \in E$ to its *end nodes* $i, j \in V$. A graph is denoted by $G = (V, E, \psi)$ or $G = (V, E)$ for short (in which case we assume that the omitted incidence function ψ is implicitly defined by the set of edges). Let $G = (V, E)$ be a graph with n nodes, $U \subset V$, and $F \subset E$. Then, $G' = (U, F)$ is the *subgraph* of G with node set U and edge set F. A subgraph $G' = (U, F)$ is called a *tree* if it is connected and $|F| = |U| - 1$; if $U = V$, then G' is called a *spanning tree* of G. The subset of edge $P = (v_1v_2, v_2v_3, \dots, v_{i-1}v_i) \subset E$ is called a *path* if $v_j \neq v_k$ for all $j, k \in \{1, \dots, i\}$, $j \neq k$. A *cycle* C is defined as $C := P \cup \{v_iv_1\} \subset E$. A subset $S \subset V$ of the nodes partitions the graph two parts S and $V \setminus S$ and is called *cut*. The subset of edges with one end node in S and the other in $V \setminus S$ is denoted by $\delta(S)$ and called *cutset*.

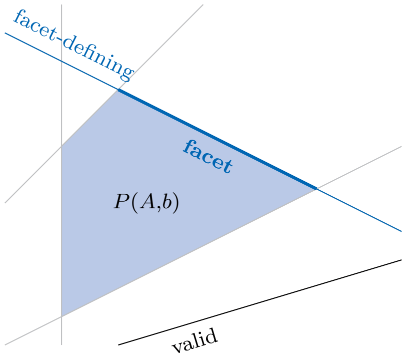


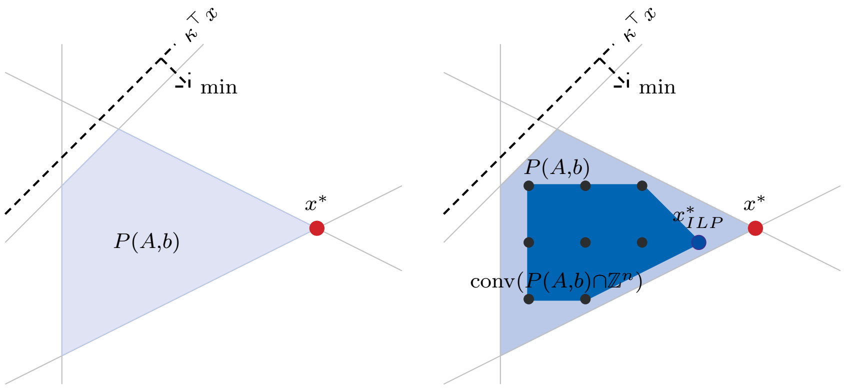
Figure 1.1: Example of a polytope $P(A, b)$ with a valid inequality, a facet-defining inequality, and a facet.

In addition to undirected graphs, there exist *directed graphs* or *digraphs* for short. A digraph D is analogously defined as the triple (V, A, ψ) of a set of nodes V, a set A of directed *arcs*, and an incidence function ψ. An undirected graph $G = (V, E)$ can be *directed* by an orientation $o : E \to V \times V$ assigning each edge $\{i, j\} \in E$ to either the arc (i, j) or the arc (j, i).

Notice, for simplicity, we also write ij for an edge or an arc if it is unambiguous in the current context.

Polyhedral theory. Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $J = \{1, 2, \dots, m\}$, and $Ax \leq b$ a system of linear inequalities. Then, the set $P(A, b) := \{x \in \mathbb{R}^n \ : \ Ax \leq b\}$ is called a *(convex) polyhedron*. W.l.o.g. we assume $P(a, b)$ to be full-dimensional in the following definitions. If $P(A, b)$ is bounded, it is called a *polytope*. The convex hull of all integer lattice points of a polyhedron, $\operatorname{conv}(P(A, b) \cap \mathbb{Z}^n)$, is called *integer hull* and a polyhedron itself. A vector $x \in P(A, b)$ is called an *extreme point* if it is not a convex combination of any vectors in $P(A, b)$. A set F is called a *face* of $P(A, b)$ if $F := \{x \in P(A, b) \ : \ \exists J' \subset J, A_{J'}.x = b_{J'}\}$ holds. A face $F \notin \{P(A, b), \emptyset\}$ is called *proper*. A proper face F of $P(A, b)$ which is not a subset of another face is called *facet*, i. e., it holds $\dim(F) = \dim(P(A, b)) - 1$. For $v \in \mathbb{R}^n$ and $w \in \mathbb{R}$, we call an inequality $v^\top x \leq w$ *valid* for $P(A, b)$ if $P(A, b) \cap P(v, w) = P(A, b)$ holds. A valid inequality is called *facet-defining* for $P(A, b)$ if there exist a facet F of $P(A, b)$ so that $F \subseteq \left\{x \in \mathbb{R}^n \ : \ v^\top x = w\right\} \neq \emptyset$. Figure 1.1 illustrates a polytope, a valid inequality, a facet-defining inequality, and a facet.

Let $Q(A, b) := \left\{x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2} \ : \ A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq b\right\}$. Then we define its *projection* onto the space of x_1 by $\operatorname{proj}_{x_1} Q(A, b) := \left\{x_1 \in \mathbb{R}^{n_1} \ : \ \exists x_2 \in \mathbb{R}^{n_2} \text{ so that } A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq b\right\}$. Let $a_1 \in \mathbb{R}^{n_1}$, $b \in \mathbb{R}$, and ${a_1}^\top x_1 \leq b$ for $x_1 \in \mathbb{R}^{n_1}$. Then, we call $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^\top \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq b$ for $x_2 \in \mathbb{R}^{n_2}$ the *lifted inequality* to the space of the x_1- and x_2-variables. For $x \geq 0$, an inequality $v^\top x \leq w$ *dominates* another inequality ${v'}^\top x \leq w^\top$ if there exists a $\lambda \in \mathbb{R}_{>0}$ so that $\lambda v' \leq v$ and $\lambda w' \geq w$ holds. The case $x \leq 0$ is analog.



(a) LP (b) LP with integer hull as ILP

Figure 1.2: Example of an LP and the ILP defined by the integer hull $\text{conv}(P(A,b) \cap \mathbb{Z}^n)$. The optimal LP solution x^* and integer solution x^*_{ILP} are shown.

Linear programming. Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\kappa \in \mathbb{R}^n$. Then, we call the optimization problem to maximize a linear function over the polyhedron $P(A,b)$ a *linear programming (LP) problem* (in standard form). It can be written as

$$\max \kappa^\top x \tag{1.1}$$
$$\text{s. t.} Ax \leq b. \tag{1.2}$$

A vector $x \in \mathbb{R}^n$ satisfying the conditions of LP (1.1) is called *feasible*; a feasible vector x^* minimizing the objective value $\kappa^\top x^*$ is called *optimal*. If there exists an optimal solution, then there exists an optimal solution which is an extreme point of $P(A,b)$. In Figure 1.2(a), an LP and its optimal solution x^* is visualized.

Given an LP in standard form (1.1), we call the associated LP

$$\min b^\top y \tag{1.3}$$
$$\text{s. t.} A^\top y = \kappa \tag{1.4}$$
$$y \geq 0 \tag{1.5}$$

its *dual LP*. The original LP (1.1) is called the *primal LP*. Note, the dual LP of the dual LP (1.3) is again the primal LP (1.1).

Theorem 1.1 (Duality of linear programming)**.** *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\kappa \in \mathbb{R}^n$. If there exist feasible solutions $\tilde{x}$ and $\tilde{y}$ of the primal LP* $\max \left\{\kappa^\top x \ : \ Ax \leq b\right\}$ *and the dual LP* $\min \left\{b^\top y \ : \ A^\top y = \kappa,\ y \geq 0\right\}$*, respectively, then there exist finite optimal solutions x^* and y^* so that $\kappa^\top x^* = b^\top y^*$ holds.*

Notice that the standard form of an LP is no restriction of generality: there exist transformations between maximization and minimization problems, equality and inequality constraints, and unbounded and bounded/nonnegative variables.

In 1951, Dantzig [59] developed the *simplex algorithm*, an iterative algorithm which starts with a feasible solution (corresponding to an extreme point of the polytope) and improves to another feasible solution (also corresponding to an extreme point) which yields a better objective value until an optimal solution is reached. Unfortunately, the simplex algorithm has exponential worst-case complexity since exponentially-many extreme points have to be evaluated in the worst-case (using known pivot rules); cf. Klee and Minty [95]. However, later it has been proven that its complexity is polynomial on average; cf. Borgwardt [43] and Spielman and Teng [148]. Another approach to is the ellipsoid method which has polynomial complexity but is useless in practice due to numerical instabilities; cf. Khachiyan [93] and Grötschel et al. [75]. A different polynomial algorithm to solve an LP has been introduced by Karmarkar [90] and is called *interior point method* or *barrier algorithm*.

(Mixed) integer linear programming. Given an LP in standard form (1.1) and with rational input data. If we restrict its feasible solutions to integer vectors only, we obtain an *integer linear programming (ILP) problem*. It reads

$$\max \kappa^\top x \tag{1.6}$$
$$\text{s.t.} Ax \leq b. \tag{1.7}$$
$$x \in \mathbb{Z}^n. \tag{1.8}$$

By relaxing the integrality constraint $x \in \mathbb{Z}^n$, we obtain the *linear (programming) relaxation* of ILP (1.6). In Figure 1.2(b), an LP and its integer hull are visualized. In addition, the optimal LP solution x^* and integer solution x^*_{ILP} are shown. Note, if only a subset of the variables is restricted to integrality, a *mixed integer linear program* (MILP) is obtained. The following algorithmic approaches to ILPs are w. l. o. g. also applicable to MILPs. In contrast to LPs, solving ILPs is known to be strongly $\mathcal{NP}$-hard; see Kannan and Monma [88].

In the 1960s, Land and Doig [108] and Dakin [57] introduced the *branch-and-bound* algorithm to solve ILPs. It is an (implicitly) enumerative algorithm following the divide-and-conquer principle used in computer science.

In this algorithm, only the LP relaxation of an ILP is solved, e. g., by using the simplex algorithm. If the LP relaxation is unbounded or has no solution, the ILP is as well or has not one either, respectively. If the LP relaxation has an integer optimal solution, then this solution is also optimal for the original ILP. If the optimal solution vector x^* of the LP relaxation has a fractional valued entry x^*_i, it is not feasible for the ILP. This fractionality is removed by splitting the LP relaxation into two new subproblems where the constraint $x_i \leq \lfloor x^*_i \rfloor$ is added to one of them and the constraint $x_i \geq \lceil x^*_i \rceil$ to the other. This split step is called *branching*. Notice, the union of the set of feasible solutions of both newly created subproblems contains all feasible integer points of the solution set of the original ILP. Figure 1.3(a) illustrates the branching geometrically. After branching, each subproblem is solved individually. If the solution vector of a subproblem has again fractional entries, the algorithm is recursively repeated for this

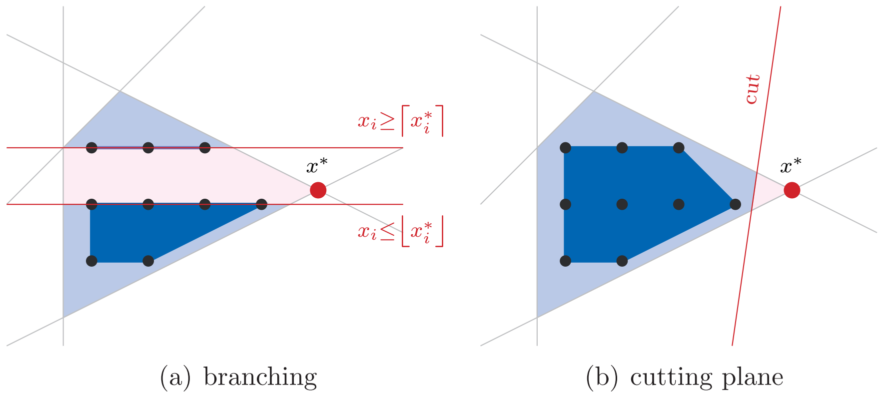


Figure 1.3: Example of branching and cutting in a branch-and-bound or branch-and-cut algorithm, respectively.

subproblem. By succeeding branchings, the so-called (binary) *branch-and-bound tree* is built. Note, the LP relaxation of the original problem is located at the root node of this tree. Its leaf nodes correspond to subproblems which are either infeasible (i.e. have no solution) or have an integer optimal solution.

Clearly, the branch-and-bound tree grows exponentially. Hence, the explicit enumeration of all tree nodes should be avoided. A way to do so is motivated by the following observations: on the one hand, every integer feasible solution found at any node of the branch-and-bound tree is feasible to the original ILP. Thus, its objective value yields an global lower bound (for a maximization problem) on the actual possible currently unknown objective value of the ILP. On the other hand, at each node, the objective value of an optimal solution of the LP relaxation is a local upper bound on the objective value of an optimal solution of the corresponding ILP at the same node. Hence, whenever the local upper bound is lower than the currently best known global upper bound while running the branch-and-bound algorithm, the node corresponding to the local upper bound cannot yield any better integer solution than the one corresponding to the current global upper bound. Thus, this node and its subproblem can be ignore. No branching takes places. It can be removed from the branch-and-bound tree; the node and its potential subtree are called *fathomed*. This overall principle is known as *bounding*.

Another approach to reduce the number of actual solved branch-and-bound nodes is to tighten the LP relaxations by adding additional inequalities which are valid for the ILP but violated for the actual fractional solution of the LP relaxation. Geometrically, these inequalities cut-off some region of the polyhedron associated with the LP relaxation including the fractional LP solution (but no integer feasible solution of the ILP). Hence, these inequalities are called *cutting planes* or *cuts* for short. Figure 1.3(b) shows an example of a cut. The problem to determine a cut given a fractional LP solution or proof that none exist, is called *separation problem*, the procedure itself we call *separation*. In 1981, Grötschel et al. [75] have shown that optimization and separation are polynomially equivalent, i. e., an optimization problem can be solved efficiently if and only if the

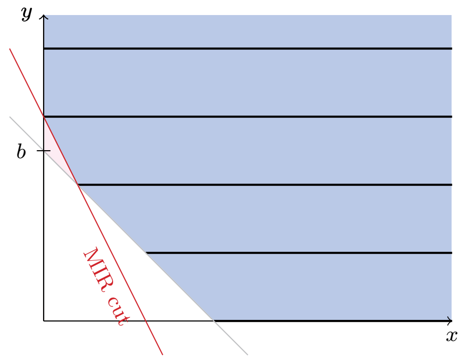


Figure 1.4: Example of a MIR cut in two dimensions.

separation problem to remove infeasible solutions can be solved efficiently. Numerous classes of cutting planes are known. Some of then are general cutting planes exploiting the structure of the coefficient matrix A, the right hand side vector b, or the variable ranges. Others are problem-specific cuts which are only valid for the specific polyhedral structure of selected combinatorial problems. Chvátal-Gomory cuts [73, 74] or mixed integer rounding (MIR) cuts [126] (for MILPs) are examples for the first type of cuts, cover inequalities to the knapsack problem for the latter [18, 80, 158].

We will use MIR to derive valid inequalities for several polyhedra in this thesis. Let us now consider the two-dimensional MIR inequalities as an example.

Lemma 1.2 (Wolsey [159]). *Let* $Q = \{(x, y) \in \mathbb{R} \times \mathbb{Z} \,:\, x + y \geq b,\ x \geq 0\}$. *Then, the* mixed integer rounding inequality

$$x + ry \geq r \lceil b \rceil \tag{1.9}$$

with $r := b - \lfloor b \rfloor$ *is valid for* Q.

In Figure 1.1 a mixed integer set and MIR cut is shown.

By integrating the separation of cutting planes into the branch-and-bound algorithm, we obtain the so-called *branch-and-cut* algorithm. Here, cuts are separated, added to the LP relaxation, and the extended LP is resolved. A problem is only branched into two subproblems if no violated cut has been found or another technically motivated abort criterion for the implemented separation algorithm is met. If the cutting planes are only applied at the root node, the resulting algorithm is called *cut-and-branch*. Achterberg [6] presents an excellent survey on the state-of-the-art algorithms and implementations of branch-and-cut algorithms.

1.2 The knapsack problem

One of the most fundamental problems in mathematical optimization is the well-known knapsack problem. In its general form the (binary or 0-1) knapsack problem asks to select a subset of valuable items such that their total value is maximized while they have to "fit into" the knapsack, i. e., their total weight must not exceed the given capacity of the knapsack. Although its apparent clearness, the knapsack problem turns out to be a hard problem, in fact weakly co-$\mathcal{NP}$-hard. Nevertheless, it occurs in the mathematical models of many applications. Oftentimes, it is a subproblem or relaxation of more complex real-world problems, e. g., (i) in telecommunications traffic has to be routed within the capacity of cables or bandwidth restrictions of base station antennas, (ii) in logistics the capacity restrictions of trucks, planes, ships have to be met, and (iii) in finance the costs of taken decisions must be within a given budget.

Furthermore, each individual constraint of a general 0-1 integer linear program (0-1 ILP) can be considered as knapsack constraint. Therefore inequalities for the knapsack polytope can be used as general cutting planes to 0-1 ILPs. In fact, many results from the 1950/60s on the polyhedral structure of the knapsack polytope were obtained when considering individual rows of 0-1 ILPs; cf. Martello and Toth [119].

The knapsack problem, its variants, and extensions have been studied for several decades. For example, Karp [91] has investigated the complexity of the knapsack problem showing its $\mathcal{NP}$-hardness. Kolesar [98] and Horowitz and Sahni [82] have considered branch-and-bound approaches to solve the knapsack problem exactly. At the same time, a polynomial time approximation scheme has been presented by Johnson [87]. One year later, a fully polynomial time approximation scheme has been published by Ibarra and Kim [83]. Whereas Salkin and De Kluyver [142] have studied the relation between ILPs and knapsack problems. Dudzinski and Walukiewicz [62] have studied LP and Lagrangian relaxations of the problem obtaining lower/dual bounds. In 1979 Martello and Toth [115] have presented an exact exponential algorithm to solve the binary knapsack problem. An algorithmic survey including dynamic programming approaches is given in Martello and Toth [118] and the books by Martello and Toth [119] and Kellerer et al. [92].

Besides its simplest form, the binary knapsack problem, many variants and extensions of the knapsack problem exist. A very famous one is the *subset-sum problem* where the item values are identical to the item weights, see Karp [91], Martello and Toth [119]. It has applications in complexity theory, cryptography and computer science. The *(un)bounded knapsack problem* allows items to be selected more than once up to an (optional) upper bound, see Martello and Toth [119]. The *multiple knapsack problem* and *multiple-choice knapsack problem* group the items such that at most one item per group may be selected, see Martello and Toth [117], Sinha and Zoltners [146]. In *multi-dimensional knapsack problems*, the selected items have to "fit" into several knapsacks at the same time while the weight of an item may differ between these knapsacks, see Weingartner and Ness [156]. *Multi-objective knapsack problems* and *min-max knapsack problems* attach several values to an item and consider a multi-objective approach to determine the combined total value of items, see Ehrgott [67]. Furthermore, variants exists where items have to be selected with minimum value and a total weight above a certain threshold (*minimum knapsack problem*), or where the total weight of the items to be selected is given (*equality knapsack problem*; if additionally all item values are the same: *change-making problem*, see Martello and Toth [116]). Sometimes certain items have to be selected before other items, leading to *precedence constraint knapsack problems*, see Boyd [45]. In addition, more non-linear variants exists, e. g., the *quadratic knapsack problem* where pairs of items are valued, see Hochbaum [81], Mathur et al. [120], Pisinger [138].

For an extensive survey and detailed description of the knapsack problem, its variants and extensions, we recommend the excellent books by Martello and Toth [119] and by Kellerer et al. [92].

In the following, we focus on the (binary or 0-1) knapsack problem and the main results on its computability and the polyhedral structure of the corresponding polytope.

Definition 1.3 (Binary or 0-1 Knapsack Problem)**.** Given a set of items N, a knapsack capacity $c \in \mathbb{Z}_{>0}$, a weight function $w : N \to \mathbb{Z}_{>0}$ with $w_j \leq c$ for all $j \in N$, and a profit function $p : N \to \mathbb{Z}_{>0}$. The *(binary or 0-1) knapsack problem* is to find a subset of items with maximum total profit whose total weight does not exceed the knapsack capacity.

The ILP formulation of the knapsack problem reads

$$\max \sum_{j \in N} p_j x_j \tag{1.10a}$$

$$\text{s.t.} \sum_{j \in N} w_j x_j \leq c \tag{1.10b}$$

$$x_j \in \{0, 1\} \qquad \forall j \in N. \tag{1.10c}$$

where $x_j = 1$ if item $j \in N$ is selected, and 0 otherwise.

Theorem 1.4 (Karp [91])**.** *The knapsack problem is $\mathcal{NP}$-hard.*

Considering the decision problem of the knapsack problem, Karp [91] gives a proof of Theorem 1.4 by polynomial reduction of the $\mathcal{NP}$-complete PARTITION decision problem [72] to the subset-sum decision problem which itself is a special case of the knapsack decision problem.

For constant profit functions, the KP can be solved polynomially as the following example illustrates.

Example 1.5. *Consider the knapsack problem with six items given by* $N = \{1, \ldots, 6\}$, $c = 10$, $w = \begin{pmatrix} 2 & 2 & 3 & 4 & 6 & 7 \end{pmatrix}$, *and* $p = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$. *An optimal solution is the subset* $\{1, 2, 4\} \subseteq N$ *with an optimal solution value (i. e., total profit) of* $1 + 1 + 1 = 3$ *and a total weight of* $2 + 2 + 4 = 8 \leq c$.

Although no efficient algorithm to solve the knapsack problem exists, unless $\mathcal{P} = \mathcal{NP}$, the knapsack problem can be solved in pseudo-polynomial time (in the number of items $|N|$ and capacity c) by the following dynamic program:

Let $\phi : N \to \{1, \ldots, |N|\}$ be an ordering of the item set. Define $f_\phi(j, u)$ as the optimal solution value of the knapsack with the item set $\{\phi(1), \ldots, \phi(j)\} \subseteq N$ and integer capacity $0 \leq u \leq c$. Define $f_\phi(j, u) = -\infty$ where no feasible solution exists. Then the optimal solution value $f_\phi(|N|, c)$ of the knapsack problem can be determined by the following dynamic program

$$f_\phi(1,u) = \begin{cases} 0 & \text{if } 0 \le u \le w_{\phi(1)} - 1 \\ p_{\phi(1)} & \text{if } w_{\phi(1)} \le u \le c \end{cases} \tag{1.11a}$$

$$f_\phi(j,u) = \max\{f_\phi(j-1,u), f_\phi(j-1,u-w_{\phi(j)}) + p_{\phi(j)}\} \qquad \forall\, j \ge 2. \tag{1.11b}$$

with complexity $\mathcal{O}(|N|\,c)$, see also [119] for details.

In addition to this exact algorithm, numerous heuristic and/or approximative algorithms exist yielding lower bounds on the optimal objective value. In this thesis, we focus on exact algorithms and refer to the literature mentioned above for more details on heuristics, approximation algorithms, and (fully) polynomial approximation schemes for the knapsack problem.

The knapsack problem occurs as subproblem in more complex problems. Oftentimes these problems are solved using ILP techniques exploiting the polyhedral problem structure. Hence the polyhedral aspects of the set of feasible solutions of the knapsack problem reoccurs in these problems. In the following, we report on important polyhedral insights of the knapsack problem. Therefore, we define the knapsack polytope as the convex hull of the set of feasible solutions to ILP (1.10); more formally:

Definition 1.6 (Knapsack Polytope)**.** The *knapsack polytope* is defined as

$$\mathcal{K} := \text{conv}\left\{x \in \{0,1\}^{|N|} \;:\; x \text{ satisfies (1.10b)}\right\}.$$

Lemma 1.7 (Balas [18])**.** *The knapsack polytope is full-dimensional, i. e.,*

$$\dim(\mathcal{K}) = |N|.$$

Proof. The zero vector and all unit vectors are feasible to $\mathcal{K}$ and define $|N|+1$ affinely independent vectors. □

Lemma 1.8 (Balas [18])**.** *Trivial facets of the knapsack polytope $\mathcal{K}$ are given by the inequalities*

$$x_j \ge 0 \qquad \forall j \in N \tag{1.12a}$$

$$x_j \le 1 \qquad \forall j \in N : w_j + \max_{i \in N \setminus \{j\}} w_i \le c. \tag{1.12b}$$

Next, we consider certain subsets of items to derive further valid or facet-defining inequalities for $\mathcal{K}$. Let $\mathcal{C} \subseteq N$ be a subset of items whose total weight exceeds the knapsack capacity, i. e., $\sum_{j\in\mathcal{C}} w_j > c$, then $\mathcal{C}$ is called a *cover*. A cover is called *minimal* if $\mathcal{C} \setminus \{j\}$ is not a cover for all $j \in \mathcal{C}$. Given a cover $\mathcal{C}$, the set $E(\mathcal{C}) := \{j \in N \setminus \mathcal{C} \;:\; w_j \ge w_i \;\forall\, i \in \mathcal{C}\} \cup \mathcal{C}$ is called the *extension* of $\mathcal{C}$ and is a cover itself.

Example 1.9. *The set $\mathcal{C} = \{2,4,5\}$ with total weight $2+4+6 = 12 > c$ is a cover for the knapsack defined in Example 1.5. Furthermore, it is minimal. The corresponding extension is $E(\mathcal{C}) = \{2,4,5,6\}$.*

Note that any feasible solution to the knapsack problem cannot contain a subset of selected items forming a cover. Based on this observation, the following valid inequalities for the knapsack polytope can be derived.

Lemma 1.10 (Balas [18], Hammer et al. [80], Wolsey [158]). *Let* $\mathcal{C} \subseteq N$ *be a cover and* $E(\mathcal{C})$ *its extension. Then, the* cover inequality

$$\sum_{j\in\mathcal{C}} x_j \leq |\mathcal{C}| - 1 \tag{1.13}$$

and the extended cover inequality

$$\sum_{j\in E(\mathcal{C})} x_j \leq |\mathcal{C}| - 1 \tag{1.14}$$

are valid for $\mathcal{K}$*. In addition, the cover inequality is facet-defining for* $\mathcal{K}$ *if* $\mathcal{C} = N$ *and* $\mathcal{C}$ *is minimal.*

In fact, the knapsack polytope can be characterized by the extensions of all minimal covers as follows:

Theorem 1.11 (Balas and Jeroslow [19]). *The point* $x \in \{0,1\}^{|N|}$ *is feasible for* $\mathcal{K}$ *if and only if it satisfies the extended cover inequality* (1.14) *for all minimal covers* $\mathcal{C} \subseteq N$.

The extension of a cover can be interpreted as a simple lifting procedure from the corresponding cover inequality to its extended cover inequality [18, 89, 133]. A more general form of lifted cover inequalities has been introduced by Van Roy and Wolsey [154] using the so-called up-lifting and down-lifting [77, 78].

Lemma 1.12 (Van Roy and Wolsey [154]). *Let* $\mathcal{C} \subseteq N$ *be a cover and* $\mathcal{C}' \subsetneq \mathcal{C}$*. Then lifting coefficients* $\alpha \geq 0$ *and* $\beta \geq 0$ *can be determined such that the* lifted cover inequality

$$\sum_{j\in\mathcal{C}\setminus\mathcal{C}'} x_j + \sum_{j\in N\setminus\mathcal{C}} \alpha_j x_j + \sum_{j\in\mathcal{C}'} \beta_j x_j \leq |\mathcal{C}\setminus\mathcal{C}'| + \sum_{j\in\mathcal{C}'} \beta_j - 1 \tag{1.15}$$

is valid for $\mathcal{K}$.

We refer to [154] for details how to determine the lifting coefficients. Note, lifted cover inequalities (1.15) generalize the class of extended cover inequalities (1.14).

A subset $N' \subseteq N$ and an element $t \in N \setminus N'$ is called a *(1,k)-configuration* with $2 \leq k \leq |N'|$, if $\sum_{j\in N'} w_j \leq c$ and $Q \cup \{t\}$ is a minimal cover for all $Q \subseteq N'$ with $|Q| = k$. For $k = |N'|$, a (1,k)-configuration is a minimal cover.

Example 1.13. *Let* $k = 2$*. The set* $N' = \{2,3,4\}$ *and the element* $t = 5$ *are a (1,k)-configuration for the knapsack defined in Example 1.5.*

The concept of (1,k)-configurations generalizes the minimal covers and yields a new class of inequalities valid for the knapsack polytope.

Theorem 1.14 (Padberg [134]). *Let $N' \subseteq N$ and $t \in N \setminus N'$ be a (1,k)-configuration. Then, the* (1,k)-configuration inequalities

$$(r-k+1)x_t + \sum_{j \in T} x_j \leq r \tag{1.16}$$

where $T \subseteq N'$ with $|T| = r$ and $k \leq r \leq |N'|$ are valid for $\mathcal{K}$. If $N' = N \setminus \{t\}$, these inequalities are facet-defining.

Note, (1,k)-configuration inequalities (1.16) are a special case of up-and-down-lifted cover inequalities (1.15).

Another class of valid inequalities, not based on the concept of covers, can be derived from the so-called pack inequality. A set $\mathcal{P} \subseteq N$ is called a *pack* if $\sum_{j \in \mathcal{P}} w_j \leq c$ holds.

Lemma 1.15 (Martello and Toth [119]). *Let $\mathcal{P} \subseteq N$ be a pack. Then, the* pack inequality

$$\sum_{j \in \mathcal{P}} w_j x_j \leq \sum_{j \in \mathcal{P}} w_j \tag{1.17}$$

is valid for $\mathcal{K}$.

The pack inequality can be derived as sum of the upper bound inequalities $x_j \leq 1$ scaled by w_j for each $j \in \mathcal{P}$. Although it is a relatively weak inequality dominated by the corresponding upper bound inequalities, lifting also yields non-dominated inequalities. A specific subclass of these *lifted pack inequalities* are the so-called weight inequalities, introduced by Weismantel [157].

Lemma 1.16 (Weismantel [157]). *Let $N' \subseteq N$ with $\sum_{j \in N'} w_j < c$ and residual capacity $c_{res} := c - \sum_{j \in N'} w_j$. Then, the* weight inequality

$$\sum_{j \in N'} w_j x_j + \sum_{j \in N \setminus N'} \max\{0, w_j - c_{res}\} x_j \leq \sum_{j \in N'} w_j \tag{1.18}$$

is valid for $\mathcal{K}$.

Most branch-and-cut approaches to solve the knapsack problem include the separation of lifted and/or extended cover inequalities. The separation of (1,k)-configurations or weight inequalities seems not to be that common, cf. [89]. However, the exact separation of violated cover inequalities, extended cover inequalities, lifted cover inequalities has been shown to be $\mathcal{NP}$-hard [68, 78, 94]. Weismantel [157] has presented a pseudo-polynomial algorithm to separate violated weight inequalities. An excellent survey on state-of-the-art separation algorithms for the 0-1 knapsack problem and detailed computational experiments can be found in Kaparis and Letchford [89]. A statistical study on knapsack instances to identify classes of knapsack instances particularly hard to solve is given by Pisinger [137].

1.3 The capacitated network design problem

Another class of well-known mathematical optimization problems are *flow problems*. Here a network is given by its nodes, edges and capacity values assigned to the edges. Furthermore, some nodes are marked as *sources* (or *source nodes*) and others as *sinks* (or *target nodes*). At each source, a certain value of a *commodity* is available. At each sink a certain value of commodity is requested. The flow problem asks to send available commodity from the sources to the sinks such that their requests are satisfied, the individual edge capacities are not exceeded and flow conservation holds at each node which is neither source nor sink. This way a so-called *flow* through the network is established. If you think of water: it flows from the sources through a network of pipelines (edges) with different diameters (edge capacities) to the sinks. Flow problems occur in many real-world application areas wherever some commodities have to be transported; for example in transportation, logistics, public transport, telecommunications, etc. Objective functions are usually to maximize the amount of flow to be send or to satisfy the requested demand while minimizing the utilization of the edges.

Of course, the basic flow problem previously sketched can be enriched by more restrictions to model the underlying real-world problems and their requirements more accurately. For example, oftentimes survivability of network connectivity and the satisfiability of the demands is required in the case that some parts of the network fail, i. e., nodes or edges become unavailable; cf. Stoer and Dahl [150] and Pióro and Medhi [136] for more details on survivability in telecommunication networks. Moreover, usually not only one type of commodity is considered but a set of different commodities and thus commodity dependent sources and sinks. Then a flow must be determined for all commodities simultaneously as each commodity contributes to the usage of the available edge capacity. This is called the *multi-commodity flow problem*.

The *capacitated network design problem* includes the multi-commodity flow problem. In addition, the edge capacities are not given (i. e., fixed and part of the input) anymore. Instead, they have to be determined together with the optimal flow. The objective is to minimize the total costs consisting of the costs of edge capacities and the edge- and volume-dependent costs of sending flow. In most practical settings, the volume-dependent costs of sending flow along an edge is zero or can be neglected with respect to the much higher costs for installing or increasing the capacities on the edges. Furthermore, in real-world applications as telecommunication networks there exists some discretization of available edge capacities. For example, the available bandwidth (capacity) on telecommunication edges depends on the installed line card hardware at the end nodes of this edge. Due to the available technological and commercial hardware products only a bounded number of line cards and only a certain set of available line card types can be installed. Thus, the installable edge capacities can be assumed to be integer (after some normalization). The resulting problem where only integer units of edge capacity can be installed and the total costs are only inflicted by capacity installation is called the *(capacitated) network loading problem.* Oftentimes the (capacitated) network loading problem is also just called the network design problem in literature. In this thesis we follow this popular slight inaccuracy and speak of the network design problem.

There exists a wide range of work on the capacitated network design problem yielding different problem formulations, intensive studies on the polyhedral structure of the set of feasible solutions, classes of valid and facet-defining inequalities strengthening these formulations, as well as separation algorithms and heuristics. This includes for examples the early work by Magnanti et al. [114], Bienstock and Günlük [39], and Günlük [79] on the network design polyhedron and classes of valid inequalities, as well as the work of Avella et al. [17] describing the capacity formulation of the network design problem, and the references therein.

It is beyond the focus of this thesis to give a complete survey of all related results. Instead we refer to Ahuja et al. [10] for a detailed survey on network flow problems, and to Magnanti and Wong [112] and Pióro and Medhi [136] on network design problems including an extensive catalog of variants of this problem. Nevertheless in the following we give a brief survey on the most important results and polyhedral insights of this problem.

Now, let us consider the capacitated network design problem more formally.

Definition 1.17 (Capacitated Network Design Problem)**.** Let $G = (V, E)$ be an undirected graph where V denotes the set of all nodes and E the set of all edges. Further let K be the set of all node-to-node commodities. For each commodity $k \in K$, let s^k and t^k denote its source and target node, respectively, and let $d^k > 0$ denote its requested demand value. Moreover let $\kappa : E \to \mathbb{R}_{\geq 0}$ be a cost function assigning to an edge the installation costs of one unit of capacity on this edge.

Then the *capacitated network design problem (NDP)* asks to determine an installment of edge capacities and a multi-commodity flow such that

1. all demands are satisfied, i. e., for each commodity, there is a flow between its source and target nodes whose flow value is at least the demand value of this commodity,

2. no capacities are exceeded, i. e., for each edge, the total value of all flows along this edge has to be at most the installed capacity on this edge,

3. and the total capacity costs are minimized

Remark 1.18. Although the NDP is defined on an undirected graph, an orientation of the edges is needed to have the notation of ingoing and outgoing edges when modeling the flow. This orientation is omitted when determining the total flow on an edge.

Notice, usually we do not explicitly state the fact that the problem is capacitated and just call it the network design problem.

Chopra et al. have investigated the complexity of the NDP.

Theorem 1.19 (Chopra et al. [52])**.** *The NDP is strongly $\mathcal{NP}$-hard.*

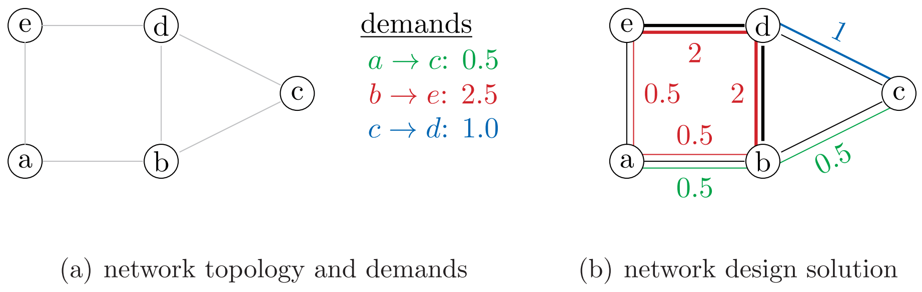


Figure 1.5: Example of network design. Given a 5-node network with three commodities, edge capacities are installed (black) and a multi-commodity flow is determined (green/red/blue) minimizing the installment costs.

Example 1.20. *Let us consider a simple 5-node example of a network design problem with three node-to-node demands. The demand values may be non-integral but the edge capacities have to be integral. Further let the edge capacity installment costs be the same for all edges. Figure 1.5(a) shows the potential network topology, the three commodities and their demand values.*

In Figure 1.5(b), a network design solution is shown for this problem instance: one unit of edge capacity is installed on all thin black edges (ab, ae, bc, cd) and two units are installed on all thick black edges (bd, de). Besides a multi-commodity flow has been determined splitting the red commodity $b \to e$ along two paths such that 2 demand units are send $b \to d \to e$ and 0.5 demand units are send $b \to a \to e$. For the green commodity $a \to c$, a single-path flow $a \to b \to c$ is established. The blue commodity $c \to d$ is satisfied by the direct flow $c \to d$. The total installation costs are minimized since the costs are the same for all edges and the total number of installed capacity units is minimized.

Several mixed integer programming formulations of the NDP exist. In the following we report on three important ones: the edge-flow formulation, the path-flow formulation, and the capacity formulation. The edge-flow formulation of the NDP reads:

$$\min \sum_{e \in E} \kappa_e x_e \tag{1.19a}$$

$$\text{s.t.} \sum_{j\,:\,ij \in E} (f_{ij}^k - f_{ji}^k) = \begin{cases} 1 & i = s^k \\ -1 & i = t^k \\ 0 & \text{else} \end{cases} \qquad \forall i \in V,\ k \in K \tag{1.19b}$$

$$\sum_{k \in K} d^k (f_{ij}^k + f_{ji}^k) \leq x_e, \qquad \forall e \in E \tag{1.19c}$$

$$f, x \geq 0 \tag{1.19d}$$

$$x \in \mathbb{Z}^{|E|} \tag{1.19e}$$

where variables f_{ij}^k denote the fraction of the demand k send along the (orientated) edge $e = \{i, j\} \in E$ from end node i to j, and variables x_e denote the number of capacity units installed on edge e at cost κ_e per unit. The constraints (1.19b) describe a multi-commodity flow using an edge-flow formulation. The flow for commodity k is directed from (its source) s^k to (its target) t^k without loss of generality. Constraints (1.19c) are edge capacity constraints guaranteeing that the flow on an edge does not exceed the installed capacity on this edge.

Since the cost for edge capacity is minimized, we may ignore cycle flows and hence assume that either $f_{ij}^k = 0$ or $f_{ji}^k = 0$ and $f_e^k \leq 1$ in any optimal solution. More precisely, given an optimal solution where a cycle flow exists, there exists another solution with the same objective value but without this cycle flow. Notice that while flow and demands are directed, the actual direction is arbitrary since we sum up the two flows in (1.19c). Moreover by model (1.19) we also see that the NDP includes a multi-commodity flow problem and a knapsack problem with variable knapsack capacity and fractional item selection.

While the previous model is based on flow variables for each directed edge, another flow-based model is often used where the flow is modeled based on paths from the source to the target node of each commodity. This so-called path-flow formulation of the NDP is the following

$$\min \sum_{e \in E} \kappa_e x_e \tag{1.20a}$$

$$\text{s.t.} \sum_{p \in P^k} f_p^k \geq 1 \qquad \forall k \in K \tag{1.20b}$$

$$\sum_{k \in K} d^k \left(\sum_{p \in P^k : e \in p} f_p^k \right) \leq x_e, \qquad \forall e \in E \tag{1.20c}$$

$$f, x \geq 0 \tag{1.20d}$$

$$x \in \mathbb{Z}^{|E|} \tag{1.20e}$$

This formulation includes flow variables $f_p^k \geq 0$ for each commodity $k \in K$ and path $p \in P^k$ where P^k is the set of all possible s^k-t^k-paths. Each variable denotes the fraction of demand of the considered commodity sent along the considered path. Constraint (1.20b) ensures that at least 100 % of the demand of each commodity is routed. It corresponds to the flow conservation constraint (1.19b) of the edge-flow formulation (1.19). The remaining constraints (1.20c)–(1.20e) are also the path-flow equivalents of the corresponding constraints of the edge-flow formulation (1.19).
Note that formulation (1.20) is of exponential size due to the exponential number of possible paths and thus f-variables. But this formulation is still computationally tractable since a column generation approach can be applied where the resulting pricing problem is a shortest-path problem. Note, the multi-commodity flow and knapsack (with fractional item selection) subproblems of the NDP can again be found in this formulation.

W. l. o. g. we assume that there are no cycle flows in any solution. Otherwise, a solution with the same or less objective value exists whose corresponding flow has no cycles and satisfies all demands. Then both, the edge-flow and the path-flow formulation are equivalent formulations of the NDP: given a solution of one of them a corresponding solution of the other can be constructed transferring the flow variables straight-forwardly. For example, given path flow variables f_p^k, the corresponding edge flow variables f_e^k can be determined as $f_e^k = \sum_{p \in P^k : e \in p} f_p^k$. Vice versa, a solution f_e^k for all $e \in E$ defines a multi-flow which can be decomposed into paths yielding corresponding path flow variables f_p^k; cf. Pióro and Medhi [136].

These two flow-based formulations both include the multi-commodity flow problem as subproblem. The feasibility of a (fixed) capacity vector x according to this subproblem can be characterized by the following sufficient and necessary condition.

Theorem 1.21 (Iri [85], Onaga and Kakusho [129]). *Given link capacities* $x \in \mathbb{R}_{\geq 0}^{|E|}$, *there exists a flow satisfying* (1.19b)–(1.19d) *if and only if for all length functions* ℓ *the* length inequality

$$\sum_{e \in E} x_e \ell(e) \geq \sum_{k \in K} \ell(s^k, t^k) d^k \tag{1.21}$$

holds where $\ell(s^k, t^k)$ *is the length of a shortest-path from* s^k *to* t^k.

This theorem can be proven directly by applying Farkas' Lemma. Further, note that Theorem 1.21 is often called the "Japanese Theorem" in literature. Its results can also be obtained by the more general technique known as Benders decomposition applied to the flow-formulation of the NDP; cf. Benders [32]. Then the length inequality (1.21) corresponds to the resulting Benders cut.

Theorem 1.21 gives rise to the capacity formulation of the network design problem:

$$\min \quad \sum_{e \in E} \kappa_e x_e \tag{1.22a}$$

$$s.t. \quad \sum_{e \in E} \ell(e) x_e \geq \sum_{k \in K} \ell(s^k, t^k) d^k \quad \text{for all length functions } \ell \tag{1.22b}$$

$$x \in \mathbb{Z}_{\geq 0}^{|E|} \tag{1.22c}$$

Notice that formulation (1.22) does not include flow-based variables but only the edge capacity variables x. The flow-variables have been projected out in the process of Benders decomposition. The feasibility of x is guaranteed by the Benders cuts, the length inequalities (1.21). Furthermore, inequalities (1.22b) obtained from non-metric length functions are dominated: moreover, Avella et al. [17] shows that all facet defining inequalities are so-called tight metric inequalities, i. e., length inequalities with metric length functions and maximal right hand side value.

For each of the three presented models, the convex hull of all its feasible solutions is a polyhedron, a *network design polyhedron*. In fact, we obtain the following three polyhedra:

Definition 1.22 (Network Design Polyhedra)**.** The *edge-flow (or link-flow) network design polyhedron*

$$\mathcal{N}^{\mathrm{LF}} := \operatorname{conv}\left\{(x, f) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{2|E||K|} \ : \ (x, f) \text{ satisfies } (1.19)\right\},$$

the *path-flow network design polyhedron*

$$\mathcal{N}^{\mathrm{PF}} := \operatorname{conv}\left\{(x, f) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{|P|} \ : \ (x, f) \text{ satisfies } (1.20)\right\},$$

and the *capacity network design polyhedron*

$$\mathcal{N}_x := \operatorname{conv}\left\{x \in \mathbb{Z}_{\geq 0}^{|E|} \ : \ x \text{ satisfies } (1.22)\right\}.$$

By construction, it holds $\mathcal{N}_x = \operatorname{proj}_x \mathcal{N}^{\mathrm{LF}} = \operatorname{proj}_x \mathcal{N}^{\mathrm{PF}}$. Network design polyhedra have been studied intensively: for example by Magnanti et al. [113], Bienstock and Günlük [39], Bienstock et al. [40], and Günlük [79].

Proposition 1.23. *If the underlying graph $G = (V, E)$ is connected, the dimension of $\mathcal{N}^{LF}$ is $2|E|(|K| + 1) - |K|(|V| - 1)$.*

Lemma 1.24 (Mattia [121])**.** *The polyhedron $\mathcal{N}_x$ is full-dimensional, i. e., $\dim(\mathcal{N}_x) = |E|$.*

Oftentimes "easier" (i. e., less-dimensional) polyhedra like relaxations and/or projections of network design polyhedra are considered to study their polyhedral faces. Afterwards the identified classes of valid or facet-defining inequalities are lifted to corresponding classes for the original network design problem itself. A good example of such an approach leads to the well-studied cutset inequalities: Given a subset $S \subsetneq V$ of nodes, the network is partitioned into two parts S and $V \setminus S$. More precisely, S defines a cut and a corresponding cutset $\delta(S)$, i. e., the set of edges with one end node in S and the other in $V \setminus S$. Then for investigation, the node sets S and $V \setminus S$ are shrunken into one node each and thus yielding a 2-node network with only the cutset edges in between. The convex hull of all feasible solutions of NDP on this small network is called the *cutset polyhedron* and has been studied among others by Atamtürk [14], Chopra et al. [52], Magnanti and Mirchandani [111], Magnanti et al. [114], and Raack et al. [141]. The valid class of cutset inequalities is based on the following simple observation: the total capacity installed on all cutset edges must be at least the total demand value of all commodities with sources in S and targets in $V \setminus S$ or vice-versa. If less capacity is installed on the cutset edges, not all demand values of the "cut-crossing" commodities can be satisfied.

Theorem 1.25 (Atamtürk [14], Barahona [20], Magnanti et al. [114])**.** *Let $K_S \subseteq K$ denote the set of cut-crossing commodities, i. e.,*

$$K_S := \left\{k \in K \ : \ (s^k \in S \wedge t^k \in V \setminus S) \ \vee \ (s^k \in V \setminus S \wedge t^k \in S)\right\}.$$

Then, the cutset inequality

$$\sum_{e \in \delta(S))} x_e \geq \left\lceil \sum_{k \in K_S} d^k \right\rceil \tag{1.23}$$

is valid for $\mathcal{N}^{LF}$, $\mathcal{N}^{PF}$ *and* $\mathcal{N}_x$. *It is facet-defining if* $\mathrm{frac}(\sum_{k \in K_S} d^k) < 1$ *and* $S \notin \{\emptyset, V\}$.

Another class of valid inequalities for the NDP are the so-called *arc residual capacity inequalities* which are obtained by studying another projected relaxation of the edge-flow network design polyhedron $\mathcal{N}^{\mathrm{LF}}$: here we apply Lagrangian relaxation to the flow conservation constraints (1.19b). The resulting problem is decomposable into $|E|$ individual problems, one per edge $e \in E$. Each of these problems is called a *single arc design problem* (SADP) and can be formulated as

$$\max \sum_{k \in K} \lambda_k f^k + \kappa x \tag{1.24}$$

$$\text{s.t.} \sum_{k \in K} d^k f \leq x \tag{1.25}$$

$$f^k \in [0,1] \qquad \forall k \in K \tag{1.26}$$

$$x \in \mathbb{Z}_{\geq 0}. \tag{1.27}$$

Notice that we have dropped the edge-dependent subscripts; f^k $(k \in K)$, κ, and x are scalars in this formulation. Its feasible solutions give rise to the *single arc design polyhedron*. It has been studied by Magnanti et al. [113] identifying the class of so-called *arc residual capacity inequalities* which are also valid for the NDP polyhedra.

Theorem 1.26 (Magnanti et al. [113]). *Let* $Q \subseteq K$. *The* arc residual capacity inequality

$$\sum_{k \in Q} d^k f^k \leq r^Q x + (\left\lceil \sum_{k \in Q} d^k \right\rceil - 1)(1 - r^Q)$$

is valid for $\mathcal{N}^{LF}$, $\mathcal{N}^{PF}$ *and* $\mathcal{N}_x$ *and facet-defining if the two following conditions are satisfied: (i)* $\sum_{k \in Q} d^k = 1$ *implies* $|Q| = 1$, *and (ii)* $r^Q = 1$ *implies* $Q = K$.

Furthermore, the single arc design polytope is in fact completely described by the arc residual capacity inequalities; cf. [113].

Besides these two classes of valid inequalities many more are known for the NDP: for example, the concept of cutset inequalities can be generalized to k-node partitions resulting in k-partition inequalities; cf. Agarwal [7], Barahona [20], Magnanti et al. [113], and Agarwal [8]. Also metric inequalities generalize k-partition (and thus 2-partition or cutset) inequalities; cf. Avella et al. [17] and Mattia [121]. Another more general concept is applied by the so-called flow-cover inequalities; cf. Padberg et al. [135] and Atamtürk [13].

1.4 Applications in telecommunication

Network design problems arise in many application areas, for example in logistics, public transport, or telecommunications. In particular, telecommunications depends on the solution of network design problems to determine the topology, (hardware) configuration, and routing of electrical or optical signals through the communication network in a cost-efficient way.

Today's telecommunication networks are rather complex in reality: knowledge covering the fields of physics, information theory, engineering, computer science and mathematics is required to completely understand the structure, functionality and operation of these networks. For the purpose of (strategic) network planning and design using mathematical optimization techniques, we consider a more abstract and less technical view of the real network architectures aggregating and hiding a lot of its complexity in high-level concepts and models. In the following, we will briefly elaborate on this more abstract structure and the arising mathematical challenges. For a detailed survey on and for a classification of optimization problems in telecommunication and their mathematical models, we refer to Pióro and Medhi [136]. More recently, Koster and Munoz [102] give a state-of-the-art survey on mathematical optimization problems in communication networks and algorithmic approaches to solve them.

We consider the topology of fixed line telecommunication networks as directed graphs. It is used to establish node-to-node connections for communication using the network edges, oftentimes called links. These connections are realized via electrical or optical signals and require specific hardware to be installed, set-up, and maintained at the nodes and links of the network.

Example 1.27. *Let us simplify a telecommunication network and consider a single communication demand and its way through the network. Its connection may start at a customer's home computer where an electrical connection is set-up via his landline. Then the electrical signal may be send to a network node where it is aggregated with other signals to facilitate the available bandwidth better. This process of aggregation is called* multiplexing. *Let us assume the multiplexed signal is then send further via several network nodes and probably de-multiplexed and multiplexed with other signals several times. At some point it may reach a node where it is transformed into an optical signal. Then, an optical transponder sends the signal along a so-called* light path *which can be seen as an optical channel passing further network nodes not interfering with the content of the channel. At the end of the light path is the receiver and another hardware transforming the optical signal back into an electrical one. After passing several nodes, de-multiplexing (and maybe intermediate multiplexing) the signal reaches its destination network node and another customer's computer establishing a communication with the first customer.*

This is a simple abstract view on the key aspects of telecommunication networks. Of course, it contains a lot of technical inaccuracies. For example, it does not reflect

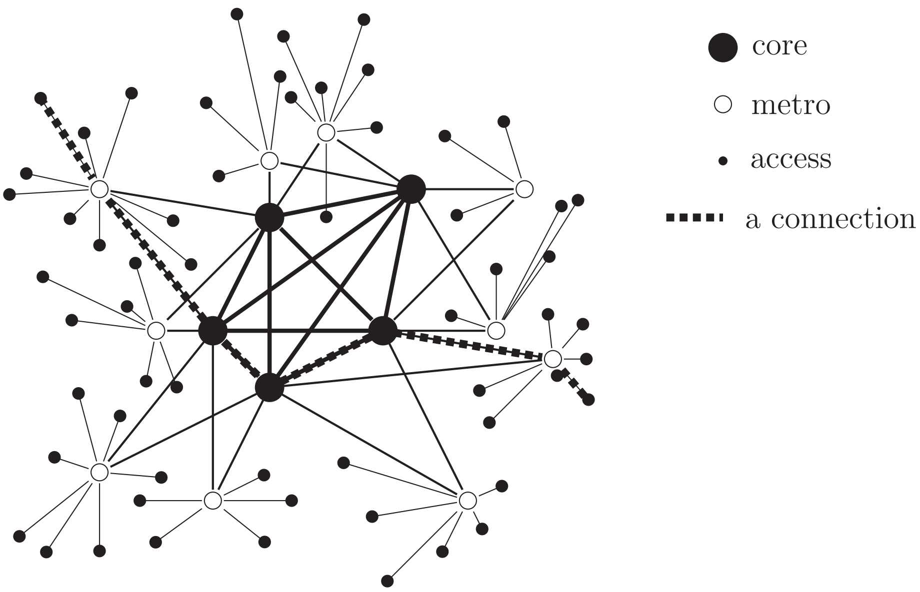


Figure 1.6: Spatial (or horizontal) hierarchy of telecommunication networks: access nodes are single-connected to regional metro nodes which are multiple-connected to nodes of the fully connected core network. The available edge capacity also increases from access to core network. In addition, an arbitrary connection between two access nodes is shown.

protocols, nor physical aspects as the degeneracy of signals and thus the need to amplify them, etc.

The structure of telecommunication networks can be classified by different hierarchical systems: First, there exists a spatial or horizontal hierarchy that distinguishes between the *access*, the *metro*, and the *core network*. The access network is the outer-most network part which connects directly to the customers. Thus the sources and destinations of demands are usually located in or directly connected to the access network. Most of the times, the access network is relatively sparse with less requirements for survivability. The metro network connects to the access network. It is denser than the access network and usually contains some basic survivability functionality. The traffic is further aggregated in the metro network. Geographically, this usually takes place on a regional level. The inner-most network part is the core network. It is connected to the metro network, is oftentimes fully connected (maximum density), and has high requirements for survivability. It forms the backbone of the communication network. Here the aggregated demands of the metro network are further aggregated and sent on long distance connections.

Figure 1.6 shows the horizontal hierarchy. From this point of view, a connection starts at the access network, goes through the metro and the core network, back to another metro network node, and to the access network.

Second, there also exists a technological or vertical hierarchy of telecommunication networks; this is illustrated in Figure 1.7. Here the network is considered as a stack of different so-called *network layers* which lie on top of each other. Each layer represents

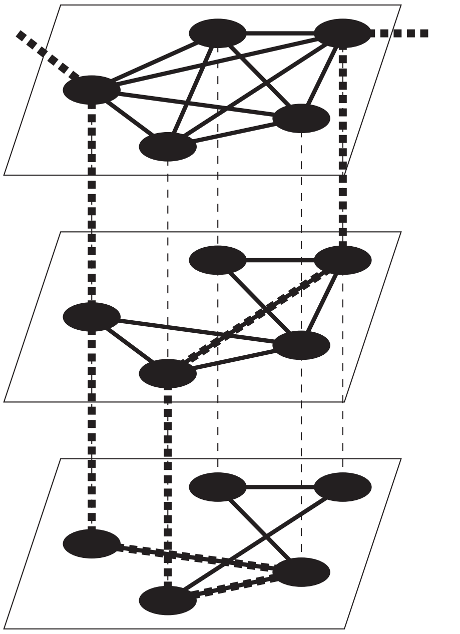


Figure 1.7: Technological (or vertical) hierarchy of telecommunication networks: the physical layer represents the real-world hardware, cables, fibers. On top of this are several logical layers without real-world counterparts. In addition, an arbitrary connection between two nodes is shown.

different technologies or protocols that are used in the network. The lowest layer is the *physical layer* whose nodes correspond to real-world locations and links are real-world cables or *fibers* between these locations. The upper layers usually share copies of the same nodes but not the same links. In general, each higher layer is on a higher aggregation level than the lower, i. e., a link between two nodes in a higher layer is realized by a path between the corresponding copies of these nodes in a lower layer. The resulting topologies of those layers do not have physical counterparts. Because of this they, are called *logical or virtual layers*. For example, a very high logical layer is the *Internet Protocol* (IP) layer which is usually highly connected. The node-to-node demands can also be seen as individual layer (*demand layer*) in which case it is usually the highest layer and each of its logical links corresponds directly to a demand between the end nodes of this link.

The basic network planning and design problem asks to find the cost-efficient hardware installation and configuration such that all demands are satisfied. Therefore, the hardware to be installed at the links (e. g., fibers, amplifiers, regenerators, etc.) and nodes (e. g., transponders, receivers, (de-)multiplexers, switches and routers, wavelength converters, etc.) must be determined as well as a routing of each node-to-node demand using this hardware.

In this thesis we consider only non-survivable single layer core networks. Figure 1.8 shows an example single layer network design instance taken from the SNDlib [130] problem library. It is reasonable to focus our investigations to such networks as our

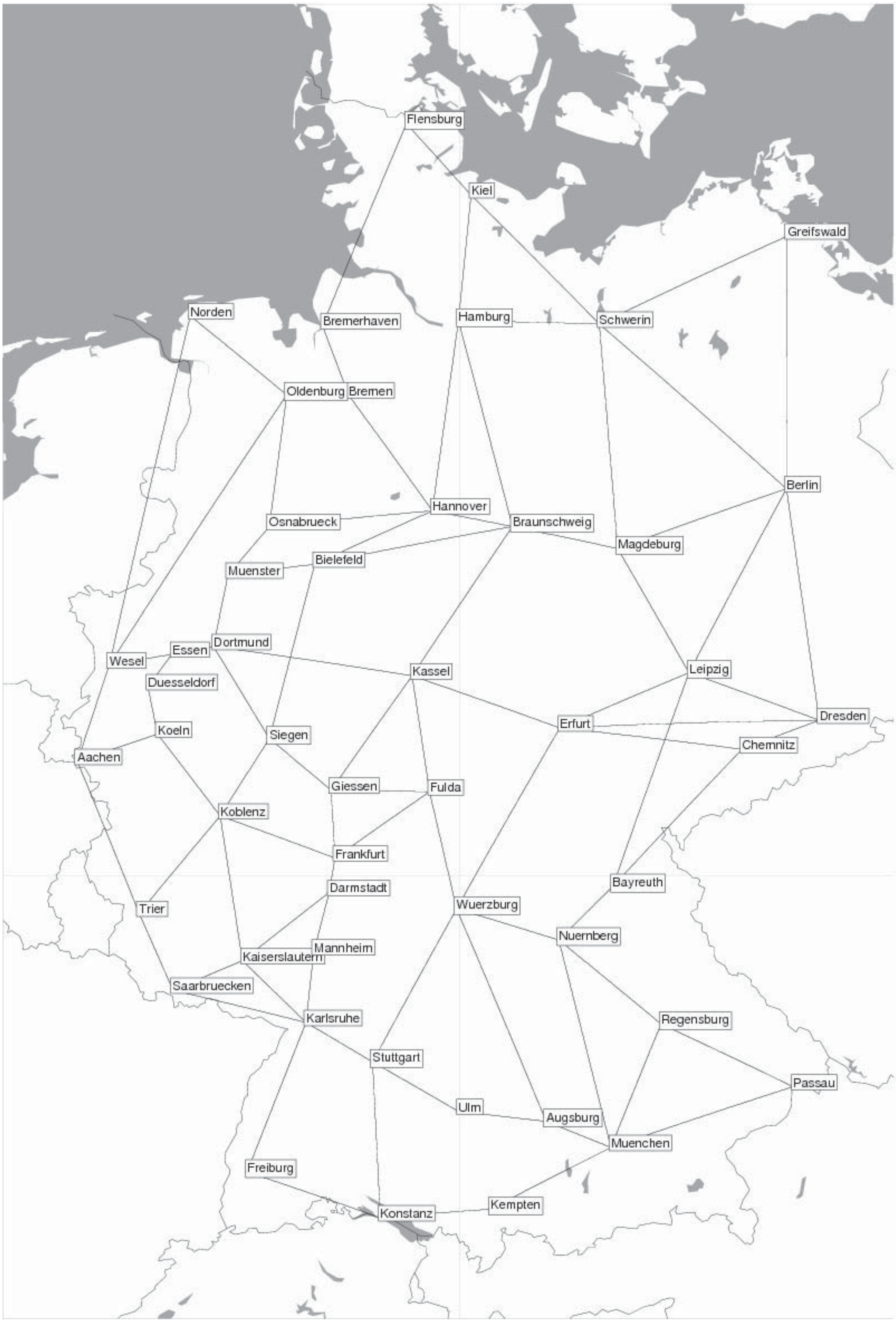
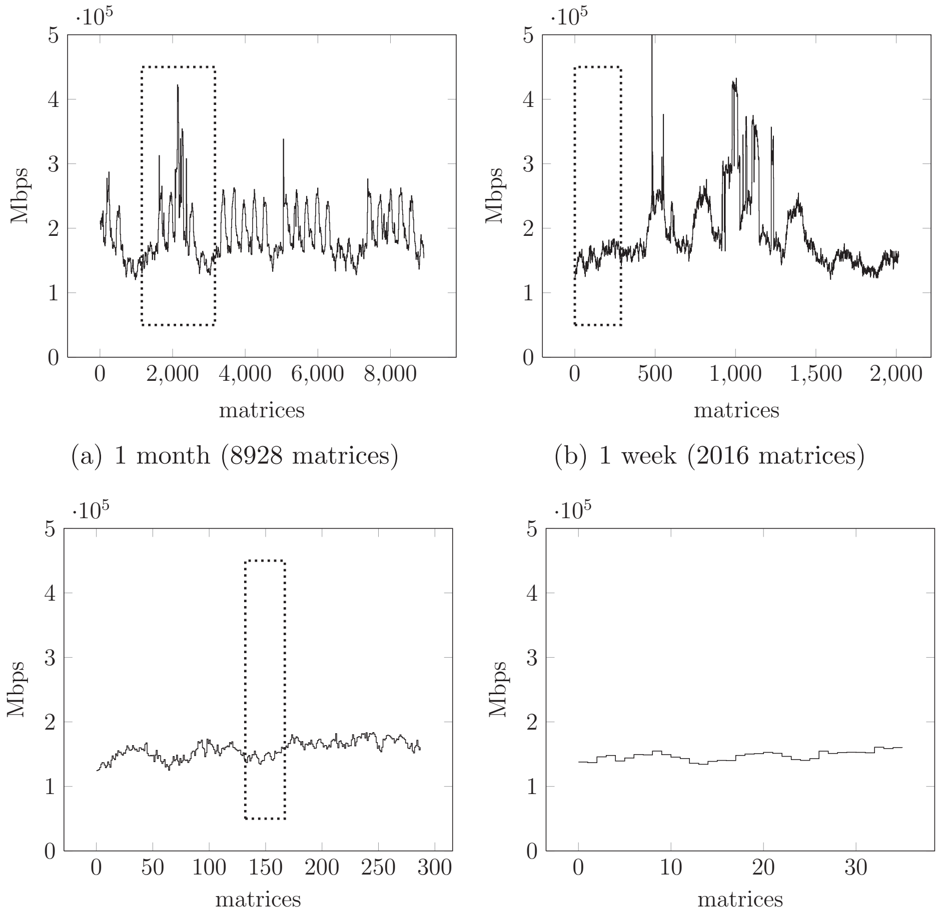


Figure 1.8: Example network design instance. The GERMANY50 instance and this picture are taken from SNDlib [130, 131], a public problem library for the design of survivable networks. This instances was originally provided to SNDlib by T-Systems International AG.



(a) 1 month (8928 matrices)
(b) 1 week (2016 matrices)
(c) 1 day (288 matrices)
(d) 3 hours (36 matrices)

Figure 1.9: Total network load over time for the Abilene network [5]. Different time periods in 2004 are shown: 1 month (July), 1 week (July 5th–11th), 1 day (July 5th), 3 hours (July 5th, 11 am–2 pm). The dotted areas illustrate the areas zoomed-in in the subsequent diagram, i. e., the dotted area in Figure 1.9(a) is shown in detail in Figure 1.9(b)

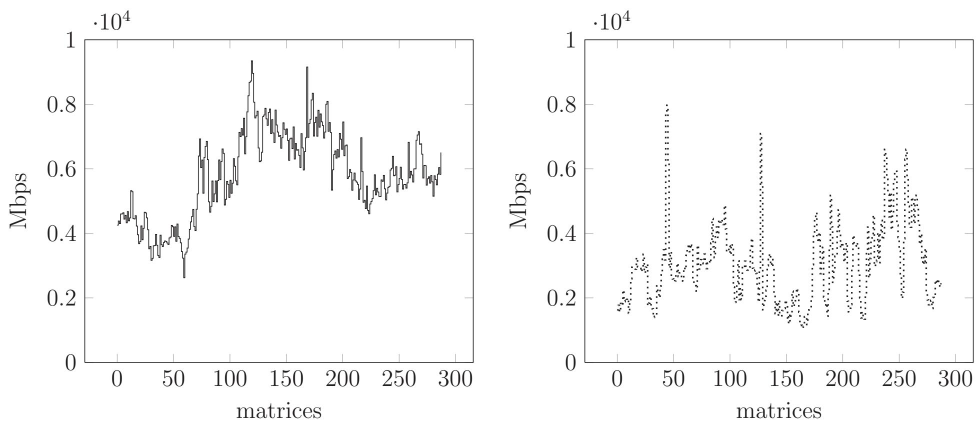
results transfer to more complex networks models (whereof our models are relaxations). For example, in Koster and Kutschka [100] we have shown the integration of survivability requirements into our models for the Γ-robust network design problem (cf. Chapter 9). In Duhovnikov et al. [64] we have extended our models to represent more technical hardware of a telecommunication network also aspects of multi-layer networks (implicitly by modeling multi-line rate planning).

Network dynamics. The planning and operation of communication networks holds several mathematical challenges. Especially the dynamics of such a complex structure are tricky to take into account in the mathematical planning. A key aspect of these dynamics is that network traffic fluctuates heavily over time. Figure 1.9 shows the total network load of ABILENE, the US Internet2 IP research backbone network [5, 162], for different time scales. We observe that it is far from constant. In fact, we can identify some trends and "patterns": first, the individual weeks of the 1-month time period can easily be identified in Figure 1.9(a) by the low-load weekend phases. If we zoom into one week of this month (see Figure 1.9(b)), we also notice the five working days with higher network load. However, we also observe that there are some — almost singular – very high peaks "on top" of this pattern. This is presumably caused by some large data transfers present in such kind of research backbone networks. Figures 1.9(c) and 1.9(d) show a more detailed zoom into time periods of one day and three hours, respectively. Here, we see that the amplitude of fluctuations decreases with smaller time periods, i. e., extreme fluctuations of the total network load are on larger time scales. For shorter time periods, there seems to be an averaging effect.

This averaging is caused by the following observation: the fluctuation peaks do not all occur at the same time at the same location. This happens only during extraordinary events like major sports events, special holidays, or greater emergencies. Usually, traffic peaks are equilibrated over time, i. e., multiple communication connections sharing the same resource tend to need less of that resource together than the sum of their individually maximal usage of this resource. This effect is known as *statistical multiplexing* in telecommunications. For example, Figure 1.10 shows the fluctuations of two different node-to-node demands of the ABILENE network for one day. We notice that the "patterns" are quite different. Furthermore, if we draw both plots into one diagram (see Figure 1.10(c)), we clearly see the effect of statistical multiplexing for these two demands. As a consequence, we do not need to reserve the sum of the maximum peaks for both demands when dimensioning capacity that is shared between both demands.

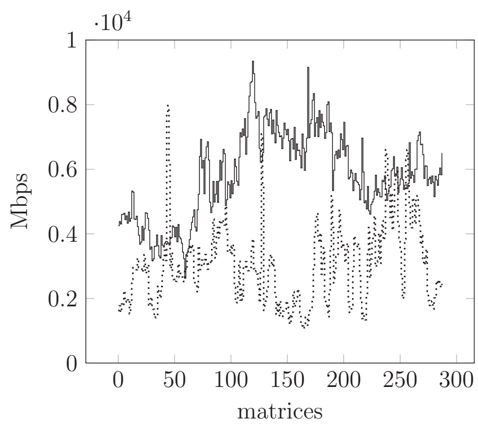
Another important aspect of network dynamics is the exponential growth of the total global network traffic. The latest forecast of consumer internet traffic by Cisco [54] predicts an average annual increase of 24 % between 2012 and 2017. Furthermore, the traffic is forecast to be more dynamic: while the average internet traffic will increase by factor 2.9 between 2012 and 2017, the internet traffic at busy hours is predicted to increase by factor 3.5. Figure 1.11 illustrates the annual forecast up to 2017. The traffic is shown in peta bytes per month; 1 peta byte is 1000 tera bytes. In addition, the forecast traffic is categorized into four types of traffic: gaming, file sharing (p2p), web, email and data transfer excluding file sharing (web/data), and internet video streaming (video). Notice the increasing importance of internet video streaming service from 56.6 % of consumer internet traffic in 2012 to 69.4 % in 2017. Also notice that the internet traffic caused by gaming can be neglected since it is less than 0.1 % of the consumer internet traffic.

These traffic fluctuations, statistical multiplexing, and the overall global traffic growth yielding an increasing scarcity of network capacity are currently not taken sufficiently into account during the planning process.



(a) Traffic fluctuations of the Chicago–Denver demand

(b) Traffic fluctuations of the Houston–Washington D.C. demand



(c) Statistical multiplexing of the Chicago–Denver and Houston–Washington D.C. demands

Figure 1.10: Example of traffic fluctuations and statistical multiplexing in the Abilene network [5]. Traffic measurements for the Chicago–Denver and Houston–Washington D.C. demands on July 5th in 2004 are shown.

So how are traffic fluctuations handled in practice? — Usually the uncertain traffic values are estimated by some rather vague measure yielding a single estimated value. For example, there exist several estimation rules based on population statistics; cf. Bley et al. [42] and Dwivedi and Wagner [65]. Then a deterministic optimization problem is solved using this single value. In a second step additional “safety resources” are included in the solution to increase the feasibility of the solution for more extreme traffic realizations. For example, additional edge capacities of about 300 % are added by some telecommunication network providers in practice. Needless to say, the resulting network is (most of the time) over-provisioned and not cost-efficient. Here new mathematical optimization methods can help.

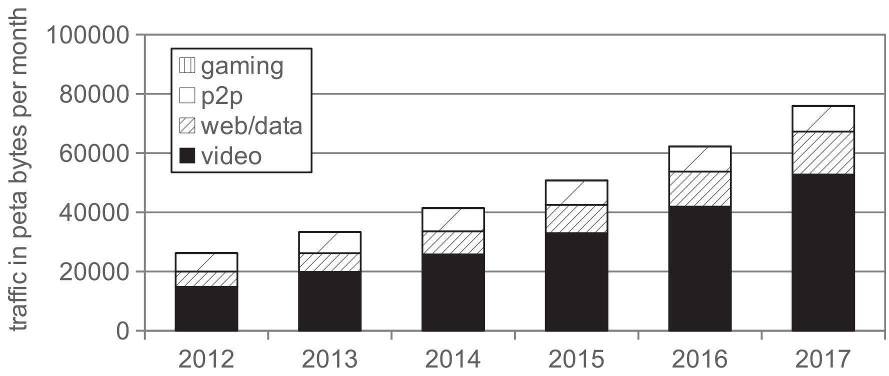


Figure 1.11: Forecast of consumer internet traffic by Cisco [54]. The traffic is categorized into gaming, file sharing (p2p), web, email and data transfer excluding file sharing (web/data), and internet video streaming (video).

In Chapter 9, we will consider one specific approach to network design taking into account such traffic fluctuations that yield cost-efficient solutions avoiding unnecessary over-provisioning of telecommunication networks. Our studies were carried out as part of the German ROBUKOM [3] project in cooperation with Nokia Siemens Networks GmbH & Co. KG, and also experimentally validated using real-life traffic measurements; cf. Chapter 11.

CHAPTER TWO

OPTIMIZATION UNDER DATA UNCERTAINTY

The scientific progress in the field of mathematical optimization has always been motivated by real-world problems. Oftentimes those problems include complicated processes and depend on numerous possible decisions. Mathematical optimization has been proven as an efficient methodology to obtain meaningful solutions to these problems.

Nevertheless, the mathematical model always simplifies the underlying real-world problem. This simplification intends to reduce the complexity of the mathematical model and thus to increase its computational tractability. One specific simplification is *data certainty*, e. g., the assumption that all input data is exactly known a-priori and can be represented exactly during the solving process.

But the optimization of real-world problems is always subject to data uncertainty. Ben-Tal et al. [31] name some reasons for data uncertainty:

prediction errors The data does not exist and can only be forecast, e. g., future demands in telecommunication networks, or future customer behavior. Oftentimes, estimations are based on historical data, if available. Sometimes the estimations are intentionally too conservative to introduce some buffering range for prediction errors.

measurement errors The data cannot be measured exactly, e. g., physical limitations to obtain real-time measurements. Furthermore, large distributed systems may imply the need for local measurements whose results cannot be propagated instantly. Delays are introduced and an exact real-time data measurement is not possible in practice.

implementation errors The real-world data cannot be represented exactly in the solving process or vice-versa. For example, on the one hand the input data is represented on a computer as precise as the precision of the floating point arithmetics. On the other hand, the precise value of a decision variable in an optimal solution might not be realizable in the real-world application the mathematical model is based on.

In Ben-Tal and Nemirovski [27], the authors present a sensitivity analysis of LP problems to data uncertainty. They study 90 instances from NETLIB [2] and their optimal solutions, which they call the nominal solutions. By perturbing the input data by 1 %, the corresponding nominal solution becomes infeasible with probability more than 5 % for 27 of these instances. In fact, for 13 instances a perturbation of 0.01 % already makes the nominal solution infeasible violating some constraints by more than 50 %. For a detailed example, see Ben-Tal and Nemirovski [27], Ben-Tal et al. [31].

From the very beginning of modern computer-aided optimization, researchers have been aware of the model limitations by assuming data certainty and of the open challenge to take data uncertainty into account properly. Several approaches based on different paradigms have been proposed, always struggling with the accuracy of the uncertainty model on the one hand and the increase in complexity and computational intractability on the other hand. Bertsimas et al. [36] give a good survey of optimization under data uncertainty with a focus on robust optimization in particular.

A very intuitive view on optimization problems under data uncertainty is given by the chance constraints: arbitrary constraints of the mathematical optimization model have to be satisfied with a given probability. For example, for the linear constraint

$$a^\top x \leq b$$

with uncertain coefficients a, the corresponding chance constraint

$$\mathbb{P}[a^\top x \leq b] \geq \varepsilon$$

requires the probability that $a^\top x \leq b$ is satisfied for the uncertain data to be at least $\varepsilon \in [0, 1]$. Of course, a chance constraint can also be formulated based on the complementary event and thus the probability that the constraint is not satisfied(i. e., violated):

$$\mathbb{P}[a^\top x > b] \leq 1 - \varepsilon.$$

In the following, we give a brief survey on different paradigms and approaches to handle data uncertainty in optimization problems. While also reporting on stochastic programming, dynamic programming and nonlinear approaches, we focus on robust optimization and its linear approaches in particular.

2.1 Stochastic optimization

In 1955, Dantzig [60] considered LPs under uncertainty coining the paradigm of *stochastic optimization*. Stochastic optimization assumes that the underlying stochastic distribution of the uncertain data is known. The aim of stochastic optimization is to solve a stochastic program such that the expected value of the objective function is optimal, cf. Shapiro et al. [145]. The general stochastic optimization problem can be formulated as follows

$$\min \ \mathbb{E}[\kappa(x, U)] \tag{2.1a}$$
$$\text{s.t.} \ f(x, U) \leq 0 \tag{2.1b}$$
$$x \geq 0 \tag{2.1c}$$

where $x \in \mathbb{R}^n_{\geq 0}$ is the vector of decision variables, $f : \mathbb{R}^{n\times k} \to \mathbb{R}$ a constraint function, and $\kappa : \mathbb{R}^n \to \mathbb{R}$ the objective function. The data uncertainty is given by the vector of random variables $U \in \mathbb{R}^k$ with known distribution $\mathcal{D}(U)$. Note, if the distribution $\mathcal{D}(U)$ is discrete, the expected value in the objective function can be written as a finite sum. Thus, if $f(x, U)$ is linear, the stochastic optimization problem can be formulated as an LP in this case.

Unfortunately, most of the times the distribution of uncertain real-world data is not known. Instead, a certain distribution is assumed, e. g., Poisson distribution for queuing problems or just a standard distribution. The resulting stochastic program might not model the data uncertainty properly anymore. Furthermore, a solution which is optimal for the expected value must not be optimal for any realization.

Moreover, it is well-known that solving stochastic optimization problems is computationally hard in practice as many scenarios or realizations have to be considered.

A detailed introduction and survey of stochastic optimization is given in Shapiro et al. [145], and applications are considered in Stein and Ziemba [149].

2.2 Robust optimization

A different paradigm, *robust optimization*, dates back to Soyster [147] in 1973. In contrast to stochastic programming, here the probability of a specific realization of uncertain data is not considered. Instead, it is assumed that all realizations of the uncertain data take place in a so-called *uncertainty set* $\mathcal{U}$. The aim of robust optimization is to solve a robust program such that its solution is optimal among all possible realizations of uncertain data, i. e., realization in the uncertainty set $\mathcal{U}$. Consider a general optimization problem

$$\min \kappa(x) \tag{2.2a}$$

$$\text{s. t. } f(x) \leq 0 \tag{2.2b}$$

with a vector of decision variables $x \in \mathbb{R}^n_{\geq 0}$, an objective function $\kappa : \mathbb{R}^n \to \mathbb{R}$, and a constraint function $f : \mathbb{R}^n \to \mathbb{R}^m$. The general robust optimization problem can be formulated as follows

$$\min \max_{u \in \mathcal{U}} \kappa(x, u) \tag{2.3a}$$

$$\text{s. t. } f^R(x, u) \leq 0 \qquad \forall u \in \mathcal{U} \tag{2.3b}$$

where $x \in \mathbb{R}^n_{\geq 0}$ is the vector of decision variables, $\kappa : \mathbb{R}^{n+k} \to \mathbb{R}$ the objective function, and $f^R : \mathbb{R}^{n+k} \to \mathbb{R}^m$ a constraint function. The possible realizations of the uncertain data u are given by the uncertainty set $\mathcal{U} \subseteq \mathbb{R}^k$. The formulation (2.3) is called the *robust counterpart* of the original formulation (2.2). We call a robust counterpart *compact* if its size is polynomial in the size of problem formulation for a single realization.

Ben-Tal et al. [31] show that any uncertain data in the objective function can be moved to the constraints by introduction of a new auxiliary variable capturing the uncertain terms. Hence, w. l. o. g., we consider the objective function to be unaffected of the data uncertainty, i. e., we write $\kappa(x)$. In addition, they show that the i-th constraint is only affected by the uncertainty of the data included in this constraint, i. e., $f_i^R(x, u_i) \leq 0 \; \forall u_i \in \mathcal{U}_i$ where $\mathcal{U}_i$ is the projection of the uncertainty set $\mathcal{U}$ to the data of constraint i. W. l. o. g., we assume $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \cdots \times \mathcal{U}_m$ with (topologically) closed sets $\mathcal{U}_i$. Then the robust counterpart of (2.2) reads

$$\min \kappa(x) \tag{2.4a}$$

$$\text{s.t. } f_i^R(x, u_i) \leq 0 \qquad \forall i = 1, \ldots, m, u_i \in \mathcal{U}_i. \tag{2.4b}$$

Note, a singleton uncertainty set, i. e., $|\mathcal{U}_i| = 1$, is used to model a constraint without data uncertainty. Furthermore, if $\mathcal{U}$ is a continuous set, infinitely many constraints (2.4b) exist.

By neglecting the probability of a specific realization, the robust solution has to be feasible even for realizations with probabilities almost zero, i. e., realizations which almost never occur. Therefore, a robust solution tends to be “cautious” by including additional slacks to constraints to guarantee feasibility for these rare realizations. This effect is called *conservatism* and one major drawback of the robust optimization paradigm. To overcome conservatism, the concept of a *budget of robustness* has been introduced where the uncertainty set $\mathcal{U}$ does not include all possible realizations but appropriate many realizations, especially those with high probability.

Following the robust optimization paradigm, we can formulate a robust version of an optimization problem which takes data uncertainty into account. In general this robust counterpart of an optimization problem has a significantly increased computational complexity compared to the original problem. In fact, Ben-Tal and Nemirovski [25] show that the robust counterpart of a convex optimization problem is in general computationally intractable. Nevertheless, there exist computationally tractable robust counterparts for special classes of functions $f_i(x)$ and types of uncertainty sets $\mathcal{U}_i$.

2.2.1 Uncertainty sets

In the previous, we have seen that in robust optimization the data uncertainty is characterized by uncertainty sets. Five important types of uncertainty sets are the following; cf. Bertsimas et al. [36].

discrete uncertainty set Let $(u_i)_{i \in \{1,\ldots,n\}}$ be a finite collection of real vectors $u_i \in \mathbb{R}^n$. Then

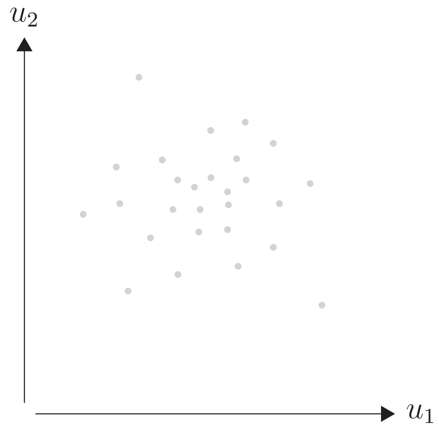
$$\mathcal{U} := \{u_1, \ldots, u_n\} \tag{2.5}$$

defines a discrete uncertainty set. An example is shown in Figure 2.1(a).

ellipsoidal uncertainty set Let $Q \in \mathbb{R}^{l \times k}$ be a matrix and $\rho \in \mathbb{R}_{\geq 0}$. Then

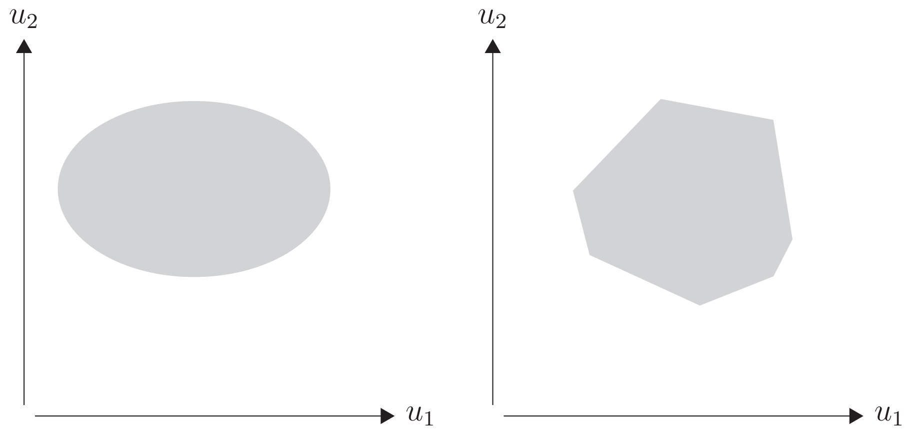
$$\mathcal{U} := \left\{u \in \mathbb{R}^k \;:\; \|Qu\| \leq \rho\right\} \tag{2.6}$$

defines an ellipsoidal uncertainty set. Figure 2.2(a) shows such a set.



(a) discrete

Figure 2.1: Example of a non-convex uncertainty set: discrete uncertainty set



(a) ellipsoidal

(b) polyhedral

Figure 2.2: Example of convex uncertainty sets: ellipsoidal and polyhedral uncertainty sets

polyhedral uncertainty set Let $D \in \mathbb{R}^{k \times l}$ be a matrix and $d \in \mathbb{R}^k$ a vector. Then

$$\mathcal{U} := P(D, d) = \left\{u \in \mathbb{R}^k \ : \ Du \leq d\right\} \tag{2.7}$$

defines a polyhedral uncertainty set. If $\mathcal{U}$ is bounded and thus the polyhedron, then it can be written as the convex hull of a finite number of extreme points corresponding to the worst-case realizations. A polyhedral uncertainty set is shown in Figure 2.2(b).

interval uncertainty set Let $I_1, I_2, \ldots, I_k \subseteq \mathbb{R}$ be closed intervals. Then

$$\mathcal{U} := \left\{u = \begin{pmatrix} u_1 & u_2 & \cdots & u_k \end{pmatrix}^\top \in I_1 \times I_2 \times \cdots \times I_k\right\} \tag{2.8}$$

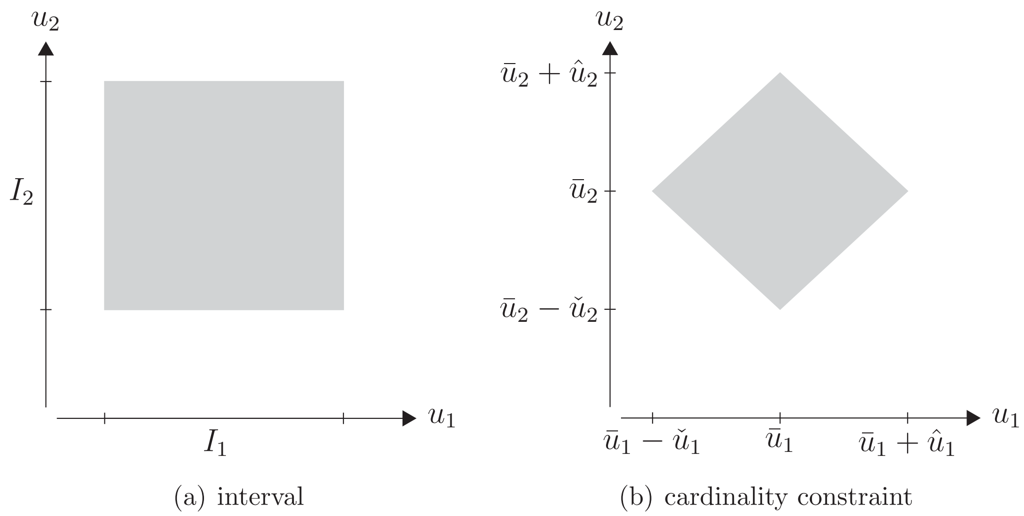


Figure 2.3: Example of special polyhedral uncertainty sets: interval and cardinality constraint uncertainty sets.

defines an interval uncertainty set; a special case of a polyhedral uncertainty set. See Figure 2.3(a) for an illustration.

cardinality constraint uncertainty set Another special case of a polyhedral uncertainty set is the cardinality constraint uncertainty set. Let $\bar{u}, \check{u}, \hat{u} \in \mathbb{R}^k$ be bounds such that $u \in [\bar{u} - \check{u}, \bar{u} + \hat{u}]$ holds for all realizations of the uncertain data u. We call $\bar{u}$ the *nominal value* of u, and $\check{u}$ and $\hat{u}$ the *negative* resp. *positive deviation* of u. Further let $\gamma \in \{0, \dots, k\}$ be the budget of robustness parameter bounding the number of data entries which deviate maximally at the same time in a realization of u. Then

$$\begin{aligned} \mathcal{U} := \Big\{ u = (u_i)_{i=1,\dots,k} \in \mathbb{R}^k \ : \ & \bar{u}_i - \check{\alpha}_i \check{u}_i \leq u_i \leq \bar{u}_i + \hat{\alpha}_i \hat{u}_i, \\ & \sum_{i=1}^{k} (\check{\alpha}_i + \hat{\alpha}_i) \leq \gamma, \\ & \check{\alpha}_i, \hat{\alpha}_i \in [0,1] \Big\} \end{aligned} \tag{2.9}$$

defines a cardinality constraint uncertainty set, where at most γ many data entries of u deviate with maximal negative or positive deviation. Note, it is possible that more than γ data entries deviate from their nominal value with only a fraction of their maximal positive or negative deviations, e. g., instead of two data entries at their maximal positive deviations, there can be four data entries deviating to half of their maximal positive deviations. Such an uncertainty set is shown in Figure 2.3(b).

A prominent cardinality constraint uncertainty set is the Γ-robust uncertainty set used in the Γ-robustness approach. Note, it is also a special case of polyhedral uncertainty. We will investigate this set and the underlying concept of Γ-robustness in detail in section 3.1.

2.2.2 Tractability of robust optimization problems

In the following, we consider different classes of optimization problems defined by specific constraint functions f in (2.2). We report on tractability results for the robust counterparts of these classes and different types of uncertainty sets.

Robust linear optimization. Let the constraint function f be of the form

$$f(x) = Ax - b, \tag{2.10}$$

with b certain. Then problem (2.2) is an LP and its robust counterpart is given by

$$\min \kappa^\top x \tag{2.11a}$$

$$\text{s. t. } A_{i.}x \leq b_i \qquad \forall A_{i.} \in \mathcal{U}_i, i = 1, \ldots, m. \tag{2.11b}$$

Note, that $A_{i.}x \leq b_i$ holds for all $A_{i.} \in \mathcal{U}_i$ if and only if $\max\{A_{i.}x \ : \ A_{i.} \in \mathcal{U}_i\} \leq b_i$ holds for all $i = 1, \ldots, m$. Thus, the possible infinitely many constraints (2.11b) can be verified by solving the maximization subproblem $\max\{A_{i.}x \ : \ A_{i.} \in \mathcal{U}_i\}$ and checking if the maximum is not greater than the smallest right-hand side b_i.

The robust counterpart of an LP with ellipsoidal uncertainty set is a second-order cone program (SOCP) as the maximization subproblem optimizes over a quadratic constraint; cf. [36]. If the uncertainty set $\mathcal{U}$ is polyhedral, the maximization subproblem is an LP and the robust counterpart (2.11) is linear as well; cf. [36]. The robust counterpart of an LP with a cardinality constrained uncertainty set is also an LP.

Robust quadratic optimization. Optimization problems with quadratic terms of decision variables in the objective or constraint functions are called *quadratic optimization problems*. The general optimization problem (2.2) with a constraint function f of the form

$$f(x) = \|A_i x\|^2 + b_i^\top x + c_i \tag{2.12}$$

is called a *quadratically constraint quadratic problem (QCQP)*. An equivalent formulation is the one obtained for constraint functions of the form

$$f(x) = \|A_i x + b^\top x_i\| - c_i^\top x - d_i \tag{2.13}$$

which is called *second-order cone program (SOCP)*. Sometimes the second-order cone is more illustratively also called the "ice cream cone".

The robust counterparts of a QCQP or a SOCP are $\mathcal{NP}$-hard if the uncertainty set $\mathcal{U}$ is the intersection of ellipsoids or even if it is polyhedral, see Ben-Tal and Nemirovski [25, 26], Ben-Tal et al. [29]. If $\mathcal{U}$ is a single ellipsoid, the robust counterparts are semidefinite optimization problems; cf. Ben-Tal et al. [29].

Robust semidefinite optimization. The optimization over symmetric positive definite matrix variables with linear cost and constraint functions is called *semidefinite optimization*. For example, quadratically constraints quadratic programming and linear programming are special cases of semidefinite programming. A general semidefinite programming formulation is given by

$$\min \; \operatorname{tr}(\kappa^\top X) \tag{2.14a}$$

$$\text{s.t.} \; \operatorname{tr}(A_i^\top X) = b_i \qquad \forall i = 1, \ldots, m \tag{2.14b}$$

$$X \succeq 0 \tag{2.14c}$$

where $X \in \mathbb{R}^{n\times n}$ is the (symmetrical) matrix decision variable, $\kappa \in \mathbb{R}^{n\times n}$ is a symmetrical objective coefficient matrix, $A_i \in \mathbb{R}^{n\times n}$ is a symmetrical constraint matrix for $i = 1, \ldots, m$, and $b \in \mathbb{R}^m$ a vector of corresponding right-hand side coefficients.

The robust counterparts of *semidefinite optimization problems (SDP)* are in general computationally intractable (see Ben-Tal and Nemirovski [25], Ben-Tal et al. [28]) even if the uncertainty set $\mathcal{U}$ is polyhedral (see Nemirovski [127]).

Robust discrete optimization. The robust counterpart to polynomially solvable discrete optimization problems may become $\mathcal{NP}$-hard, as shown for selected discrete optimization problems by Kouvelis and Yu [107]. When considering only uncertain data in the objective function, the robust counterpart becomes computationally tractable in some cases, e. g., the cardinality constrained data uncertainty in the objective function by Bertsimas and Sim [33, 34], or Altin et al. [12].

2.3 Multi-stage optimization under data uncertainty

In the previous sections, we considered single-stage optimization under data uncertainty, i. e., the optimal decision of all decision variables has to be made before any uncertainty is realized. Multi-stage optimization under data uncertainty, also called robust adaptable optimization, introduces a sequence of stages. A subset of the decision variables is assigned to each stage. The decisions have to be made sequentially, one stage at a time, and may take realizations of uncertain data in previous stages into account.

There exist several approaches to multi-stage optimization and extensions taking data uncertainty into account have been proposed for some of them. In the following, we highlight some of these.

Receding horizon. The receding horizon approach is an iterative approach. In each stage, a single-stage optimization problem including all decision variables of the current stage and all subsequent stages is solved. Realizations of data uncertainty in previous stages is taken into account. Thus, a static solution feasible to the current and all subsequent stages is determined and implemented for the current stage. Due to its non-adaptive static nature, this solution tends to be conservative and sub-optimal compared to an adaptive approach. Nevertheless, oftentimes the receding horizon approach is computationally tractable and used by operations research practitioners in production planning, etc. [36].

Dynamic programming. The recursive nature of dynamic programming can be seen as a sequential process where the outcome of the n-th step (or stage) depends only on the input of the $(n-1)$-th step. In this sense, it includes aspects of multi-stage optimization. The integration of uncertain data leads to robust dynamic programming and also to robust Markov decision processes. For details on those approaches we refer to Iyengar [86], Nilim and El Ghaoui [128], Xu et al. [160] and the references therein. Finally, we remark that the computational tractability of dynamic programs with certain data is highly influenced by the dimension of the recursion space. In the setting with uncertain data further assertions have to be met to preserve this property and obtain a computationally tractable robust dynamic program; cf. Xu et al. [160].

2.3.1 Multi-stage stochastic optimization

In multi-stage stochastic optimization, the decisions of later stages are called recourse (actions) as they may change decisions of earlier stages; sometimes even restoring feasibility lost due to a specific realization of uncertain data. If a complete change of earlier decisions is allowed, this is called *complete recourse*; cf. Birge and Louveaux [41]. In this case, a completely new/different solution can be obtained in each stage. If the recourse must preserve some parts of the earlier decisions, it is called *incomplete* or *limited recourse*. Note that in this case infeasibilities due to data realizations might not be restorable because of the limitations of the recourse action.

Multi-stage stochastic optimization problems with limited recourse are more difficult to solve in practice than those with complete recourse [36]. In general, multi-stage stochastic optimization problems are computationally more tractable if only a finite number of scenarios have to be considered. Then decomposition techniques as Bender's decomposition can be applied; cf. Bertsimas and Tsitsiklis [35]. Yet, a very large number of realizations may have to be considered to capture the structure of the uncertainty set. Furthermore, the overall complexity increases rapidly with the number of stages increasing.

2.3.2 Multi-stage robust optimization

Although we have learned that single-stage robust optimization is oftentimes computationally tractable for reasonable uncertainty sets, extending the concepts to multiple stages makes it rather tricky. Ben-Tal et al. [30] have shown that already the two-stage linear problem with deterministic uncertainty is $\mathcal{NP}$-hard in general, i. e., for general rules how future decisions depend on past realizations.

Affine adaptability. However Ben-Tal et al. [30] also point out that there exist special cases, i. e., classes of dependencies of the decision between the individual stages, which might be easier to solve. They introduce the concept of *affine adaptivity* where future decisions depend affinely on past decisions. Therefore, decisions of later stages can be parametrized by affine functions taking arguments of earlier stages. This can be done iteratively. Thus, only the first-stage decisions remain and the optimization problem becomes again a single-stage problem. Yet, the resulting problem includes non-linearities involving the uncertain data coefficients because of the recursive parametrizations; loosing its computational tractability.

If only the first-stage coefficients are uncertain and the uncertainty set is conic, then the affine adaptable robust counterpart is an LP [36].

Recoverable robustness A recent approach to multi-stage robust optimization is the so-called *recoverable robustness* concept, a two-stage deterministic approach similar to non-deterministic stochastic optimization with (limited) recourse. In recoverable robustness, a first-stage decision is taken without the knowledge of the uncertain data. Then after the realization of the uncertainty this first-stage decision may be adjusted in a second-stage to obtain feasibility ("recover" from a solution that became infeasible due to the data realization). The overall optimization problem is to determine the best first-stage decision taking the possible limited (and costly) second-stage adjustments/recovery into account while minimizing the total costs.

Recoverable robustness has been introduced in the context of applications in train scheduling by Liebchen et al. [110] and intensively studied in Büsing [46]. We discuss this robustness concept in detail in Section 3.4.

CHAPTER THREE

ROBUSTNESS CONCEPTS

In this chapter, we present four different concepts of robustness, which we will investigate in detail in this thesis. We discuss their usability and highlight their advantages and limitations.

At first, we consider the concept of Γ-robustness. It is well-known and thus can be seen as a starting point for our investigation of further robustness concepts. In addition, we consider the experimental properties of Γ-robust optimization problems as benchmark in the computational comparison of the robustness concepts we consider in this thesis.

Second, we present the concept of multi-band robustness, which generalizes the single deviation interval of the Γ-robust approach to multiple "histogram-like" intervals providing a more detailed modeling potential.

At third, recoverable robustness, a robustness concept implementing a two-stage robust optimization approach is discussed. Its key idea consists of a first stage decision and a limited second stage modification of the first stage decision after the realization of the uncertain data is known. The first stage decision may even become infeasible for a specific data realization such that its feasibility has to be "recovered" by the subsequent modification to restore its feasibility. Hence it is called recoverable robustness.

Fourth, we describe the concept of submodular robustness which generalizes multi-band robustness (and thus Γ-robustness) even further.

The last section of this chapter does not focus on further robustness concepts but on two other important aspects: determination of robustness parameters and evaluation of robustness. All presented robustness concepts include several robustness parameters and/or modeling decisions which highly affect the accuracy of the resulting model. We discuss how to determine meaningful parameter settings based on historical data and which trade-offs have to be considered. Furthermore, we present different robustness measures to evaluate the realized robustness of an optimal robust solution with respect to the historical data/sample the uncertainty set is based on.

3.1 Γ-robustness

With the revival of robust optimization in the last years, one specific robustness concept has become in particular popular: the concept of Γ-*robustness*. Introduced by Bertsimas and Sim [33, 34] in 2003, it uses a cardinality constrained uncertainty set. The concept of Γ-robustness implements a budget of robustness to control conservatism using a so-called robustness parameter, denoted by Γ. Bertsimas and Sim [34] have shown that the robust counterpart of an LP with a Γ-robust uncertainty set remains computational tractable. Furthermore, they have derived probabilistic guarantees that a constraint, affected by uncertain data from a Γ-robust uncertainty set, is fulfilled. These are only some reasons why Γ-robustness has become popular and widely applied to many optimization problems and real-world application areas, e. g., facility location, inventory management, supply chain management, revenue management, telecommunication network problems, vehicle routing problems, (train) shunting, and time tabling problems; as of summer 2013, Bertsimas' and Sim's original work has been cited about a thousand times.

3.1.1 The concept of Γ-robustness

In the following, we will define the concept of Γ-robustness formally, report on important results and investigate the Γ-robust counterparts of selected optimization problems under data uncertainty.

Let $\Gamma \in [0, n]$. The concept of Γ-robustness makes the following assumptions on the vector of uncertain data $u \in \mathbb{R}^n$:

- Each entry u_i of u is modeled as an independent random variable with unknown bounded symmetrical distribution.

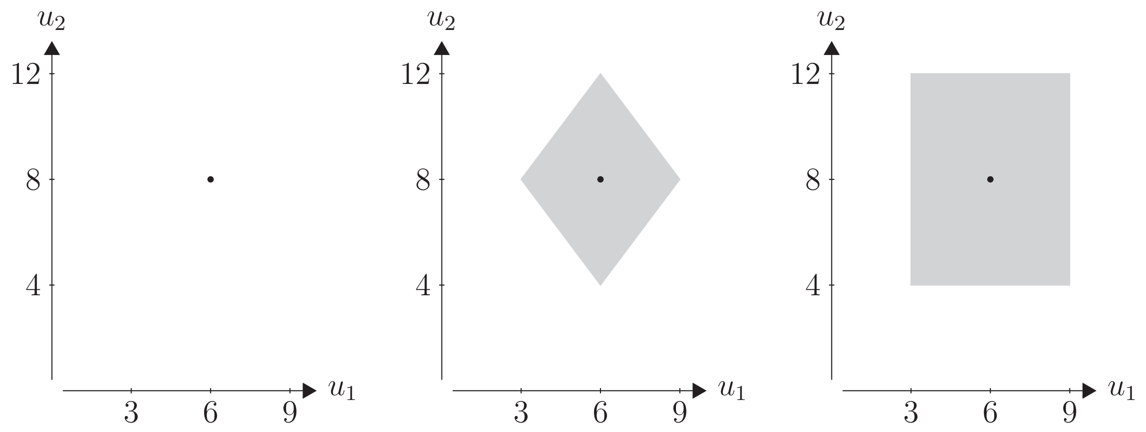
- The distribution is assumed to be symmetrical around a nominal value $\bar{u}_i$ with a maximal deviation of $\hat{u}_i \geq 0$, i. e., $u_i \in [\bar{u}_i - \hat{u}_i, \bar{u}_i + \hat{u}_i]$ holds for every realization of u.

- Every realization vector of u has at most $\lfloor \Gamma \rfloor$-many entries u_i simultaneously deviating from their nominal values $\bar{u}_i$ to either their minimum $\bar{u}_i - \hat{u}_i$ or maximum values $\bar{u}_i + \hat{u}_i$. In addition, another entry u_i may deviate to $\Gamma - \lfloor \Gamma \rfloor$ to its minimum or maximum value at the same time.

Although, the Γ-robustness concept of Bertsimas and Sim allows fractional values of Γ, we restrict ourselves w. l. o. g. to integer values of Γ in this thesis. Note, our results can be generalized to fractional values following the interpretation of those values presented above.

This leads us to the subsequent definition of the Γ-robust uncertainty set.

Definition 3.1 (Γ-robust Uncertainty Set)**.** Let $\bar{u}, \hat{u} \in \mathbb{R}^n$, $\hat{u} \geq 0$ be the nominal and the deviation values of the uncertain data $u = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^\top$ such that $u_i \in [\bar{u}_i - \hat{u}_i, \bar{u}_i + \hat{u}_i]$ holds for every realization of u. Further, let $\Gamma \in \{0, 1, \ldots, n\}$ be a parameter. Then the Γ-*robust uncertainty set* $\mathcal{U}^\Gamma$ is defined as



(a) $\Gamma = 0$ (b) $\Gamma = 1$ (c) $\Gamma = 2$

Figure 3.1: Γ-robust uncertainty sets for the problem instance defined in Example 3.2 and $\Gamma = 0, 1, 2$. Also the nominal data value $(6, 8)$ is shown. While only a single point $(6, 8)$ for $\Gamma = 0$, the volume of the convex uncertainty set and thus its conservatism grows with increasing value of Γ.

$$\mathcal{U}^\Gamma := \Big\{ u = (u_i)_{i=1,\dots,n} \in \mathbb{R}^n \ : \bar{u}_i - \hat{\alpha}_i \hat{u}_i \leq u_i \leq \bar{u}_i + \hat{\alpha}_i \hat{u}_i, \qquad (3.1)$$
$$\sum_{i=1}^{n} \hat{\alpha}_i \leq \Gamma, \ \hat{\alpha}_i \in [0, 1] \Big\}.$$

Note, the Γ-robust uncertainty set includes realizations where more than Γ entries $\bar{u}_i$ deviate from their nominal values as long as the sum of their relative deviations (i. e., $\hat{\alpha}_i$) is at most Γ. Clearly, the most deviated realizations are those on the border of the polyhedra and thus, vertices or a convex combination of them. Hence, it suffices to consider extreme points, i. e., realizations with $\hat{\alpha} \in \{0, 1\}^n$ and $\sum_{i=1}^{n} \hat{\alpha}_i = \Gamma$.

Example 3.2. *Let us consider uncertain data* $u \in \mathbb{R}^2$ *with nominal values* $\bar{u} = \begin{pmatrix} 6 & 8 \end{pmatrix}^\top$ *and deviation values* $\hat{u} = \begin{pmatrix} 3 & 4 \end{pmatrix}^\top$*. Figure 3.1 illustrates some possible data realizations of such a set. Furthermore let* $\Gamma \in \{0, 1, 2\}$*; cf. Figure 3.1(a)–3.1(c). With increasing value of* Γ*, more realizations are in the uncertainty set and thus protected (i. e., considered in the worst-case determination to derive the* Γ*-robust counterpart). At the same time the level of conservatism increases: for* $\Gamma = 2$ *the worst-case is the realization* $(9, 12)$*, although this realization may only occur with very low probability.*

3.1.2 The Γ-robust counterpart

Let $\kappa \in \mathbb{R}^n_{>0}$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. Consider a general LP in its maximization form

$$\max\ \kappa^\top x \tag{3.2a}$$
$$\text{s.t. } A_{i\cdot} x \le b_i \qquad \forall i = 1, \ldots, m \tag{3.2b}$$
$$x \ge 0. \tag{3.2c}$$

Its Γ-robust counterpart reads

$$\max\ \kappa^\top x \tag{3.3a}$$
$$\text{s.t. } A_{i\cdot} x \le b_i \qquad \forall i = 1, \ldots, m, A_{i\cdot}^\top \in \mathcal{U}_i^\Gamma \tag{3.3b}$$
$$x \ge 0. \tag{3.3c}$$

Note, there are exponential many constraints (3.3b). We can rewrite (3.3) using a polynomial number of constraints as

$$\max\ \kappa^\top x \tag{3.4a}$$
$$\text{s.t. } \max_{A_{i\cdot}^\top \in \mathcal{U}_i^\Gamma} A_{i\cdot} x \le b_i \qquad \forall i = 1, \ldots, m \tag{3.4b}$$
$$x \ge 0 \tag{3.4c}$$

with only m constraints (3.4b). Unfortunately, formulation (3.4) is not linear anymore due to its max-term. However, the max-term of constraint $1 \le i \le m$ is an optimization problem itself and can – for a fixed value $\tilde{x}$ of x – be formulated as an ILP

$$\max_{A_{i\cdot}^\top \in \mathcal{U}_i^\Gamma} A_{i\cdot} \tilde{x} \quad = \quad \max \sum_{j=1}^{n} \left(\bar{A}_{ij} \tilde{x}_j + \hat{A}_{ij} \tilde{x}_j z_j \right) \tag{3.5a}$$
$$\text{s.t. } \sum_{j=1}^{n} z_j \le \Gamma \tag{3.5b}$$
$$z_j \in \{0, 1\} \qquad \forall j = 1, \ldots, n \tag{3.5c}$$

with indicator variables $z_j = 1$ if and only if the j-th entry A_{ij} is at its maximal/peak value $\bar{A}_{ij} + \hat{A}_{ij}$. Negative deviations are ignored in this formulation as (3.5) is a maximization problem and its optimum will always include positive and exclude negative deviations (since $x \ge 0$). Note, the objective function term $\sum_{j=1}^{n} \bar{A}_{ij} \tilde{x}_j$ is constant. Further, the coefficient matrix defined by constraints (3.5b) is totally unimodular, i. e., the set of feasible solutions of the linear relaxation of (3.5) is a polytope with only integer vertices.

Hence, we can relax the integrality constraint obtaining a LP. By strong duality theory of linear programming

$$\begin{aligned}
\max \sum_{j=1}^{n} \hat{A}_{ij}\tilde{x}_j z_j & \quad = \min \Gamma\pi_i + \sum_{j=1}^{n} \rho_{ij} \\
\text{s. t.} \sum_{j=1}^{n} z_j \leq \Gamma & \quad \text{s. t.} \pi_i + \rho_{ij} \geq \hat{A}_{ij}\tilde{x}_j \qquad \forall j = 1, \dots, n \\
z_j \in [0,1] \qquad \forall j = 1, \dots, n & \quad \pi_i, \rho_{ij} \geq 0 \qquad \forall j = 1, \dots, n
\end{aligned}$$

holds. The variables π_i and ρ_i are dual variables to constraints (3.5b) and the upper bound constraints of the relaxation of (3.5c), respectively. This yields the following nonlinear formulation of the Γ-robust counterpart of (3.2):

$$\max \kappa^\top x \tag{3.7a}$$

$$\text{s. t.} \sum_{j=1}^{n} \bar{A}_{ij}x_j + \min\left(\Gamma\pi_i + \sum_{j=1}^{n} \rho_{ij}\right) \leq b_i \qquad \forall i = 1, \dots, m \tag{3.7b}$$

$$\pi_i + \rho_{ij} \geq \hat{A}_{ij}x_j \qquad \forall i = 1, \dots, m, j = 1, \dots, n \tag{3.7c}$$

$$x, \pi, \rho \geq 0. \tag{3.7d}$$

We can relax the min-operator in constraint (3.7b) because if a non-minimal solution of the minimization subproblem fulfills this constraint, the minimum does it as well. Furthermore, due to the objective sense of (3.7) and the non-negativity of x, the term tends to its minimum in the overall optimization process. In summary this gives us the following compact LP formulation of the Γ-robust counterpart of the general LP (3.2)

$$\max \kappa^\top x \tag{3.8a}$$

$$\text{s. t.} \sum_{j=1}^{n} \bar{A}_{ij}x_j + \Gamma\pi_i + \sum_{j=1}^{n} \rho_{ij} \leq b_i \qquad \forall i = 1, \dots, m \tag{3.8b}$$

$$\pi_i + \rho_{ij} \geq \hat{A}_{ij}x_j \qquad \forall i = 1, \dots, m, j = 1, \dots, n \tag{3.8c}$$

$$x, \pi, \rho \geq 0. \tag{3.8d}$$

In contrast to the exponential formulation (3.3), its size is polynomial in the size of the non-robust LP (3.2). In fact, it has $m(n + 1)$ additional variables and $m \cdot n$ additional constraints.

This observation is crucial as it implies that applying the Γ-robustness concept to linear programming does not increase the complexity of the problem. Nevertheless, we can expect an impact on its computational tractability in practice as the problem size increases polynomially.

In contrast to the compact reformulation (3.8), Fischetti and Monaci [69] propose to use the exponential-sized formulation (3.3) together with a lazy constraint approach (i.e., dynamic row generation) in practice. In their approach, violated constraints (3.3b) are separated on-the-fly as so-called *robustness cuts*. The computational effectiveness of those cuts is discussed for the uncertain set covering problem in [69]. The authors conclude that this approach seems promising for LPs but not computationally tractable for ILPs as the solver produces too many infeasible solutions which must be cut-off quite inefficiently by robustness cuts each. This comes along with our computational results in Koster et al. [103, 106] where the compact robust counterpart outperforms other approaches.

3.1.3 Probabilistic analysis and feasibility guarantees

Bertsimas and Sim [33, 34] also investigate the probability that a constraint subject to a Γ-robust uncertainty set is violated. They derive probability bounds on the constraint violation under mild conditions to the underlying unknown probability distribution. In the following, we report on these results.

Given $\Gamma \in \{0, \ldots, n\}$ and a Γ-robust linear constraint

$$A_{i\cdot}x \leq b_i \quad \forall A_{i\cdot}^\top \in \mathcal{U}_i^\Gamma, \tag{3.9}$$

we are interested in the probability

$$\mathbb{P}[A_{i\cdot}\tilde{x} > b_i] \tag{3.10}$$

that this constraint is violated for an optimal solution $\tilde{x}$. The next theorem yields upper bounds on this probability. Notice that these bounds may be fractional since the definition of Γ may also be extended to fractional values; cf. the introduction of this chapter.

Theorem 3.3 (Bertsimas and Sim [33, 34]). *Let $\tilde{x}$ be an optimal solution of problem* (3.8)*, then the following holds:*

1. *Suppose the Γ-robust data uncertainty set is used in the model, then the probability that the i-th constraint is violated satisfies:*

$$\mathbb{P}[A_{i\cdot}\tilde{x} > b_i] \leq B(n, \Gamma_i) = \frac{1}{2^j}\left((1-\mu)\sum_{\ell=\lfloor\nu\rfloor}^{n}\binom{n}{\ell} + \mu\sum_{\ell=\lfloor\nu\rfloor+1}^{n}\binom{n}{\ell}\right) \tag{3.11}$$

   where $n = \dim(\mathcal{U}_i^\Gamma)$ is the number of coefficients affected by data uncertainty, $\nu = \frac{\Gamma_i+n}{2}$, and $\mu = \nu - \lfloor\nu\rfloor$. Moreover, the bound is tight.

2. *The bound* (3.11) *satisfies*

$$B(n, \Gamma_i) \leq (1-\mu)C(n, \lfloor\nu\rfloor) + \sum_{\ell=\lfloor\nu\rfloor+1}^{n} C(n, \ell) \tag{3.12}$$

   where

$$C(n,\ell) = \begin{cases} \frac{1}{2^n}, & \text{if } \ell \in \{0, n\}, \\ \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(n-\ell)\ell}}\exp\left(n\log\left(\frac{n}{2(n-\ell)}\right) + \ell\log\left(\frac{n-\ell}{\ell}\right)\right), & \text{otherwise.} \end{cases}$$

n	bounds (3.11), (3.12)	approximation (3.14)
5	5.0	5.0
10	8.2	8.4
100	24.3	24.3
200	33.9	33.9
2000	105.0	105.0

Table 3.1: Bertsimas and Sim [34]. Choice of Γ_i as a function of $n(=\dim(\mathcal{U}^\Gamma))$ so that the probability of constraint violation is less than 1 %.

3. For $\Gamma_i = \theta\sqrt{n}$,

$$\lim_{n\to\infty} B(n, \Gamma_i) = 1 - \Phi(\theta) \tag{3.13}$$

where

$$\Phi(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta} \exp\left(-\frac{y^2}{2}\right) dy$$

is the cumulative distribution function of the standard normal.

Bertsimas and Sim [33] remark that the bound (3.12) is easier to compute than bound (3.11) because it avoids to evaluate sums of binomials for large values n. Further, they derive the following approximation of the bound (3.11) by using the De Moive-Laplace approximation of the Binomial distribution

$$B(n, \Gamma_i) \approx 1 - \Phi\left(\frac{\Gamma_i - 1}{\sqrt{n}}\right). \tag{3.14}$$

Table 3.1 states the bounds on (and approximation of) Γ_i for selected values of n as determined in Bertsimas and Sim [34]. It shows that a Γ-robust linear constraint (3.9) with a sufficiently large number of uncertain coefficients is already fulfilled with probability greater than 99 % for comparatively small values of Γ_i. For example, if there are 100 uncertain coefficients, $\Gamma_i = 24.3$ already yields a probability greater than 99 % that this constraint is satisfied.

3.2 Multi-band robustness

In the following, we present a generalization of the concept of Γ-robustness: multi-band robustness. It uses multiple deviation intervals, so-called bands. For each band and each uncertain data coefficient an associated deviation value is assumed. In addition, bounds on the total number of realizations in each band are included in the concept. Thus, a "histogram-like" discrete distribution can be specified. Following this approach, the concept of Γ-robustness is the special case with one nominal value (band) and one negative and one positive deviation band. The idea of multi-band (or histogram) robustness goes back to the work of Bienstock [37] where the key idea has been applied to portfolio optimization problems in finance but not formulated as a more general abstract robustness concept. Later, this idea has been applied to wireless network design problems by Bienstock and D'Andreagiovanni [38], D'Andreagiovanni [58]. Recently, Büsing and D'Andreagiovanni [47, 48] introduced the concept of multi-band robustness as a theoretical framework presenting fundamental investigations on its properties, multi-band robust counterparts of LPs, and preliminary probabilistic studies on feasibility guarantees. Shortly after this work, Mattia [122] presented a very similar robustness concept in a technical report. But in contrast to multi-band robustness it is less general: it assumes symmetrical random variables and does not have lower bounds on the number of realizations. In particular, the lack of the lower bounds implies a higher conservatism.

3.2.1 The concept of multi-band robustness

Let us consider the concept of multi-band robustness more formally. It makes the following assumptions on the vector $u \in \mathbb{R}^n$ of uncertain data: Each entry u_i of u is modeled as an independent random variable with unknown distribution. There is a set B of negative and positive deviation bands

$$B = \{\underline{B}, \dots, -1, 0, 1, \dots, \overline{B}\} \tag{3.15}$$

where $-\underline{B}$ and $\overline{B}$ denote the number of negative and positive deviation bands, respectively. The band 0 corresponds to no deviation and thus the nominal value. It is included as a band for notational reasons. Hence, there are $\overline{B} - \underline{B} + 1$ bands in total. For each uncertain coefficient u_i, there exists a nominal value $\bar{u}$ and a deviation value $\hat{u}^b$ for each band $b \in B$ so that

$$-\infty < \hat{u}_i^{\underline{B}} < \dots < \hat{u}_i^0 = 0 < \dots < \hat{u}_i^{\overline{B}} < \infty \tag{3.16}$$

holds. We say a realization $\tilde{u}_i$ lies in band $b \in B$ if and only if

$$\tilde{u}_i \in (\bar{u}_i + \hat{u}_i^{b-1}, \bar{u}_i + \hat{u}_i^b].$$

In addition each band $b \in B$ has two robustness parameters $\gamma^b, \Gamma^b \in \{0, 1, \dots, n\}$ with $0 \leq \gamma^b \leq \Gamma^b \leq n$ bounding the total number of realizations in b from below (γ^b) and above (Γ^b). To allow feasible realizations $\sum_{b \in B} \gamma^b \leq n$ must hold. Further, the band 0 robustness parameters are defined as $\gamma^0 = 0$ and $\Gamma^0 = n$. All bands, nominal and deviation values, and robustness parameters are assumed to model the underlying probability distribution. In particular, we assume that

$$u \in (\bar{u} + \hat{u}^{\underline{B}}, \bar{u} + \hat{u}^{\overline{B}}] \tag{3.17}$$

holds for all realizations of the uncertain data vector $u \in \mathbb{R}^n$.

Using the previous notation and definitions, we can now formally define the multi-band robust uncertainty set as

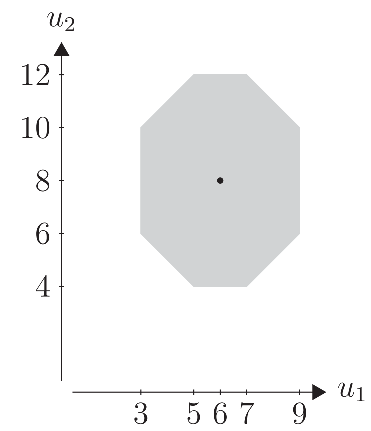


Figure 3.2: Multi-band robust uncertainty set of problem instance defined in Example 3.5 and for bounds $\gamma^b = 0$ for all $b \in B$, $\Gamma^b = 1$ for all $b \in B \setminus \{0\}$ and $\Gamma^0 = 2$

Definition 3.4 (Multi-band Robust Uncertainty Set)**.** Let the vector of uncertain data be denoted by $u = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^\top \in \mathbb{R}^n$. Let B be the set of bands and $\bar{u}, \hat{u}^b \in \mathbb{R}^n$, $\gamma^b, \Gamma^b \in \{0, 1, \ldots, n\}$ for all $b \in B$ such that $u \in (\bar{u} + \hat{u}^{\underline{B}}, \bar{u} + \hat{u}^{\overline{B}}]$ holds for every realization as described above.

Then, the *multi-band robust uncertainty set* $\mathcal{U}^{\text{mb}}$ is defined as

$$\mathcal{U}^{\text{mb}} := \text{conv}\Big\{ u = (u_i)_{i=1,\ldots,n} \in \mathbb{R}^n \; : \\ \bar{u}_i + \sum_{b \in B} \hat{\alpha}_i^b \hat{u}_i^{b-1} < u_i \leq \bar{u}_i + \sum_{b \in B} \hat{\alpha}_i^b \hat{u}_i^b \quad \forall i = 1, \ldots, n, \\ \gamma^b \leq \sum_{i=1}^{n} \hat{\alpha}_i^b \leq \Gamma^b \quad \forall b \in B, \\ \sum_{b \in B} \hat{\alpha}_i^b = 1, \; \hat{\alpha}_i^b \in \{0, 1\} \quad \forall i = 1, \ldots, n, \; b \in B \Big\}.$$

with $\hat{u}^{\underline{B}-1} := \hat{u}^{\underline{B}}$ for technical reasons and notational simplicity.

Example 3.5. *Let us consider an example of an uncertain problem where two coefficients u_1, u_2 are uncertain. We decide to model five bands: the nominal band, two positive and two negative deviation bands. Formally, let $n = 2$ and $B := \{-2, -1, 0, 1, 2\}$ with the following deviation values $\hat{u}_1^{-2} = -3$, $\hat{u}_1^{-1} = -1$, $\hat{u}_1^1 = 1$, $\hat{u}_1^2 = 3$, $\hat{u}_2^{-2} = -4$, $\hat{u}_2^{-1} = -2$, $\hat{u}_2^1 = 2$, $\hat{u}_2^2 = 4$, and $\hat{u}_1^0 = \hat{u}_2^0 = 0$. For each deviation band we assume at least 0 and at most 1 realization to fall into this band: i. e., $\gamma^b = 0$ for all $b \in B$, $\Gamma^b = 1$ for all $b \in B \setminus \{0\}$ and $\Gamma^0 = 2$.*

The corresponding multi-band robust uncertainty set is shown in Figure 3.5. Note that the example corresponds to Example 3.2 for the Γ-robust uncertainty set. The deviation value of the maximal deviation bands is the same. Here, we have introduced additional intermediate deviation bands to model the uncertainty more precise. By *comparing Figures 3.5 and 3.1(c), we observe that the uncertainty set $\mathcal{U}^{mb}$ is a subset of the corresponding uncertainty set $\mathcal{U}^{\Gamma}$ and thus, it is also less conservative.*

Remark 3.6. In the following, we assume that the uncertain data u affects the left-hand side of linear constraint of the for $ux \leq d$. Thus positive deviations contribute most to a worst-case realization with respect to the feasibility of this constraint. For constraints with different relation signs, the following definition of the frequency profile has to be adjusted accordingly or the constraints have to be transformed into the assumed form.

Given a multi-band robust uncertainty set $\mathcal{U}^{\text{mb}}$, Büsing and D'Andreagiovanni [48] define its *(frequency) profile* p, $\{\vartheta^b\}_{b\in B}$ as follows

$$p := \min \left\{ b \in B \ : \ \sum_{i=\underline{B}}^{b} \gamma^i + \sum_{i=b+1}^{\overline{B}} \Gamma^i \leq n \right\}$$

$$\vartheta^b := \begin{cases} \gamma^b & \text{if } b \leq p-1 \\ \Gamma^b & \text{if } b \geq p+1 \\ n - \sum_{b\in B\setminus\{p\}} \vartheta^b & \text{if } b = p \end{cases}$$

where p denotes the smallest band index where the lower bound on data realization may be exceeded in the worst-case. It is used to determine the values ϑ^b which determine the exact number of data realizations in band b in a worst-case realization; see also Example 3.8. Note that since $\Gamma^0 = n$, it holds $p \geq 0$. Besides, $\sum_{b\in B} \vartheta^b = n$ by definition.

Further, Büsing and D'Andreagiovanni [48] have shown that only realization vectors $u = (u_i)_{i=1,\dots,n} \in \mathcal{U}^{\text{mb}}$ with $u_i \in \{\bar{u} + \hat{u}^{\underline{B}+1}, \dots, \bar{u} + \hat{u}^0, \dots, \bar{u} + \hat{u}^{\overline{B}}\}$ have to be considered when determining a worst-case data realization. Furthermore, there exists a worst-case realization for which we can determine the number of realizations u_i for each band $b \in B$ as follows.

Lemma 3.7 (Büsing and D'Andreagiovanni [48])**.** *Consider uncertain data modeled using a multi-band robust uncertainty set $\mathcal{U}^{mb}$. Let p, $\{\vartheta^b\}_{b\in B}$ be its corresponding frequency profile and $u \in \mathcal{U}^{mb}$ an arbitrary realization vector.*

If there exists a band $b \in B$ with not exactly ϑ^b-many realizations u_i, then u is dominated by a different realization vector $u' \in \mathcal{U}^{mb}$ with ϑ^b-many realizations u'_i in band b.

The following example illustrates the idea of the frequency profile and its contribution to a worst-case scenario realization.

Example 3.8. *Consider a constraint with $n = 10$ uncertain coefficients modeled by a multi-band robust uncertainty set with three negative and three positive bands, and bounds γ and Γ given by Table 3.2. In addition, this table shows the value $\sum_{i=\underline{B}}^{b} \gamma^i + \sum_{i=b+1}^{\overline{B}} \Gamma^i$ used to determine the band p. It follows $p = 1$. The resulting values ϑ^b are shown in the last row of Table 3.2. We observe that the frequency profile basically follows the idea of first satisfying all lower bounds γ^b and second filling up bands from high bands to low bands until the upper bounds Γ^b are reached or in total n uncertain coefficients have been assigned.*

band b:	-3	-2	-1	0	1	2	3
lower bound γ^b on #realizations:	0	1	2	0	2	2	0
upper bound Γ^b on #realizations:	1	2	3	10	4	3	1
$\sum_{i=\underline{B}}^{b} \gamma^i + \sum_{i=b+1}^{\overline{B}} \Gamma^i$	24	23	21	11	$\underline{9}$	8	7
number ϑ^b of realizations in worst-case:	0	1	2	0	3	3	1

Table 3.2: Example of frequency profile for multi-band robust uncertainty sets

From Lemma 3.7, we know that the frequency profile determines for each band the number of coefficient realizations u_i of a worst-case realization vector $u \in \mathcal{U}^{\text{mb}}$. But it does not state which coefficient u_i realizes in exactly which band $b \in B$ and thus also the total value $\sum_{i=1}^{n} u_i$ cannot be evaluated directly. Büsing and D'Andreagiovanni [47] have shown that this assignment problem (uncertain coefficients must be assigned to bands in which they realize such that the total sum is maximized) subject to the side constraints of multi-band robustness can be solved optimally by a combinatorial algorithm in polynomial time.

Lemma 3.9 (Büsing and D'Andreagiovanni [47])**.** *The problem to determine a worst-case realization vector, i. e., an assignment of uncertain coefficient to bands such that $\sum_{i=1}^{n} u_i$ is maximized, is equivalent to solving a min-cost flow problem on an appropriate auxiliary graph.*

3.2.2 The multi-band robust counterpart

Let $\kappa \in \mathbb{R}^n_{>0}$, $A \in \mathbb{R}^{m \times n}$, and $d \in \mathbb{R}^m$. Consider a general LP in its maximization form

$$\max \kappa^\top x \tag{3.18a}$$
$$\text{s.t. } A_{i\cdot} x \leq d_i \qquad \forall i = 1, \dots, m \tag{3.18b}$$
$$x \geq 0. \tag{3.18c}$$

Then its multi-band robust counterpart reads

$$\max \kappa^\top x \tag{3.19a}$$
$$\text{s.t. } A_{i\cdot} x \leq d_i \qquad \forall i = 1, \dots, m, A_{i\cdot}^\top \in \mathcal{U}_i^{\text{mb}} \tag{3.19b}$$
$$x \geq 0. \tag{3.19c}$$

and can be reformulated using only a polynomial number of constraints of type (3.19b) by considering the worst-case realization. This is analogous to the special case of Γ-robustness. The resulting formulation is

$$\max \kappa^\top x \tag{3.20a}$$
$$\text{s.t. } \max_{A_{i\cdot}^\top \in \mathcal{U}_i^{\text{mb}}} A_{i\cdot} x \leq d_i \qquad \forall i = 1, \dots, m \tag{3.20b}$$
$$x \geq 0. \tag{3.20c}$$

This formulation is nonlinear since $\max_{A_{i\cdot}^\top \in \mathcal{U}_i^{\mathrm{mb}}} A_{i\cdot}x$ has to be determined which is itself a maximization problem. Again, as for the Γ-robust counterpart, we can formulate this inner maximization problem as a separate integer LP. At this point the frequency profile pays off in particular: by using it, we avoid to characterize the number of realizations in a band b by its bounds γ^b and Γ^b because the resulting two constraints would later yield additional dual variables in the dualized problem. Considering a constraint $1 \leq i \leq m$ and a fixed vector $\tilde{x}$ for x, the term $\max_{A_{i\cdot}^\top \in \mathcal{U}_i^{\mathrm{mb}}} A_{i\cdot}\tilde{x}$ can be formulated as

$$\max_{A_{i\cdot}^\top \in \mathcal{U}_i^{\mathrm{mb}}} A_{i\cdot}\tilde{x} \quad = \quad \max \sum_{j=1}^{n} \left(\bar{A}_{ij}\tilde{x}_j + \sum_{b\in B} \hat{A}_{ij}^b \tilde{x}_j z_j^b \right) \tag{3.21a}$$

$$\text{s. t.} \sum_{j=1}^{n} z_j^b \leq \vartheta^b \qquad \forall b \in B \tag{3.21b}$$

$$\sum_{b\in B} z_j^b = 1 \qquad \forall j = 1,\ldots,n \tag{3.21c}$$

$$z_j^b \in \{0,1\} \qquad \forall j = 1,\ldots,n, b \in B \tag{3.21d}$$

with indicator variables $z_j^b = 1$ if and only if the j-th entry A_{ij} is assigned to band $b \in B$, i. e., $A_{ij} \in (\bar{A}_{ij} + \hat{A}_{ij}^{b-1}, \bar{A}_{ij} + \hat{A}_{ij}^b]$. W. l. o. g. we can ignore the constant term $\sum_{j=1}^n \bar{A}_{ij}\tilde{x}_j$. Let us consider the remaining problem. Büsing and D'Andreagiovanni [48] have shown the totally unimodularity of the coefficient matrix resulting from formulation (3.21). In their model, the constraint (3.21c) is relaxed to the inequality $\sum_{b\in B} z_j^b \leq 1$. Nevertheless their proof holds to our formulation (3.21) as the coefficient matrix is not changed. Due to the totally unimodularity, the integrality constraint (3.21d) can be relaxed resulting in the constraint

$$0 \leq z_j^b \leq 1.$$

Notice, the upper bound on z_j^b is dominated by constraint (3.21c) and thus redundant. The resulting formulation is a pure LP. Hence, we can apply strong LP duality. This yields

$$\begin{array}{llllll}
\max & \displaystyle\sum_{b\in B} \hat{A}_{ij}^b \tilde{x}_j z_j^b & & = \min & \displaystyle\sum_{b\in B} \vartheta^b \pi_i^b + \sum_{j=1}^n \sigma_{ij} & \\
\text{s. t.} & \displaystyle\sum_{j=1}^n z_j^b \leq \vartheta^b & \forall b \in B & \text{s. t.} & \pi_i^b + \sigma_{ij} \geq \hat{A}_{ij}^b \tilde{x}_j & \forall j = 1,\ldots,n, \; b \in B \\
& \displaystyle\sum_{b\in B} z_j^b = 1 & \forall j = 1,\ldots,n & & \pi_i^b \geq 0 & \forall b \in B \\
& z_j^b \geq 0 & \forall j = 1,\ldots,n, b \in B & & \sigma_{ij} \text{ free} & \forall j = 1,\ldots,n
\end{array}$$

with dual variables z_j^b and σ_{ij} corresponding to constraints (3.21b) and (3.21c), respectively. Note, in contrast to Büsing and D'Andreagiovanni [48] we do not relax constraint (3.21c) and thus, obtain free dual variables σ. This difference is needed to exploit the information of minimum numbers of realizations in the negative deviation bands. Otherwise, these realizations would be accounted by 0 and not contribute with a negative value to the objective function. Therefore, the dualized subproblem determining the worst-case deviation in [48] is not correct as it does not model the multi-band robust uncertainty set $\mathcal{U}^{\text{mb}}$ but its restriction to positive deviations.

Analogously to the Γ-robust setting, we can use the minimization formulation to reformulate the multi-band robust counterpart of the general LP (3.18) in a compact way. This yields the *compact multi-band robust counterpart*

$$\max \kappa^\top x \tag{3.23a}$$

$$\text{s.t.} \sum_{j=1}^{n} \bar{A}_{ij} x_j + \sum_{b \in B} \vartheta^b \pi_i^b + \sum_{j=1}^{n} \sigma_{ij} \leq d_i \quad \forall i = 1, \dots, m \tag{3.23b}$$

$$\pi_i^b + \sigma_{ij} \geq \hat{A}_{ij}^b x_j \quad \forall i = 1, \dots, m, j = 1, \dots, n, b \in B \tag{3.23c}$$

$$x,\ \pi \geq 0,\ \sigma \text{ free} \tag{3.23d}$$

Notice that we have relaxed the inner min-operator. This can be done since the term tends to its minimum by the objective sense, the objective coefficients of (3.23a) are positive, and the relation sign "≤" of constraint (3.23b). Formulation (3.23) is compact as its size is polynomial in the size of the non-robust LP (3.18): it contains $m(|B|+n)$ additional variables and $mn|B|$ additional constraints. Furthermore, the compact multi-band robust counterpart of an LP is again linear and thus does not increase the computational complexity.

Alternatively to the compact counterpart, a separation approach can be followed by solving the non-robust original LP (3.18) and separating violated model constraints (3.19b). These inequalities are called *robustness cuts*. They ensure the feasibility with respect to the uncertain data; cf. Büsing and D'Andreagiovanni [47].

3.2.3 Probabilistic analysis

Since multi-band robustness is a generalization of Γ-robustness, we expect similar probabilistic results on the feasibility of a constraint containing uncertain coefficient in the multi-band robust counterpart. Both, Mattia [122] and Büsing and D'Andreagiovanni [48] present theoretical bounds on the constraint violation based on probabilistic analysis. These bounds seem rather rough and less handy than the (already complicated) bounds in [34] for the special case of Γ-robustness. We refer to Büsing and D'Andreagiovanni [48] for further details as it is not in the focus of this work.

3.3 Submodular robustness

Now we investigate a more general robustness concept where the constraints of the robust counterparts can be described by submodular functions resulting in so-called submodular knapsack constraints (in Chapter 6, we will investigate the submodular knapsack problem itself). The concept of submodular robustness generalizes the Γ-robustness and multi-band robustness concepts. The underlying uncertainty set of this general robustness concept is the polymatroid of the corresponding submodular function.

Polymatroids have been introduced by Edmonds [66] and have been studied extensively; for example, see Frank and Tardos [70] and the references therein. One particular interesting result shown by Edmonds is that a linear function can efficiently be optimized over a polymatroid by using a greedy algorithm. Although many studies exist, to our knowledge polymatroids have not been related to robust optimization except for the work by Atamtürk and Narayanan [15] on mean-risk minimization where submodular functions and polymatroids are considered for a stochastic optimization problem. A relation to discrete robust optimization and Γ-robustness is drafted for a special case of this mean-risk minimization problem. Investigations of polymatroidal uncertainty sets have not been published. The submodular robust counterpart includes a submodular knapsack constraint. The related submodular knapsack problem has been introduced by Atamtürk and Narayanan [16]. Their investigation includes mathematical formulations of the problem, basic properties of the corresponding polyhedron, and the class of extended submodular robust cover inequalities.

3.3.1 The concept of submodular robustness

Let us recall submodularity: Given a base set $N := \{1, \dots, n\}$, a function $f : 2^N \to \mathbb{R}_{\geq 0}$ is called *submodular* if

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

holds for all $X, Y \subseteq N$. Alternatively, f is submodular if for all $X, Y \subseteq N$ with $X \subseteq Y$ and $j \in N \setminus Y$

$$f(X \cup \{j\}) - f(X) \geq f(Y \cup \{j\}) - f(Y)$$

holds, or — equivalently — if

$$f(X \cup \{j_1\}) + f(X \cup \{j_2\}) \geq f(X \cup \{j_1, j_2\}) + f(X)$$

is satisfied for all $X \subseteq N$ and $j_1, j_2 \in N \setminus X$. Note that we will sometimes abuse the notation and write $f(x)$ instead of the correct set notation $f(\{j \in N \; : \; x_j = 1\})$ if the vector $x \in \{0,1\}^n$ acts as an incidence vector of subsets of N.

Definition 3.10 (Submodular Robust Uncertainty Set)**.** Let $N := \{1, \dots, n\}$ be a base set and $f : 2^N \to \mathbb{R}_{\geq 0}$ be submodular. Then the *submodular robust uncertainty set* $\mathcal{U}^f$ is defined as follows

$$\mathcal{U}^f := \left\{ u \in \mathbb{R}^n \; : \; \sum_{j \in X} u_j \leq f(X) \text{ for all } X \subseteq N \right\}. \tag{3.24}$$

The set $\mathcal{U}^f$ is the *polymatroid of* f.

Example 3.11. *Let us consider the function* $f^{\Gamma 1} : 2^N \to \mathbb{R}_{\geq 0}$ *defined by* $f^{\Gamma 1}(X) := \sum_{j \in X} \bar{u}_j + \max_{j \in X} \hat{u}_j$ *for some vectors* $\bar{u}, \hat{u} \in \mathbb{R}^n$.

Next, we show the submodularity of $f^{\Gamma 1}$. *Therefore we consider the following transformations and relaxations.*

$$\begin{aligned}
& f^{\Gamma 1}(X) + f^{\Gamma 1}(Y) \\
=& \sum_{j \in X} \bar{u}_j + \max_{j \in X} \hat{u}_j + \sum_{j \in Y} \bar{u}_j + \max_{j \in Y} \hat{u}_j \\
=& \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \max_{j \in X} \hat{u}_j + \max_{j \in Y} \hat{u}_j \\
=& \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \max\{\max_{j \in X} \hat{u}_j, \max_{j \in Y} \hat{u}_j\} + \min\{\max_{j \in X} \hat{u}_j, \max_{j \in Y} \hat{u}_j\} \\
\geq& \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \max_{j \in X \cup Y} \hat{u}_j + \max_{j \in X \cap Y} \hat{u}_j \\
=& f^{\Gamma 1}(X \cup Y) + f^{\Gamma 1}(X \cap Y)
\end{aligned}$$

Hence the function $f^{\Gamma 1}$ *is submodular. The corresponding submodular robust uncertainty set* $\mathcal{U}^{f^{\Gamma 1}}$ *is*

$$\mathcal{U}^{f^{\Gamma 1}} := \left\{ u \in \mathbb{R}^{|N|} \ : \ \sum_{j \in X} u_j \leq \sum_{j \in X} \bar{u}_j + \max_{j \in X} \hat{u}_j \right\}$$

which is the Γ*-robust uncertainty set* $\mathcal{U}^{\Gamma}$ *where* $\bar{u}$ *is the nominal and* $\hat{u}$ *the deviation data vector and* $\Gamma = 1$*; cf. Section 3.1.*

3.3.2 The submodular robust counterpart

Let $\kappa \in \mathbb{R}^n_{>0}$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. Consider a binary ILP in its maximization form

$$\max \ \kappa^\top x \tag{3.25a}$$
$$\text{s.t.} \ A_{i \cdot} x \leq b_i \qquad \forall i = 1, \ldots, m \tag{3.25b}$$
$$x \in \{0,1\}^n. \tag{3.25c}$$

Given a submodular function $f : 2^N \to \mathbb{R}_{\geq 0}$, its submodular robust counterpart reads

$$\max \ \kappa^\top x \tag{3.26a}$$
$$\text{s.t.} \ A_{i \cdot} x \leq b_i \qquad \forall A_{i \cdot}^\top \in \mathcal{U}_i^f, i = 1, \ldots, m \tag{3.26b}$$
$$x \in \{0,1\}^n. \tag{3.26c}$$

Note, there are exponential many constraints (3.26b). We can rewrite (3.26) using a polynomial number of constraints as

$$\max \ \kappa^\top x \tag{3.27a}$$
$$\text{s.t.} \ \max_{A_{i \cdot}^\top \in \mathcal{U}_i^f} A_{i \cdot} x \leq b_i \qquad \forall i = 1, \ldots, m \tag{3.27b}$$
$$x \in \{0,1\}^n \tag{3.27c}$$

with only m constraints (3.27b). Formulation (3.27) is not linear. By definition of $\mathcal{U}^f$ the following holds.

$$\begin{aligned}\mathcal{U}^f &= \left\{u \in \mathbb{R}^n \;:\; \sum\nolimits_{j\in X} u_j \leq f(X) \;\forall X \subseteq N\right\}\\ &= \left\{u \in \mathbb{R}^n \;:\; u^\top x \leq f(x) \;\forall x \in \{0,1\}^n\right\}\end{aligned}$$

Hence, $u^\top x \leq f(x)\ \forall u \in \mathcal{U}^f$ also holds and implies $\max_{u\in\mathcal{U}^f} u^\top x = f(x)$. Setting $u = A_{i\cdot}$ and $\mathcal{U}^f = \mathcal{U}_i^f$, we can reformulate (3.27) as

$$\max\ \kappa^\top x \tag{3.28a}$$
$$\text{s.t. } f(x) \leq b_i \qquad \forall i = 1,\dots,m \tag{3.28b}$$
$$x \in \{0,1\}^n \tag{3.28c}$$

using the submodular function f. Notice, the resulting submodular robust counterpart (3.28) does not include the coefficient matrix A anymore. Instead the constraints are solely defined by the submodular function whose definition has to capture the intended restrictions modeled by the original coefficient matrix.

3.3.3 Submodular functions: Γ- and multi-band robustness

In this section, we state selected submodular functions including those describing the robustness concepts Γ-robustness and multi-band robustness. Therefore, we consider the ILP 3.25 and define $X := \{i \in N \;:\; x_i = 1\} \subseteq N$.

Let $\bar{u}$, $\hat{u} \in \mathbb{R}^n$, $\Gamma \in \{0,\dots,n\}$ be the vector of nominal values, the vector of deviation values, and the robustness parameter as defined for the concept of Γ-robustness. We define

$$f^\Gamma(X) := \sum_{j\in X} \bar{u}_j + \max_{\substack{X'\subseteq X\\ |X'|\leq\Gamma}} \sum_{j\in X'} \hat{u}_j.$$

Then constraint (3.28b) of the submodular robust counterpart with $f = f^\Gamma$ is a formulation of the corresponding constraint (3.4b) of the Γ-robust counterpart.

Lemma 3.12. *The function f^Γ as defined above is submodular.*

Proof. Define $\hat{v}_j^S$ as the j-th largest element in set $S \subseteq N$ w.r.t. $\hat{u}$ and with $\hat{v}_j^S = 0$ for $j > |S|$. For $S_1, S_2 \subseteq N$, it clearly holds $\hat{v}_j^{S_1\cup S_2} = \max\{\hat{v}_j^{S_1}, \hat{v}_j^{S_2}\}$, $\hat{v}_j^{S_1\cap S_2} \leq \hat{v}_j^{S_1}$, $\hat{v}_j^{S_1\cap S_2} \leq \hat{v}_j^{S_2}$, and thus $\hat{v}_j^{S_1\cap S_2} \leq \min\{\hat{v}_j^{S_1}, \hat{v}_j^{S_2}\}$. Now, let $X, Y \subseteq N$. Considering f^Γ, the following holds.

$$\begin{aligned}
f^{\Gamma}(X) + f^{\Gamma}(Y) &= \sum_{j \in X} \bar{u}_j + \max_{\substack{X' \subseteq X \\ |X'| \leq \Gamma}} \sum_{j \in X'} \hat{u}_j + \sum_{j \in Y} \bar{u}_j + \max_{\substack{Y' \subseteq Y \\ |Y'| \leq \Gamma}} \sum_{j \in Y'} \hat{u}_j \\
&= \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \max_{\substack{X' \subseteq X \\ |X'| \leq \Gamma}} \sum_{j \in X'} \hat{u}_j + \max_{\substack{Y' \subseteq Y \\ |Y'| \leq \Gamma}} \sum_{j \in Y'} \hat{u}_j \\
&= \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \sum_{j=1}^{\Gamma} \hat{v}_j^X + \sum_{j=1}^{\Gamma} \hat{v}_j^Y \\
&= \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \sum_{j=1}^{\Gamma} \max\{\hat{v}_j^X, \hat{v}_j^Y\} + \sum_{j=1}^{\Gamma} \min\{\hat{v}_j^X, \hat{v}_j^Y\} \\
&\geq \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \sum_{j=1}^{\Gamma} \hat{v}_j^{X \cup Y} + \sum_{j=1}^{\Gamma} \hat{v}_j^{X \cap Y} \\
&= \sum_{j \in X \cup Y} \bar{u}_j + \sum_{j \in X \cap Y} \bar{u}_j + \max_{\substack{Z' \subseteq X \cup Y \\ |Z'| \leq \Gamma}} \sum_{j \in Z'} \hat{u}_j + \max_{\substack{Z' \subseteq X \cap Y \\ |Z'| \leq \Gamma}} \sum_{j \in Z'} \hat{u}_j \\
&= f^{\Gamma}(X \cup Y) + f^{\Gamma}(X \cap Y)
\end{aligned}$$

Hence, f^{Γ} is submodular. □

Let B, ϑ^b, $\bar{u}$, $\hat{u}^b$ be the set bands, the number of realizations in band $b \in B$ in the worst-case, the vector of nominal, and the vector of deviation values for band $b \in B$ as defined for the concept of multi-band robustness. Further let $(X^b)_{b \in B}$ be a partition of X. We define

$$f^{\mathrm{mb}}(X) := \sum_{j \in X} \bar{u}_j + \max_{\substack{X = \bigcup X^b \\ |X^b| = \vartheta^b}} \sum_{b \in B} \sum_{j \in X^b} \hat{u}_j^b.$$

Then constraint (3.28b) of the submodular robust counterpart with $f = f^{\mathrm{mb}}$ is a formulation of the corresponding constraint (3.20b) of the multi-band robust counterpart.

Lemma 3.13. *The function f^{mb} as defined above is submodular.*

Proof. Define $\hat{v}_j^{b,S}$ as the j-th largest element in set $S \subseteq X$ w.r.t. $\hat{u}^b$ and with $\hat{v}_j^{b,S} = 0$ for $j > |S|$. Then the proof is analogous to the proof of Lemma 3.12. □

3.4 Recoverable robustness

Recoverable robustness is a recent two-stage approach to optimization under uncertainty and can be seen as a deterministic alternative to stochastic programming with limited recourse. It has been introduced in the context of robust railway optimization by Liebchen et al. [110] in 2009.

The concept of recoverable robustness can be described as follows: after the first-stage decision the realization of the uncertain data is observed. Then, the previous decision may be altered according to a given adjustment rule taking the data realization into account. This second-stage adjustment is called recovery as the first-stage decision may become infeasible due to the data realization. If there exists a realization such that the first-stage solution cannot be recovered, then the recoverable robust problem has no solution. Both, the first stage decision and its second stage adjustment inflict costs. An optimal recoverable robust solution minimizes the overall costs, i. e., the first stage costs and the worst-case second stage costs.

We already mentioned that recoverable robustness can be seen as a robust optimization version of stochastic optimization with (limited) recourse. In contrast to stochastic optimization with (limited) recourse, recoverable robustness does not optimize the costs in expectation but the first-stage costs plus the worst-case recovery costs. It does not depend on the knowledge of the underlying probability distribution.

In fact, the concept of recoverable robustness is a rather general framework which allows some flexibility in the structure of the uncertainty. Each possible realization of uncertain data is called a scenario. All scenarios are (implicitly) given in the set of scenarios. Several types of sets of scenarios for recoverable robust problems have been studied by different authors, e.g., discrete scenarios [49], interval scenarios [49], Γ-robust scenarios [50], and others [46, 110]. Scenario sets are related to uncertainty sets in robust optimization: each scenario is associated to a single realization of uncertain data. Thus, each scenario set induces a corresponding uncertainty set.

3.4.1 The concept of recoverable robustness

Let us define the concept of recoverable robustness more formally.

Definition 3.14 (Recoverable Robustness Concept)**.** Let $\mathcal{S}$ be the *set of scenarios* of uncertain data whose realizations define the *recoverable robust uncertainty set* $\mathcal{U}^{\mathcal{S}}$. Further, let the optimization problem Π consist of the following two stages:

1. a first-stage decision before the realization of the uncertain data

2. a second-stage decision modifying the first-stage decision according to a recovery rule and possibly recovering the feasibility of the first-stage decision according to the data realization.

We call Π a *recoverable robust optimization problem.* Furthermore, $\mathcal{R}(x^0)$ is the set of all values x^S, which are feasible as second-stage decision with respect to the first-stage decision x^0. The set $\mathcal{R}(x^0)$ is called *recovery set (of* x^0*).* Therefore it (implicitly) defines the so-called *recovery rule* which characterizes the adaptability of the second stage decision from the first stage decision.

In general, the scenario set (and thus the corresponding uncertainty set) can be chosen arbitrarily. For example, the set of all realizations of any of the uncertainty sets defined in Section 2.2.1 can be used. Some explicit examples of scenario sets are the following three:

discrete scenarios The set $\mathcal{S}_D$ of discrete scenarios defines an explicit vector for each possible realization of the uncertain data. The corresponding uncertainty set $\mathcal{U}^{\mathcal{S}_D}$ consists of all realizations.

interval scenarios The set $\mathcal{S}_I$ of interval scenarios defines intervals for each uncertain data such that all realizations are within these intervals. The corresponding uncertainty set is denoted by $\mathcal{U}^{\mathcal{S}_I}$.

Γ-scenarios The set $\mathcal{S}_\Gamma$ of Γ-scenarios defines the data uncertainty according to the Γ-robustness concept; cf. Section 3.1. The corresponding uncertainty set is denoted by $\mathcal{U}^{\mathcal{S}_\Gamma}$.

3.4.2 The recoverable robust counterpart

Let $\kappa \in \mathbb{R}^n_{>0}$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. Consider a general LP in its maximization form

$$\max \kappa^\top x \tag{3.29a}$$

$$\text{s. t. } A_{i\cdot} x \le b_i \qquad \forall i = 1, \dots, m \tag{3.29b}$$

$$x \ge 0. \tag{3.29c}$$

To model the two-stages of the recoverable robustness concept some more notation is needed. Let the first-stage be modeled by $\kappa^0 = \kappa$, $A^0 = A$, and $b^0 = b$. Further let $\kappa^S \in \mathbb{R}^n_{>0}$, $A^S \in \mathbb{R}^{m^S \times n}$, and $b^S \in \mathbb{R}^{m^S}$ model the second-stage for scenario s. Then, the recoverable robust counterpart of LP (3.29) can be formulated as

$$\max \left({\kappa^0}^\top x^0 + \min_{S \in \mathcal{S}} {\kappa^S}^\top x^S \right) \tag{3.30a}$$

$$\text{s. t. } A^0_{i\cdot} x^0 \le b^0_i \qquad \forall i = 1, \dots, m^0 \tag{3.30b}$$

$$A^S_{i\cdot} x^S \le b^S_i \qquad \forall {A^S_{i\cdot}}^\top \in \mathcal{U}^{\mathcal{S}},\ i = 1, \dots, m^S \tag{3.30c}$$

$$x^S \in \mathcal{R}(x^0) \qquad \forall S \in \mathcal{S} \tag{3.30d}$$

$$x^0, x^S \ge 0 \qquad \forall S \in \mathcal{S} \tag{3.30e}$$

where $\mathcal{R}(x^0)$ is the set of all solutions recoverable from x^0. The objective (3.30a) maximizes total first-stage profit plus the worst-case maximum second-stage profit. Constraint (3.30b) models the first, (3.30c) the second-stage constraints. The recovery rule is enforced by (3.30d), non-negativity is ensured by (3.30e).

The size of this formulation highly depends on the formulation of the uncertainty set $\mathcal{U}^{\mathcal{S}}$ and the formulation of the recovery rule $x^S \in \mathcal{R}(x^0)$.

Example 3.15. *Let us consider a LP with uncertain input data and determine its recoverable robust counterpart. Therefore we have to decide on a scenario set and recovery rule. Let us consider the discrete scenario set $\mathcal{S}_D$ and the following recovery rule: for each scenario the sum of second-stage decision variables may change only up to 25 % to the sum of first-stage decision variables. Furthermore, let only the first-stage* *decision contribute to the objective value. The resulting recoverable robust counterpart can be formulated as*

$$\max \sum_{j=1}^{n} \kappa_j^0 x_j^0 \tag{3.31a}$$

$$\text{s. t. } A_{i\cdot}^0 x^0 \leq b_i^0 \qquad \forall i = 1, \ldots, m^0 \tag{3.31b}$$

$$A_{i\cdot}^S x^S \leq b_i^S \qquad \forall {A^S_{i\cdot}}^\top \in \mathcal{U}^{\mathcal{S}_D},\ i = 1, \ldots, m^S \tag{3.31c}$$

$$(1 - 0.25) \sum_{j=1}^{n} x_j^0 \leq \sum_{j=1}^{n} x_j^S \qquad \forall S \in \mathcal{S}_D \tag{3.31d}$$

$$\sum_{j=1}^{n} x_j^S \leq (1 + 0.25) \sum_{j=1}^{n} x_j^0 \qquad \forall S \in \mathcal{S}_D \tag{3.31e}$$

$$x^0, x^S \geq 0 \qquad \forall S \in \mathcal{S}_D \tag{3.31f}$$

3.5 Evaluation of robustness

In the previous sections, we have considered different approaches to take data uncertainty into account when modeling and solving optimization problems. We have focused on robust optimization and presented different robustness concepts to define the uncertainty set modeling the data uncertainty. Although we have presented theoretical bounds on the probability that a Γ-robust constraint is violated, we have excluded the question how to evaluate robustness and compare different robustness concepts in practice. Of course, this question is important in practice as a basis for the decision which robustness concept should be applied for a given application. Moreover, each presented robustness concept also includes a wide range of parameters whose settings have to be determined and tuned according to the actual problem, e. g., the nominal and deviation data in Γ-robustness or multi-band robustness, or the recovery rule of recoverable robustness. We address the second problem of parameter determination at the end of this section.

In the following, we present several measures to evaluate robustness, compare the concepts, or even different parameter settings. Considering a robust optimization problem following a certain robustness concept, there exist two key questions to evaluate its quality:

1. What is the ratio of realizations of uncertain data in the underlying real-world application that is actually modeled by the uncertainty set? – Or shorter: How robust is the model or a specific robust solution?

2. What is the change in the objective value compared to the objective value of the corresponding non-robust model or solution? – Or shorter: How much does the given robustness cost?

Let us start with the second question. Bertsimas and Sim [34] coined the phrase *price of robustness* to describe the increase in the objective function value by introducing

Γ-robustness to the nominal non-robust model. It is associated with the fact that oftentimes the objective models total costs to be minimized and Γ-robustness requires additional (costly) resources to guarantee feasibility for the worst-case data realization of the uncertainty set. Formally, let z be the objective value of a non-robust optimization problem with data vector u. Let z^Γ be the objective value of the Γ-robust counterpart with nominal data vector $\bar{u} = u$, deviation data vector $\hat{u}$ and robustness parameter Γ. Then the *price of robustness* ($PoR(\Gamma)$) is defined by

$$PoR(\Gamma) := \frac{z^\Gamma}{z}.$$

This measure is quite intuitive and can easily be used as decision support in practice, e. g., it allows statements like "the Γ-robust counterpart of this problem is 60 % more expensive for $\Gamma = 12$". Nevertheless, there exists also one major flaw: to determine the PoR the objective is compared with the non-robust problem where the data vector is the nominal data vector of the robust problem. Clearly, the nominal problem does not sufficiently reflect the real problem under data uncertainty; otherwise the consideration of optimization under data uncertainty and robust optimization in particular would not be necessary at all. Hence, when determining the PoR, the comparison should be with a non-robust setting that takes the worst-case data realization into account, e. g., the data vector $\bar{u} + \hat{u}$. This view is very conservative and the resulting solution is expensive. The corresponding robust solution is expected to be less conservative, more resource-efficient, and thus less expensive. This way the ratio between z^Γ and z is less than 1 and the PoR becomes rather a "gain of robustness" describing the relative cost saving according to the most-conservative setting. In summary, both approaches to determine the PoR are plausible, give meaningful measures and answer our question about determining the costs of a robust solution.

Note that for the concept of recoverable robustness we also use the term *gain of recovery* instead of price of robustness. This relates to the fact that this concept allows first-stage solutions to become infeasible as long as they can be recovered in the second stage. Thus, the first-stage constraints are "softer" as in the setting without recovery. This "relaxation" of constraints yields a better optimal value and the idea behind the term gain of recovery.

Next we consider the question how to quantify robustness of a given a solution of a robust problem. Let $\mathcal{R}$ be the set of all possible realizations of the uncertain data. Let $\mathcal{R}^{\text{feas}} \subseteq \mathcal{R}$ be the subset of realizations for which the given solution is still feasible. The ratio

$$\frac{|\mathcal{R}^{\text{feas}}|}{|\mathcal{R}|}$$

denotes the *(exact) robustness* of this solution. In theory, the sets $\mathcal{R}$ or $\mathcal{R}^{\text{feas}}$ may be of infinite cardinality. But in practice, oftentimes only a finite sample or historical data is known and can be used as set $\mathcal{R}$. Then $\mathcal{R}^{\text{feas}}$ is finite as well and the robustness is well-defined.

Notice that the binary evaluation whether the given solution is feasible or not for a certain realization, may be too restrictive in practice. Instead the level of (in)feasibility should be determined. Let us illustrate this idea by an example:

Example 3.16. *When considering a multi-commodity flow problem with uncertain demands, a robust flow may become infeasible with respect to a specific realization vector just because of one entry in this vector, i. e., because of one deviating demand value of a single commodity while it is still feasible for all other commodities. In this case it would be better to regard this as "almost feasible" rather than infeasible.*

To obtain a well-defined measure of "almost feasibility", we propose the following: determine the largest α, $0 \leq \alpha \leq 1$ such that the solution is feasible for the realization vector scaled by α. Hence for a specific realization vector, $\alpha = 1$ corresponds to a feasible solution and $0 \leq \alpha < 1$ to an infeasible one. But the actual value of α gives a measure of the level of (in)feasibility. In a second step the average, minimum, or maximum value of α for all realizations in $\mathcal{R}$ can be determined to give a measure of the overall robustness independent of specific realizations.

In Koster et al. [103] and Koster and Kutschka [101], we have considered a wide range of different robustness measures in our work on (Γ-)robust network design problems in telecommunications. In this setting, data uncertainty is given as the uncertain future demands between the network nodes. Most of the times historical data is available as snapshots of measurements of the actual point-to-point traffic. Each snapshot is called a traffic matrix. In Koster et al. [101, 103], we propose several robustness measures for a robust network design $\mathcal{D}$ w.r.t a given set of traffic matrices $\mathcal{M}$. For example, we define the realized robustness (w.r.t. the supported traffic matrices) as follows

$$r_{\text{suppTM}}(\mathcal{D}, \mathcal{M}) := \frac{|\mathcal{M}_{\text{supp}}(\mathcal{D})|}{\mathcal{M}} \tag{3.32}$$

where $\mathcal{M}_{\text{supp}}(\mathcal{D}) \subseteq \mathcal{M}$ denotes the subset of traffic matrices which can be completely routed using the routing template and link dimensions of the robust network design $\mathcal{D}$. This is a rather conservative measure as it disqualifies a traffic matrix as soon as it cannot support a single demand completely regardless the total network load. However, it gives an idea of robustness measures in practice.

To emphasize that these measures are comparing the robustness of a given solution with the set of historical traffic matrices, we call them *realized robustness (measures)*. These different realized robustness measures reflect different modeling intentions and assumed levels of conservatism. Hence, the actual values vary depending on the used measure and have to be interpreted by the decision maker according to the original real-world application.

Furthermore, a single realized robustness value averaging the robustness values of each individual traffic matrix, etc. may not be sufficient as decision support. In particular if detailed historical data is available, the following approach yields a more meaningful evaluation of realized robustness. We propose to use the so-called *robustness profile* which visualizes bar-code-like the robustness values of all traffic matrices of the evaluated

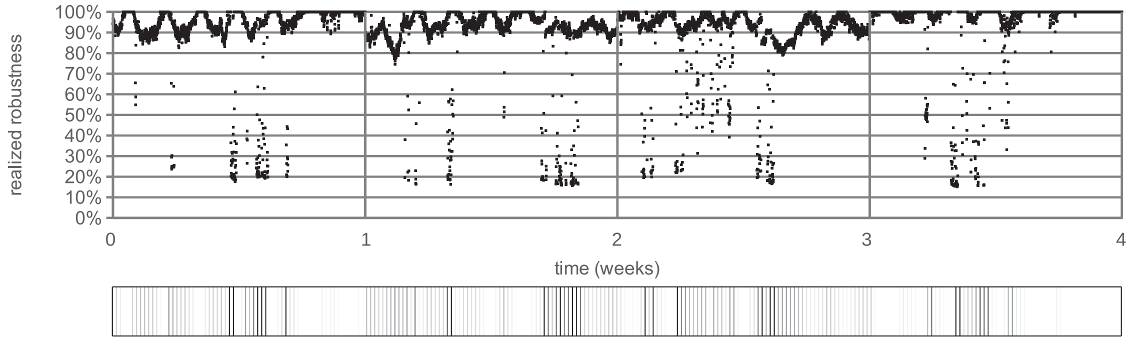


Figure 3.3: Example of robustness profile. Given a set of historical data spanning four weeks, the realized robustness value for each individual data is shown in the diagram on top. Below the corresponding robustness profile is drawn. Its bars are darker the lower the related realized robustness value $0 \leq r(d) \leq 1$ is; compare with diagram on top.

historical data. More precisely, let a robust solution, a data set $\mathcal{H}$ of available historical data, and a realized robustness value $r(d) \in [0, 1]$ for each data $d \in \mathcal{H}$ be given. Then the robustness profile visualizes each individual $r(d)$ value as a vertical bar of some height. The color of such a bar is gray-scaled and determined by the value $r(d)$ reaching from white ($r(d) = 1$) to black ($r(d) = 0$). Hence, profiles with fewer and lighter bars are better.

Figure 3.3 illustrates the construction of a robustness profile. On top the realized robustness values $r(d)$ of some four-week spanning historical data set $\mathcal{H}$ are shown in a diagram. Below the corresponding robustness profile is visualized. Notice that the bars are darker the lower the corresponding $r(d)$ value is in the diagram above.

Each robustness concept comes with its own set of robustness parameters, e. g., the Γ parameter of Γ-robustness, the number of bands and their bounds in the setting of multi-band robustness, or the choice of scenario sets and recovery rule and their parametrization for recoverable robustness. Unfortunately, there does not exist a general rule how to set these parameters. It usually depends on the application and the available historical data. Nevertheless, determining the right parameter settings is crucial.

Therefore, we propose the following approach in practice. Although the theoretical probability bounds may give some suggestions on the order of the parameter values, they are usually less useful in practice. Hence, we propose to utilize the robustness measures we described above. Given a set of historical data and some (preliminary) robust solution for an (arbitrary) parameter setting, we can evaluate its realized robustness using some of the previously presented measures (or a better suited new one). In addition, we can also evaluate the cost of robustness of this solution. Hence the trade-off between robustness and costs can be determined. By repeating this analysis for different parameter settings, we can do a practical sensitivity analysis of a (reasonable) subset of parameter settings. Furthermore, pareto-optimal settings can be identified. Then a decision maker can use

these pareto-optimal settings as decision support to determine the most robust setting given a certain cost budget
or — vice- versa — the cost-minimal parameter setting realizing a requested minimal robustness level. Of course, this approach is biased as only historical data is used. But in many real-world applications it is suitable to assume that the same unknown source of uncertainty behaves similarly in trend in the future as it has in the past. In Koster et al. [105] and Koster and Kutschka [101] we have successfully applied this approach to determine the robustness parameters of Γ-robust network design of telecommunication networks using real-world historical data of research backbone networks.

PART TWO

ROBUST KNAPSACK PROBLEMS

CHAPTER FOUR

THE Γ-ROBUST KNAPSACK PROBLEM

In this chapter, we consider the (binary) knapsack problem under data uncertainty. Here we use the Γ-robust uncertainty concept and the corresponding uncertainty set $\mathcal{U}^\Gamma$ to model the uncertain data. The resulting problem is called the (Γ-)robust knapsack problem (Γ-RKP) and has been investigated by different authors. Bertsimas and Sim [33, 34] have given the Γ-RKP as an example when introducing the concept of Γ-robustness and evaluating the price of robustness. In 2008, Klopfenstein and Nace [96] have considered the chance-constraint knapsack problem and the Γ-RKP as a possible implementation of it pointing out the relation between feasible and optimal solutions of these two problems. A work focusing on the Γ-RKP and the properties of the related polyhedron has been published by the same authors [97] in 2012. They present results on the dimension of the polytope, introduce the classes of Γ-robust cover inequalities and extended Γ-robust cover inequalities, proof that these classes are facet-defining under some conditions, and present an exact ILP-based and a heuristic separation algorithm. Monaci and Pferschy [124] consider the solution of Γ-RKP: they investigate the worst-case price of robustness and performance of greedy algorithms to solve Γ-RKP heuristically. Moreover, let N denote the set of items, they characterize the optimal solution of the fractional Γ-RKP and give a $\mathcal{O}(|N| \cdot \log |\, gkpitemset|)$ algorithm for a special case. Monaci et al. [125] present an exact solution algorithm for the Γ-RKP using a dynamic program with running time $\mathcal{O}(\Gamma \cdot |N| \cdot c)$ where c denotes the knapsack capacity.

In the following, we investigate the Γ-RKP and its related polyhedron. On the one hand, we state important results focusing on the polyhedral structure. On the other hand, we extend the known results, e. g., by introducing a stronger class of extended cover inequalities.

Definition 4.1 (Γ-Robust Knapsack Problem)**.** Given a set of items N, a knapsack capacity $c \in \mathbb{Z}_{>0}$, a profit function $p : N \to \mathbb{Z}_{>0}$, and a parameter $\Gamma \in \mathbb{Z}_{>0}$. For each item $j \in N$, let the item weight w_j be uncertain such that $w_j \in [\bar{w}_j, \bar{w}_j + \hat{w}_j]$ holds for given nominal weight function $\bar{w} : N \to \mathbb{Z}_{>0}$ and deviation weight function $\hat{w} : N \to \mathbb{Z}_{>0}$.

The Γ-*robust knapsack problem* (Γ-RKP) is to find a subset of items with maximum total profit whose total weight does not exceed the knapsack capacity in any realization of the uncertain weights with at most Γ item weights set to their maximum.

W. l. o. g., we assume integer realizations of the uncertain weights. Note, for $\Gamma = 0$, the Γ-RKP is the classic non-robust knapsack problem; cf. Definition. 1.3. Since the Γ-RKP contains the classic knapsack problem as a special case, its complexity is at least weakly $\mathcal{NP}$-hard. Monaci et al. [125] have shown that the Γ-RKP can be solved in pseudo-polynomial time by dynamic programming.

Example 4.2. *Consider the Γ-robust knapsack problem with four items given by* $N = \{1, \dots, 4\}$, $c = 8$, $\bar{w} = \begin{pmatrix} 3 & 4 & 5 & 4 \end{pmatrix}$, $\hat{w} = \begin{pmatrix} 1 & 1 & 1 & 4 \end{pmatrix}$, $\Gamma = 1$, *and* $p = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$. *An optimal solution is given by the subset* $\{1, 2\} \subseteq N$ *with an optimal solution value (i. e., total profit) of* $1 + 1 = 2$ *and a total weight of* $3 + 4 + 1 = 8$.

4.1 Formulations

The Γ-robust knapsack problem can be formulated as

$$\max \sum_{j \in N} p_j x_j \tag{4.1a}$$

$$\text{s. t.} \sum_{j \in N} \bar{w}_j x_j + \max_{N' \subseteq N : |N'| = \Gamma} \sum_{j \in N'} \hat{w}_j x_j \leq c \tag{4.1b}$$

$$x_j \in \{0, 1\} \qquad \forall j \in N \tag{4.1c}$$

whereas (4.1a) maximizes the profit while satisfying the robust knapsack constraint (4.1b). Note, this formulation is the Γ-robust counterpart of the classic knapsack problem formulation (1.10). The linearization of (4.1) is given by

$$\max \sum_{j \in N} p_j x_j \tag{4.2a}$$

$$\text{s. t.} \sum_{j \in N} \bar{w}_j x_j + \sum_{j \in N'} \hat{w}_j x_j \leq c \qquad \forall N' \subseteq N : |N'| = \Gamma \tag{4.2b}$$

$$x_j \in \{0, 1\} \qquad \forall j \in N. \tag{4.2c}$$

The compact reformulation of (4.1) reads as follows

$$\max \sum_{j \in N} p_j x_j \tag{4.3a}$$

$$\text{s. t.} \sum_{j \in N} \bar{w}_j x_j + \Gamma \pi + \sum_{j \in N} \rho_j \leq c \tag{4.3b}$$

$$\pi + \rho_j \geq \hat{w}_j x_j \qquad \forall j \in N \tag{4.3c}$$

$$x_j \in \{0, 1\}, \ \pi, \rho_j \geq 0 \qquad \forall j \in N \tag{4.3d}$$

with nonnegative dual variables π and ρ and dual constraints (4.3c).

4.2 Polyhedral study

In the following, we will report on the main polyhedral results from literature. We will extend the results on extended cover inequalities by providing a stronger type of extension. We define the corresponding polytope as follows

Definition 4.3 (Γ-Robust Knapsack Polytope)**.** The Γ-*robust knapsack polytope* is defined as

$$\mathcal{K}^\Gamma := \operatorname{conv}\left\{x \in \{0,1\}^{|N|} \;:\; x \text{ satisfies (4.2b)}\right\}.$$

By construction of the compact Γ-robust counterpart

$$\mathcal{K}^\Gamma = \operatorname{conv}\left\{x \in \{0,1\}^{|N|} \;:\; \exists\, \pi, \rho \geq 0 \text{ so that } x, \pi, \rho \text{ satisfy (4.3b)–(4.3c)}\right\}$$

holds.

4.2.1 Basic characteristics

Lemma 4.4 (Klopfenstein and Nace [97])**.** *For $\Gamma \geq 1$, the Γ-robust knapsack polytope $\mathcal{K}^\Gamma$ is full-dimensional if and only if $\bar{w}_j + \hat{w}_j \leq c$ for all $j \in N$.*

In the following, we assume w. l. o. g. that $\mathcal{K}^\Gamma$ is full-dimensional.

Lemma 4.5. *Let $j \in N$. The constraints*

$$x_j \geq 0 \tag{4.4}$$

and

$$x_j \leq 1 \qquad \text{if and only if } \bar{w}_j + \bar{w}_i + \hat{w}_j + \hat{w}_i \leq c \;\; \forall i \in N \setminus \{j\} \tag{4.5}$$

are trivial facets of $\mathcal{K}^\Gamma$.

Proof. For $x_j \geq 0$, the unit vectors e_i $(i \neq j)$ and the zero vector are valid for $\mathcal{K}^\Gamma$, satisfy the non-negativity constraint (4.4) with equality, and are affinely independent.

For $x_j \leq 1$, the unit vector e_j and the vectors e_j+e_i $(i \neq j)$ satisfy the upper bound (4.5) with equality, are feasible for the full-dimensional $\mathcal{K}^\Gamma$, and affinely independent.

Vice versa, let $x_j \leq 1$ be a facet of $\mathcal{K}^\Gamma$. Then there exist $|N|$ affinely independent points $x^\tau \in \mathcal{K}^\Gamma$ on this facet. Suppose there exists $i \in N \setminus \{j\}$ such that $\bar{w}_j + \bar{w}_i + \hat{w}_j + \hat{w}_i > c$ holds. Further suppose there exists a τ' with $x_i^{\tau'} = 1$. Then it follows

$$c < \bar{w}_j + \bar{w}_i + \hat{w}_j + \hat{w}_i \;=\; \bar{w}_j x_j^{\tau'} + \bar{w}_i x_i^{\tau'} + \hat{w}_j x_j^{\tau'} + \hat{w}_i x_i^{\tau'} \leq c$$

a contradiction. Hence, $x_i^{\tau'} = 0$ follows for all values τ'. This implies that x^τ cannot be affinely independent contradicting our first supposition. This completes the proof. □

4.2.2 Valid inequalities

In the following, we will consider selected classes of valid inequalities which are based on subsets of items with specific properties. For example, Γ-robust cover inequalities are such a class.

Similar to the classic knapsack, we define covers, minimal covers, and extensions as follows. A subset of items $\mathcal{C} \subseteq N$ is called a *(Γ-robust-)cover* if there exists a a realization of uncertain weights such that its total weight exceeds the knapsack capacity, i. e.,

$$\sum_{j \in \mathcal{C}} \bar{w}_j + \max_{\mathcal{C}' \subseteq \mathcal{C} : |\mathcal{C}'| = \Gamma} \sum_{j \in \mathcal{C}'} \hat{w}_j > c.$$

A cover is called *minimal* if $\mathcal{C} \setminus \{j\}$ is not a cover for all $j \in \mathcal{C}$. Given a cover $\mathcal{C}$, Klopfenstein and Nace [97] define its *(Γ-robust-)extension* $E(\mathcal{C})$ as

$$E(\mathcal{C}) := \begin{cases} \left\{ j \in N \ : \ \bar{w}_j + \hat{w}_j \geq \max_{i \in \mathcal{C}} (\bar{w}_i + \hat{w}_i) \right\} \cup \mathcal{C}, & \text{if } |\mathcal{C}| \leq \Gamma \\ \left\{ j \in N \ : \ \bar{w}_j \geq \max_{i \in \mathcal{C}} \bar{w}_i \wedge \bar{w}_j + \hat{w}_j \geq \max_{i \in \mathcal{C}} (\bar{w}_i + \hat{w}_i) \right\} \cup \mathcal{C}, & \text{if } |\mathcal{C}| > \Gamma. \end{cases} \tag{4.6}$$

Analogously to the classic knapsack problem, the extension is a cover itself.

We can strengthen this extension by exploiting the structure of an optimal solution of the inner maximum as follows. In the case of the Γ-RKP, a maximum weight subset $N' \subseteq N$ has the weight

$$\sum_{j \in N'} \bar{w}_j + \max_{N'' \subseteq N' : |N''| = \Gamma} \sum_{j \in N''} \hat{w}_j$$

which can be determined easily for a given subset $N' \subseteq N$. Therefore, let $\mathcal{C}$ be a cover with partition $\mathcal{C} = \bar{\mathcal{C}} \cup \hat{\mathcal{C}}$ where $\bar{\mathcal{C}} \cap \hat{\mathcal{C}} = \emptyset$ and $\sum_{j \in \bar{\mathcal{C}}} \bar{w}_j + \sum_{j \in \hat{\mathcal{C}}} \hat{w}_j$ is maximal, i. e., items $j \in \bar{\mathcal{C}}$ take their nominal and items $j \in \hat{\mathcal{C}}$ take their peak weights. Then, the extension $E(\mathcal{C})$ can be strengthen to

$$E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}}) := \left\{ j \in N \ : \ \bar{w}_j \geq \max_{i \in \bar{\mathcal{C}}} \bar{w}_i \wedge \bar{w}_j + \hat{w}_j \geq \max_{i \in \hat{\mathcal{C}}} (\bar{w}_i + \hat{w}_i) \right\} \cup \bar{\mathcal{C}} \cup \hat{\mathcal{C}}. \tag{4.7}$$

We call $E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}})$ the *strengthened (Γ-robust-)extension*. By definition, $E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}})$ is a cover itself.

Lemma 4.6. *Let $\mathcal{C} \subseteq N$ be a cover, $E(\mathcal{C})$ its extension, and $E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}})$ its strengthened extension. Then, the* cover inequality

$$\sum_{j \in \mathcal{C}} x_j \leq |\mathcal{C}| - 1 \tag{4.8}$$

and the extended Γ-robust cover inequalities

$$\sum_{j \in E(\mathcal{C})} x_j \leq |\mathcal{C}| - 1 \tag{4.9}$$

and

$$\sum_{j \in E^+(\bar{\mathcal{C}},\hat{\mathcal{C}})} x_j \leq |\mathcal{C}| - 1 \tag{4.10}$$

are valid for $\mathcal{K}^\Gamma$.

Proof. Klopfenstein and Nace [97] have shown the validity of (4.8) and (4.9).

Suppose the extended cover inequality (4.10) is not valid. Then there exist an integer point $\tilde{x} \in \mathcal{K}^\Gamma \cap \{0,1\}^{|N|}$ (integer since $\mathcal{K}^\Gamma$ is defined as an integer hull) and a strengthened extension $E^+(\bar{\mathcal{C}},\hat{\mathcal{C}})$ for which

$$\sum_{j \in E^+(\bar{\mathcal{C}},\hat{\mathcal{C}})} \tilde{x}_j \geq |\mathcal{C}|$$

holds. Define $N_{E^+} := \left\{ j \in E^+(\bar{\mathcal{C}},\hat{\mathcal{C}}) \;:\; \tilde{x}_j = 1 \right\}$. From the feasibility of $\tilde{x}$ follows

$$\begin{aligned}
c &\geq \sum_{j \in N} \bar{w}_j \tilde{x}_j + \max_{N' \subseteq N : |N'| = \Gamma} \sum_{j \in N'} \hat{w}_j \tilde{x}_j \\
&\geq \sum_{j \in E^+(\bar{\mathcal{C}},\hat{\mathcal{C}})} \bar{w}_j \tilde{x}_j + \max_{N' \subseteq E^+(\bar{\mathcal{C}},\hat{\mathcal{C}}) : |N'| = \Gamma} \sum_{j \in N'} \hat{w}_j \tilde{x}_j \\
&= \sum_{j \in N_{E^+}} \bar{w}_j + \max_{N' \subseteq N_{E^+} : |N'| = \Gamma} \sum_{j \in N'} \hat{w}_j \\
&\geq \min_{N_{\mathcal{C}} \subseteq N_{E^+} : |N_{\mathcal{C}}| = |\mathcal{C}|} \left\{ \sum_{j \in N_{\mathcal{C}}} \bar{w}_j + \max_{N' \subseteq N_{\mathcal{C}} : |N'| = \Gamma} \sum_{j \in N'} \hat{w}_j \right\} \\
&= \sum_{j \in \mathcal{C}} \bar{w}_j + \sum_{j \in \hat{\mathcal{C}}} \hat{w}_j > c.
\end{aligned}$$

This is a contradiction and completes the proof. □

Example 4.7. *The set* $\mathcal{C} = \{1,3\}$ *with total weight* $3+5+1 = 9 > c$ *is a cover for the* Γ*-RKP instance defined in Example 4.2. Furthermore, it is minimal. Although this cover cannot be extended in the sense of Klopfenstein and Nace, i. e.,* $E(\mathcal{C}) = \mathcal{C}$*, there exists a strengthened extension* $E^+(\bar{\mathcal{C}},\hat{\mathcal{C}}) = \{1,3,4\}$ *for* $\bar{\mathcal{C}} = \{1\}$ *and* $\hat{\mathcal{C}} = \{3\}$*. This is the only way to obtain the (strengthened) extended cover inequality* $x_1 + x_3 + x_4 \leq 1$*. It dominates* $x_1 + x_3 + x_4 \leq 2$*, the cover inequality for the cover* $\{1,3,4\}$*. Furthermore, there does not exist any cover such that this strengthened extended cover can be obtained as an normal extension of another cover.*

Example 4.7 demonstrates that stronger inequalities can be obtained by preferring our strengthened extensions to the common extensions by Klopfenstein and Nace.

The conditions under which extended robust cover inequalities are facet defining have been studied by Klopfenstein and Nace [97]. The following results give an example.

Lemma 4.8 (Klopfenstein and Nace [97])**.** *Without loss of generality, suppose that* N *is ordered non-increasingly w. r. t. the item peak weights, i. e.,* $j < i \Rightarrow \bar{w}_j + \hat{w}_j \geq \bar{w}_i + \hat{w}_i$*. Let* $\mathcal{C} = \{j_1, \ldots, j_r\}$ *be a minimal cover with* $j_1 < j_2 < \cdots < j_r$*. If any of the following conditions holds, then* (4.9) *is a facet of* $\mathcal{K}^\Gamma$*:*

1. $\mathcal{C} = N$,

2. $E(\mathcal{C}) = N$, $|\mathcal{C}| \leq \Gamma + 1$, *and* $(\mathcal{C} \setminus \{j_1, j_2\}) \cup \{1\}$ *is not a cover,*

3. $E(\mathcal{C}) = \mathcal{C}$, $|\mathcal{C}| \leq \Gamma$, *and* $(\mathcal{C} \setminus \{j_1\}) \cup \{p\}$ *is not a cover, where* $p := \min\{j \in N \setminus E(\mathcal{C})\}$,

4. $\mathcal{C} \subset E(\mathcal{C}) \subset N$, $|\mathcal{C}| \leq \Gamma$, *and neither* $(\mathcal{C} \setminus \{j_1, j_2\}) \cup \{1\}$ *nor* $(\mathcal{C} \setminus \{j_1\}) \cup \{p\}$ *with* $p := \min\{j \in N \setminus E(\mathcal{C})\}$ *are covers.*

4.3 Algorithms

In this section, we present algorithmic approaches to solve the Γ-RKP. Furthermore, we describe algorithms solving the related separation problems of finding maximally violated Γ-robust (extended) cover inequalities.

4.3.1 Separation of violated Γ-robust (extended) cover inequalities

In Klopfenstein and Nace [97], the authors present two approaches to solve the separation problem for Γ-robust cover inequalities: an exact ILP-based algorithm and a heuristic greedy algorithm. In the following, we will briefly report on these algorithms and refer to [97] for more details. At the end of this section, we present an ILP formulation of the separation problem for Γ-robust strengthened extended cover inequalities. Therefore, let x^* be a fractional LP solution of the Γ-RKP.

Similar to the classic KP, Klopfenstein and Nace [97] formulated separation problem for violated cover inequalities (4.8) as ILP as follows. Let $\bar{y}_j$ and $\hat{y}_j$ defined as binary decision variables so that $\bar{y}_j = 1$ if and only if $j \in \mathcal{C}$, and $\hat{y}_j = 1$ if and only if $j \in \mathcal{C}$ and j is set to its peak weight, respectively. Then the resulting ILP reads.

$$
\begin{aligned}
Z := \min \quad & \sum_{j \in N} (1 - x_j^*) \bar{y}_j && && (4.11a)\\
\text{s.t.} \quad & \sum_{j \in N} (\bar{w}_j \bar{y}_j + \hat{w}_j \hat{y}_j) \geq c + 1 && && (4.11b)\\
& \sum_{j \in N} \hat{y}_j \leq \Gamma && && (4.11c)\\
& \hat{y}_j \leq \bar{y}_j && \forall j \in N && (4.11d)\\
& \bar{y}_j, \hat{y}_j \in \{0, 1\} && \forall j \in N && (4.11e)
\end{aligned}
$$

where the objective function (4.11a) minimizes the feasible of the resulting cover inequality. The Constraint (4.11b) ensures the covering condition whereas constraints (4.11c) and (4.11d) model the Γ-robustness. Finally, the variable ranges are given by (4.11e).

Let $(\bar{y}^*, \hat{y}^*)$ be an optimal solution of (4.11) and define $\mathcal{C} := \{j \in N \,:\, \bar{y}_j^* = 1\}$. Then, the cover inequality for $\mathcal{C}$ is violated if and only if $Z < 1$ holds.

In addition to this exact ILP-based separation algorithm, Klopfenstein and Nace [97] present a greedy heuristic to separate violated Γ-robust cover inequalities (4.8). There, they follow a similar approach as known for the KP. However in contrast to the classic problem, the relative profits of the items are first sorted w.r.t. the peak weights of the items. Then, up to Γ items are selected greedily (while not exceeding the knapsack capacity). Afterwards, the remaining items are resorted w.r.t. the relative profit according to their nominal weight. Then, the items are selected greedily if they still fit into the already partially filled knapsack; cf. Klopfenstein and Nace [97] for further details and computational experiments.

Finally, we present an exact ILP-based separation algorithm to separate violated strengthened extended cover inequalities (4.10). Therefore, let $\bar{y}_j$, $\hat{y}_j$ and α_j be binary decision variables so that $\bar{y}_j = 1$ if and only if $j \in \bar{\mathcal{C}}$, $\bar{y}_j = 1$ if and only if $j \in \hat{\mathcal{C}}$, and $\alpha_j = 1$ if and only if j is added while extending the cover, i. e., $j \in E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}}) \setminus (\bar{\mathcal{C}} \cup \hat{\mathcal{C}})$. Next, we formulate the resulting ILP. It extends ILP (4.11) by incorporating the partition of the cover and the conditions for the strengthened extension. It reads as follows.

$$
\begin{aligned}
Z := \max \; & \sum_{j \in N} (x_j^* - 1)(\bar{y}_j + \hat{y}_j) + \sum_{j \in N} x_j^* \alpha_j && && \text{(4.12a)} \\
\text{s. t.} \; & \sum_{j \in N} \bar{w}_j \bar{y}_j + \sum_{j \in N} (\bar{w}_j + \hat{w}_j) \hat{y}_j \geq c + 1 && && \text{(4.12b)} \\
& \sum_{j \in N} \hat{y}_j \leq \Gamma && && \text{(4.12c)} \\
& \bar{y}_j + \hat{y}_j + \alpha_j \leq 1 && \forall j \in N && \text{(4.12d)} \\
& \bar{y}_i + \alpha_j \leq 1 && \forall j, i \in N \, : \, \bar{w}_i > \bar{w}_j && \text{(4.12e)} \\
& \hat{y}_i + \alpha_j \leq 1 && \forall j, i \in N \, : \, \bar{w}_i + \hat{w}_i > \bar{w}_j + \hat{w}_j && \text{(4.12f)} \\
& \bar{y}_j, \hat{y}_j, \alpha_j \in \{0, 1\} && \forall j \in N && \text{(4.12g)}
\end{aligned}
$$

where the objective function (4.12a) maximizes the violation of the resulting Γ-robust strengthened extended cover inequality. Constraint (4.12b) ensures the covering condition. Constraint (4.12d) guarantees that an item may either be in the cover at their nominal weight, or in the cover at their peak weight, or in the set of items added in the extension, or not in the extended cover at all. At most Γ items may be at their maximum peak; this is modeled by constraint (4.12c). The constraints (4.12e) and (4.12f) model the conditions for an item added into the strengthened extension. Finally, the variable ranges are given by (4.12g).

Let $(\bar{y}^*, \hat{y}^*, \alpha^*)$ be a solution of ILP (4.12). Then define

$$
E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}}) := \left\{ j \in N \; : \; \bar{y}_j^* + \hat{y}_j^* + \alpha_j^* = 1 \right\}
$$

with $\bar{\mathcal{C}} = \left\{ j \in N \; : \; \bar{y}_j^* = 1 \right\}$ and $\hat{\mathcal{C}} = \left\{ j \in N \; : \; \hat{y}_j^* = 1 \right\}$. By construction, the Γ-robust strengthened extended cover inequality corresponding to $E^+(\bar{\mathcal{C}}, \hat{\mathcal{C}})$ is violated if and only if $Z > -1$.

Note that ILP (4.12) is a special case of the separation ILP (7.23) to find violated recoverable robust strengthened extended cover inequalities for the k/Γ-RRKP we will present in Section 7.3.2.

4.3.2 Solving the Γ-RKP

There exist different approaches to solve the Γ-RKP. First, the presented exponential (4.2) or compact ILP formulation (4.3) can be solved by an out-of-the-shelf MIP solver as ILOG Cplex [84]. Further, the solution process may be improved by integrating the separation of violated Γ-robust (extended) cover inequalities using any of the algorithms presented in Section 4.3.1. Klopfenstein and Nace [97] have followed this approach and carried out computational studies.

Second, Fischetti and Monaci [69] suggest to solve Γ-robust problems using robustness cuts, i. e., to solve the deterministic non-robust problem (here: the Γ-RKP using the nominal item weights of the Γ-RKP instance) and add violated cuts ensuring robustness on the fly. Robustness is ensured by cutting off solution vectors that may become infeasible w.r.t. the uncertainty set $\mathcal{U}^{\Gamma}$; cf. Fischetti and Monaci [69] and Monaci et al. [125] who also present experimental results on the usage of robustness cuts for the Γ-RKP.

Third, Monaci et al. [125] provide a dynamic programming approach to solve the Γ-RKP in 2013. Their algorithm has a pseudo-polynomial time complexity of $\mathcal{O}(\Gamma|N|c)$. We refer to Monaci et al. [125] for details on this algorithm, suggested refinements to decrease its recursion space in practice, and computational results.

CHAPTER FIVE

THE MULTI-BAND ROBUST KNAPSACK PROBLEM

The knapsack problem is one of the most fundamental problems in mathematical optimization, it is a reoccurring subproblem to many other optimization problems, and has been studied intensively. Hence, it is obvious to consider this problem and its robust counterpart when investigating and comparing (new) robustness concepts. The Γ-RKP as a robust approach to the knapsack problem under data uncertainty has been well-studied in the literature; cf. Chapter 4. Nevertheless, its quality is bounded by the limitations of the concept of Γ-robustness. To overcome some of these, we have introduced the concept of multi-band robustness in Section 3.2.

In the following, we apply this concept to the knapsack problem to take uncertain data into account. The resulting multi-band robust knapsack problem (mb-RKP) generalizes the Γ-RKP considered in Chapter 4.

So far, the mb-RKP has only been studied implicitly: Büsing and D'Andreagiovanni [47] and Mattia [122] describe the multi-band robust counterpart of a general linear constraint, thus a multi-band robust knapsack constraint. They do not introduce the problem itself nor study its polyhedral structure. When applying the histogram model of Bienstock [37] to wireless network design (WND), D'Andreagiovanni [58] considers one specific constraint of the WND model as a dyadic knapsack constraint under histogram uncertainty (a special case of the mb-RKP) and describes a corresponding cover inequality for this special case without further proof of its validity or investigation of the facial structure of the knapsack polytope.

Definition 5.1 (Multi-Band Robust Knapsack Problem)**.** Given a set of items N, a knapsack capacity $c \in \mathbb{Z}_{>0}$, a profit function $p : N \to \mathbb{Z}_{>0}$, and uncertain item weights w modeled by a multi-band robust uncertainty set $\mathcal{U}^{\text{mb}}$, i. e., $w \in \mathcal{U}^{\text{mb}}$.

The *multi-band robust knapsack problem (mb-RKP)* is to find a subset of items maximizing the total profit whose total weight does not exceed the knapsack constraint for all realizations in $\mathcal{U}^{\text{mb}}$.

Example 5.2. *Let us consider the mb-RKP with four items, two negative and two positive deviation bands, and robustness parameters defined by the following table*

j	$\bar{w}_j$	$\hat{w}_j^{-2}$	$\hat{w}_j^{-1}$	$\hat{w}_j^{0}$	$\hat{w}_j^{1}$	$\hat{w}_j^{2}$	p_j
1	3	-4	-2	0	2	4	1
2	3	-3	-2	0	2	3	1
3	2	-3	-1	0	1	2	1
4	2	-2	-1	0	1	2	1
γ^b	-	0	1	0	1	0	-
Γ^b	-	2	3	4	2	1	-
ϑ^b	-	0	1	0	2	1	-

By its frequency profile $\{\vartheta^b\}_{b\in B}$*, we observe that a maximum-weight realization has one item deviating in band -1, two items in band 1, and one item in band 2. Assume a knapsack capacity of 15. Then the solution* $x_1 = x_2 = x_3 = 1$*,* $x_4 = 0$ *is optimal with objective value 3. See also Section 3.2 for further information on the parameters defining* $\mathcal{U}^{mb}$*.*

Since KP and Γ-RKP are special cases of the mb-RKP its complexity is at least weakly $\mathcal{NP}$-hard. So far, no dynamic program solving the mb-RKP is known in the literature. The development of an applicable dynamic program also depends on a better characterization and in particular a closed formula of worst-case realizations in $\mathcal{U}^{\mathrm{mb}}$ that can be exploited by a dynamic program. The state-of-the-art algorithm based on solving a min-cost flow might not be suited for this application; cf. Büsing and D'Andreagiovanni [47, 48] and Mattia [122].

5.1 Formulations

In this section, we present integer programming formulations for the mb-RKP. For each item $j \in N$, we introduce a binary variable x_j which is 1 if and only if item j is selected to be in the subset of items maximizing the total profit.

First, we present a straight-forward exponential ILP formulation for the mb-RKP. It reads

$$\max \sum_{j\in N} p_j x_j \tag{5.1a}$$

$$\text{s.t.} \sum_{j\in N} \bar{w}_j x_j + \sum_{b\in B}\sum_{j\in N^b} \hat{w}_j^b x_j \le c \qquad \forall N^b \subseteq N, b\in B : N = \bigcup_{b\in B} N^b, |N^b| = \vartheta^b \tag{5.1b}$$

$$x_j \in \{0,1\} \qquad \forall j \in N. \tag{5.1c}$$

The objective (5.1a) maximizes the total profit and the exponentially many knapsack constraints (5.1b) ensure that no worst-case realization in $\mathcal{U}^{\mathrm{mb}}$ exceeds the capacity. The worst-case realizations are characterized using the frequency profile ϑ^b (for $b \in B$).

Second, the following compact ILP formulation of the mb-RKP can be obtained by exploiting LP duality analogously to Section 3.2.2:

$$\max \sum_{j \in N} p_j x_j \tag{5.2a}$$

$$\text{s.t.} \sum_{j \in N} \bar{w}_j x_j + \sum_{b \in B} \vartheta^b \pi^b + \sum_{j \in N} \sigma_j \le c \tag{5.2b}$$

$$\pi^b + \sigma_j \ge \hat{w}_j^b x_j \qquad \forall j \in N, b \in B \tag{5.2c}$$

$$x_j \in \{0,1\},\ \pi^b \ge 0,\ \sigma_j \text{ free} \qquad \forall j \in N, b \in B \tag{5.2d}$$

where (5.2a) maximizes the total profit subject to the multi-band robust knapsack constraint (5.2b) and the dual constraint (5.2c). Variables π and σ are dual variables modeling the deviation; cf. Section 3.2.2.

5.2 Polyhedral study

In this section, we investigate the polyhedral structure of the multi-band robust knapsack polytope. This section starts with the definition of the polytope followed by a report on its basic properties. Afterwards, we consider classes of valid inequalities.

Definition 5.3 (Multi-Band Robust Knapsack Polytope)**.** Given a mb-RKP. The *multi-band robust knapsack polytope* is defined as

$$\mathcal{K}^{\text{mb}} := \text{conv}\left\{x \in \{0,1\}^{|N|} \ : \ x \text{ satisfies } (5.1\text{b})\right\}$$

By construction of the compact multi-band robust counterpart,

$$\mathcal{K}^{\text{mb}} = \text{conv}\left\{x \in \{0,1\}^{|N|} \ : \ \exists\, \pi \ge 0, \sigma \text{ so that } x, \pi, \sigma \text{ satisfy } (5.2\text{b})\text{–}(5.2\text{c})\right\}$$

holds.

5.2.1 Basic characteristics

The investigation of the dimension of $\mathcal{K}^{\text{mb}}$ and its trivial facets is straight-forward. The results are as follows.

Lemma 5.4. *For $\Gamma^{\overline{B}} > 0$, the polytope $\mathcal{K}^{mb}$ is full-dimensional if and only if $\bar{w}_j + \hat{w}_j^{\overline{B}} \le c$ for all $j \in N$.*

Proof. All unit vectors and the zero vector are feasible to $\mathcal{K}^{\text{mb}}$ and affinely independent.

Let $\mathcal{K}^{\text{mb}}$ be full-dimensional. Assume there exists a $i \in N$ such that $\bar{w}_j + \hat{w}_j^{\overline{B}} > c$. Then $e_i \notin \mathcal{K}^{\text{mb}}$. Furthermore, by definition of $\mathcal{K}^{\text{mb}}$ it holds $x_i = 0$ for all $x \in \mathcal{K}^{\text{mb}}$. Thus, $\mathcal{K}^{\text{mb}}$ is not full-dimensional; a contradiction completing the proof. □

W. l. o. g. we assume $\mathcal{K}^{\text{mb}}$ is full-dimensional in the following.

Lemma 5.5. *Let $j \in N$. The constraints*

$$x_j \geq 0 \tag{5.3}$$

and

$$x_j \leq 1 \qquad \text{if and only if } \bar{w}_j + \bar{w}_i + \hat{w}_j^{\overline{B}} + \hat{w}_i^{\overline{B}} \leq c \;\; \forall i \in N \setminus \{j\} \tag{5.4}$$

are trivial facets of $\mathcal{K}^{mb}$.

Proof. For $x_j \geq 0$, the unit vectors e_i $(i \neq j)$ and the zero vector are valid for $\mathcal{K}^{\text{mb}}$, satisfy the non-negativity constraint (5.3) with equality, and are affinely independent.

For $x_j \leq 1$, the unit vector e_j and the vectors $e_j + e_i$ $(i \neq j)$ satisfy the upper bound (5.4) with equality, are feasible for the full-dimensional $\mathcal{K}^{\text{mb}}$, and affinely independent. Vice versa, let $x_j \leq 1$ be a facet of $\mathcal{K}^{\text{mb}}$. Then there exist $|N|$ affinely independent points $x^\tau \in \mathcal{K}^{\text{mb}}$ on this facet. Suppose there exists $i \in N \setminus \{j\}$ such that $\bar{w}_j + \bar{w}_i + \hat{w}_j^{\overline{B}} + \hat{w}_i^{\overline{B}} > c$ holds. Further suppose there exists a τ' with $x_i^{\tau'} = 1$. Then it follows

$$\begin{aligned} c < \bar{w}_j + \bar{w}_i + \hat{w}_j^{\overline{B}} + \hat{w}_i^{\overline{B}} &= \bar{w}_j x_j^{\tau'} + \bar{w}_i x_i^{\tau'} + \hat{w}_j^{\overline{B}} x_j^{\tau'} + \hat{w}_i^{\overline{B}} x_i^{\tau'} \\ &\leq \sum_{j \in N} \bar{w}_j x_j^{\tau'} + \sum_{b \in B} \sum_{j \in N^b} \hat{w}_j^b x_j^{\tau'} \leq c; \end{aligned}$$

a contradiction. Hence, $x_i^{\tau'} = 0$ follows for all values τ'. This implies that x^τ cannot be affinely independent contradicting our first supposition. This completes the proof. □

5.2.2 Valid inequalities

In the following, we investigate (extended) covers and the corresponding classes of valid inequalities for the mb-RKP generalizing the results for the KP and Γ-RKP.

A subset of items $\mathcal{C} \subseteq N$ is called a *(multi-band robust-)cover* if there exists a a realization of uncertain weights such that its total weight exceeds the knapsack capacity. A cover is called *minimal* if $\mathcal{C} \setminus \{j\}$ is not a cover for all $j \in \mathcal{C}$.

Given a cover $\mathcal{C} \subseteq N$, we can define its extension $E(\mathcal{C})$ similar to the special case of Γ-RKP:

$$E(\mathcal{C}) := \left\{ j \in N \;:\; \bar{w}_j + \hat{w}_j^b \geq \max_{i \in \mathcal{C}} (\bar{w}_i + \hat{w}_i^b) \; \forall b \in B \right\} \tag{5.5}$$

Furthermore, an assignment of items $j \in \mathcal{C}$ to bands $b \in B$ which maximizes the total weight of the cover $\mathcal{C}$ can be obtained by the min-cost flow algorithm pointed out in Section 3.2 in particular cf. Lemma 3.9. Let $\{\mathcal{C}^b\}_{b \in B}$ be a family of disjoint subsets of $\mathcal{C}$ defining such an assignment. Then a strengthened extension $E^+(\mathcal{C}^{\underline{B}}, \ldots, \mathcal{C}^{\overline{B}})$ can be defined as

$$E^+(\mathcal{C}^{\underline{B}}, \ldots, \mathcal{C}^{\overline{B}}) := \left\{ j \in N \;:\; \bar{w}_j + \hat{w}_j^b \geq \max_{i \in \mathcal{C}^b} (\bar{w}_i + \hat{w}_i^b) \; \forall b \in B \right\} \tag{5.6}$$

By definition, both extensions are covers themselves.

Lemma 5.6. *Let $\mathcal{C} \subseteq N$ be a cover, $E(\mathcal{C})$ its extension and $E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}})$ its strengthened extension. Then, the* cover inequality

$$\sum_{j \in \mathcal{C}} x_j \leq |\mathcal{C}| - 1 \tag{5.7}$$

and the extended cover inequalities

$$\sum_{j \in E(\mathcal{C})} x_j \leq |\mathcal{C}| - 1 \tag{5.8}$$

and

$$\sum_{j \in E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}})} x_j \leq |\mathcal{C}| - 1 \tag{5.9}$$

are valid for $\mathcal{K}^{mb}$.

Proof. Let $\mathcal{C} \subseteq N$ be a cover, $E(\mathcal{C})$ its extension and $E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}})$ its strengthened extension. Further, let $x \in \mathcal{K}^{\text{mb}}$ be a feasible point of the multi-band robust knapsack polytope.

Suppose x does not satisfy (5.7). Then $\sum_{j \in \mathcal{C}} x_j \geq |\mathcal{C}|$ implies $x_j = 1$ for all $j \in \mathcal{C}$. Hence

$$\begin{aligned} &\sum_{j \in N} \bar{w}_j x_j + \max_{N^b \subseteq N, b \in B: N = \bigcup_{b \in B} N^b, |N^b| = \vartheta^b} \sum_{b \in B} \sum_{j \in N^b} \hat{w}_j^b x_j \\ \geq &\sum_{j \in \mathcal{C}} \bar{w}_j + \max_{N^b \subseteq \mathcal{C}, b \in B: \mathcal{C} = \bigcup_{b \in B} N^b, |N^b| \leq \vartheta^b} \sum_{b \in B} \sum_{j \in N^b} \hat{w}_j^b \\ \geq &c + 1 \end{aligned}$$

and thus $x \notin \mathcal{K}^{\text{mb}}$; a contradiction.

Let us consider the extended cover inequalities next. Define $\Delta := E(\mathcal{C}) \setminus \mathcal{C}$ and $\Delta^+ := E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}}) \setminus \mathcal{C}$. By definition, $\mathcal{C} \setminus \{j\} \cup \{i\}$ is a cover for all $j \in \mathcal{C}$ and $i \in \Delta$ or $i \in \Delta^+$, respectively. Hence the corresponding cover inequalities are valid. The extended cover inequalities (5.8) and (5.9) are conic combinations of these cover inequalities with right hand side values rounded down (due to $x \in \{0,1\}^{|N|}$) and thus valid. This completes the proof. □

5.3 Algorithms

In this section, we present an ILP-based separation algorithm to find violated multi-band robust strengthened extended cover inequalities.

5.3.1 Separation of multi-band robust strengthened extended cover inequalities

We consider a fractional LP solution x^* of the mb-RKP. Let y_j^b and α_j be binary decision variables so that $y_j^b = 1$ if and only if $j \in \mathcal{C}^b$ and $\alpha_j = 1$ if and only if j is added while extending the cover, i. e., $j \in E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}}) \setminus \bigcup_{b \in B} \mathcal{C}^b$. Using this notation, we formulate the separation problem as ILP as follows.

$$
\begin{aligned}
Z := \min \; & \sum_{j \in N} \sum_{b \in B} (1 - x_j^*) y_j^b + \sum_{j \in N} x_j^* \alpha_j && && (5.10a)\\
\text{s. t.} \; & \sum_{j \in N} \sum_{b \in B} (\bar{w}_j + \hat{w}_j^b) y_j^b \geq c + 1 && && (5.10b)\\
& \sum_{b \in B} y_j^b + \alpha_j \leq 1 && \forall j \in N && (5.10c)\\
& \sum_{j \in N} y_j^b = \vartheta^b && \forall b \in B && (5.10d)\\
& y_i^b + \alpha_j \leq 1 && \forall j, i \in N,\, b \in B \,:\, \bar{w}_i + \hat{w}_i^b > \bar{w}_j + \hat{w}_j^b && (5.10e)\\
& y_j^b, \alpha_j \in \{0, 1\} && \forall j \in N,\, b \in B && (5.10f)
\end{aligned}
$$

where the objective function (5.10a) minimizes the feasibility of the resulting Γ-robust strengthened extended cover inequality. Constraint (5.10b) ensures the covering condition. Constraint (5.10c) guarantees that an item may either be in the cover, or in the set of items added in the extension, or not in the extended cover at all. We know by the definition of the frequency profile, that ϑ^b items with fall in band $b \in B$ in a worst case realization; this is modeled by constraint (5.10d). The constraints (5.10e) models the conditions for an item added into the strengthened extension. Finally, the variable ranges are given by (5.10f).

Let (y^{b^*}, α^*) be a solution of ILP (5.10). Then define

$$
E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}}) := \left\{ j \in N \,:\, \sum_{b \in B} y_j^b + \alpha_j = 1 \right\}
$$

with $\mathcal{C}^b := \left\{ j \in N \,:\, y_j^b = 1 \right\}$. By construction, the Γ-robust strengthened extended cover inequality corresponding to $E^+(\mathcal{C}^{\underline{B}}, \dots, \mathcal{C}^{\overline{B}})$ is violated if and only if $Z < 1$.

Note that ILP (5.10) generalizes ILP (4.12) from the Γ-robust setting to multi-band robustness.

CHAPTER SIX

THE SUBMODULAR KNAPSACK PROBLEM

In this chapter, we consider knapsack problems under data uncertainty for which submodular formulations for the corresponding robust counterparts exist. The resulting problem is called the submodular knapsack problem (SMKP). For example, it generalizes the Γ-robust knapsack problem.

The SMKP has been introduced by Atamtürk and Narayanan [16]. Their investigation covers a polyhedral study on the problem focusing on the class of (extended) cover inequalities and the related lifting problem. Further, they consider a special subclass of submodular functions and present theoretical bounds on the lifting coefficients as well as a practical separation algorithm for lifted extended cover inequalities for this subclass. Koster and Kutschka [99] applied the SMKP to routing problems in the context of telecommunication networks and also identified new valid and facet-defining classes of inequalities for the SMKP.

Before the work by Atamtürk and Narayanan [16], the submodular knapsack problem was not covered in the literature. Instead in most works, only the min-/maximization of a super-/submodular objective function over a linear (i.e. classic) knapsack constraint is considered. For example, Sviridenko [151] gives an approximation algorithm to the maximization of a submodular function subject to a knapsack constraint. Atamtürk and Narayanan [15] and Ahmed and Atamtürk [9] consider the minimization/maximization of a special subclass of submodular functions (containing a concave term) using a cutting plane approach. The supermodular knapsack problem has been investigated by Gallo and Simeone [71] applying a Lagrangian relaxation to the problem of maximizing a supermodular function over a linear knapsack constraint.

Definition 6.1 (Submodular Knapsack Problem)**.** Given a set of items N, a knapsack capacity $c \in \mathbb{Z}_{>0}$, a profit function $p : N \to \mathbb{Z}_{\geq 0}$, and a submodular weight function $f : 2^N \to \mathbb{Z}_{>0}$. The *submodular (robust) knapsack problem* (SMKP) is to find a subset of items N' with maximum total profit whose total weight $f(N')$ does not exceed the knapsack capacity.

We define the *difference function* $\varrho_j : 2^N \to \mathbb{Z}_{\geq 0}$ as $\varrho_j(N') := f(N' \cup \{j\}) - f(N')$, i. e., the difference function ϱ_j returns the increase in weight by adding item $j \in N$ to the set $N' \in 2^N$.

We follow Atamtürk and Narayanan [16] and assume that

1. f is non-decreasing, i. e., $f(N') \leq f(N'')$ for all $N' \subseteq N'' \subseteq N$,

2. $f(\emptyset) = 0$,

3. $0 < \varrho_j(\emptyset) \leq c$ for all $j \in N$.

Note, the first two assumptions imply that the system of subsets of items fitting into the submodular knapsack is an independence system. Vice versa, given an independence system $\mathcal{I} \subseteq 2^N$, we can construct a corresponding SMKP instance as follows: define the capacity as $c_{\mathcal{I}} := 1$, and the weight function $f_{\mathcal{I}}(\emptyset) := 0$, $f_{\mathcal{I}}(N') := 1$ for all $N' \in \mathcal{I} \setminus \{\emptyset\}$, and 2 otherwise.

Example 6.2. *Let us consider an SMKP instance with four items defined as follows. Let $N := \{1, 2, 3, 4\}$ be the set of items, $p_1 = p_2 = p_3 = p_4 = 1$ be constant profits (thus maximizing the number of items), and $c = 11$ be the knapsack capacity. Next, we have to define a submodular function f. For $N' \in 2^N$ define*

$$f(N') = \begin{cases} 0 & \textit{if } N' = \emptyset, \\ 6 & \textit{if } |N'| = 1, \\ 11 & \textit{if } N' = \{1,2\}, \\ 12 & \textit{if } |N'| = 2 \textit{ and } N' \neq \{1,2\}, \\ 15 & \textit{if } N' = \{1,2,3\} \textit{ or } N' = \{1,2,4\}, \\ 18 & \textit{if } N' = \{1,3,4\} \textit{ or } N' = \{2,3,4\}, \\ 20 & \textit{if } N' = N. \end{cases}$$

Then, $N' = \{1, 2\}$ with weight 11 satisfies the knapsack capacity of 11 and maximizes the profit function with value 2 since it is the largest subset of N fitting into the submodular knapsack.

6.1 Formulations

The SMKP can be formulated as the following mathematical program

$$\max \sum_{j \in N} p_j x_j \tag{6.1a}$$

$$\text{s. t. } f(\{j \in N \ : \ x_j = 1\}) \leq c \tag{6.1b}$$

$$x_j \in \{0,1\} \qquad \forall j \in N. \tag{6.1c}$$

An ILP formulation exists if f can be linearized. For example, for $f = f^\Gamma$ with

$$f^\Gamma : 2^N \to \mathbb{Z}_{>0},\ f^\Gamma(N') := \sum_{j \in N'} \bar{w}_j + \max_{N'' \subseteq N' : |N'| \leq \Gamma} \sum_{j \in N''} \hat{w}_j$$

for all $N' \subseteq N$, the formulation (6.1) reads

$$\max \sum_{j \in N} p_j x_j \tag{6.2a}$$

$$\text{s. t.} \sum_{j \in N} \bar{w}_j x_j + \max_{N' \subseteq N \,:\, |N'| \leq \Gamma} \sum_{j \in N'} \hat{w}_j x_j \leq c \tag{6.2b}$$

$$x_j \in \{0, 1\} \qquad \forall j \in N. \tag{6.2c}$$

This formulation is the Γ-robust counterpart of the Γ-RKP for which a compact linear reformulation exists; cf. Chapter 4.

6.2 Polyhedral study

In this section, we describe polyhedral insights to the set of feasible solutions of the SMKP. We start with the following basic definition. Note, we write $f(x)$ instead of the correct $f(\{j \in N \,:\, x_j = 1\})$ for a simpler notation and better readability.

Definition 6.3 (Submodular Knapsack Polytope)**.** Given a set of items N, a positive capacity $c \in \mathbb{Z}_{>0}$, and a submodular weight function $f : 2^N \to \mathbb{Z}_{\geq 0}$, the *submodular knapsack polytope* is defined as

$$\mathcal{K}^f := \operatorname{conv}\left\{x \in \{0, 1\}^{|N|} \,:\, f(x) \leq c\right\}.$$

6.2.1 Basic characteristics

Some basic properties of the SMKP polytope are the following.

Lemma 6.4 (Atamtürk and Narayanan [16])**.** *The submodular knapsack polytope $\mathcal{K}^f$ is full-dimensional.*

In the following, we assume w. l. o. g. that $\mathcal{K}^f$ is full-dimensional.

Lemma 6.5 (Atamtürk and Narayanan [16])**.** *Trivial facets of the submodular knapsack polytope $\mathcal{K}^f$ are given by the inequalities*

$$x_j \geq 0 \qquad \forall j \in N \tag{6.3a}$$

$$x_j \leq 1 \qquad \forall j \in N : \max_{i \in N \setminus \{j\}} f(\{j, i\}) \leq c. \tag{6.3b}$$

6.2.2 Valid inequalities

Many classes of valid inequalities for the classic knapsack problem are derived from special structured subsets of items, e. g., covers. In this section, we present results for the generalizations of these structures. The results on (extended) covers have been obtained by Atamtürk and Narayanan [16] and are stated for completeness.

A subset of items $\mathcal{C} \subseteq N$ is called a *(submodular) cover* if its total weight exceeds the knapsack capacity, i. e., $f(\mathcal{C}) \geq c + 1$. A cover is called *minimal* if $\mathcal{C} \setminus \{j\}$ is not a cover for all $j \in \mathcal{C}$.

Let $\mathcal{K}^f(N') := \text{proj}_{x_j \,:\, j \in N'} \mathcal{K}^f$ be the submodular knapsack polytope restricted to the items in $N' \subseteq N$.

Lemma 6.6 (Atamtürk and Narayanan [16])**.** *Let $\mathcal{C} \subseteq N$ be a cover. Then, the* cover inequality

$$\sum_{j \in \mathcal{C}} x_j \leq |\mathcal{C}| - 1 \tag{6.4}$$

is valid for $\mathcal{K}^f$. Furthermore, it is facet-defining for $\mathcal{K}^f(\mathcal{C})$ if and only if $\mathcal{C}$ is minimal.

Let $\pi = (\pi_1, \ldots, \pi_{|N \setminus \mathcal{C}|})$ be a permutation of the items not in the cover $\mathcal{C}$. Let $\mathcal{C}_t := \mathcal{C} \cup \{\pi_1, \ldots, \pi_t\}$ for all $t \in \{1, \ldots, |N \setminus \mathcal{C}|\}$ and let $\mathcal{C}_0 := \mathcal{C}$. The *extension of $\mathcal{C}$ (w.r.t. π)* is given by

$$E_\pi(\mathcal{C}) := \mathcal{C} \cup \{\pi_t \in N \setminus \mathcal{C} \;:\; \varrho_{\pi_t}(\mathcal{C}_{t-1}) \geq \varrho_j(\emptyset) \text{ for all } j \in \mathcal{C}\}\,. \tag{6.5}$$

In general, the extension $E_\pi(\mathcal{C})$ of a cover $\mathcal{C}$ is sequence-dependent due to the submodularity of f.

Example 6.7. *Let us consider the SMKP instance with four items introduced in Example 6.2. Then, the set $\mathcal{C} := \{3, 4\}$ is a cover since $f(\mathcal{C}) = 12 > 11 = c$. The set of remaining items $N \setminus \mathcal{C}$ is $\{1, 2\}$. Thus, there exist two permutations of this two item set: $\pi = (1, 2)$ and $\pi' = (2, 1)$. The two extensions w.r.t. these permutations are $E_\pi(\mathcal{C}) = \{1, 3, 4\}$ and $E_{\pi'}(\mathcal{C}) = \{2, 3, 4\}$. This illustrates the sequence-dependency of the cover extension in the submodular robust setting.*

Lemma 6.8 (Atamtürk and Narayanan [16])**.** *Let $\mathcal{C} \subseteq N$ be a cover, π a permutation of the items $N \setminus \mathcal{C}$, and $E_\pi(\mathcal{C})$ the corresponding extension of the cover. Then, the* extended cover inequality

$$\sum_{j \in E_\pi(\mathcal{C})} x_j \leq |\mathcal{C}| - 1 \tag{6.6}$$

is valid for $\mathcal{K}^f$. Furthermore, it is facet-defining for $\mathcal{K}^f(E_\pi(\mathcal{C}))$ if and only if $\mathcal{C}$ is minimal and for each $j \in \{\pi_t \in N \setminus \mathcal{C} \;:\; \varrho_{\pi_t}(\mathcal{C}_{t-1}) \geq \varrho_j(\emptyset)$ for all $j \in \mathcal{C}\}$ there exist distinct $i, i' \in \mathcal{C}$ such that $f(\mathcal{C} \cup \{j\} \setminus \{i, i'\}) \leq c$.

Further, Atamtürk and Narayanan [16] consider the lifting of cover inequalities to strengthen the inequalities and define facets of the higher-dimensional polytope $\mathcal{K}^f$. Consider the *lifted cover inequality*

$$\sum_{j \in \mathcal{C}} x_j + \sum_{j \in N \setminus \mathcal{C}} \alpha_j x_j \leq |\mathcal{C}| - 1 \tag{6.7}$$

with lifting coefficients α_j. These lifting coefficients are sequence-dependent. Given a (possible empty) subset $N' \subseteq N$ of items for which the lifting coefficients have already been determined, the lifting coefficient α_i with $i \in N \setminus (\mathcal{C} \cup N')$ can be calculated as follows

$$\alpha_i := |\mathcal{C}| - 1 - \max_{N'' \subseteq (\mathcal{C} \cup N')} \left\{ |\mathcal{C} \cap N''| + \sum_{j \in N' \cap N''} \alpha_j \ : \ f(N'' \cup \{i\}) \leq c \right\}. \tag{6.8}$$

Note, determining the optimal lifting coefficients is again submodular knapsack problem.

Proposition 6.9 (Atamtürk and Narayanan [16]). *If $\mathcal{C}$ is a minimal cover for SMKP, then for any lifting sequence the lifted cover inequality* (6.7) *with lifting coefficients α_i determined by* (6.8) *is facet-defining for $\mathcal{K}^f$.*

Atamtürk and Narayanan [16] state sufficient conditions to bound the lifting coefficients from below and above. They consider a subclass of submodular functions and show the upper bound can be computed in polynomial time for this subclass whereas the determination of the lower bound is $\mathcal{NP}$-complete.

Next, we consider structures more general than covers. The concept of a $(1,k)$-*configuration* has been introduced by Padberg [134] for the knapsack, see Section 1.2. We generalize this concept to submodular knapsacks. A subset $N' \subseteq N$ and an element $t \in N \setminus N'$ is called a $(1,k)$-*configuration* if $f(N') \leq c$ and $Q \cup \{t\}$ is a minimal cover for $2 \leq k \leq |N'|$ and for all $Q \subseteq N'$ with $|Q| = k$. Note, for $k = |N'|$ a $(1,k)$-configuration is a minimal cover; cf. Section 1.2 for $(1,k)$-configurations for KP.

Padberg [134] has shown that $(1,k)$-configurations yield valid inequalities for the knapsack polytope $\mathcal{K}$ which are also facet-defining under mild conditions. This result can be extended to the submodular knapsack polytope $\mathcal{K}^f$ as follows.

Theorem 6.10. *Given a submodular knapsack polytope $\mathcal{K}^f$ with elements $N = \{1,\ldots,n\}$. Let $t \in N$ and $N' \subseteq N \setminus \{t\}$ be a $(1,k)$-configuration. The* $(1,k)$-configuration inequalities

$$(r-k+1)x_t + \sum_{j \in T} x_j \leq r \qquad \forall T \subseteq N' \ : \ |T| = r, \ k \leq r \leq |N'| \tag{6.9}$$

are valid for $\mathcal{K}^f$ and facet-defining for $\mathcal{K}^f(N' \cup \{t\})$.

Proof. First, we show the validity of the inequalities (6.9). Given a $(1,k)$-configuration and a related inequality (6.9), we consider the two possible values of x_t: If $x_t = 0$, then the inequality reduces to

$$\sum_{j \in T} x_j \leq r.$$

This is valid for $\mathcal{K}^f$ by definition of T. If $x_t = 1$, then inequality (6.9) reduces to

$$\sum_{j \in T} x_j \leq k - 1$$

which is valid for $\mathcal{K}^f$ since $Q \cup \{t\}$ is a minimal cover for all $Q \subseteq N'$ with $|Q| = k$ and so it holds that $f(Q \cup \{t\}) > c$ for each k-element subset Q of T.

Second, we show that for the case $N' = N \setminus \{t\}$, the face defined by inequality (6.9) contains n affinely independent vectors. W.l.o.g., we assume $T = \{1, \ldots, r\}$ and $t = n$. For $i = 1, \ldots, r$ we define the sets $U_i := \{((i-1) \bmod r) + 1,\ (i \bmod r) + 1,\ ((i+1) \bmod r) + 1,\ \ldots,\ ((i+k-3) \bmod r) + 1\}$. Since $|U_i| = k-1$, it holds $f(U_i \cup \{t\}) \leq c$ by definition. Further, we define sets $W_j := \{1, \ldots, k-2, j, n\}$ for $j = r+1, \ldots, n-1$. For these sets, $f(W_j) \leq c$ holds because $t \in W_j$ and $|W_j \setminus \{t\}| = k-1$. Finally, we consider the set T for which $f(T) \leq c$ holds by definition. The n characteristic vectors of $U_i \cup \{t\}$ $(1 \leq i \leq r)$, W_j $(r \leq j \leq n-1)$ and T are clearly affinely independent and on the face defined by inequality (6.9). □

An even stronger result for $(1,k)$-configurations can be obtained: the submodular knapsack polytope $\mathcal{K}^f$ is completely described by the trivial facets, the knapsack constraint, and all $(1,k)$-configuration inequalities. The following theorem states this result more formally.

Theorem 6.11. *Given the submodular knapsack polytope $\mathcal{K}^f$ and its linear relaxation ${}^{LP}\mathcal{K}^f$ defined by*

$$ {}^{LP}\mathcal{K}^f := \operatorname{conv}\left\{x \in [0,1]^{|N|} \ :\ f(x) \leq c\right\}. \tag{6.10}$$

If for $N' = \{2, \ldots, n\}$ and $t = 1$ the following two conditions hold

1. *$f(N') \leq c$,*

2. *$Q \cup \{t\}$ is a minimal cover for all $Q \subseteq N'$ with $|Q| = k$ and $2 \leq k \leq |N'|$,*

then the polytope obtained by intersecting ${}^{LP}\mathcal{K}^f$ with the $(1,k)$-configuration inequalities

$$(r-k+1)x_t + \sum_{j \in T} x_j \leq r \qquad \forall T \subseteq N' \ :\ |T| = r,\ k \leq r \leq |N'| \tag{6.11}$$

has zero-one vertices only, i. e.,

$$\mathcal{K}^f = {}^{LP}\mathcal{K}^f \cap \left\{x \in \mathbb{R}^{|N|} \ :\ x \text{ satisfies all inequalities } (6.11)\right\}. \tag{6.12}$$

Proof. W.l.o.g., let

$$\pi^\top x \leq 1 \tag{6.13}$$

define a non-trivial facet F (with right-hand side normalized to 1) of $\mathcal{K}^f$. In the following, we show that F is in fact of the form (6.11).

From $f(N') \leq c$ it follows that $\sum_{j=2}^n \pi_j = 1$. Suppose this is not true, i.e., $\sum_{j=2}^n \pi_j < 1$. Then for $x \in F$, $1 = \pi^\top x = \pi_1 x_1 + \sum_{j=2}^n \pi_j x_j \leq \pi_1 x_1 + \sum_{j=2}^n \pi_j < \pi_1 x_1 + 1$ which holds if and only if $0 < \pi_1 x_1$. This implies that the facet F is identical to $x_1 = 1$; a contradiction to F being non-trivial.

W.l.o.g., let the variables be indexed such that $\pi_j \geq \pi_{j+1}$ for $j = 2, \ldots, n-1$. Since, $Q \cup \{t\}$ is a minimal cover for all $Q \subseteq N'$ with $|Q| = k$ and $2 \leq k \leq |N'|$, it follows that $\sum_{j=1}^k \pi_j = 1$. Suppose this is not true, i.e., $\sum_{j=1}^k \pi_j < 1$. Let $R \subseteq N'$ such that $f(R \cup \{t\}) \leq c$ holds. Then $|R| \leq k-1$. It follows for x with $x_j = 1$ if and only if $j \in R \cup \{t\}$ that $\pi^\top x = \pi_1 + \sum_{j \in R} \pi_j \leq \pi_1 + \sum_{j=2}^{k} \pi_j < 1$. Therefore, no solution with $x_1 = 1$ exists on the facet F. Thus, the facet F is contained in the face $\left\{x \in \mathcal{K}^f \ : \ x_1 = 0\right\} \subsetneq \mathcal{K}^f$ contradicting F being a facet of $\mathcal{K}^f$.

Furthermore, it holds $\sum_{j=1}^{k-1} \pi_j + \pi_{k+1} = 1$. Suppose this is not true, i. e., $\sum_{j=1}^{k-1} \pi_j + \pi_{k+1} < 1$. Let $R \subseteq N'$ such that $f(R \cup \{t\} \cup \{k+1\}) \leq c$ holds. Then $|R| \leq k-2$. For x with $x_j = 1$ if and only if $j \in R \cup \{t\} \cup \{k+1\}$, it follows $\pi^\top x = \pi_1 + \sum_{j \in R} \pi_j + \pi_{k+1} \leq \pi_1 + \sum_{j=2}^{k-1} \pi_j + \pi_{k+1} < 1$. Therefore, no solution with $x_{k+1} = 1$ exists on the facet F and $F \subseteq \left\{x \in \mathcal{K}^f \ : \ x_{k+1} = 0\right\} \subsetneq \mathcal{K}^f$; a contradiction to F being a facet of $\mathcal{K}^f$.

On the other hand, it holds $\pi_1 + \sum_{j=3}^{k+1} \pi_j = 1$. Suppose this is not true, i. e., $\pi_1 + \sum_{j=3}^{k+1} \pi_j < 1$. Let $R \subseteq N' \setminus \{2\}$ such that $f(R \cup \{t\}) \leq c$ holds. Then $|R| \leq k-1$. For x with $x_j = 1$ if and only if $j \in R \cup \{t\}$, it follows $\pi^\top x = \pi_1 + \sum_{j \in R} \pi_j \leq \pi_1 + \sum_{j=3}^{k+1} \pi_j < 1$. Hence, $x_2 = 1$ holds for all points on F. Thus, F is identical to the trivial facet $x_2 = 1$; a contradiction.

Repeating the argument for 3, 4, ..., $k-1$, it follows that there exists a value μ with $0 \leq \mu \leq 1$ such that $\pi_j = \mu$ for all $j = 2, \ldots, k+1$. Since $\mu = 0$ implies $\pi_1 = 1$, $\mu = 1$ implies $\pi_1 = 0$, and F being non-trivial, it holds $0 < \mu < 1$.

Further, it holds $x_1 + x_j \geq 1$ $(j = 2, \ldots, n)$ for every point $x \in \mathcal{K}^f$ satisfying $\pi^\top x = 1$. Suppose this is not true, i. e., there exists a $x \in F$ with $x_1 + x_j < 1$. Then, $x_1 = x_j = 0$. W. l. o. g., we assume $j = 2$. Then $1 = \pi^\top x = \sum_{j=1}^{n} \pi_j x_j = \sum_{j=3}^{n} \pi_j x_j \leq \sum_{j=3}^{n} \pi_j = \sum_{j=2}^{n} \pi_j - \pi_2 = 1 - \pi_2$. This implies $\pi_2 = 0$; a contradiction to $\pi_2 = \mu > 0$.

Suppose $\pi_j = \mu$ for all $j = 2, \ldots, q$ and $\pi_j = 0$ for all $j \geq q+1$. Then $1 = \sum_{j=2}^{n} \pi_j = (n-1)\mu$ if and only if $\mu = 1/(n-1)$, and $\pi_1 = 1 - (k-1)/(n-1) = (n-k)/(n-1)$. Thus, the facet F is of the form

$$(n-k)x_1 + \sum_{j=2}^{n} x_j \leq n-1 \tag{6.14}$$

which is a $(1,k)$-configuration inequality with $r = n-1$.

Suppose $\pi_j = \mu$ for all $j = 2, \ldots, q$ and $\pi_j = \mu$ for all $j \geq q+1$. Then $1 = \sum_{j=2}^{n} \pi_j = (q-1)\mu$ if and only if $\mu = 1/(q-1)$, and $\pi_1 = 1 - (k-1)/(q-1) = (q-k)/(q-1)$. Thus, the facet F is of the form

$$(q-k)x_1 + \sum_{j=2}^{n} x_j \leq q-1 \tag{6.15}$$

which is a $(1,k)$-configuration inequality with $r = q-1$. □

Notice, Theorem 6.11 generalizes Theorem 2 in Padberg [134] to the submodular setting.

Next, we consider the submodular robust counterpart of weight inequalities. In contrast to the classic setting, it turns out that submodular robust weight inequalities are sequence-dependent in general; similarly to the extension of submodular robust covers we presented earlier.

Lemma 6.12. *Let $N' = \{j_1, \ldots, j_k\} \subseteq N$ with $f(N') \leq c$ and let $\pi = (\pi_1, \ldots, \pi_{n-k})$ be a permutation of the items $N \setminus N'$. Define $c_{res} := c - f(N')$ as the residual capacity w. r. t. N' and the following sets of items*

$$\begin{aligned}
&N'_0 := \emptyset, \\
&N'_t := \{j_1, \ldots j_t\} && \forall t = 1, \ldots, k, \\
&\bar{N}_0 := N'_k = N', \\
&\bar{N}_t := N' \cup \{\pi_1, \ldots, \pi_t\} && \forall t = 1, \ldots, n-k.
\end{aligned}$$

Then, the weight inequality

$$\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} + \sum_{t=1}^{n-k} \max\{0, \varrho_{\pi_t}(\bar{N}_{t-1}) - c_{res}\} x_{\pi_t} \leq f(N') \tag{6.16}$$

is valid for $\mathcal{K}^f$.

Proof. Let $x^* \in \mathcal{K}^f \cap \mathbb{Z}^n$. Define $X := \{j \in N \ : \ x^*_j = 1\}$. To prove the validity of inequality (6.16), we consider the following two cases.

First, let $X \setminus N' = \emptyset$. Then, for the left-hand side of inequality (6.16) holds

$$\begin{aligned}
&\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} + \underbrace{\sum_{t=1}^{n-k} \max\{0, \varrho_{\pi_t}(\bar{N}_{t-1}) - c_{\mathrm{res}}\} x_{\pi_t}}_{=0} = \sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} \\
\leq &\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) = f(X \cap N') \leq f(N').
\end{aligned}$$

Second, let $X \setminus N' \neq \emptyset$. Then, for the left-hand side of inequality (6.16) holds

$$\begin{aligned}
&\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} + \sum_{t=1}^{n-k} \max\{0, \varrho_{\pi_t}(\bar{N}_{t-1}) - c_{\mathrm{res}}\} x_{\pi_t} \\
= &\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} + \sum_{\substack{t=1,\ldots,n-k: \\ \varrho_{\pi_t}(\bar{N}_{t-1}) > c_{\mathrm{res}}}} \varrho_{\pi_t}(\bar{N}_{t-1}) x_{\pi_t} \\
&- \underbrace{\left| \left\{ t = 1, \ldots, n-k \ : \ \varrho_{\pi_t}(\bar{N}_{t-1}) > c_{\mathrm{res}} \right\} \right|}_{:=\tau} \cdot c_{\mathrm{res}} \\
= &\sum_{t=1}^{k} \varrho_{j_t}(N'_{t-1}) x_{j_t} + \sum_{\substack{t=1,\ldots,n-k: \\ \varrho_{\pi_t}(\bar{N}_{t-1}) > c_{\mathrm{res}}}} \varrho_{\pi_t}(\bar{N}_{t-1}) x_{\pi_t} - \tau \cdot c_{\mathrm{res}}.
\end{aligned} \tag{6.17}$$

Let us consider $\tau = 0$, i. e., $\varrho_{\pi_t}(\bar{N}_{t-1}) \leq c_{\mathrm{res}}$ for all items in $j \in N \setminus N'$. Hence, it follows

$$(6.17) \leq f(X \cap N') \leq f(N').$$

Now, let us consider $\tau \geq 1$. Because of $\emptyset = N'_0 \supseteq \cdots \supseteq N'_k = N' = \bar{N}_0 \supseteq \bar{N}_1 \supseteq \cdots \supseteq \bar{N}_{n-k} = N$, we relax the term (6.17) as follows

$$\begin{aligned}(6.17) &\leq f(X \cap N') + \sum_{t=1}^{n-k} \varrho_{\pi_t}(\bar{N}_{t-1}) x_{\pi_t} - \tau \cdot c_{\text{res}} \\ &\leq f(X \cap N) - \tau \cdot c_{\text{res}} = f(X) - \tau \cdot c_{\text{res}} \leq f(X) - c_{\text{res}} \leq c - c_{\text{res}} = f(N')\end{aligned}$$

where $f(X) \leq c$ holds because of the definition of X and $x^* \in \mathcal{K}^f$. This completes the proof. □

Note, the weight inequalities are sequence-dependent similar to the submodular robust extended cover inequalities.

CHAPTER SEVEN

THE RECOVERABLE ROBUST KNAPSACK PROBLEM

Previously, we have considered one-stage approaches to solve the knapsack problem under data uncertainty. Now, we follow the two-stage concept of recoverable robustness and investigate the corresponding recoverable robust knapsack problem (RRKP). In a first-stage decision, this problem asks to find a most profitable subset of items fitting into a given first-stage knapsack capacity taking the future second stage decisions into account. In a second-stage adjustment, some of the items may be removed or new items may be added to satisfy a given second-stage knapsack capacity or to increase the profit further.

The author of this thesis together with the co-authors Christina Büsing and Arie Koster has investigated special cases of the general RRKP: the k, ℓ-recoverable robust knapsack problem with discrete scenarios ($k, \ell/D$-RRKP), the k, ℓ-recoverable robust knapsack problem with interval scenarios, and the k-recoverable knapsack problem with Γ-scenarios (k/Γ-RRKP). Some of the results for special cases of the RRKP in this chapter have been previously published; cf. [46, 49, 50]. The special case of the $k, \ell/D$-RRKP for which the item weights are certain but the scenario knapsack capacity is uncertain has been studied by Bouman et al. [44] focusing on its computability by applying column generation methods.

To our best knowledge, most work on uncertain knapsack problems is limited to uncertain item weights or profits and/or knapsack capacity without the capability of recovery. Thus, we are not aware of other work on the RRKP besides the mentioned ones. A survey on previous work on the Γ-robust knapsack problem (without recovery) and other approaches on uncertain knapsack problems can be found in the introduction of Chapter 4.

Although we are also going to present main results of our work on the $k, \ell/D$-RRKP, the main focus will be on new results for the k/Γ-RRKP. First, let us define the RRKP in general.

Definition 7.1 (Recoverable Robust Knapsack Problem)**.** Let N be a set of n items with first-stage profits $p_j^0 \in \mathbb{N}$ and weights $w_j^0 \in \mathbb{N}$, and second-stage profits $p_j^S \in \mathbb{N}$ and weights $w_j^S \in \mathbb{N}$. Let $\mathcal{S}$ be the scenario set and $\mathcal{R}(N')$ the recovery set of $N' \subseteq N$.

Given first-stage capacity $c^0 \in \mathbb{N}$ and scenario (or second-stage) capacity $c^S \in \mathbb{N}$, the *recoverable robust knapsack problem* is to find a set $N' \in N$ with maximum profit

$$p^0(N') + \min_{S \in \mathcal{S}} \max_{N'' \in \mathcal{R}(N')} p^S(N'')$$

such that $w^0(N') \leq c^0$ and for every scenario $S \in \mathcal{S}$, there exists a set $N'' \in \mathcal{R}(N')$ with $w^S(N'') \leq c^S$ and N'' maximizing $p^S(N'')$.

Next, we define three scenario sets and two recovery rules which we are going to investigate.

Definition 7.2. Let N be a set of n items. We define the following scenario sets:

discrete scenarios The finite set $\mathcal{S}_D$ consists of scenarios S, each defining a weight function $w^S : N \to \mathbb{N}$, a profit function $p^S : N \to \mathbb{N}$, and capacity $c^S \in \mathbb{N}$.

interval scenarios The set $\mathcal{S}_I$ consists of scenarios S, each defining a weight function $w^S : N \to \mathbb{N}$ such that $w_j^S \in [w_j^{\min}, w_j^{\max}]$ where $w_j^{\min}$ defines the minimum and $w_j^{\max}$ maximum weight. Further, let $p^S : N \to \mathbb{N}$ be the profit function of scenario S and $c^S = c \in \mathbb{N}$ be the capacity. Note, the latter is the same for all scenarios.

Γ-scenarios For a given $\Gamma \in \mathbb{N}$, the set $\mathcal{S}_\Gamma$ consists of all scenarios S which define a weight function $w^S : N \to \mathbb{N}$ such that $w_j^S \in [\bar{w}_j, \bar{w}_j + \hat{w}_j]$ for all $j \in \mathbb{N}$ and $|\left\{j \in N \ : \ w_j^S > \bar{w}_j\right\}| \leq \Gamma$ where $\bar{w}_j$ is the nominal (or default) and $\hat{w}_j$ the maximum deviation weight. Again, let $p^S : N \to \mathbb{N}$ be the profit function of scenario S and $c^S = c \in \mathbb{N}$ be the capacity. Note, the capacity is the same for all scenarios.

Furthermore, let $N' \subseteq N$ be the subset of items selected as first-stage decision. We consider the following two recovery rules:

k, ℓ-recovery For $k, \ell \in \mathbb{N}$, the recovery set $\mathcal{R}^{k,\ell}(N')$ consists of all subsets of N' with at least $|N'| - k$ elements combined with all subsets of $N \setminus N'$ with at most ℓ elements, i. e.,

$$\mathcal{R}^{k,\ell}(N') := \{N'' \subseteq N \ : \ |N' \setminus N''| \leq k, \ |N'' \setminus N'| \leq \ell\} .$$

This recovery rule allows to remove up to k elements from N' and add up to ℓ additional elements to N'.

k-recovery For $k \in \mathbb{N}$, the recovery set $\mathcal{R}^k(N')$ consists of all subsets of N' with at least $|N'| - k$ elements, i. e.,

$$\mathcal{R}^k(N') := \{N'' \subseteq N' \ : \ |N' \setminus N''| \leq k\} .$$

This recovery rule allows to remove up to k elements from N'. It is the special case of the k, ℓ-recovery rule with ℓ set to 0.

Now we can define the special cases of the RRKP mentioned in the introduction of this chapter.

Definition 7.3. We call the RRKP with $\mathcal{S} = \mathcal{S}_D$ and $\mathcal{R}(N') = \mathcal{R}^{k,\ell}(N')$ the *k,ℓ-recoverable robust knapsack problem.* We denote it by $k,\ell/D$-RRKP.

Further, we call the RRKP with $\mathcal{S} = \mathcal{S}_\Gamma$, $\mathcal{R}(N') = \mathcal{R}^k(N')$, and $p_j^S = 0$ for all $j \in N$ the *k-recoverable Γ-robust knapsack problem.* We denote it by k/Γ-RRKP.

In Büsing et al. [49], we show that the RRKP with interval scenarios and k,ℓ-recovery ($\mathcal{S} = \mathcal{S}_I$, $\mathcal{R}(N') = \mathcal{R}^{k,\ell}(N')$) can be reduced to the $k,\ell/D$-RRKP with a single discrete scenario where for each item only the maximum weight of its weight interval is considered.

Note, the k/Γ-RRKP with $\hat{w} = 0$ or $\Gamma = 0$ is identical to the $k,\ell/D$-RRKP with $p^S = 0$, $\ell = 0$, and a single scenario S defined by $w^S = \bar{w}$.

In Büsing et al. [49] we prove the following complexity results.

Lemma 7.4 ([49]). *Considering the k,ℓ/D$-$RRKP, the decision if the total profit of a feasible first stage solution is greater or equal to a constant for just one given scenario, is weakly $\mathcal{NP}$-hard, even if $k = 0$ or $\ell = 0$.*

Theorem 7.5 ([49]). *The $k,\ell/D$-RRKP is strongly $\mathcal{NP}$-hard for an unbounded number of discrete scenarios even if $p^0 = 0$.*

Furthermore, the k,ℓ/D$-RRKP$ cannot be approximated within $(\ell + 1)/\ell$, unless $\mathcal{P} = \mathcal{NP}$. In particular, the problem is inapproximable for $\ell = 0$.

Theorem 7.6 ([49]). *The $k,\ell/D$-RRKP is strongly $\mathcal{NP}$-complete for an unbounded number of discrete scenarios even if $p^S = 0$ for all $S \in \mathcal{S}_D$.*

Theorem 7.7 ([49]). *The $k,\ell/D$-RRKP can be solved in pseudo-polynomial time for a bounded number of scenarios.*

Note, this last result is obtained by generalizing a dynamic program for the Γ-robust knapsack problem presented by Yu [161] to the recoverable robust setting.

7.1 Formulations

In this section, we present (compact) ILP formulations for the $k,\ell/D$-RRKP and k/Γ-RRKP. To obtain a compact ILP formulation for the k/Γ-RRKP, we have to consider the so-called maximum weight set problem to determine the worst-case realization and thus the left-hand side value. This way, we develop a compact ILP reformulation of the k/Γ-RRKP exploiting a new combinatorial algorithm solving the related maximum weight set problem.

$k, \ell/D$-RRKP. To formulate the $k, \ell/D$-RRKP as ILP, we introduce binary variables $x_j^0 = 1$ if and only if $j \in N$ is selected in the first-stage solution, and 0 otherwise. Similarly, binary variable $x_j^S = 1$ if and only if item j is part of the second-stage solution after the recovery rule has been applied, and 0 otherwise. To keep track of the recovery actions, i. e., the addition and removal of items from the first-stage decision, the formulation includes the binary variables y_j^S and z_j^S to indicate whether the item j has been added or removed from the first-stage solution in scenario $S \in \mathcal{S}_D$, respectively. The rational variable ω is used to linearly model the min-max term of the objective function, i. e., the worst-case scenario profit. Thus, an ILP formulation of the $k, \ell/D$-RRKP is given by

$$\max \sum_{j \in N} p_j^0 x_j^0 + \omega \tag{7.1a}$$

$$\text{s. t. } \omega \leq \sum_{j \in N} p_j^S x_j^S \qquad \forall S \in \mathcal{S}_D \tag{7.1b}$$

$$\sum_{j \in N} w_j^0 x_j^0 \leq c^0 \tag{7.1c}$$

$$\sum_{j \in N} w_j^S x_j^S \leq c^S \qquad \forall S \in \mathcal{S}_D \tag{7.1d}$$

$$x_j^S \leq y_j^S + x^0 \qquad \forall S \in \mathcal{S}_D, j \in N \tag{7.1e}$$

$$x_j^S \geq x^0 - z_j^S \qquad \forall S \in \mathcal{S}_D, j \in N \tag{7.1f}$$

$$\sum_{j \in N} y_j^S \leq \ell \qquad \forall S \in \mathcal{S}_D \tag{7.1g}$$

$$\sum_{j \in N} z_j^S \leq k \qquad \forall S \in \mathcal{S}_D \tag{7.1h}$$

$$\omega \in \mathbb{Q}_{\geq 0},\ x_j^0, x_j^S, y_j^S, z_j^S \in \{0,1\} \qquad \forall S \in \mathcal{S}_D, j \in N \tag{7.1i}$$

where (7.1a) and (7.1b) model the objective function of the $k, \ell/D$-RRKP. Constraints (7.1c) and (7.1d) enforce the first-stage and scenario knapsack capacities, respectively. The recovery rule is implemented by constraints (7.1e)–(7.1h). The variable bounds are given by (7.1i).

k/Γ-RRKP. Next, we develop a compact ILP formulation of the k/Γ-RRKP. We model the fist-stage decision by binary variables x_j^0 as in the case of discrete scenarios. But we do not model the second-stage decision explicitly. This is due to the fact that for this problem $p_j^S = 0$ holds for all $S \in \mathcal{S}_\Gamma$ and $j \in N$. Hence, we do not need to keep track of the second-stage decision to determine the scenario profit and thus we can omit additional variables to model the second stage. In addition, the number of scenarios would be anyway exponential in the worst case and thus, we avoid to handle them explicitly. Hence, we formulate the k/Γ-RRKP as follows

$$\max \sum_{j\in N} p_j^0 x_j^0 \tag{7.2a}$$

$$\text{s. t.} \sum_{j\in N} w_j^0 x_j^0 \le c^0 \tag{7.2b}$$

$$\sum_{j\in N} \bar{w}_j x_j^0 + \max_{\substack{N'\subseteq N\\ |N'|\le\Gamma}} \Big(\sum_{j\in N'} \hat{w}_j x_j^0 - \max_{\substack{N''\subseteq N\\ |N''|\le k}} \big(\sum_{j\in N''} \bar{w}_j x_j^0 + \sum_{j\in N''\cap N'} \hat{w}_j x_j^0\big)\Big) \le c \tag{7.2c}$$

$$x_j^0 \in \{0,1\} \qquad \forall j \in N. \tag{7.2d}$$

The objective (7.2a) maximizes the total first-stage profit. Constraint (7.2b) ensures the first-stage knapsack capacity. The variable bounds are given by (7.2d). The more interesting constraint is (7.2c) which models both, the recovery rule and the scenario knapsack capacity constraint. Its left-hand side consists of the term $\sum_{j\in N} \bar{w}_j x_j^0$ capturing the nominal item weights plus the remaining and possible negative maximum term which determines the change of weight due to Γ-robustness and k-recovery. Formulation (7.2) is nonlinear because of the two nested maximum terms. Its straight-forward linearization contains exponentially many constraints.

To obtain a compact reformulation, we investigate the outer maximum term of constraint (7.2c). It can be considered as an optimization problem itself and we call it the *maximum weight set problem*. More formally, we define the following.

Definition 7.8 (Maximum Weight Set Problem)**.** Given a k/Γ-RRKP instance with first-stage solution $X \subseteq N$, $X := \{j \in N \,:\, x_j^0 = 1\}$. Let $\Gamma \in \mathbb{N}$ and $k \in \mathbb{N}$ be fixed.

The *maximum weight set problem* (MWSP) is to find a subset $X' \subseteq X$ with $|X'| \le \Gamma$ maximizing its weight

$$\sum_{j\in X'} \hat{w}_j - \max_{\substack{X''\subseteq X\\ |X''|\le k}} \Big(\sum_{j\in X''} \bar{w}_j + \sum_{j\in X''\cap X'} \hat{w}_j\Big)$$

according to the given set X, and the parameters Γ and k.

In Büsing et al. [50], we point out that if $k > 0$, there is no inclusion relation between optimal solutions of the MWSP for different values of Γ, i. e., a maximum weight set for a value Γ cannot be extended to a maximum weight set for $\Gamma + 1$ in general. The following example illustrates this observation.

Example 7.9. *Let us consider a k/Γ-RRKP instance with 3 items, $k = 1$, $\Gamma \in \{1, 2\}$, and nominal and deviation weights are given by the following table*

			maximum-sized sets of ***deviating*** *item(s)*					
item			$\Gamma = 1$			$\Gamma = 2$		
j	$\bar{w}_j$	$\hat{w}_j$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
1	*1*	*2*	***3***	*1*	*1*	***3***	***3***	*1*
2	*3*	*4*	*3*	***7***	*3*	***7***	*3*	***7***
3	*4*	*4*	*4*	*4*	***8***	*4*	***8***	***8***
removed items by recovery			*{3}*	*{2}*	*{3}*	*{2}*	*{3}*	*{3}*
total weight after recovery			*6*	*5*	*4*	*7*	*6*	*8*

For both values of Γ*, all extremal scenarios in* $\mathcal{S}_\Gamma$ *are shown. The deviating item weights are highlighted in bold. The bottom rows state the items removed by the recovery action and the total weight of the remaining items after the recovery has taken place. For* $\Gamma = 1$*, the scenario where item 1 deviates yields the maximum weight set with total weight 6. For* $\Gamma = 2$*, the maximum weight of 8 is realized in the scenario where items 2 and 3 are deviating from the nominal weights. In contrast to* $\Gamma = 1$*, item 1 does not deviate. The optimal solutions of the two MWSP for the two values of* Γ *are not related.*

Nevertheless, the MWSP can be solved efficiently. Let us recall that the maximum term in constraint (7.2c) is a mathematical formulation of the MWSP based on which we obtain the following equivalent mathematical formulation:

$$\max_{y \in \{0,1\}^{|X|}} \left\{ \sum_{j \in X} \hat{w}_j y_j - \max_{z \in \{0,1\}^{|X|}} \left\{ \sum_{j \in X} (\bar{w}_j + \hat{w}_j y_j) z_j \ : \ \sum_{j \in X} z_j \leq k \right\} : \sum_{j \in X} y_j \leq \Gamma \right\} \quad (7.3)$$

Indicator variables y_j model the Γ-robustness part, i. e., $y_j = 1$ if the weight of item $j \in N$ deviates to its maximum, and 0 otherwise. Similarly, indicator variables z_j are used to model the k-recovery part, i. e., $z_j = 1$ if item $j \in X$ is removed from X when the recovery rule is applied, and 0 otherwise.

The coefficient matrix of the inner maximization problem in Formulation (7.3) is totally unimodular. Hence, (7.3) can be solved by its linear relaxation for a fixed value of y. Exploiting LP duality, cf. Section 3.1.2, yields a compact ILP formulation for the MWSP. It reads

$$\max \sum_{j \in X} \hat{w}_j y_j - k \cdot u - \sum_{j \in X} v_j \quad (7.4a)$$

$$\text{s. t.} \sum_{j \in X} y_j \leq \Gamma \quad (7.4b)$$

$$\hat{w}_j y_j - u - v_j \leq -\bar{w}_j \qquad \forall j \in X \quad (7.4c)$$

$$y_j \in \{0,1\}, \ u, v_j \geq 0 \qquad \forall j \in X \quad (7.4d)$$

with dual variables u and v_j corresponding to the constraints $\sum_{j \in X} z_j \leq k$ and $z_j \leq 1$, respectively. The following two lemmas characterizes optimal solutions of (7.4) differently providing a basis for an efficient algorithm to solve the MWSP.

Lemma 7.10. *For a fixed parameter u', define $w_j(u') := \min\{\hat{w}_j, -\bar{w}_j + u'\}$ for all $j \in X$, $X^-(u') := \{j \in X : w_j(u') < 0\}$, and let $X(u') \subseteq X \setminus X^-(u')$ maximizing $\sum_{j \in X(u')} w_j(u')$ with $|X(u')| \leq \Gamma$. Then, the optimal solution value $z(u)$ of (7.4) equals*

$$\sum_{j \in X(u')} w_j(u') + \sum_{j \in X^-(u')} w_j(u') + k \cdot u'.$$

Proof. Let $u' \geq 0$ be some value and let us define

$$c(u') := \sum_{j \in X(u')} w_j(u') + \sum_{j \in X^-(u')} w_j(u') - k \cdot u'.$$

We show $z(u') = c(u')$. Let $y_j^* \in \{0,1\}$ and $v_j^* \geq 0$, for all $j \in X$, be an optimal solution of the ILP (7.4) with $u = u'$. If $-\bar{w}_j + u' < 0$ for some $j \in X$, $\hat{w}_j \cdot y_j^* - v_j^* = -\bar{w}_j + u'$ due to the coefficients of y_j^* and v_j^* in the objective. If $y_i^* = 1$ for an item i with $-\bar{w}_i + u' < 0$, the solution $\overline{y}_i = 0$, $\overline{y}_j = \overline{y}_j^*$ for $i \neq x$, $\overline{v}_i = \bar{w}_i - u'$, and $\overline{v}_j = \overline{v}_j^*$ for $i \neq x$ is also an optimal solution. Therefore, we can assume that w.l.o.g. $y_j^* = 0$ for all $j \in X^-(u')$.

Let $X(y^*) = \{j \in X \mid y_j^* = 1\}$. Then $Y^* \subseteq X \backslash X^-(u)$ with $|Y^*| \leq \Gamma$. If $y_j^* = 0$ for $j \in X \backslash X^-(u)$, then $v_j^* = 0$. If $y_j^* = 1$ for some $j \in X$, $\hat{w}_j - v_j \leq -\bar{w}_j + u'$. Hence, if $\hat{w}_j > -\bar{w}_j + u'$, $v_j^* = \hat{w}_j + \bar{w}_j - u'$ and $v_j^* = 0$ otherwise. Therefore,

$$\begin{aligned} z(u') &= \sum_{j \in X^-(u')} (-\bar{w}_j + u') + \sum_{j \in Y^*} \min\{-\bar{w}_j + u', \hat{w}_j\} - k \cdot u' \\ &\leq \sum_{j \in X^-(u')} w_j(u') + \sum_{j \in X(u')} w_j(u') - k \cdot u' = c(u'). \end{aligned}$$

□

Lemma 7.11. *There always exists an optimal solution (u^*, y^*, v^*) of (7.4) with $u^* \in U$ where $U := \{0\} \cup \{\bar{w}_j : j \in X\} \cup \{\bar{w}_j + \hat{w}_j : j \in X\}$.*

Proof. We show that there always exists an optimal solution (u^*, y^*, v^*) of (7.4) with $u^* \in U$, where the set U consists of the values $0, \bar{w}_j, \bar{w}_j + \hat{w}_j$, $j \in X$.

Let (u^*, y^*, v^*) be an optimal solution of (7.4) with $u^* \notin U$. Then, we consider $\underline{u} \in U$ with $\underline{u} = \arg\min\{u^* - u \mid u \in U, u < u^*\}$ and $\overline{u} \in U$ with $\overline{u} = \arg\min\{u - u^* \mid u \in U, u > u^*\}$. Define $\tilde{X} = \{j \in \{1, \ldots, n'\} \mid w_j(u^*) < \hat{w}_j\}$ and $r = |\tilde{X}|$. Since $u^* \notin U$, we obtain $-\bar{w}_j + u^* \neq \{\hat{w}_j, 0\}$ for $j \in X$, and for $j \in \tilde{X}$

$$w_j(\underline{u}) + (u^* - \underline{u}) = w_j(u^*) = w_j(\overline{u}) + (\overline{u} - u^*).$$

If $r < k$,

$$\begin{aligned} z(u^*) &= \sum_{j \in X(u^*)} w_j(u^*) + \sum_{j \in X^-(u^*)} w_j(u^*) - k \cdot u^* \\ &= \sum_{j \in X(\underline{u})} w_j(\underline{u}) + \sum_{j \in X^-(\underline{u})} w_j(\underline{u}) + r \cdot (u^* - \underline{u}) - k \cdot u^* \\ &= z(\underline{u}) + (r - k)(u^* - \underline{u}) \leq z(\underline{u}). \end{aligned}$$

A similar argument provides the result for $r \geq k$. □

Let U, $w_j(u)$ for $j \in N$ and $u \in U$, and $X^-(u)$ for $u \in U$ defined as in Lemmas 7.10 and 7.11. By the same lemma, constraint (7.2c) can equivalently be written as

$$\sum_{j\in N} \bar{w}_j x_j^0 + \max_{u\in U} \left\{ \sum_{j\in X^-} w_j(u)x_j^0 - ku + \max_{\substack{N'\in N\\|N'|\leq\Gamma}} \sum_{j\in N'} w_j(u)x_j^0 \right\} \leq c. \tag{7.5}$$

Linearizing the outer maximum and simplifying the inner maximum results in the following set of constraints:

$$\sum_{j\in N} \bar{w}_j x_j^0 + \sum_{j\in X^-} w_j(u)x_j^0 + \max \sum_{j\in N} w_j(u)x_j^0 y_j^u \leq c + ku \qquad \forall u \in U \tag{7.6a}$$

$$\sum_{j\in N} y_j^u \leq \Gamma \qquad \forall u \in U \tag{7.6b}$$

$$y_j^u \in \{0,1\} \qquad \forall j \in N, u \in U. \tag{7.6c}$$

where y_j^u indicates if the item j is set to its peak value in the scenario characterized by u. Only the maximum term $\max \sum_{j\in N} w_j(u)x_j^0 y_j^u$ is left, which given x^0 can be linearized by dualization due to its totally unimodular structure. Based on the resulting set of linear constraints, we give a formulation of the k/Γ-RRKP differently from (7.2):

$$\max \sum_{j\in N} p_j^0 x_j^0 \tag{7.7a}$$

$$\text{s. t.} \sum_{j\in N} w_j^0 x_j^0 \leq c^0 \tag{7.7b}$$

$$\sum_{\substack{j\in N:\\ \bar{w}_j<u}} \bar{w}_j x_j^0 + \sum_{\substack{j\in N:\\ \bar{w}_j\geq u}} u x_j^0 + \Gamma\xi^u + \sum_{j\in N} \theta_j^u \leq c + ku \qquad \forall u \in U \tag{7.7c}$$

$$\min\{\hat{w}_j, -\bar{w}_j + u\} x_j^0 - \xi^u - \theta_j^u \leq 0 \qquad \forall j \in N, u \in U \tag{7.7d}$$

$$x_j^0 \in \{0,1\},\ \xi^u, \theta_j^u \geq 0 \qquad \forall j \in N, u \in U. \tag{7.7e}$$

Constraint (7.7b) models the first-stage knapsack capacity. The second-stage knapsack capacity together with the recovery rule are enforced by constraints (7.7c) and (7.7d). The variable bounds are given by (7.7e). Formulation (7.7) is compact: it contains $\mathcal{O}(n^2)$ variables and $\mathcal{O}(n^2)$ constraints depending on the number of different values of $\bar{w}_j$ and $\bar{w}_j + \hat{w}_j$ for all $j \in N$.

7.2 Polyhedral study

In this section, we investigate the polyhedral structure of the RRKP and its special cases, the $k, \ell/D$-RRKP and the k/Γ-RRKP. Therefore, we define the general recoverable robust knapsack polytope as the convex hull over all valid first-stage solutions.

Definition 7.12 (Recoverable Robust Knapsack Polytope)**.** Let $\mathcal{S}$ be a scenario set and $\mathcal{R}$ a recovery set. Then, the *recoverable robust knapsack polytope* is defined as

$$\mathcal{K}^{RR} := \operatorname{conv} \left\{ \begin{array}{l} x \in \{0,1\}^n \ : \ \sum_{j \in N} w_j^0 x_j \leq c^0, \\ \qquad \min_{N' \subseteq \mathcal{R}(\{j \in N \,:\, x_j = 1\})} \sum_{j \in N'} w_j^S x_j \leq c^S \ \forall S \in \mathcal{S} \end{array} \right\}.$$

The corresponding polytopes for the special cases $k,\ell/D$-RRKP and k/Γ-RRKP are defined as follows.

Definition 7.13 (selected recoverable robust knapsack polytopes)**.** The *k,ℓ-recoverable robust knapsack polytope* is defined as

$$\mathcal{K}^{k,\ell/D} := \operatorname{conv} \left\{ \begin{array}{l} x \in \{0,1\}^n \ : \ \sum_{j \in N} w_j^0 x_j \leq c^0, \\ \qquad \min_{\substack{N' \subseteq N \\ |N'| \leq k}} \sum_{j \in N \setminus N'} w_j^S x_j \leq c^S \ \forall S \in \mathcal{S}_D \end{array} \right\}.$$

The *k-recoverable Γ-robust knapsack polytope* is defined as

$$\mathcal{K}^{k/\Gamma} := \operatorname{conv} \left\{ \begin{array}{l} x \in \{0,1\}^n \ : \ \sum_{j \in N} w_j^0 x_j \leq c^0, \\ \qquad \min_{\substack{N' \subseteq N \\ |N'| \leq k}} \sum_{j \in N \setminus N'} w_j^S x_j \leq c \ \forall S \in \mathcal{S}_\Gamma \end{array} \right\}$$

Note, that we defined the recoverable robust polytopes in the space of the first-stage variables analogously to the polytopes of the KP, Γ-RKP, mb-RKP, and SMKP.

7.2.1 Basic characteristics

First, we consider the dimensions of the polytopes defined above. Second, we investigate the conditions for which the variable bounds imply trivial facets for these polytopes.

Lemma 7.14. *$\mathcal{K}^{RR}$ is full-dimensional if and only if the following two conditions hold*

1. *$w_j^0 \leq c^0$ for all $j \in N$,*
2. *$w_j^S \leq c^S$ for all $j \in \{i \in N \ : \ \{i\} \in N'$ for all $N' \in \mathcal{R}(\{i\})\}$, $S \in \mathcal{S}$.*

Proof. All unit vectors and the zero vector are feasible for RRKP and affinely independent. □

Corollary 7.15. *$\mathcal{K}^{k,\ell/D}$ is full-dimensional if and only if the following two conditions hold*

1. *$w_j^0 \leq c^0$ for all $j \in N$,*
2. *if $k = 0$, then $w_j^S \leq c^S$ for all $j \in N$, $S \in \mathcal{S}_D$.*

$\mathcal{K}^{k/\Gamma}$ is full-dimensional if and only if the following two conditions hold

1. *$w_j^0 \leq c^0$ for all $j \in N$,*

2. *if* $k = 0$, *then* $\bar{w}_j + \hat{w}_j \le c^S$ *for all* $j \in N$, $S \in \mathcal{S}_\Gamma$.

Proof. All unit vectors and the zero vector are feasible for $k, \ell/D$-RRKP and k/Γ-RRKP. Furthermore, they are affinely independent.

Let $\mathcal{K}^{k,\ell/D}$ ($\mathcal{K}^{k/\Gamma}$) be full-dimensional. Suppose there exists a $i \in N$ with $w_i^0 > c^0$, then $x_i = 0$ for all $x \in \mathcal{K}^{k,\ell/D}$ ($\mathcal{K}^{k/\Gamma}$). This is a contradiction to the polytope being full-dimensional. Next, suppose there exists another $i \in N$ with $w_i^S > c^S$ for $PrrkpKLDiscrete$ (or $\bar{w}_i + \hat{w}_i > c^S$ for $\mathcal{K}^{k/\Gamma}$, respectively). Then $x_i = 0$ for all $x \in \mathcal{K}^{k,\ell/D}$ ($\mathcal{K}^{k/\Gamma}$); again contradicting the precondition that the polytope is full-dimensional. This completes the proof. □

W. l. o. g. we assume RRKP to be full-dimensional in the rest. By studying the variable bounds $x_j \ge 0$ and $x_j \le 1$, we obtain the following results.

Lemma 7.16. *Trivial facets of the recoverable robust knapsack polytope* $\mathcal{K}^{RR}$ *are*

$$x_j \ge 0 \quad \forall j \in N \tag{7.8}$$

and

$$x_j \le 1 \quad \forall j \in N \,:\, w_j^0 + w_i^0 \le c^0 \;\forall i \in N \setminus \{j\}. \tag{7.9}$$

Proof. For $x_j \ge 0$, the zero vector and all unit vectors e_i for $i \in N \setminus \{j\}$ are feasible for the RRKP, affinely independent, and satisfy inequality (7.8) with equality.

For $x_j \le 1$, the unit vector e_j and the points $e_j + e_i$ for all $i \in N \setminus \{j\}$ are feasible, affinely independent and satisfy inequality (7.9) with equality. □

Corollary 7.17. *The inequalities*

$$x_j \ge 0 \quad \forall j \in N \tag{7.10}$$

are trivial facets of $\mathcal{K}^{k,\ell/D}$. *Furthermore, the inequality*

$$x_j \le 1 \tag{7.11}$$

is a trivial facet of $\mathcal{K}^{k,\ell/D}$ *for all* $j \in N$ *for which the following conditions hold*

1. $w_j^0 + \max\limits_{i \in N \setminus \{j\}} w_i^0 \le c^0$,

2. *if* $k = 0$, *then* $w_j^S + \max\limits_{i \in N \setminus \{j\}} w_i^S \le c^S$ *for all* $S \in \mathcal{S}_D$,

3. *if* $k = 1$, *then* $\min\{w_j^S, \max\limits_{i \in N \setminus \{j\}} w_i^S\} \le c^S$ *for all* $S \in \mathcal{S}_D$.

Proof. The zero vector and all unit vectors e_i for $i \in N \setminus \{j\}$ are feasible for the $k, \ell/D$-RRKP, affinely independent, and satisfy inequality (7.10) with equality.

The unit vector e_j and all vectors $e_j + e_i$ for $i \in N \setminus \{j\}$ are feasible for the $k, \ell/D$-RRKP, affinely independent, and satisfy inequality (7.11) with equality. □

Lemma 7.18. *The inequalities*

$$x_j \geq 0 \quad \forall j \in N \tag{7.12}$$

are trivial facets of $\mathcal{K}^{k/\Gamma}$. *Furthermore, the inequality*

$$x_j \leq 1 \tag{7.13}$$

is a trivial facet of $\mathcal{K}^{k/\Gamma}$ *for all* $j \in N$ *for which the following conditions hold*

1. $w_j^0 + \max_{i \in N \setminus \{j\}} w_i^0 \leq c^0$,

2a. *if* $k = 0$, $\Gamma = 0$, *then* $\bar{w}_j + \max_{i \in N \setminus \{j\}} \bar{w}_i \leq c$

2b. *if* $k = 0$, $\Gamma = 1$, *then* $\bar{w}_j + \max_{i \in N \setminus \{j\}} (\bar{w}_i + \max\{\hat{w}_j, \hat{w}_i\}) \leq c$

2c. *if* $k = 0$, $\Gamma \geq 2$, *then* $\bar{w}_j + \hat{w}_j + \max_{i \in N \setminus \{j\}} (\bar{w}_i + \hat{w}_i) \leq c$

3a. *if* $k = 1$, $\Gamma = 0$, *then* $\min\{\bar{w}_j, \max_{i \in N \setminus \{j\}} \bar{w}_i\} \leq c$

3b. *if* $k = 1$, $\Gamma \geq 1$, *then* $\min\{\bar{w}_j + \hat{w}_j, \max_{i \in N \setminus \{j\}} (\bar{w}_i + \hat{w}_i)\} \leq c$

Proof. The zero vector and all unit vectors e_i for $i \in N \setminus \{j\}$ are feasible for the k/Γ-RRKP, affinely independent, and satisfy inequality (7.12) with equality.

The unit vector e_j and all vectors $e_j + e_i$ for $i \in N \setminus \{j\}$ are feasible for the k/Γ-RRKP, affinely independent, and satisfy inequality (7.13) with equality. □

7.2.2 Valid inequalities

In this section, we continue our polyhedral study and investigate valid or facet-defining inequalities besides the trivial ones. Motivated by the classic knapsack problem and its well-studied polyhedral structure, we also consider specially structured subsets of items of the RRKP generalizing (extended) covers and the corresponding valid inequalities to the recoverable robust setting. We base our study on the definitions introduced in our previous work; cf. [49, 50].

For the RRKP, a subset $\mathcal{C} \subseteq N$ is called a *(recoverable robust) cover* if at least one of the following conditions holds

1. $\mathcal{C}$ is a first-stage cover, i. e., its total first-stage weights exceed the first-stage capacity:
$$\sum_{j \in \mathcal{C}} w_j^0 \geq c^0 + 1$$

2. $\mathcal{C}$ is a second-stage or scenario cover, i. e., there exists a scenario such that the total weight of this scenario exceeds the capacity of this scenario:

$$\exists S \in \mathcal{S} : \min_{\mathcal{C}' \in \mathcal{R}(\mathcal{C})} \sum_{j \in \mathcal{C}'} w_j^S \geq c^S + 1.$$

As for to the classic knapsack problem, a cover $\mathcal{C}$ is called *minimal* if $\mathcal{C} \setminus \{j\}$ is not a cover for all $j \in \mathcal{C}$. Furthermore, it holds

Lemma 7.19. *Let $\mathcal{C} \subseteq N$ be a cover. Then, the* cover inequality

$$\sum_{j \in \mathcal{C}} x_j \leq |\mathcal{C}| - 1$$

is valid for $\mathcal{K}^{RR}$.

Proof. Suppose $x \in \mathcal{K}^{RR}$ does not satisfy the cover inequality. Then $|\mathcal{C}| \leq \sum_{j \in \mathcal{C}} x_j$ implies $x_j = 1$ for all $j \in \mathcal{C}$. If $\mathcal{C}$ is a first-stage cover, then

$$\sum_{j \in N} w_j^0 x_j = \sum_{j \in \mathcal{C}} w_j^0 x_j + \sum_{j \in N \setminus \mathcal{C}} w_j^0 x_j \geq c^0 + 1.$$

A contradiction to $x \in \mathcal{K}^{RR}$. If $\mathcal{C}$ is a second-stage cover, an analogous contradiction can be obtained completing the proof. □

Theorem 7.20. *Let $x \in \{0,1\}^n$. Then $x \in \mathcal{K}^{RR}$ if and only if x satisfies all minimal cover inequalities.*

Proof. Let $x \in \{0,1\}^n$ satisfy all minimal cover inequalities and $N_x \subseteq N$ be its support, i. e., $N_x := \{j \in N : x_j = 1\}$. Suppose $x \notin \mathcal{K}^{RR}$, i. e., it violates the first or second stage knapsack constraint. Thus, the set N_x is a cover. By iteratively removing the item with minimum weight, N_x becomes a minimal cover whose corresponding minimal cover inequality is not satisfied by x; a contradiction.

Conversely, consider a binary feasible point $x \in \mathcal{K}^{RR}$. Then it satisfies all cover inequalities since they are valid. □

Our previous definition of a cover has two variable aspects not included in the definition: if the cover is a scenario cover, the scenario for which the capacity is exceeded is not known beforehand. Moreover, the best recovery action and thus the recovered cover is only implicitly given by the min-term. To overcome these aspects and to give a more explicit description of a cover, we introduce an alternative notation where the scenario of a scenario cover as well as the recovery is fixed in the definition of the cover. Therefore, let $\mathcal{C} \subseteq N$, $K \subseteq \mathcal{C}$, and $S \in \mathcal{S}$. Then, $(\mathcal{C}, S, K)$ is a cover if at least one of the following conditions is true.

1. $\mathcal{C}$ is a first-stage cover, i. e., $\sum_{j \in \mathcal{C}} w_j^0 \geq c^0 + 1$,

2. $(\mathcal{C}, S, K)$ is a second-stage cover with $\mathcal{C} \setminus K = \arg\min_{\mathcal{C}' \in \mathcal{R}(\mathcal{C})} \sum_{j \in \mathcal{C}'} w_j^S$, i.e., $\sum_{j \in \mathcal{C} \setminus K} w_j^S \geq c^S + 1$.

We apply this alternative definition of a cover to the special cases $k, \ell/D$-RRKP and k/Γ-RRKP of the RRKP. The resulting definitions take the specific aspects of the special cases better into account.

$k, \ell/D$-RRKP. We consider the $k, \ell/D$-RRKP, subsets of its items $\mathcal{C} \subseteq N$, $K \subseteq \mathcal{C}$, and a scenario $S \in \mathcal{S}$. Then, the tuple $(\mathcal{C}, S, K)$ is a cover if at least one of the following conditions is true.

1. $\mathcal{C}$ is a first-stage cover, i.e., $\sum_{j \in \mathcal{C}} w_j^0 \geq c^0 + 1$
2. $(\mathcal{C}, S, K)$ is a scenario cover, i.e., all of the following conditions hold
   a) $|K| = k$
   b) $w_j^S \geq w_i^S$ for all $j \in K$ and $i \in \mathcal{C} \setminus K$
   c) $\sum_{j \in \mathcal{C} \setminus K} w_j^S \geq c^S + 1$

Corollary 7.21. *Let $(\mathcal{C}, S, K)$ be a cover. Then, the* cover inequality

$$\sum_{j \in \mathcal{C}} x_j \leq |\mathcal{C}| - 1$$

is valid for $\mathcal{K}^{k,\ell/D}$.

Proof. Follows directly from Lemma 7.19. □

In the following, we present two methods to extend a given cover and thereby to strengthen its implied cover inequality. Note, that both extension methods are scenario-dependent. Let $(\mathcal{C}, S, K)$ be a cover. Then, a canonical extension is given by

$$E(\mathcal{C}, S, K) = \mathcal{C} \cup \left\{ j \in N \ : \ w_j^0 \geq \max_{i \in \mathcal{C} \setminus K} w_i^0, \ w_j^S \geq \max_{i \in \mathcal{C} \setminus K} w_i^S \right\} \tag{7.14}$$

To obtain a strengthened extension, we make the following observations. Consider the non-recovered items $\mathcal{C} \setminus K$. It is not necessary that an item of the extension has a weight at least maximum weight of $\mathcal{C} \setminus K$ since it may replace the item with maximum weight by the exchange argument of an extension. Instead it is necessary that the weight of the newly added item is at least the second largest weight of an item in $\mathcal{C} \setminus K$. However, by replacing the maximum weight item in $\mathcal{C} \setminus K$ the total weight of $\mathcal{C} \setminus K$ might be lowered so that it does not exceed the knapsack capacity anymore. To avoid this, a second condition is needed relating the maximum weight item in $\mathcal{C} \setminus K$ to the knapsack capacity. In summary, a strengthened extension can be obtained in the following way: When considering the non-recovered items $\mathcal{C} \setminus K$, by only adding all items whose weight is at least the residual capacity according to the total weight of the first $|\mathcal{C}| - k - 1$ items with lowest weight and whose weight is at least the second-largest weight in $\mathcal{C} \setminus K$. This yields the following definition of a strengthened extension

$$E^+(\mathcal{C}, S, K) = \mathcal{C} \cup \left\{ \begin{array}{lll} j \in N & : & w_j^S \geq c^S - \sum\limits_{i \in \mathcal{C} \setminus K} w_i^S + \max\limits_{i \in \mathcal{C} \setminus K} w_i^S + 1, \\ & & w_j^S \geq \max\limits_{\substack{\mathcal{C}' \subseteq \mathcal{C} \\ |\mathcal{C}'| = k+2}} \sum\limits_{i \in \mathcal{C}'} w_i^S - \max\limits_{\substack{\mathcal{C}' \subseteq \mathcal{C} \\ |\mathcal{C}'| = k+1}} \sum\limits_{i \in \mathcal{C}'} w_i^S \end{array} \right\} \tag{7.15}$$

Of course, an extended cover is a cover itself.

Example 7.22. *Consider a scenario S of a $k, \ell/D$-RRKP instance with 5 items, $k = 1$, and $\ell = 0$. Let the scenario weights be $w^S = \begin{pmatrix} 4 & 6 & 6 & 5 & 6 \end{pmatrix}$ and the scenario capacity $c^S = 8$. Then, $(\mathcal{C}, S, K)$ with $\mathcal{C} = \{1, 2, 3\}$ and $K = \{3\}$ is a scenario cover for S with total weight 10 exceeding the scenario capacity. Its canonical extension is obtained by addition of all items with scenario weight at least $\max_{j \in \mathcal{C}} w_j^S = 6$. Hence, $E(\mathcal{C}, S, K) = \mathcal{C} \cup \{5\}$. In contrast, its strengthened extension $E^+(\mathcal{C}, S, K) = \mathcal{C} \cup \{4, 5\}$ allows the addition of item 4 as well (the first condition in the definition of the strengthened extension requires the weight of the additional item to be at least 5 in this case, the second condition asks for an item with weight at least 4 here).*

Extended covers give rise to the following class of valid inequalities.

Lemma 7.23. *Let $(\mathcal{C}, S, K)$ be a cover together with its corresponding canonically extended cover $E(\mathcal{C}, S, K)$ and its strengthened extended cover $E^+(\mathcal{C}, S, K)$. Then, the* extended cover inequalities

$$\sum_{j \in E(\mathcal{C}, S, K)} x_j \leq |\mathcal{C}| - 1$$

and

$$\sum_{j \in E^+(\mathcal{C}, S, K)} x_j \leq |\mathcal{C}| - 1$$

are valid for $\mathcal{K}^{k, \ell/D}$.

Proof. This proof is similar to the proof of Lemma 4.6 for the Γ-RKP.

Let $E^* \in \{E(\mathcal{C}, S, K), E^+(\mathcal{C}, S, K)\}$ be an extended cover. Suppose the corresponding (strengthened) extended cover is not valid. Then there exists a $N' \subseteq E^*$, $|N'| \geq |C|$ so that

$$\sum_{j \in N'} w_j^S - \max_{\substack{K' \subseteq N': \\ |K'| \leq k}} \sum_{j \in K'} w_j^S \leq c^S$$

holds. Let $\tilde{x} \in \mathcal{K}^{k, \ell/D} \cap \mathbb{Z}^{|N|}$ the characteristic vector of N', i. e., $\tilde{x}_j = 1$ if and only if $j \in N'$. Then, it follows

$$\begin{aligned}
c^S &\geq \sum_{j \in N} w_j^S \tilde{x}_j - \max_{\substack{K' \subseteq N,: \\ |K'| \leq k}} \sum_{j \in K'} w_j^S \tilde{x}_j \\
&= \sum_{j \in N'} w_j^S - \max_{\substack{K' \subseteq N': \\ |K'| \leq k}} \sum_{j \in K'} w_j^S \\
&\geq \min_{\substack{N^{\min} \subseteq E^*: \\ |N^{\min}| = |C|}} \left\{ \sum_{j \in N^{\min}} w_j^S - \max_{\substack{K' \subseteq N^{\min}: \\ |K'| \leq k}} \sum_{j \in K'} w_j^S \right\} \\
&= \sum_{j \in C \setminus K} w_j^S > c.
\end{aligned}$$

This is a contradiction to $\tilde{x} \in \mathcal{K}^{k,\ell/D}$ and completes the proof. □

k/Γ-RRKP. Next, we identify classes of valid inequalities for the k/Γ-RRKP. We start with the definition of a cover for this problem. Let $\bar{\mathcal{C}} \subseteq N$, $\bar{K} \subseteq \bar{\mathcal{C}}$, $\hat{\mathcal{C}} \subseteq N$, $\hat{K} \subseteq \hat{\mathcal{C}}$ with $\bar{\mathcal{C}} \cap \hat{\mathcal{C}} = \emptyset$. Then, the tuple $(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ is a cover if at least one of the following conditions is true

1. $\bar{\mathcal{C}} \cup \hat{\mathcal{C}}$ is a first-stage cover, i. e., $\sum_{j \in \bar{\mathcal{C}} \cup \hat{\mathcal{C}}} w_j^0 \geq c^0 + 1$
2. $(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ is a scenario cover, i. e., all of the following conditions hold
   a) $|\hat{\mathcal{C}}| \leq \Gamma$
   b) $\sum_{j \in \bar{\mathcal{C}} \setminus \bar{K}} \bar{w}_j + \sum_{j \in \hat{\mathcal{C}} \setminus \hat{K}} (\bar{w}_j + \hat{w}_j) \geq c + 1$
   c) $|\bar{K}| + |\hat{K}| = k$
   d) $\bar{w}_j \geq \bar{w}_i$ for all $j \in \bar{K}$ and $i \in \bar{\mathcal{C}} \setminus \bar{K}$
   e) $\bar{w}_j \geq \bar{w}_i + \hat{w}_i$ for all $j \in \bar{K}$ and $i \in \hat{\mathcal{C}} \setminus \hat{K}$
   f) $\bar{w}_j + \hat{w}_j \geq \bar{w}_i + \hat{w}_i$ for all $j \in \hat{K}$ and $i \in \hat{\mathcal{C}} \setminus \hat{K}$
   g) $\bar{w}_j + \hat{w}_j \geq \bar{w}_i$ for all $j \in \hat{K}$ and $i \in \bar{\mathcal{C}} \setminus \bar{K}$

Next, we define the corresponding recoverable robust cover inequality for the k/Γ-RRKP.

Corollary 7.24. *Let $(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ be a cover. Then, the* cover inequality

$$\sum_{j \in \bar{\mathcal{C}} \cup \hat{\mathcal{C}}} x_j \leq |\bar{\mathcal{C}} \cup \hat{\mathcal{C}}| - 1$$

is valid for $\mathcal{K}^{k/\Gamma}$.

Proof. Follows directly from Lemma 7.19. □

As for the classic knapsack problem, recoverable robust covers may be extended and thus the corresponding cover inequality strengthened. Therefore let

$$\bar{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) := \left\{ j \in N \ : \ \bar{w}_j \geq \max_{\substack{i \in \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \\ i \notin \bar{K} \cup \hat{K}}} \bar{w}_i \right\} \tag{7.16}$$

$$\hat{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) := \left\{ j \in N \ : \ \bar{w}_j + \hat{w}_j \geq \max_{\substack{i \in \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \\ i \notin \bar{K} \cup \hat{K}}} \bar{w}_i + \hat{w}_i \right\} \tag{7.17}$$

be the set $\bar{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ of all items not in the cover whose nominal weight is at least the maximum nominal weight of the not-recovered items of the cover and the set $\hat{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ of all items not in the cover whose peak weight is at least the maximum peak weight of the not-recovered items of the cover, respectively. Klopfenstein and Nace [97] propose the following extension of covers for Γ-RKP which we adopted to the recoverable robust setting.

$$E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) = \begin{cases} \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \cup \hat{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) & \text{if } |\bar{\mathcal{C}}| \leq \Gamma \\ \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \cup (\bar{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) \cap \hat{N}(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})) & \text{if } |\bar{\mathcal{C}}| \geq \Gamma + 1 \end{cases} \tag{7.18}$$

This extension can be strengthened by adding all items whose nominal weight is at least the maximum nominal weight in $\bar{\mathcal{C}} \setminus \bar{K}$ and whose peak weight is at least the maximum peak weight in $\hat{\mathcal{C}} \setminus \hat{K}$. This leads to the following definition of a strengthened extended cover.

$$E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) := \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \cup \left\{ j \in N \ : \ \bar{w}_j \geq \max_{i \in \bar{\mathcal{C}} \setminus \bar{K}} \bar{w}_i, \ \bar{w}_j + \hat{w}_j \geq \max_{i \in \hat{\mathcal{C}} \setminus \hat{K}} \bar{w}_i + \hat{w}_i \right\} \tag{7.19}$$

Note, a (strengthened) extended cover is a cover itself.

Example 7.25. *Consider an instance of k/Γ-RRKP with 5 items, $k = 1$, $\Gamma = 1$, and $c = 10$. Let the nominal and deviation weights of the items be $\bar{w} = \begin{pmatrix} 8 & 5 & 3 & 6 & 7 \end{pmatrix}$ and $\hat{w} = \begin{pmatrix} 1 & 2 & 3 & 6 & 8 \end{pmatrix}$, respectively. Then $(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ with $\bar{\mathcal{C}} = \{1, 2\}$, $\bar{K} = \{1\}$, $\hat{\mathcal{C}} = \{3\}$, and $\hat{K} = \emptyset$ is a cover for this instance. Its weight of 11 exceeds the scenario capacity by 1. This cover can be canonically extended by adding all items with nominal weight of at least 5, and peak weight of at least 7. Therefore, $E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) = \bar{\mathcal{C}} \cup \hat{\mathcal{C}} \cup \{5\}$. The strengthened extension allows the addition of one more item since items with peak weight 6 may be added as well, thus $E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) = E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) \cup \{4\}$.*

Based on extended covers, we formulate the following class of valid inequalities.

Lemma 7.26. *Let $(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ be a cover together with its corresponding extended cover $E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ and its strengthened extended cover $E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$. Then, the* extended cover inequalities

$$\sum_{j \in E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})} x_j \leq |\bar{\mathcal{C}} \cup \hat{\mathcal{C}}| - 1$$

and

$$\sum_{j \in E^+(\bar{\mathcal{C}},\bar{K},\hat{\mathcal{C}},\hat{K})} x_j \leq |\bar{\mathcal{C}} \cup \hat{\mathcal{C}}| - 1$$

are valid for $\mathcal{K}^{k/\Gamma}$.

Proof. This proof is similar to the proof of Lemma 4.6 for the Γ-RKP (the special case of $k = 0$).

Let $E^* \in \{E(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}), E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})\}$ be an extended cover. Suppose the corresponding (strengthened) extended cover is not valid. Then there exists a $N' \subseteq E^*$, $|N'| \geq |\bar{C} \cup \hat{C}|$ so that

$$\sum_{j \in N'} \bar{w}_j + \max_{\substack{\hat{N}' \subseteq N':\\ |\hat{N}'| \leq \Gamma}} \left(\sum_{j \in \hat{N}'} \hat{w}_j - \max_{\substack{K' \subseteq N':\\ |K'| \leq k}} \sum_{j \in K'} (\bar{w}_j + \hat{w}_j) \right) \leq c$$

holds. Let $\tilde{x} \in \mathcal{K}^{k/\Gamma} \cap \mathbb{Z}^{|N|}$ the characteristic vector of N', i. e., $\tilde{x}_j = 1$ if and only if $j \in N'$. Then, it follows

$$\begin{aligned}
c &\geq \sum_{j \in N} \bar{w}_j \tilde{x}_j + \max_{\substack{\hat{N} \subseteq N:\\ |\hat{N}| \leq \Gamma}} \left(\sum_{j \in \hat{N}'} \hat{w}_j \tilde{x}_j - \max_{\substack{K \subseteq N:\\ |K| \leq k}} \sum_{j \in K} (\bar{w}_j + \hat{w}_j) \tilde{x}_j \right) \\
&= \sum_{j \in N'} \bar{w}_j + \max_{\substack{\hat{N}' \subseteq N':\\ |\hat{N}'| \leq \Gamma}} \left(\sum_{j \in \hat{N}'} \hat{w}_j - \max_{\substack{K' \subseteq N':\\ |K'| \leq k}} \sum_{j \in K'} (\bar{w}_j + \hat{w}_j) \right) \\
&\geq \min_{\substack{N^{\min} \subseteq E^*:\\ |N^{\min}| = |\bar{C} \cup \hat{C}|}} \left\{ \sum_{j \in N^{\min}} \bar{w}_j + \max_{\substack{\hat{N}' \subseteq N^{\min}:\\ |\hat{N}'| \leq \Gamma}} \left(\sum_{j \in \hat{N}'} \hat{w}_j - \max_{\substack{K' \subseteq N^{\min}:\\ |K'| \leq k}} \sum_{j \in K'} (\bar{w}_j + \hat{w}_j) \right) \right\} \\
&= \sum_{j \in \bar{C} \setminus \bar{K}} \bar{w}_j + \sum_{j \in \hat{C} \setminus \hat{K}} \hat{w}_j > c.
\end{aligned}$$

This is a contradiction to $\tilde{x} \in \mathcal{K}^{k/\Gamma}$ and completes the proof. □

7.3 Algorithms

In this section, we present algorithms for the RRKP, the MWSP, and the separation problem associated with the considered classes of valid inequalities. These algorithms are based on the theoretical results of our previous investigation.

7.3.1 Solving the Maximum Weight Set Problem

The MWSP asks for determining a subset $X' \subseteq X \subseteq N$ with maximum weight taking all possible realizations in $\mathcal{U}^{\mathcal{S}}$ and recovery actions defined by $\mathcal{R}^{k,\ell}(X)$ into account. By Lemma 7.10, we can determine the objective value of an optimal solution if u is fixed.

Furthermore by Lemma 7.11, there exists always a solution where u takes a value of a given finite set of possible values. Combining both results, we propose the following algorithm: first, determine the set U of possible values for u. Then, for each value, determine the corresponding objective value $z(u)$ of the MWSP with fixed u. Finally, return the maximum u and the corresponding item subset. This procedure is formalized as pseudo-code in Algorithm 7.3.1.

Algorithm 1: Maximum Weight Set Problem

Input : item set X, for all $j \in X$ nominal item weights $\bar{w}_j \in \mathbb{N}$, deviation item weights $\hat{w}_j \in N$, robustness parameter $\Gamma \in \mathbb{N}$, and recovery parameter $k \in \mathbb{N}$

Output : maximum weight set $X' \subseteq X$

```
set U = {0} ∪ ⋃_{j∈X} {w̄_j} ∪ ⋃_{j∈X} {w̄_j + ŵ_j} ;
forall u ∈ U do
    forall j ∈ X do
        set w_j(u) = min{ŵ_j, −w̄_j + u} ;
    set X⁻(u) = {j ∈ X : w_j(u) < 0} ;
    set X(u) = argmax_{X'⊆X\X⁻, |X'|≤Γ} Σ_{j∈X} w_j(u) ;
    set z(u) = Σ_{j∈X(u)} w_j(u) + Σ_{j∈X⁻} w_j(u) − k · u ;
return X(u_max) with u_max = argmax_{u∈U} z(u)
```

The algorithm has a run-time of $\mathcal{O}(|X|^2)$ because $|U| \leq 2|X| + 1$.

7.3.2 Separation of violated recoverable robust extended cover inequalities

The definition of covers for the RRKP always includes two cases: the cover may be a first-stage cover or a second-stage cover. First-stage covers are classic knapsack covers with respect to the first-stage weights and capacities. They can be separated using the well-known methods and algorithms for the KP; cf. Martello and Toth [119].

Therefore in this section, we focus on the new challenge to separate violated valid inequalities based on second-stage (extended) covers. We present exact separation algorithms to separate violated (extended) cover inequalities for both, the $k, \ell/D$-RRKP and the k/Γ-RRKP. In addition to ILP formulations, we develop dynamic programs solving these separation problems in pseudo-polynomial time.

$k, \ell/D$-RRKP. The extended covers we have presented for $k, \ell/D$-RRKP are scenario-dependent. Hence, the separation problem has to be solved for each scenario (or until a scenario with a violated extended scenario cover is found).

Let $(\mathcal{C}, S, K)$ be a cover and $E(\mathcal{C}, S, K)$ its canonical extension. Notice, $E(\mathcal{C}, S, K) \setminus \mathcal{C}$ consists of those items with weights at least the maximum weight of the items in $\mathcal{C} \setminus K$.

In Büsing et al. [49] we refer to the items $\mathcal{C} \setminus K$ as the *core of the cover* since they are not recovered and remain in the cover after the recovery rule has been applied. These core items contribute to the knapsack capacity and must satisfy the cover condition (i. e., exceed the knapsack capacity). Hence, our separation approaches focus on determining the core items. Further, we observe that if an item is added to the recovery K or the extension $E(\mathcal{C}, S, K) \setminus \mathcal{C}$, all items with larger weight can also be added to the extension as they are exchangeable by the first one. We exploit this fact of transitivity by grouping items with same weights in buckets and treating them uniformly. Therefore, let $\Omega = \bigcup_{j \in N}\{w_j^S\}$ be the set of different weight values in scenario S and let $T := \{1, \ldots, |\Omega|\}$ be an index set. Furthermore, let the different weight values be ordered increasingly, i. e., $0 \leq w_{j_1}^S < w_{j_2}^S < \cdots < w_{j_{|\Omega|}}^S \leq c^S$ and define $N_t := \left\{j \in N \ : \ w_j^S = w_{j_t}^S\right\}$ as the set of items in bucket t for all $t \in T$.

To formulate the separation problem of extended covers as an ILP, we introduce binary variables y_j and α_j for all $j \in N$ indicating whether an item is a core item ($j \in \mathcal{C} \setminus K$) or added to the cover as part of the extension ($j \in E(\mathcal{C}, S, K) \setminus \mathcal{C}$), respectively. Given an LP solution x^{*S}, we now formulate the separation problem as follows

$$\max \sum_{j \in N}(x_j^{*S} - 1)y_j + \sum_{j \in N} x_j^{*S}\alpha_j - k \tag{7.20a}$$

$$\text{s. t.} \sum_{j \in N} w_j^S y_j \geq c^S + 1 \tag{7.20b}$$

$$y_j + \alpha_j \leq 1 \qquad \forall j \in N \tag{7.20c}$$

$$\alpha_j \leq \alpha_i \qquad \forall j \in N_t, i \in N_{t+1}, t \in T \tag{7.20d}$$

$$y_j, \alpha_j \in \{0, 1\} \qquad \forall j \in N \tag{7.20e}$$

The objective (7.20a) maximizes the violation of the resulting extended cover inequality. The constraint (7.20b) ensures that the items with $y_N = 1$ exceed the scenario capacity. The remaining constraints (7.20c) and (7.20d) model the extension rule exploiting the simplification by using buckets of items.

A violated extended cover inequality is found if the objective value (7.20a) is greater than -1. Otherwise, a proof is given that such an inequality does not exist. Furthermore, since the extension determined by (7.20) may be empty, this proof is given for violated cover inequalities, too.

To speed up the solving of (7.20), several preprocessing rules can be applied to reduce the number of variables and constraints. In addition, it suffices to find a (non-optimal) feasible solution with objective values > 1. Found (extended) covers can be extended further greedily.

Note, by fixing $\alpha_j = 0$ for all $j \in N$, we force $E(\mathcal{C}, S, K) \setminus \mathcal{C} = \emptyset$ and thus restrict ourselves to the separation of cover inequalities.

Given a solution with objective greater than -1, a violated extended cover $E(\mathcal{C}, S, K)$ can be constructed in the following way: define $\mathcal{C} := \{j \in N \ : \ y_j = 1\}$, the removed items $K := \arg\max_{N' \subseteq \mathcal{C},\, |N'| \leq k} \sum_{j \in N'} w_j^S$, and the extension $E(\mathcal{C}, S, K) := \mathcal{C} \cup \{j \in N \ : \ \alpha_j = 1\}$.

Based on the ILP (7.20), we formulate a dynamic program solving the separation problem. Therefore, we define the function

$$
\begin{aligned}
f_\omega(t,d) := \max \quad & \sum_{j=1}^{t}(x_j^{*S}-1)y_j + \sum_{j=1}^{t} x_j^{*S}\alpha_j \\
\text{s.t.} \quad & \sum_{j\in N} w_j^S y_j = d \\
& y_j + \alpha_j \le 1 && \forall j = 1,\dots,t \\
& \alpha_j = 0 && \forall j = 1,\dots,t : w_j^S < \omega \\
& y_j = 0 && \forall j = 1,\dots,t : w_j^S > \omega \\
& y_j, \alpha_j \in \{0,1\} && \forall j = 1,\dots,t.
\end{aligned}
$$

determining the optimal objective value of the separation ILP by restricting to the first t items, setting the total sum of scenario weights of items in $\mathcal{C}\setminus K$ to d, and limiting the maximal (minimal) weight of an item in $\mathcal{C}\setminus K$ (resp. the extension) to ω. Then, the optimal solution of the separation problem (7.20) can be obtained as

$$\max_{\omega\in\Omega,\ d=c^S+1,\dots,2c^S+1} f_\omega(n,d). \tag{7.22}$$

For $t = 1$ (i. e., only the first item), the function $f_\omega(1,d)$ can easily be evaluated considering the following three cases

$$
f_\omega(1,d) = \begin{cases} x_1^{*S}-1 & \text{if } w_1^S = d \text{ and } \omega > d \\ x_j^{*S} & \text{if } d = 0 \text{ and } w_1^S \ge \omega \\ -\infty & \text{otherwise} \end{cases}
$$

For $t \ge 2$ and $d = 0,\dots,2c^S+1$, the function $f_\omega(t,d)$ can be evaluated by the following recursive formula generalizing the three cases considered above.

$$
\begin{aligned}
f_\omega(t,d) = \max\{ & f_\omega(t-1,d), \\
& f_\omega(t-1,d-w_t^S) + (x_t^S - 1) && \text{if } w_t^S \le \omega, \\
& f_\omega(t-1,d) + x_t^S && \text{if } w_t^S \ge \omega\}
\end{aligned}
$$

In other words we decide, whether t is not in the core and not added to the extension, is part of the core or the cover, or is added to the extension taking the weight bound imposed by ω into account. Obviously, we can construct via an optimal solution of $f_\omega(t-1,d)$, $f_\omega(t-1,d-w_t^S)$, and $f_\omega(t-1,d)$ three feasible solutions for $f_\omega(t,d)$.

Because of $|\Omega| \le n$, the run-time of the dynamic program (7.22) is $\mathcal{O}(n^2c^S)$.

k/Γ-RRKP. Next, we consider the separation problems arising for the k/Γ-RRKP. To model the separation of violated strengthened extended covers for k/Γ-RRKP as ILP, we introduce five binary decision variables reflecting whether an item is in the cover at nominal/peak weight, or in the recovered set at nominal/peak weight, or not in the cover but in the extension. More formally, for all $j \in N$ let

- $\bar{y}_j = 1$ if and only if the item is at its nominal weight and not removed (i.e. $j \in \bar{\mathcal{C}} \setminus \bar{K}$),
- $\hat{y}_j = 1$ if and only if the item is at its peak weight and not removed (i.e. $j \in \hat{\mathcal{C}} \setminus \hat{K}$),
- $\bar{z}_j = 1$ if and only if the item is at its nominal weight and removed (i.e. $j \in \bar{K}$),
- $\hat{z}_j = 1$ if and only if the item is at its peak weight and removed (i.e. $j \in \hat{K}$), and
- $\alpha_j = 1$ if and only if the item is added to the extension and has not been in the cover (i.e. $j \in E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) \setminus (\bar{\mathcal{C}} \cup \hat{\mathcal{C}})$).

This gives rise to the following ILP formulation of this separation problem.

$$\max \sum_{j\in N} (x_j^{*0} - 1)(\bar{y}_j + \hat{y}_j + \bar{z}_j + \hat{z}_j) \quad + \sum_{j\in N} x_j^{*0} \alpha_j \tag{7.23a}$$

$$\text{s.t.} \sum_{j\in N} \bar{w}_j \bar{y}_j + \sum_{j\in N} (\bar{w}_j + \hat{w}_j)\hat{y}_j \quad \geq c + 1 \tag{7.23b}$$

$$\sum_{j\in N} (\bar{z}_j + \hat{z}_j) = k \tag{7.23c}$$

$$\sum_{j\in N} (\hat{y}_j + \hat{z}_j) = \Gamma \tag{7.23d}$$

$$\bar{y}_j + \hat{y}_j + \bar{z}_j + \hat{z}_j + \alpha_j \leq 1 \qquad \forall j \in N \tag{7.23e}$$

$$\bar{y}_i + \bar{z}_j + \alpha_j \leq 1 \qquad \forall j, i \in N : \bar{w}_i > \bar{w}_j \tag{7.23f}$$

$$\hat{y}_i + \bar{z}_j + \alpha_j \leq 1 \qquad \forall j, i \in N : \bar{w}_i + \hat{w}_i > \bar{w}_j \tag{7.23g}$$

$$\hat{y}_i + \hat{z}_j + \alpha_j \leq 1 \qquad \forall j, i \in N : \bar{w}_i + \hat{w}_i > \bar{w}_j + \hat{w}_j \tag{7.23h}$$

$$\bar{y}_i + \hat{z}_j + \alpha_j \leq 1 \qquad \forall j, i \in N : \bar{w}_i > \bar{w}_j + \hat{w}_j \tag{7.23i}$$

$$\bar{y}_j, \hat{y}_j, \bar{z}_j, \hat{z}_j, \alpha_j \in \{0, 1\} \qquad \forall j \in N \tag{7.23j}$$

where x^{*0} is the current LP solution. The violation of the resulting extended cover inequality is maximized by the objective (7.23a). The scenario covering property is enforced by (7.23b) whereas constraints (7.23c) and (7.23d) bound the size of the recovery set and the number of deviating item weights according to the recovery rule and scenario set. Furthermore, constraints (7.23e)–(7.23i) model the scenario set and recovery rule by implementing logical rules on the settings of the five indicator variables.

Note, by fixing $\alpha_j = 0$ for all $j \in N$, we force $E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K}) \setminus (\bar{\mathcal{C}} \cup \hat{\mathcal{C}}) = \emptyset$ and thus restrict to the separation of cover inequalities.

If ILP (7.23) has an optimal value greater than -1, a violated extended cover inequality is found. The corresponding extended cover $E^+(\bar{\mathcal{C}}, \bar{K}, \hat{\mathcal{C}}, \hat{K})$ is defined by the sets $\bar{\mathcal{C}} := \{j \in N \ : \ \bar{y}_j + \bar{z}_j = 1\}$, $\hat{\mathcal{C}} := \{j \in N \ : \ \hat{y}_j + \hat{z}_j = 1\}$, $\bar{K} := \{j \in N \ : \ \bar{z} = 1\}$, and $\hat{K} := \{j \in N \ : \ \hat{z} = 1\}$.

If the objective value of (7.23) is equal or less than -1, a proof is given that no violated cover inequality exists. Note that since the ILP formulation allows an empty extension, the proof of non-existence is really given not only for an extended cover inequality but also for a violated cover inequality.

Next, we present a pseudo-polynomial algorithm solving the separation problem. Therefore, let $\bar{\Omega} := \bigcup_{j\in N}\{\bar{w}_j\}$ be the set of different values of nominal item weights and $\hat{\Omega} := \bigcup_{j\in N}\{\bar{w}_j + \hat{w}_j\}$ be the set of different values of peak item weights. Define

$$
\begin{aligned}
&f_{\bar{\omega},\hat{\omega}}(t,\kappa,\gamma,d) := \\
&\qquad \max \sum_{j=1}^{t}(x_j^{*0}-1)(\bar{y}_j+\hat{y}_j+\bar{z}_j+\hat{z}_j) + \sum_{j=1}^{t} x_j^{*0}\alpha_j \\
&\qquad \text{s.t.} \sum_{j=1}^{t}\bar{w}_j\bar{y}_j + \sum_{j=1}^{t}(\bar{w}_j+\hat{w}_j)\hat{y}_j = d \\
&\qquad\qquad \sum_{j=1}^{t}(\bar{z}_j+\hat{z}_j) = \kappa \\
&\qquad\qquad \sum_{j=1}^{t}(\hat{y}_j+\hat{z}_j) = \gamma \\
&\qquad\qquad \bar{y}_j+\hat{y}_j+\bar{z}_j+\hat{z}_j+\alpha_j \le 1 && \forall j=1,\dots,t \\
&\qquad\qquad \bar{y}_i+\bar{z}_j+\alpha_j \le 1 && \forall j=1,\dots,t: \bar{w}_i > \bar{w}_j \\
&\qquad\qquad \hat{y}_i+\bar{z}_j+\alpha_j \le 1 && \forall j=1,\dots,t: \bar{w}_i+\hat{w}_i > \bar{w}_j \\
&\qquad\qquad \hat{y}_i+\hat{z}_j+\alpha_j \le 1 && \forall j=1,\dots,t: \bar{w}_i+\hat{w}_i > \bar{w}_j+\hat{w}_j \\
&\qquad\qquad \bar{y}_i+\hat{z}_j+\alpha_j \le 1 && \forall j=1,\dots,t: \bar{w}_i > \bar{w}_j+\hat{w}_j \\
&\qquad\qquad \bar{y}_j,\hat{y}_j,\bar{z}_j,\hat{z}_j,\alpha_j \in \{0,1\} && \forall j=1,\dots,t.
\end{aligned}
$$

as the function determining the optimal objective value (7.23a) when restricting this separation problem to the first t items, the removal of at most κ items due to recovery, the deviation of at most γ-many items from their nominal weights, setting the total weight of items in $\bar{\mathcal{C}} \cup \hat{\mathcal{C}}$ to d, and limiting the maximal nominal (peak) weight of an item to $\bar{\omega}$ ($\hat{\omega}$, respectively).

Then, the optimal solution of the separation ILP (7.23) can be determined by

$$
\max_{\bar{\omega}\in\bar{\Omega},\ \hat{\omega}\in\hat{\Omega},\ d=c+1,\dots,\sum_{j\in N}(\bar{w}_j+\hat{w}_j)} f_{\bar{\omega},\hat{\omega}}(n,k,\Gamma,d). \tag{7.25}
$$

When considering only one item, i. e., $t = 1$, the function $f_{\bar{\omega},\hat{\omega}}(1, \kappa, \gamma, d)$ can be evaluated directly as

$$f_{\bar{\omega},\hat{\omega}}(1, \kappa, \gamma, d) = \begin{cases} x_1^{*0} - 1 & \text{if } \bar{w}_1 \leq \bar{\omega},\ d = \bar{w}_1,\ \kappa = 0 \\ x_1^{*0} - 1 & \text{if } \bar{w}_1 \geq \bar{\omega},\ d = 0,\ \kappa = 1 \\ x_1^{*0} - 1 & \text{if } \bar{w}_1 + \hat{w}_1 \leq \hat{\omega},\ d = \bar{w}_1 + \hat{w}_1,\ \kappa = 0,\ \gamma \geq 1 \\ x_1^{*0} - 1 & \text{if } \bar{w}_1 + \hat{w}_1 \geq \hat{\omega},\ d = 0,\ \kappa = 1,\ \gamma \geq 1 \\ x_1^{*0} & \text{if } \bar{w}_1 \geq \bar{\omega},\ \bar{w}_1 + \hat{w}_1 \geq \hat{\omega},\ d = 0,\ \kappa = 0 \\ -\infty & \text{otherwise} \end{cases}$$

For $t \geq 2$, $\kappa = 0, \ldots, k$, $\gamma = 0, \ldots, \Gamma$, and $d = 0, \ldots, \sum_{j \in N}(\bar{w}_j + \hat{w}_j)$, the function $f_{\bar{\omega},\hat{\omega}}(t, \kappa, \gamma, d)$ can be evaluated by using the following recursive formula based on the different cases above.

$$\begin{aligned} f_{\bar{\omega},\hat{\omega}}(t, \kappa, \gamma, d) = & \\ \max\{\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa, \gamma, d), \\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa, \gamma, d - \bar{w}_t) + (x_t^0 - 1) && \text{if } \bar{w}_t \leq \bar{\omega}, \\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa - 1, \gamma, d) + (x_t^0 - 1) && \text{if } \bar{w}_t \geq \bar{\omega}, \\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa, \gamma - 1, d - \bar{w}_t) + (x_t^0 - 1) && \text{if } \bar{w}_t + \hat{w}_t \leq \hat{\omega}, \\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa - 1, \gamma - 1, d) + (x_t^0 - 1) && \text{if } \bar{w}_t + \hat{w}_t \geq \hat{\omega}, \\ & f_{\bar{\omega},\hat{\omega}}(t-1, \kappa, \gamma, d) + x_t^0 && \text{if } \bar{w}_t \geq \bar{\omega} \text{ and } \bar{w}_t + \hat{w}_t \geq \hat{\omega}\}. \end{aligned}$$

Note, that $|\bar{\Omega}| \leq n$ and $|\hat{\Omega}| \leq n$. Hence, the run-time of this dynamic program is $\mathcal{O}(n^5 \sum_{j \in N}(\bar{w}_j + \hat{w}_j))$.

7.3.3 Solving the RRKP: an approach using robustness cuts

The initial exponential formulation of the recoverable robust counterpart can be tackled by means of robustness cuts. For the special case of the Γ-RKP, Fischetti and Monaci [69] suggest such an approach. In the following, we describe how to solve the k/Γ-RRKP using robustness cuts. Although this approach can be applied to the $k, \ell/D$-RRKP as well, we focus on the more interesting case of the Γ-robust scenario set $\mathcal{S}_\Gamma$ and the k/Γ-RRKP.

Solving the k/Γ-RRKP by robustness cuts. The compact formulation (7.7) allows to solve the k/Γ-RRKP in an one-step integrated approach. An optimal solution takes the worst-case scenario and its recovery into account while maximizing the total profit. Although this formulation is compact, $\mathcal{O}(n^2)$ additional variables and $\mathcal{O}(n^2)$ additional constraints are needed to model the second-stage decision of the k/Γ-RRKP. Depending on n, the increased size of the formulation may decrease its computational tractability in practice.

To overcome this potential slow-down, an approach by so-called robustness cuts is oftentimes used in robust optimization: only the nominal first stage decision is included in the formulation and the feasibility according to the second-stage decision (and in our case, recovery) is guaranteed by additional constraints separated on-the-fly during the solution process. These constraints are called robustness cuts since they separate first-stage feasible solutions that are not feasible in the robust sense; cf. Fischetti and Monaci [69] for the case $k = 0$.

In the context of k/Γ-RRKP we consider the incomplete formulation defined by (7.7a), (7.7b), and $x_j^0 \in \{0,1\}$ for all $j \in N$. This models the first-stage decision/knapsack only. Let x^* be a solution of this system. To ensure its feasibility w.r.t. the second stage, we propose two alternative separation procedures for two types of robustness cuts:

1. *Separation of discrete scenarios.* This procedure identifies a violated scenario, constructs a corresponding discrete scenario, and adds it to the incomplete formulation as follows:

   Algorithm 7.3.1 is applied to determine a subset X of items with maximum weight and thus defining the deviating items in a worst-case scenario $S \in \mathcal{S}_\Gamma$. If this maximum weight exceeds the second-stage capacity, the system

$$\sum_{j\in N} \bar{w}_j x_j^S + \sum_{j\in X} \hat{w}_j x_j^S \leq c \tag{7.26a}$$

$$\sum_{j\in N}(x_j - x_j^S) \leq k \tag{7.26b}$$

$$x_j^S \in \{0,1\} \qquad \forall j \in N \tag{7.26c}$$

   defines a discrete scenario violated for $x = x^*$. By adding it to the incomplete formulation, the solution x^* is separated.

2. *Separation of model constraints* (7.7c)–(7.7e). This procedure identifies a $u^* \in U$ for which the system (7.7c)–(7.7e) restricted to u^* is violated: For fixed x^* and arbitrary $u^* \in U$, define

$$c(\xi,\theta) := \min\ \Gamma\xi^{u^*} + \sum_{j\in N}\theta_j^{u^*} \tag{7.27a}$$

$$\text{s.t. } \xi^{u^*} + \theta_j^{u^*} \geq \beta_j \qquad \forall j \in N \tag{7.27b}$$

$$\xi^{u^*}, \theta_j^{u^*} \geq 0 \qquad \forall j \in N \tag{7.27c}$$

   with $\beta_i := \min\{\hat{w}_j, -\bar{w}_j + u^*\}x_j^*$. Then, the system (7.7c)–(7.7e) is satisfied if and only if

$$c(\xi,\theta) \leq c + ku^* - \sum_{j\in N:\bar{w}_j<u^*} \bar{w}_j x_j^* - \sum_{j\in N:\bar{w}_j\geq u^*} u^* x_j^* \tag{7.28}$$

   holds. The separation problem can be solved by solving ILP (7.27) for each value $u \in U$. If condition (7.28) is not satisfied for a $u^* \in U$, constraints (7.7c)–(7.7e) (for u^*) are added to the incomplete formulation and separate x^*.

Moreover, there exists a combinatorial algorithm alternatively to solving (7.27) as an ILP. For an optimal solution (ξ^*, θ^*) of (7.27), it holds $\theta_j^* = \max\{\beta_j - \xi^*, 0\}$. Therefore, the objective function $c(\xi, \theta)$ is only a function of ξ. Furthermore $c(\xi)$ is convex. For $\xi > 0$, it holds $c(\xi + 1) - c(\xi) = \Gamma - |\{j \in N : \beta_j - \xi > 0\}| + |\{j \in N : \beta_j - (\xi + 1) > 0\}|$. Let the values β_j be sorted non-decreasingly and β_Γ its Γ-largest value, then $c(\xi + 1) - c(\xi) \leq 0$ for $\xi < \beta_\Gamma$ and $c(\xi + 1) - c(\xi) \geq 0$ for $\xi > \beta_\Gamma$. Hence, $c(\beta_\Gamma)$ equals the optimum of (7.27).

CHAPTER EIGHT

COMPUTATIONAL STUDIES

We will finalize our investigation of robust knapsack problems with computational studies. Since we have considered four different robustness concepts in the previous, excessive computational studies and detailed analyses for each of them go beyond the scope of this thesis. In addition, some problems such as the Γ-RKP have also been studied experimentally by other authors before. Therefore, we give a brief survey on related computational studies in the following. Then, we present exemplarily the detailed results of our computational studies for the RRKP and in particular the special cases of the $k, \ell/D$-RRKP and the k/Γ-RRKP, respectively. By definition of the k/Γ-RRKP this also includes the Γ-RKP as special case for $k = 0$.

Computational studies for robust knapsack problems include the following. We explicitly also refer to the references therein.

For the Γ-RKP, Bertsimas and Sim [34] give a very short exemplary computational study in their initial paper introducing the Γ-robustness concept itself. Klopfenstein and Nace [97] consider the polyhedral aspects of the Γ-RKP introducing (extended) cover inequalities and thus, evaluate these inequalities experimentally in a branch-and-cut approach in the context of the robust bandwidth packing problem; an application arising in telecommunications where different traffic flows share the same link. Klopfenstein and Nace consider an exact (IP based) and an heuristic approach to separate violated cuts in their study. In 2012, Monaci et al. [125] present a dynamic program solving the Γ-RKP and test it experimentally.

For the mb-RKP, only computational studies for applications from wireless telecommunications exist which include mathematical covering and assignment problems and are thus not primary related to the mb-RKP.

For the SMKP, Atamtürk and Narayanan [16] present a computational study on the effectiveness of submodular robust cover, extended cover, and lifted cover inequalities for a test set of generated instances.

For the RRKP, Büsing [46] presents the results of computational studies based on joint work with the author of this thesis and partially published before in Büsing et al. [49] (for the $k, \ell/D$-RRKP) and Büsing et al. [50] (for the k/Γ-RRKP).

Environment. All algorithms were implemented in C++ using ILOG Cplex 12.1 [84] with ILOG Concert as MIP solver and branch-and-cut framework, respectively.

The computations were carried out using a single thread of Intel Xeon W3540 CPU at 2.93 GHz and 12 GB RAM. If not stated differently, all other solver settings were left at their defaults. A time limit of 1 hour was set for solving each problem instance.

8.1 Instances

To our knowledge, no public available RRKP instances exist. Hence, we considered multi-dimensional knapsack instances from the ORLIB [21] problem library as a starting point to obtain realistic test instances for the $k, \ell/D$-RRKP. For the k/Γ-RRKP, we followed the generation rules used by Klopfenstein and Nace [97] for the Γ-RKP and based on the analysis for the KP by Pisinger [137]. We discuss our problem instances in more detail in the following.

$k, \ell/D$-RRKP. The considered $k, \ell/D$-RRKP instances are slight modifications of multi-dimensional KP instances taken from the ORLIB [21] created by Chu and Beasley. The ORLIB provides instances with three different knapsack tightness ratios (i. e., the quotient of the knapsack capacity and the the sum of all item weights), where from we selected all instances with a medium tightness ratio of 0.5 for all combinations of n and m, where $n \in \{100, 250, 500\}$ denotes the number of items and $m \in \{5, 10, 30\}$ the number of constraints. This yields 90 instances. For each $k, \ell/D$-RRKP instance, the first knapsack is treated as first stage constraint, and each remaining knapsack as individual discrete scenario. For each item, the profit of the corresponding multi-dimensional knapsack KP is scaled by 0.7 and used as first stage profit. The scenario profits are also determined by scaling these values but the scaling factor is uniformly randomly generated in $[0.2, 0.4]$.

k/Γ-RRKP. Pisinger [137] identifies general classes of computationally hard knapsack instances. In particular, the KP instances which Pisinger classifies as "weakly correlated" are hard to solve in practice. Despite their naming there is a high correlation between the item weights and the item profits, such that highly weighted items tend to be more profitable. Based on this, Klopfenstein and Nace [97] generated hard instances of the Γ-RKP when computationally investigating its solvability and the impact of extended cover inequalities in a branch-and-cut approach. We followed their generation approach. But we observed that the generated instances are too easy to solve in the recoverable robust setting. This is primary caused by the interaction of the setting of the knapsack tightness ratio and our recovery rule (the limited removal of items). Klopfenstein and Nace's setting oftentimes leads to second-stage knapsack constraints that are trivially satisfied and thus reducing the overall problem hardness in practice more or less to its first-stage knapsack. To overcome this effect, we start again at Pisinger's results and generate instances similar to Klopfenstein and Nace [97] without the described flaws as follows: For each number of items $n \in \{10, 25, 50, 100, 250, 500, 1000\}$ and knapsack tightness ratio $\tau \in \{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}\}$, we randomly generate first-stage item weights $w_i^0 \in [1, 1500]$, nominal second-stage item weights $\bar{w}_i \in [1, 1000]$, deviation second-stage item weights $\hat{w}_i \in [1, 500]$ and item profits $p_i \in [w_i^0 + 150 - 30, w_i^0 + 150 + 30]$. The first-stage and second-stage knapsack capacities are determined as $c^0 = \tau \sum_{i \in N} w_i^0$ and $c = \tau \sum_{i \in N} \bar{w}_i$, respectively. All generated values are rounded to the nearest integer.

8.2 Robustness parameters

Following the concept of recoverable robustness the robustness parameters depend on the chosen scenario set and recovery rule. For the $k, \ell/D$-RRKP, we have to specify the maximal number of removed items k, the maximal number of added items ℓ, and for each scenario $S \in \mathcal{S}_D$ the corresponding vector w^S of item weights and the knapsack capacity c^S. The item weights and the capacity are part of the instance, thus k and ℓ remain. For the k/Γ-RRKP, we have to specify the maximal number of removed items k, the robustness parameter Γ, the vector of nominal weights $\bar{w}$, the vector of deviation weights $\hat{w}$, and the scenario knapsack capacity c. Again, the item weights and the capacity is given as part of the instances, thus k and Γ remain in this case.

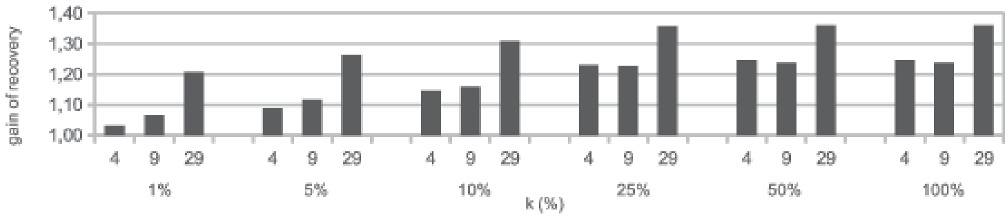
In order to analyze the impact of the recovery parameters k and ℓ (the latter only for the $k, \ell/D$-RRKP), we chose them as a fraction of the number of items, e. g., $k = 0.25$ means that 25% of the total set of items may be removed in each scenario. All combinations of $k, \ell \in \{0\%, 1\%, 5\%, 10\%, 25\%, 50\%, 100\%\}$ are tested. Furthermore, for k/Γ-RRKP, we set $\Gamma \in \{0\%, 1\%, 5\%, 10\%, 25\%\}$.

8.3 Results for the $k, \ell/D$-RRKP

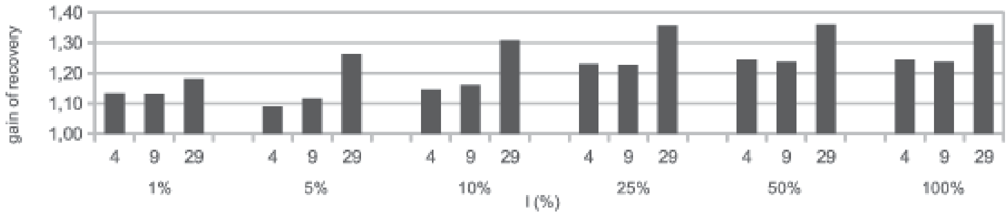
We solve the IP formulation (7.1) of the $k, \ell/D$-RRKP in the following. First, we investigate the gain of recovery, i. e., the increase in the objective profit function due to the presence of the recovery action. In addition, we report on the solvability of the $k, \ell/D$-RRKP in practice. Second, we investigate the strength of (extended) cover inequalities by evaluating the additional closure of the integrality gap at the root node.

Gain of recovery. In comparison to the classic KP, we achieve an increase in the objective function by the flexibility of the recovery action. In the following, we evaluate this gain of recovery averaged for all instances with 4, 9 and 29 discrete scenarios and w.r.t. different values of k and ℓ. The corresponding results are shown in Figures 8.1(a) and 8.1(b). Since not all instances could be solved optimally within the time limit we take the best known primal solution for our calculations and thus, the obtained gains of recovery are lower bounds on the actual gains. Later, we report on the solvability of the instances in detail.

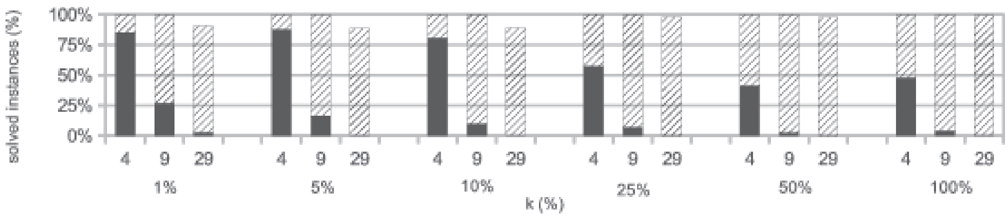
Considering the gain of recovery, we observe that it is higher the larger the number of scenarios is, e. g., it ranges from 15 % ($k = 10$ %, 4 scenarios) to 31 % ($k = 10$ %, 29 scenarios). Furthermore, the gain also increases with increasing value of k, e. g., for 4 scenarios it increases from 3 % for $k = 1$ % to 25 % for $k = 100$ %. We observe similar



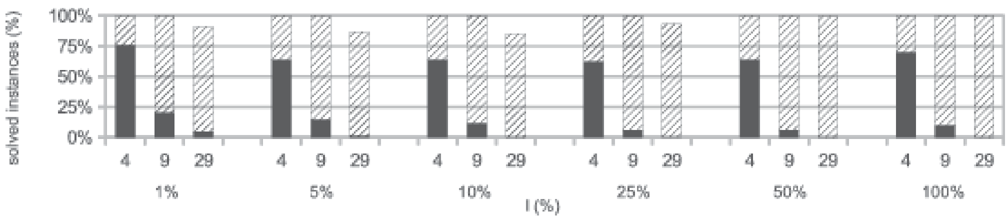
(a) average gain of recovery w.r.t. k



(b) average gain of recovery w.r.t. ℓ

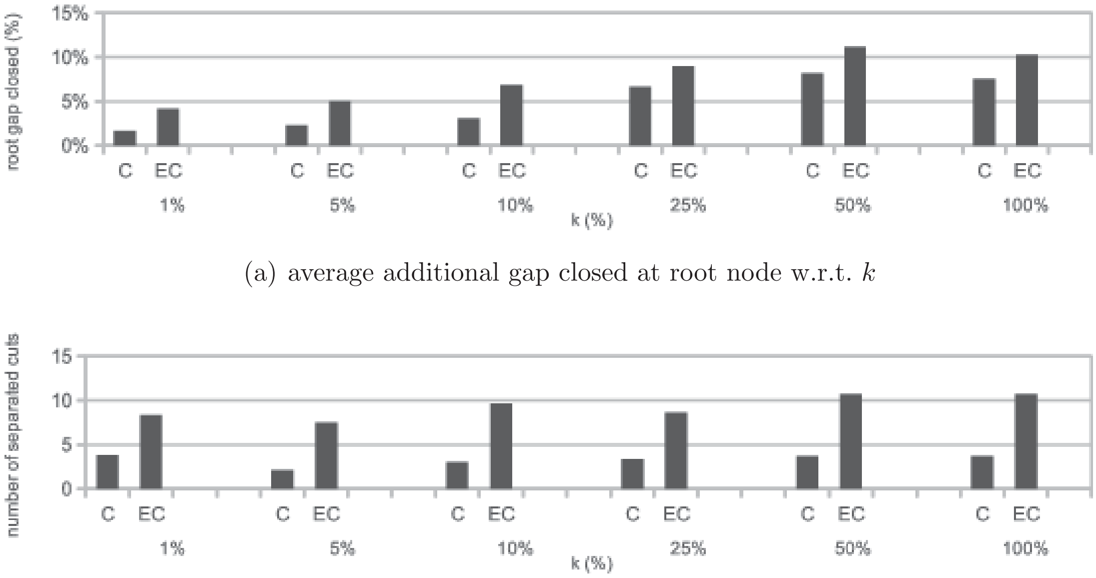


(c) average number of (almost-)optimally solved instances w.r.t. k



(d) average number of (almost-)optimally solved instances w.r.t. ℓ

Figure 8.1: Gain of recovery (normalized to $k = \ell = 0$), and number of optimally solved instances (solid blue) and instances solved with gap less than 0.5 % to optimum (shaded blue), respectively. All results are averaged and shown for instances with 4, 9 and 29 scenarios, and $k = 1\,\%, 5\,\%, 10\,\%, 25\,\%, 50\,\%, 100\,\%$ and $\ell = 1\,\%, 5\,\%, 10\,\%, 25\,\%, 50\,\%, 100\,\%$, respectively.



(a) average additional gap closed at root node w.r.t. k

(b) average number of separated cuts at root node w.r.t. k

Figure 8.2: Additional integrality gap closed at root node by separation of covers (C) and extended covers (EC). In addition the number of separated cuts is presented. All results are averaged and shown for $k = 1\,\%, 5\,\%, 10\,\%, 25\,\%, 50\,\%, 100\,\%$.

trends for the results w.r.t. the value of ℓ. The effects can be explained as follows: larger numbers of scenarios enforce more conservative solutions in the deterministic non-robust case ($k = \ell = 0$). Thus, it allows a higher gain if this restriction is relaxed by the recovery action and in particular for higher values of k or ℓ. The resulting solutions are more profitable.

Next, we address the practical computability of the $k, \ell/D$-RRKP. Figures 8.1(c) and 8.1(d) visualize the averaged relative number of optimally solved instances with 4, 9 and 29 discrete scenarios for selected values of k and ℓ as solid blue bars, respectively. For larger number of scenarios, the instances could often not be solved to optimality. Nevertheless almost-optimal solutions with remaining optimality gaps at most 0.5 % to optimality could be in 98 % of all cases. Therefore, we report on the averaged relative number of these instances as well. They are shown as shaded blue bars stacked on the solid blue bars in Figures 8.1(c) and 8.1(d). First, we observe that 98 % of all instances can be solved to 0.5 % of optimality within the time limit. Second, the number of actually optimally solved instances decreases with increasing number of scenarios whereas the instances with 29 scenarios could not be solved to optimally in 97 % of all cases. Moreover, the solution times are on average 50 to 100 times, 16 to 20 times, and about 12 times larger than the corresponding non-robust KP ($k = \ell = 0$) instances with 4, 9, and 29 scenarios, respectively.

Impact of recoverable robust extended covers. Next, we consider the cover and extended cover inequalities for the $k, \ell/D$-RRKP. In particular, the strength of these inequalities is of interest to us. Therefore, we implemented the separation problem in its ILP formulation (7.20) to separate violated extended cover inequalities exactly. Whenever our separator is called the first stage knapsack is checked. If no violated extended cover is found, all scenarios are tested beginning with the last scenario which provided a violated cut, until a violation is determined. This inequality is then added to the LP and the separation round is aborted. Only the root node of the $k, \ell/D$-RRKP is solved in this study. We consider three algorithms which differ by the IP solved (i.e., whether the canonical extension is integrated or not) and how the solution is strengthened in a post processing step (i.e., items with $x_j^* = 0$ are not necessarily selected in the separation ILP): Algorithm 0 solves the IP (7.1) with Cplex at its default settings, Algorithm C additionally separates cover inequalities with canonical greedy extension, and Algorithm EC additionally separates extended cover inequalities with canonical greedy extension. All results are evaluated normalized to the results of Algorithm 0 which serves as a benchmark (and thus allows to determine the added value).

Figure 8.2(a) shows the averaged additional integrality gap closed (w.r.t. Algorithm 0) at the end of the root node for Algorithms C and EC, and $k = 1\,\%, 5\,\%, 10\,\%, 25\,\%, 50\,\%, 100\,\%$. The results are in line with those for the classical KP. The separation of extended cover inequalities always closes the integrality gap more than covers only. On average 5 % to 10 % can be closed by extended cover inequalities. The averaged numbers of separated cuts are shown for both algorithms and selected values of k in Figure 8.2(b). Here, we observe that Algorithm EC separates at least twice the number of cuts on average. In general, only a small number of cuts is separated (on average less than 5 cuts by Algorithm C, and about 10 cuts by Algorithm EC). Hence, it is remarkable that such a small number of cuts achieve additional gap closures of the observed values.

Conclusions. We have carried out extensive tests evaluating the gain of recovery observing an average profit increase of up to 26 % ($k = 50\,\%$, 29 scenarios). Although many instances could not be solved to optimality, solutions with a guaranteed relative distance to the optimum of at most 0.5 % have been achieved in 98 % of all cases. Furthermore, we have investigated the effectiveness of recoverable robust (extended) cover inequalities showing that the integrality gap at the root node could be closed by up to 10 % on average by these cuts in our experiments. These are promising results on the quality of the cuts. Hence, fast heuristic separation approaches should be the focus of future investigations.

8.4 Results for the k/Γ-RRKP

We consider different algorithmic approaches to solve the k/Γ-RRKP, cf. Section 7.3 for details. The following list gives an overview.

CmpIP: Algorithm CmpIP solves the compact IP (7.7).

SepU: Algorithm SepU solves (7.7a), (7.7b), (7.7e), and separates model constraints (7.7c) and (7.7d) on-the-fly by the combinatorial algorithm described in Section 7.3.

SepEC: Algorithm SepEC solves (7.7a), (7.7b), (7.7e), and separates extended cover inequalities on-the-fly by solving IP (7.23).

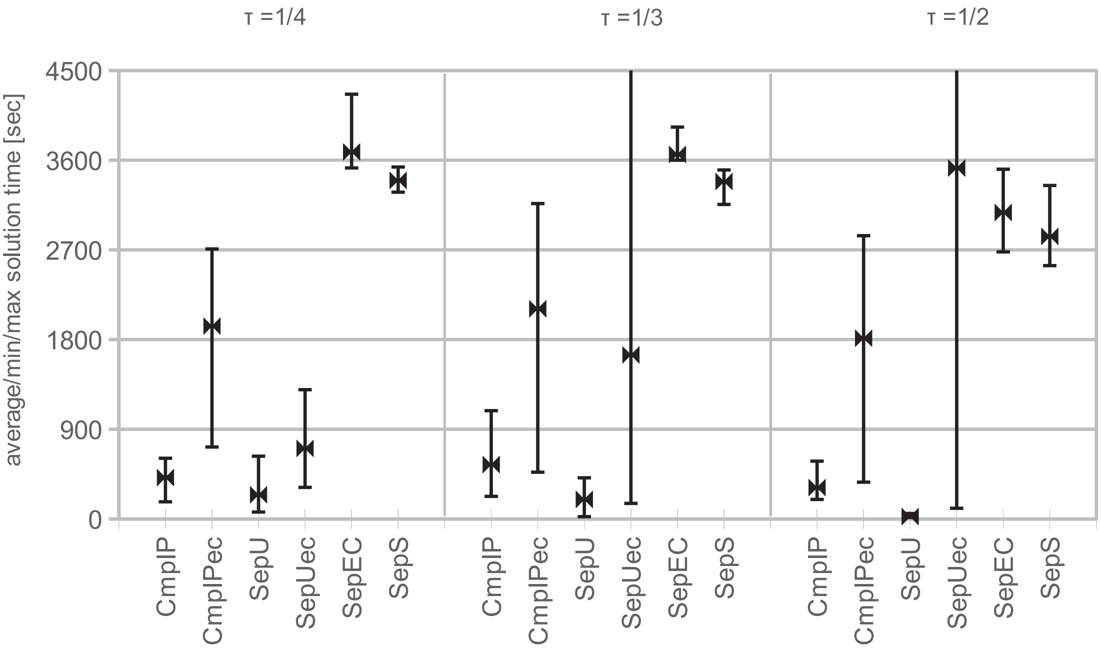
SepS: Algorithm SepS solves (7.7a), (7.7b), (7.7e), and separates scenarios (7.26) on-the-fly by the combinatorial algorithm described in Section 7.3.

CmpIPec/SepUec: For CmpIP and SepU we also consider variants which include the additional separation of violated extended cover inequalities in a cut-and-branch approach by solving IP (7.23). We denote them by the suffix -ec.

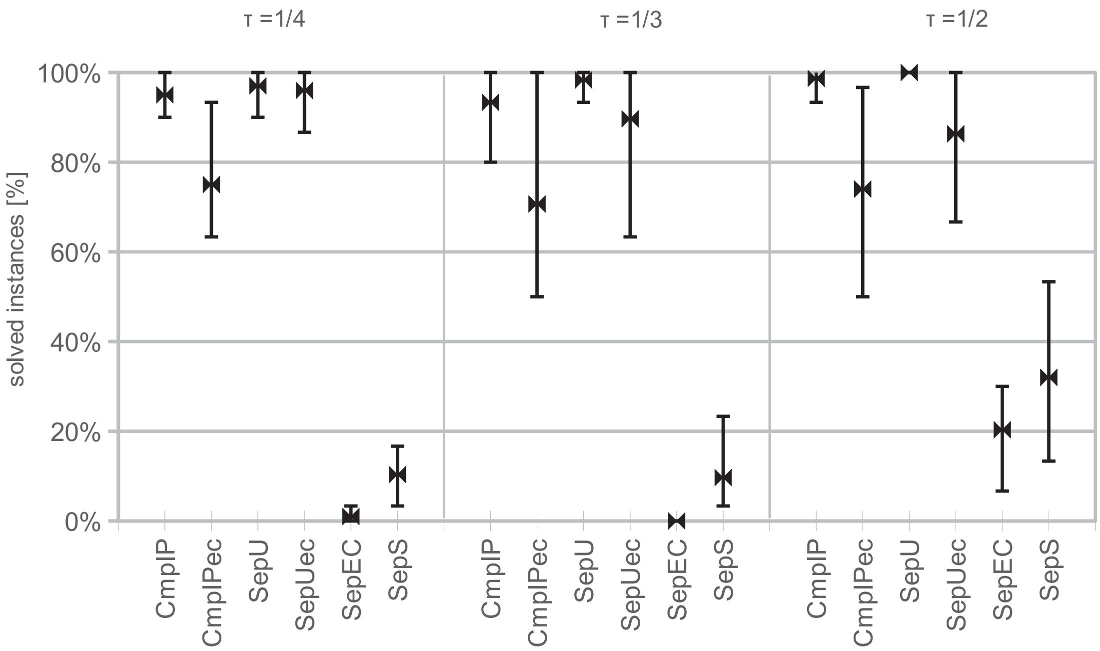
First observations. The k/Γ-RRKP generalizes the classic KP (special case: $\Gamma = k = 0$), the RRKP without data uncertainty (special case: $\Gamma = 0$), as well as the Γ-RKP (special case: $k = 0$). Notice that these special cases are not representative for the k/Γ-RRKP in general and other more specific algorithms to solve these may exist or our algorithms may perform much better. For example, our compact model (7.7) reduces to the well-known multi-dimensional KP for $\Gamma = 0$. In this case, we observed a drastic decrease in the solution times of our algorithm CmpIP compared to the robust case where $\Gamma > 0$.

In an initial study, we tested the computational tractability of the generated instances (for $\Gamma = k$ and $\tau = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}$) within the time limit depending on the number of items. We observe the following: instances with 100 items are the most interesting for our study as on the one hand smaller instances could be solved by all algorithms oftentimes with no effort within a few seconds and on the other hand larger instances were not solvable by most algorithms within the time limit of one hour. Hence, in the following we focus on instances with 100 items to compare the proposed algorithms within our experimental set-up.

Analysis per tightness ratio. First, we investigate the algorithmic performances solving all instances with 100 items and knapsack tightness ratios of $\frac{1}{4}$, $\frac{1}{3}$, and $\frac{1}{2}$ for each algorithm. Therefore, we evaluate the average, minimal and maximal solution times for each algorithm and tightness ratio. The results are shown in Figure 8.3. Notice, some instances did not terminate within the time limit (due to the technical implementation of the time limit) resulting in some higher solution times. In Figure 8.3(a), we observe that Algorithm SepU is the fastest on average. But also the range of its observed minimal and maximal solution time is good compared to the other algorithms. Solving the compact ILP, CmpIP is the second-fasted algorithm. The algorithms CmpIPec and SepUec are slower than the corresponding algorithms without the additional separation of extended cover inequalities. In fact, it turns out that the exact separation of these inequalities by solving an ILP is also quite slow in practice. Among all considered algorithms SepEC is the slowest; reaching the time limit in 60 % of all cases (80 %, 100 % and 0 % of all cases for $\tau = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}$, respectively) whereas the algorithms CmpIP and SepU are able to



(a) average/minimal/maximal solution times



(b) percentage of optimally solved instances

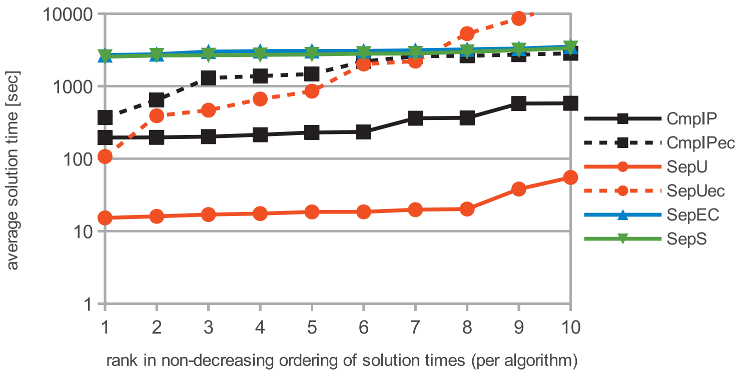
Figure 8.3: Comparison of algorithmic performance according to average/minimal/maximal solution times and the percentage of optimally solved instances. The results are shown for tightness ratios $\tau = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}$.

solve most instances in the time limit. The second-worst algorithm is SepS. For $\tau < \frac{1}{2}$ it is the second slowest, only slightly better for $\tau = \frac{1}{2}$. Besides it solves less than half the instances to optimality, for $\tau = \frac{1}{2}$ even less than 25 %.

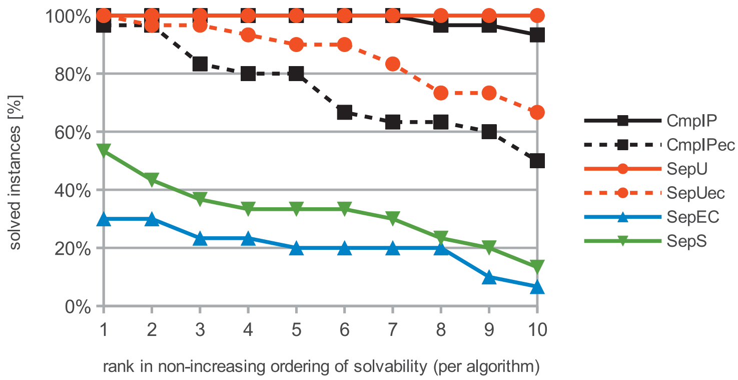
Comparing the results of different tightness ratios for a fixed algorithm, we do not observe any clear trend. One difference is remarkable: the difference between maximal and minimal solution time increases significantly with higher tightnesses for SepUec while SepS remains stable. An explanation of this effect is unknown to us. In the following we focus on the instances with tightness ratio 0.5.

Analysis per instance. Second, we compare the performance of the algorithms per instance. Therefore we evaluate the average solution times, the percentage of optimally solved instances, and the maximal remaining optimality gaps. For each instance and algorithm, The results are averaged for all considered values of the parameters k and Γ. Figure 8.4 visualizes the results of this analysis. To obtain meaningful results in a normalized way, the x-axes do not state the individual instances but an ordering of the evaluated values. For example, in Figure 8.4(a) the solution times of each algorithm are sorted non-decreasingly and thus 1 on the x-axis relates to the smallest value, 2 to the second-smallest value and so on, e.g., for SepU the smallest value is 106 sec. and the second-smallest is 390 sec. Notice that Figures 8.4(a) and 8.4(c) have logarithmically scaled axes. Again, Algorithm SepU performs best followed by Algorithm CmpIP which is already about 10 times slower. The slowest is algorithm SepEC which is about 100 times slower than SepU. Algorithm SepS is similarly slow. The variants CmpIPec and SepUec are also about 10 times slower than CmpIP and SepU, respectively. Both algorithms CmpIP and SepU solve 85 % resp. 100 % of all instances to optimality. SepEC solves the least instances to optimality; at least 70 % remain unsolved. Algorithm SepS achieves only slightly better results. Considering the maximal optimality gaps, we observe that Algorithm SepEC performs worst with optimality gaps between 29 % and 57 % when reaching the time limit. Although SepS yields results similar to SepEC w.r.t. solution times and solvability, it achieves optimality gaps ten times better than the ones by SepEC. In contrast, the gaps of the unsolved instances by CmpIPec and SepUec are less than 2.5 %. All other optimality gaps — in particular of SepU— are zero or negligible.

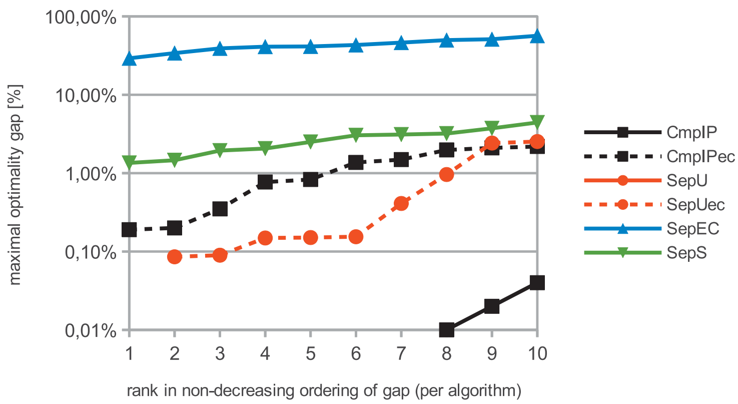
Analysis per $(\boldsymbol{k}, \boldsymbol{\Gamma})$. Third, we report on the solution times and percentage of solved instances dependent on the values of k and Γ. Figures 8.5(a) and 8.5(b) show the results of this evaluation. First we observe that the results for $k = 0$ are differently shaped than for $k > 0$. For increasing Γ, the solution times trends remain similar, as does the number of optimally solved instances for all algorithms except for SepS. For algorithm SepS, less instances are solved for increasing values of Γ, e.g., 70 % for $\Gamma = 5$ down to 0 % for $\Gamma = 25$. Second, for $k > 0$ we observe the following trends: if Γ increases, then the solution times increase as well and thus the ratio of optimally solved instances decreases (by trend). With increasing value of k the solution time tends to decrease whereas there



(a) average solution times

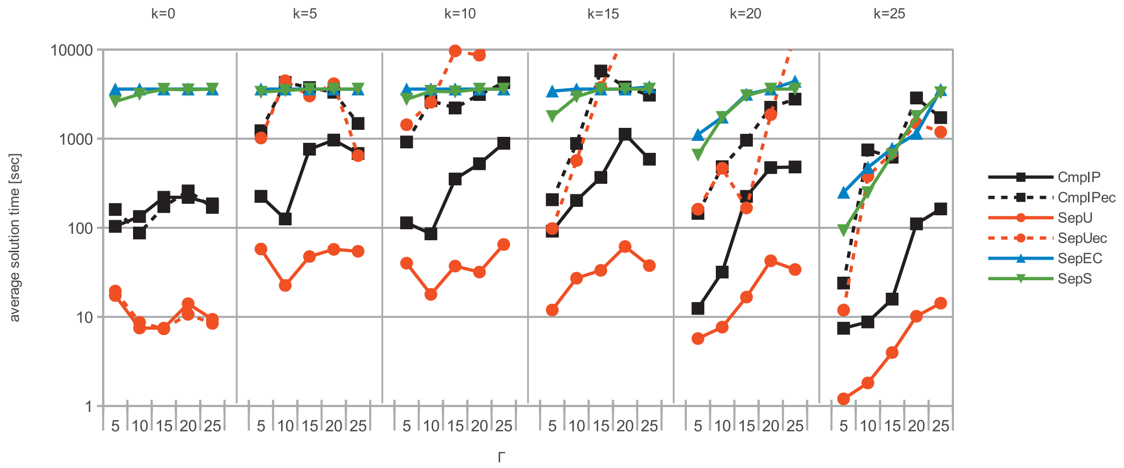


(b) percentage of optimally solved instances

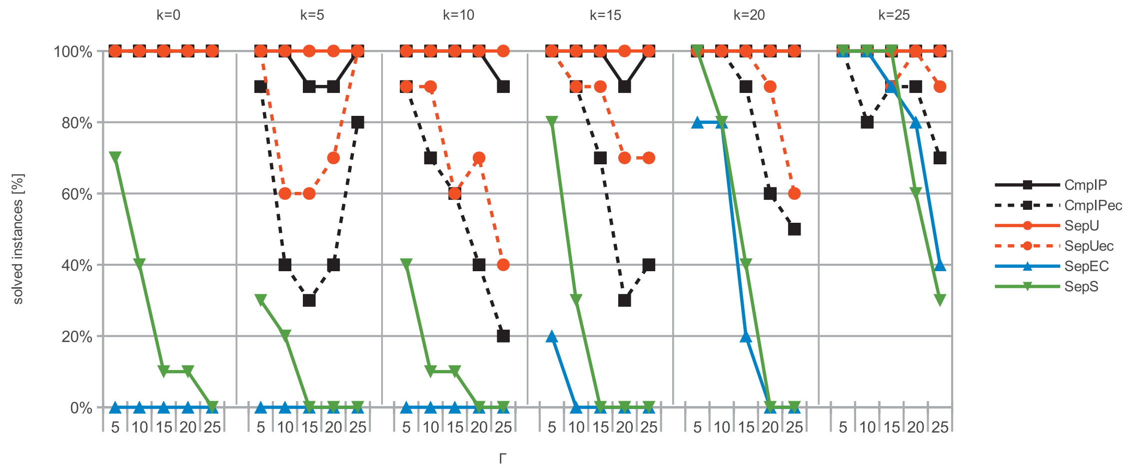


(c) maximal optimality gap

Figure 8.4: Comparison of algorithmic performance according to average solution times, the percentage of optimally solved instances, and maximal optimality gaps. All results are computed for ten instances with 100 items each. For each algorithm the considered values (time/#solved/gap%) are sorted and shown w.r.t. this ordering.



(a) average solution times

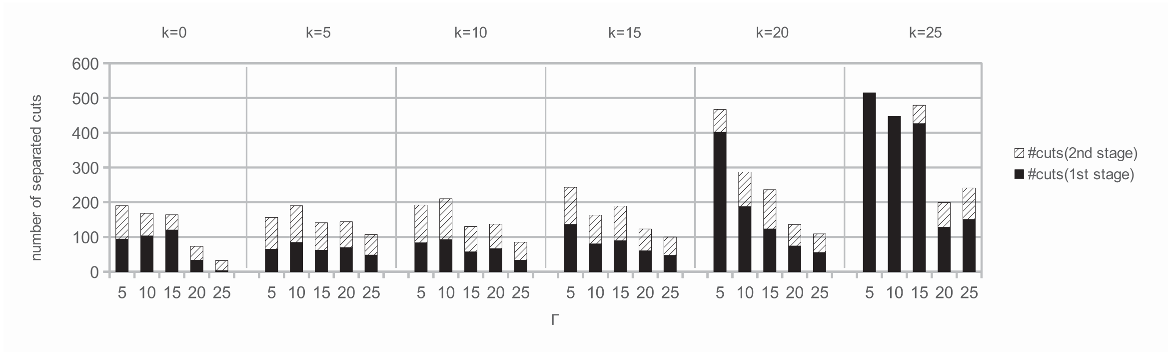


(b) percentage of optimally solved instances

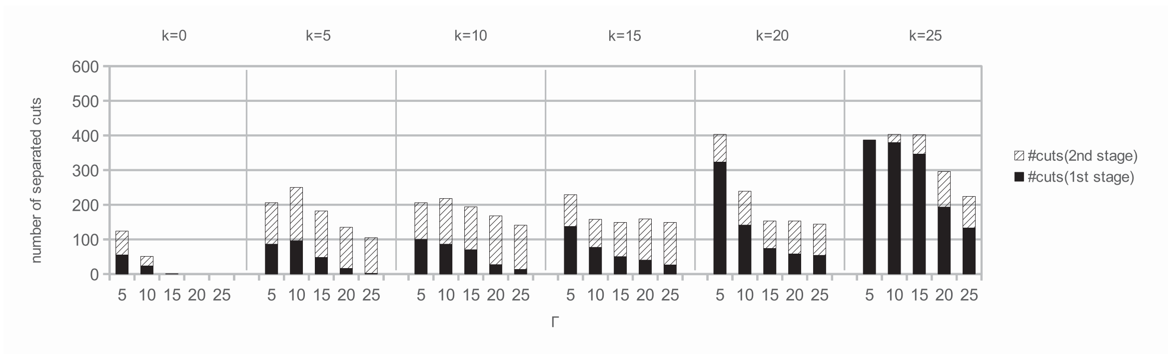
Figure 8.5: Comparison of algorithmic performance according to average solution times and the percentage of optimally solved instances. All results are averaged for ten instances with 100 items each. They are shown for $k = 0, 5, 10, 15, 20, 25$ and $\Gamma = 5, 10, 15, 20, 25$.

is no clear trend for the number of optimally solved instances. Only for SepEC and SepS there is an increase in the ratio of optimally solved instances for higher values of k.

Additionally, we report on the average number of separated extended cover inequalities for algorithms CmpIPec and SepUec. Remember that a recoverable robust cover is defined as a cover for the first-stage knapsack or the second-stage / scenario knapsack. Thus, we distinguish between first- and second-stage covers in our evaluation. Figure 8.6 shows the average numbers of first-stage and second-stage extended cover inequalities



(a) Algorithm CmpIPec



(b) Algorithm SepUec

Figure 8.6: The total numbers of separated extended covers inequalities for algorithm CmpIPec and SepUec. All results are averaged for ten instances with 100 items each, $k = 0, 5, 10, 15, 20, 25$ and $\Gamma = 5, 10, 15, 20, 25$.

(cuts). We observe no clear trend depending on k or Γ. It seems that more cuts have been separated for higher values of k and lower values of Γ. Further, we notice that — in line with our separation strategy which first separates first-stage cuts and second-stage cuts only if no first-stage cuts have been found — more first-stage covers have been found than second-stage covers but still second-stage covers have been separated indicating that in these case no further violated first-stage covers have existed. Furthermore, comparing the results for CmpIPec and SepUec more second-stage cuts have been separated when using SepUec than when using CmpIPec. This indicates that extended cover inequalities may contribute more to the feasibility of the problem in an algorithm using lazy constraints for the second stage than in an algorithm solving a complete IP.

Conclusions. We have carried out extensive computational studies comparing k/Γ-RRKP instances with seven different numbers of items and three different tightness ratios for 36 different values of k and Γ each. We have evaluated the performance w.r.t. the solution times, the number of optimally solved instances, and the remaining optimality gaps. All results strongly indicate that our new approach leading to the combinatorial separation of violated model constraints (7.7c)–(7.7d) is best-suited to solve the k/Γ-RRKP: the corresponding algorithm SepU clearly outperforms the others. The compact ILP formulation (7.7), corresponding to Algorithm CmpIP, is second-best and the Algorithm SepEC is the worst (at least 10 times slower). Also algorithm SepS, the approach to separate scenarios which is oftentimes suggested in literature, performs bad in our studies and should not be used to solve RRKP in practice. Although violated recoverable robust extended cover inequalities have been separated, their effect on the overall solving process is disadvantageous because the proposed exact separation by IP is too slow. Here an alternative heuristic separation should be considered.

PART THREE

ROBUST NETWORK DESIGN PROBLEMS

CHAPTER NINE

THE Γ-ROBUST NETWORK DESIGN PROBLEM

In this chapter, we consider a robust version of the classic network design problem with zero routing costs and integral link capacities. Sometimes this classic network design problem is more precisely called the network loading problem. Nevertheless, we use the more common term network design problem. The robustness concept which we apply to take demand uncertainties into account is Γ-robustness. We call the resulting problem the Γ-robust network design problem (Γ-RNDP). Due to the popularity of Γ-robustness, this problem, its formulations, and related structures are oftentimes just called *robust* neglecting alternative robustness concepts. For example, network design under data uncertainty has been consider in telecommunications using multi-period or multi-hour approaches by Lardeux et al. [109] and Terblanche et al. [152], where an optimal solution is determined for each data realization of a given finite set of possible realizations. Ben-Ameur and Kerivin [24] or Duffield et al. [63] follow approaches closer to robust optimization by introducing a polyhedral uncertainty sets and determining a single solution feasible for all possible realizations of the uncertainty set. A prominent special case of a polyhedral uncertainty set is the so-called *hose model*, where the sum of out-going and in-going traffic is given for each traffic node; cf. Altin et al. [11], Chekuri et al. [51] and Altin et al. [12]. The Γ-robustness concept, also a special case of a polyhedral uncertainty set, has been applied to the Γ-RNDP by Altin et al. [11] in the context of the planning of virtual private networks. Further investigations by others include a simplification of the Γ-RNDP for $\Gamma = 1$ studied by Belotti et al. [22], the Γ-RNDP with alternative routing rules studied by Ouorou and Vial [132] and Poss and Raack [139], and the integration of technological decisions studied by Belotti et al. [23]. Note that although several robust approaches exist, telecommunication networks are typically designed using historical traffic measurements or population statistics as in Bley et al. [42], Dwivedi and Wagner [65], Uhlig et al. [153] and Zhang [162] in practice. Thus deterministic NDP are obtained which neglect the data uncertainty of the underlying real-world applications and thus have to be overestimations to ensure feasibility. Clearly, these are cost-inefficient approaches.

In contrast, we consider the Γ-RNDP in the following. Parts of this chapter are based on joint work and have been published by the author of this thesis and the corresponding co-authors or have been submitted for publication. These works are Koster and Kutschka [101], Koster et al. [103, 105], Raack [140] and Koster et al. [106] on different formulation, basic polyhedral properties and cutset-based inequalities, and Claßen et al. [55] on the capacity formulation and Γ-robust metric inequalities.

We consider the following robust network design problem. We are given an undirected connected graph $G = (V, E)$ representing a potential network topology. On each of the links $e \in E$ capacity can be installed in integral units and costs κ_e per unit. A set of commodities K represents potential traffic demands. More precisely, a commodity $k \in K$ corresponds to node pair (s^k, t^k) and a demand $d^k \geq 0$ for traffic from $s^k \in V$ to $t^k \in V$. The actual demand values are considered to be uncertain. We model the uncertain demand using the Γ-robustness concept, i. e., the demand vector $d \in \mathcal{U}^\Gamma$ corresponding to the demand values $d^k, k \in K$ w. r. t. given nominal and deviation demand values: $\bar{d} \in \mathbb{R}^{|K|}_{\geq 0}$ and $\hat{d} \in \mathbb{R}^{|K|}_{\geq 0}$, respectively. Furthermore, let $\Gamma \in \{0, 1, \dots, |K|\}$ be the associated robustness parameter determining the maximal number of maximally deviating demands at the same time. Notice that we could also consider fractional values $\Gamma \in [0, |K|] \setminus \mathbb{N}$. This would correspond to at most $\lfloor \Gamma \rfloor$ simultaneously maximally deviating demands and an additional demand deviating to $(\Gamma - \lfloor \Gamma \rfloor)$-times its maximal deviation. Although we focus on integer values for Γ in this thesis, the results can be generalized following this previous interpretation.

The traffic for commodity k is realized by a multi-path flow between s^k and t^k which is not restricted to one single path and hence also called *splittable*. Of course, the actual multi-commodity flow depends on the realization of the demand $d \in \mathcal{U}^\Gamma$. In this context, the literature roughly distinguishes two main routing principles. We either choose an arbitrary flow for every realization of the demand in $\mathcal{U}^\Gamma$, which is known as *dynamic routing* or we fix a *routing template* for every commodity, that is, every realization of the uncertain demand has to use the same set of paths between $s^k \in V$ and $t^k \in V$ with the same fraction of the total s^k-t^k-flow assigned to these paths. This latter principle is known as *oblivious routing* (or *static routing*) and is considered in this paper. We refer the reader to Mattia [121] and Poss and Raack [139] for solution approaches considering dynamic routing.

Definition 9.1 (Γ-Robust Network Design Problem)**.** Given a potential network topology, a Γ-robust uncertainty set $\mathcal{U}^\Gamma$ of the demand, and installation costs as described above, the *Γ-robust network design problem (Γ-RNDP)* using oblivious/static, splittable routing is to find a minimum-cost installation of integral capacities and a routing template (i. e., a partition of the total flow of a commodity into paths) for every commodity such that actual flow does not exceed the link capacities independent of the realization of demands in $\mathcal{U}^\Gamma$.

This chapter is structured as follows: Next, we present formulations of the Γ-RNDP. Then it follows a polyhedral investigation, starting with the definition of the related polyhedra, their basic properties, and followed by classes of valid inequalities based on three approaches: the study of the Γ-robust cutset polyhedron, the Γ-robust single arc design polyhedron, and the capacity formulation of the Γ-RNDP. We conclude with reports on representative computational studies using real-life traffic measurements of telecommunication networks as uncertain traffic/demand data.

Next, we formulate the Γ-RNDP as ILP. As for the classic network design problem, NDP, there also exist link-flow and path-flow formulations for the Γ-RNDP.

9.1 Formulations

In this section, we present ILP formulations of the Γ-RNDP. Therefore, let us recall the ILP formulation (1.19) of the classic NDP.

Given link-based flow variables f_{ij}^k denoting the fraction of demand k routed along arc ij and link capacity installment variables x_e determining the number of capacity modules to be installed on edge e, the link-flow formulation of the NDP is

$$\min \quad \sum_{e \in E} \kappa_e x_e \tag{1.19a}$$

$$\text{s.t.} \quad \sum_{j\,:\,ij \in E} (f_{ij}^k - f_{ji}^k) = \begin{cases} 1 & i = s^k \\ -1 & i = t^k \\ 0 & \text{else} \end{cases} \qquad \forall i \in V,\, k \in K \tag{1.19b}$$

$$\sum_{k \in K} d^k f_e^k \ \le x_e, \qquad \forall e \in E \tag{1.19c}$$

$$f, x \ge 0 \tag{1.19d}$$

$$x \in \mathbb{Z}^{|E|} \tag{1.19e}$$

with $f_e^k := f_{ij}^k + f_{ji}^k$; flow conservation (1.19b), link capacity (1.19c), nonnegativity(1.19d) and integrality (1.19e) constraints; cf. Section 1.3.

By introducing path-based flow variables f_p^k denoting the fraction of demand k routing along path $p \in P^k$ (where P^k is the set of all possible s^k-t^k-paths), we obtain the path-flow ILP formulation of the NDP:

$$\min \quad \sum_{e \in E} \kappa_e x_e \tag{1.20a}$$

$$\text{s.t.} \quad \sum_{p \in P^k} f_p^k \ge 1 \qquad \forall k \in K \tag{1.20b}$$

$$\sum_{k \in K} d^k \left(\sum_{p \in P^k\,:\,e \in p} f_p^k \right) \le x_e, \qquad \forall e \in E \tag{1.20c}$$

$$f, x \ge 0 \tag{1.20d}$$

$$x \in \mathbb{Z}^{|E|} \tag{1.20e}$$

with path covering constraints (1.20b), link capacity (1.20c), nonnegativity(1.20d) and integrality (1.20e) constraints; cf. Section 1.3.

Further, in Section 1.3 we also describe how a solution of the link-flow formulation can be transformed into a solution of the path-flow formulation and vice-versa showing the equivalence of both formulations.

Although the Γ-robustness concept assumes a symmetric distribution around a nominal value, we can focus only on the positive deviations w. l. o. g. because the worst-case realization does not include negative deviations: for each realizations with a negative demand deviation for commodity k a dominating realization exists without a deviation or a positive demand deviation for k using the following definition of dominance. A realization dominates another if its contribution to the left-hand side of the link capacity constraint is greater than or equal to the corresponding contribution of the other realization.

Routing templates and realized flows under demand uncertainty. Given a capacity allocation $x \in \mathbb{R}^{|E|}_{\geq 0}$, and routing templates f^k for all $k \in K$, we say that (x, f) *supports* $d \in \mathcal{U}^\Gamma$ in case (1.19c) is satisfied for d. Fixing $d \in \mathcal{U}^\Gamma$, the realized flow $f^k_e(d)$ for commodity k on edge e amounts to

$$f^k_e(d) := d^k f^k_e. \tag{9.3}$$

This means that we allow the flow to change with the demand fluctuations d but we restrict the flow dynamics to the linear functions given by (9.3). The realized flow $f^k_p(d)$ on a path in the path-flow formulation can be determined analogously. Note hat (9.3) is a special case of so-called *affine recourse* introduced by Ben-Tal et al. [30] in the context of adjustable robust solutions of LPs with uncertain data. Ouorou and Vial [132] apply affine recourse to network design introducing a new affine routing scheme. Poss and Raack [139] provide a conceptual discussion of the three routing schemes: oblivious, affine, and dynamic.

Exponential flow formulations of the Γ-RNDP. The Γ-robust counterparts of the classic link-flow (1.19) and path-flow formulation (1.20) are

$$\begin{aligned} &(1.19a), (1.19b) - (1.19e) \\ &\sum_{k \in K} d^k f^k_e \leq x_e \qquad \forall e \in E, d \in \mathcal{U}^\Gamma \end{aligned} \tag{9.4a}$$

and

$$\begin{aligned} &(1.20a), (1.20b) - (1.20e) \\ &\sum_{k \in K} d^k \left(\sum_{p \in P^k : e \in p} f^k_p \right) \leq x_e \qquad \forall e \in E, d \in \mathcal{U}^\Gamma, \end{aligned} \tag{9.5a}$$

respectively. In both formulations infinitely many realizations $d \in \mathcal{U}^\Gamma$ are considered in the corresponding link capacity constraints. In fact, not infinitely but exponentially many extremal realizations are non-dominated resp. unobtainable as convex combination of other realizations and have to be taken into account to determine the worst-case value in $\mathcal{U}^\Gamma$; cf. Section 3.1.2. Hence, both formulations are of exponential size.

Compact flow formulations of the Γ-RNDP. Following the explanations in Section 3.1.2, formulations (9.4) and (9.5) can be linearized and reformulated in a compact way by exploiting strong duality of linear programming theory. Thus, the *compact link-flow formulation* of the Γ-RNDP reads

$$
\begin{aligned}
&(1.19\text{a}), (1.19\text{b}), (1.19\text{d}), (1.19\text{e}) && \\
&\Gamma \pi_e + \sum_{k \in K} \bar{d}^k f_e^k + \sum_{k \in K} \rho_e^k \le x_e, && \forall e \in E \quad (9.6\text{a}) \\
&-\pi_e + \hat{d}^k f_e^k - \rho_e^k \le 0 && \forall e \in E, k \in K \quad (9.6\text{b}) \\
&\rho, \pi \ge 0 && (9.6\text{c})
\end{aligned}
$$

with nonnegative dual variables ρ and π and the dual constraints (9.6b). Similarly, the *compact path-flow formulation* of the Γ-RNDP is given by

$$
\begin{aligned}
&(1.20\text{a}), (1.20\text{b}), (1.20\text{d}), (1.20\text{e}) && \\
&\Gamma \pi_e + \sum_{k \in K} \bar{d}^k \left(\sum_{p \in P^k : e \in p} f_p^k \right) + \sum_{k \in K} \rho_e^k \le x_e, && \forall e \in E \quad (9.7\text{a}) \\
&-\pi_e + \hat{d}^k \sum_{p \in P^k : e \in p} f_p^k - \rho_e^k \le 0 && \forall e \in E, k \in K \quad (9.7\text{b}) \\
&\rho, \pi \ge 0 && (9.7\text{c})
\end{aligned}
$$

with nonnegative dual variables ρ and π and the dual constraints (9.7b). Notice that we have assumed a common source of uncertainty affecting all paths of a commodity simultaneously. Hence, we can model the uncertainty with a single max-term in the first step, dualize and obtain (9.7a)–(9.7c) with a single robustness parameter Γ and path-independent dual variables; cf. Bertsimas and Sim [33] and Bertsimas and Sim [34] for a discussion of this assumption.

Capacity formulation of the Γ-RNDP. Given the link capacities $x \in \mathbb{Z}_{\ge 0}^{|E|}$, we can characterize the existence of a feasible flow satisfying the constraints (1.20b), (1.20d), and (9.7a)–(9.7c) by applying Farkas' lemma. In fact, this characterization holds even for fractional capacities:

Lemma 9.2. *Given* $\tilde{x} \in \mathbb{R}_{\ge 0}^{|E|}$*, there exists a flow satisfying* (1.20b)*,* (1.20d)*, and* (9.7a)–(9.7c) *if and only if for all lengths functions* $\ell : E \to \mathbb{R}_{\ge 0}$

$$\sum_{e \in E} \tilde{x}_e \ell(e) \ge b_\ell \quad (9.8)$$

holds, where b_ℓ is defined by the following LP.

$$b_\ell := \quad \max \sum_{k \in K} b^k \tag{9.9a}$$

$$s.t.\ b^k - \sum_{e \in p} \hat{d}^k m^k(e) \leq \bar{d}^k \ell(p) \qquad \forall k \in K,\, p \in P^k \tag{9.9b}$$

$$\sum_{k \in K} m^k(e) \leq \Gamma \ell(e) \qquad \forall e \in E \tag{9.9c}$$

$$m^k(e) \leq \ell(e) \qquad \forall k \in K,\, e \in E \tag{9.9d}$$

$$b^k,\, m^k(e) \geq 0 \qquad \forall k \in K,\, e \in E \tag{9.9e}$$

We call (9.8) *the* Γ-robust length inequality (RLI)*, and if ℓ is metric,* (9.8) *is called a* Γ-robust metric inequality (RMI)*.*

Proof. By setting $x = \tilde{x}$ in (9.7), the problem is turned into an LP. We introduce dual variables $\{b^k\}_{k \in K}$, $\{\ell(e)\}_{e \in E}$, $\{m^k(e)\}_{e \in E, k \in K}$ for constraints (1.20b), (9.7a) and (9.7b), respectively, and apply Farkas' lemma yielding conditions (9.9b)–(9.9e) and $\ell(e) \geq 0$. Thus, a solution to (9.7) exists if and only if

$$\sum_{e \in E} \tilde{x}_e \ell(e) \geq \sum_{k \in K} b^k \tag{9.10}$$

is valid for all $b^k,\ \ell(e),\ m^k(e) \geq 0$ satisfying (9.9b)–(9.9d). Therefore, for a given length function $\{\tilde{\ell}(e)\}_{e \in E}$ inequality (9.10) is valid for the convex hull of all feasible solutions of the linear relaxation of (9.7) if constraints (9.9b)–(9.9e) are fulfilled. To determine the strongest valid inequality for $\tilde{\ell}$, we maximize the sum over b^k for $k \in K$. This can be formulated as LP (9.9), where $\ell = \tilde{\ell}$, completing the proof. □

Although we use the path-flow formulation of the Γ-RNDP in Lemma 9.2 and its proof, the result holds for the link-flow formulation as well due to the equivalence of both formulations.

The LP (9.9) to compute the right-hand side of a RLI is comparable to the problem ($\text{subsep}_{\max}$) of Mattia [121]. However, we specified the general constraint $Ad \leq b$ according to the compact formulation of the Γ-RNDP. Lemma 9.2 generalizes the "Japanese Theorem" 1.21 of the classic NDP to the Γ-robust setting. Compared to this theorem, b^k is the length of the shortest path with respect to ℓ adjusted by values $m^k(e)$ for which additional constraints (9.9c) and (9.9d) hold. The previous lemma gives rise to the *capacity formulation* of the Γ-RNDP:

(1.20a)

$$\sum_{e \in E} \ell(e) x_e \geq b_\ell \qquad \forall \ell \in \mathcal{L} \tag{9.11a}$$

$$x \in \mathbb{Z}_{\geq 0}^{|E|} \tag{9.11b}$$

where $\mathcal{L}$ is the set of all length functions $E \to \mathbb{R}_{\geq 0}$ and the right-hand sides b_ℓ of the Γ-robust length inequalities (9.11a) are determined as described in Lemma 9.2. Note, the convex hull of feasible points of (9.11) is the projection of the convex hull of feasible points of the formulations (9.6) and (9.7) onto the space of x.

9.2 Polyhedral study

In this section, we investigate the polyhedral structure of the convex hull of solutions feasible to the Γ-RNDP. First, we define the corresponding polyhedra for the link-flow and path-flow formulations and introduce notation for selected projections of these polyhedra. Second, we state basic characteristics of these objects. Finally, we investigate further valid inequalities based on cutsets, the residual capacity of arcs, and metric inequalities. To obtain strong inequalities, we follow the path of projecting the original polyhedra, characterizing classes of facet-defining inequalities and lifting them back to the original problem space.

We start with the definitions of the basic polyhedra:

Definition 9.3 (Γ-robust Network Design Flow Polyhedra)**.** We define the Γ-*robust network design link-flow polyhedron* $\mathcal{N}^{\mathrm{LF},\Gamma}$ as the convex hull of all feasible solutions of the compact link-flow formulation (9.6) of the Γ-RNDP, i. e.,

$$\mathcal{N}^{\mathrm{LF},\Gamma} := \mathrm{conv}\left\{ \begin{array}{r} (x,f,\pi,\rho) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{2|E||K|} \times \mathbb{R}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{|E||K|} \ : \\ (x,f,\pi,\rho) \text{ satisfies (9.6)} \end{array} \right\}. \tag{9.12}$$

We denote by $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}$ the projection of $\mathcal{N}^{\mathrm{LF},\Gamma}$ onto the space of the x and π variables, i. e.,

$$\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma} = \mathrm{proj}_{x,\pi}\, \mathcal{N}^{\mathrm{LF},\Gamma}, \tag{9.13}$$

and by $\mathcal{N}_{x}^{\mathrm{LF},\Gamma}$ the projection of $\mathcal{N}^{\mathrm{LF},\Gamma}$ onto the space of the x variables, i. e.,

$$\mathcal{N}_{x}^{\mathrm{LF},\Gamma} = \mathrm{proj}_{x}\, \mathcal{N}^{\mathrm{LF},\Gamma}. \tag{9.14}$$

Similarly, we define the Γ-*robust network design path-flow polyhedron* $\mathcal{N}^{\mathrm{PF},\Gamma}$ as the convex hull of all feasible solutions of the compact path-flow formulation (9.7) of the Γ-RNDP, i. e.,

$$\mathcal{N}^{\mathrm{PF},\Gamma} := \mathrm{conv}\left\{ \begin{array}{r} (x,f,\pi,\rho) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{|P|} \times \mathbb{R}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{|E||K|} \ : \\ (x,f,\pi,\rho) \text{ satisfies (9.7)} \end{array} \right\}. \tag{9.15}$$

We denote by $\mathcal{N}_{x}^{\mathrm{PF},\Gamma}$ the projection of $\mathcal{N}^{\mathrm{PF},\Gamma}$ onto the space of the x variables, i. e.,

$$\mathcal{N}_{x}^{\mathrm{PF},\Gamma} = \mathrm{proj}_{x}\, \mathcal{N}^{\mathrm{PF},\Gamma}. \tag{9.16}$$

Definition 9.4 (Γ-robust Network Design Capacity Polyhedron)**.** We define the Γ-*robust network design capacity polyhedron* $\mathcal{N}_x^\Gamma$ as the convex hull of all feasible solutions of the capacity formulation (9.11) of the Γ-RNDP, i. e.,

$$\mathcal{N}_x^\Gamma := \operatorname{conv}\left\{x \in \mathbb{Z}_{\geq 0}^{|E|} \;:\; x \text{ satisfies } (9.11)\right\}. \tag{9.17}$$

By construction of the capacity formulation (9.11), the following corollary holds

Corollary 9.5. $\mathcal{N}_x^\Gamma = \mathcal{N}_x^{LF,\Gamma} = \mathcal{N}_x^{PF,\Gamma}$.

9.2.1 Basic characteristics

In this section, we report on the dimensions of the Γ-robust network design link-flow and capacity polyhedra. The proofs follow similar proofs for the NDP.

Lemma 9.6. *The dimension of* $\mathcal{N}^{LF,\Gamma}$ *equals* $2|E| + 3|E||K| - (|V| - 1)|K|$ *whereas* $\mathcal{N}_{x,\pi}^{LF,\Gamma}$ *is full-dimensional.*

Proof. For $\mathcal{N}^{\mathrm{LF},\Gamma}$, there are $2|E|+3|E||K|$ variables and $(|V|-1)|K|$ linearly independent flow conservation constraints (1.19b). We show that there are no additional implied equations. Let

$$\sum_{e\in E}\alpha_e x_e + \sum_{e\in E}\beta_e\pi_e + \sum_{e\in E}\sum_{k\in K}\delta_e^k\rho_e^k + \sum_{e=\{i,j\}\in E}\sum_{k\in K}(\mu_{ij}^k f_{ij}^k + \mu_{ji}^k f_{ji}^k) = \gamma \tag{9.18}$$

be an equation satisfied by all points in $\mathcal{N}^{\mathrm{LF},\Gamma}$ and let $\hat{\varrho} = (\hat{x}, \hat{f}, \hat{\pi}, \hat{\rho}) \in \mathcal{N}^{\mathrm{LF},\Gamma}$. For all $e \in E$, we can modify $\hat{\varrho}$ by increasing the capacity without leaving $\mathcal{N}^{\mathrm{LF},\Gamma}$. Hence, $\alpha_e = 0$ for all $e \in E$. Once we have increased the capacity we can also increase variables π_e and ρ_e^k for every $e \in E$ and $k \in K$ which gives $\beta_e = \delta_e^k = 0$ for all $e \in E$ and $k \in K$. Now, we choose a spanning tree $T \subseteq E$ in G which exists since G is connected. By adding a linear combination of the flow conservation constraints (1.19b) to (9.18), we can assume that either μ_{ij}^k or $\mu_{ji}^k = 0$ for all $e = \{i, j\} \in T, k \in K$. Sending a small flow in both directions on every e in T gives $\mu_{ij}^k = \mu_{ji}^k = 0$. Now choosing an arbitrary edge $e \in E \setminus T$ there is an unique circuit consisting of e and edges in T. Sending small circulation flows on this circuit finally results in $\mu_{ij}^k = \mu_{ji}^k = 0$ for all $e = \{i, j\} \in E, k \in K$. It follows that (9.18) is a linear combination of flow conservation constraints which gives the desired results. By projecting all constructed points, we also conclude that $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}$ has dimension $2|E|$. □

Lemma 9.7 (Mattia [121])**.** *The polyhedron* $\mathcal{N}_x^\Gamma$ *is full-dimensional.*

9.2.2 Cutset-based inequalities

In the following, we investigate classes of valid inequalities based on cuts and the corresponding cutset edge sets. First, we generalize the classic cutset inequalities to their Γ-robust counterparts. Second, we introduce a new class of valid inequalities by further generalizing the Γ-robust cutset inequalities. The polyhedral study of cutset-based inequalities is based on joint work and has previously be published in Koster et al. [106]. An overview of all investigated polyhedra is shown in Figure 9.2.4.

Γ-robust cutset polyhedron. We consider a proper and non-empty subset S of the nodes V and the corresponding cutset $\delta(S)$ and denote by $Q_S \subseteq K$ the subset of commodities with source s^k and target t^k not in the same shore of the cut. Since we may always reverse single demands without changing the model, we may assume in this description $s^k \in S$ for all $k \in Q_S$. We denote by $\bar{d}_S := \sum_{k \in Q_S} \bar{d}^k$ the aggregated *nominal cut-demand* with respect to S. We will throughout assume that $|Q_S| \geq \Gamma \geq 1$. Notice that we can always reduce Γ to $|Q_S|$ without changing the problem on the cut. It follows $\bar{d}_S > 0$. Contracting both shores of the cut $\delta(S)$, we consider the following Γ-robust two-node formulation corresponding to (9.6):

$$\sum_{\{i,j\} \in \delta(S)} (f_{ij}^k - f_{ji}^k) = 1 \qquad \forall k \in Q_S \tag{9.19a}$$

$$\sum_{\{i,j\} \in \delta(S)} (f_{ij}^k - f_{ji}^k) = 0 \qquad \forall k \in K \setminus Q_S \tag{9.19b}$$

$$\Gamma \pi_e + \sum_{k \in K} \bar{d}^k f_e^k + \sum_{k \in K} \rho_e^k \leq x_e \qquad \forall e \in \delta(S) \tag{9.19c}$$

$$-\pi_e + \hat{d}^k f_e^k - \rho_e^k \leq 0 \qquad \forall e \in \delta(S), k \in K \tag{9.19d}$$

$$x, f, \rho, \pi \geq 0 \tag{9.19e}$$

We define the *Γ-robust cutset polyhedron* $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ with respect to S to be

$$\mathcal{N}^{\mathrm{LF},\Gamma}(S) := \operatorname{conv} \left\{ \begin{array}{r} (x, f, \pi, \rho) \in \mathbb{Z}_{\geq 0}^{|\delta(S)|} \times \mathbb{R}_{\geq 0}^{2|\delta(S)||K|} \times \mathbb{R}_{\geq 0}^{|\delta(S)|} \times \mathbb{R}_{\geq 0}^{|\delta(S)||K|} : \\ (x, f, \pi, \rho) \text{ satisfies (9.19)} \end{array} \right\}.$$

The projection of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ onto the space of the variables x and π is denoted by

$$\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S) := \operatorname{conv} \left\{ \begin{array}{r} (x, \pi) \in \mathbb{Z}_{\geq 0}^{|\delta(S)|} \times \mathbb{R}_{\geq 0}^{|\delta(S)|} : \exists\, (f, \rho) \in \mathbb{R}_{\geq 0}^{2|\delta(S)||K|} \times \mathbb{R}_{\geq 0}^{|\delta(S)||K|} \\ \text{so that } (x, f, \pi, \rho) \in \mathcal{N}^{\mathrm{LF},\Gamma}(S) \end{array} \right\}.$$

The following follows from Lemma 9.6 as $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defines a two-node Γ-robust network design problem.

Corollary 9.8. *The dimension of $\mathcal{N}^{LF,\Gamma}(S)$ equals $2|\delta(S)| + 3|\delta(S)||K| - |K|$ whereas $\mathcal{N}_{x,\pi}^{LF,\Gamma}(S)$ is full-dimensional.*

Γ-robust cutset inequalities. Independent of the realization of demand, all cut commodities Q_S have to be realized across the cut $\delta(S)$, that is $f^k(\delta(S)) \geq 1$ for all $k \in Q_S$. It follows that we have to provide sufficient cut capacity $x(\delta(S))$. Thus the following base cutset inequality must hold:

$$x(\delta(S)) \geq d_0 := \sum_{k \in Q_S} \bar{d}^k + \max_{Q'_S \subseteq Q_S : |Q'_S| \leq \Gamma} \sum_{k \in Q'_S} \hat{d}^k \tag{9.20}$$

It states that the capacity on the cut should be at least the nominal cut demand plus the Γ largest deviations among Q_S. Note that the right-hand side is independent of the realized flow. The value d_0 only depends on the cut $\delta(S)$ and the value of Γ. As the left-hand side is integral, we may round up the right-hand side yielding

$$x(\delta(S)) \geq \lceil d_0 \rceil . \tag{9.21}$$

This already generalizes the classic cutset inequality for network design; cf. Magnanti and Mirchandani [111]. Since no dual variables π_e appear in this inequality, it is also valid for the exponential formulation (9.4) and capacity formulation (9.11). We will prove in Corollary 9.24, inequality (9.21) defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ if $d_0 < \lceil d_0 \rceil$ and either $|\delta(S)| = 1$ or $d_0 > 1$. It also defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}$ if additionally the graphs defined by the two shores S and $V \setminus S$ are connected. In the rest of this section, we will generalize this essential result to a more general class of inequalities in the space of the x and π variables.

Generalizing Γ-robust cutset inequalities. Let us start by generalizing the base inequality (9.20). Let Q be an arbitrary but non-empty subset of the cut-commodities Q_S. From the flow-conservation constraints (9.19a) follows that

$$\sum_{k \in Q} \bar{d}^k f^k(\delta(S)) \geq \bar{d}(Q) \text{ and } \sum_{k \in Q} \hat{d}^k f^k(\delta(S)) \geq \hat{d}(Q). \tag{9.22}$$

Aggregating all capacity constraints (9.19c), adding all constraints (9.19d) for $e \in \delta(S)$ and $k \in Q$, using (9.22), and relaxing the backward flow variables results in

$$x(\delta(S)) + (|Q| - \Gamma)\pi(\delta(S)) \geq \bar{d}_S + \hat{d}(Q) . \tag{9.23}$$

The left-hand side of (9.23) is not changing as long as the cardinality of the subset Q is constant. Hence among all subsets of Q with cardinality $|Q|$ the one maximizing $\hat{d}(Q)$ gives the strongest inequality (9.23). To state this inequality, we have to sort the commodities non-increasingly w.r.t. the maximum deviation $\hat{d}^k$ and define subsets of Q_S corresponding to large deviations. This needs some new notation. Let $\varphi : Q_S \mapsto \{0, \ldots, |Q_S|\}$ be a permutation of the commodities in Q_S such that $\hat{d}^{\varphi^{-1}(1)} \geq \hat{d}^{\varphi^{-1}(2)} \geq \cdots \geq \hat{d}^{\varphi^{-1}(|Q_S|)}$ and let $J = \{-\Gamma, \ldots, |Q_S| - \Gamma\}$. Fixing the cut, we define $Q_i := \{k \in Q_S : \varphi(k) \leq i + \Gamma\}$ for $i \in J$ as the commodities corresponding to the $i + \Gamma$ largest $\hat{d}^k$ values with respect to Q_S. Hence, the demand $d_i := \bar{d}_S + \hat{d}(Q_i)$ denotes the total nominal demand plus the $i+\Gamma$ largest peak demands across the cut. This definition is consistent with the definition of d_0 in (9.20) since $|Q_S| \geq \Gamma$ and hence $\bar{d}_S = \sum_{k\in Q_S} \bar{d}^k$ and $\hat{d}(Q_0) = \max_{Q'_S \subseteq Q_S : |Q'_S| \leq \Gamma} \sum_{k \in Q'_S} \hat{d}^k$.

Using this notation, inequality (9.23) reduces to

$$x(\delta(S)) + i\pi(\delta(S)) \geq d_i. \tag{9.24}$$

It is valid for all $i \in J$ and by setting $i = 0$, we get inequality (9.21). In the sequel, we consider the polyhedron

$$X^\Gamma(S) = \text{conv}\left\{(x,\pi) \in \mathbb{Z}_{\geq 0}^{|\delta(S)|} \times \mathbb{R}_{\geq 0}^{|\delta(S)|} \ : \ (x,\pi) \text{ satisfies (9.24) } \forall i \in J\right\}$$

Every valid inequality for $X^\Gamma(S)$ is also valid for the Γ-robust formulation (9.6). In the following, we will completely describe $X^\Gamma(S)$ providing all facet-defining inequalities. Since all coefficients in (9.24) are identical for all edges in $\delta(S)$, it suffices to study the two-dimensional case with base inequalities

$$x + i\pi \geq d_i \tag{9.25}$$

for $i \in J$ and the related polyhedron

$$X^\Gamma = \text{conv}\left\{(x,\pi) \in \mathbb{Z}_{\geq 0} \times \mathbb{R}_{\geq 0} \ : \ (x,\pi) \text{ satisfy (9.25) for all } i \in J\right\}.$$

Notice that $X^\Gamma(S)$ is obtained from X^Γ by copying variables and forcing nonnegativity for the copied variables. It follows that every facet for X^Γ translates into a facet for $X^\Gamma(S)$ and vice versa except for the nonnegativity constraints. In fact, a complete description of X^Γ determines a complete description of $X^\Gamma(S)$ and vice versa. We formalize this observation as follows

Lemma 9.9. *Let $\alpha \in \mathbb{R}^{|Q_S|}$, $\beta \in \mathbb{R}^{|Q_S|}$ and $\gamma \in \mathbb{R}$. Let*

$$\sum_{e\in E} \alpha_e x_e + \sum_{e \in E} \beta_e \pi_e \geq \gamma$$

be different from a nonnegativity constraint and facet-defining for $X^\Gamma(S)$ with $|\delta(S)| \geq 2$. Then it holds $\alpha_{e_1} = \alpha_{e_2}$ and $\beta_{e_1} = \beta_{e_2}$ for all $e_1, e_2 \in \delta(S)$.

Proof. The statement is obviously true if $|\delta(S)| = 1$. Let $|\delta(S)| \geq 2$ and $e_1, e_2 \in \delta(S)$ with $e_1 \neq e_2$. There exists a point $(\bar{x}, \bar{\pi})$ on the facet such that $\bar{\pi}_{e_1} > 0$ since it is not a nonnegativity constraint. This point still satisfies all nonnegativity constraints and (9.24) if we move the x-values from e_1 to e_2. Hence, $(\alpha_{e_2} - \alpha_{e_1}) \geq 0$. But there is also a point $(\hat{x}, \hat{\pi})$ on the facet with $\hat{\pi}_{e_2} > 0$ from which we conclude that $(\alpha_{e_1} - \alpha_{e_2}) \geq 0$ and thus $\alpha_{e_1} = \alpha_{e_2}$. In a similar way, we argue that $\beta_{e_1} = \beta_{e_2}$. □

Lemma 9.10. *Let $\alpha, \beta, \gamma \in \mathbb{R}$ and $|\delta(S)| > 1$. Inequality $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ defines a facet for $X^\Gamma(S)$ if and only if $\alpha x + \beta\pi \geq \gamma$ defines facet of X^Γ and is not a nonnegativity constraint.*

Proof. First assume $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ defines a facet for $X^\Gamma(S)$. Hence, it is not the sum of nonnegativity constraints. Hence $\alpha x + \beta\pi \geq \gamma$ is not a nonnegativity constraint for X^Γ. Consider $2|\delta(S)|$ affinely independent points $(x^i, \pi^i) \in X^\Gamma(S)$ for $i = 1, \dots, 2|\delta(S)|$ satisfying $\alpha x^i(\delta(S)) + \beta\pi^i(\delta(S)) = \gamma$. We construct a point $(y^i, \omega^i) \in \mathbb{R}^2$ for every $i = 1, \dots, 2|\delta(S)|$ by setting $y^i := \sum_{e\in\delta(S)} x_e^i$ and $\omega^i := \sum_{e\in\delta(S)} \pi_e^i$. The points (y^i, ω^i) are valid for X^Γ and they satisfy $\alpha y^i(\delta(S)) + \beta\omega^i(\delta(S)) = \gamma$. There must exist at least two affinely independent points among (y^i, ω^i) otherwise the points (x^i, π^i) cannot be affinely independent.

Now assume $\alpha x + \beta\pi \geq \gamma$ defines a facet of X^Γ. There are two affinely independent points (y^i, ω^i), $i = 1, 2$ on the facet. From these points, we can construct two points in $X^\Gamma(S)$ on the face defined by $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ by setting the corresponding values to a single edge e only. Varying e, we get $2|\delta(S)|$ valid points. These are affinely independent since $\alpha x + \beta\pi \geq \gamma$ was not $x \geq 0$ and not $\pi \geq 0$, thus $y^1 > 0$ or $y^2 > 0$, and similarly either $\omega^1 > 0$ or $\omega^2 > 0$. □

Corollary 9.11. *Every facet-defining inequality $\alpha x + \beta\pi \geq \gamma$ for X^Γ with $\alpha, \beta, \gamma \in \mathbb{R}$ different from a nonnegativity constraint translates into a facet-defining inequality $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ for $X^\Gamma(S)$. All facets of $X^\Gamma(S)$ defined by inequalities different from nonnegativity constraints are of the form $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ for $\alpha, \beta, \gamma \in \mathbb{R}$ and correspond to a facet-defining inequality $\alpha x + \beta\pi \geq \gamma$ for X^Γ.*

Of course, we have $X^\Gamma = X^\Gamma(S)$ if and only if $|\delta(S)| = 1$. In the following, we will not distinguish facet-defining inequalities of X^Γ and $X^\Gamma(S)$ as long as they are different from nonnegativity constraints.

Envelope inequalities. Let us partition the index set J into three sets as follows: $J_- := \{-\Gamma, \dots, -1\}$, $J_+ := \{1, \dots, |Q_S| - \Gamma\}$, and $\{0\}$. Then, it holds $J = J_- \cup \{0\} \cup J_+$. The *upper region* of X^Γ corresponds to indices in J_- and the *lower region* of X^Γ corresponds to indices in J_+; see Figure 9.1(a). More precisely, the upper region is given by $X^\Gamma \cap \{\pi \geq d_0 - d_{-1}\}$ whereas we define the lower region of X^Γ as $X^\Gamma \cap \{\pi \leq d_1 - d_0\}$. Notice that the upper region is always non-empty if $\Gamma > 0$. The lower region is non-empty if and only if $|Q_S| > \Gamma$ and $d_1 > d_0$. Let us assume $\Gamma > 0$ in the following.

We call valid inequalities for X^Γ trivial if they are nonnegativity constraints or if they are of type (9.25). In the following we are only interested in non-trivial facets of X^Γ as these will translate to facets of $\mathcal{N}^{\mathrm{LF},\Gamma}$. We will call such inequalities *envelope inequalities* as they describe the envelope of X^Γ. Lower and upper regions are similar in structure. The lower region, however, is cut by $\pi \geq 0$ which leads to one additional type of facet. We will see that besides the vertical facet $x \geq \lceil d_0 \rceil$, there are two classes of non-trivial inequalities describing the lower region and one class of non-trivial inequalities describing the upper region facets.

Setting $r_i := \mathrm{frac}(d_i)$ and applying mixed integer rounding to (9.25) yields

$$r_i x + \max(0, i)\pi \geq r_i \lceil d_i \rceil, \tag{9.26}$$

which is valid for X^Γ; see Lemma 1.2. If $r_i = 1$, inequality (9.26) reduces to the base inequality $x + i\pi \geq d_i$. For $r_i < 1$ and $i = 0$, this inequality reduces to $x \geq \lceil d_0 \rceil$ which is the Γ-robust cutset inequality (9.21):

Lemma 9.12. *Inequality* (9.26) *defines a facet of* X^Γ *if* $i = 0$ *and* $r_i < 1$.

Proof. Consider $\epsilon > 0$ and the two affinely independent points $(\lceil d_0 \rceil, d_0 - d_{-1})$ and $(\lceil d_0 \rceil, d_0 - d_{-1} + \epsilon)$ which both satisfy (9.26) with equality. To see feasibility, notice that $d_0 - d_{-1} \geq 0$ gives the Γ largest deviation demand among Q_S. Setting

$$x_e = \lceil d_0 \rceil, \pi_e = d_0 - d_{-1}, \rho_e^k = \max(\hat{d}^k - \pi_e, 0), \text{ and } f_{uv}^k = 1 \text{ for } k \in Q_S \tag{9.27}$$

for some edge $e = \{u, v\} \in \delta(S)$ gives a feasible point for $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ which has a slack of $1 - r_0$ in the capacity constraint (9.19c) since $\Gamma\pi_e + \sum_{k\in K} \rho_e^k + \sum_{k \in K} \bar{d}^k f_e^k = d_0$. Hence, $(\lceil d_0 \rceil, d_0 - d_{-1})$ is feasible for X^Γ. As $r_0 < 1$ it follows also that the second point $(\lceil d_0 \rceil, d_0 - d_{-1} + \epsilon)$ is feasible for X^Γ for $\epsilon < 1 - r_0$. □

For $i \in J_-$, inequalities (9.26) are obviously dominated by (9.21). For $i \in J_+$, inequality (9.26) connects the two points $(\lfloor d_i \rfloor, r_i/i)$ and $(\lceil d_i \rceil, 0)$ in case $r_i < 1$. We get:

Lemma 9.13. *Assume* $J_+ \neq \emptyset$ *and* $\lceil d_{|Q_S|-\Gamma} \rceil > \lceil d_0 \rceil$. *Set* $i = \arg\max(r_\ell/\ell : \ell \in J_+ \text{ with } \lceil d_\ell \rceil = \lceil d_{|Q_S|-\Gamma} \rceil)$. *Inequality* (9.26) *defines a facet of* X^Γ *if* $r_i < 1$.

Proof. Let $\pi' := \max(r_\ell/\ell : \ell \in J_+ \text{ with } \lceil d_\ell \rceil = \lceil d_{|Q_S|-\Gamma} \rceil)$. Since $\lfloor d_i \rfloor \geq \lceil d_0 \rceil$, the two points $(\lfloor d_i \rfloor, \pi')$ and $(\lceil d_i \rceil, 0)$ are feasible . They satisfy (9.26) with equality and are affinely independent. □

Inequality (9.26) for $i = 0$ reduces to a cutset inequality in the space of the capacity variables. Inequality (9.26) for $i \in J_+$ is called a *lower envelope inequality* as by Lemma 9.13, it may define a facet of the lower region; see also Figure 9.1(a). However, in general, the two inequalities from Lemma 9.12 and Lemma 9.13 do not suffice to provide a complete description of X^Γ. To get a complete description of the lower region of X^Γ we have to consider two arbitrary base inequalities $x + i\pi \geq d_i$ and $x + j\pi \geq d_j$ with $i, j \in J_+, i < j$. Their intersection has x-value

$$b_{i,j} := (jd_i - id_j)/(j - i).$$

Now we can "connect" the two points $(\lfloor b_{i,j} \rfloor, (d_i - \lfloor b_{i,j} \rfloor)/i)$ and $(\lceil b_{i,j} \rceil, (d_j - \lceil b_{i,j} \rceil)/j)$ to obtain a valid inequality. Therefore let $r_{i,j} := (j - i)r(b_{i,j})$. Recall that $r_{i,j}$ defined this way is the remainder of the division of $jd_i - id_j$ by $(j - i)$ with $r_{i,j} = (j - i)$ in case $b_{i,j}$ is not fractional; see Lemma 1.2. Then we obtain the following inequality.

Lemma 9.14. *For* $i, j \in J_+$ *with* $i < j$, *the following inequality is valid for* X^Γ.

$$(i + r_{i,j})x + ij\pi \geq r_{i,j} \lceil b_{i,j} \rceil + id_j \tag{9.28}$$

Proof. We scale the two base inequalities with j and i, respectively:

$$jx + ji\pi \geq jd_i \quad \text{and} \quad ix + ij\pi \geq id_j.$$

Introducing the slack $s_j := ix + ij\pi - id_j \geq 0$ of the second constraint and combining the two inequalities gives

$$(j-i)x + s_j \geq jd_i - id_j.$$

Applying MIR and re-substituting results in (9.28). □

We also call (9.28) a *lower envelope inequality*. In a similar way, we combine two base constraints for $i, j \in J_-$ to get valid inequalities for the upper region of X^Γ.

Lemma 9.15. *For $i, j \in J_-$ with $i < j$, the following inequality is valid for X^Γ.*

$$(-j + r_{i,j})x - ij\pi \geq r_{i,j}\lceil b_{i,j}\rceil - jd_i \tag{9.29}$$

Proof. We multiply the base constraints for i and j by $-j$ and $-i$, respectively:

$$-jx - ji\pi \geq -jd_i \quad \text{and} \quad -ix - ij\pi \geq -id_j.$$

Introducing the slack $s_i := -jx - ji\pi + jd_i \geq 0$ for the first constraint and combining the two inequalities gives

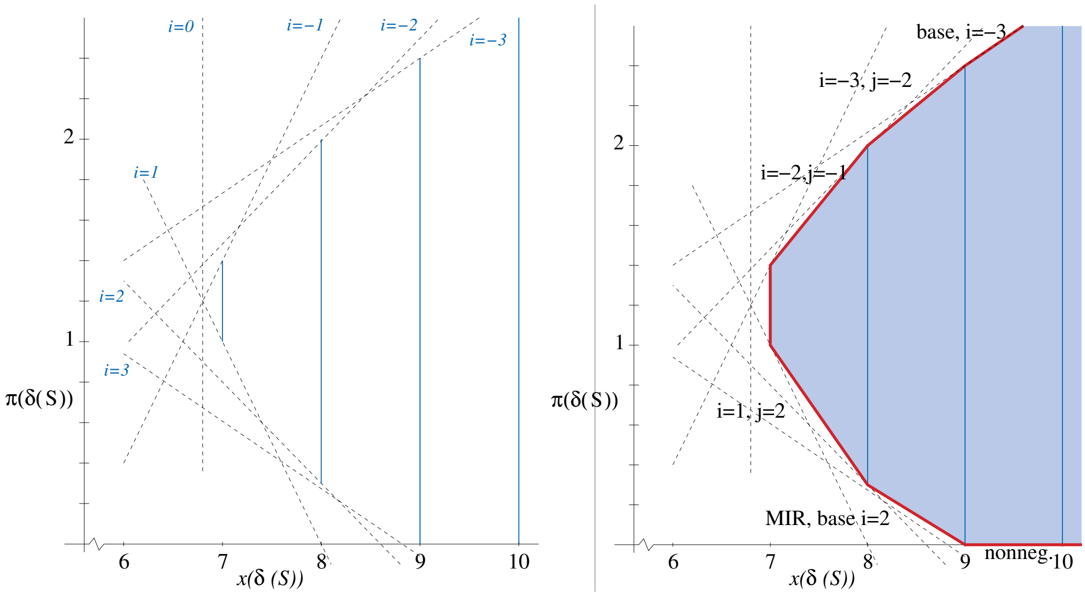
$$(j-i)x + s_i \geq jd_i - id_j.$$

Applying MIR and back substituting results in (9.29). □

We call (9.29) an *upper envelope inequality* as it defines a facet of the upper region of X^Γ. In case $b_{i,j}$ is fractional, inequalities (9.28) resp. (9.29) defined above cut off the fractional intersection point $(b_{i,j}, \pi)$ with $\pi = (d_i - b_{i,j})/i$ of the two base inequalities (9.25) corresponding to i and j. Note that by construction of the demand values d_i, it holds that $b_{i,i+1} \geq b_{i+1,i+2}$ for $0 > i \in J_-$ and $b_{i,i+1} \leq b_{i+1,i+2}$ for $0 < i \in J_+$. Also note that if $b_{i,j}$ is not fractional, then inequality (9.28) and (9.29) reduces to the base inequality for i and j, respectively. Of course, not every pair (i, j) results in a facet. In fact, only linearly many of the inequalities (9.28) and (9.29) are non-redundant. Let us define the function

$$\pi(\ell, x) := \frac{d_\ell - x}{\ell} \qquad \text{for all } \ell \in J_- \cup J_+ \text{ and } x \in \mathbb{R}_{\geq 0}.$$

We now consider an arbitrary interval $[a, a+1]$ with $a \in \mathbb{Z}, a \geq \lceil d_0 \rceil$ and easily determine the indices i, j that yield an inequality of (9.28) and (9.29) respectively dominating all others of this type on the chosen interval by simply maximizing (or minimizing respectively) the value $\pi(\ell, a)$ and $\pi(\ell, a+1)$. Doing so for all relevant values of a, we get all (non-trivial) facets of the lower and upper region, respectively:



(a) base inequalities and feasible set (b) convex envelope

Figure 9.1: Example of X^Γ and its completely described convex envelope with Γ-robust cutset, Γ-robust upper and lower envelope inequalities. Let $\Gamma = 3$, $|Q_S| = 6$, $\bar{d}_S = \frac{9}{5}$, and $\hat{d} = \frac{1}{5}(11\,8\,6\,6\,3\,1)^\top$. The upper Γ-robust envelope inequalities (9.29) for $i = -3$, $j = -2$ ($2x - 5\pi \geq 6$) and for $i = -2$, $j = -3$ ($3x - 5\pi \geq 14$), the lower Γ-robust envelope inequality (9.28) for $i = 1$, $j = 2$ ($7x + 10\pi \geq 59$), the lower Γ-robust envelope inequality (9.26) for $i = 2$ ($3x + 10\pi \geq 27$), the Γ-robust cutset inequality (9.21) $x \geq 7$, the base inequality (9.25) $x - 3\pi \geq \frac{9}{5}$ and nonnegativity $\pi \geq 0$ completely describe the convex hull.

Lemma 9.16. *Assume $J_+ \neq \emptyset$ and $\lceil d_{|Q_S|-\Gamma} \rceil > \lceil d_0 \rceil + 1$. For $a \in \mathbb{Z}$ with $\lceil d_0 \rceil \leq a \leq \lceil d_{|Q_S|-\Gamma} \rceil - 1$ let $i := \arg\max_{\ell \in J_+} \pi(\ell, a)$ and $j := \arg\max_{\ell \in J_+} \pi(\ell, a+1)$. If $i \neq j$, then inequality* (9.28) *defines a facet of X^Γ. If otherwise $i = j$, then the base inequality* (9.25) *defines a facet of X^Γ.*

Proof. If $i \neq j$ or $i = j$ then inequality (9.28) or (9.25) respectively connects the two affinely independent points $(a, \pi(i,a))$ and $(a+1, \pi(j, a+1))$, that is, they satisfy inequality (9.28) or (9.25) at equality, respectively. To see feasibility of the first point, we check that for $k \in J_+$ it holds $a + \ell\pi(i,a) \geq \ell\pi(\ell, a) = d_\ell$ by definition of i. For $\ell \in J_-$, we have $a + \ell\pi(i,a) \geq d_0 + \ell(d_i - d_0)/i \geq d_0 + \ell(d_k - d_0)/\ell = d_\ell$ where the first inequality follows from $a \geq d_0$ and the second inequality follows from $\ell < 0 < i$ and the definition of the demands d_i. The difference $d_i - d_{i-1}$ is non-increasing with i. Feasibility of the second point can be shown in a similar way. □

Note that for the lower region and $\lceil d_{|Q_S|-\Gamma} \rceil - 1 \leq x \leq \lceil d_{|Q_S|-\Gamma} \rceil$, we get a facet of type (9.26) by Lemma 1.2. For $x \geq \lceil d_{|Q_S|-\Gamma} \rceil$, we have $\pi \geq 0$ as a facet. Together, these inequalities completely describe the lower region. A complete description of the upper region of X^Γ is obtained with the following lemma which is proved similar to the proof of Lemma 9.16.

Lemma 9.17. *For $a \in \mathbb{Z}$ with $a \geq \lceil d_0 \rceil$, let $i = \arg\min_{\ell \in J_-} \pi(\ell, a+1)$ and $j = \arg\min_{\ell \in J_-} \pi(\ell, a)$. If $i \neq j$, then inequality* (9.29) *defines a facet of X^Γ. If otherwise $i = j$, then the base inequality* (9.25) *defines a facet of X^Γ.*

For $x \geq \lceil b_{\Gamma-1,\Gamma} \rceil$, the base inequality (9.25) for $i = -\Gamma$ is the only facet. Also notice that the pairs $\{i, j\}$ in Lemma 9.16 and Lemma 9.17 respectively are not unique. However, the resulting facet-defining inequalities are of course unique.

We have established different classes of facet-defining inequalities for X^Γ. In fact, all these inequalities together with the trivial facets completely describe X^Γ. This essentially follows already from the above since we stated the dominant inequalities for all intervals $[a, a+1]$ with $a \geq \lceil d_0 \rceil$. Figure 9.1 illustrates the convex envelope by example.

Completeness also follows from a result of Miller and Wolsey [123] who study a two-dimensional set (Model W) similar to X^Γ. Applying [123, Theorem 3] for the lower and upper region (using an appropriate variable transformation) respectively, we get:

Corollary 9.18. *The polyhedron X^Γ is completely described by the inequalities* (9.25), (9.26), (9.28), (9.29), *and $\pi \geq 0$.*

Lifting results. We have provided a complete and non-redundant description of X^Γ and thus of $X^\Gamma(S)$ in the previous paragraph. Next, we show how facets of $X^\Gamma(S)$ translate to facets of the cutset polyhedron $\mathcal{N}^{\text{LF},\Gamma}(S)$ and the original network design polyhedron $\mathcal{N}^{\text{LF},\Gamma}$. We also prove that the set $X^\Gamma(S)$ is identical to $\mathcal{N}^{\text{LF},\Gamma}_{x,\pi}(S)$, the projection of the cutset polyhedron $\mathcal{N}^{\text{LF},\Gamma}(S)$ to the space of the x and π variables, if the cut contains a single edge.

Lemma 9.19. *$\mathcal{N}^{LF,\Gamma}_{x,\pi}(S) \subseteq X^\Gamma(S)$. Moreover, $\mathcal{N}^{LF,\Gamma}_{x,\pi}(S) = X^\Gamma(S)$ if and only if $|\delta(S)| = 1$.*

Proof. For $(x, \pi) \in \mathcal{N}^{\text{LF},\Gamma}_{x,\pi}(S)$, let $(x, f, \pi, \rho) \in \mathcal{N}^{\text{LF},\Gamma}(S)$. Inequalities (9.24) are valid for $\mathcal{N}^{\text{LF},\Gamma}(S)$ which gives $(x, \pi) \in X^\Gamma(S)$ and $\mathcal{N}^{\text{LF},\Gamma}_{x,\pi}(S) \subseteq X^\Gamma(S)$.

Let $\delta(S) = \{e\}$ with $e = \{i, j\}$ for $i, j \in V$. Given $(x, \pi) \in X^\Gamma(S)$, we set $f^k_{ij} := 1, f^k_{ji} := 0$, and $\rho^k_e := \max(0, \hat{d}^k - \pi_e)$ for all $k \in Q_S$. Now (x, f, π, ρ) obviously satisfies (9.19a), (9.19d), and (9.19e). Moreover,

$$\begin{aligned}
\Gamma\pi_e + \sum_{k \in Q_S} \bar{d}^k f^k_e + \sum_{k \in Q_S} \rho^k_e &= \Gamma\pi_e + \sum_{k \in Q_S} \bar{d}^k + \sum_{k \in Q_S} \max(0, \hat{d}^k - \pi_e) \\
&= \Gamma\pi_e + \bar{d}_S + \hat{d}(Q_i) - (i + \Gamma)\pi_e \\
&\leq x_e
\end{aligned}$$

for some $i \in J$ using the introduced ordering of demands and (9.24). It follows that (x, f, π, ρ) satisfies (9.19c) and hence, $(x, \pi) \in \mathcal{N}^{\text{LF},\Gamma}_{x,\pi}(S)$.

It remains to show that $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S) \neq X^{\Gamma}(S)$ if $|\delta(S)| > 1$. Let $e_1, e_2 \in \delta(S)$. There is a point (x, f, π, ρ) in $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ with $x_{e_1}, \pi_{e_1} > 0$ and $x_{e_2}, \pi_{e_2} = 0$. We simply route all traffic on e_1 and set x_{e_1}, π_{e_1} large enough. Clearly, $(x, \pi) \in X^{\Gamma}(S)$ as already shown. We modify this point by shifting the capacity from e_1 to e_2 but keeping the value π_{e_1} such that $x_{e_1} = 0$ and $\pi_{e_1} > 0$. This gives a vector $(x, \pi) \in X^{\Gamma}(S) \setminus \mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$ since inequalities (9.24) are still satisfied but (9.19c) is violated for e_1. □

Notice, from Lemma 9.19 follows that any point (x, π) which is defined on a single edge, that is, there exists $e \in \delta(S)$ such that $x_f = \pi_f = 0$ for all $f \in \delta(S)$, $f \neq e$, is valid for $X^{\Gamma}(S)$ if and only if it is valid for $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$. We will use this fact several times below.

Lemma 9.20. *Every facet-defining inequality for $X^{\Gamma}(S)$ different from a nonnegativity constraint defines a facet of $\mathcal{N}_{x,\pi}^{LF,\Gamma}(S)$.*

Proof. By Lemma 9.9, we can assume that the facet of $X^{\Gamma}(S)$ is defined by $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ for $\alpha, \beta, \gamma \in \mathbb{R}$. Consider $2|\delta(S)|$ affinely independent points $(x^i, \pi^i) \in X^{\Gamma}(S)$ for $i = 1, \ldots, 2|\delta(S)|$ satisfying $\alpha x^i(\delta(S)) + \beta\pi^i(\delta(S)) = \gamma$. Given an arbitrary edge $f \in \delta(S)$ we construct a point $(\tilde{x}^i, \tilde{\pi}^i)$ for every $i = 1, \ldots, 2|\delta(S)|$ by shifting all entries to edge f, more precisely $\tilde{x}^i_f := \sum_{e\in\delta(S)} x^i_e$ and $\tilde{\pi}^i_f := \sum_{e\in\delta(S)} \pi^i_e$. All other entries are set to zero: $\tilde{x}^i_e := \tilde{\pi}^i_e := 0$ for all $e \in \delta(S) \setminus \{f\}$. The points $(\tilde{x}^i, \tilde{\pi}^i)$ are valid for $X^{\Gamma}(S)$ and they satisfy $\alpha\tilde{x}^i(\delta(S)) + \beta\tilde{\pi}^i(\delta(S)) = \gamma$. Moreover, since $(\tilde{x}^i, \tilde{\pi}^i)$ is defined on a single edge, we get $(\tilde{x}^i, \tilde{\pi}^i) \in \mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$. Notice, that $(\tilde{x}^i, \tilde{\pi}^i) \neq 0$ as there is at least one cut demand. There must exist at least two affinely independent points among $(\tilde{x}^i, \tilde{\pi}^i)$, otherwise the points (x^i, π^i) cannot be affinely independent. Assume these points are $(\tilde{x}^1, \tilde{\pi}^1)$ and $(\tilde{x}^2, \tilde{\pi}^2)$. The proof is complete for $|\delta(S)| = 1$. In case $|\delta(S)| > 1$, we can assume that either $\tilde{x}^1_f > 0$ or $\tilde{x}^2_f > 0$ and similarly either $\tilde{\pi}^1_f > 0$ or $\tilde{\pi}^2_f > 0$. Otherwise, the original points (x^i, π^i) are all contained in the face defined by $x(\delta(S)) \geq 0$ or $\pi(\delta(S)) \geq 0$ respectively which is a contradiction as the sum of nonnegativity constraints cannot define a facet. Now we vary $f \in \delta(S)$ which gives $2|\delta(S)|$ affinely independent points, both in $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$ and on the face defined by $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$. □

Lemma 9.21. *Let $\alpha, \beta, \gamma \in \mathbb{R}$. If $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ defines a facet for $\mathcal{N}_{x,\pi}^{LF,\Gamma}(S)$ then it also defines a facet for $X^{\Gamma}(S)$.*

Proof. Since $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S) \subseteq X^{\Gamma}(S)$ is full-dimensional, we only have to show that $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$ is valid for $X^{\Gamma}(S)$. Assume the contrary. We take a point in $X^{\Gamma}(S)$ which violates $\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma$. Now we modify this point by shifting everything to one edge. The constructed point is also valid for $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$ as shown above but violates the facet-defining inequality which is a contradiction. □

We call facet-defining inequalities for $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$ non-trivial if they are non-trivial for $X^{\Gamma}(S)$, i. e., they are different from nonnegativity constraints and different from (9.24).

Theorem 9.22. *Every non-trivial facet-defining inequality*

$$\alpha x(\delta(S)) + \beta\pi(\delta(S)) \geq \gamma \tag{9.30}$$

for $\mathcal{N}_{x,\pi}^{LF,\Gamma}(S)$ also defines a facet of $\mathcal{N}^{LF,\Gamma}(S)$ if one of the following conditions holds:

- $|\delta(S)| = 1$ *and there exists a feasible point* (x, f, π, ρ) *on the face of* $\mathcal{N}^{LF,\Gamma}(S)$ *defined by* (9.30) *such that the link capacity constraint* (9.6a) *is not tight.*
- $|\delta(S)| \geq 2$ *and there exists a feasible point* (x, f, π, ρ) *on the face of* $\mathcal{N}^{LF,\Gamma}(S)$ *defined by* (9.30) *such that the capacity constraint* (9.6a) *is not tight for at least two different edges.*

Proof. We assume that (9.30) does not define a facet for $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$. Hence, every point $(x, f, \pi, \rho) \in \mathcal{N}^{\mathrm{LF},\Gamma}(S)$ satisfying (9.30) with equality must be contained in a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defined by

$$\sum_{e\in\delta(S)} \alpha_e x_e + \sum_{e\in\delta(S)} \beta_e \pi_e + \sum_{e\in\delta(S)} \sum_{k\in K} \delta_e^k \rho_e^k + \sum_{e=\{i,j\}\in\delta(S)} \sum_{k\in K} (\mu_{ij}^k f_{ij}^k + \mu_{ji}^k f_{ji}^k) \geq \gamma \tag{9.31}$$

By adding flow conservation constraints to (9.31), we conclude that $\mu_{ij}^k = 0$ for an arbitrary edge $e = ij \in \delta(S)$ and all $k \in K$. We may hence assume that for the same edge the capacity constraint is not tight for the point (x, f, π, ρ) on the face of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defined by (9.30). By increasing ρ_e^k we see that $\delta_e^k = 0$ for all $k \in K$. Similarly, sending a small circulation flow on e, we conclude $\mu_{ji}^k = 0$. Notice that by these perturbations we never leave the face.

Now assume that $|\delta(S)| \geq 2$. There is a second edge $e' \neq e$ such that the corresponding capacity constraint is not tight. Since we may exchange variable values of two different edges without leaving the face (9.30), edge $e' \neq e$ is in fact arbitrary. By sending circulation flow using edges e and e' and by increasing $\rho_{e'}^k$, it follows that $\delta_e^k = \mu_{ij}^k = \mu_{ji}^k = 0$ for all edges $e \in \delta(S)$ and commodities $k \in K$.

Since (9.30) defines a facet of $\mathcal{N}_{x,\pi}^{\mathrm{LF},\Gamma}(S)$, $2|\delta(S)|$ affinely independent points exist. These points can be lifted to points in $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ remaining affinely independent in the (x, π) space and satisfying (9.30) as well as (9.31) with equality. We showed that only the $2|\delta(S)|$ coefficients in (9.31) corresponding to the x and π variables are nonzero. Hence (9.31) equals (9.30) up to scaling and up to a linear combination of flow conservation constraints. It follows that (9.30) defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$. □

We call a valid inequality for $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ non-trivial if it is different from the constraints (9.19a)-(9.19e) defining $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$. The following result is a straightforward generalization of the corresponding result for the deterministic case from Raack et al. [141]; also see Agarwal [7].

Lemma 9.23. *Every non-trivial facet-defining inequality of* $\mathcal{N}^{LF,\Gamma}(S)$ *defines a facet of* $\mathcal{N}^{LF,\Gamma}$ *if both cut shores are connected.*

The proof of Lemma 9.23 is based on the fact that in case both shores are connected, then the flow for commodities in $K \setminus Q_S$ can be routed in the two shores without using cut edges. This means that we can construct feasible points for $\mathcal{N}^{\mathrm{LF},\Gamma}$ from points valid for $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ without changing the cut values. This is done by assigning sufficiently large values for x_e, π_e, and ρ_e^k for edges $e \in E \setminus \delta(S)$ and then decomposing the problem w.r.t. the two graphs defined by S and $V \setminus S$, respectively.

Corollary 9.24. *Given a node set $S \subset V$ such that the two shores of the corresponding cut $\delta(S)$ are connected, the cutset inequality (9.21) defines a facet of $\mathcal{N}^{LF,\Gamma}$ if $\operatorname{frac}(d_0) < 1$ and either $|\delta(S)| = 1$ or $d_0 > 1$.*

Proof. By Lemma 9.12 and Lemma 9.20, inequality (9.21) defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$. In this case inequality (9.21) is also non-trivial for $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$. Fixing edge $e = \{i,j\} \in \delta(S)$, we consider the point (x, f, π, ρ) defined in (9.27) which is on the face of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defined by (9.21). All other variables are set to zero. Recall that the capacity constraint of e has a slack of $1 - \operatorname{frac}(d_0)$. In case $|\delta(S)| \geq 2$ and $d_0 > 1$ and hence $\lceil d_0 \rceil \geq 2$, we can shift one unit of capacity to a second edge. Also a fraction of $1/\lceil d_0 \rceil$ of all other variables is shifted to the second edge. This way, we construct a point on the face with two edges not being tight in the capacity constraint. Hence, using Theorem 9.22 and Lemma 9.23 we get the desired result. □

Corollary 9.25. *Given a node set $S \subset V$ such that the two shores of the corresponding cut $\delta(S)$ are connected and $J_+ \neq \emptyset$ as well as $\lceil d_{|Q_S|-\Gamma} \rceil > \lceil d_0 \rceil$, the lower envelope inequality*

$$r_i x(\delta(S)) + \max(0,i)\pi(\delta(S)) \geq r_i \lceil d_i \rceil \tag{9.32}$$

defines a facet of $\mathcal{N}^{LF,\Gamma}$ if $i = \arg\max(r_\ell/\ell : \ell \in J_+$ with $\lceil d_\ell \rceil = \lceil d_{|Q_S|-\Gamma} \rceil)$ and $r_i < 1$.

Proof. By Lemma 9.13 and Lemma 9.20, inequality (9.32) defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$ if $r_i < 1$. In this case, inequality (9.32) is also non-trivial for $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$. Fixing $e \in \delta(S)$ we consider the following point (x, f, π, ρ) on the face of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defined by (9.32)

$$x_e = \lceil d_i \rceil = \lceil d_{|Q_S|-\Gamma} \rceil, \pi_e = 0, \rho_e^k = \hat{d}^k, \text{ and } f_{ij}^k = 1 \text{ for } k \in Q_S.$$

All other variables are set to zero. The point protects against all demands across the cut at their peak. There is a slack of at least $1 - r_i$. Now $\lceil d_{|Q_S|-\Gamma} \rceil \geq 2$ since $\lceil d_0 \rceil \geq 1$. If $|\delta(S)| \geq 2$, we can hence shift one unit of capacity and a fraction of all other variables to a second edge such that two edges are non-tight in the capacity constraint. Hence, using Theorem 9.22 and Lemma 9.23 we get the desired result. □

Corollary 9.26. *Given a node set $S \subset V$ such that the two shores of the corresponding cut $\delta(S)$ are connected, the upper envelope inequality*

$$(-j + r_{i,j})x(\delta(S)) - ij\pi(\delta(S)) \geq r_{i,j} \lceil b_{i,j} \rceil - jd_i \tag{9.33}$$

defines a facet of $\mathcal{N}^{LF,\Gamma}$ if $i, j \in J_-$, $i < j$, such that $i = \arg\min_{\ell \in J_-} \pi(\ell, a+1)$ and $j = \arg\min_{\ell \in J_-} \pi(\ell, a)$ with $r_{i,j} < 1$ and $a \in \mathbb{Z}$ with $a \geq \lceil d_0 \rceil$ having either $|\delta(S)| = 1$ or $a \geq 2$.

Proof. By Lemma 9.17 and Lemma 9.20, inequality (9.33) defines a facet of $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$. From $r_{i,j} < 1$ it follows that $i < j$ and the break point $b_{i,j}$ is fractional and hence (9.33) is non-trivial for $\mathcal{N}^{\mathrm{LF},\Gamma}_{x,\pi}(S)$. Let F be the face of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ defined by (9.33). There is a point $(\bar{x}, \bar{\pi})$ with $a < \bar{x} < a + 1$ cut off in the linear relaxation of X^Γ by (9.33). Using a single edge e we may of course lift this point to a valid point $(\bar{x}, \bar{f}, \bar{\pi}, \bar{\rho})$ of the linear relaxation of $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$. Set $\alpha = (-j + r_{i,j})$, $\beta = -ij$, and $\gamma = r_{i,j}\lceil b_{i,j}\rceil - jd_i$. The point $(\dot{x}, \bar{\pi})$ with $\dot{x} = \frac{\gamma - \beta\bar{\pi}}{\alpha} > \bar{x}$ is in X^Γ and lies on the facet. Moreover $p_1 := (\dot{x}, \bar{f}, \bar{\pi}, \bar{\rho}) \in F$ is such that for the selected single edge e the capacity constraint is not tight. However, p_1 is not feasible because $\dot{x}$ with $a < \dot{x} < a+1$ is not integral. Consider the two points $(a, \pi(j,a))$ and $(a+1, \pi(i, a+1))$ on the facet of X^Γ and denote by p_2 and p_3 the two corresponding points lifted to $\mathcal{N}^{\mathrm{LF},\Gamma}(S)$ on the face F. We can assume that p_2 and p_3 have nonzero values only on edge e and that p_1 is a convex combination of p_2 and p_3. Hence at least one of p_2 or p_3 is not tight in the capacity constraint of e. The proof is complete in case $|\delta(S)| = 1$. Assume $|\delta(S)| \geq 2$. By shifting one unit of capacity to a second edge e_2 we construct points p_4 and p_5 from p_2 and p_3 similar to the proof of Corollary 9.24. As long as $a \geq 2$, at least one of these points is not tight in the capacity constraint of at least two edges. By Theorem 9.22 and Lemma 9.23, the claim follows. □

Corollary 9.27. *Let $S \subset E$ be a node set such that the two shores of the corresponding cut $\delta(S)$ are connected and $J_+ \neq \emptyset$ as well as $\lceil d_{|Q_S|-\Gamma}\rceil > \lceil d_0 \rceil + 1$. The lower envelope inequality*

$$(i + r_{i,j})x(\delta(S)) + ij\pi(\delta(S)) \geq r_{i,j}\lceil b_{i,j}\rceil + id_j \tag{9.34}$$

defines a facet of $\mathcal{N}^{LF,\Gamma}$ if $i, j \in J_+$, $i < j$, such that $i := \arg\max_{\ell \in J_+} \pi(\ell, a)$ and $j := \arg\max_{\ell \in J_+} \pi(\ell, a+1)$ with $r_{i,j} < 1$ and $a \in \mathbb{Z}$ with $\lceil d_0 \rceil \leq a \leq \lceil d_{|Q_S|-\Gamma}\rceil - 1$ having either $|\delta(S)| = 1$ or $a \geq 2$.

Proof. Similar to the proof of Corollary 9.26. □

9.2.3 Γ-robust arc residual capacity inequalities

In this section, we investigate a relaxation of the Γ-RNDP, the Γ-robust single arc design problem and its associated polyhedron. We identify classes of facet-defining inequalities for this polyhedron which are also valid for Γ-robust network design polyhedra. An overview of all investigated polyhedra is shown in Figure 9.2.4.

Our results generalize the polyhedral study of the classic single arc design problem and polyhedron to the Γ-robust setting; see Magnanti et al. [113] for results on the classic problem.

Γ-robust single arc design polyhedron. We consider the compact link-flow formulation of the Γ-RNDP (9.6) and apply Lagrangian relaxation to its flow conservation constraints (1.19b). The resulting ILP reads

$$\max_{\lambda} \min \quad \sum_{e\in E} \kappa_e x_e + \sum_{k\in K} \sum_{i\in V} \lambda_{k,i} \left(\sum_{j\,:\,ij\in E} (f_{ij}^k - f_{ji}^k) - \delta_{k,i} \right) \tag{9.35a}$$

$$\Gamma\pi_e + \sum_{k\in K} \bar{d}^k f_e^k + \sum_{k\in K} \rho_e^k \leq x_e, \qquad \forall e \in E \tag{9.35b}$$

$$-\pi_e + \hat{d}^k f_e^k - \rho_e^k \leq 0 \qquad \forall e \in E, k \in K \tag{9.35c}$$

$$f, x, \rho, \pi \geq 0 \tag{9.35d}$$

$$x \in \mathbb{Z}^{|E|} \tag{9.35e}$$

with Lagrangian multipliers $\lambda_{k,i}$ penalizing the violation of the flow conservation constraint for $k \in K$ and $i \in V$, and $\delta_{k,i} := 1$ if $i = s^k$, -1 if $i = t^k$, and 0 otherwise. The objective (9.35a) can be rewritten as follows

$$\max_{\lambda} \min \sum_{e\in E} \kappa_e x_e + \sum_{ij\in E} \sum_{k\in K} (\lambda_{k,ij} f_{ij}^k + \lambda_{k,ji} f_{ji}^k) - \sum_{k\in K} \sum_{i\in V} \lambda_{k,i} \delta_{k,i} \tag{9.35a'}$$

where $\lambda_{k,ij} := \lambda_{k,i} - \lambda_{k,j}$. Clearly, formulation (9.35) with its rewritten objective (9.35a') decomposes for a given λ into $|E|$ separate problems. For a fixed edge $e = ij \in E$, we can assume by a flow cancellation argument that there is a flow f_e^k only in one direction of the edge with value $|f_{ij}^k - f_{ji}^k|$. Thus we drop the edge-related variable subscripts. Furthermore, we can assume w. l. o. g. that this flow is at most 1 since thus all demands still can be satisfied. Hence, each individual problem can be written as

$$\min \quad \kappa x + \sum_{k\in K} \lambda_k f^k \tag{9.36a}$$

$$\text{s.t.} \quad \Gamma\pi + \sum_{k\in K} \bar{d}^k f^k + \sum_{k\in K} \rho^k \leq x \tag{9.36b}$$

$$-\pi + \hat{d}^k f^k - \rho^k \leq 0 \qquad \forall k \in K \tag{9.36c}$$

$$f \leq 1 \tag{9.36d}$$

$$f, x, \rho, \pi \geq 0 \tag{9.36e}$$

$$x \in \mathbb{Z} \tag{9.36f}$$

We call (9.36) the *Γ-robust single arc design problem (Γ-RSADP)*. The associated polyhedron, the *Γ-robust single arc design polyhedron*, is defined as the convex hull of all feasible solutions to (9.36), i. e.,

$$\mathcal{S}^{\mathrm{LF},\Gamma} := \left\{ \begin{array}{r} (x, f, \pi, \rho) \in \mathbb{Z}_{\geq 0} \times [0,1]^{|K|} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{|K|} \; : \\ (x, f, \pi, \rho) \text{ satisfies (9.36b)–(9.36c)} \end{array} \right\} \tag{9.37}$$

The dimension and trivial facets of $\mathcal{S}^{\mathrm{LF},\Gamma}$ are presented first.

Lemma 9.28. *The polyhedron $\mathcal{S}^{LF,\Gamma}$ is full-dimensional, i. e.,* $\dim(\mathcal{S}^{LF,\Gamma}) = 2|K| + 2$.

Proof. The following $2|K| + 3$ points are feasible to $\mathcal{S}^{\mathrm{LF},\Gamma}$ and affinely independent: $p^k_\rho := (1, 0, 0, e_k)$ for $k \in K$, $p^k_f := (1, \frac{1}{\bar{d}^k+\hat{d}^k} e_k, 0, \frac{\hat{d}^k}{\bar{d}^k+\hat{d}^k} e_k)$ for $k \in K$, $p_\pi := (1, 0, \frac{1}{\Gamma}, 0)$, the point $p_x := (1, 0, 0, 0)$, and $p_0 := (0, 0, 0, 0)$. □

Lemma 9.29. *The following inequalities define trivial facets of* $\mathcal{S}^{\mathrm{LF},\Gamma}$*:*

1. *nonnegativity constraints* $f^k \geq 0$ *for all* $k \in K$
2. *nonnegativity constraint* $\pi \geq 0$
3. *nonnegativity constraints* $\rho^k \geq 0$ *for all* $k \in K$
4. *dual constraints* (9.36c) *for all* $k \in K$
5. *capacity constraint* (9.36b)

Proof. Let the points p^k_ρ, p^k_f, p_π, and p_x be constructed as in the proof of Lemma 9.28. For each of the six types of constraints in Lemma , we construct $2|K| + 2$ affinely independent points that satisfy the corresponding constraint with equality and are feasible to $\mathcal{S}^{\mathrm{LF},\Gamma}$:

1. For $k' \in K$, consider the points p^k_ρ for all $k \in K$, p^k_f for all $k \in K \setminus \{k'\}$, p_π, p_x, and p_0.
2. Consider the points p^k_ρ for all $k \in K$, p^k_f for all $k \in K$, p_x, and p_0.
3. For $k' \in K$, consider the points p^k_ρ for all $k \in K \setminus \{k'\}$, p^k_f for all $k \in K$, p_π, p_x, the point $(1, \frac{1}{\bar{d}_{k'}+\Gamma\hat{d}_{k'}} e_{k'}, \frac{\hat{d}_{k'}}{\bar{d}_{k'}+\Gamma\hat{d}_{k'}}, 0)$, and p_0.
4. For $k' \in K$, consider the points p^k_f for all $k \in K$, the points p_x, p_0, the points $(\left\lceil \bar{d}_{k'} + |K|\hat{d}_{k'} \right\rceil, e_{k'}, 0, e_k)$ for all $k \in K \setminus \{k'\}$, and $(\bar{d}_{k'} + \Gamma\hat{d}_{k'}, e_{k'}, \hat{d}_{k'}, 0)$.
5. Consider the points p^k_ρ for all $k \in K$, p^k_f for all $k \in K$, p_π, and p_0.

This completes the proof. □

We denote by $\mathcal{S}^{\mathrm{LF},\Gamma}(Q)$ the Γ-robust single arc design polyhedron $\mathcal{S}^{\mathrm{LF},\Gamma}$ restricted to the subset $Q \subseteq K$ of commodities. The following lifting result motivates the study the polyhedral structure of the lower-dimensional $\mathcal{S}^{\mathrm{LF},\Gamma}(Q)$ instead of $\mathcal{S}^{\mathrm{LF},\Gamma}$.

Lemma 9.30. *Let* $\mathcal{S}^{LF,\Gamma}(Q)$ *be the polyhedron* $\mathcal{S}^{LF,\Gamma}$ *restricted to* $Q \subseteq K$*. Let* $\alpha^\top f + \beta\pi + \gamma^\top\rho + \delta x \leq 1$ *define a facet* F *of* $\mathcal{S}^{LF,\Gamma}(Q)$*. If* F *is not defined by* (9.36b)*, then it also defines a facet of* $\mathcal{S}^{LF,\Gamma}$*.*

Proof. As F is a facet of $\mathcal{S}^{\mathrm{LF},\Gamma}(Q)$, there exist $i = 0, \ldots, 2|Q| + 2$ affinely independent points $p^i_Q \in \mathcal{S}^{\mathrm{LF},\Gamma}(Q)$ with $\alpha^\top f + \beta\pi + \gamma^\top\rho + \delta x = 1$. For each point p^i_Q, we construct a corresponding point $p^i \in \mathcal{S}^{\mathrm{LF},\Gamma}$ by copying the entries and filling entries $k \in K \setminus Q$ with 0. The points p^i are affinely independent.

Since F is not defined by (9.36b), the inequality (9.36b) is not satisfied with equality by at least one of the points p_Q^i. W.l.o.g. let p_Q^0 be such a point. Define $s_0 := x^0 - \Gamma\pi^0 + \sum_{k\in K} \bar{d}^k f^{0,k} + \sum_{k\in K} \rho^{0,k} = x_Q^0 - \Gamma\pi_Q^0 + \sum_{k\in K} \bar{d}^k f_Q^{0,k} + \sum_{k\in K} \rho_Q^{0,k}$ as the corresponding slack value. It holds $s_0 > 0$. For $k \in K \setminus Q$, define

$$\zeta^k := \min\{1, s_0\}/((|K| - |Q|)(\bar{d}^k + \hat{d}^k)).$$

Then, for $k \in K \setminus Q$ the points $q^k := p^0 + (0, \zeta^k e_k, 0, \zeta^k \hat{d}^k e_k)$ and the points $\bar{q}^k := p^0 + (0, 0, 0, \min\{1, s_0\}e_k)$ are feasible for $\mathcal{S}^{\mathrm{LF},\Gamma}$ by construction. Furthermore, all points q^k and $\bar{q}^k$ are on the facet F as p^0 is on the facet. Finally, the points p^i, q^k, and $\bar{q}^k$ are affinely independent by construction completing the proof. □

Γ-robust arc residual capacity inequalities. In the following, we identify classes of valid or facet-defining inequalities for the $\mathcal{S}^{\mathrm{LF},\Gamma}$. First, we consider the simple case $|Q| = \Gamma$.

Lemma 9.31. *Let $Q \subseteq K$, $|Q| = \Gamma$, $d(Q) := \bar{d}(Q) + \hat{d}(Q)$, $r^Q := d(Q) - (\lceil d(Q)\rceil - 1)$. The* simple Γ-robust arc residual capacity inequality

$$\sum_{k\in Q} (\bar{d}^k + \hat{d}^k) f^k \leq r^Q x + (\lceil d(Q)\rceil - 1)(1 - r^Q) \tag{9.38}$$

is valid for $\mathcal{S}^{LF,\Gamma}$. Furthermore, it defines a facet of $\mathcal{S}^{LF,\Gamma}$ if and only if $r^Q < 1$.

Proof. First we show the validity of inequality (9.38) by applying mixed integer rounding. Therefore, we relax the Γ-robust capacity constraint (9.36b) by restricting it to the subset $Q \subseteq K$ of commodities. Then it reads

$$\Gamma\pi + \sum_{k\in Q} \bar{d}^k f^k + \sum_{k\in Q} \rho^k \leq x.$$

Next, we add the dual constraint (9.36c) for all $k \in Q$ to the relaxed constraint. This yields

$$\sum_{k\in Q} (\bar{d}^k + \hat{d}^k) f^k \leq x. \tag{9.39}$$

Due to the upper bounds on the flow variables, the left-hand side is bounded: $\sum_{k\in Q}(\bar{d}^k + \hat{d}^k) f^k \leq d(Q)$. By introducing a slack variable $s_Q := d(Q) - \sum_{k\in Q}(\bar{d}^k + \hat{d}^k) f^k \geq 0$ and combining it with (9.39), we obtain

$$x + s_Q \geq d(Q).$$

Applying mixed integer rounding to this inequality and back substitution yields the simple Γ-robust arc residual capacity inequality (9.38) and completes the proof of its validity.

Second, we investigate the conditions that inequality (9.38) defines a facet of $\mathcal{S}^{\mathrm{LF},\Gamma}$. If $r^Q = 1$, then the simple Γ-robust arc residual capacity inequality (9.38) is dominated by its base inequality (9.39).

If $r^Q < 1$, then there exist solutions $(x^*, f^*, \pi^*, \rho^*)$ and $(x^{**}, f^{**}, \pi^{**}, \rho^{**})$ with $x^* = \lfloor d(Q) \rfloor$ and $x^{**} = \lceil d(Q) \rceil$ satisfying the inequality (9.38) with equality. We can vary the solution $(x^*, f^*, \pi^*, \rho^*)$ in $|Q| - 1$ ways while still satisfying (9.38) with equality.

The solution $(x^{**}, f^{**}, \pi^{**}, \rho^{**})$ does not satisfy the capacity constraint (9.36b) with equality. Hence, we can increase π^{**} and ρ^{**} (for all $k \in Q$) while still satisfying (9.38) with equality.

The vector $(x^*, f^*, \pi^*, \rho^*)$, its $|Q| - 1$ variations, the vector $(x^{**}, f^{**}, \pi^{**}, \rho^{**})$, and its $1 + |Q|$ variations, are feasible for $\mathcal{S}^{\mathrm{LF},\Gamma}$ and affinely independent. This completes the proof. □

Next, we discuss a more general case.

Lemma 9.32. *Let $Q \subseteq K$, $R \subseteq Q$, $d(Q, R) := \bar{d}(Q) + \hat{d}(R)$, $r^{Q,R} := d(Q, R) - (\lceil d(Q, R) \rceil - 1)$. The* Γ-robust arc residual capacity inequality

$$\sum_{k \in Q} \bar{d}^k f^k + \sum_{k \in R} \hat{d}^k f^k + (\Gamma - |R|)^- \pi \leq r^{Q,R} x + (\lceil d(Q, R) \rceil - 1)(1 - r^{Q,R}) \tag{9.40}$$

is valid for $\mathcal{S}^{LF,\Gamma}$.

Proof. The proof is similar to the proof of Lemma 9.31. First, we relax the Γ-robust capacity constraint (9.36b) by restricting it to $Q \subseteq K$. Then we add the dual constraint (9.36c) for all $k \in R$ and the nonnegativity constraint $\rho^k \geq 0$ for $k \in Q \setminus R$ to the relaxed constraint. This yields

$$\sum_{k \in Q} \bar{d}^k f^k + \sum_{k \in R} \hat{d}^k f^k + (\Gamma - |R|)^- \pi \leq x \tag{9.41}$$

where we only consider nonpositive values of $(\Gamma - |R|)$ by adding $(\Gamma - |R|)$ times the nonnegativity constraint $\pi \geq 0$ if $(\Gamma - |R|)$ is positive. Note, the sum $\sum_{k \in Q} \bar{d}^k f^k + \sum_{k \in R} \hat{d}^k f^k$ is bounded by $d(Q, R)$. We introduce a slack variable $s_{Q,R} := d(Q, R) - \sum_{k \in Q} \bar{d}^k f^k - \sum_{k \in R} \hat{d}^k f^k$ and combine it with inequality (9.41) to obtain

$$x + s \geq d(Q, R) \tag{9.42}$$

where $s := s_{Q,R} - (\Gamma - |R|)^- \pi$ is now a a single continuous variable and apply mixed integer rounding to this inequality. By back substitution this yields the Γ-robust arc residual capacity inequality (9.40) and completes the proof. □

9.2.4 Γ-robust metric inequalities

In this section, we investigate the capacity formulation of the Γ-RNDP and the associated polyhedron $\mathcal{N}_x^\Gamma$. First, we consider the linear relaxation ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ of $\mathcal{N}_x^\Gamma$ and obtain a complete description of this polyhedron. Based on this result, we completely describe $\mathcal{N}_x^\Gamma$ itself in a second step. The polyhedral study of Γ-robust metric inequalities is based on joint work and has been submitted for publication (Claßen et al. [55]). An overview of all investigated polyhedra is shown in Figure 9.2.4.

Now, let us consider the polyhedron ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ of the linear relaxation of $\mathcal{N}_x^\Gamma$ and its facial structure. We present a constructive proof which is specific to the Γ-RNDP with static routing and splittable flows. If we proved the theorem analogously to Avella et al. [17], the right-hand side value b could not be specified.

Theorem 9.33. *Let $\bar{\ell}x \geq b_{\bar{\ell}}$ be any valid inequality for ${}^{LP}\mathcal{N}_x^\Gamma$. Then there exists a metric $\ell_M \in \mathrm{Met}(G)$ with*

- *$\ell_M x \geq b_{\bar{\ell}}$ is valid for ${}^{LP}\mathcal{N}_x^\Gamma$,*
- *$\ell_M(ij) \leq \bar{\ell}(ij)$ for all edges ij.*

Proof. Let $\bar{\ell}$ be a non-metric length function and $b_{\bar{\ell}}$ the corresponding optimal solution of (9.9). The length function $\bar{\ell}$ is transformed to a metric by setting $\ell_M(ij) := \bar{\ell}(i,j)$ for all $ij \in E$. Since only values are changed that are not already the shortest path value, all changes can be done simultaneously and the resulting function is a metric. To prove that the right-hand sides $b_{\bar{\ell}}$ and b_{ℓ_M} are the same, we however assume that the lengths are adapted one at a time. Hence, without loss of generality we assume that there exists only one edge uv with $\bar{\ell}(uv) > \bar{\ell}(u,v)$.

Let $p_1 = (u,v)$ be the direct path from u to v. Since $\bar{\ell}(uv) > \bar{\ell}(u,v)$, there must exist a path $p_2 \neq p_1$ such that $\bar{\ell}(p_2) = \bar{\ell}(u,v)$ and

$$\ell_M(p_1) = \ell_M(p_2) = \ell_M(uv) = \bar{\ell}(u,v).$$

Given a length function ℓ, the dual of (9.9) reads as follows.

$$\min \sum_{k\in K} \bar{d}^k \sum_{p\in P^k} \ell(p)\mu^k(p) + \Gamma \sum_{e\in E} \nu(e)\ell(e) + \sum_{k\in K}\sum_{e\in E} z^k(e)\ell(e) \tag{9.43a}$$

$$\text{s.t.} \sum_{p\in P^k} \mu^k(p) \geq 1 \qquad \forall k \in K \tag{9.43b}$$

$$\nu(e) + z^k(e) - \hat{d}^k\ell(e) \sum_{p\in P^k: e\in p} \mu^k(p) \geq 0 \qquad \forall k \in K,\, e \in E \tag{9.43c}$$

$$\mu^k(p),\, \nu(e),\, z^k(e) \geq 0 \qquad \forall k \in K,\, p \in P^k,\, e \in E, \tag{9.43d}$$

where $\{\mu^k(p)\}_{k\in K, p\in P^k}$, $\{\nu(e)\}_{e\in E}$ and $\{z^k(e)\}_{k\in K, e\in E}$ are the dual variables for (9.9b), (9.9c), and (9.9d), respectively. Let $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$ be the optimal solution of this dual LP (9.43) with $\ell = \ell_M$. Hence, $\ell_M x \geq b_{\ell_M}$ is a valid inequality. If we show $b_{\ell_M} \geq b_{\bar{\ell}}$, then $\ell_M x \geq b_{\bar{\ell}}$ is valid and the proof is completed.

Therefore, we construct a feasible as well as optimal solution (μ, ν, z) of the dual LP (9.43) with $\ell = \ell_M$, prove that (μ, ν, z) is also feasible for the dual LP (9.43) with $\ell = \bar{\ell}$ and that the objective values coincide. Given $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$, define

$$\nu(e) := \begin{cases} \tilde{\nu}(e) + \tilde{\nu}(uv) & e \in p_2 \\ 0 & e = uv \\ \tilde{\nu}(e) & \text{otherwise,} \end{cases}$$

$$z^k(e) := \begin{cases} \tilde{z}^k(e) + \tilde{z}^k(uv) & e \in p_2 \\ 0 & e = uv \\ \tilde{z}^k(e) & \text{otherwise.} \end{cases}$$

For the definition of μ, we first define two specific subsets of paths $P^k_{uv} := \{p \in P^k : uv \in p\} \subseteq P^k$ and $P^k_{p_2} := \{p \in P^k : p = p_{uv} \setminus \{uv\} \cup p_2\ ,\ p_{uv} \in P^k_{uv}\} \subseteq P^k$, where $P^k_{p_2}$ is the set of paths which use p_2 instead of the edge uv. Hence, for each $p \in P^k_{p_2}$ exists exactly one path $p_{uv} \in P^k_{uv}$. We can now define μ as follows.

$$\mu^k(p) := \begin{cases} \tilde{\mu}^k(p) + \tilde{\mu}^k(p_{uv}) & p \in P^k_{p_2} \\ 0 & p \in P^k_{uv} \\ \tilde{\mu}^k(p) & \text{otherwise.} \end{cases}$$

Thus, we shift the value $\tilde{\mu}^k(p_{uv})$ for every path using edge uv to the corresponding path using p_2 instead of uv. It is easy to see that the objective values for $(\tilde{\mu}, \tilde{\nu}, \tilde{z})$ and (μ, ν, z) are equal regarding ℓ_M. Furthermore, (9.43b) is fulfilled for all $k \in K$. If $e = uv$ then (9.43c) is valid since

$$\nu(uv) + z^k(uv) = 0 + 0 \geq \hat{d}^k \ell_M(uv) \sum_{p \in P^k_{uv}} \mu^k(p) = 0.$$

If $e \in p$ with $p \in P^k_{p_2}$, then (9.43c) is equivalent to

$$\begin{aligned} & \nu(e) + z^k(e) - \hat{d}^k \ell_M(e) \mu^k(p) \\ = \ & \tilde{\nu}(e) + \tilde{\nu}(uv) + \tilde{z}^k(e) + \tilde{z}^k(uv) - \hat{d}^k \ell_M(e) \left(\tilde{\mu}^k(p_{uv}) + \tilde{\mu}^k(p)\right) \\ = \ & \underbrace{\tilde{\nu}(e) + \tilde{z}^k(e) - \hat{d}^k \ell_M(e) \tilde{\mu}^k(p)}_{\geq 0} + \tilde{\nu}(uv) + \tilde{z}^k(uv) - \hat{d}^k \underbrace{\ell_M(e)}_{\leq \ell_M(p_2) = \ell_M(uv)} \tilde{\mu}^k(p_{uv}) \\ \geq \ & 0 + \tilde{\nu}(uv) + \tilde{z}^k(uv) - \hat{d}^k \ell_M(uv) \tilde{\mu}^k(p_{uv}) \\ \geq \ & 0, \end{aligned}$$

thus, also fulfilled. So (μ, ν, z) is feasible (and also optimal) for (9.43) with $\ell = \ell_M$. Moreover, $\mu^k(p) = 0$ for all $p \in P^k$ with $uv \in p$. Therefore, the objective value does not change for $\bar{\ell}$ and (μ, ν, z) is also feasible for (9.43) with $\ell = \bar{\ell}$ since $\bar{\ell}(e) = \ell_M(e)$ for all $e \in E \setminus \{uv\}$. Thus, $b_{\ell_M} \geq b_{\bar{\ell}}$ and the proof is completed. Note that we even have $b_{\ell_M} = b_{\bar{\ell}}$ since $b_{\ell_M} \leq b_{\bar{\ell}}$ holds as $\ell_M(e) \leq \bar{\ell}(e)\ \forall e \in E$. □

Corollary 9.34. *RMIs together with the nonnegativity completely describe* ${}^{LP}\mathcal{N}^{\Gamma}_x$*, i.e.,*

$${}^{LP}\mathcal{N}^{\Gamma}_x = \left\{ x \in \mathbb{R}^{|E|}_{\geq 0} : \sum_{e \in E} \ell(e) x_e \geq b_\ell,\ \forall \ell \in \mathrm{Met}(G) \right\}.$$

Based on the preceding result, we can now introduce a complete description of the Γ-RNDP in the capacity space.

$$
\begin{aligned}
&(1.20a)\\
\text{s.t.}\ &\sum_{e\in E} \ell_M(e)x_e \geq b && \forall \ell_M \in \mathrm{Met}(G) && (9.44a)\\
&x_e \geq 0 && \forall e \in E && (9.44b)
\end{aligned}
$$

Due to the fact that Farkas' lemma can only be applied to LPs, Theorem 9.33 cannot be directly transferred to the integer case since we do not always have a good characterization of the right-hand side b and $\mathcal{N}_x^\Gamma$ itself. Therefore, the following theorem extends Theorem 9.33 to integer capacity variables. (Here, b is not necessarily defined by (9.9).)

Theorem 9.35. *Let $\ell x \geq b$ be any valid inequality for $\mathcal{N}_x^\Gamma$. Then there exists a metric $\ell_M \in \mathrm{Met}(G)$ with*

- *$\ell_M x \geq b$ valid for $\mathcal{N}_x^\Gamma$,*
- *$\ell_M(ij) \leq \ell(ij)$ for all edges ij.*

Proof. Let ℓ be a non-metric length function and b the corresponding right-hand side of the valid inequality. We define the metric ℓ_M as $\ell_M(ij) := \ell(i,j)$. Now assume there exists an $\bar{x} \in \mathcal{N}_x^\Gamma$ with $\ell\bar{x} \geq b$ but $\ell_M\bar{x} < b$. Thus, there must exist an edge $ij \in E$ with $\ell_M(ij) < \ell(ij)$ and $\bar{x}_{ij} > 0$.

Let b_ℓ be the optimal solution of (9.9) regarding ℓ. Then $\ell x \geq b_\ell$ is a valid inequality for ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ and by Theorem 9.33, $\ell_M x \geq b_\ell$ is also valid for ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$. We determine a feasible solution $\bar{y}$ of ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ with $\bar{y}_e \leq \bar{x}_e$ $\forall e \in E$ based on a feasible flow $\bar{f} \in [0,1]^{|P|}$ computed as follows.

By definition of ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$, there exist $(\bar{f}, \bar{\pi}, \bar{\rho})$ such that $(\bar{x}, \bar{f}, \bar{\pi}, \bar{\rho})$ satisfies (9.7). We set for all $e \in E$

$$\bar{y}_e := \sum_{k\in K} \bar{d}^k \bar{f}^k(e) + \Gamma\bar{\pi}_e + \sum_{k\in K} \bar{\rho}_e^k,$$

the left-hand side of the capacity constraint (9.7a). Hence, $\bar{y} \in {}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ and $\ell\bar{y} \geq b_\ell$ is valid. Now replace the objective of (9.7) by the all zero function and exploit LP duality. Compared to (9.9), the objective now reads

$$\max\ -\sum_{e\in E} \bar{x}_e \ell(e) + \sum_{k\in K} b_k.$$

This value constitutes the violation of the inequality $\ell\bar{x} \geq b_\ell$ and, by the strong duality, is equal to 0. Hence, $\ell\bar{x} = b_\ell$. Since $\ell\bar{y} \geq b_\ell$ and $\bar{x} \geq \bar{y}$, it follows $\ell\bar{y} = b_\ell$. Since $\ell_M x \geq b_\ell$ is valid for all $x \in {}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$, it follows that $\bar{y}_{ij} = 0$ for all $ij \in E$ with $\ell_M(ij) < \ell(ij)$. Set $\tilde{x}_e = \lceil \bar{y}_e \rceil$ $\forall e \in E$. Clearly, $\tilde{x} \in \mathcal{N}_x^\Gamma$ but $\ell\tilde{x} = \ell_M\tilde{x} \leq \ell_M\bar{x} < b$, which implies that $\ell x \geq b$ is not valid for $\tilde{x} \in \mathcal{N}_x^\Gamma$, a contradiction. □

It is possible to prove Theorem 9.35 analogously to [17] or [121]. However, we have chosen this proof since it is constructive and makes an explicit use of the compact formulation of the Γ-RNDP with static routing and splittable flows. A further way to prove Theorem 9.35 is by Chvátal-Gomory derivations as follows. By Theorem 9.33, ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ can be completely described by RMIs. Additionally, $\mathcal{N}_x^\Gamma \subseteq {}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ holds. All facet-defining inequalities defining $\mathcal{N}_x^\Gamma$ can be derived by a sequence of Chvátal-Gomory derivations from the facet-defining inequalities of ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$. Since $\mathrm{Met}(G)$ is a cone, these derived inequalities are also RMIs. Thus, $\mathcal{N}_x^\Gamma$ can be completely described by RMIs.

The right-hand side of valid inequalities $\ell_M x \geq b$ in Theorem 9.35 cannot be derived by (9.9) or a similar procedure. A first step can be done by Chvátal-Gomory rounding. For a RMI (9.44a), we can assume $\ell_M \in \mathbb{Z}_{\geq 0}^{|E|}$. (If $\ell_M(e) \notin \mathbb{Z}_{\geq 0}$ for $e \in E$, the metric can be scaled such that all lengths are integer.) Since $x \in \mathbb{Z}_{\geq 0}$, the *rounded* Γ*-robust metric inequality* (rounded RMI)

$$\sum_{e\in E} \ell_M(e) x_e \geq \lceil b_{\ell_M} \rceil \tag{9.45}$$

is also a valid inequality for $\mathcal{N}_x^\Gamma$. The strongest RMI is defined as follows.

Definition 9.36. Let $\ell_M \in \mathrm{Met}(G)$ and let β_{ℓ_M} be the optimal solution of the Γ-RNDP where the cost vector κ in the objective function (1.20a) is replaced by ℓ_M. Thus, $\beta_{\ell_M} = \min\{\ell_M x : x \in \mathbb{Z}_{\geq 0},\ x \in \mathcal{N}_x^\Gamma\}$. Obviously, any valid RMI $\ell_M x \geq b$ for $\mathcal{N}_x^\Gamma$ is dominated by $\ell_M x \geq \beta_{\ell_M}$. Hence, any inequality of the form

$$\ell_M x \geq \beta_{\ell_M} \tag{9.46}$$

is denoted as *tight* Γ*-robust metric inequality* (tight RMI).

Corollary 9.37. *Tight RMIs $\ell_M x \geq \beta_{\ell_M}$ with $\ell_M \in \mathrm{Met}(G)$ together with nonnegativity completely describe $\mathcal{N}_x^\Gamma$.*

The right-hand side value β_{ℓ_M} of a tight RMI can be determined easily without solving a minimization problem over $\mathcal{N}_x^\Gamma$ if the metric ℓ_M is an extreme ray of the metric cone. This result is known for the classic NDP [17] and can be generalized to the Γ-RNDP with static routing and splittable flows as follows.

Theorem 9.38. *If $\ell_M : E \to \mathbb{Z}_{\geq 0}$ is an extreme ray of the metric cone* $\mathrm{Met}(G)$ *such that the greatest common divisor of ℓ_M is 1, then $\beta_{\ell_M} = \lceil b_{\ell_M} \rceil$.*

Proof. We recall that ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$ can be completely described by RMIs due to Theorem 9.33 and the subset correlation $\mathcal{N}_x^\Gamma \subseteq {}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$. All valid inequalities defining $\mathcal{N}_x^\Gamma$ can be derived by a sequence of Chvátal-Gomory derivations from the facet-defining inequalities of ${}^{\mathrm{LP}}\mathcal{N}_x^\Gamma$. Consider a metric length function ℓ_M that satisfies the conditions of the theorem and suppose that, for all $x \in \mathcal{N}_x^\Gamma$, ℓ_M satisfies the inequality $\ell_M x \geq \alpha$ with $\alpha > \lceil b_{\ell_M} \rceil$. This inequality cannot be derived by a Chvátal-Gomory procedure from the single RMI $\ell_M x \geq b_{\ell_M}$. Thus, it can only be derived by combining two or more inequalities in at least one step of the sequence of Chvátal-Gomory derivations. But this contradicts the assumption that ℓ_M is an extreme ray of $\mathrm{Met}(G)$. □

Note, following a two-stage approach to solve the Γ-RNDP, an analogous proof has been done by [121] in the context of the Γ-RNDP with dynamic routing.

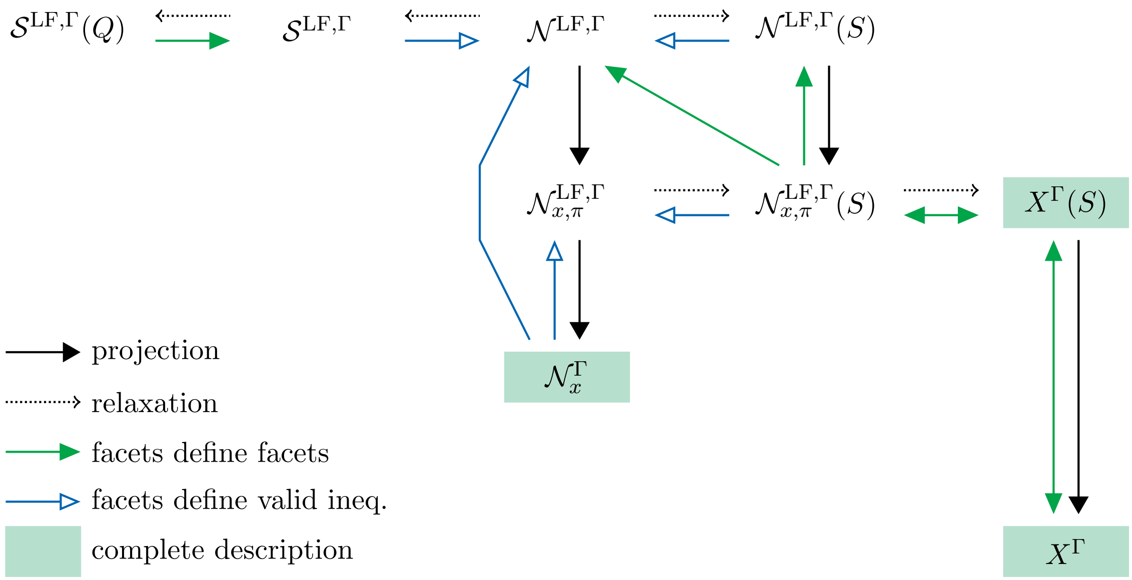


Figure 9.2: Overview of investigated polyhedra

Selected subclasses of Γ-robust metric inequalities Some well-known classes of valid inequalities for the Γ-RNDP are generalized by RMIs.

We define *(L-)bounded RMIs* as the subclass of RMIs consisting of all rounded RMIs obtained from metrics with integer link lengths bounded by $L \in \mathbb{Z}_{>0}$, i. e., $\mathrm{Met}_L(G) := \{\ell_M \in \mathrm{Met}(G) : \ell_M \in \{0, 1, .., L\}^{|E|}\}$.

The class of 1-bounded RMIs consists of all Γ-robust partition inequalities, i.e., the Γ-robust counterparts of partition inequalities (see [7, 8] and the references therein). A well-known subclass of Γ-robust partition inequalities are the Γ-robust 2-partition inequalities, better known as the Γ-robust cutset inequalities (9.24).

9.3 Algorithms

In this section, we present separation algorithms for selected classes of valid inequalities. We discuss both, exact algorithms as well as separation heuristics.

9.3.1 Separation of cutset-based inequalities

We have investigated the Γ-robust cutset inequalities and the envelope inequalities in Section 9.2.2. Here, we describe several approaches to separate violated inequalities of these types. Note, the Γ-robust cutset inequalities are a special case of the Γ-robust metric inequalities. Thus, complementing algorithms can be found in Section 9.2.4.

Exact separation of Γ-robust cutset inequalities (Enumeration). To study the effectiveness of the valid inequalities, we propose an exact separation algorithm which enumerates all network cuts explicitly and generates *all* violated inequalities of type (9.21), (9.32), (9.33), and (9.34). Clearly, this approach is suited for small networks only as the number of network cuts that must be enumerated increases exponentially. Still, for small networks this enumerative algorithm can be used to investigate the maximal effectiveness of these inequalities that can be achieved by separating all existing violated inequalities in terms of improving the root node dual bound.

Exact separation of Γ-robust cutset inequalities (ILP). Further, we present another exact separation algorithm which solves an ILP to separate a most violated inequality of type (9.21). It was introduced by us in Koster et al. [104] and is computationally tractable for larger networks as well.

We define binary variables δ_i $(i \in V)$ with $\delta_i = 1$ if and only if $i \in S$ determining the cut, α^k with $\alpha^k = 1$ if and only if $k \in Q_S$ determining the cut-crossing commodities, γ^k with $\gamma^k = 1$ if and only if commodity $k \in Q_S$ deviates from its nominal, and $\bar{\delta}_{ij}$ $(ij \in E)$ with $\bar{\delta}_{ij} = 1$ if and only if $ij \in \delta(S)$ determining the cutset. In addition, let d determine the worst-case total demand value crossing the cut, and let R be the right-hand side value of the corresponding cutset inequality (9.21). Given an LP solution x^*, we minimize the feasibility (i.e., maximize the violation) of inequality (9.21) such that a negative objective value yields a violated cut. Then, the ILP formulation of the separation problem is given by

$$\min \sum_{ij \in E} x^*_{ij} \bar{\delta}_{ij} - R$$

$$\text{s.t.} \ \max\{\delta_i - \delta_j, \delta_j - \delta_i\} \leq \bar{\delta}_{ij} \leq \min\{\delta_i + \delta_j, 2 - \delta_i - \delta_j\} \quad \forall ij \in E \tag{9.47a}$$

$$\max\{\delta_{s^k} - \delta_{t^k}, \delta_{t^k} - \delta_{s^k}\} \leq \alpha^k \leq \min\{\delta_{s^k} + \delta_{t^k}, 2 - \delta_{s^k} - \delta_{t^k}\} \quad \forall k \in Q \tag{9.47b}$$

$$\gamma^k \leq \alpha^k \quad \forall k \in Q \tag{9.47c}$$

$$\sum_{k \in Q} \gamma^k \leq \Gamma \tag{9.47d}$$

$$\sum_{k \in Q} (\bar{d}^k \alpha^k + \hat{d}^k \gamma^k) = d \tag{9.47e}$$

$$d \leq R \leq d + 1 - \varepsilon \tag{9.47f}$$

$$\alpha^k, \delta_i, \gamma^k, \bar{\delta}_{ij} \in \{0,1\}, R \in \mathbb{Z}_{\geq 0}, d \geq 0 \quad \forall k \in Q, \forall ij \in E, \forall i \in V \tag{9.47g}$$

where constraints (9.47a), (9.47b), and (9.47c) define the logical dependencies between the indicator variables α^k, δ_i, and γ^k. Constraint (9.47d) limits the number of deviating commodities to Γ. The total demand d is calculated by (9.47e). Constraint (9.47f) guarantees the round-up of the right-hand side variable R using $0 < \varepsilon \ll 1$ to avoid rounding R to $\lceil d \rceil + 1$ or higher. Note, by setting a node limit (or time limit) for solving ILP (9.47), we obtain a (non-deterministic) heuristic ILP-based separation algorithm.

Heuristic separation of Γ-robust cutset inequalities (Shrinking). Complementing the exact separation algorithms, we propose the following heuristic separation algorithm: Violated inequalities are separated for all single node network cuts (i.e., $\delta(S)$ with

$|\delta(S)| = 1$) as well as a set of network cuts resulting from a graph shrinking heuristic. This graph shrinking heuristic generalizes a shrinking heuristic dating back to Bienstock et al. [40], Günlük [79] and used by Raack et al. [141] for the deterministic model (1.19).

The idea of this extended graph shrinking heuristic is the following: The base inequality of the Γ-robust cutset inequality (9.21) is the sum of flow conservation constraints (1.19b), capacity constraints (9.6a), and constraints (9.6b). For violated cutset inequalities we need (almost) tight base inequalities. Hence we wish to have edges e in the cut $\delta(S)$ that have (almost) no slack in the constraints (9.6a) and (9.6b). In the shrinking heuristic, we hence shrink edges whose corresponding model constraints have large slacks. Technically, we try to minimize the sum of weights w_e for edges e on the cut: Given the solution of the current LP relaxation, we use $w_e := s_e^{(9.6a)} + \sum_{k\in K} s_{e,k}^{(9.6b)}$ where $s_e^{(9.6a)}$ denotes the slack of the capacity constraint (9.6a) for edge e and the $s_{e,k}^{(9.6b)}$ the slack of constraint (9.6b) for edge e and commodity k. By contracting edges in non-increasing order of w_e, we shrink the network until only η nodes or no edges with positive weight are left. Based on empirical values of previous computational studies, we propose to set $\eta = 5$.

Let $\bar{N}(\bar{V}, \bar{E})$ be the remaining shrunken network with node set $\bar{V}$ and edge set $\bar{E}$. Then, the set of network cuts returned by the shrinking heuristic consists of all network cuts corresponding to single node network cuts in $\bar{N}$ as well as to at most $|\bar{V}|^2$ additional network cuts in $\bar{N}$ obtained by enumeration.

9.3.2 Separation of Γ-robust arc residual capacity inequalities

Next, we present an exact algorithm to separate Γ-robust arc residual capacity inequalities efficiently. It is based on two necessary conditions on the positive violation of such inequalities.

Lemma 9.39. *Given a solution $(x^*, f^*, \pi^*, \rho^*)$ of ${}^{LP}\mathcal{S}^{LF,\Gamma}$. A Γ-robust arc residual capacity inequality (9.40) can only be violated if $\lfloor x^* \rfloor < d(Q,R) < \lceil x^* \rceil$ holds.*

Proof. Suppose $d(Q,R) \le \lfloor x^* \rfloor$, then inequality (9.40) is dominated by

$$\sum_{k\in Q} \bar{d}^k f^{*k} + \sum_{k\in R} \hat{d}^k f^{*k} + (\Gamma - |R|)^- \pi^* \le x^*.$$

Suppose $d(Q,R) \ge \lceil x^* \rceil$, then inequality (9.40) is dominated by

$$\sum_{k\in Q} \bar{d}^k f^{*k} + \sum_{k\in R} \hat{d}^k f^{*k} + (\Gamma - |R|)^- \pi^* \le d(Q,R).$$

Figure 9.3 illustrates these observations. □

Given a solution $(x^*, f^*, \pi^*, \rho^*)$ of ${}^{\mathrm{LP}}\mathcal{S}^{\mathrm{LF},\Gamma}$, we consider the violation of a Γ-robust arc residual capacity inequality as a function of the sets Q and R:

$$\begin{aligned} &\text{violation}(Q,R) \\ :=& \sum_{k\in Q} \bar{d}^k f^{*k} + \sum_{k\in R} \hat{d}^k f^{*k} + (\Gamma - |R|)^- \pi^* - r^{Q,R} x^* - (\lceil d(Q,R) \rceil - 1)(1 - r^{Q,R}) \end{aligned}$$

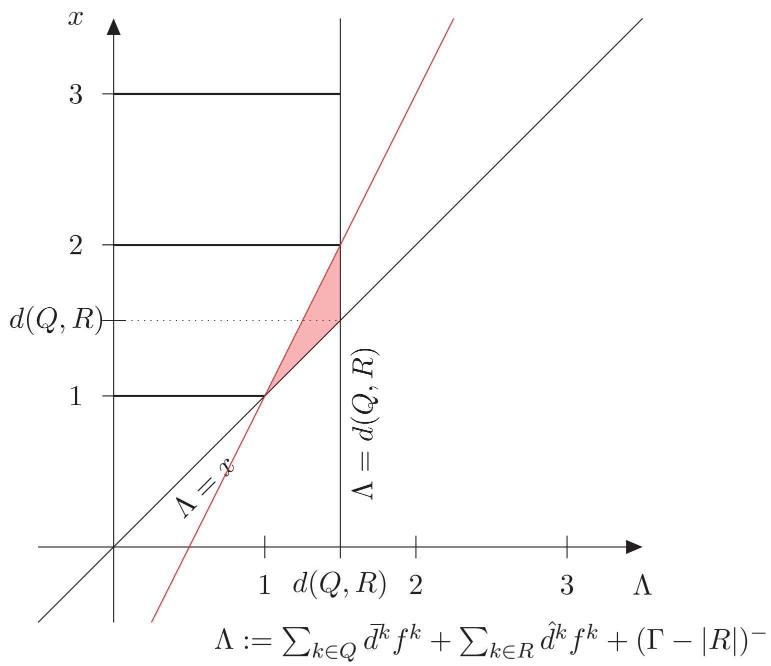


Figure 9.3: Example of Γ-robust arc residual capacity inequality. On the x-axis the value of the term $\sum_{k \in Q} \bar{d}^k f^k + \sum_{k \in R} \hat{d}^k f^k + (\Gamma - |R|)^-$ (or Λ for short) is shown. On the y-axis the integer value of x is shown. This yields the mixed integer set illustrated by thick black lines. The Γ-robust arc residual capacity inequality (red) cuts-off the reddish shaded area and thus tightens the convex hull of the feasible set.

For a violated Γ-robust arc residual capacity inequality, $\lceil d(Q,R) \rceil = \lceil x^* \rceil$ and $r^{Q,R} = d(Q,R) - \lfloor x^* \rfloor$ hold by Lemma 9.39. Hence, we can determine the violation as

$$\begin{aligned}
&\text{violation}(Q,R) \\
=& \sum_{k \in Q} \bar{d}^k f^{*k} + \sum_{k \in R} \hat{d}^k f^{*k} + (\Gamma - |R|)^- \pi^* - d(Q,R)(x^* - \lfloor x^* \rfloor) - \lfloor x^* \rfloor (\lceil x^* \rceil - x^*) \\
=& \sum_{k \in Q} \bar{d}^k (f^{*k} - (x^* - \lfloor x^* \rfloor)) + \sum_{k \in R} \hat{d}^k (f^{*k} - (x^* - \lfloor x^* \rfloor)) \\
&+ \underbrace{(\Gamma - |R|)^- \pi^* - \lfloor x^* \rfloor (\lceil x^* \rceil - x^*)}_{<0}.
\end{aligned}$$

We observe that a positive violation can only be achieved by the contribution of commodities k with $f^{*k} - (x^* - \lfloor x^* \rfloor) > 0$.

Lemma 9.40. *Given a solution $(x^*, f^*, \pi^*, \rho^*)$ of ${}^{LP}\mathcal{S}^{LF,\Gamma}$, a Γ-robust arc residual capacity inequality* (9.40) *with corresponding sets Q, R and violation* violation(Q,R)*. Let $k^* \in Q$ with $f^{*k^*} \leq (x^* - \lfloor x^* \rfloor)$.*

Then violation$(Q,R) \leq$ violation$(Q \setminus \{k^*\}, R \setminus \{k^*\})$*, i. e., the violation does not increase by including k^*.*

Proof. If $k \notin R$, then holds

$$\begin{aligned}&\text{violation}(Q \setminus \{k^*\}, R \setminus \{k^*\}) = \text{violation}(Q \setminus \{k^*\}, R)\\&= \text{violation}(Q, R) - \bar{d}_{k'}(f^{*k^*} - (x^* - \lfloor x^* \rfloor)) > \text{violation}(Q, R),\end{aligned}$$

since $f^{*k^*} \leq (x^* - \lfloor x^* \rfloor)$ by assumption. □

Based on these two observations, we are now able to describe an efficient exact separation algorithm. Let us introduce the following notation $\tau_k := \hat{d}^k(f^{*k} - (x^* - \lfloor x^* \rfloor)) + \pi^*$. The presented algorithm is a greedy algorithm. First, it assigns all commodities k with $f^k > x^* - \lfloor x^* \rfloor$ to Q as those commodities contribute to the violation with positive coefficients. Second, the subset $R \subseteq Q$ is determined. If $|Q| \leq \Gamma$, $R := Q$ is set since this contributes most and the term with π^* is zero. If $|Q| > \Gamma$, the set R is determined greedily by adding all commodities k with $\tau_k < 0$ as those are the commodities adding a nonnegative value to the violation when taking the term $(\Gamma - |R|)^- \pi^*$ into account. Algorithm 2 states the described procedure in pseudo-code.

Algorithm 2: Separation of Γ-robust arc residual capacity inequalities

Input : solution $(x^*, f^*, \pi^*, \rho^*)$ of ${}^{\text{LP}}\mathcal{S}^{\text{LF},\Gamma}$
Output : sets Q, R defining the most-violated Γ-robust arc residual capacity inequality or $\emptyset$ if none exists

set $Q = \left\{k \in K \ : \ f^{*k} > x^* - \lfloor x^* \rfloor\right\}$;
if $|Q| \leq \Gamma$ **then**
  set $R = Q$;
if $|Q| > \Gamma$ **then**
  sort Q non-decreasingly with ordering $\varphi : K \to \mathbb{N}$ such that $\tau_{k_1} \leq \tau_{k_2}$ holds for all $k_1, k_2 \in Q$ with $\varphi(k_1) < \varphi(k_2)$;
  set $\varphi_0 := \max\{i = 1, \ldots, |Q| \ : \ \tau_{\varphi^{-1}}(i) < 0\}$;
  set $R = \{k \in Q \ : \ \varphi(k) \leq \max\{\varphi_0, \Gamma\}\}$;
if *violation*$(Q, R) > 0$ **then**
  return Q, R ;
return $\emptyset$;

Lemma 9.41. *Algorithm 2 is an exact separation algorithm.*

Proof. Notice, the algorithm returns $\emptyset$ if and only if $\lfloor x^* \rfloor < d(Q, R) < \lceil x^* \rceil$ does not hold or *violation*$(Q, R) \leq 0$. By Lemma (9.39), the first condition is necessary for a violated cut. Therefore, it is correct that the algorithm returns $\emptyset$ if no violated cut exists.

To show that the algorithm is exact it is sufficient to show that the algorithm always determines subsets Q and R for which the violation of the corresponding Γ-robust arc residual capacity inequality is maximum.

Let Q and R be determined by Algorithm 2. Suppose there exist sets $R^* \subseteq Q^* \subseteq K$ with $Q^* \neq Q$ or $R^* \neq R$ defining a maximum violated Γ-robust arc residual capacity inequality with violation$(Q^*, R^*) >$ violation(Q, R).

If $Q \not\subseteq Q^*$, then there exists $k \in Q \setminus Q^*$ with $\bar{d}^k(f^{*k} - (x^* - \lfloor x^* \rfloor)) > 0$ and thus violation$(Q^* \cup \{k\}, R^*) >$ violation(Q^*, R^*); a contradiction to the maximum violation of the inequality defined by Q^*, R^*. Hence, $Q \subseteq Q^*$.

If $Q^* \not\subseteq Q$, then there exists $k \in Q^* \setminus Q$ with either $f^{*k} = x^* - \lfloor x^* \rfloor$ or $f^{*k} < x^* - \lfloor x^* \rfloor$. If $f^{*k} = x^* - \lfloor x^* \rfloor$ holds, the coefficient $\bar{d}^k(f^{*k} - (x^* - \lfloor x^* \rfloor))$ is zero and $Q^* \setminus \{k\}$ and Q^* yield the same Γ-robust arc residual capacity inequality for a fixed R^* and $k \notin R^*$. If $k \in R^*$, violation$(Q^* \setminus \{k\}, R^* \setminus \{k\}) \geq$ violation(Q^*, R^*). Hence, w. l. o. g. let $f^{*k} < x^* - \lfloor x^* \rfloor$ hold. Then violation$(Q^* \setminus \{k\}, R^*) >$ violation(Q^*, R^*) by Lemma 9.40; a contradiction. So $Q^* = Q$ holds.

Next we consider R^*. If $R \not\subseteq R^*$, then there exists $k \in R \setminus R^*$ with $\tau_k < 0$ and violation$(Q^*, R^* \cup \{k\}) >$ violation(Q^*, R^*) holds; a contradiction. Hence, $R \subseteq R^*$.

If $R^* \not\subseteq R$, then there exists $k \in R^* \setminus R$. By assumption, Q^* and R^* define a maximum violated inequality. Therefore, violation$(Q^*, R^*) \geq$ violation$(Q^*, R^* \setminus \{k\})$ holds. If the last relation hold with equality, R^* can be set to $R \setminus \{k\}$ without reducing the violation. By iterating this step, we either construct a set $R^* = R$ a contradiction and thus implying $R^* = R$, or we end up with a set R^* and commodity k with violation$(Q^*, R^*) >$ violation$(Q^*, R^* \setminus \{k\})$. In this case violation$(Q^*, R^*) =$ violation$(Q^*, R^* \setminus \{k\}) + \tau_k$ with positive τ_k. By definition, this implies $k \in R$; a contradiction. Hence, $R^* = R$. This completes the proof. □

The complexity of Algorithm 2 is dominated by the sorting of all commodities $k \in Q$. In addition, each commodity has to be checked twice: once when setting Q and once for R. This gives a total worst-case complexity of $\mathcal{O}(|K| \log |K|)$.

9.3.3 Separation of Γ-robust metric inequalities

In this section, we investigate the separation of RMIs. First, we present a polynomial time exact separation algorithm to separate RMIs as model inequalities. Second, we present an exact separation procedure to determine violated rounded RMIs in a cut-and-branch approach for solving the Γ-RNDP. Third, we address tight RMIs pointing out the differences to the non-robust setting.

The separation algorithms presented in this section separate violated RLIs in a first step. If the length function ℓ is not metric, a metric ℓ_M and hence a corresponding violated metric inequality can be constructed by $\ell_M(e) := \min\{\ell(e), \ell(u,v)\}$ for all $e = uv \in E$ (cf. Theorem 9.33). Note, the right-hand side of the RLI does not change and the violation of the RMI is at least the violation of the previous RLI.

Γ-robust metric inequalities. In the following, we describe the exact separation of violated RMIs analogously to the non-robust case considered by Avella et al. [17] and analogously to the Γ-RNDP with dynamic routing by Mattia [121]. Therefore, we define $f_e^k := \sum_{p \in P^k : e \in p} f^k(p)$ as the flow on edge $e \in E$ for commodity $k \in K$. Given a capacity vector $\tilde{x} \in \mathbb{R}_{\geq 0}^{|E|}$, the feasibility problem can be formulated as

$$\max \alpha \tag{9.48a}$$

$$\text{s.t.} \sum_{uv \in E} (f^k_{uv} - f^k_{vu}) \geq \begin{cases} \alpha & u = s^k \\ -\alpha & u = t^k \\ 0 & \text{otherwise} \end{cases} \qquad \forall u \in V,\ k \in K \tag{9.48b}$$

$$\sum_{k \in K} \bar{d}^k f^k_e + \Gamma \pi_e + \sum_{k \in K} \rho^k_e \leq \tilde{x}_e \qquad \forall e \in E \tag{9.48c}$$

$$\hat{d}^k f^k_e - \pi_e - \rho^k_e \leq 0 \qquad \forall e \in E,\ k \in K \tag{9.48d}$$

$$f^k_{uv}, \pi_e,\ \rho^k_e,\ \alpha \geq 0 \qquad \forall e = uv \in E,\ k \in K. \tag{9.48e}$$

The value of an optimal solution $(\alpha^*, f^*, \pi^*, \rho^*)$ of (9.48) is at least 1 if and only if the point $(\tilde{x}, f^*, \pi^*, \rho^*)$ is feasible for ${}^{\text{LP}}\mathcal{N}^{\Gamma}_x$. The dual of (9.48) reads

$$\min \sum_{e \in E} \tilde{x}_e \ell(e) \tag{9.49a}$$

$$\text{s.t. } \beta^k_u - \beta^k_v \leq \bar{d}^k \ell(e) + \hat{d}^k m^k(e) \qquad \forall e = uv \in E, k \in K \tag{9.49b}$$

$$\sum_{k \in K} m^k(e) \leq \Gamma \ell(e) \qquad \forall e \in E \tag{9.49c}$$

$$m^k(e) \leq \ell(e) \qquad \forall e \in E,\ k \in K \tag{9.49d}$$

$$\sum_{k \in K} (\beta^k_{s^k} - \beta^k_{t^k}) \geq 1 \tag{9.49e}$$

$$\beta^k_u,\ \ell(e),\ m^k(e) \geq 0 \qquad \forall u \in V, e \in E,\ k \in K. \tag{9.49f}$$

By strong duality, $\tilde{x}$ is feasible for ${}^{\text{LP}}\mathcal{N}^{\Gamma}_x$ if and only if the optimal value of (9.49) is at least 1. If the objective value of a solution (β^*, ℓ^*, m^*) is strictly less than 1, constraint (9.49e) implies the RLI

$$\sum_{e \in E} \ell^*(e) x_e \geq \sum_{k \in K} ({\beta^*}^k_{s^k} - {\beta^*}^k_{t^k}) \tag{9.50}$$

to be violated for $\tilde{x}$.

Corollary 9.42. *Violated RMIs can be separated in polynomial time.*

Proof. Since (9.49) is a pure LP with polynomial size and ℓ_M is constructable from ℓ in polynomial time, violated RMIs can be exactly separated in polynomial time. □

Despite this result, experiments have shown that the separation of these inequalities can be very time consuming in practice; cf. Section 11.7.

Rounded Γ-robust metric inequalities. To cut-off a fractional solution of the LP relaxation of the Γ-robust network design polyhedron, the right-hand side of a corresponding RLI (with integer lengths) has to be rounded up to yield a violation; cf. derivation of (9.45). Hence, formulation (9.49) has to be modified to take the integrality of ℓ and the rounding into account. The exact separation of rounded RLIs can be formulated as the following ILP.

$$\max \phi - \sum_{e\in E} \tilde{x}_e \ell(e) \tag{9.51a}$$

$$\text{s.t.}(9.49b) - (9.49d)$$

$$\sum_{k\in K} (\beta^k_{t^k} - \beta^k_{s^k}) + \phi \leq 1 - \varepsilon \tag{9.51b}$$

$$\phi \in \mathbb{Z} \tag{9.51c}$$

$$\ell(e) \in \mathbb{Z}_{\geq 0} \qquad \forall e \in E \tag{9.51d}$$

$$\beta^k_u,\ m^k(e) \geq 0 \qquad \forall u \in V, e \in E,\ k \in K, \tag{9.51e}$$

where, given a small constant $\varepsilon > 0$, constraint (9.51b) determines the rounded up right-hand side value ϕ of the resulting RLI. The objective value of (9.51) equals the violation of the rounded RLI. Hence, a nonpositive objective value gives a proof that no such inequality exists.

Given a violated rounded RLI, a corresponding violated rounded RMI (9.45) can be constructed as described above.

Bounded Γ-robust metric inequalities. Bounded RMIs can be separated by solving formulation (9.51), where constraint (9.51d) is replaced by

$$\ell(e) \in \{0, 1, .., L\}, \tag{9.51d'}$$

and strengthening the obtained RLI to a RMI as before.

Tight Γ-robust metric inequalities. The strongest RMIs are those with a tight right-hand side; cf. Corollary 9.37. By definition, a minimization problem over the Γ-robust network loading polyhedron has to be solved to determine the best right-hand side β_{ℓ_M} to obtain a tight RMI. Clearly, this is as hard as solving the original Γ-RNDP. In this paragraph, we present a preprocessing method to speed up the computation of β_{ℓ_M} in practice.

Avella et al. [17] suggest a shrinking heuristic to reduce the network size before determining the best right-hand side β_{ℓ_M} given a ℓ_M in the non-robust setting. In the following, we describe a version of this shrinking that can be used in the Γ-robust setting.

Theorem 9.43. *Let $\ell_M \in \text{Met}(G)$ and $\{i,j\} \in E$ with $\ell_M(ij) = 0$. Define the shrunken graph $G^h = (V^h, E^h)$ by*

$$V^h := V \setminus \{i,j\} \cup \{h\}$$
$$E^h := E \setminus \{uv \in E \ :\ u \in \{i,j\} \vee v \in \{i,j\}\} \cup \{uh \ :\ ui \in E \vee uj \in E\}.$$

Then $\beta^h_{\ell_M} = \beta_{\ell_M}$.

Proof. The proof is analog to the proof of Theorem 3.3 in [17].

Suppose $\beta_{\ell_M} < \beta^h_{\ell_M}$. Let x satisfy $\ell_M x = \beta_{\ell_M}$ and x^h be the mapping of x on G^h. Then $\beta^h_{\ell_M} \leq \ell^h_M x^h = \ell_M x = \beta_{\ell_M} < \beta^h_{\ell_M}$ holds; a contradiction.

Suppose $\beta_{\ell_M} > \beta^h_{\ell_M}$. Let x^h satisfy $\ell^h_M x^h = \beta^h_{\ell_M}$. Define x as follows.

$$x_e := \left\lceil \sum_{k \in K} \bar{d}^k + \max_{Q \subseteq K, |Q| \leq \Gamma} \hat{d}^k, \right\rceil$$

$x_{uh} = x^h_{ui} + x^h_{uj}$ for all $u \in V \setminus \{i, j\}$, and $x_{uv} := x^h_{uv}$ for all $uv \in E, u, v \in V \setminus \{i, j\}$. Then x is feasible for $\mathcal{N}^\Gamma_x$ and $\ell_M x = \ell^h_M x^h$ holds by construction. Further, it holds $\ell_M x = \ell^h_M x^h = \beta^h_{\ell_M} < \beta_{\ell_M}$; a contradiction to the minimality of β_{ℓ_M}. □

In contrast to the non-robust setting [17], commodities cannot be aggregated in the shrinking procedure for the Γ-robust setting. Instead, the shrunken graph may have "parallel" commodities with same source and destination nodes corresponding to different nodes in the original graph. The following example illustrates the difference to the non-robust setting.

Example 9.44. *Consider the complete graph G with nodes $V := \{n_1, n_2, n_3\}$, edges $E := \{n_1n_2, n_1n_3, n_2n_3\}$, commodities K (with $|K| = 2$), and $\Gamma = 1$. Further, let $\ell_M(n_1n_2) = 0$, $\ell_M(n_1n_3) = \ell_M(n_2n_3) = 1$ be a metric. In the Γ-robust setting, Γ many commodities may deviate on each edge in this network. Applying the shrinking heuristic of [17] results in the reduced graph G^h with nodes $V^h := \{h, n_3\}$, edges $E := \{hn_3\}$, and new metric $\ell^h_M(hn_3) = 1$. In contrast to the non-robust setting, the commodities cannot be aggregated since the deviations of a single aggregated commodity from h to n_3 cannot reflect the independent demand deviations on the two edges n_1n_3 and n_2n_3. Furthermore, this cannot be overcome by increasing the Γ for G^h as the resulting solution might violate the limit of $\Gamma = 1$ on the original edges n_1n_3 and n_2n_3.*

Theorem 9.43 can be exploited to obtain a processing procedure shrinking the graph. This might increase the computational tractability of the problem to determine the right-hand side of a tight RMI in practice.

CHAPTER TEN

THE MULTI-BAND ROBUST NETWORK DESIGN PROBLEM

Network design problems are well-studied and occurring in many real-world applications, e. g., in the field of telecommunications, public transport or logistics. Due to limitations in the computational tractability, data uncertainty has mostly been ignored. Robust network design problems under data uncertainty have come to the center of attention with the rising popularity of the Γ-robustness concept. Following this concept, data uncertainty is characterized by two values only: a nominal data value and a maximal (symmetrical) deviation value. This simplification of the unknown underlying probability variable is rather rough. In fact, Γ-robustness is a special case of the more general multi-band robustness concept, which characterizes the uncertain data by a nominal value and a set of deviation bands each with its own deviation value and further bounds on the number of realizations in this band. This concept allows to model a discrete distribution in a histogram-like fashion capturing the unknown underlying probability distribution in more detail than the Γ-robustness concept. A survey on network design under data uncertainty and Γ-RNDP in particular is given in the introduction of Chapter 9. To our knowledge the concept of multi-band robustness had not been applied to the network design problem in general. Only a simplified preliminary version of multi-band robustness has been considered in the planning of wireless networks by Bienstock and D'Andreagiovanni [38] and D'Andreagiovanni [58]. Recently, Mattia [122] published a technical report about a network design problem with multiple intervals which follows a robustness concept very similar to multi-band robustness. However, her concept is less general: it assumes symmetrical random variables and does not have lower bounds on the number of realizations. In particular the latter implies a higher conservatism. In her technical report, she gives link-flow ILP formulations of the network design problem with multiple intervals, analyzes the probability of constraint satisfaction w.r.t. to uncertainty and given robustness parameter settings. Further, she considers solving the problems by a robustness cuts approach. Neither a capacity formulation of the problem, nor the polyhedral structure of the problem (including valid inequalities) are studied.

In this chapter, we introduce the *multi-band robust network design problem (mb-RNDP)* generalizing the Γ-RNDP to the multi-band robust setting. Therefore let us introduce some notation: like the Γ-RNDP this problem is defined on an undirected graph $G = (V, E)$, has link capacity installment costs κ_e for all $e \in E$, and commodities $k \in K$

with sources s^k, targets t^k, and uncertain demand values $d \in \mathcal{U}^{\text{mb}}$. Following the concept of multi-band robustness, the uncertain demand of commodity $k \in K$ is specified by its nominal demand value $\bar{d}^k$, and deviation bands $B = \{\underline{B}, \underline{B}+1, \ldots, 0, \ldots, \overline{B}-1, \overline{B}\}$ each with an associated deviation demand value $\hat{d}^{b,k}$, lower bounds $\gamma^b \geq 0$ and upper bounds $\Gamma^b \leq |K|$ on the number of realizations in band $b \in B$. The deviation demand values of a commodity $k \in K$ are required to be increasing, i. e., $\hat{d}^{\underline{B},k} < \hat{d}^{\underline{B}+1,k} < \cdots < \hat{d}^{0,k} < \cdots < \hat{d}^{\overline{B}-1,k} < \hat{d}^{\overline{B},k}$, and to include all assumed realizations, i. e., $d^k \in [\bar{d} + \hat{d}^{\underline{B},k}, \bar{d} + \hat{d}^{\overline{B},k}]$. Note that $\hat{d}^{0,k} = 0$. Furthermore, to be feasible $\sum_{b \in B} \gamma^b \leq |K|$, $\gamma^0 = 0$, and $\Gamma^0 = |K|$ are required.

The traffic for commodity $k \in K$ is realized by a multi-commodity flow. Analogously to the Γ-RNDP, we focus on fixed routing templates and thus static or oblivious routing in this thesis.

Definition 10.1 (Multi-Band Robust Network Design Problem)**.** Given a potential network topology, a multi-band robust uncertainty set $\mathcal{U}^{\text{mb}}$ of the demand, and installation costs as described above.

The *multi-band robust network design problem (mb-RNDP)* with oblivious/static, splittable routing is to find a minimum-cost installation of integral capacities and a routing template for every commodity such that the actual flow does not exceed the link capacities independent of the realization of demands in $\mathcal{U}^{\text{mb}}$; cf. Definition 3.4 of $\mathcal{U}^{\text{mb}}$.

Remark 10.2 (Generalization of the Γ-RNDP)**.** As mentioned in the introduction, the Γ-RNDP is a special case of the mb-RNDP with only the nominal and one deviation band. The explicit construction is as follows. Let an instance of the Γ-RNDP be given by a graph $G = (V, E)$, a link capacity installment cost function $\kappa : E \to \mathbb{R}_{\geq 0}$, a commodity set K with nominal demand values $\bar{d}^k$ and deviation demand values $\hat{d}^k$ for every commodity $k \in K$, and a robustness parameter Γ. Then the corresponding mb-RNDP is given by the same graph G, same cost function κ, same set K of commodities with nominal demand values $\bar{d}^k$ and deviation demand values $\hat{d}^{0,k} = 0$ and $\hat{d}^{1,k} = \hat{d}^k$ for all $k \in K$, bounds $\gamma^0 = \gamma^1 = 0$ and $\Gamma^0 = |K|$ and $\Gamma^1 = \Gamma$, and two bands $B := \{0, 1\}$.

Furthermore, the results of our investigation of the Γ-RNDP can be generalized to the multi-band robust setting. Although these generalizations require more notations to capture the additional uncertainty bands, the polyhedra associated to the mb-RNDP can be investigated quite similarly and often straightforward generalizations to the multi-band robust setting can be obtained. Therefore, we demonstrate these generalizations not as detailed as our investigation of the Γ-RNDP and limit ourselves to cutset-based inequalities when investigating classes of non-trivial valid inequalities. Nevertheless, we conjecture that our further results on Γ-robust arc residual capacity and Γ-robust metric inequalities can be generalized for the mb-RNDP in an analogous way.

10.1 Formulations

In this section, we present formulations for the mb-RNDP. As for the classic network design problem NDP and its Γ-robust counterpart, modeling alternatives exist: link- and path-flow based formulations and a capacity formulation.

Exponential flow formulations of the mb-RNDP Following the multi-band robustness concept, we can formulate the multi-band robust counterpart of formulation (1.19) as

$$\begin{aligned}
&(1.19a), (1.19b) - (1.19e) \\
&\sum_{k \in K} d^k f_e^k \leq x_e \qquad \forall e \in E, d \in \mathcal{U}^{\text{mb}}. \qquad (10.1a)
\end{aligned}$$

The multi-band robust counterpart of formulation (1.20) reads

$$\begin{aligned}
&(1.20a), (1.20b) - (1.20e) \\
&\sum_{k \in K} d^k \left(\sum_{p \in P^k : e \in p} f_p^k \right) \leq x_e \qquad \forall e \in E, d \in \mathcal{U}^{\text{mb}}. \qquad (10.2a)
\end{aligned}$$

Although infinitely many realizations may exists in $\mathcal{U}^{\text{mb}}$ only exponentially many are non-dominated as described in Section 3.2.2. Only those non-dominated realizations contribute to the worst-case left-hand side value and thus to the feasibility of the constraints (10.1a) and (10.2a), respectively. Hence, for both formulations (10.1) and (10.2) finite but exponentially sized reformulations exist.

Compact flow formulations of the mb-RNDP Compact ILP reformulations of the exponential link-flow (10.1) and path-flow formulation (10.2) can be obtained by exploiting LP duality; cf. Section 3.2.2. In this process the contribution of worst-case realization to the link capacity constraint is determined. In Lemma 3.7 we have shown that the number of data realizations in each band for a worst-case realization can be determined a-priori. This has led us to the (frequency) profile of a multi-band robust problem and the coefficients ϑ^b denoting the number of realizations in band b of a worst-case realization. Taking this consideration into account, the resulting *compact ILP link-flow formulation* of the mb-RNDP is

$$\begin{aligned}
&(1.19a), (1.19b), (1.19d), (1.19e) \\
&\sum_{k \in K} \bar{d}^k f_e^k + \sum_{b \in B} \vartheta^b \pi_e^b + \sum_{k \in K} \sigma_e^k \leq x_e, \qquad \forall e \in E \qquad (10.3a) \\
&- \pi_e^b - \sigma_e^k + \hat{d}^{b,k} f_e^k \leq 0 \qquad \forall e \in E, k \in K, b \in B \qquad (10.3b) \\
&\pi \geq 0 \qquad (10.3c)
\end{aligned}$$

with dual variables π and σ. Note that σ is a free variable, i. e., no lower/upper bounds on its value exist, especially no nonnegativity constraint.

The corresponding *compact ILP path-flow formulation* of the mb-RNDP reads

$$
\begin{aligned}
& (1.20a), (1.20b), (1.20d), (1.20e) \\
& \sum_{k \in K} \bar{d}^k \left(\sum_{p \in P^k : e \in p} f_p^k \right) + \sum_{b \in B} \vartheta^b \pi_e^b + \sum_{k \in K} \sigma_e^k \leq x_e \qquad \forall e \in E && (10.4a) \\
& -\pi_e^b - \sigma_e^k + \hat{d}^{b,k} \sum_{p \in P^k : e \in p} f_p^k \leq 0 \qquad \forall e \in E, k \in K, b \in B && (10.4b) \\
& \pi \geq 0 && (10.4c)
\end{aligned}
$$

with dual variables π and σ.

In both formulations, we apply Lemma 3.7 to determine the number of realizations in each uncertainty band in the worst-case. In this way the subproblem determining the worst-case is formulated more efficiently and thus, the number of dual variables in the compact reformulation is reduced.

Capacity formulation of the mb-RNDP Similar to the classic NDP, a capacity formulation of the mb-RNDP can be derived. Therefore, we consider the link capacities $x \in \mathbb{Z}_{\geq 0}$ as given. Then, the existence of a feasible flow satisfying the constraints (1.20b), (1.20d), and (10.4a)–(10.4c) can be characterized by applying Farkas' lemma as follows. Notice, that this result also holds for fractional capacities.

Lemma 10.3. *Given* $\tilde{x} \in \mathbb{R}_{\geq 0}^{|E|}$*, there exists a flow satisfying* (1.20b), (1.20d)*, and* (9.7a)–(9.7c) *if and only if for all lengths functions* ℓ

$$\sum_{e \in E} \tilde{x}_e \ell(e) \geq z_\ell \tag{10.5}$$

holds, where z_ℓ *is defined by the following LP.*

$$
\begin{aligned}
z_\ell := \quad & \max \sum_{k \in K} z^k && (10.6a) \\
& s.t.\ z^k - \sum_{b \in B} \sum_{e \in p} \hat{d}^{b,k} m^{b,k}(e) \leq \bar{d}^k \ell(p) \qquad \forall k \in K,\ p \in P^k && (10.6b) \\
& \sum_{k \in K} m^{b,k}(e) \leq \vartheta^b \ell(e) \qquad \forall e \in E, b \in B && (10.6c) \\
& \sum_{b \in B} m^{b,k}(e) = \ell(e) \qquad \forall k \in K,\ e \in E && (10.6d) \\
& z^k,\ m^{b,k}(e) \geq 0 \qquad \forall k \in K,\ e \in E && (10.6e)
\end{aligned}
$$

We call (10.5) *the* multi-band robust length inequality (RLI)*, and if* ℓ *is metric,* (10.5) *is called a* multi-band robust metric inequality (RMI)*.*

Proof. For fixed link capacities $\tilde{x}$, formulation (10.4) is an LP. Corresponding to constraints (1.20b), (10.4a) and (10.4b), we introduce the dual variables z^k, $\ell(e)$ and $m^{b,k}(e)$ for all $e \in E, k \in K, b \in B$, respectively. The application of Farkas' lemma yields the conditions (10.6b)–(10.6e) and $\ell(e) \geq 0$. Therefore, there exists a feasible flow satisfying (1.20b), (1.20d), and (9.7a)–(9.7c) if and only if

$$\sum_{e \in E} \tilde{x}_e \ell(e) \geq z_\ell$$

is valid for all $z^k, \ell(e), m^{b,k}(e) \geq 0$ satisfying (10.6b)–(10.6e). For all length functions ℓ, this implies that the multi-band robust length inequality (10.5) is valid for the convex hull of all feasible solutions of the linear relaxation of (10.4) if constraints (10.6b)–(10.6e) are fulfilled. To determine the strongest multi-band robust length inequality for a given length function ℓ, the sum $\sum_{k \in K} z^k$ has to be maximized. This can be formulated as LP (10.6) completing the proof. □

This lemma generalizes the "Japanese Theorem" 1.21 to the multi-band robust setting. A special case of Lemma 10.3 is given by Lemma 9.2 for the Γ-RNDP.

Now, we formulate the mb-RNDP in the space of the capacity variables as follows

$$\begin{aligned} & (1.20\text{a}) \\ & \sum_{e \in E} \ell(e) x_e \geq z_\ell \qquad \forall \ell \in \mathcal{L} \qquad (10.7\text{a}) \\ & x \in \mathbb{Z}_{\geq 0}^{|E|}. \qquad (10.7\text{b}) \end{aligned}$$

where $\mathcal{L}$ denotes the set of all length functions $\ell : E \to \mathbb{R}_{\geq 0}$. We call ILP (10.7) the *capacity formulation* of the mb-RNDP.

10.2 Polyhedral study

In this section, we define the multi-band robust network design polyhedra and study their polyhedral structure. We emphasize the fact that multi-band robustness is a generalization of the well-studied Γ-robustness concept. Therefore, we do not present a polyhedral study as detailed as for the Γ-RNDP (cf. Section 9.2) but study valid inequalities for the generalized polyhedra using the example of cutset-based inequalities.

Definition 10.4 (Multi-Band Robust Network Design Flow Polyhedra)**.** We define the *multi-band robust network design link-flow polyhedron* $\mathcal{N}^{\text{LF,mb}}$ as the convex hull of all feasible solutions of the compact link-flow formulation (10.3) of the mb-RNDP, i. e.,

$$\mathcal{N}^{\text{LF,mb}} := \text{conv} \left\{ \begin{array}{r} (x, f, \pi, \sigma) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{2|E||K|} \times \mathbb{R}_{\geq 0}^{|E||B|} \times \mathbb{R}^{|E||K|} \;: \\ (x, f, \pi, \sigma) \text{ satisfies } (10.3) \end{array} \right\}. \qquad (10.8)$$

We denote by $\mathcal{N}_{x,\pi}^{\text{LF,mb}}$ the projection of $\mathcal{N}^{\text{LF,mb}}$ onto the space of the x and π variables, i. e.,

$$\mathcal{N}_{x,\pi}^{\text{LF,mb}} = \text{proj}_{x,\pi}\,\mathcal{N}^{\text{LF,mb}}, \tag{10.9}$$

and by $\mathcal{N}_{x}^{\text{LF,mb}}$ the projection of $\mathcal{N}^{\text{LF,mb}}$ onto the space of the x variables, i. e.,

$$\mathcal{N}_{x}^{\text{LF,mb}} = \text{proj}_{x}\,\mathcal{N}^{\text{LF,mb}}. \tag{10.10}$$

Similarly, we define the *multi-band robust network design path-flow polyhedron* $\mathcal{N}^{\text{PF,mb}}$ as the convex hull of all feasible solutions of the compact path-flow formulation (10.4) of the mb-RNDP, i. e.,

$$\mathcal{N}^{\text{PF,mb}} := \text{conv}\left\{ \begin{aligned} &(x,f,\pi,\sigma) \in \mathbb{Z}_{\geq 0}^{|E|} \times \mathbb{R}_{\geq 0}^{|P|} \times \mathbb{R}_{\geq 0}^{|E||B|} \times \mathbb{R}^{|E||K|} \ : \\ &\qquad\qquad\qquad\qquad (x,f,\pi,\sigma) \text{ satisfies } (10.4) \end{aligned} \right\}. \tag{10.11}$$

We denote by $\mathcal{N}_{x}^{\text{PF,mb}}$ the projection of $\mathcal{N}^{\text{PF,mb}}$ onto the space of the x variables, i. e.,

$$\mathcal{N}_{x}^{\text{PF,mb}} = \text{proj}_{x}\,\mathcal{N}^{\text{PF,mb}}. \tag{10.12}$$

Definition 10.5 (Multi-Band Robust Network Design Capacity Polyhedron)**.** We define the *multi-band robust network design capacity polyhedron* $\mathcal{N}_{x}^{\text{mb}}$ as the convex hull of all feasible solutions of the capacity formulation (10.7) of the mb-RNDP, i. e.,

$$\mathcal{N}_{x}^{\text{mb}} := \text{conv}\left\{ x \in \mathbb{Z}_{\geq 0}^{|E|} \ : \ x \text{ satisfies } (10.7) \right\}. \tag{10.13}$$

By construction of the capacity formulation (10.7), the following corollary holds.

Corollary 10.6. $\mathcal{N}_{x}^{mb} = \mathcal{N}_{x}^{LF,mb} = \mathcal{N}_{x}^{PF,mb}$.

10.2.1 Basic characteristics

In this section, we briefly report on the dimension of the multi-band robust link-flow and capacity polyhedra.

Lemma 10.7. *The dimension of* $\mathcal{N}^{LF,mb}$ *equals* $3|E||K| + |E||B| + |E| - (|V|-1)|K|$ *whereas* $\mathcal{N}_{x,\pi}^{LF,mb}$ *is full-dimensional.*

Proof. The proof is analog to the proof of Lemma 9.6 for the Γ-RNDP. □

Lemma 10.8. *The polyhedron* $\mathcal{N}_{x}^{mb}$ *is full-dimensional, i. e.,* $\dim(\mathcal{N}_{x}^{mb}) = |E|$.

Proof. Let $d^{max} := \lceil \max_{d \in \mathcal{U}^{\text{mb}}} \sum_{k \in K} d^k \rceil$ be the maximum total demand. Define v_0 as the one vector multiplied by d^{max}. Then, v_0 and $(v_0 + e_e)_{e \in E}$ are $|E|+1$ affinely independent vectors in $\mathcal{N}_{x}^{\text{mb}}$. □

10.2.2 Cutset-based inequalities

In the past, cutset-based inequalities have been proven to be effective in a branch-and-cut approach to solve instances of network design problems, both in the classic non-robust and the Γ-robust setting. Therefore, we focus our investigation of the polyhedral structure of the mb-RNDP on these types of inequalities.

Multi-band robust cutset polyhedron. Analogously to the Γ-RNDP, we consider a non-empty subset $S \subsetneq V$ of the nodes, its corresponding cutset $\delta(S) \subseteq E$, and the induced set of cut-crossing commodities $Q_S \subseteq K$. We assume w. l. o. g. that $s^k \in S$ for all $k \in Q_S$.

As in the classic or Γ-robust setting, the capacities installed on the cutset edges must be large enough to support the worst-case demand realization in $\mathcal{U}^{\text{mb}}$ of all cut-crossing cuts Q_S. Therefore the *multi-band robust cutset inequality*

$$x(\delta(S)) \geq \lceil d_\delta^{\max}(S) \rceil \tag{10.14}$$

must hold where $d_\delta^{\max}(S)$ denotes the worst-case total demand value crossing the cut. In contrast to the Γ-robust uncertainty set $\mathcal{U}^\Gamma$ no closed formula is known for all terms and sets needed to determine $d_\delta^{\max}(S)$. Instead Büsing and D'Andreagiovanni [47] give a polynomial-time combinatorial algorithm to determine $d_\delta^{\max}(S)$ by solving a min-cost flow problem on an auxiliary graph (in fact, the underlying assignment problem of assigning commodities to deviation bands), cf. Lemma 3.9.

In the following we show the validity of the multi-band robust cutset inequality (10.14) by deriving it from valid model constraints. Moreover, we define valid inequalities which give rise to generalizations of the envelope inequalities for the Γ-RNDP.

Let us begin by restricting the mb-RNDP to a cutset $\delta(S)$: contracting both shores of the cut $\delta(S)$, we consider the resulting multi-band robust two-node formulation corresponding to (10.3):

$$\sum_{\{i,j\}\in\delta(S)} (f_{ij}^k - f_{ji}^k) = 1 \qquad \forall k \in Q_S \tag{10.15a}$$

$$\sum_{\{i,j\}\in\delta(S)} (f_{ij}^k - f_{ji}^k) = 0 \qquad \forall k \in K \setminus Q_S \tag{10.15b}$$

$$\sum_{k\in K} \bar{d}^k f_e^k + \sum_{b\in B} \vartheta^b \pi_e^b + \sum_{k\in K} \sigma_e^k \leq x_e, \qquad \forall e \in \delta(S) \tag{10.15c}$$

$$-\pi_e^b - \sigma_e^k + \hat{d}^{b,k} f_e^k \leq 0 \qquad \forall e \in \delta(S), k \in K, b \in B \tag{10.15d}$$

$$x, f, \pi \ \geq 0 \tag{10.15e}$$

We define the *multi-band robust cutset polyhedron* $\mathcal{N}^{\text{LF,mb}}(S)$ with respect to S as

$$\mathcal{N}^{\text{LF,mb}}(S) := \operatorname{conv}\left\{ \begin{array}{r} (x, f, \pi, \sigma) \in \mathbb{Z}_{\geq 0}^{|\delta(S)|} \times \mathbb{R}_{\geq 0}^{2|\delta(S)||K|} \times \mathbb{R}_{\geq 0}^{|\delta(S)||B|} \times \mathbb{R}^{|\delta(S)||K|} : \\ (x, f, \pi, \sigma) \text{ satisfies (10.15)} \end{array} \right\}.$$

The projection of $\mathcal{N}^{\text{LF,mb}}(S)$ onto the space of the variables x and π is denoted by

$$\mathcal{N}^{\text{LF,mb}}_{x,\pi}(S) := \left\{ \begin{array}{ll} (x,\pi) \in \mathbb{Z}^{|\delta(S)|}_{\geq 0} \times \mathbb{R}^{|\delta(S)||B|}_{\geq 0} & : \exists (f,\sigma) \in \mathbb{R}^{2|\delta(S)||K|}_{\geq 0} \times \mathbb{R}^{|\delta(S)||B|} \\ & \text{so that } (x,f,\pi,\sigma) \in \mathcal{N}^{\text{LF,mb}}(S) \end{array} \right\}.$$

We can apply Lemma 10.7 as $\mathcal{N}^{\text{LF,mb}}(S)$ defines a two-node multi-band robust network design problem. This yields the following.

Corollary 10.9. *It holds* $\dim(\mathcal{N}^{LF,mb}(S)) = 3|\delta(S)||K| + |\delta(S)||B| + |\delta(S)| - |K|$ *and* $\dim(\mathcal{N}^{LF,mb}_{x,\pi}(S)) = |\delta(S)||B| + |\delta(S)|$.

Multi-band robust cutset inequalities. Let $Q \subseteq Q_S$ be a non-empty subset of the cut commodities and $\{Q^b\}_{b\in B}$ be a family of non-empty subsets of Q. Further let $\mu^k := |\{b \in B \;:\; k \in Q^b\}|$ denote the number of subsets of $\{Q^b\}_{b\in B}$ containing commodity $k \in K$. Next, we add and relax constraints of formulation (10.15) to obtain a valid (base) inequality which is used in a subsequent application of MIR to derive a new class of valid inequalities. This approach is similar to the one we followed for Γ-RNDP.

First, note that

$$f^k_e(\delta(S)) = \sum_{e=\{i,j\}\in\delta(S)} (f^k_{ij} + f^k_{ji}) \geq \sum_{e=\{i,j\}\in\delta(S)} f^k_{ij} \geq \sum_{e=\{i,j\}\in\delta(S)} (f^k_{ij} - f^k_{ji}) = 1$$

holds by (10.15a) and (10.15b). Thus, the following two relaxations also holds

$$\sum_{k\in K} \bar{d}^k f^k(\delta(S)) \geq \sum_{k\in K} \bar{d}^k \text{ and } \sum_{b\in B}\sum_{k\in Q^b} \hat{d}^{b,k} f^k(\delta(S)) \geq \sum_{b\in B}\sum_{k\in Q^b} \hat{d}^{b,k}.$$

Second, by adding all link capacity constraints (10.15c) for all $e \in \delta(S)$, dual constraints (10.15d) for all $e \in \delta(S)$, $b \in B$, $k \in Q^b$, we obtain the inequality

$$\begin{aligned} & x(\delta(S)) + \sum_{b\in B}\sum_{k\in Q^b} \pi^b(\delta(S)) - \sum_{b\in B} \vartheta^b \pi^b(\delta(S)) + \sum_{b\in B}\sum_{k\in Q^b} \sigma^k(\delta(S)) - \sum_{k\in K} \sigma^k(\delta(S)) \\ & \geq \sum_{k\in K} \bar{d}^k f^k(\delta(S)) + \sum_{b\in B}\sum_{k\in Q^b} \hat{d}^{b,k} f^k(\delta(S)). \end{aligned}$$

Using the two relaxations introduced before, we can simplify and relax this inequality to

$$x(\delta(S)) + \sum_{b\in B}(|Q^b| - \vartheta^b)\pi^b(\delta(S)) + \sum_{k\in K}(\mu^k - 1)\sigma^k(\delta(S)) \geq \sum_{k\in K} \bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b} \hat{d}^{b,k} \tag{10.16}$$

which is valid for $\mathcal{N}^{\text{LF,mb}}_{x,\pi}(S)$. We call (10.16) the *non-disjoint base inequality.*

Lemma 10.10. *Let* $Q \subseteq Q_S$ *and let* $\{Q^b\}_{b\in B}$ *and* $\{\bar{Q}^b\}_{b\in B}$ *be two families of subsets* $Q^b, \bar{Q}^b \subseteq Q_S$. *Further let* μ^k *be defined as above and let* $\bar{\mu}^k$ *be defined analogously to* μ^k *but according to* $\{\bar{Q}^b\}_{b\in B}$.

If $\bar{Q}^b = Q^b$ *for all* $b \in B \setminus \{b^*\}$ *and* $\bar{Q}^{b^*} = Q^{b^*} \cup \{k^*\}$ *with* $k^* \in \bigcup_{b\in B\setminus\{b^*\}} Q^b$ *holds for an arbitrary but unique* $b^* \in B$, *then the non-disjoint base inequality* (10.16) *for* $\{Q^b\}_{b\in B}$ *dominates the non-disjoint base inequality for* $\{\bar{Q}^b\}_{b\in B}$.

Proof. We consider the left-hand side of the non-disjoint base inequality (10.16) for $\{\bar{Q}^b\}_{b\in B}$. It holds

$$\begin{aligned} &x(\delta(S)) + \sum_{b\in B}(|\bar{Q}^b| - \vartheta^b)\pi^b(\delta(S)) + \sum_{k\in K}(\bar{\mu}^k - 1)\sigma^k(\delta(S)) \\ =&x(\delta(S)) + \sum_{b\in B}(|Q^b| - \vartheta^b)\pi^b(\delta(S)) + \pi^{b^*}(\delta(S)) + \sum_{k\in K}(\mu^k - 1)\sigma^k(\delta(S)) + \sigma^{k^*}(\delta(S)). \end{aligned}$$

Next, we transform the corresponding right-hand side of the non-disjoint base inequality for $\{\bar{Q}^b\}_{b\in B}$ as follows

$$\sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in\bar{Q}^b}\hat{d}^{b,k} = \sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k} + \hat{d}^{b^*,k^*}.$$

From constraints (10.15a) and (10.15d) it follows $\hat{d}^{b^*,k^*} - \pi^{b^*}(\delta(S)) - \sigma^{k^*}(\delta(S)) \leq 0$. Thus, the non-disjoint base inequality (10.16) for $\{Q^b\}_{b\in B}$ dominates the one for $\{\bar{Q}^b\}_{b\in B}$; this completes the proof. □

Because of this lemma, we restrict our further investigation to families $\{Q^b\}_{b\in B}$ of disjoint subsets of Q, i. e., $Q^{b_1} \cap Q^{b_2} = \emptyset$ for all $b_1, b_2 \in B$, $b_1 \neq b_2$. Then, inequality (10.16) can be simplified and reads

$$x(\delta(S)) + \sum_{b\in B} i^b\pi^b(\delta(S)) \geq \sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k}. \tag{10.17}$$

with $i^b := |Q^b| - \vartheta^b$. We refer to inequality (10.17) as the *base inequality*. Applying mixed integer rounding yields

$$r_{Q^b}x(\delta(S)) + \sum_{b\in B} i^{b+}\pi^b(\delta(S)) \geq r_{Q^b}\left\lceil \sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k} \right\rceil \tag{10.18}$$

with $r_{Q^b} := \text{frac}(\sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k})$.

If the family $\{Q^b\}_{b\in B}$ of disjoint subsets of Q is in fact a partition of Q and $i^b = 0$ for all $b \in B$, then the inequality (10.18) reduces to

$$x(\delta(S)) \geq \left\lceil \sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k} \right\rceil. \tag{10.19}$$

In the beginning, we have mentioned that $d_\delta^{\max}(S)$ can be determined combinatorially by a min-cost flow on an auxiliary graph; cf. Lemma 3.9. Such an optimal flow also defines sets Q^b such that $d_\delta^{\max}(S) = \sum_{k\in Q}\bar{d}^k + \sum_{b\in B}\sum_{k\in Q^b}\hat{d}^{b,k}$ holds but does not yield a closed formula to determine the sets Q^b algebraically. Nevertheless by construction, inequality (10.19) is in fact the multi-band robust cutset inequality and by construction feasible for $\mathcal{N}^{\text{LF,mb}}(S)$.

Generalizing multi-band robust cutset inequalities. When we were considering the cutset-based inequalities for Γ-RNDP, we were able to parametrize the corresponding 2-dimensional base inequality (9.25) on a single edge by a parameter $i \in J$. This has yield more general inequalities, in particular the class of envelope inequalities.

Here, we can restrict our investigation to a single edge analogously to the problem for the Γ-RNDP. This yields the *single edge base inequality*

$$x + \sum_{b \in B} i^b \pi^b \geq \sum_{k \in Q} \bar{d}^k + \sum_{b \in B} \sum_{k \in Q^b} \hat{d}^{b,k} \tag{10.20}$$

which still includes $|B| + 1$ variables in contrast to the resulting 2-dimensional set in the Γ-robust setting. Moreover, to parametrize (10.20) by $\{i^b\}_{b \in B}$ all combinations of $i^b \in \{-\vartheta^b, \ldots, |Q_S| - \vartheta^b\}$ are feasible that satisfy $|Q_S| = \sum_{b \in B} |Q^b| = \sum_{b \in B}(i^b + \vartheta^b) = \sum_{b \in B} i^b + \sum_{b \in B} \vartheta^b$ or equivalently $\sum_{b \in B} i^b = 0$. In fact, there are as many feasible combinations of $\{i^b\}_{b \in B}$ as there are combinations of values for i^b that sum up to zero. This problem is a number partitioning problem and known to have exponentially many combinations. Hence, a general parametrization approach as for the Γ-robust setting is not promising and most likely computationally intractable. Instead, we propose a more tractable yet less general parametrization in the following.

Let $b_1, b_2 \in B$, $b_1 \neq b_2$ and $Q^b \subset Q$ be a partition of Q such that $i^b = 0$ for all $b \in B \setminus \{b_1, b_2\}$, $i^{b_1} = i$, $i^{b_2} = -i$ with $i \in \{-\vartheta^{b_1}, \ldots, \vartheta^{b_2}\}$. Then the single edge base inequality (10.20) reduces to

$$x + i\pi^{b_1} - i\pi^{b_2} \geq \sum_{k \in Q} \bar{d}^k + \sum_{b \in B} \sum_{k \in Q^b} \hat{d}^{b,k}. \tag{10.21}$$

Applying mixed integer rounding gives

$$r_{Q^b} x + i^+ \pi^{b_1} \geq r_{Q^b} \left\lceil \sum_{k \in Q} \bar{d}^k + \sum_{b \in B} \sum_{k \in Q^b} \hat{d}^{b,k} \right\rceil \qquad \text{if } i > 0 \tag{10.22a}$$

$$r_{Q^b} x \geq r_{Q^b} \left\lceil \sum_{k \in Q} \bar{d}^k + \sum_{b \in B} \sum_{k \in Q^b} \hat{d}^{b,k} \right\rceil \qquad \text{if } i = 0 \tag{10.22b}$$

$$r_{Q^b} x - i\pi^{b_2} \geq r_{Q^b} \left\lceil \sum_{k \in Q} \bar{d}^k + \sum_{b \in B} \sum_{k \in Q^b} \hat{d}^{b,k} \right\rceil \qquad \text{if } i < 0 \tag{10.22c}$$

where either the π^{b_1}- or the π^{b_2}-term has a non-zero coefficient if $i \neq 0$. If $i = 0$, inequality (10.22) reduces to the multi-band robust cutset inequality (10.14).

Notice, inequality (10.21) generalizes the base inequality (9.25) of the 2-dimensional polyhedron X^Γ for the Γ-RNDP: set bands $B := \{0, 1\}$, $b_1 = 1$, $b_2 = 0$, $\gamma^0 = \gamma^1 = 0$, $\Gamma^0 = |Q_S|$ and $\Gamma^1 = \Gamma$, and deviations $\hat{d}^{0,k} = 0$ and $\hat{d}^{1,k} = \hat{d}^k$ for all $k \in Q_S$. This gives $\vartheta^1 = \Gamma$ and $\vartheta^0 = |Q_S| - \Gamma$ and thus $i \in \{-\Gamma, \ldots, |Q_S| - \Gamma\}$. Besides, constraint (10.15d) for $b = 0$ and $\hat{d}^{0,k} = 0$ implies that $\pi^{b_1} = 0$ is feasible. Hence, for the Γ-robust setting inequality (10.21) reads

$$x + i\pi^{b_1} \geq \sum_{k \in Q} \bar{d}^k + \sum_{k \in Q^1} \hat{d}^{1,k}$$

where $\sum_{k \in Q^1} \hat{d}^{1,k}$ is maximal for the sum of the $|Q^1| = i + \vartheta^1 = i + \Gamma$ largest deviations $\hat{d}^{1,k}$; cf. inequality (9.25).

10.3 Algorithms

In this section, we present exact ILP-based separation algorithms for the multi-band robust cutset inequalities as well as for multi-band robust length and multi-band robust metric inequalities. These algorithms are generalizations of the corresponding ones for the Γ-RNDP.

10.3.1 Separation of multi-band robust cutset inequalities.

To obtain an exact approach to separate violated multi-band robust cutset inequalities (10.19) we generalize the ILP (9.47), which we used for the separation of Γ-robust cutset inequalities, to the multi-band robust setting. Therefore, we define binary variables δ_i $(i \in V)$ with $\delta_i = 1$ if and only if $i \in S$ determining the cut, $\alpha^{b,k}$ with $\alpha^{b,k} = 1$ if and only if $k \in Q^b$ determining the cut-crossing commodities in band $b \in B$, and $\bar{\delta}_{ij}$ $(ij \in E)$ with $\bar{\delta}_{ij} = 1$ if and only if $ij \in \delta(S)$ determining the cutset. In addition, let d determine the worst-case total demand value crossing the cut, and let R be the right-hand side value of the corresponding multi-band robust cutset inequality (10.19). Given an LP solution x^*, we minimize the feasibility (i.e., maximize the violation) of inequality (10.19) such that a negative objective value yields a violated cut. Then, the ILP formulation of the separation problem is given by

$$\min \sum_{ij \in E} x^*_{ij}\bar{\delta}_{ij} - R$$

$$\text{s.t.}\ \max\{\delta_i - \delta_j, \delta_j - \delta_i\} \le \bar{\delta}_{ij} \le \min\{\delta_i + \delta_j, 2 - \delta_i - \delta_j\} \qquad \forall ij \in E \tag{10.23a}$$

$$\max\{\delta_{s^k} - \delta_{t^k}, \delta_{t^k} - \delta_{s^k}\} \le \sum_{b \in B} \alpha^{b,k} \le \min\{\delta_{s^k} + \delta_{t^k}, 2 - \delta_{s^k} - \delta_{t^k}\} \qquad \forall k \in Q \tag{10.23b}$$

$$\sum_{k \in Q} \alpha^{b,k} \le \vartheta^b \qquad \forall b \in B \tag{10.23c}$$

$$\sum_{b \in B}\sum_{k \in Q} (\bar{d}^k + \hat{d}^{b,k})\alpha^{b,k} = d \tag{10.23d}$$

$$d \le R \le d + 1 - \varepsilon \tag{10.23e}$$

$$\alpha^{b,k}, \delta_i, \bar{\delta}_{ij} \in \{0,1\}, R \in \mathbb{Z}_{\ge 0}, d \ge 0 \qquad \forall b \in B, k \in Q, \forall ij \in E, \forall i \in V \tag{10.23f}$$

where constraints (9.47a) and (9.47b) define the logical dependencies between the indicator variables α^k, δ_i, and $\bar{\delta}_{ij}$. Constraint (10.23c) bounds the number of realizations in band b. Notice, that this constraint is no equality since there might be commodities which are not cut-crossing. The total demand d is calculated by (9.47e). Constraint (9.47f) guarantees the round-up of the right-hand side variable R using $0 < \varepsilon \ll 1$ to avoid rounding R to $\lceil d \rceil + 1$ or higher. Note, by setting a node limit (or time limit) for solving ILP (9.47), we obtain a (non-deterministic) heuristic ILP-based separation algorithm.

10.3.2 Separation of multi-band robust length inequalities

In the following, we generalize the polynomial exact separation of Γ-robust metric inequalities to the multi-band robust setting. Therefore, we follow an analogous approach as we did for the Γ-robust setting. First, we formulate the feasibility problem for a given capacity vector as an LP. Second, we dualize the LP to derive a valid multi-band robust length inequality for which the violation is determined by the objective value of the dual LP.

Given a capacity vector $\tilde{x} \in \mathbb{R}^{|E|}_{\geq 0}$, the feasibility problem can be formulated as

$$\max \alpha \tag{10.24a}$$

$$\text{s. t.} \sum_{uv \in E} (f^k_{uv} - f^k_{vu}) \geq \begin{cases} \alpha & u = s^k \\ -\alpha & u = t^k \\ 0 & \text{otherwise} \end{cases} \qquad \forall u \in V,\, k \in K \tag{10.24b}$$

$$\sum_{k \in K} \bar{d}^k f^k_e + \sum_{b \in B} \vartheta^b \pi^b_e + \sum_{k \in K} \sigma^k_e \leq \tilde{x}_e, \qquad \forall e \in E \tag{10.24c}$$

$$-\pi^b_e - \sigma^k_e + \hat{d}^{b,k} f^k_e \leq 0 \qquad \forall e \in E, k \in K, b \in B \tag{10.24d}$$

$$f^k_{uv}, \pi_e,\ \sigma^k_e,\ \alpha \geq 0 \qquad \forall e = uv \in E,\ k \in K. \tag{10.24e}$$

Thus, the point $(\tilde{x}, f^*, \pi^*, \rho^*)$ is feasible for ${}^{\mathrm{LP}}\mathcal{N}^{\mathrm{mb}}_x$ if and only if the objective value $(\alpha^*, f^*, \pi^*, \rho^*)$ of (10.24) is at least 1, i. e., if at least 100 % of the demand can be satisfied for each commodity. Next, we dualize (10.24) and obtain the following LP

$$\min \sum_{e \in E} \tilde{x}_e \ell(e) \tag{10.25a}$$

$$\text{s. t. } \beta^k_i - \beta^k_j \leq \bar{d}^k \ell(e) + \sum_{b \in B} \hat{d}^{b,k} m^{b,k}(e) \qquad \forall e = ij \in E, k \in K \tag{10.25b}$$

$$\sum_{k \in K} m^{b,k}(e) \leq \vartheta^b \ell(e) \qquad \forall e \in E \tag{10.25c}$$

$$\sum_{b \in B} m^{b,k}(e) = \ell(e) \qquad \forall e \in E,\ k \in K \tag{10.25d}$$

$$\sum_{k \in K} (\beta^k_{s^k} - \beta^k_{t^k}) \geq 1 \tag{10.25e}$$

$$\beta^k_i,\ \ell(e),\ m^{b,k}(e) \geq 0 \qquad \forall i \in V, e \in E,\ k \in K, b \in B. \tag{10.25f}$$

By strong duality, $\tilde{x}$ is feasible for ${}^{\mathrm{LP}}\mathcal{N}^{\mathrm{mb}}_x$ if and only if the optimal value of (10.24) is at least 1. If the objective value of a solution (β^*, ℓ^*, m^*) is strictly less than 1, constraint (10.25e) implies the multi-band robust length inequality

$$\sum_{e \in E} \ell^*(e) x_e \geq \sum_{k \in K} (\beta^{*k}_{s^k} - \beta^{*k}_{t^k}) \tag{10.26}$$

to be violated for $\tilde{x}$. Given a violated multi-band robust length inequality (10.26) (for length function ℓ), a violated multi-band robust metric inequality (for metric ℓ_M) can be constructed by defining $\ell_M(e) := \min\{\ell(e), \ell(u,v)\}$ for all $e = uv \in E$. Note, the right-hand side is not changed and the violation of the multi-band robust metric inequality is at least the violation of the multi-band robust length inequality. Furthermore, we claim that the obtained multi-band robust metric inequality is still valid.

Corollary 10.11. *Violated multi-band robust length inequalities and violated multi-band robust metric inequalities can be separated in polynomial time.*

Proof. Since (10.24) is a pure LP with polynomial size and ℓ_M is constructable from ℓ in polynomial time, violated multi-band robust metric inequalities can be exactly separated in polynomial time. □

Note, experiments for the special case of Γ-robust length inequalities have shown that the separation of these inequalities can be very time consuming in practice.

CHAPTER ELEVEN

COMPUTATIONAL STUDIES

We will finalize our investigation of robust network design problems with computational studies. In Chapter 9, we have presented a detailed study of the Γ-RNDP including various problem formulations, three different approaches to derive valid inequalities (cutset-based, arc residual capacity, and metric), and several exact and heuristic algorithms. Furthermore, we have motivated and shown how the Γ-RNDP can be generalized to the mb-RNDP in Chapter 10.

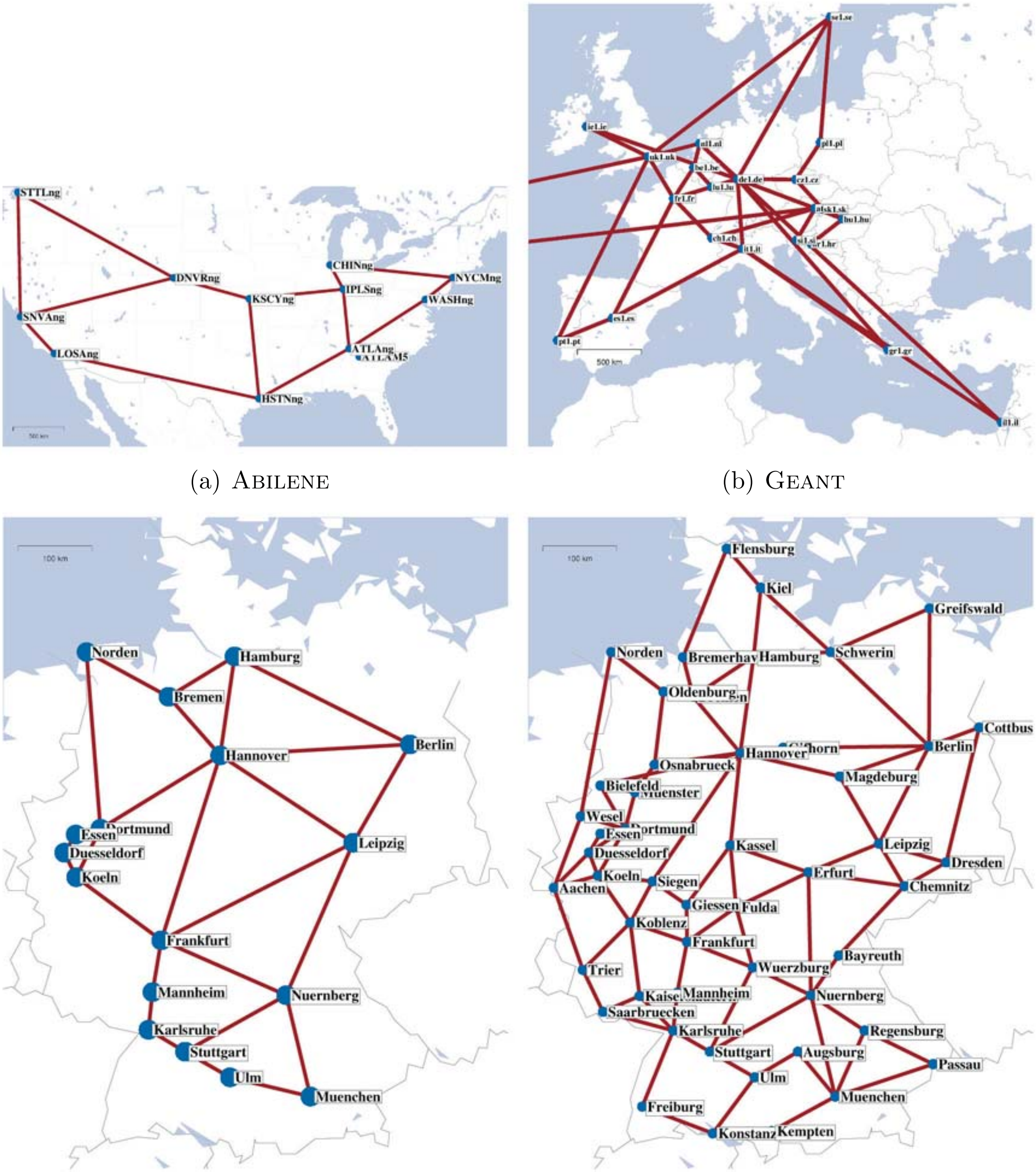
In the following, we present the detailed results of our excessive computational studies for the Γ-RNDP using historical real-life traffic measurements of telecommunication back-bone networks to build the Γ-robust uncertainty set. Our computational studies are based on and extend the computational experiments carried out by the author of the thesis while working with the ROBUKOM project [3]. The ROBUKOM project started in 10/2010 and has finished in 06/2013. It involved five research groups at different German universities and research institutions (RWTH Aachen University, TU Chemnitz, TU Berlin, Zuse Institute Berlin) as well as two industrial vendors (Nokia Siemens Networks GmbH & Co. KG, and DFN-Verein). The experimental results have partially been published by the author of this thesis and his co-authors in [103, 105, 106] and submitted for publication [55].

Environment. All algorithms were implemented in C++ using ILOG Cplex 12.1 [84] with ILOG Concert as MIP solver and branch-and-cut framework, respectively.

The computations were carried out using a single thread of Intel Xeon W3540 CPU at 2.93 GHz and 12 GB RAM. If not stated differently, all other solver settings were left at their defaults. A time limit of 12 hours was set for solving each problem instance.

11.1 Instances

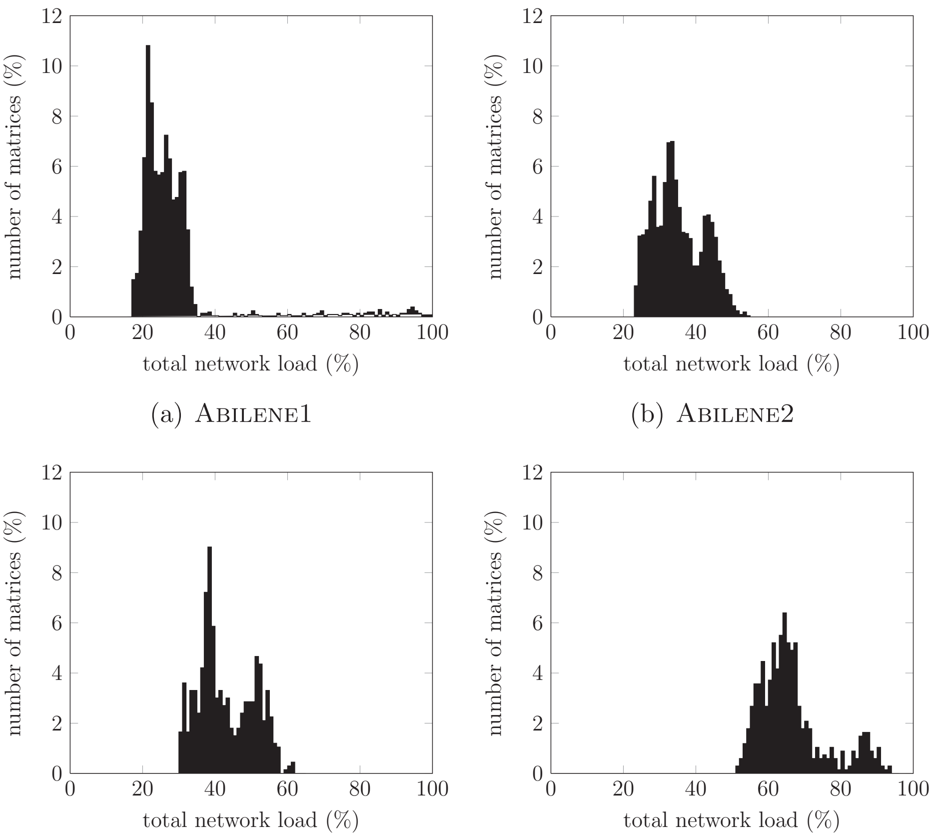
We consider problem instances based on live traffic data from different sources: the U.S. Internet2 Network (Abilene) [5, 162], the pan-European research backbone network Géant [1], and the national research backbone network operated by the German DFN-Verein [155] mapped on the network (Germany17) defined by the NOBEL project [4], and in addition mapped on a larger network (Germany50) [131]. The



(a) Abilene (b) Geant

(c) Germany17 (d) Germany50

Figure 11.1: Network topologies for the Abilene, Géant, Germany17, and Germany50 networks; cf. SNDlib [130, 131]. Notice, the Géant network contains two transatlantic links to New York City, USA.



(a) Abilene1 (b) Abilene2

(c) Geant1 (d) Geant2

Figure 11.2: Distribution of total network load over time for the considered traffic measurements of the Abilene1, Abilene2, Geant1, and Geant2 networks.

network topologies of all four networks are shown in Figure 11.1. For each network the live traffic data is given as a set of measured traffic matrices with a granularity of 5 minutes (Abilene, Germany17, Germany50) or 15 minutes (Géant). Recently, the live traffic measurements of these networks have also become available in the SNDlib [130, 131]. For Abilene and Géant we consider two time periods of one week each, resulting in two instances each: Abilene1, Abilene2, Geant1, and Geant2, respectively. For Germany17 and Germany50 we consider one day each. Figure 11.2 visualizes the distribution of the total network loads (i.e. the sum of all measured demands in the network at a certain moment) of the planning week for Abilene1, Abilene2, Geant1, and Geant2. We observe that most of the time three of the four networks are moderately loaded: 20–40 % (Abilene1), 20–50 % (Abilene2), and 30–60 % (Geant2). In contrast, the planning week of Geant1 includes many high loaded traffic matrices resulting in total network loads from 50 % to 90 %. Moreover, in particular for Abilene1, we see that there exist few matrices with very high total loads of up to 100 %.

We call the one-week time periods used for Abilene and Géant the *planning week(s)*. In addition to each planning week, the three following weeks are used in the evaluation

Network	Abilene	Géant	Germany17	Germany50
# nodes	12	22	17	50
# links	15	36	26	89
# demands	66	231	136	1044
traffic data:				
- period	6 months	4 months	1 day	1 day
- granularity	5 min	15 min	5 min	5 min
- # matrices (TMs)	48 095	10 737	288	288
instances	Abilene1	Geant1	Germany17	Germany50
from	2004/05/03	2005/05/26	2005/02/15	2005/02/15
to	2004/05/09	2005/06/01	2005/02/15	2005/02/15
instances (cont.)	Abilene2	Geant2		
from	2004/06/28	2005/07/28		
to	2004/07/04	2005/08/03		
per instance:				
- planning period	1 week	1 week	1 day	1 day
	2016 TMs	762 TMs	288 TMs	288 TMs
- evaluation period	4 weeks	4 weeks	1 day	1 day
	8064 TMs	2688 TMs	288 TMs	288 TMs

Table 11.1: Network and traffic properties of considered data sets. For each instance, the beginning and end of the planning period are shown in rows from and to. The format yyyy/mm/dd is used for all dates.

to simulate uncertain future traffic. Hence, in total four weeks of traffic measurements are used for each Abilene and Géant instance. Table 11.1 summarizes the network and traffic properties of all considered data sets.

Input instances are derived as follows: There are six data sets. For Abilene and Géant, we consider the traffic measurements of twice one planning weeks each. For Germany17 and Germany50, we consider all available traffic measurements (one day each). For each data set, let T denote the considered time period and let $d^k_{(t)}$ be the measured demand for commodity $k \in K$ at time step $t \in T$. In a first step, we determine a scaling factor $\sigma \in \mathbb{R}_{\geq 0}$ used to scale the traffic data in such a way that the sum of all peak demands $\max_{t \in T}(d^k_{(t)})$ over all commodities $k \in K$ amounts to 1 Tbps. That is, assuming that demands are given in Mbps, we set $\sigma := 10^6 / (\sum_{k \in K} \max_{t \in T}(d^k_{(t)}))$.

Notice that the mentioned data sets contain traffic measurements that are only weakly correlated w.r.t the different source-destination pairs: The averages of the correlations between the measured point-to-point traffic volumes are 0.21 (Abilene1), 0.18 (Abilene2), 0.10 (Geant1), and 0.06 (Geant2), respectively. For 95% of the traffic pairs, the correlation ranges from -0.49 to 0.50 (Abilene1), from -0.25 to 0.42 (Abilene2), from -0.40 to 0.40 (Geant1), and from -0.28 to 0.28 (Geant2). However, regardless of the correlation between some traffic pairs, we apply the Γ-robust approach and consider the Γ-RNDP. In a post-processing step, we evaluate the realized robustness of our Γ-robust solutions in Section 11.8. The mentioned correlations, a less conservative setting of the peak demands, and historical data including traffic matrices with more than Γ-many simultaneous peaks, influence the realized robustness we evaluate.

We set the link capacity module size, that is one unit of capacity, to 40 Gbps in our computational studies. This is in line with capacities in today's IP core networks; cf. citepROBUKOM.

11.2 Robustness parameters

Following the concept of Γ-robustness and given the historical data (traffic measurements), we have to determine the robustness parameter Γ, the nominal $\bar{d}^k$ and the deviation demand value $\hat{d}^k$ for each $k \in K$. Then, $\bar{d}^k + \hat{d}^k$ is the peak demand value. Alternatively, it is also sufficient to specify only $\bar{d}^k$ and $\bar{d}^k + \hat{d}^k$.

Given historical data, i. e., the planning week traffic measurements, we assume that individual demand values are uncorrelated. Therefore, we can consider each demand individually. The nominal value $\bar{d}^k$ should relate to some average realization value. We propose to use the arithmetic mean, geometric mean, or median of the available demand values for k. The peak value corresponds to the maximum demand value for k. Considering the fact, that only a few matrices with maximum total network load (and thus presumably in general also only a few matrices with maximum demand value k) exist, setting the peak demand value to the actual maximum demand value is too conservative.

Instead, we propose to eliminate statistical outliers to obtain less conservative values. For example, this can be done by neglecting the top 5 % demand values, i. e., we consider the 95 %-percentile. Of course, this introduces some kind of approximation into our modeling of the uncertainty set. On the one hand, a resulting Γ-robust solution may never achieve 100 % robustness because we omit the necessary high-demand matrices. However on the other hand, the level of conservatism may be reduced significantly. As for Γ-robustness, this is another kind of budgeting or trade-off decision to be made.

To evaluate this potential trade-off, we carried out a detailed computational study in Koster et al. [105]. We solved the compact link flow formulation (9.6) using Cplex as stand-alone solver for different combinations of values of nominal demand, peak demand, and Γ. We considered five possible ways to determine the nominal demand values (arithmetic mean, geometric mean, median, 60 %- and 70 %-percentile), nine possible ways to determine the peak demand values (the 80 %-, 85 %-, 90 %-, 95 %-, 96 %-, 97 %-, 98 %-, 99 %-, 100 %-percentile), and eleven possible Γ-values ($\Gamma = 0, 1, \ldots, 10$).

We determined the level of conservatism by the total cost (objective value) and evaluated the realized robustness of all 495 (nominal, peak, Γ)-combinations. Figure 11.3 visualizes the normalized cost versus the realized robustness for Abilene1. Each combination is marked by a cross. In addition, the lower convex envelope of the points is shown. Clearly,

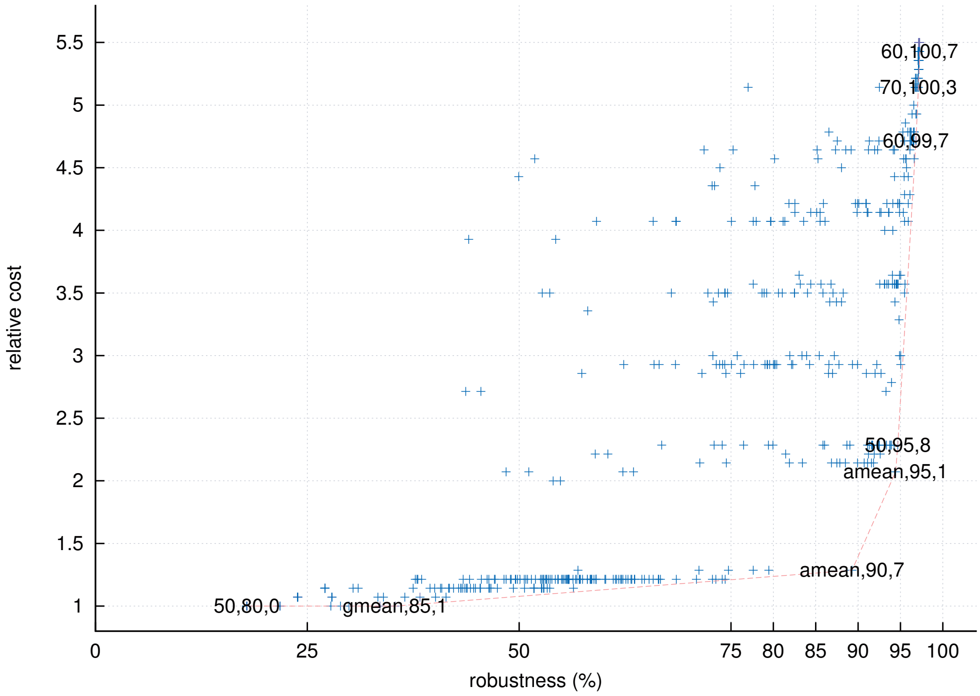


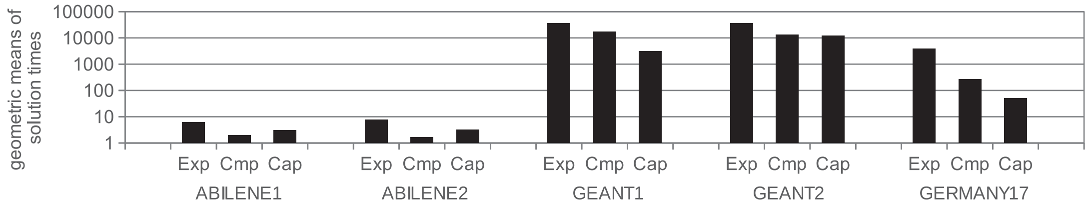
Figure 11.3: Determining robustness parameters. Normalized cost vs. realized robustness for the Abilene1 network is shown for different combinations of (nominal, peak, Γ) parameter settings. In addition, the dominant settings defining the lower envelope are labeled.

there exist (nominal, peak, Γ)-combinations which are dominated by some others, i. e., which are less robust for the same or higher costs. Following our analysis, pareto-optimal settings could be determined defining the lower envelope. We refer to [105] for further details.
Following our observations in [105], we determine the nominal values $\bar{d}^k$ and peak values $\bar{d}^k + \hat{d}^k$ formally as follows. For every commodity $k \in K$, we first order the demands $d^k_{(t)}$ non-decreasingly such that $d^k_{(t_i)}$ refers to the i-th element in the list with $d^k_{(t_1)} \leq d^k_{(t_2)} \leq \cdots \leq d^k_{(t_{|T|})}$. We then set $\bar{d}^k := |T| \cdot \sum_{t \in T} \sigma d^k_{(t)}$ and $\hat{d}^k := \sigma d^k_{(t_{\lceil 0.95|T| \rceil})} - \bar{d}^k$. Furthermore, we consider eleven values of Γ, i. e., $\Gamma \in \{0, 1, \ldots, 10\}$. Thus, we obtain in total 66 realistic test instances.

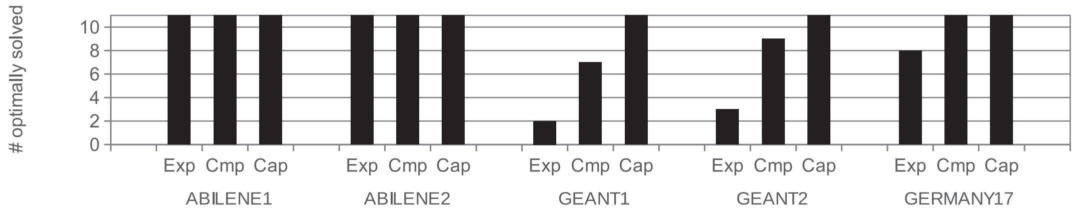
11.3 Comparison of formulations

In our first computational study, we compare the exponential link flow formulation (9.4), the compact link flow formulation (9.6), and the capacity (9.11) formulation of the Γ-RNDP. We abbreviate these formulations as *Exp*, *Cmp*, and *Cap*, respectively.

For *Exp*, the exponentially many inequalities (9.4a) are treated implicitly: violated ones are separated as so-called lazy constraints during the solving process. Figure 11.4



(a) solution times averaged over $\Gamma = 0, \ldots, 10$



(b) number of optimally solved instances

Figure 11.4: Comparison of different MILP formulations of the Γ-RNDP. The exponential (9.4), compact (9.6), and capacity (9.11) formulation are solved by Cplex as stand-alone solver. The Germany50 instance could not be solved in any case within the time limit.

visualizes the computational behavior of the three formulations: First, for each network a comparison of the solution times averaged over $\Gamma = 0, \ldots, 10$ is shown in Figure 11.4(a). Second, for each network and formulation, the number of optimally solved instances out of eleven is shown in Figure 11.4(b).

Impact on the solution time. In our computational study, the capacity formulation is (among) the fastest in 54.5 % of all 66 test instances; for the *Cmp* and *Exp* formulation it is 43.6 % and 9.1 %, respectively. Only for Abilene1 with $\Gamma = 0$ and Abilene2 with $\Gamma = 1$ the solution time of *Exp* has been slightly faster than the others by less than a second. Nevertheless, it is the overall slowest formulation in our study outperformed by *Cmp* and *Cap* for 75.8 % and 68.2 % of all instances. Its average solution times range from 2.1 (Geant1) to 14.5 (Germany17) times the averages obtained for *Cap*. In addition, very often the time limit has been reached while solving *Exp*.

Germany50 could not be solved in any case within the time limit of 12 hours. For Germany50, the optimality gap has been in the range from 43 % to 59 % (*Exp*), from 51 % to 71 % (*Cmp*), and from 47 % to 104 % (*Cap*).

It turns that out the MIP solver obtains many solutions feasible for the incomplete *Exp* formulation which are infeasible to the complete problem. Thus, these solutions are separated by additional model inequalities (lazy constraints). In our studies, the number of these non-redundant cuts slows down the solution processes of this formulation significantly. These results are in contrast to those in Fischetti and Monaci [69] for the set covering problem, since their instances for the set covering problem does only

abbreviation	class	reference
RCI	Γ-robust cutset inequality	(9.21)
REI	Γ-robust envelope inequality	(9.26), (9.28), (9.29)
upper REI	upper Γ-robust envelope inequality	(9.29)
lower REI	lower Γ-robust envelope inequality	(9.26), (9.28)
RMI	Γ-robust metric inequality	(9.8)
1-bounded RMI	1-bounded Γ-robust metric inequality	(9.45) : $\ell_M \in \{0,1\}^{\vert E\vert}$
rounded RMI	rounded Γ-robust metric inequality	(9.45)
tight RMI	rounded Γ-robust metric inequality with minimum right hand side	(9.46)
RARCI	Γ-robust arc residual capacity inequality	(9.40)

Table 11.2: List of considered classes of valid inequalities with abbreviations used in this chapter

include very small deviations compared to the measured real-world deviations of the telecommunication networks we consider.

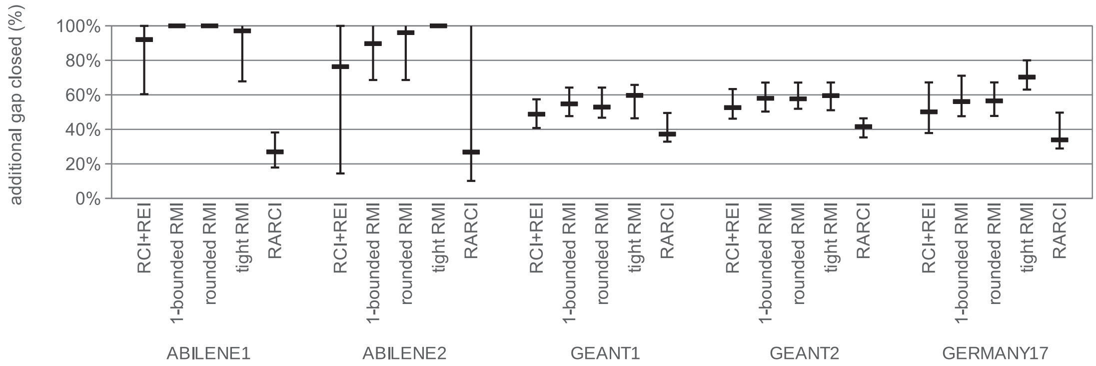
Impact on solvability. In the following, we consider the number of optimally solved instances within the time limit. First, we observe that the Abilene1 and Abilene2 instances could be solved optimally by all three formulations. The instance Germany50 could not be solved optimally by any model. For all 33 remaining instances, *Cap* solved all instances to optimality, whereas *Cmp* and *Exp* only solved 82 % and 39 %, respectively. Every instance solved by *Exp* was also solved by *Cmp*.

Conclusions. Clearly for all solved instances, *Exp* is outperformed by the others. Therefore, we focus on the compact and capacity formulations in the following studies. However, we want to remark that additional special purpose primal heuristics might improve on the computational behavior of the *Exp*.

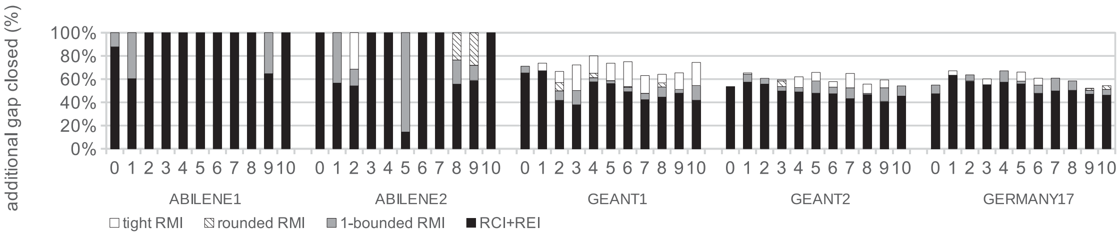
11.4 Strength of valid inequalities

In our next computational study, we investigate the strengthening of the linear relaxation of the compact link flow formulation (9.6) of the Γ-RNDP at the root node by nine different classes of valid inequalities. Therefore, we consider the exact separation of violated inequalities for these classes. Table 11.2 gives an overview of the considered classes and also introduces abbreviations for better readability. Notice, some of the inequalities generalize some of the others: RCIs are 1-bounded RMIs which are rounded RMIs and thus RMIs themselves. Obviously, tight RMIs are also RMIs, and upper and lower REIs are REIs.

Using the callback functionality of Cplex, we add an exact separation algorithm solving (9.51) including the option to separate bounded RMIs or tighten the right hand



(a) minimal/average/maximal additional gap closed at the root node for $\Gamma = 0, \dots, 10$



(b) additional gap closed at the root node by RMIs

Figure 11.5: Additional gap closed at the root node by cutset-based inequalities (RCIs+REIs), 1-bounded RMIs, rounded RMIs, tight RMIs, and RARCIs, for $\Gamma = 0, \dots, 10$. The gap is computed w.r.t. the LP dual bound and the best known primal bound.

side. Furthermore, we solve the exact separation ILP (9.47) to find violated Γ-robust cutset-based inequalities. Therefore, we solve the separation problem for all possible cutsets which we enumerate. In this study, only the root node is solved, all Cplex cuts are turned off, and the gap closed at the end of the root node is evaluated. Then, the additional gap closed is determined for each considered class of valid inequalities.

The additional gap closed is the ratio $(\mathrm{DB}_{\mathrm{root}} - \mathrm{DB}_{\mathrm{LP}})/(\mathrm{PB}_{\mathrm{best}} - \mathrm{DB}_{\mathrm{LP}})$ where $\mathrm{DB}_{\mathrm{root}}$ denotes the dual bound after solving the root node before branching, $\mathrm{DB}_{\mathrm{LP}}$ the objective value of the LP relaxation at the end of the root node, and $\mathrm{PB}_{\mathrm{best}}$ the (overall) best known primal bound. A gap closed of 100% is observed if the instance could be solved to optimality at the root node.

Figure 11.5 shows the (minimal/average/maximal) gap closed for each network (except Germany50), value of Γ, and considered classes of inequalities. For Germany50, the root node could not be solved within the time limit in any setting. This is due to the fact that the standard cuts of Cplex were turned off in this study. A further evaluation of Germany50 in a full cut-and-branch approach is presented later in Section 11.7.

Figure 11.5(a) shows the average gap closed by separating cutset-based inequalities (RCIs+REIs), 1-bounded RMIs, rounded RMIs, tight RMIs, or RARCIs exactly. Considering the averages, we observe that already RCIs+REIs close the gap on average at least by 48.8 % (Geant1). This concurs with the conclusion of the effectiveness of RCIs+REIs in Koster et al. [106]. When considering more general RMIs, the average gap can be closed by additional 7.7 % by 1-bounded RMIs, further 1.3 % by rounded RMIs, and even further 4.2 % by tight RMIs. In this way, the gap can be closed in total by additional 13.3 % by tight RMIs compared to RCIs+REIs. Furthermore, the total gap closed by tight RMIs is 100 % for both Abilene instances and $\Gamma = 0, \ldots, 10$, solving these instances optimally at the root node. For RARCIs, we observe that the average additional gap closed is always less than the corresponding values for the other classes of inequalities. It ranges from 26.9 % for Abilene1 to 41.6 % for Geant2 on average. Considering the minimal and maximal observed values for the additional gap closed, we notice that except for two cases (Abilene2 and RCIs+REIs and RARCIs, respectively), the minimal and maximal values are relatively close to the average and deviate less than 32 %.

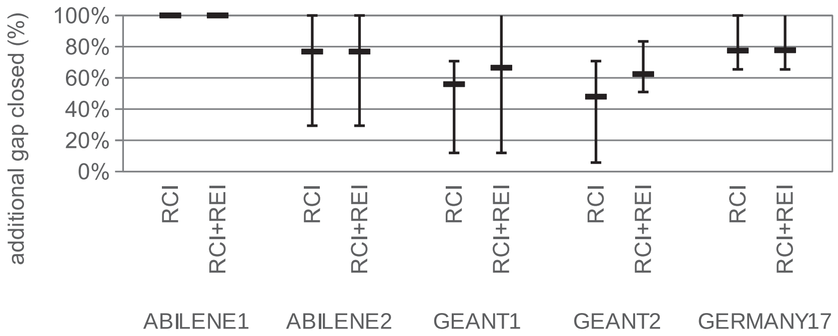
Figure 11.5(b) provides a break down of the additional gap closed by individual subclasses of RMIs per instance and value of Γ. RARCIs are left out since they are not RMIs. Besides, due to their low potential of closing the integrality gap as seen above, they are of little interest for us. In Figure 11.5(b), we observe that RCI contribute most to the gap closed. Nevertheless, more general RMIs as (1-bounded) RMIs are needed to completely close the gap at the root node for all Abilene instances. In addition, tight RMIs are needed for Abilene2 and $\Gamma = 2$ to achieve 100 % gap closed.

So far, we have considered the additional gap closed at the root node w.r.t the overall best known primal solution and the LP dual bound. In particular, the cut generating methods of Cplex have been switched off. Next, we change the latter and allow Cplex to separate its own cuts in addition to ours. Thus, we can evaluate the added value of our cuts in another more realistic setting. For this extra analysis, we focus on cutset-based inequalities: RCIs, upper REIs, and lower REIs. Similar to Figure 11.5, the (minimal/average/maximal) additional gap closed for this set-up is shown in Figures 11.6(a) and 11.6(b). Moreover, the number of separated RCIs, lower and upper REIs are reported in Figure 11.6(c).

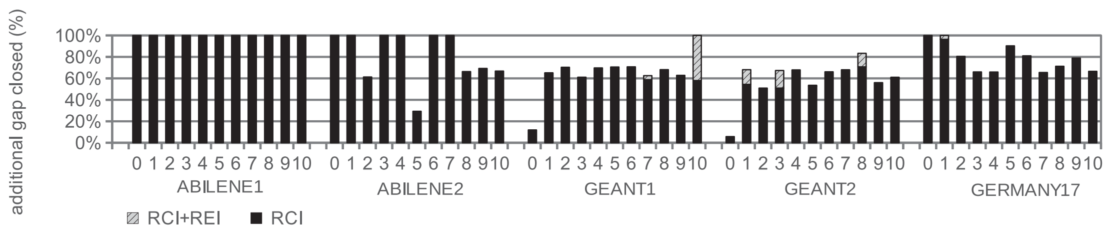
We observe that the RCIs together with the cuts of Cplex perform remarkable well. Abilene1 is solved optimally at the root node for all considered values of Γ. Also for Abilene2, the gap can be closed in most of the time. In total, for more than 90 % of all instances, the optimality gap can be closed at least by 50 % at the root node.

Only for 7 of 66 instances, we observe an additional gain of 2.5 % on average by separating violated REIs. In fact, Figure 11.6(c) shows that for Geant1, Geant2, and Germany17 hundreds or thousands of violated RCIs are separated while only a few violated lower REIs (always less than 10 cuts, except for Geant2 and $\Gamma = 3, 8, 10$) and no upper REIs are separated.

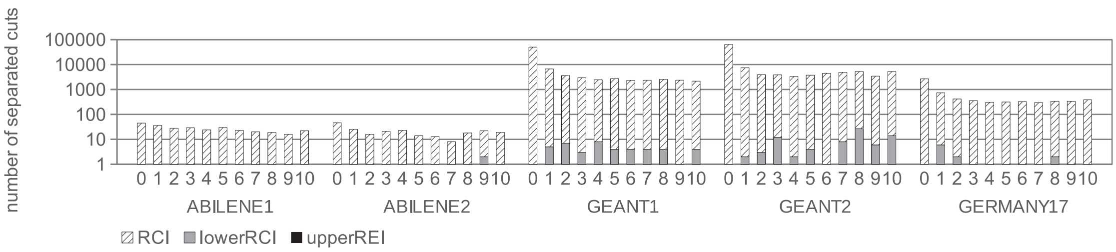
Conclusions. The results of our first computational study show that separating violated RMIs closes the gap significantly at the root node. In particular cutset-based inequalities are contributing a lot to the tightening of the formulations. Furthermore, when switching on the cut generation of Cplex, RCIs+REIs still have a positive impact. In contrast,



(a) minimal/average/maximal additional gap closed at the root node for $\Gamma = 0, \ldots, 10$



(b) additional gap closed at the root node



(c) number of separated RCIs and REIs

Figure 11.6: Additional gap closed at the root node by RCIs and REIs, and RCIs alone, for $\Gamma = 0, \ldots, 10$. The gap is computed w.r.t. the dual bound obtained by standard CPLEX and the best known primal bound. Notice, the dual bound is not necessarily the LP bound. In addition, the number of separated RCIs and REIs is shown.

RARCIs have the lowest contribution to the tightening of the formulation and thus should not be applied in a cut-and-branch approach.

11.5 Speed-up of the compact link flow formulation

In the next computational study, we investigate the speed-up by integrating the separation of violated RMIs in a cut-and-branch approach to solve the compact flow formulation (9.6)

	Algorithm	Description
I	Cplex	Cplex as stand-alone solver; no additional separation of RMIs.
IIs	RCI/s	Heuristic separation of violated RCIs using the shrinking heuristic described in Section 9.3.1. The network is shrunken with respect to the slack values of inequalities (9.6a) and (9.6b). If no violated cut is found heuristically, an ILP-based exact separation algorithm is run.
IIc	RCI/c	Same as algorithm IIs, but the network is shrunken with respect to the LP value of the capacity variables x.
IIIs	1. RCI/s, 2. rounded RMI/s	First, algorithm IIs is run. Second, if no cut has been found, the network is shrunken with respect to the slack values of (9.6a) and (9.6b). Then violated RMIs are separated by solving the ILP (9.51) for the shrunken network.
IIIc	1. RCI/c, 2. rounded RMI/c	Same as algorithm IIIs, but all network shrinkings are done with respect to the LP value of the capacity variables x.
IVc	1. RCI/c, 2. tight RMI/c	Same as algorithm IIIc. If a violated RMI is found, its best right hand side β_{ℓ_M} is determined by solving an ILP. Only the resulting tight RMI is separated.

Table 11.3: Overview of considered algorithms to solve the Γ-RNDP

of the Γ-RNDP. In our previous studies, we observed that on the one hand violated REIs are almost never separated and on the other hand violated RARCIs perform bad w.r.t. additional gap closed by the other classes. Therefore, we focus on RMIs and their promising subclasses in this study. We consider several heuristic algorithms summarized in Table 11.3.

Figure 11.7 shows the average speed-up factors obtained by the individual algorithms. The factors are normalized to the solution time of Cplex as stand-alone solver (algorithm I). For example, a solution time of 60 seconds compared to a corresponding solution time of 120 seconds of Cplex yields a speed-up factor of 2. For the Abilene network, we observe that all average speed-up factors are less than 1.0 except for algorithm IIs and Abilene1. In fact, the solution times of Abilene1 and Abilene2 instances are in the range of only a few seconds. Therefore, the overhead introduced by all separation algorithms results in a slow-down on average. For all larger instances, there is a significant average speed-up factor of at least 1.85 (algorithm IIIc, Geant2). Separating only

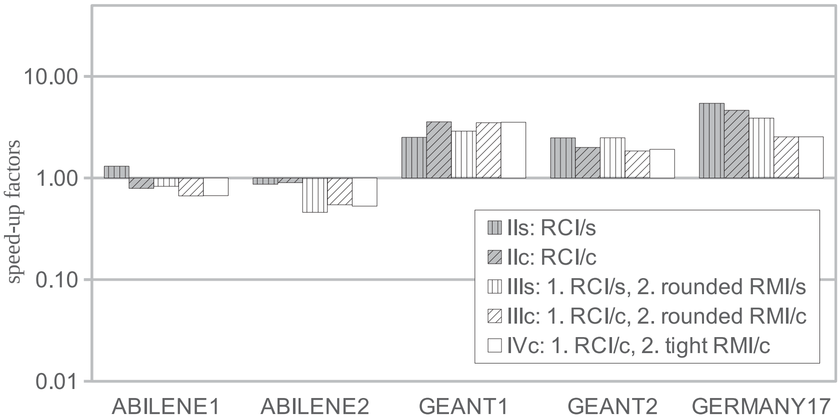


Figure 11.7: Speed-up factors of algorithms IIs, IIIs, IIc–IVc in a cut-and-branch approach to solve the compact flow formulation (9.6) of the Γ-RNDP. All factors are normalized to solution times of algorithm I (CPLEX as stand-alone solver).

violated RCIs (algorithms IIs and IIc) yields good speed-up factors for GEANT1, GEANT2, and GERMANY17 of 2.5 (IIs) and 3.6 (IIc), 2.5 (IIs) and 2.0 (IIc), 5.5 (IIs) and 4.6 (IIc), respectively. Most of the times these are better than the corresponding speed-up factors obtained by algorithms IIIc and IVc where additionally tight RMI are separated heuristically. This can be explained by the additional computational effort to solve the ILP (9.51) to separate violated RMIs. Although the network has been shrunken, the remaining integer problem may still be computationally hard reducing the speed-up factor compared to the RCI-only algorithm IIc.

Conclusions. In our study, we have observed a substantial speed-up in solving the compact flow formulation of the Γ-RNDP on mid-sized instances (e.g., GEANT1, GEANT2, and GERMANY17) by separating RMIs in a cut-and-branch approach. Furthermore, the highest speed-up factors are achieved by separating RCIs.

11.6 Speed-up of the capacity formulation

In our next computational study, we compare the compact flow formulation (9.6) to the capacity formulation (9.11) in a cut-and-branch approach. The exponentially many robust metric model inequalities of the capacity formulation are handled implicitly: violated RMIs (10.26) separated on-the-fly as lazy constraints using the callback capabilities of CPLEX whenever an integer solution is found. We start without any RMI.

We consider algorithms I, IIc, and IVc introduced before and summarized in Table 11.3. Algorithms IIs and IIIs are not applicable to the incomplete capacity formulation because the slack values of inequalities (9.6a) and (9.6b) cannot be evaluated as these constraints

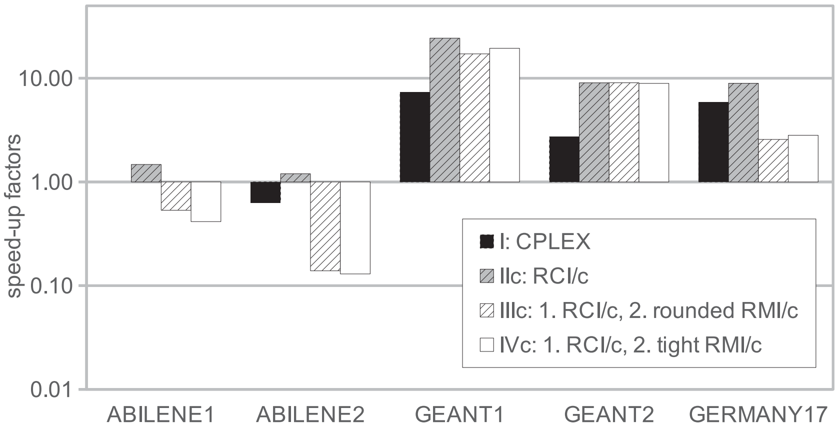


Figure 11.8: Speed-up factors of algorithms IIc–IVc in a cut-and-branch approach to solve the capacity formulation of the Γ-RNDP. Algorithm I: CPLEX corresponds to solving the capacity formulation (9.11) with CPLEX as a stand-alone solver. All factors are normalized to solution times of CPLEX solving the compact flow formulation as stand-alone-solver.

do not exist. Notice, RARCIs are also not applicable to the capacity formulation as their definition includes flow and dual variables.

Figure 11.8 shows the average speed-up factors of algorithms IIc–IVc normalized to the solution times of CPLEX solving the compact flow formulation as stand-alone solver. Because the solution times of CPLEX as stand-alone solver solving the compact flow formulation and the capacity formulation differ, we report also on the normalized solution times of CPLEX for the capacity formulation. Similarly to the compact flow formulation, we observe that on the small-sized ABILENE1 and ABILENE2 instances no speed-up can be achieved. The computational effort to separate RMI slows down the instances compared to their fast solution times. For GEANT1, GEANT2, and GERMANY17, we notice significant average speed-up factors of 2.7 (GEANT1), 5.9 (GERMANY17), and 7,3 (GEANT2) obtained by CPLEX solving the capacity formulation (algorithm I). Introducing the separation of RMIs yields even higher speed-up factors on average whereas the best are achieved by algorithm IIc: 1.5 (ABILENE1), 1.2 (ABILENE2), 24.4 (GEANT1), 9.0 (GEANT2), and 8.9 (GERMANY17). For each network, the best average speed-up factor for the capacity formulation is higher than the corresponding one for the compact flow formulation (cf. Figure 11.7). In particular, the algorithms IIs (only compact flow formulation) and IIc perform well achieving the highest average speed-up factors in most cases. In summary, we conclude that the capacity formulation is computationally more tractable than the compact flow formulation for mid-sized instances. A significant speed-up could be observed in our computationally studies.

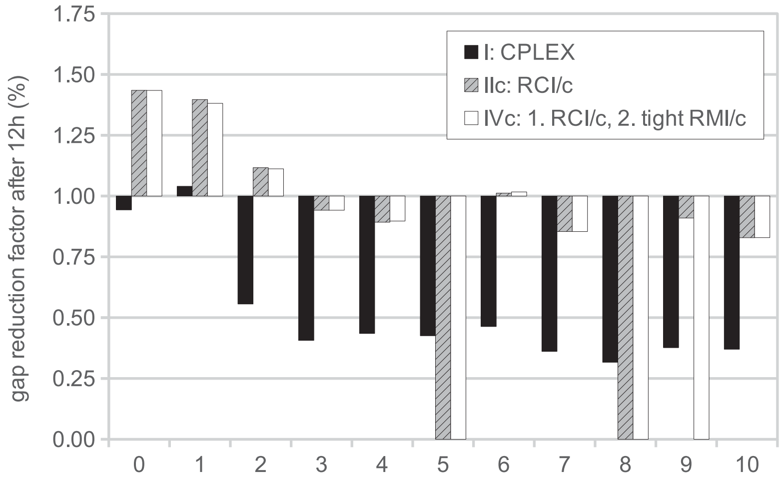
11.7 Handling large instances using the capacity formulation

To evaluate the impact of RMIs in a cut-and-branch approach to larger instances such as Germany50 which cannot be solved within the time limit of 12 hours, we consider the optimality gap after 12 hours. In particular, we evaluate the gap reduction factor compared to the optimality gap left by Cplex as stand-alone solver when solving the compact flow formulation to the time limit. For example, let Cplex solve the Germany50 instance to 60 % of optimality and let another algorithm obtain a optimality gap of 40 %. Then the gap reduction factor of this other algorithm is $0.6/0.4 = 1.5$ compared to Cplex. Therefore, a gap reduction factor less than 1.0 is given if the optimality gap after 12 hours is larger than the corresponding one by Cplex and the compact flow formulation. The cuts obtained by algorithm IIIc may be strengthened to tight RMIs and hence are dominated by those obtained by algorithm IVc. Thus, we do not consider algorithm IIIc in this study.

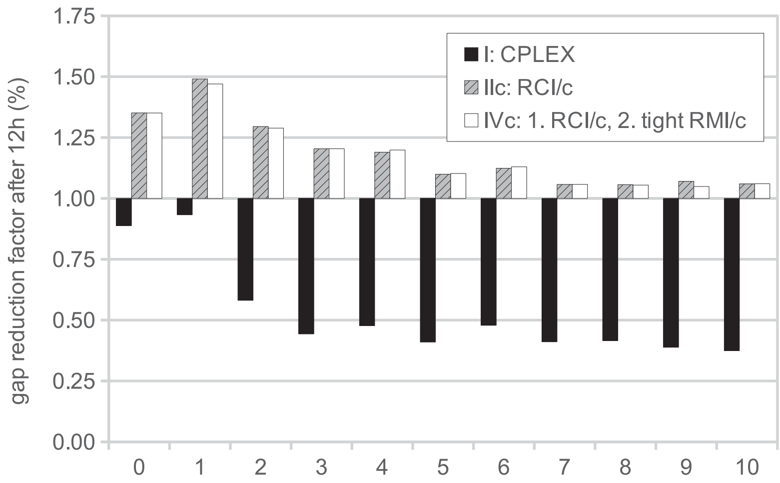
Figure 11.9 shows the gap reduction factors for Germany50 compared to Cplex as stand-alone solver solving the compact flow formulation. The figure consists of two diagrams. Figure 11.9(a) reports the gap reduction factors using the actual gap value obtained by the MIP solver. In contrast, the gap reduction factors in Figure 11.9(b) are calculated based on the best known primal bound values (and thus normalizing the influence of the primal bounds on the gaps). Hence, the latter figure focuses on the improvement in the dual bound value.

In Figure 11.9(a) we observe that the optimality gaps of the capacity formulation are worse than the corresponding ones for the compact flow formulation. Oftentimes the gaps are in fact as twice as large. This can mainly be explained by the fact that integer solutions found for the incomplete capacity formulation are actually not feasible for the complete formulation and thus discarded at a later point during the solving process. Hence they are slowing down the solving process. For Algorithm IIc and IVc and $\Gamma = 5, 8, (9)$, no primal solution has been found resulting in gap reduction factors of 0. Nevertheless, notice that for $\Gamma \leq 2$ positive gap reduction factors could be achieved for algorithms IIc and IVc. Moreover, except for the cases where no primal solution could be determined the gap reduction factors of algorithm I is always the worst.

If we do not limit our analysis on the gap but consider the dual bound values, we get a different picture. Therefore we normalize the primal bound values by taking the best known primal bound value for each value of Γ when determining the remaining gap after 12 hours. The resulting gaps are shown in Figure 11.9(b). First, we notice that algorithm I (Cplex) for the capacity formulation cannot close the optimality gap better than Cplex for the compact flow formulation. In fact, in most cases algorithm I finishes with optimality gaps at least twice as high. This bad performance can be explained by the fact that during the solution process of algorithm I temporary solutions are obtained which are in fact infeasible to the complete capacity formulation but not the current incomplete one. To lower the computational effort, these solutions are only separated if they are integer. In contrast, algorithms IIc and IVc separate violated



(a) gap reduction factors according to obtained primal bounds



(b) gap reduction factors according to best known primal bounds

Figure 11.9: Gap reduction factors of algorithms I, IIc, and IVc in a cut-and-branch approach to solve the capacity formulation of the Γ-RNDP for the Germany50 instance. Algorithm I: Cplex corresponds to solving the capacity formulation with Cplex as a stand-alone solver. All factors are normalized to the optimality gap left by Cplex as a stand-alone solver at the time limit of 12 hours when solving the compact flow formulation.

rounded RMIs for fractional solutions in addition to the model constraints. So both finish with smaller optimality gaps compared to Cplex and the compact flow formulation. The gap reduction factor ranges from 1.1 ($\Gamma = 10$) to 1.5 ($\Gamma = 1$). For $\Gamma \leq 4$ the gap reduction factor is at least 1.2. For larger values of Γ it decreases. The gap reduction factors do not differ much between the algorithms IIc and IVc except for $\Gamma = 1$ and $\Gamma = 9$ where algorithm IIc clearly outperforms the others. In summary, we observe that a cut-and-branch approach with rounded RMIs on the capacity formulation of the Γ-RNDP significantly closes the optimality gap of the large Germany50 instance in our study.

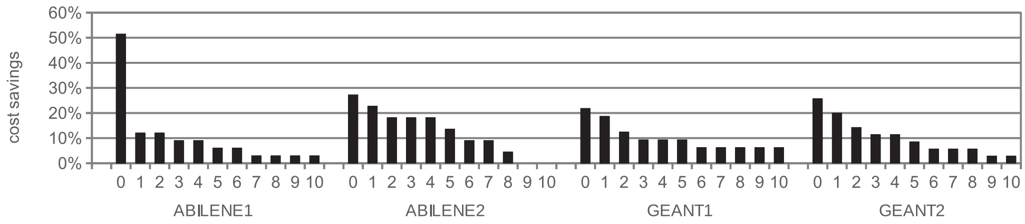


Figure 11.10: Cost savings of robust ABILENE and GÉANT network design compared to classical network design with peak demand values (i.e., corresponding to Γ-Robust Network Design with $\Gamma = |K|$), for $\Gamma = 0, \ldots, 10$.

Comparing Figures 11.9(a) and 11.9(b), we conclude that algorithms IIc and IVc significantly improve the dual bounds. At the same time this has only little effect on the overall performance because the straight-forward approach to solve the incomplete capacity formulation by separating model inequalities is slowed down by disadvantageous primal solutions. Here, additional good primal heuristics would be promising.

Conclusions. In conclusion, in our studies it turns out that the capacity formulation of the Γ-RNDP is computationally more tractable on mid-sized instances. When used in a cut-and-branch approach, it offers higher speed-up factors on mid-sized instances and might lowers the optimality gap left at the time limit for larger instances when combined with strong primal heuristics.

11.8 Quality of optimal robust network designs

Our final computational study focuses on the quality of optimal robust network designs. We investigate two aspects as quality criteria: the cost of an optimal robust network design and the realized robustness with respect to a given set of traffic matrices.

For the latter, we only consider the ABILENE and GÉANT networks because only for these networks, traffic measurements spanning several weeks are available and thus, a meaningful robust evaluation can be carried out. Given the traffic measurements of one week as input data (as described above), we include additional weeks of traffic measurements in our evaluation of the realized robustness to simulate uncertain future traffic.

Cost savings. In contrast to the price of robustness which compares the objective value of an optimal robust solution to the same of an optimal non-robust solution, we report on the *cost savings* by comparing to the value of the most conservative solution, i. e.,

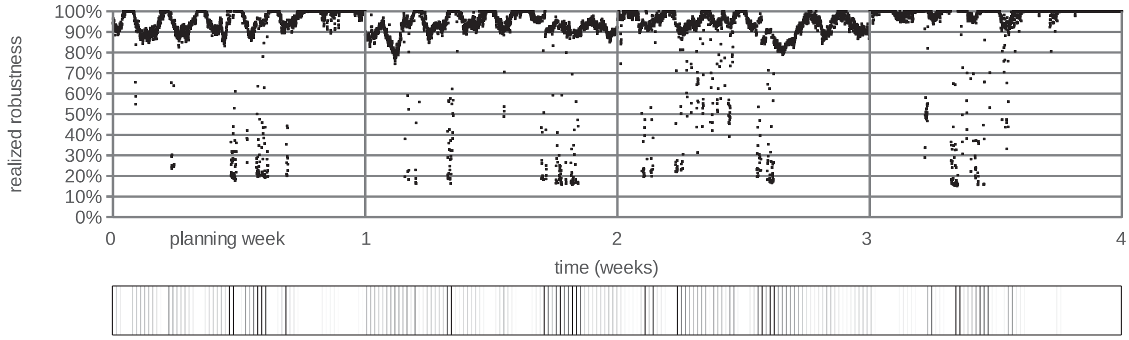


Figure 11.11: Realized robustness of ABILENE1 and $\Gamma = 0$. Additionally, the corresponding traffic loss profile is shown below the realized robustness diagram.

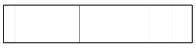
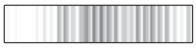
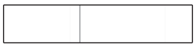
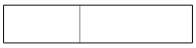
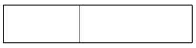
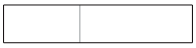
Γ	ABILENE1	ABILENE2	GEANT1	GEANT2
0				
1				
5				
10				
$\|K\|$				

Table 11.4: Traffic loss profiles of ABILENE and GÉANT networks and selected values of Γ. Note, the traffic loss profile of ABILENE1 and $\Gamma = 0$ is the same as in Figure 11.11.

to the setting $\Gamma = |K|$. This corresponds to the (conservative) classical network design where capacities are optimized against the worst peak scenario.

Figure 11.10 shows the relative cost savings of ABILENE1, ABILENE2, GEANT1, and GEANT2 for $\Gamma = 0, 1, \dots, 10$. Clearly, the cost savings (compared to $\Gamma = |K|$) decrease with increasing value of Γ as the costly additional installment of link capacity modules is implied. Still, for $\Gamma = 5$ about 10% of the cost can be saved in all considered networks. We also see that the advantage of a robust design in terms of cost is relatively small already for $\Gamma = 10$. That is, allowing for 10 commodities being at the peak simultaneously gives capacity designs at a cost similar to networks that are designed against the all-peak scenario.

Realized robustness. Given a traffic matrix d, a capacity design x, and a static routing f of all commodities, the *realized robustness* is determined as the maximal fraction of the total demand $\sum_{k \in K} d^k$ that can be realized as flow within the given capacities x and using the routing defined by f; cf. Section 3.5. To calculate this value, we solve a linear program that takes d, x, and f as input and maximizes the fraction of total demand that can be realized. For each instance, we evaluate the realized robustness in every time step for traffic measurements of four consecutive weeks. The first week within the time period refers to the planning week. Recall that only traffic measurements of the planning week are used to parametrize the Γ-model. The traffic measurements of the remaining three weeks are used to simulate uncertain future traffic.

Figure 11.11 shows the result for Abilene1 and $\Gamma = 0$. The average of the realized robustness in this case is 88.2 %, that is, on average over the considered time period we are able to realize 88.2 % of the demand in the given capacities using the given static routing. Clearly, such a value does not catch the change of the realized robustness over time. We observe most of the time a realized robustness of 85–100 %. But there exist several traffic matrices for which the realized robustness is as worse as 15 %. To capture this temporal aspect of robustness, we propose a different visualization which we call the *traffic loss profile*. The corresponding traffic loss profile of Abilene1 and $\Gamma = 0$ is shown below the diagram in Figure 11.11. This profile visualizes each traffic matrix by a vertical line whose gray scale value corresponds to the relative traffic amount that cannot be routed (i.e., 100 % minus the realized robustness of the considered traffic matrix). The darker the line, the more traffic is lost, i.e., the less robustness is realized. Hence, a profile without lines is best and corresponds to a totally robust network design.

Table 11.4 shows the traffic loss profiles of optimal robust network designs for Abilene1, Abilene2, Geant1, and Geant2 for selected values of Γ. Notice that the traffic profiles for $\Gamma = |K|$ correspond to the best robustness that can be achieved with the given parametrization of nominal and peak demand values $\bar{d}$ and $\bar{d} + \hat{d}$. These values are derived from traffic measurements in the planning week. Moreover, as we used 95 % percentiles to determine the peak demand values $\bar{d}^k + \hat{d}^k$ for each commodity $k \in K$, even w.r.t the planning week the profiles are not necessarily totally robust.

Fixing $\Gamma = 0$, we observe that the realized robustness of optimal network designs of the four instances are quite different: 88.2 % (Abilene1), 99.9 % (Abilene2), 93.2 % (Geant1), and 96.3 % (Geant2). Comparing $\Gamma = 0$ to $\Gamma = 1$ already shows a significant improvement for Abilene1 (from 88.2% to 94.9 %) and Geant2 (from 96.3 % to 98.6 %). By trend, the realized robustness of a network design increases when Γ increases, i.e., the corresponding traffic loss profile has fewer vertical lines or the gray scales of the lines are brighter. A decrease can only occur due to a different and disadvantageous traffic routing. For example, this can be observed for Geant2 where the realized robustness decreases from 99.8 % to 99.7 %; compare the traffic loss profiles. Note, for Abilene2 the classical network design ($\Gamma = 0$) achieves already a realized robustness of almost 100 %.

For $\Gamma = |K|$, the realized robustness is 95.8 % (Abilene1), 99.9 % (Abilene2), 98.0 % (Geant1), and 99.9 % (Geant2). This is best for the given choice of nominal and peak demand values. Note that by choosing the 95% percentile to determine the peak demand values, some strong peaks in the traffic data of Abilene1 are ignored, resulting in the shown traffic profile with some dark lines still present for $\Gamma = |K|$. The evaluation of the realized robustness for all instances and considered values of Γ yields that for $\Gamma \geq 1$ (Abilene1), $\Gamma \geq 0$ (Abilene2), $\Gamma \geq 10$ (Geant1), and $\Gamma \geq 5$ (Geant2), at least 99 % of the corresponding realized robustness value for $\Gamma = |K|$ is achieved. The traffic loss profiles for these cases hence basically coincide with the corresponding profiles for $\Gamma = |K|$; compare with Table 11.4.

Further, by comparing the first quarter of each traffic loss profile to its remaining part, we can evaluate the realized robustness of the planning week (the first week of the four-week time period) compared to the remaining three weeks representing uncertain future traffic. For example, we observe that the network design of Geant1 realizes high robustness during the planning week but is significantly less robust in the following weeks. Clearly, the network load has been larger in the three weeks following the planning week.

Conclusions. Optimal robust network designs provide high potential for significant cost savings compared to the conservative setting where only peak demand values are considered. Further, the traffic loss due to peaks drops significantly already for relatively small values of Γ. In particular, for $\Gamma \leq 5$ a remarkable increase in the realized robustness can be achieved. A good value for Γ seems to depend on the size of the instance. For Abilene a value $\Gamma = 1$ is sufficient for high robustness while for Géant choosing $\Gamma = 5$ gives a good trade-off between cost and robustness. With these values we get almost totally robust networks at a cost of roughly 10–20 % less than the cost for a network designed for the all-peak scenario; see Figure 11.10.

CONCLUSIONS

In this thesis, optimization problems under data uncertainty have been studied by means of mathematical optimization methods, polyhedral combinatorics, and computational experiments. Strongly influenced by our work with the ROBUKOM project and our industrial partners from telecommunications, we aimed to develop and investigate robustness concepts suited to be applied to network planning problems arising in telecommunications and ILPs in general.

In total, we have studied four different robustness concepts: the well-known Γ-robustness, its recent refinement multi-band robustness, the new and more general submodular robustness, and the adaptive two-staged recoverable robustness. Applying these concepts, we have investigated selected robust variants of the KP and NDP in detail. The study of robust KPs has given us a fundamental understanding of the properties of each robustness concept and its theoretical applicability to general ILPs. By the investigation of robust NDPs, we have considered a challenging application of robust concepts to a more complex problem while also investigating the underlying mathematical problem of planning problems which arise in telecommunications and many other application areas where capacities for multi-commodity flow problems have to be dimensioned and a corresponding flow has to be determined.

For each robustness concept, we have considered the corresponding robust KP stating its ILP formulations and have given insights to its polyhedral structure. In particular, the structure of covers has been identified for all robust KPs. Furthermore, we have given canonical and strengthened cover extensions yielding stronger extended cover inequalities than the best known until now. For the SMKP, we have generalized $(1, k)$-configurations, $(1, k)$-configuration inequalities, and weight inequalities; including a generalization of the strong results on $(1, k)$-configurations inequalities by Padberg [134]. Additionally, we have related the determination of the worst-case realization of the submodular robust uncertainty set to the maximization of a linear function over a polymatroid. For the RRKP, we have given ILP formulations and polyhedral results while focusing on the special cases of the $k, \ell/D$-RRKP and k/Γ-RRKP. In particular, we have considered the MWSP subproblem of the k/Γ-RRKP, provided a combinatorial polynomial-time algorithm solving it, and thus presented a new compact reformulation of the k/Γ-RRKP. Moreover and besides developing several separation algorithms, we have discussed how to solve the k/Γ-RRKP by means of robustness cuts. Our studies of robust KPs have been completed by computational studies for both considered variants of the RRKP showing the gain of recovery, the effectiveness of (strengthened (extended)) cover inequalities, and comparing the different approaches based on robustness cuts.

Our studies show that a significant increase in profit (gain of recovery) is achieved by applying the recoverable robustness concept. Although, extended cover inequalities are effectively tightening the LP formulation, our proposed ILP-based separation is too slow in practice. Nevertheless for the k/Γ-RRKP, our algorithm SepU is the fastest clearly outperforming standard ILP solvers.

In addition to the robust KPs, we have studied the Γ-RNDP in great detail. We have presented the following: several MILP formulations including the recent capacity formulation, an intensive study of the polyhedral structure including new classes of valid and to some extend facet-defining inequalities (Γ-robust cutset inequalities, Γ-robust envelope inequalities, Γ-robust arc residual capacity inequalities, and Γ-robust metric inequalities with several subclasses), and algorithms solving the corresponding separation problems. In addition and for the first time, the mb-RNDP has been studied and we have exemplarily shown how results for the Γ-RNDP can be generalized to the mb-RNDP. In particular, our results include a capacity ILP formulation of the mb-RNDP generalizing the so-called "Japanese Theorem" to the multi-band robust setting, as well as the generalization of cutset inequalities and the derivation of Γ-robust envelope inequalities. Our studies of robust NDPs have also been completed by extensive computational experiments of the Γ-RNDP. Here, we have used real-life traffic measurements of telecommunication backbone networks to derive representative realistic instances. In our experiments, we have addressed the problem of parametrizing the uncertainty set by historical data, compared three different formulations of the Γ-RNDP, and investigated the effectiveness of nine different (sub)classes of valid inequalities. Furthermore, we have proposed several algorithms to solve the Γ-RNDP and evaluated their performance for both, the compact link flow and the capacity formulation of the problem. Moreover, we have highlighted the handling of larger network instances in practice. Finally, we have evaluated the realized robustness of the obtained optimal robust solutions w.r.t. the given historical data.

In particular, our computational results show the effectiveness of Γ-robust cutset inequalities in a cut-and-branch approach. Moreover, we have observed that the capacity formulation is oftentimes more tractable than the other formulations.

Altogether, the crucial question remains which robustness concept should be used. Based on our work, we are convinced that there does not exist a best concept in general. Instead, the choice of the robustness concept must be in line with the considered problem. We have investigated robustness concepts which are widely applicable, especially Γ-robust and multi-band robust uncertainty sets are well-suited if historical data is given. The recoverable robustness concept offers additional adaptability which allows the modeling of different time scales (first stage and second stage correspond to long-term and short-term, respectively), or different planning options as "rent" (first stage decision) or "buy" (second stage decision). The concept of submodular robustness is less applicable in practice but highlights the underlying common structures of many robustness concepts.

For each choice of concept, we have highlighted its individual properties and provided mathematical programming formulations, polyhedral insights, and methods to solve the considered robust KPs and robust NDPs in this thesis.

Outlook As a final outlook, some research directions for future investigations are mentioned. The concepts of submodular robustness and recoverable robustness should be applied to the NDP, the related polyhedra should be investigated, and their structure exploited to develop effective algorithms. The study of a submodular robust NDP should in particular include the investigation of the possible generalizations of results from the Γ-RNDP. Furthermore, the relation between submodular robustness and polymatroids should be the focus of further investigation.

LIST OF TABLES

LIST OF FIGURES

BIBLIOGRAPHY

[1] GÉANT project. URL `http://www.geant.net`. Cited on page *189*.

[2] NETLIB. URL `http://www.netlib.org/`. Cited on page *35*.

[3] ROBUKOM - Robust Communication Networks. BMBF grant 03MS616A. URL `http://www.robukom.de`. Cited on pages *34 and 189*.

[4] NOBEL – Next generation Optical networks for Broadband European Leadership, 2004–2007. URL `http://www.ist-nobel.org`. Cited on page *189*.

[5] Abilene Internet2 Network. URL `http://www.internet2.edu/network`. Cited on pages *31, 32, 33, and 189*.

[6] Tobias Achterberg. *Constraint integer programming*. PhD thesis, Technische Universität Berlin, 2007. Cited on pages *7 and 14*.

[7] Yogesh K. Agarwal. k-Partition-based facets of the network design problem. *Networks*, 47(3):123–139, 2006. Cited on pages *26, 154, and 165*.

[8] Yogesh K. Agarwal. Polyhedral structure of the 4-node network design problem. *Networks*, 54:139–149, October 2009. ISSN 0028-3045. doi: 10.1002/net.v54:3. Cited on pages *26 and 165*.

[9] Shabbir Ahmed and Alper Atamtürk. Maximizing a class of submodular utility functions. *Mathematical Programming*, 128:149–169, 2011. Cited on page *85*.

[10] Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. Network flows: theory, algorithms, and applications. 1993. Cited on pages *7 and 21*.

[11] Ayşegül Altin, Eduardo Amaldi, Pietro Belotti, and Mustafa Ç. Pinar. Provisioning virtual private networks under traffic uncertainty. *Networks*, 49(1):100–155, 2007. Cited on page *137*.

[12] Ayşegül Altin, Hande Yaman, and Mustafa Ç. Pinar. The Robust Network Loading Problem under Hose Demand Uncertainty: Formulation, Polyhedral Analysis, and Computations. *INFORMS Journal on Computing*, 23(1):77–89, 2011. Cited on pages *42 and 137*.

[13] Alper Atamtürk. Flow Pack Facets of the Single Node Fixed Charge Flow Polytope. *Operations Research*, 29:107–114, 2001. Cited on page *26*.

[14] Alper Atamtürk. On capacitated network design cut-set polyhedra. *Mathematical Programming*, 92:425–437, 2002. Cited on page *25*.

[15] Alper Atamtürk and Vishnu Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008. Cited on pages *58 and 85*.

[16] Alper Atamtürk and Vishnu Narayanan. The submodular knapsack polytope. *Discrete Optimization*, 6(4):333–344, 2009. Cited on pages *58, 85, 87, 88, 89, and 121*.

[17] Pasquale Avella, Sara Mattia, and Antonio Sassano. Metric inequalities and the network loading problem. *Discrete Optimization*, 4:103–114, 2006. Cited on pages *21, 24, 26, 161, 164, 170, 172, and 173*.

[18] Egon Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8: 146–164, 1975. Cited on pages *14, 17, and 18*.

[19] Egon Balas and Robert Jeroslow. Canonical cuts on the unit hypercube. *SIAM Journal on Applied Mathematics*, 23(1):61–69, 1972. Cited on page *18*.

[20] Francisco Barahona. Network design using cut inequalities. *SIAM Journal on Optimization*, 6:823–837, March 1996. ISSN 1052-6234. Cited on pages *25 and 26*.

[21] John E. Beasley. OR-Library: distributing test problems by electronic mail. *Journal of the OR Society*, 41:1069–1072, 1990. Cited on page *122*.

[22] Pietro Belotti, Antonio Capone, Giuliana Carello, and Frederico Malucelli. Multi-layer MPLS network design: The impact of statistical multiplexing. *Computer Networks*, 52(6):1291–1307, 2008. Cited on page *137*.

[23] Pietro Belotti, Kireetti Kompella, and Lloyd Noronha. A comparison of OTN and MPLS networks under traffic uncertainty. *Working paper*, 2011. URL `http://myweb.clemson.edu/~pbelott/papers/robust-opt-network-design.pdf`. Cited on page *137*.

[24] Walid Ben-Ameur and Hervé Kerivin. Routing of uncertain demands. *Optimization and Engineering*, 3:283—313, 2005. Cited on page *137*.

[25] Aharon Ben-Tal and Arkadi Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:769–805, 1998. Cited on pages *38, 41, and 42*.

[26] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–14, 1999. Cited on page *41*.

[27] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88: 411–424, 2000. Cited on pages *35 and 36*.

[28] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. Robustness. In *Handbook of Semidefinite Programming*, pages 139–162, 2000. Cited on page *42*.

[29] Aharon Ben-Tal, Arkadi Nemirovski, and Cornelis Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13:535–560, 2002. Cited on page *41*.

[30] Aharon Ben-Tal, Alexander Goryashko, E. Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004. Cited on pages *43 and 140*.

[31] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, 2009. ISBN 9780691143684. Cited on pages *35, 36, and 37*.

[32] Jacques F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, (1):238–252, 1962. Cited on page *24*.

[33] Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98(1):49–71, 2003. Cited on pages *42, 46, 50, 51, 71, and 141*.

[34] Dimitris Bertsimas and Melvyn Sim. The Price of Robustness. *Operations Research*, 52(1):35–53, 2004. Cited on pages *42, 46, 50, 51, 57, 64, 71, 121, and 141*.

[35] Dimitris Bertsimas and John N. Tsitsiklis. Introduction to linear optimization. 1997. Cited on page *43*.

[36] Dimitris Bertsimas, David B. Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011. Cited on pages *36, 38, 41, 43, and 44*.

[37] Daniel Bienstock. Histogram models for robust portfolio optimization. *Journal on Computational Finance*, 11:1–64, 2007. Cited on pages *51 and 79*.

[38] Daniel Bienstock and Fabio D'Andreagiovanni. Robust Wireless Network Planning. In *Proceedings AIRO 2009, the 40th Annual Conference of the Italian Operational Research Society*, pages 131–132, 2009. Cited on pages *52 and 175*.

[39] Daniel Bienstock and Oktay Günlük. Capacitated network design – polyhedral structure and computation. *INFORMS Journal on Computing*, 8:243–259, 1996. Cited on pages *21 and 25*.

[40] Daniel Bienstock, Sunil Chopra, Oktay Günlük, and Chih-Yang Tsai. Minimum cost capacity installation for multicommodity network flows. *Mathematical Programming*, 81:177–199, 1998. Cited on pages *25 and 167.*

[41] John R. Birge and Francois Louveaux. *Introduction to stochastic programming.* Springer Verlag, 1997. Cited on page *43.*

[42] Andreas Bley, Ulrich Menne, Roman Klähne, Christian Raack, and Roland Wessäly. Multi-layer network design – a model-based optimization approach. In *Proc. of the Polish-German Teletraffic Symposium (PGTS) 2008*, pages 107–116, Berlin, Germany, 2008. Polish-German Teletraffic Symposium. Cited on pages *33 and 137.*

[43] Karl-Heinz Borgwardt. The average number of pivot steps required by the simplex-method is polynomial. *Mathematical Methods of Operations Research*, 26(1): 157–177, 1982. Cited on page *12.*

[44] Paul C. Bouman, J. Marjan Van Den Akker, and J. A. Hoogeveen. Recoverable robustness by column generation. In *Proceedings of the 19th European conference on Algorithms*, ESA'11, pages 215–226, Berlin, Heidelberg, 2011. Springer-Verlag. ISBN 978-3-642-23718-8. Cited on page *95.*

[45] E. Andrew Boyd. Polyhedral results for the precedence constraint knapsack problem. *Discrete Applied Mathematics*, 41:185–201, 1993. Cited on page *15.*

[46] Christina Büsing. *Recoverable Robustness in Combinatorial Optimization.* PhD thesis, Technische Universität Berlin, 2011. Cited on pages *44, 62, 95, and 121.*

[47] Christina Büsing and Fabio D'Andreagiovanni. New Results about Multi-band Uncertainty in Robust Optimization. In R. Klasing, editor, *Experimental Analysis — SEA 2012*, volume 7276 of *LNCS*, pages 63–74. Springer, 2012. Cited on pages *52, 55, 57, 79, 80, and 181.*

[48] Christina Büsing and Fabio D'Andreagiovanni. Robust Optimization under Multi-band Uncertainty — Part I: Theory. January 2013. URL `http://arxiv.org/abs/1301.2734v1`. Cited on pages *52, 54, 56, 57, and 80.*

[49] Christina Büsing, Arie M.C.A. Koster, and Manuel Kutschka. Recoverable Robust Knapsacks: the Discrete Scenario Case. *Optimization Letters*, 5(3):379–392, 2011. Cited on pages *62, 95, 97, 105, 113, and 121.*

[50] Christina Büsing, Arie M.C.A. Koster, and Manuel Kutschka. Recoverable Robust Knapsacks: *Gamma*-Scenarios. In *Proceedings of INOC 2011, International Network Optimization Conference*, volume 6701 of *Lecture Notes on Computer Science*, pages 583–588. Springer, 2011. Cited on pages *62, 95, 99, 105, and 121.*

[51] Chandra Chekuri, Gianpaolo Oriolo, Maria G. Scutellá, and F. Bruce Shepherd. Hardness of robust network design. *Networks*, 50(1):50–54, 2007. Cited on page *137.*

[52] Sunil Chopra, Itzhak Gilboa, and S. Trilochan Sastry. Source sink flows with capacity installation in batches. *Discrete Applied Mathematics*, 86:165–192, 1998. Cited on pages *21 and 25*.

[53] Vasek Chvatal. *Linear programming*. WH Freeman, 1983. Cited on page *7*.

[54] Cisco Systems Inc. Cisco Visual Networking Index: Forecast and Methodology, 2012-2017, 2013. URL `http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html`. Cited on pages *32 and 34*.

[55] Grit Claßen, Arie M.C.A. Koster, Manuel Kutschka, and Issam Tahiri. Robust Metric Inequalities for the Γ-Robust Network Loading Problem. *Networks*, 2012. submitted. Cited on pages *138, 161, and 189*.

[56] Stephen Cook. *The millennium prize problems*, chapter The P versus NP Problem, pages 87–106. American Mathematical Society, 2006. Cited on page *9*.

[57] R. J. Dakin. A tree-search algorithm for mixed integer programming problems. *Computer Journal*, 8(3):250–255, 1965. Cited on page *12*.

[58] Fabio D'Andreagiovanni. *Pure 0-1 Programming approaches to Wireless Network Design*. PhD thesis, Universita "La Sapienza" di Roma, via Ariosto, 25, 00185 Rome, Italy, 2011. Cited on pages *52, 79, and 175*.

[59] George B. Dantzig. Maximization of a linear function of variables subject to linear inequalities. In *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13, pages 339–347. John Wiley & Sons Inc., New York, N. Y., 1951. Cited on page *12*.

[60] George B. Dantzig. Linear programming under uncertainty. *Management Science*, pages 197–206, 1955. Cited on page *36*.

[61] George Bernard Dantzig. *Linear programming and extensions*. Princeton university press, 1998. Cited on page *7*.

[62] Krystof Dudzinski and Stanisław Walukiewicz. Exact methods for the knapsack problem and its generalizations. *European Journal of Operational Research*, 28: 3–21, 1987. Cited on page *15*.

[63] Nick G. Duffield, Pawan Goyal, Albert G. Greenberg, Partho Pratim Mishra, K. K. Ramakrishnan, and Jacobus E. van der Merive. A flexible model for resource management in virtual private networks. In *Proc. of the ACM SIGCOMM*, pages 95–108, 1999. Cited on page *137*.

[64] Svetoslav Duhovnikov, Arie M.C.A. Koster, Manuel Kutschka, Frank Rambach, and Dominic Schupke. Γ-Robust Network Design for Mixed-Line-Rate-Planning of Optical Networks. In *Proceedings of the National Fiber Optic Engineers Conference (OFC 2013)*. Optical Society of America, 2013. Cited on page *31*.

[65] Anurag Dwivedi and Richard E. Wagner. Traffic model for USA long distance optimal network. In *Proc. of the Optical Fiber Communication Conference*, pages 156–158, 2000. Cited on pages *33 and 137*.

[66] Jack Edmonds. *Combinatorial Structures and their Applications*, chapter Submodular functions, matroids, and certain polyhedra, pages 69–87. Gordon and Breach, 1970. Cited on page *58*.

[67] Matthias Ehrgott. *Multicriteria Optimization*. Springer, 2005. ISBN 3540213988. Cited on page *15*.

[68] Carlos E. Ferreira, Alexander Martin, and Robert Weismantel. Solving multiple knapsack problems by cutting planes. *SIAM Journal on Optimization*, 6:858–877, 1996. Cited on page *19*.

[69] Matteo Fischetti and Michele Monaci. Cutting plane versus compact formulations for uncertain (integer) linear programs. *Mathematical Programming Computation*, 4(3):239–273, 2012. Cited on pages *49, 50, 78, 117, 118, and 195*.

[70] Andras Frank and Eva Tardos. Generalized Polymatroids and Submodular Flows. *Mathematical Programming*, (42):489–563, 1988. Cited on page *58*.

[71] Giorgio Gallo and B. Simeone. On the supermodular knapsack problem. *Mathematical Programming*, 45:295–309, 1989. Cited on page *85*.

[72] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. ISBN 0716710447. Cited on pages *7, 8, 9, and 16*.

[73] Ralph E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64:275–278, 1958. Cited on page *14*.

[74] Ralph E. Gomory. An Algorithm for Integer Solutions to Linear Programs. In R. L. Graves and P. Wolfe, editors, *Recent Advances in Mathematical Programming*, pages 269–302. McGraw-Hill, New York, 1963. Cited on page *14*.

[75] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981. Cited on pages *12 and 13*.

[76] Martin Grötschel, László Lovász, and Alexander Schrijver. Geometric algorithms and combinatorial optimization. 1988. Cited on page *7*.

[77] Zonghao Gu, George L. Nemhauser, and Martin W.P. Savelsbergh. Lifted cover inequalities for 0-1 integer programs: computation. *INFORMS Journal on Computing*, 10:427–437, 1998. Cited on page *18*.

[78] Zonghao Gu, George L. Nemhauser, and Martin W.P. Savelsbergh. Lifted cover inequalities for 0-1 integer programs: complexity. *INFORMS Journal on Computing*, 11:117–123, 1999. Cited on pages *18 and 19*.

[79] Oktay Günlük. A branch and cut algorithm for capacitated network design problems. *Mathematical Programming*, 86:17–39, 1999. Cited on pages *21, 25, and 167*.

[80] Peter L. Hammer, Ellis L. Johnson, and Uri N. Peled. Facet of regular 0-1 polytopes. *Mathematical Programming*, 8:179–206, 1975. Cited on pages *14 and 18*.

[81] Dorit S. Hochbaum. A nonlinear knapsack problem. *Operations Research Letters*, 17:103–110, 1995. Cited on page *16*.

[82] Ellis Horowitz and Sartaj Sahni. Computing partitions with applications to the knapsack problem. *Journal of ACM*, 21:277–292, 1974. Cited on page *15*.

[83] Oscar H. Ibarra and Chul E. Kim. Fast approximation algorithms for the knapsack and sum of subset problems. *Journal of ACM*, 22:463–468, 1975. Cited on page *15*.

[84] IBM ILOG. CPLEX version 12.1, 2009. URL `http://www.ilog.com/products/cplex`. Cited on pages *78, 122, and 189*.

[85] M. Iri. On an Extension of the Maximum-flow Minimum-cut Theorem to Multicommodity Flows. *Journal of the Operations Research Society of Japan*, 13(3): 129–135, 1971. Cited on page *24*.

[86] Garud N Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2):257–280, 2005. Cited on page *43*.

[87] David S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9:256–278, 1974. Cited on page *15*.

[88] Ravindran Kannan and Clyde L Monma. *On the computational complexity of integer programming problems*. Springer, 1978. Cited on page *12*.

[89] Konstantinos Kaparis and Adam N. Letchford. Separation algorithms for 0-1 knapsack polytopes. *Mathematical Programming*, 124(1–2):69–91, 2010. Cited on pages *18 and 19*.

[90] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311. ACM, 1984. Cited on page *12*.

[91] Richard M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972. Cited on pages *15 and 16*.

[92] Hans Kellerer, Ulrich Pferschy, and David Pisinger. *Knapsack Problems*. Springer, 2004. ISBN 3540402861. Cited on pages *15 and 16*.

[93] Leonid G. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR*, (224):1093–1096, 1979. (English translation: Soviet Math. Dokl. **20**, 191–194.). Cited on page *12*.

[94] Diego Klabjan, George L. Nemhauser, and Craig A. Tovey. The complexity of cover inequality separation. *Operations Research Letters*, 23:35–40, 1998. Cited on page *19*.

[95] Victor Klee and George J. Minty. How good is the simplex algorithm. Technical report, DTIC Document, 1970. Cited on page *12*.

[96] Oliver Klopfenstein and Dritan Nace. A robust approach to the chance-constraint knapsack problem. *Operations Research Letters*, 36(5):628–532, September 2008. Cited on page *71*.

[97] Oliver Klopfenstein and Dritan Nace. Cover inequalities for robust knapsack sets — Application to the robust bandwidth packing problem. *Networks*, 59(1):59–72, January 2012. Cited on pages *71, 73, 74, 75, 76, 77, 78, 110, 121, and 122*.

[98] P.J. Kolesar. A branch and bound algorithm for the knapsack problem. *Management Science*, 13:723–735, 1967. Cited on page *15*.

[99] Arie M.C.A. Koster and Manuel Kutschka. Multi-commodity Single Path Routing with submodular bandwidth consumption. In *Proceedings of INOC 2009*, 2009. Cited on page *85*.

[100] Arie M.C.A. Koster and Manuel Kutschka. An integrated model for survivable network design under demand uncertainty. In *Design of Reliable Communication Networks (DRCN), 2011 8th International Workshop on the*, pages 54–61. IEEE Xplore, 2011. Cited on page *31*.

[101] Arie M.C.A. Koster and Manuel Kutschka. Network Design Under Demand Uncertainties: A Case Study on the Abilene and GEANT network data. In *Tagungsband der 12. ITG-Fachtagung Photonische Netze*, volume 228, pages 154–161, Leipzig, Germany, 2011. VDE Verlag GmbH. Cited on pages *66, 68, and 137*.

[102] Arie M.C.A. Koster and X. Munoz, editors. *Graphs and Algorithms in Communication Networks*. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-02249-4. Cited on page *27*.

[103] Arie M.C.A. Koster, Manuel Kutschka, and Christian Raack. Towards Robust Network Design using Integer Linear Programming Techniques. In *Proceedings Next Generation Internet, NGI 2010*, pages 1–8. IEEE Xplore, 2010. Cited on pages *50, 66, 137, and 189*.

[104] Arie M.C.A. Koster, Manuel Kutschka, and Christian Raack. Cutset Inequalities for Robust Network Design. In *Proceedings of INOC 2011, International Network Optimization Conference*, volume 6701 of *Lecture Notes on Computer Science*, pages 118–123. Springer, 2011. Cited on page *166*.

[105] Arie M.C.A. Koster, Manuel Kutschka, and Christian Raack. On the robustness of optimal network designs. In *Proceedings IEEE International Conference on Communications, ICC 2011*, pages 1–5. IEEE Xplore, 2011. Cited on pages *68, 137, 189, 193, and 194*.

[106] Arie M.C.A. Koster, Manuel Kutschka, and Christian Raack. Robust network design: Formulations, valid inequalities, and computations. *Networks*, 61(2):128–149, 2013. Cited on pages *50, 137, 145, 189, and 198*.

[107] Panos Kouvelis and Gang Yu. *Robust discrete optimization and its applications.* Kluwer Academics Publishers, Norwell, MA, 1997. Cited on page *42*.

[108] A. H. Land and A. G. Doig. An automatic method for solving discrete programming problems. *Econometrica*, 28:493–520, 1960. Cited on page *12*.

[109] Benoit Lardeux, Dritan Nace, and Jérôme Geffard. Multiperiod network design with incremental routing. *Networks*, 50(1):109–117, 2007. Cited on page *137*.

[110] Christian Liebchen, Marco E. Lübbecke, Rolf H. Möhring, and Sebastian Stiller. The concept of recoverable robustness, linear programming recovery, and railway applications. In Ravindra Ahuja, Rolf Möhring, and Christos Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *LNCS*, pages 1–27. Springer, 2009. ISBN 978-3-642-05464-8. Cited on pages *44, 61, and 62*.

[111] Thomas L. Magnanti and Prakash Mirchandani. Shortest paths, single origin-destination network design and associated polyhedra. *Networks*, 33:103–121, 1993. Cited on pages *25 and 146*.

[112] Thomas L. Magnanti and Richard T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18(1):1–55, 1984. Cited on page *21*.

[113] Thomas L. Magnanti, Prakash Mirchandani, and Rita Vachani. The convex hull of two core capacitated network design problems. *Mathematical Programming*, 60: 233–250, 1993. Cited on pages *25, 26, and 156*.

[114] Thomas L. Magnanti, Prakash Mirchandani, and Rita Vachani. Modelling and solving the two-facility capacitated network loading problem. *Operations Research*, 43:142–157, 1995. Cited on pages *21 and 25*.

[115] Silvano Martello and Paolo Toth. The 0-1 knapsack problem. In N. Christofides, A. Mingozzi, Paolo Toth, and C. Sandi, editors, *Combinatorial Optimization*, pages 237–279. Wiley, 1979. Cited on page *15*.

[116] Silvano Martello and Paolo Toth. Optimal and canonical solutions of the change-making problem. *European Journal of Operational Research*, 4:322–329, 1980. Cited on page *15*.

[117] Silvano Martello and Paolo Toth. Solution of the zero-one multiple knapsack problem. *European Journal of Operational Research*, 4:322–329, 1980. Cited on page *15*.

[118] Silvano Martello and Paolo Toth. Algorithms for knapsack problems. In Silvano Martello, G. Laporte, and M. Minoux, editors, *Surveys in Combinatorial Optimization*, Annals of Discrete Mathematics 31, pages 213–257. North-Holland, 1987. Cited on page *15*.

[119] Silvano Martello and Paolo Toth. *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & Sons, 1990. ISBN 0471924202. Cited on pages *15, 16, 17, 19, and 112*.

[120] Kamlesh Mathur, Harvey M. Salkin, and Bidhu B. Mohanty. A note on a general non-linear knapsack problem. *Operations Research Letters*, 5:79–81, 1986. Cited on page *16*.

[121] Sara Mattia. The robust network loading problem with dynamic routing. *Computational Optimization and Applications*, 2012. doi: 0.1007/s10589-012-9500-0. Cited on pages *25, 26, 138, 142, 144, 164, and 170*.

[122] Sara Mattia. Robust optimization with multiple intervals. technical report R. 7,2012, Istituto di Analisi dei Sistemi ed Informatica (IASI), Consiglio Nazionale delle Ricerche (CNR), viale Manzoni 30, 00185 Rome, Italy, 2012. Cited on pages *52, 57, 79, 80, and 175*.

[123] Andrew J. Miller and Laurence A. Wolsey. Tight formulations for some simple mixed integer programs and convex objective integer programs. *Math. Programming*, 98:73–88, 2003. Cited on page *152*.

[124] Michele Monaci and Ulrich Pferschy. On the Robust Knapsack Problem. April 2011. URL `http://www.optimization-online.org/DB_HTML/2011/04/3019.html`. Cited on page *71*.

[125] Michele Monaci, Ulrich Pferschy, and Paolo Serafini. Exact solution of the robust knapsack problem. 2013. URL `http://www.sciencedirect.com/science/article/pii/S0305054813001342`. Cited on pages *71, 72, 78, and 121*.

[126] George L. Nemhauser and Laurence A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience Series in Discrete Mathematics and Optimization. New York, 1988. Cited on pages *7 and 14*.

[127] Arkadi Nemirovski. Several np-hard problems arising in robust stability analysis. *Mathematics of Control, Signals, and Systems*, 6:99–105, 1993. Cited on page *42*.

[128] Arnab Nilim and Laurent El Ghaoui. *Robust markov decision processes with uncertain transition matrices.* PhD thesis, University of California, Berkeley, 2004. Cited on page *43*.

[129] K. Onaga and O. Kakusho. On feasibility conditions of multicommodity flows in networks. *Circuit Theory, IEEE Transactions on*, 18(4):425 – 429, July 1971. doi: 10.1109/TCT.1971.1083312. Cited on page *24*.

[130] Sebastian Orlowski, Michal Pióro, Arthur Tomaszewski, and Roland Wessäly. SNDlib 1.0–Survivable Network Design Library. In *Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium*, April 2007. URL `http://sndlib.zib.de`. Cited on pages *29, 30, 190, and 191*.

[131] Sebastian Orlowski, Roland Wessäly, Michal Pióro, and Arthur Tomaszewski. SNDlib 1.0–Survivable Network Design Library. *Networks*, published online, 2009. URL `http://dx.doi.org/10.1002/net.20371`. Cited on pages *30, 189, 190, and 191*.

[132] Adam Ouorou and Jean-Philippe Vial. A model for robust capacity planning for telecommunications networks under demand uncertainty. In *Design and Reliable Communication Networks, 2007. DRCN 2007. 6th International Workshop on*, pages 1 –4, oct. 2007. doi: 10.1109/DRCN.2007.4762287. Cited on pages *137 and 140*.

[133] Manfred W. Padberg. A note on zero-one programming. *Operations Research*, 23: 833–837, 1975. Cited on page *18*.

[134] Manfred W. Padberg. $(1, k)$-configurations for facets for packing problems. *Mathematical Programming*, 18:94–99, 1980. Cited on pages *19, 89, 91, and 209*.

[135] Manfred W Padberg, Tony J Van Roy, and Laurence A. Wolsey. Valid linear inequalities for fixed charge problems. *Operations Research*, 33(4):842–861, 1985. Cited on page *26*.

[136] Michal Pióro and Deepankar Medhi. *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan-Kaufmann Publishers (Elsevier), 2004. ISBN 0-12-557189-5. Cited on pages *20, 21, 24, and 27*.

[137] David Pisinger. Where are the hard knapsack problems? *Computers & Operations Research*, 32:2271–2284, 2005. Cited on pages *19 and 122*.

[138] David Pisinger. The quadratic knapsack problem – a survey. *Discrete Applied Mathemtics*, 155(5):623–648, March 2007. Cited on page *16*.

[139] Michael Poss and Christian Raack. Affine recourse for the robust network design problem: between static and dynamic routing. Technical Report ZR-11-03, Zuse Institute Berlin (ZIB), Berlin, Germany, 2011. URL `http://www.`

optimization-online.org/DB_HTML/2011/02/2917.html. Cited on pages *137, 138, and 140*.

[140] Christian Raack. *Capacitated Network Design - Multi-Commodity Flow Formulations, Cutting Planes, and Demand Uncertainty*. PhD thesis, Technische Universität Berlin, 2012. Cited on page *137*.

[141] Christian Raack, Arie M.C.A. Koster, Sebastian Orlowski, and Roland Wessäly. On cut-based inequalities for capacitated network design polyhedra. *Networks*, 2 (57):141–156, 2011. Cited on pages *25, 154, and 167*.

[142] Harvey M. Salkin and Cornelis A. De Kluyver. The knapsack problem: a survey. *Naval Research Logistics Quarterly*, 22:127–144, 1975. Cited on page *15*.

[143] Alexander Schrijver. *Theory of linear and integer programming*. Wiley, 1998. Cited on page *7*.

[144] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Verlag, 2003. Cited on page *7*.

[145] Alexander Shapiro, Darinka Dentcheva, and Adrezej Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory*. MPS-SIAM Series on Optimization. SIAM – Society for Industrial and Applied Mathematics, 2009. ISBN 9780898716870. Cited on pages *36 and 37*.

[146] Prabhakant Sinha and Andris A. Zoltners. The multiple-choice knapsack problem. *Operations Research*, 27(3):503–515, 1979. Cited on page *15*.

[147] Allen L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21:1154–1157, 1973. Cited on page *37*.

[148] Daniel Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 296–305. ACM, 2001. Cited on page *12*.

[149] W. Wallace Stein and William T. Ziemba. *Applications of Stochastic Programming*. MPS-SIAM Series on Optimization. MPS-SIAM Series on Optimization, 2005. ISBN 9780898715552. Cited on page *37*.

[150] Mechthild Stoer and Geir Dahl. A Polyhedral Approach to Multicommodity Survivable Network Design. *Numerische Mathematik*, 68:149–167, 1993. Cited on page *20*.

[151] Maxim Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, 2004. Cited on page *85*.

[152] S. E. Terblanche, Roland Wessäly, and Johannes M. Hattingh. Lagrangian relaxation as a solution approach to solving the survivable multi-hour network design problem. In *Proc. of the South African Telecommuniations and Networks Applications Conference - SATNAC 2007*, 2007. Cited on page *137*.

[153] Steve Uhlig, Bruno Quoitin, Jean Lepropre, and Simon Balon. Providing public intradomain traffic matrices to the research community. *ACM SIGCOMM Computer Communication Review*, 36(1):83–86, 2006. Cited on page *137*.

[154] Tony J. Van Roy and Laurence A. Wolsey. Solving mixed integer programming problems using automatic reformulation. *Operations Research*, 35:45–57, 1987. Cited on page *18*.

[155] Verein zur Förderung eines Deutschen Forschungsnetzes e. V. (DFN-Verein). URL `http://www.dfn.de`. Cited on page *189*.

[156] H. Martin Weingartner and David N. Ness. Methods for the solution of the multidimensional 0/1 knapsack problem. *Operations Research*, 15(1):83–103, 1967. Cited on page *15*.

[157] Robert Weismantel. On the 0/1 knapsack polytope. *Mathematical Programming*, 77:49–68, 1997. Cited on page *19*.

[158] Laurence A. Wolsey. Facets for a linear inequality in 0-1 variables. *Mathematical Programming*, 8(1):165–178, 1975. Cited on pages *14 and 18*.

[159] Laurence A. Wolsey. Integer programming. *IIE Transactions*, 32:273–285, 2000. Cited on pages *7 and 14*.

[160] Huan Xu, Shie Mannor, B Schölkopf, J Platt, and T Hofmann. The robustness-performance tradeoff in Markov decision processes. *Advances in Neural Information Processing Systems*, 19:1537, 2007. Cited on page *43*.

[161] Gang Yu. On the max-min 0-1 knapsack problem with robust optimization applications. *Operations Research*, 44:407–415, 1996. Cited on page *97*.

[162] Yin Zhang. 6 months of Abilene traffic matrices, August 2009. URL `http://www.cs.utexas.edu/~yzhang/research/AbileneTM/`. Cited on pages *32, 137, and 189*.

INDEX

F

G

I

J

K

L

M

ABOUT THE AUTHOR

Manuel Kutschka received his diploma degree in Mathematics and Business Studies (Wirtschaftsmathematik) from Technische Universität Berlin in 2007 and was awarded his Ph.D. with summa cum laude from RWTH Aachen University in 2013.

His research interests are in robust optimization, (telecommunication) network design, and methods and algorithms of (mixed) integer programming. He is a member of the Mathematical Optimization Society, former member of INFORMS and IEEE, and reviewer for several international journals.

During his studies in Berlin, Manuel was a student research assistant at the Optimization department of Zuse Institute Berlin focussing on cutting plane algorithms for solving mixed integer programs. In 2008 he worked as graduate research assistant on dynamic network admission algorithms in the Networks Research Centre of British Telecommunications in Ipswich, UK. In the same year he began his investigation of robust optimization as a Doctoral researcher at the Warwick Business School working in the Centre for Discrete Mathematics and its Applications (DIMAP), University of Warwick, UK. In 2009 he transferred his research to RWTH Aachen University and became a research assistant in the discrete optimization group of Arie Koster at RWTH Aachen University. Here, Manuel continued his research on robust optimization studying new robustness concepts, investigating the polyhedral structure of robust knapsack and network design problems, developing and evaluating (separation) algorithms, and adding telecommunications as application area while working in the German ROBUKOM project in 2010–2013.