Optimisation under Data Uncertainty in Wireless Communication Networks

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

vorgelegt von

Diplom-Computermathematikerin
Grit Claßen

aus Köthen

Berichter: Univ.-Prof. Dr. Ir. Arie M.C.A. Koster
Univ.-Prof. Dr.-Ing. Anke Schmeink

Tag der mündlichen Prüfung: 04. März 2015

Vorwort

Diese Dissertation entstand während meiner Zeit am Lehrstuhl II für Mathematik, Lehr- und Forschungsgebiet „Diskrete Optimierung“, und am Lehrstuhl für Theoretische Informationstechnik, Lehr- und Forschungsgebiet „Informationstheorie und Systematischer Entwurf von Kommunikationssystemen“, an der RWTH Aachen. Für die Unterstützung, die ich in dieser Zeit erfahren habe, möchte ich mich bei einigen Menschen bedanken.

Zunächst möchte ich mich bei Arie Koster und Anke Schmeink für das Ermöglichen einer interdisziplinären Promotion bedanken. Auch wenn die Aufteilung zwischen zwei Lehrstühlen nicht immer ganz einfach war, bin ich dankbar für die abwechslungsreiche Zusammenarbeit und die erfahrene Unterstützung.

Ein besonderer Dank gilt meinen Bürokollegen Simon Görtzen, Manuel Kutschka und Stephan Lemkens für die unzähligen konstruktiven fachlichen und privaten Diskussionen. Danke für eure Begleitung durch sämtliche Hochs und Tiefs der Promotion! Meinen Kolleginnen und Kollegen am Lehrstuhl II für Mathematik, vom UMIC Research Centre und am Lehrstuhl für Theoretische Informationstechnik danke ich für die angenehme Arbeitsatmosphäre und den inhaltlichen Austausch.

Für das Korrekturlesen dieser Arbeit und für die wertvollen Verbesserungsvorschläge auch bezüglich meines Abschlussvortrags bedanke ich mich bei Christina Büsing, Arie Koster und Manuel Kutschka.

Meiner Familie und meinen Freunden danke ich für ihre Unterstützung, ihr offenes Ohr und ihre Nachsicht, wenn ich mal wieder wochenlang nichts von mir hören lies. Danke für eure Begleitung durch eine schöne, aufregende, aber auch schwierige Zeit.

Mein größter Dank jedoch gilt meinem „Mann“ Michael Ecker. Danke für deine Geduld, deine Nachsicht und dass du meinem allabendlichen Blubbern immer wieder Aufmerksamkeit schenkst.

Aachen, im März 2015 Grit Claßen

Contents

Introduction

Optimisation is an everyday term used frequently to describe all sorts of improvements and enhanced modifications. However, optimisation also describes one specific field of mathematics which provides theory, models and methods to formalise, handle, and solve complex problems emerging from various applications such as logistics, production or telecommunication.

One sector of telecommunication are wireless communication networks which still gain more and more importance, for instance, due to already six billion mobile cellular telephone subscribers world wide [40] with an upward trend; the world population numbers seven billion. Mathematical optimisation is in particular capable of tackling one major challenge inherent to all sorts of wireless communication networks, which are scarce resources. These resources such as frequency spectrum or data rate have to be subdivided as efficiently as possible. An optimisation problem that models a limited budget constraint is the so-called knapsack problem, which is one of the most fundamental and diverse problems in discrete optimisation.

A crucial aspect of modelling real-life problems mathematically are the simplifications necessary to standardise complex processes. A common simplification is the disregard of dynamics and uncertainties intrinsic to numerous real-world applications. For example, in wireless communication networks the performance of radio links is prone to variations due to external factors like weather, users are not static but move, and also demands such as bit rate requirements fluctuate especially by the daytime.

Stochastic and robust optimisation are special fields of mathematical optimisation providing methodologies to handle data uncertainty. These two approaches are suitable for different types of randomness. If the uncertain data obeys a previously known probability distribution, this version of uncertainty can be modelled by stochastic optimisation. Conversely, robust optimisation can handle uncertain data which is, for instance, based on a finite discrete set of historical information. The knowledge of a probability distribution is not essential for this methodology.

The aim of robust optimisation is to find an optimal solution, which is feasible for every realisation of uncertain data that lies in a previously defined uncertainty set. For many variants, the theoretical complexity of the non-robust problem is thereby not increased since there exist compact reformulations, which are only polynomially larger but still computationally tractable.

The object of research of this thesis is the study of diverse robust optimisation approaches as well as one variant of stochastic optimisation, which is related to robust optimisation, in the context of wireless network planning. The application of robustness to radio communication requires the development of precise sophisticated models and algorithms to solve the problems. Apart from this modelling aspect, the performance improve-

ments proposed in this thesis have a wide application spectrum for a variety of problems. Additionally, we provide a novel complexity result for a robust version of the knapsack problem. All theoretical investigations are supported by several (comprehensive) computational studies performed on realistic test instances.

Contributions Parts of this thesis are based on prior publications and numerous conference talks which originated from various collaborations. Those previously published works comprise the following. In [46, 47, 51], we have studied chance constraints in the framework of fixed broadband wireless networks, while we have investigated Γ-robustness for the planning of wireless access networks in [48, 49, 50]. Moreover, in the work [52], which is currently under peer-review, we have achieved new results for the multi-band robust knapsack problem. Finally, in [53] we have investigated robust metric inequalities for fixed networks.

The main contributions of this thesis comprise theoretical as well as applied achievements and are briefly summarised in the following listing.

- The modification of existing and the development of novel approaches to model interference in wireless access networks approximately as well as exactly.
- The modelling of link outages in fixed broadband wireless networks via chance constraints obtaining reliable solutions and the development of valid inequalities and a primal heuristic improving miscellaneous formulations.
- The incorporation of demand uncertainties in the wireless network planning problem by means of Γ-robust optimisation economising base stations and a detailed study of this robustness concept and the enhancements gained by valid inequalities.
- The design of a branch-and-price algorithm for the Γ-robust network planning problem, including the development of various performance improvements, which yields a better linear program solution than the corresponding compact formulation.
- The modelling of multi-band robustness in the context of a wireless network planning problem including a numerical study of this robustness approach, which demonstrates the achievement of further economisations of base stations.
- The proof of novel complexity results for the multi-band robust knapsack problem via dynamic programming algorithms.
- The development of a Lagrangian relaxation approach for a subproblem of the two-band robust wireless network planning problem which exploits a dynamic program for the multi-band robust knapsack problem and provides a fast algorithm to obtain good approximate solutions.
- The application of recoverable robustness to the wireless network planning problem with the result of energy savings by the temporary deactivation of base stations during low traffic times.

Structure This thesis is subdivided into the following five parts.

In *Part I. Mathematical Preliminaries and Wireless Communication Networks*, we introduce the mathematical as well as technical concepts, which form the basis of this thesis, including respective literature overviews on related work and present formulations for the two types of wireless communication networks investigated in this thesis. In Chapter 1, we introduce the notation and the basics of mathematical optimisation. Additionally, we summarise optimisation techniques applied in this thesis and state the classical knapsack problem including important known results. Chapter 2 gives a review of optimisation under data uncertainty comprising stochastic as well as robust methodologies. To apply the chance constraints and robust optimisation concepts introduced in this chapter to wireless communication networks, we summarise the relevant technical background in Chapter 3 while basic formulations of wireless cellular as well as fixed broadband wireless networks are given in Chapter 4. A crucial topic in wireless cellular networks is interference. Hence, in Chapter 5, we develop and study various approaches to model interference. This chapter represents a digression, which is not related to optimisation under data uncertainty but which is necessary for a complete description of cellular wireless networks.

In the subsequent part, *Part II. Chance Constraints*, we investigate chance constraints which typify a stochastic optimisation methodology to handle data uncertainty. We give a survey of related work and formalise this concept in Chapter 6. Then, in Chapter 7, we apply chance constraints to the fixed broadband wireless network planning problem to achieve reliable solutions, develop linear formulations and derive valid inequalities as well as a primal heuristic. These formulations and improvements are investigated in computational studies performed on the one hand, on grid networks and on the other hand, on realistic network instances. The computational studies reveal the gains of the valid inequalities via improved solving times or optimality gaps, the effectiveness of the primal heuristic, and a significant gain in reliability by our joint probability model.

We study one of the first and most commonly applied robust optimisation approaches in *Part III. Γ-Robustness*, which constitutes the most elaborate part of this thesis and is subdivided into three chapters. We first introduce the general concept of Γ-robust optimisation including the Γ-robust counterpart, possibilities to evaluate robustness and probability bounds in Chapter 8. Moreover, this chapter contains an introduction to the Γ-robust knapsack problem. In Chapter 9, we apply the Γ-robustness to the wireless network planning problem, on the one hand, in case of uncertain demands, and on the other hand, in case of uncertain spectral efficiencies. For the formulation with uncertain demands, we further derive valid inequalities, investigate their performance and the price of robustness and compare robust solutions to conventional planning, while we study the second formulation in terms of validity for interference modelling. We propose a reformulation of the uncertain demand model in Chapter 10 by means of a branch-and-price approach. Additionally, we develop performance improvements whose effectiveness is investigated in a computational study and compare the two formulations of the wireless network planning problem with uncertain demands in a further computational study, which concludes the part on Γ-robustness.

In *Part IV. Multi-Band Robustness*, we study a more general robustness concept. In Chapter 11, we summarise the basics and state the robust counterpart. Afterwards, we

present a theoretical study of the complexity of the multi-band robust knapsack problem in Chapter 12, which in particular entails a dynamic programming algorithm of practical applicability. In Chapter 13, we apply the multi-band robustness to the wireless network planning problem and evaluate the gain of multi-band versus Γ-robustness. Moreover, we use the dynamic program for the multi-band robust knapsack problem to efficiently solve the Lagrangian relaxation of a subproblem of the wireless network planning problem.

In *Part V. Recoverable Robustness*, we study a two-stage robust optimisation concept. The main idea including the recoverable robust counterpart and possibilities to evaluate the recovery action are summarised in Chapter 14. We apply this methodology to the wireless network planning problem in Chapter 15 and analyse the gain of recovery via a computational study revealing significant savings via the deactivation of base stations during low traffic times.

Finally, we present a critical discussion of the interpretation of computational studies and give general concluding remarks and topics for future research.

Part I.

Mathematical Preliminaries and Wireless Communication Networks

1. Basic Mathematical Concepts

In this chapter, we give a brief introduction to linear optimisation and the techniques applied in this thesis along with the used notation, some related work and some important results utilised later on. Furthermore, we present one of the most important optimisation problems, the knapsack problem, in the classic deterministic setting. This problem occurs consistently in the investigation of various uncertainty concepts and in applications considered in the remainder of this thesis.

1.1. Notation and Basics

First, we state some basic notation for sets of numbers. Let $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$ denote the sets of *real, integer* and *natural* numbers, respectively. The subsets of positive, non-negative, negative, or non-positive numbers are denoted by the subscripts > 0, ≥ 0, < 0, ≤ 0, respectively. For example, $\mathbb{Z}_{>0} = \mathbb{N}$.

Complexity theory As we will consider the complexity of some optimisation problems in the remainder of this thesis, we give a brief informal introduction to complexity theory in the following. Details and a formal introduction can be found, e. g., in the book of Garey and Johnson [80].

To characterise the (time) complexity of an algorithm, we use the following notation. A function $f(n)$ is in $O(g(n))$ if there exists a constant c such that $|f(n)| \leq c|g(n)|$ for all values $n \geq 0$. If the time complexity function $f(n)$ of an algorithm is $O(p(n))$ for a polynomial function p, the algorithm is called a *polynomial time algorithm*. If the time complexity function is bounded by the dimension and the magnitudes of the input data, the algorithm is called a *pseudo-polynomial time algorithm*.

Definition 1.1. *(Problem)* A *problem* is a general question that is to be answered and that possesses several parameters. By an *instance* we denote the specification of particular values for all parameters. If the only possible solutions (answers) are "yes" and "no", the problem is called a *decision problem*. If the answers are the minimum or the maximum value of a given objective function, the problem is an *optimisation problem*.

Note, for each minimisation (maximisation) problem, there exists a corresponding decision problem deciding whether there can be found a solution that is less (greater) than a given numerical bound.

In the following definition, we give a brief overview on important complexity classes.

Definition 1.2. *(Complexity classes)*

P	The class of all decision problems for which a polynomial time algorithm exists that decides for every instance if the solution is "yes" or "no".
NP	The class of all decision problems that can be solved by a polynomial time non-deterministic algorithm. A non-deterministic algorithm first guesses a solution and then verifies it in a second step. Moreover, a "yes" instance of a decision problem in NP can be verified in polynomial time.
NP-complete	The class of decision problems Π being in NP and every problem in NP can be polynomially transformed to Π. A polynomial transformation is also called polynomial time reduction and describes a polynomial time algorithm which attributes one problem π_1 to a second problem π_2. If there exists an algorithm to solve π_2, then also π_1 can be solved via the reduction.
strongly NP-complete	The class of NP-complete problems that cannot be solved by a pseudo-polynomial time algorithm unless P = NP.
NP-hard	The class of problems for which any NP-complete problem can be reduced to in polynomial time.

Polyhedral theory For $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, we define the *polyhedron*

$$P(A, b) := \{x \in \mathbb{R}^n \mid Ax \leq b\},$$

where $Ax \leq b$ is a system of linear inequalities. A bounded polyhedron is called *polytope*. By $\text{conv}(P(A, b) \cap \mathbb{Z})$, we denote the *convex hull* of all integer points of the polyhedron $P(A, b)$ which encompasses the set of all convex combinations of integer vectors in the polyhedron. The dimension $\dim(P(A, b))$ is the maximum number of affinely independent vectors in the polyhedron minus one. An inequality $\pi^t x \leq \rho$ with $\pi \in \mathbb{R}^n$ and $\rho \in \mathbb{R}$, is *valid* for $P(A, b)$ if $P(A, b) \subseteq P(\pi, \rho)$. The inequality is *facet-defining* if $\dim(\tilde{P}(\pi, \rho) \cap P(A, b)) = \dim(P(A, b)) - 1$ with $\tilde{P}(\pi, \rho) := \{x \in \mathbb{R}^n \mid \pi^t x = \rho\}$. We are interested in facet-defining (valid) inequalities as they describe the polyhedron $P(A, b)$ completely.

A linear optimisation problem that maximises a (linear) objective $c \in \mathbb{R}^n$ over the polyhedron $P(A, b)$ is called a *linear program* (LP) and has the form

$$\begin{aligned} &\max\ c^t x \\ &\text{s.t. } x \in P(A, b). \end{aligned}$$

The standard form of this LP plus non-negativity which is most commonly used in this

work is given as

$$\begin{aligned} &\max\ c^t x && (1.1a)\\ (P)\quad &\text{s.t. } Ax \leq b && (1.1b)\\ &x \in \mathbb{R}^n_{\geq 0}. && (1.1c) \end{aligned}$$

This problem has m constraints and n variables and is called the *primal problem (P)*. A solution vector $x \in \mathbb{R}^n_{\geq 0}$ is called *feasible* if it satisfies (1.1b) and *optimal* if it also maximises the objective (1.1a). There exists a corresponding problem to (P) called the *dual* problem and it can be written as follows.

$$\begin{aligned} &\min\ b^t y && (1.2a)\\ (D)\quad &\text{s.t. } A^t y \geq c && (1.2b)\\ &y \in \mathbb{R}^m_{\geq 0}. && (1.2c) \end{aligned}$$

This problem has n constraints and m variables. Note, the dual problem of (D) is again the primal problem (P).

Theorem 1.3. (Weak and strong duality) *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. For a feasible solution $x \in \mathbb{R}^n$ of the primal LP* (1.1) *and a feasible solution $y \in \mathbb{R}^m$ of the dual LP* (1.2) *it holds*

$$c^t x \leq b^t y. \tag{1.3}$$

Furthermore, there exist optimal solutions x^* and y^* for (1.1) *and* (1.2)*, respectively, with finite objective values so that*

$$c^t x^* = b^t y^*. \tag{1.4}$$

If the feasible solutions of the LP (1.1) are restricted to integer values, i. e., $x \in \mathbb{Z}^n_{\geq 0}$, the resulting optimisation problem is an *integer linear program* (ILP). The solving of ILPs is strongly NP-hard [103] and remains so even if some of the variables are not required to be integer yielding a *mixed integer linear program* (MILP). A generalisation of ILPs and MILPs are *integer programs* (IPs) and *mixed integer programs* (MIP) which are not necessarily linear. Relaxing all integrality constraints, we obtain the *linear programming relaxation* of the ILP (LP relaxation). The following example depicts the corresponding polytopes of an ILP and its LP relaxation.

Example 1.4. Let

$$A = \begin{pmatrix} 5 & 4 \\ -1 & 0 \\ 0 & -1 \end{pmatrix},\ b = \begin{pmatrix} 24 \\ -\frac{3}{2} \\ -\frac{1}{2} \end{pmatrix}$$

and define the polytopes $P(A,b) \coloneqq \{x \in \mathbb{R}^2_{\geq 0} \,|\, Ax \leq b\}$ and $P_{\text{int}}(A,b) \coloneqq \{x \in \mathbb{Z}^2_{\geq 0} \,|\, Ax \leq b\}$. The convex hulls of these two polytopes are displayed in Figure 1.1.

As an ILP is directly connected to its corresponding polytope, we also refer to the valid inequalities introduced in Section 1.1 as valid inequalities for an ILP.

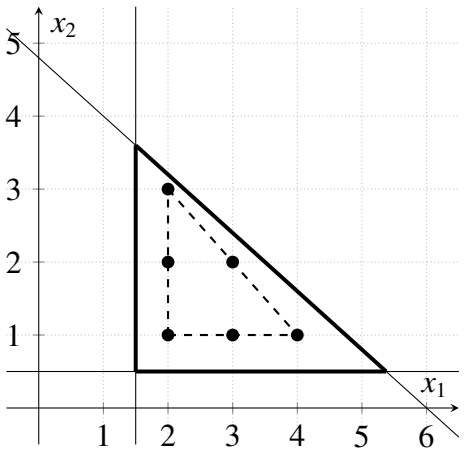


Figure 1.1.: The polytope $P(A, b)$ defined in Example 1.4 (thick outer triangle), its integer points (dots) and the convex hull of these points (dashed inner triangle).

1.2. Optimisation Techniques

In this section, we briefly present some important optimisation techniques applied in this thesis.

1.2.1. Branch-and-Bound

A well-established optimisation framework to solve ILPs is the *branch-and-bound* (B&B) algorithm first introduced by Land and Doig [119]. This algorithm enumerates all candidate solutions systematically and discards unprofitable subsets of solutions on the fly by means of upper and lower bounds. The solutions are constructed in a (binary) tree where the root node is the optimum solution of the LP relaxation. If the LP relaxation is unbounded or has no solution, the same holds for the ILP, while an integer optimal solution of the LP is also optimal for the ILP. In case of an optimal solution vector x^* with at least one non-integer entry $x_i^* \notin \mathbb{Z}$, the solution space is split into two solution subsets where x_i^* is bounded above by the next smaller integer ($x_i^* \leq \lfloor x_i^* \rfloor$) in one of them and bounded below by the next larger integer ($x_i^* \geq \lceil x_i^* \rceil$) in the other. This splitting step is called *branching*. The two newly created subproblems (nodes) are solved individually and the branching procedure is repeated recursively until all solution vectors in the leaves of the tree are integer. However, this enumeration leads to an exponential number of nodes. Hence, lower and upper bounds are used to remove complete subproblems from the tree without solving them. Every integer feasible solution found in any node of the tree is feasible for the original ILP, hence yields a global lower bound in case of a maximisation problem. Furthermore, the optimal solution of the LP relaxation in each node is a local upper bound for the current subproblem. If this upper bound is below the global best known lower bound, this node (and all its potential child nodes) can be discarded from the solution tree as this subproblem cannot yield a better integer solution. This procedure is called *bounding*. A survey on the early applications of the B&B approach can be found in Lawler and Wood [121].

1.2.2. Cutting Planes

An approach to tighten the LP relaxation of an ILP, which is for example computed in a B&B algorithm, is to add inequalities which are valid for the ILP but are not satisfied by the current (fractional) LP solution. Such inequalities are called *cutting planes* or simply *cuts* as they cut an area of the polyhedron associated with the LP relaxation off, including the current fractional solution. Though, no integer solution is excluded from the solution space by adding these inequalities. One of the earliest paper introducing a complete cutting plane algorithm was published in 1958 by Gomory [86]. Given a fractional LP solution, a cut or a proof that no violated cut exists is determined by a so-called *separation problem*. Grötschel et al. [90] show that a separation problem is polynomially solvable if and only if the corresponding optimisation problem can be solved in polynomial time. There exist numerous classes of cutting planes. We will present the class of cover inequalities which are problem-specific cuts for the knapsack problem in Section 1.3.

The integration of the separation of cutting planes for a LP relaxation into a B&B algorithm is called *branch-and-cut* (B&C) or cut-and-branch if cuts are computed only at the root node of the B&B tree. The generation of cutting planes is sometimes also referred to as *row generation*.

1.2.3. Column Generation

The complementary approach of generating valid inequalities or rows in the primal problem is the generation of columns (variables) in the dual problem by means of *column generation*. We give a brief introduction to the general concept of column generation in combination with a Dantzig-Wolfe decomposition [62] in the following, where a comprehensive overview can be found in Barnhart et al. [14], Desaulniers et al. [64] and the references therein.

Let the compact or original minimisation LP be defined as

$$z^* \coloneqq \min c^t x \tag{1.5a}$$
$$\text{s.t. } Ax \geq b \tag{1.5b}$$
$$Dx \geq d \tag{1.5c}$$
$$x \in \mathbb{R}^n_{\geq 0}, \tag{1.5d}$$

where constraints (1.5c) and (1.5d) can be replaced by $x \in X$ with $X \coloneqq \{x \in \mathbb{R}^n_{\geq 0} \mid Dx \geq d\}$ defining the polyhedron of the subproblem (1.5c), (1.5d). We assume henceforth that X is bounded.

The basic idea of Dantzig-Wolfe decomposition is to substitute the variables x by a convex combination of the extreme points of the subproblem described by X. Thus, x is replaced by

$$\sum_{k=1}^{K} \lambda_k x^k \text{ with } \sum_{k=1}^{K} \lambda_k = 1,\ \lambda_k \geq 0\ \forall\, k = 1, \ldots, K$$

and K denoting the number of extreme points x^k. By this substitution, we obtain the

following LP called *master problem* (MP).

$$z^*_{\text{MP}} \coloneqq \min \sum_{k=1}^{K} (c^t x^k)\lambda_k \tag{1.6a}$$

$$\text{s.t.} \sum_{k=1}^{K} (Ax^k)\lambda_k \geq b \tag{1.6b}$$

$$\sum_{k=1}^{K} \lambda_k = 1 \tag{1.6c}$$

$$\lambda_k \geq 0 \qquad \forall\, k = 1, \ldots, K. \tag{1.6d}$$

Since the number of extreme points K can be huge constituting a huge number of variables, the MP is restricted to a subset $K' \subseteq K$ and further variables are determined in a *pricing problem* (PP). More precisely, the *restricted master problem* (RMP) with objective value $\bar{z}$ is obtained by restricting the MP (1.6) to K'. Note that such a restriction is justified as many variables will be set to zero in an optimal solution anyway. The optimality of a solution of such a RMP is checked in the PP defined as follows.

$$\bar{c}^* \coloneqq \min (c^t - \pi^t A)y - \pi_0 \tag{1.7a}$$

$$\text{s.t. } Dy \geq d \tag{1.7b}$$

$$y \in \mathbb{R}_{\geq 0}, \tag{1.7c}$$

where π denotes the dual of constraint (1.6b) and π_0 of (1.6c). The problem (1.7) determines further necessary variables λ_k in $K \setminus K'$ if the objective $\bar{c}^*$, which is called *reduced cost*, is negative. If the matrix D has a block diagonal structure, this can be exploited in the PP by dividing it into several subproblems; see [64] for details. Furthermore, the PP is a separation problem for the dual of the RMP which reflects the complementarity of generating rows and columns. If no further column with negative reduced cost can be computed by the pricing problem, the current solution of the RMP is optimal for the MP and the column generation routine is completed. This approach is usually applied to problems with a huge number of variables and displayed schematically in Figure 1.2.

In case of integer variables, the column generation approach has to be combined with a branching routine yielding the so-called *branch-and-price* (B&P) algorithm. A further reason to apply B&P apart from the capability of handling huge numbers of variables is that the LP solution found by the column generation approach can be significantly better than the solution found for the LP relaxation of the corresponding original ILP formulation.

Proposition 1.5. *(Vanderbeck [174])*

$$z^*_{LP} \leq z^*_{MP} \leq z^*,$$

where z^*_{LP} denotes the LP relaxation of (1.5) *in case $x \in \mathbb{Z}^n_{\geq 0}$. The first inequality is strict unless the extreme points of the LP relaxation of the PP* (1.7) *with $y \in \mathbb{Z}^n_{\geq 0}$ are integral, i. e., conv$(\{y \in \mathbb{Z}^n_{\geq 0} \mid Dy \geq d\}) = \{y \in \mathbb{R}^n_{\geq 0} \mid Dy \geq d\}$.*

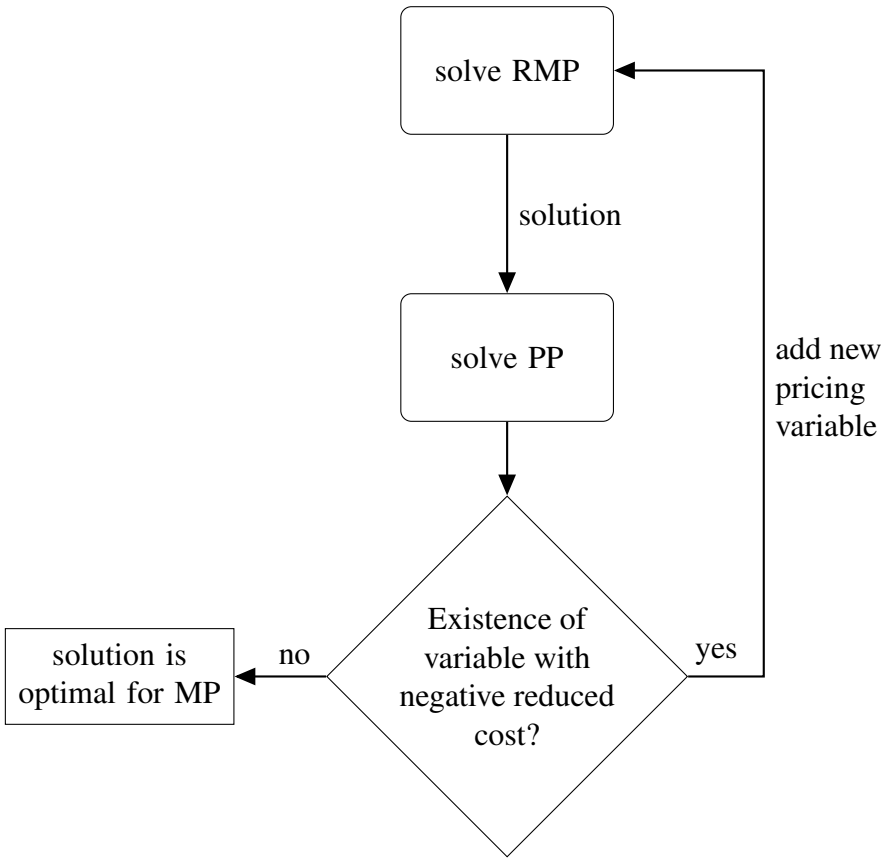


Figure 1.2.: Schematic diagram of the column generation concept.

1.2.4. Lagrangian Relaxation

A dual approach to solve optimisation problems is Lagrangian relaxation, which has been proposed in 1974 by Geoffrion [81]. In this approach, constraints are relaxed and their violation is penalised in the objective function. The following summary of the main concept in combination with a subgradient algorithm is based on Geoffrion [81], Nemhauser and Wolsey [142], and Wolsey [179].

Consider a general maximisation ILP given as

$$\chi_{\text{ILP}} \coloneqq \max\ c^t x \tag{1.8a}$$
$$\text{s.t. } Ax \leq b \tag{1.8b}$$
$$Dx \leq d \tag{1.8c}$$
$$x \in \mathbb{Z}^n_{\geq 0}, \tag{1.8d}$$

where constraints (1.8b) are “nice” in the sense that a problem containing only these constraints can be solved easier than (1.8) and (1.8c) are “complicating” constraints. The

Lagrangian relaxation of (1.8) is defined as follows.

$$\chi_{\mathrm{LR}}(\mu) \coloneqq \max\ c^t x + \mu(d - Dx) \tag{1.9a}$$
$$\text{s.t } Ax \leq b \tag{1.9b}$$
$$x \in \mathbb{Z}^n_{\geq 0} \tag{1.9c}$$

with $\mu \in \mathbb{R}^m_{\geq 0}$ being the dual variable of (1.8c) which is also called *Lagrange multiplier*. The complicating constraints (1.8c) are included in the objective function by means of the penalty term $\mu(d - Dx)$. If a constraint (1.8c) is violated by a solution of (1.9), then the penalty term is negative since $\mu \geq 0$ and decreases the objective value. Thus, if μ is sufficiently large, it is beneficial to satisfy (1.8c) in an optimal solution of (1.9).

Remark 1.6. The Lagrangian relaxation yields an upper bound of the ILP (1.8), thus

$$\chi_{\mathrm{ILP}} \leq \chi_{\mathrm{LR}}(\mu) \quad \forall\, \mu \geq 0.$$

The best upper bound for (1.8) is $\chi_{\mathrm{LR}}(\mu^*)$ with μ^* being the optimal solution of the *Lagrangian dual problem*

$$\chi_{\mathrm{LD}} \coloneqq \min_{\mu \geq 0} \chi_{\mathrm{LR}}(\mu). \tag{1.10}$$

The following proposition specifies the strength of the Lagrangian dual.

Proposition 1.7. (Nemhauser and Wolsey [142])

$$\chi_{ILP} \leq \chi_{LD} \leq \chi_{LP}, \tag{1.11}$$

where χ_{LP} is the LP relaxation of (1.8)*.*

If the extreme points of $\{x \in \mathbb{R}^n_{\geq 0} \,|\, Ax \leq b\}$ are integral, i. e., conv$(\{x \in \mathbb{Z}^n_{\geq 0} \,|\, Ax \leq b\}) = \{x \in \mathbb{R}^n_{\geq 0} \,|\, Ax \leq b\}$, equality holds for the second inequality.

The Lagrangian relaxation approach is especially suitable for applications for which a slightly modified infeasible solution of the Lagrangian dual becomes feasible for the original problem with only a small degradation of the objective value.

The crucial task is now to solve (1.10), i. e., to find the optimal Lagrange multiplier μ^*. One approach is a subgradient algorithm which we briefly present in the following.

Subgradient algorithm The *subgradient algorithm* is an iterative method to solve the Lagrangian dual problem with initial vector μ^0, where $\mu^0 = 0$ is a natural choice. For an optimal solution x^i of $\chi_{\mathrm{LR}}(\mu^i)$ at iteration i,

$$\xi^i \coloneqq d - Dx^i \tag{1.12}$$

is a subgradient of $\chi_{\mathrm{LR}}(\mu)$ at $\mu = \mu^i$; see [142]. The Lagrange multiplier μ^{i+1} of iteration $i+1$ is determined by the multiplier of the previous iteration i:

$$\mu^{i+1} = \left(\mu^i - \alpha_i \xi^i\right)^+.$$

Furthermore, $\alpha_i \geq 0$ denotes the step length which has to be defined for every iteration i. The most commonly used definition of the step length is

$$\alpha_i := \delta_i \frac{\chi_{\mathrm{LR}}(\mu^i) - \chi^i}{\|\xi^i\|^2},$$

with χ^i denoting the primal lower bound and $0 < \delta^i \leq 2$ a scaling parameter at iteration i.

Possible stop criteria for the described iteration are the following, where ε_1, ε_2, $\varepsilon_3 > 0$ are predefined precision parameters.

- $(\chi_{\mathrm{LR}}(\mu^i) - \chi^i)/\chi_{\mathrm{LR}}(\mu^i) \leq \varepsilon_1$; the desired quality ε_1 is reached,
- $\chi_{\mathrm{LR}}(\mu^i) - \chi_{\mathrm{LR}}(\mu^{i-1}) \leq \varepsilon_2$; the upper bound was improved only marginally in one step of iteration,
- $\alpha_i \leq \varepsilon_3$; the step length is too small.

If one of these criteria catches, the subgradient algorithm stops.

1.3. The Knapsack Problem

One of the most fundamental problems in mathematical optimisation is the classical (binary) knapsack problem (KP). It asks to select a subset of items $i \in N = \{1, \ldots, n\}$ having a (positive) weight w_i and a (positive) profit p_i such that a given capacity B is not exceeded and the total profit is maximised.

Definition 1.8. *(Knapsack Problem)* The classical (binary or 0-1) *knapsack problem* can be formalised as

$$\max \sum_{i \in N} p_i x_i \tag{1.13a}$$

$$\text{s.t.} \sum_{i \in N} w_i x_i \leq B \tag{1.13b}$$

$$x_i \in \{0, 1\} \qquad \forall\, i \in N. \tag{1.13c}$$

This problem is well studied, see [106, 135] for comprehensive surveys, since it is a quite general problem that occurs in many applications such as scheduling, telecommunication and logistics. However, it is difficult to solve in general as the corresponding decision problem, which determines if a subset of items exists such that the sum of the weights is less than or equal to a given capacity and the total profit exceeds a certain threshold, is NP-complete. This characteristic of the KP was first shown in 1972 by Karp [105]. Since then quite some research has been carried out solving the KP. We state just the earliest publications of different approaches and refer to the books Kellerer et al. [106] and Martello and Toth [135] for surveys on algorithms. The first B&B algorithm to solve the KP exactly was proposed by Kolesar [114] in 1967. A variant of the B&B algorithm

is a primal method which adds items to an empty knapsack until the capacity is exceeded. The first primal framework was presented in 1974 by Horowitz and Sahni [96]. A further exact solution approach is the concept of *dynamic programming*. In this method, only a (small) subproblem of the KP is solved and then this solution is extended iteratively to a solution of the complete problem. Two of the earliest works on dynamic programming algorithms for the subset sum problem, which is a KP where the profits are equal to the weights of the items, are Faaland [69] and Ahrens and Finke [5] published in 1973 and 1975, respectively. Even though many improvements of the just stated exact algorithms exist and many standard knapsack instances from the literature can be solved, there exists a variety of instances for which the knapsack is still hard to solve in practice; see Pisinger [154].

As mentioned in Section 1.2.2, one approach to improve the upper bounds for the KP in a B&B algorithm are cutting planes. A problem specific class of cutting planes for the KP are *cover inequalities*.

Definition 1.9. *(Cover)* A subset $C \subseteq N$ is called a *cover* if the weights of the items in the cover exceed the knapsack capacity:

$$\sum_{i \in C} w_i > B.$$

It is *minimal* if the removal of a single item j accounts for $C \setminus \{j\}$ not being a cover anymore. This means,

$$\sum_{i \in C \setminus \{j\}} w_i \leq B$$

for every $j \in C$.

Definition 1.10. *(Knapsack Polytope)* The *knapsack polytope* P_{KP} is defined as the convex hull of the set of feasible solutions of the KP (1.13):

$$P_{KP} := \operatorname{conv}\left\{ x \in \{0,1\}^n \;\middle|\; \sum_{i \in N} w_i x_i \leq B \right\}.$$

Lemma 1.11. (Balas [11], Hammer et al. [93], Wolsey [178]) *For every (minimal) cover C, the following inequality is valid for the knapsack polytope P_{KP}.*

$$\sum_{i \in C} x_i \leq |C| - 1 \tag{1.14}$$

Such an inequality is referred to as (minimal) cover inequality *and is facet-defining if the cover C is minimal and $C = N$.*

The knapsack constraint (1.13b) can be replaced by all cover inequalities. This means, an integer solution is feasible for the KP if and only if it satisfies every cover inequality. However, since there exist (exponentially) many cover inequalities, it is not beneficial to compute all inequalities in advance. Instead, they are usually separated on the fly as described in the following.

Separation of cover inequalities. The problem to decide whether there exists a violated cover inequality for a given LP solution is NP-hard [107]. A survey on exact and heuristic separation algorithms of different types of valid inequalities for the (binary) KP is given in Kaparis and Letchford [104]. Here, we present a well-established exact separation routine; see, e.g., Balas [11].

Let x^{LP} be the solution of the LP relaxation with at least one $x_i^{\mathrm{LP}} \notin \{0, 1\}$. Without loss of generality, we assume that all weights w_i as well as the capacity B are integer values henceforth. We intend either to show that the solution x^{LP} satisfies all cover inequalities or to compute a most violated cover inequality. For this purpose, we solve the following separation problem.

$$\min \sum_{i \in N'} (1 - x_i^{\mathrm{LP}}) y_i \tag{1.15a}$$

$$\text{s.t.} \sum_{i \in N'} w_i y_i \geq B + 1 \tag{1.15b}$$

$$y_i \in \{0, 1\} \qquad \forall i \in N', \tag{1.15c}$$

with $N' \coloneqq N \setminus \{i \in N \mid x_i^{\mathrm{LP}} = 1\}$. If the objective value corresponding to a solution y^* is greater than or equal to 1, then the LP solution x^{LP} satisfies all cover inequalities. Otherwise, we construct a violated cover inequality which has to be added to the problem. The minimal cover C is defined as the set of items chosen in the solution y^*, i.e., $C \coloneqq \{i \in N' \mid y_i^* = 1\}$. It is minimal, since the removal of an item cannot retain the cover property as this would give a smaller solution value. Then a most violated cover inequality is given by

$$\sum_{i \in C} x_i \leq |C| - 1.$$

The violation of this inequality can be computed as

$$\sum_{i \in C} x_i^{\mathrm{LP}} - |C| + 1$$

which is maximised by the objective function (1.15a). Hence, the determined inequality is a most violated.

Note, the separation problem (1.15) is a minimisation KP with real objective coefficients $(1 - x_i^{\mathrm{LP}})$. It is equivalent to the KP

$$\sum_{i \in N'} (1 - x_i^{\mathrm{LP}}) - \max \sum_{i \in N'} (1 - x_i^{\mathrm{LP}}) z_i$$

$$\text{s.t.} \sum_{i \in N'} w_i z_i \leq \sum_{i \in N'} w_i - B - 1$$

$$z_i \in \{0, 1\} \qquad \forall i \in N'$$

with $y_i = 1 - z_i$. Hence, (1.15) is also NP-hard.

In general, cover inequalities are not facet-defining for P_{KP} unless the cover is minimal

and $C = N$. Nevertheless, they can be lifted/strengthened to define a facet of the knapsack polytope [11, 93, 178]. One concept of strengthened inequalities which are not necessarily facet-defining are so-called extended covers.

Definition 1.12. *(Extended Cover)* For a cover C we define the corresponding *extended cover* as

$$E(C) := C \cup \left\{ i \in N \setminus C \,\middle|\, w_i \geq \max_{j \in C} w_j \right\}.$$

The *extended cover inequality*

$$\sum_{i \in E(C)} x_i \leq |C| - 1 \tag{1.16}$$

is valid for the knapsack polytope P_{KP}.

Necessary and sufficient conditions for (1.16) to be facet-defining are given in [11, 93, 178]. By Balas and Jeroslow [12], we know that a vector $x \in \{0,1\}^n$ is in P_{KP} if and only if it satisfies the extended cover inequality (1.16) for every minimal cover $C \subseteq N$. Note, the extended cover inequalities cannot describe the knapsack polytope completely. A complete description of a special case of the binary knapsack polytope with specific conditions on the weights can be found in Weismantel [176].

1.3.1. A Dynamic Program

As mentioned in the previous section, an exact algorithm to solve the KP is a dynamic programming algorithm (DP). Such an algorithm solves the problem in pseudo-polynomial time in the number of items n and the capacity $B \in \mathbb{N}$. Hence, the KP is *NP-hard in the weak sense* (or weakly NP-hard). We state a basic DP in the following from Toth [170] which can also be found, e.g., in the books Kellerer et al. [106] and Martello and Toth [135].

Let $f(j, b)$ denote the highest profit for a feasible solution of the KP (1.13) with total weight equal to $b \in \{0, 1, \ldots, B\}$ and in which only the set of items $\{1, \ldots, j\} \subseteq N$ with $j \in N$ are considered, i.e.,

$$f(j, b) := \max \left\{ \sum_{i=1}^{j} p_i x_i \,\middle|\, \sum_{i=1}^{j} w_i x_i = b, x_i \in \{0,1\} \, \forall i \leq j \right\}.$$

The DP then consists of the computation of all values of f by the recursive equation

$$f(j, b) = \max \left\{ f(j-1, b), f(j-1, b-w_j) + p_j \right\} \tag{1.17}$$

with initial values

$$f(1, b) = \begin{cases} 0, & \text{if } b = 0 \\ p_1, & \text{if } b = w_1 \\ -\infty, & \text{otherwise} \end{cases}$$

and $j \in N$, $b \in \{0, \dots, B\}$. Formally, the optimal solution value z^* of (1.13) is determined by Algorithm 1.

Algorithm 1 DP for KP

Input: A set of items N, capacity $B \in \mathbb{Z}$, weights $w \in \mathbb{R}^n_{\geq 0}$, and profits $p \in \mathbb{R}^n_{\geq 0}$.
Output: An optimal solution value z^* of the KP (1.13).

for $b = 0 \dots B$ **do**

$$f(1, b) = \begin{cases} 0, & \text{if } b = 0 \\ p_1, & \text{if } b = w_1 \\ -\infty, & \text{otherwise} \end{cases}$$

end for
for $j = 2 \dots, n$ **do**
  for $b = 0, \dots, w_j - 1$ **do**
    $f(j, b) = f(j-1, b)$
  end for
  for $b = w_j, \dots, B$ **do**
    if $f(j-1, b-w_j) + p_j > f(j-1, b)$ **then**
      $f(j, b) = f(j-1, b-w_j) + p_j$
    else
      $f(j, b) = f(j-1, b)$
    end if
  end for
end for
$z^* := \max_{b \in \{0, \dots, B\}} f(n, b)$

Lemma 1.13. *The DP 1 solving the KP* (1.13) *has an overall running time of* $O(nB)$.

Proof. Each iteration over the n items $j \in N$ contains $B + 1$ iterations for the capacity $b \in \{0, \dots, B\}$ yielding an overall running time of $O(nB)$. □

We would like to remark that Algorithm 1 only computes the optimal solution value z^* and not the corresponding solution vector x^*. Nevertheless, it can be extended to compute also the solution at the expense of a higher running time; cf. algorithm "DP-3" in [106].

2. Optimisation under Uncertainty

Optimisation models, which formalise real-world problems mathematically, typically include assumptions and simplifications to reduce their complexity. One prevalent assumption is the certainty of data. Many optimisation problems ignore fluctuations in the input data. However, in real-world applications even a small variance in the input values can strongly affect the quality of the computed solution, demonstrated by Ben-Tal and Nemirovski [15]. A solution found while ignoring data uncertainty can actually be infeasible or useless from a practical point of view. Therefore, a methodology which is capable of handling data uncertainty is needed.

2.1. Data Uncertainty

Ben-Tal et al. [16] name the following three errors which are the most common reasons for data uncertainty.

- *prediction*: Since some data is not known when the problem is solved, forecasts, often based on historical data, are used instead.
- *measurement*: If some data cannot be measured exactly due to, e. g., physical limitations, the actual values deviate from the measured values used in the problem formulation.
- *implementation*: A decision variable cannot be implemented in the real-world application exactly as computed. Or vice versa, the data cannot be represented in the model as precise as given.

Sensitivity analysis is a traditional method of identifying data uncertainty in optimisation problems. By such an analysis, the continuance property of an optimal solution of the original problem, which is also called nominal solution, is investigated; see [15] for a sensitivity analysis of LPs. In contrast, the following methodologies, which are discussed in this thesis, deal with data uncertainty by building solutions that are resistant to uncertainty.

2.2. Stochastic Optimisation

If the uncertain data is random obeying a probability distribution which is known in advance, stochastic optimisation is suitable to handle this type of uncertainty. One of the earliest works on optimisation under uncertainty was presented by Dantzig [61] in 1955

laying the foundations for stochastic optimisation. The proposed stochastic programs minimise the expected value of an objective function. Such a problem is usually referred to as an one-stage stochastic program and can be written as

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

where $x \in \mathbb{R}^n$ is the vector of decision variables, $D \in \mathbb{R}^k$ a vector of random variables representing, e. g., uncertain demand, $f : \mathbb{R}^{n \times k} \to \mathbb{R}$ is the objective function and $h : \mathbb{R}^n \to \mathbb{R}$ is a constraint function; cf. also Shapiro et al. [164]. As an example, if f represents a cost function, then the stochastic program minimises the expected total cost with respect to the different realisations of D. If D has a finite number of realisations, then the stochastic program (2.1) can be modelled as a deterministic optimisation problem where the expected value is given as the weighted sum

$$\mathbb{E}[f(x, D)] = \sum_{i=1}^{I} p_i f(x, d^i),$$

where I is the number of realisations and realisation d^i occurs with probability p_i.
The stochastic program (2.1) is an one-stage concept, which can be extended to a multi-stage stochastic program. Multi-stage programming is an approach of sequential decision making where decisions of later stages may depend on determinations of earlier stages and may also change those former decisions. One example of a two-stage stochastic optimisation model is the following.

$$\begin{aligned} &\min cx + \mathbb{E}[\phi(x, D)] \\ &\text{s.t. } h(x) \leq 0 \\ &x \geq 0, \end{aligned}$$

where $\phi(x, d)$ is an optimal value of a subproblem depending on realisation d of D and on x. The expectation is with respect to the probability distribution of D. Here, the decisions in the first stage (x) are made to meet the uncertain but known distribution of demand D occurring in the second stage (subproblem). Hence, the expected value of the present two-stage objective is the sum over the known cost c plus the expected value of costs which depend on the demands of the second stage. A common solution method to solve such a two-stage model is based on so-called *scenarios* [26]. When assuming that the realisations d of D can be specified in the form of K scenarios $d^1, d^2, \ldots, d^K$ occurring with probabilities $\rho^1, \rho^2, \ldots, \rho^K$ the problem can be formulated as

$$\min cx + \sum_{k=1}^{K} \rho^k \phi(x, d^k)$$

$$\text{s.t. } h(x) \leq 0$$
$$x \geq 0.$$

Moreover, if the second stage is also divided into several stages, we have a multi-stage optimisation problem.

2.3. Chance Constraints

In the one-stage stochastic program (2.1), the expected optimal value $\mathbb{E}[f(x, D)]$ can be considerably far from the actual cost $f(x, D)$ for a particular realisation of the demand D. Hence, we ask for an approach to limit the cost $f(x, D)$ instead of averaging over all realisations. One possibility to limit the cost is to require $f(x, D) \leq \delta$, where $\delta > 0$ is a threshold. To include this inequality for every realisation d of D is quite restrictive, especially for a large number of realisations. Instead, we limit the probability that $f(x, D)$ exceeds the threshold δ by $\varepsilon \in (0, 1)$. This *chance or probabilistic constraint* can be written as

$$\mathbb{P}\{f(x, D) > \delta\} \leq \varepsilon$$

or equivalently

$$\mathbb{P}\{f(x, D) \leq \delta\} \geq 1 - \varepsilon; \tag{2.2}$$

see [164]. Adding the chance constraint (2.2) to the optimisation problem (2.1) leads to a stochastic program that minimises the cost on average while ensuring that the probability of $f(x, D)$ staying below the threshold δ is large. Note, probabilistic constraints can also be incorporated in an optimisation problem with an objective function $c(x)$ which is not an expectation. We discuss chance constraints in more detail in Part II.

2.4. Robust Optimisation

For the chance constraint (2.2), we can define the feasible set

$$\{x \geq 0 \,|\, \mathbb{P}\{f(x, D) \leq \delta\} \geq 1 - \varepsilon\}$$

which can be written equivalently as

$$\{x \geq 0 \,|\, f(x, d) \leq \delta, d \in \mathcal{D}, \mathbb{P}(d) \geq 1 - \varepsilon\}$$

by abusing the notation of the function $f(x, \cdot)$; see [164]. The set $\mathcal{D}$ is any measurable subset of $\mathbb{R}^n$ such that $\mathbb{P}(D \in \mathcal{D}) \geq 1 - \varepsilon$. We can simplify this formulation by choosing a fixed set $\mathcal{D}_\varepsilon$ with $\mathbb{P}(\mathcal{D}_\varepsilon) \geq 1 - \varepsilon$, which represents a subset of realisations of the random variable D and is also referred to as *uncertainty set*. Using this set and minimising the

objective function $c(x)$, we can rewrite the optimisation problem as

$$\begin{aligned} &\min c(x) \\ &\text{s.t.} f(x,d) \leq \delta && \forall\, d \in \mathcal{D}_\varepsilon \\ &\quad h(x) \leq 0 \\ &\quad x \geq 0. \end{aligned}$$

This formulation is a robust optimisation problem which includes a deterministic and set-based uncertainty model. Since we are considering only linear problems in this thesis, we only refer to the linear case henceforth. Robust linear optimisation goes back to Soyster [167] who considered inexact linear programs in 1973. We refer to [16, 20] for the following introduction on robust linear optimisation.

Definition 2.1. An *uncertain linear optimisation problem* is a collection of linear optimisation problems

$$\min c^t x \tag{2.3a}$$
$$\text{s.t. } Ax \leq b \tag{2.3b}$$
$$x \in \mathbb{R}^n \tag{2.3c}$$

with $(c, A, b) \in \mathcal{U}$ and $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^{m\times n} \times \mathbb{R}^m$ a given uncertainty set. A vector $x \in \mathbb{R}^n$ is *robust feasible* if it satisfies all realisations of the constraints (2.3b) with $(c, A, b) \in \mathcal{U}$, thus, if

$$Ax \leq b \quad \forall\, (c, A, b) \in \mathcal{U}.$$

The aim is to find the best objective value among all robust feasible solutions. This can be done by solving the so-called robust counterpart of the problem (2.3).

Definition 2.2. The *robust counterpart* of the uncertain LP (2.3) is defined as the optimisation problem

$$\min \sup_{(c,A,b)\in\mathcal{U}} c^t x \tag{2.4a}$$
$$\text{s.t. } Ax \leq b \qquad \forall\, (c, A, b) \in \mathcal{U} \tag{2.4b}$$
$$x \in \mathbb{R}^n. \tag{2.4c}$$

An optimal solution of (2.4) is called a *robust optimal solution* of (2.3).

The uncertainty in the objective function can always be shifted to the constraints by introducing an auxiliary variable; see Ben-Tal et al. [16]. Hence, we restrict ourselves to uncertain LPs with certain objectives from now on and thus, $(A, b) \in \mathcal{U} \subset \mathbb{R}^{m\times n} \times \mathbb{R}^m$.

Since robust optimisation constructs solutions which are resistant to the worst-case realisation of the uncertainty in every constraint, we have to consider all worst-case realisations separately for each constraint unless we apply a multi-stage approach that allows for modifications of earlier decisions. Therefore, the uncertainty set is extended to the direct

product $\mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_m$ of its projections on the uncertain data of each constraint i and the robust counterpart reads

$$\min c^t x \tag{2.5a}$$
$$\text{s.t. } a_i^t x \leq b_i \qquad \forall\, i = 1, \ldots, m, \forall (a_i, b_i) \in \mathcal{U}_i \tag{2.5b}$$
$$x \in \mathbb{R}^n, \tag{2.5c}$$

where a_i^t is the i-th row in matrix A and b_i is the i-th entry in vector b. Moreover, Ben-Tal et al. [16] show that every robust feasible solution of (2.5) remains feasible if the uncertainty sets $\mathcal{U}_i$ are extended to their convex hull $\text{conv}(\mathcal{U}_i)$ and to the closure of this set. Therefore, the uncertainty set $\mathcal{U}$ is the direct product of closed and convex sets. Examples of uncertainty sets (see Bertsimas et al. [20] and Poss [155]) are the following.

- Ellipsoidal: All uncertain vectors are described by an ellipsoid.
- Polyhedral: All uncertain vectors are described by a polyhedron.
- Cardinality constrained/budgeted/Γ-robustness: A special case of a polyhedral uncertainty set, where the uncertain vectors are defined by nominal and deviation values and the number of simultaneous deviations is limited by a robustness parameter; see Part III for a detailed description of the concept and applications.
- Variable budgeted: Generalises the budgeted uncertainty set so that the robustness parameter depends on the solution; see Poss [155].

Light robustness A heuristic approach to model uncertainty is the combination of robust optimisation with a simplified two-stage stochastic programming approach, introduced as *light robustness* in Fischetti and Monaci [70]. First, the maximum impairment of the objective value of the nominal problem is fixed by means of an additional constraint. Afterwards, robustness requirements, e. g., via Γ-robustness constraints, are introduced. However, the resulting model is most likely infeasible. Slack variables are then introduced to handle this infeasibility by allowing limited violations of the robustness requirements. The objective of this approach is to minimise these slack variables.

Light robustness can be seen as the "flexible counterpart of robust optimisation" [70] and is so far mostly used in timetabling problems; see, e. g., Fischetti et al. [74], Goerigk et al. [82].

Recoverable robustness Just as in stochastic optimisation, there exist multi-stage concepts in robust optimisation. A quite recent two-stage approach is *recoverable robustness*. There, a first stage decision is made regardless of the uncertain data of the second stage. Then for a realisation of the uncertainty, a recovery action can take place to adjust a first-stage decision. The objective is to minimise the total cost of the first-stage decision taking limited second-stage adjustments into account.

The notion of recoverable robustness was first introduced by Liebchen et al. [124] for railway applications and will be discussed in more detail in Part V.

3. Technical Background

In this chapter, we give a brief summary of the technical background on wireless communication networks which is needed for the remainder of this thesis. For more details on wireless communication in general, we refer to the books by Goldsmith [83] and Tse and Viswanath [171] which also form the basis of this chapter.

Wireless communication dates back to 1897 when Marconi sent the first wireless telegraphy over open sea via a distance of six kilometres. The first radio transmission across the Atlantic Ocean was performed in 1901. Since then different generations of networks were developed. The first generation (1G) of mobile telephony introduced in the 1980s was an analogue telecommunications standard. From the second generation (2G) on, the wireless systems are digital and comprise also data services. Since the introduction of the third generation (3G), wireless networks provide data rates sufficiently high for mobile internet access. Fourth generation (4G) networks, which are IP-based systems meaning that telephony is conducted via Voice over IP (VoIP), are the latest deployed mobile networks. The employed techniques vary between the generations due to different requirements for voice (high latency requirement) and data (transmitted intermittently). Currently, a new standard (named 5G as a working title) is being developed; see news of standardisation authorities such as [1, 100, 145]. According to Next Generation Mobile Networks (NGMN) [145], the 5G network requirements comprise greater throughput, lower latency, ultra-high reliability, higher connectivity density, and higher mobility range to provide support, e. g., also to specific use cases such as Internet of Things.

In 4G, a complete communication network mainly consists of three sub-networks. The *access network* connects users to transceivers, which are combined transmitters and receivers. These transceivers are then connected to each other by copper or fibre wired connections or by microwave links. This network is denoted as the *backhaul network* and is usually represented by a mobile telephone switching office or mobile switching center (MSC). Additionally, the backhaul network takes the traffic from a transceiver and backhauls it to the *core network*, the third main part of a communication network. The core network is constituted by the public switched telephone network (PSTN) or the Internet. A complete communication network and the interrelations between the different sub-networks are illustrated in Figure 3.1; see also [144].

3.1. Cellular Networks

A cellular network, representing an access network, consists of a (large) number of subscribers with cellular phones called (mobile) users and a fixed number of base stations (BSs) which provide coverage of the subscribers. A BS can unite several antennae with

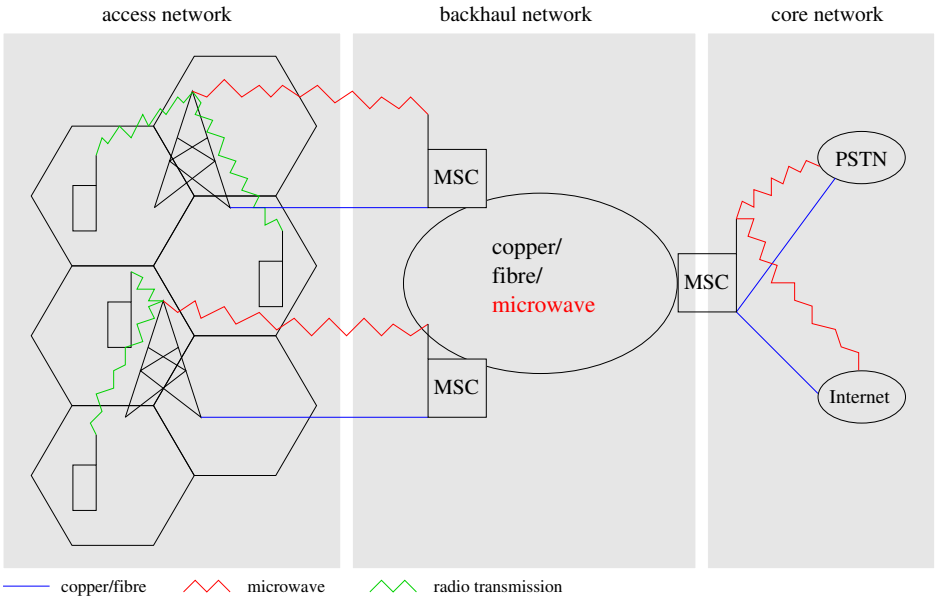


Figure 3.1.: A complete communication network

different configurations such as azimuth and downtilt. Furthermore, the area covered by a BS is called *cell* and cells are usually sub-divided into (three) sectors established by directional antennae. Every BS is connected to the backhaul network by high-speed wired connections or microwave links.

The movement of users from one cell to another is called *hand-off* or *handover* as the communication has to be switched. Such a switching from one BS to another is initiated if the current signal quality decreases below a threshold. In 3G networks, soft handover is performed: A user is dropped by a BS only if a link to a neighbouring BS is already established. In contrast, hard handover is applied in 4G networks implying that a user cannot be connected to more than one BS at the same time. Communication between BSs and users is performed in two directions called downlink (DL), i. e., from BS to user, and uplink (UL), i. e., from user to BS, and operates on a licensed frequency spectrum. The radio spectrum is regulated worldwide by the International Telecommunications Union (ITU) [100], the United Nations specialised agency for information and communication technologies (ICTs), whose main functions are to allocate global radio spectrum, develop technical standards and improve access to ICTs worldwide. A license or authorisation is then given by an administration, such as the government, and allows the assignment of a frequency spectrum to a BS.

For initial cellular systems, the cost for BSs were high and hence, the networks comprised only a small number of cells but used high transmission powers. In such a network, the coverage area of a BS is large and the corresponding cell is called a *macro cell*. Nowadays, BSs are close to street level and transmit with lower power establishing *micro or pico cells*. In 4G networks, even smaller cells, so-called *femto cells*, are deployed which can also be user-operated and are connected to the backhaul network via optical fibre or DSL. The combination of different types of cells, which are also called tiers, is called a *hetero-*

geneous network (HetNet). Reasons to deploy smaller cells are on the one hand, the need for higher capacity due to an increase in user demand and density and on the other hand, the size and cost for BS electronics have decreased significantly during the last decades. This is why modern networks comprise a large number of (different types of) cells leading to quite complex systems.

3.2. Signal Propagation

Large-scale propagation effects occurring in wireless communication networks are shadowing and path loss. *Shadow fading* describes the random variation of the transmitted signal due to blockage from objects in the signal path, reflecting surfaces and scattering objects.

The *linear path loss* P_{L} is the ratio of transmit power P_{t} to receive power P_{r}:

$$P_{\mathrm{L}} = \frac{P_{\mathrm{t}}}{P_{\mathrm{r}}}.$$

Extending this definition, the *path loss* is defined as the value of the linear path loss in decibels (dB) which is the difference in dB between transmitted and received signal power.

$$P_{\mathrm{L}}^{\mathrm{dB}} = 10 \log_{10} \frac{P_{\mathrm{t}}}{P_{\mathrm{r}}} \mathrm{dB}$$

In general, it holds $P_{\mathrm{L}}^{\mathrm{dB}} > 0$. However, sometimes the negative of the path loss is used which is called *path gain*, for example if the previous equation is solved for P_{r}:

$$P_{\mathrm{r}} = P_{\mathrm{t}} \cdot 10^{\frac{-P_{\mathrm{L}}^{\mathrm{dB}}}{10}}.$$

In a *distance based path loss model* (see [102]), the received signal power P_{r} is inversely proportional to the distance d between transmitter and receiver: $P_{\mathrm{r}} \propto \left(\frac{1}{d}\right)^{\alpha}$, where α is the *path loss exponent* and it is generally assumed $2 \leq \alpha \leq 4$, with $\alpha = 2$ in the free space.

3.3. Channel Capacity and Interference

The (radio) transmission of information is described by a channel which is modelled theoretically by a channel model. A commonly used channel model is *Additive White Gaussian Noise* (AWGN). Theoretically applied, it produces simple and tractable mathematical models. *White noise* is a random signal with a constant power spectral density, which means that the power of the signal is the same at each frequency. In an AWGN channel, the impairment to communication is caused by the addition of white noise which has a normally (Gaussian) distributed signal amplitude.

To measure the impact of noise to the signal transmission, the *signal-to-noise ratio* (SNR), which is the channel input divided by noise, is used. For a received power P_{r} and

a background noise η, which depends on the bandwidth or frequency range B given in Hz, the SNR γ is computed as

$$\gamma = \frac{P_\mathrm{r}}{\eta}. \tag{3.1}$$

The channel *capacity C_{Sh} by Shannon [163]* is computed as follows.

$$C_\mathrm{Sh} = B \log_2(1 + \gamma) \tag{3.2}$$

It is given in bps and generally used as an upper bound on the data rates that can be achieved under real system constraints. The ratio C_Sh/B, gives a measurement of the maximum achievable *spectral efficiency* as a function of the SNR. It is given in bps/Hz and denotes the information rate that can be transmitted over 1 Hz of bandwidth through the channel.

The communication in a wireless network is affected by *interference*. We distinguish between two main types of interference, *intra-cell* and *inter-cell*. By transmitting signals simultaneously to different users in the same cell, intra-cell interference can occur, whereas inter-cell interference is experienced by cell-edge users from two or more neighbouring BSs. It is also called co-channel interference. Common techniques to reduce interference and to provide higher data rates are multiple antenna techniques such as Multiple-Input and Multiple-Output (MIMO), and multi-user detection, which denotes the detection of desired signals from interference and noise.

The amount of inter- and intra-cell interference experienced by a user is measured in the so-called *signal-to-interference-plus-noise ratio* (SINR). For a received power P_r, a bandwidth B and noise η, the SINR is computed as

$$\mathrm{SINR} = \frac{P_\mathrm{r}}{\eta + P_\mathrm{I}}, \tag{3.3}$$

where P_I denotes the power associated with both types of interferences. The higher the interference the lower the SINR and the higher the bit error rate, that is the number of received bits that have been modified by noise, interference or further transmission errors.

We would like to point out that the presented ratios are the linear versions. This means, all powers are given in Watt (W). In case that a value is given on a logarithmic scale, in dB, it has to be transformed to W to be used in (3.3) and similar formulas. Assume x is given in dB, then the power P given in W can be computed as

$$P\,(\mathrm{W}) = 10^{\frac{x\,(\mathrm{dB})}{10}}$$

and vice versa

$$x\,(\mathrm{dB}) = 10 \cdot \log_{10}(P\,(\mathrm{W})).$$

Furthermore, the unit dBm is commonly used which is dB but referenced to one milliwatt (mW). Thus, 30 dBm $\hat{=}$ 0 dB.

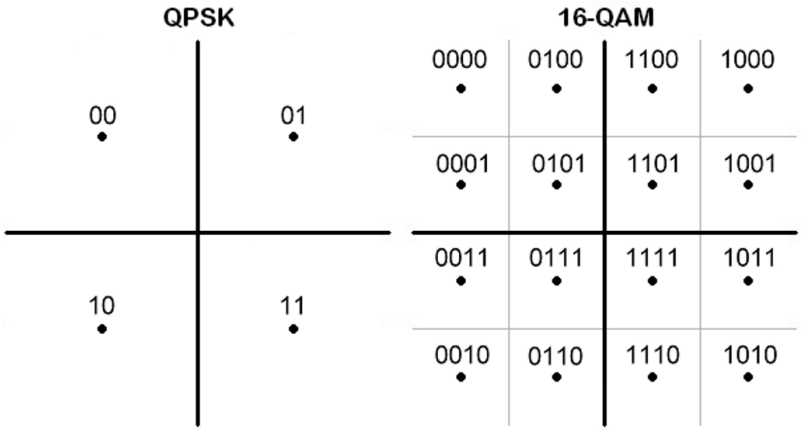


Figure 3.2.: Signal constellations for QPSK (4-QAM) and 16-QAM, where each dot denotes a constellation state with the above binary identifier. (Taken from [57].)

3.4. Modulation

By means of modulation, a signal (music, voice etc.) is sent on a carrier signal which can be transmitted physically. For example, the phase or amplitude of the carrier signal which has a periodic waveform is modulated. A commonly used modulation is digital Quadrature Amplitude Modulation (QAM). An m-QAM scheme consists of m symbols where each symbol is a combination of amplitude and phase and represents an n-bit pattern with $n = \log_2 m$ and $n \in \mathbb{Z}$. This means, instead of transmitting one bit at a time, n bits are transmitted simultaneously. The signal constellation states for 4-QAM, which is also called Quadrature Phase-Shift Keying (QPSK), and 16-QAM are displayed in Figure 3.2. The capacity C of a channel with bandwidth B using an m-QAM scheme can be computed as

$$C\,(\text{bps}) = n \cdot B\,(\text{Hz}) \leq C_{\text{Sh}}.$$

Hence, high-order QAM schemes have a higher bandwidth (spectral) efficiency. However, at the same time they are more susceptible to noise or errors caused by channel impairments since the constellation states are closer together, see Figure 3.2. Furthermore, to preserve a certain bit error rate, a better (higher) SNR is needed for modulation schemes with high bandwidth efficiency. These correlations of modulation scheme, spectral efficiency, minimum required SNR and capacity for different bandwidths are exemplarily shown in Table 3.1. Note, the increase of transmission power improves the SNR but entails higher system cost.

3.5. LTE Specifications

In this section, which is mainly based on the book by Kahn [102], we summarise some characteristics specific for the 4G networks considered in this thesis. Basically, there exist two types of 4G networks, WiMAX and LTE. WiMAX stands for Worldwide Interoper-

modulation scheme	e (bps/Hz)	min. SNR (dB)	C for 7 MHz (Mbps)	C for 14 MHz (Mbps)	C for 28 MHz (Mbps)
QPSK	2	14.21	14	28	56
16-QAM	4	21.02	28	56	112
32-QAM	5	25.24	35	70	140
64-QAM	6	27.45	42	84	168
128-QAM	7	31.10	49	98	196
256-QAM	8	33.78	56	112	224

Table 3.1.: Spectral efficiency (e), SNR requirement, and capacity (C) for three different bandwidths per modulation scheme [47].

ability of Microwave Access, while the official name of this type of network is Wireless MAN and it is standardised in the IEEE 802.16 specifications [99]. The second type of 4G network *Long-Term Evolution* (LTE) is a further development of the 3G network *Universal Mobile Telecommunications System* (UMTS) and is specified by the 3rd Generation Partnership Project (3GPP) [1]. Common goals of both types of networks are to enhance the system spectral efficiency, provide significantly higher data rates and additionally to support backward compatibility. In the remainder of this chapter and thesis, we focus on LTE.

In a LTE network, a BS is called *evolved Node-B* (eNodeB or eNB) since the radio network controller's functions are now incorporated in the BS. Tasks of a radio network controller include radio protocols, mobility management, and retransmissions. ENBs operate with a bandwidth between 1.25 and 20 MHz. User mobility is supported until up to 350 km/h whereupon the network is optimised for low speeds up to 15 km/h. For 20 MHz bandwidth, peak data rates of 326 Mbps in DL with a 4x4 MIMO eNB configuration (4 transmitting and 4 receiving antennae) can be achieved.

The modulation scheme used for DL in LTE is *Orthogonal Frequency-Division Multiple Access* (OFDMA). It is based on Orthogonal Frequency-Division Multiplexing (OFDM), where a bitstream of information is divided into substreams which are sent over different subchannels (carrier frequencies). This technique is known as frequency-division multiplexing. Moreover, OFDM uses a multitude of narrowband ($\sim$ 18 kHz) subcarrier signals which are arranged to be mutually orthogonal to transmit data streams in parallel without experiencing intra-cell interference. OFDMA is the multi-user version of OFDM, where subsets of different numbers of subcarriers are assigned to users. This is a so-called multiple access method. Despite frequency division, adjacent cells share the same frequency, which is why inter-cell interference occurs also in OFDMA based systems.

4. Wireless Communication Networks

4.1. The Wireless Access Network Planning Problem

In this section, we introduce the used notation and present a formulation to plan a cellular access network as described in Section 3.1. This network planning problem serves as a basis to apply various robustness concepts studied in this thesis.

The proposed formulation is mainly taken from our work Claßen et al. [49]. The wireless network we consider in this work consists of BSs and mobile users. The location and all configurations of a BS are consolidated in a BS candidate site $s \in \mathcal{S}$. For simplicity, we speak of BSs instead of BS candidate sites henceforth. Each BS has cost c_s and provides a total DL bandwidth b_s which constitutes the capacity of the BS in Hz. Here, "cost" should be regarded as a generalised term standing for either operational cost or power consumption depending on the specific aims of the network planning.

A mobile user, also called user equipment (UE) or just user, has to be covered by (exactly) one BS and requests a certain demand. To incorporate user mobility and to reduce the number of users which have to be taken into account in the optimisation models, we merge demands of users in a small area to a single *traffic demand node* (TN) based on the concept presented by Tutschku et al. [172]. We then denote the set of TNs by $\mathcal{T}$ and each TN $t \in \mathcal{T}$ requests a data rate w_t.

The *Wireless Network Planning Problem* (WNPP) aims, on the one hand, at deciding which BS to deploy and, on the other hand, at assigning TNs to the installed BSs. A BS placement decision consists of the selection of a site and a configuration. Both tasks, BS placement and TN assignment, are subject to several constraints such as interference and resource restrictions. Since future wireless networks utilize advanced transmission techniques such as OFDMA, we do not regard intra-cell interference, which occurs within one cell among different users, in this thesis.

Depending on the aim of the network, several different objectives are possible. On the one hand, the number of deployed BSs should be as small as possible to reduce the cost of the network as well as the total power consumption. On the other hand, a high number of covered TNs is desirable and also a high total throughput or sum rate. Hence, the objective of a WNPP is usually a multi-objective with two or more contrary functions.

To guarantee a certain link quality, we introduce the parameter e_{st} denoting the spectral efficiency for the link from BS s to TN t. This parameter gives the ratio between achievable data rate and available bandwidth; see Section 3.3. To establish a transmission link, the spectral efficiency must exceed a certain threshold $e_{\min} > 0$. Based on this constraint, we

Symbol	Parameter Description
$\mathcal{S}$	index set of BS candidate sites
$s \in \mathcal{S}$	representative index for a BS
$\mathcal{T}$	index set of TNs
$t \in \mathcal{T}$	representative index for a TN
c_s	cost to deploy BS s
b_s	total DL bandwidth available at BS s
w_t	requested data rate at TN t
e_{st}	supported spectral efficiency from BS s to TN t
$e_{\min}$	minimum required spectral efficiency to establish a link
	Variable Description
$x_s \in \{0, 1\}$	BS placement decision, with $x_s = 1$ if s is deployed
$z_{st} \in \{0, 1\}$	TN assignment decision with $z_{st} = 1$ if t is served by s
$u_t \in \{0, 1\}$	TN non-coverage decision, with $u_t = 1$ if t is **not** served

Table 4.1.: General parameters and variables for the formulation of the WNPP.

define the following auxiliary sets of indices.

$$\begin{aligned} \mathcal{S} * \mathcal{T} &:= \{(s,t) \in \mathcal{S} \times \mathcal{T} \mid e_{st} \geq e_{\min}\}, \\ \mathcal{S}_t &:= \{s \in \mathcal{S} \mid (s,t) \in \mathcal{S} * \mathcal{T}\} && \forall\, t \in \mathcal{T}, \\ \mathcal{T}_s &:= \{t \in \mathcal{T} \mid (s,t) \in \mathcal{S} * \mathcal{T}\} && \forall\, s \in \mathcal{S}. \end{aligned}$$

The set $\mathcal{S} * \mathcal{T}$ consists of all BS-TN pairs for which the establishment of a link is possible. Based on this set, $\mathcal{S}_t$ is the set of all BSs which can provide the minimum required spectral efficiency to TN t. Similarly, $\mathcal{T}_s$ denotes the set of TNs for which a link to BS s has sufficient spectral efficiency.

Based on the spectral efficiency, the amount of bandwidth to be allocated to TN t from BS s to serve the total required bit rate w_t is given by w_t/e_{st}. All parameters introduced so far are summarised in Table 4.1.

The variables we utilise in our problem formulations are denoted as follows and also listed in Table 4.1. Let $x_s \in \{0, 1\}$ indicate whether or not BS $s \in \mathcal{S}$ is deployed and $z_{st} \in \{0, 1\}$ whether TN t is assigned to BS s with $(s, t) \in \mathcal{S} * \mathcal{T}$. Furthermore, we introduce a slack variable u_t which is equal to one if TN t is *not* served by *any* BS. Thus, the correlation between variables z and u can be formulated as the following *coverage constraints*.

$$\sum_{s \in \mathcal{S}_t} z_{st} + u_t = 1 \quad \forall\, t \in \mathcal{T} \tag{4.1}$$

They ensure that all TNs are either covered by exactly one BS or not covered at all.

Further constraints inevitable in any model of the WNPP are the capacity constraints,

which ensure that the bandwidth allocated to the TNs served by a single BS does not exceed its capacity. This can be modelled as the following *knapsack constraint* with variable right hand side which is also related to a bin packing problem.

$$\sum_{t \in \mathcal{T}_s} \frac{w_t}{e_{st}} z_{st} \leq b_s x_s \quad \forall\, s \in \mathcal{S}. \tag{4.2}$$

Note, these constraints particularly guarantee that a TN can be assigned to a BS if and only if this BS is deployed.

In this thesis, we focus on cost minimisation as an objective, where the installation of BSs and unsatisfied users entail cost. Hence, the WNPP referred to in this thesis minimises the number of deployed BSs while the number of served TNs is maximised. To combine the two conflicting objectives, we introduce a scaling parameter $\lambda > 0$. The objective function then reads

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \lambda \sum_{t \in \mathcal{T}} u_t. \tag{4.3}$$

The complete *basic formulation of the WNPP* is given as the following ILP.

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \lambda \sum_{t \in \mathcal{T}} u_t \tag{4.4a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{4.4b}$$

$$\sum_{t \in \mathcal{T}_s} \frac{w_t}{e_{st}} z_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{4.4c}$$

$$x_s, z_{st}, u_t \in \{0, 1\} \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T},\ t \in \mathcal{T} \tag{4.4d}$$

Due to the capacity constraints (4.4c), the problem is NP-hard.

Valid inequalities such as Gomory cuts are internally generated by state-of-the-art ILP solvers such as CPLEX [98]. However, ILP solvers cannot take advantage of the particular problem structure known to the user. One type of problem-specific cutting planes for the presented model of the WNPP are *variable upper bounds* [173]. Constraints (4.2) implicitly ensure that a TN can be assigned to a BS if and only if this BS is deployed. It is well-known that the problem formulation can be strengthened by adding these constraints explicitly as

$$z_{st} \leq x_s \quad \forall\, (s,t) \in \mathcal{S} * \mathcal{T}. \tag{4.5}$$

Henceforth, we denote these inequalities as *vub constraints*.

4.2. Fixed Broadband Wireless Networks

The wireless network presented in the previous section represents the access network in a communication network. As explained in Chapter 3, the backhaul network provides interconnectivity between the access and the core network. In this chapter, we introduce

a backhaul network transmitting via microwave links. Microwave radio transmission is a preferred technology for backhaul networks, especially when fibre or copper lines are too costly or impracticable, e. g., in emerging countries [8], as it can be deployed rapidly and cost-efficiently.

The type of microwave transmission we consider in this work is terrestrial fixed point-to-point digital radio communications or briefly fixed broadband wireless communications. The employed antennae at a radio base station (RBS) are highly directional transmitting and receiving energy mainly in/from one specific direction and are positioned in clear line-of-sight. Thus, these networks are typically not affected adversely by interference. Additionally, the antennae operate in licensed frequency bands of 6 to 38 GHz, but the frequency range at which commercial communication systems can be deployed is continuously expanded [134].

One major challenge is the planning of capacity in these wireless networks which is fairly different from wired network planning. On the one hand, the frequency spectrum is a limited natural resource restricting the available bandwidth. On the other hand, environmental conditions such as weather can lead to channel fluctuations and hence, to capacity variations. This is a reason for the common practice of conservative planning, which deploys as much bandwidth as needed in a worst-case planning scenario. We state a MILP formulation for such a conservative planning that assigns bandwidth to RBSs while respecting the traffic requirements and minimising the cost in the following paragraph. This MILP forms the basis for our investigations of reliable formulations, which incorporate channel fluctuations but are less conservative, in Chapter 7.

A MILP Formulation The model presented in the following is based on [57] and assumes digital QAM, which is the most commonly used modulation in microwave systems; see Section 3.4.

To formulate the minimum cost design of a *fixed broadband wireless network* (FBWN) mathematically, we model each RBS as a node in a directed graph $G = (\mathcal{V}, \mathcal{A})$ where each arc $uv \in \mathcal{A}$ represents a microwave link from RBS u to RBS v with $u \neq v$. Henceforth, we use "arc" and "link" interchangeably for $uv \in \mathcal{A}$. By $\delta^-(v)$ ($\delta^+(v)$), we denote the set of incoming (outgoing) neighbours of v. For every link $uv \in \mathcal{A}$, let P_{uv} be the number of available bandwidth choices with index $p = 1, \ldots, P_{uv}$, capacity $B^p_{uv} > 0$ (determined from Table 3.1), and cost $c^p_{uv} > 0$. In the planning, we assume that a modulation scheme is fixed for every bandwidth choice p. The traffic requirements are expressed as commodities and summarised in the set $\mathcal{K}$. For every $k \in \mathcal{K}$, s^k denotes the source (node), t^k the target (node), and d^k the demand value.

To determine the bandwidth assignment and the network flows such that the total cost is minimised, we introduce two types of variables, x^k_{uv} and y^p_{uv}. The flow variable $x^k_{uv} \geq 0$ denotes the amount of demand d^k of commodity k routed on arc $uv \in \mathcal{A}$ while the decision variable $y^p_{uv} \in \{0, 1\}$ indicates whether bandwidth choice p with capacity B^p_{uv} is assigned to

arc $uv \in \mathcal{A}$ or not. The FBWN planning problem can then be formulated as:

$$\min \sum_{uv \in \mathcal{A}} \sum_{p=1}^{P_{uv}} c_{uv}^p y_{uv}^p \tag{4.6a}$$

$$\text{s.t.} \sum_{u \in \delta^-(v)} x_{uv}^k - \sum_{u \in \delta^+(v)} x_{vu}^k = \begin{cases} -d^k, & v = s^k \\ d^k, & v = t^k \\ 0, & \text{otherwise} \end{cases} \qquad \forall\, v \in \mathcal{V},\, k \in \mathcal{K} \tag{4.6b}$$

$$\sum_{k \in \mathcal{K}} x_{uv}^k \leq \sum_{p=1}^{P_{uv}} B_{uv}^p y_{uv}^p \qquad \forall uv \in \mathcal{A} \tag{4.6c}$$

$$\sum_{p=1}^{P_{uv}} y_{uv}^p \leq 1 \qquad \forall uv \in \mathcal{A} \tag{4.6d}$$

$$x_{uv}^k \geq 0, y_{uv}^p \in \{0, 1\} \qquad \forall uv \in \mathcal{A},\, k \in \mathcal{K},\, p = 1, \ldots, P_{uv}. \tag{4.6e}$$

The objective function (4.6a) represents the total bandwidth cost we aim at minimising. The flow conservation constraints (4.6b) guarantee that every flow entering a node which is neither source nor target also leaves the node. Additionally, for every source and target, the total traffic requirement has to be fulfilled. Constraints (4.6c) are the capacity constraints which ensure that the chosen bandwidth is sufficient to route the total assigned flow on each link. Finally by means of constraints (4.6d), at most one bandwidth is chosen for each arc, which implies that if no bandwidth is chosen, the link is not operated.

Model (4.6) is a minimum cost multi-commodity flow problem which is widely used to formulate problems arising in telecommunication, see [140] for a survey. Additionally, (4.6) is a special case of a network design problem and thus, strongly NP-hard [45].

Since a modulation scheme per bandwidth choice is fixed, the network modelled by (4.6) cannot react to outage events. This means, if a link is established with a high-level modulation scheme, this link is more likely to fail and then traffic might be lost. To handle such link outages, adaptive modulation and coding (AMC) is usually employed in modern microwave links [84]. Hence, the radio configuration is a random factor which has to be taken into account already during the planning of a reliable network. One possibility to model FBWNs under outage probability constraints are so-called chance constraints which are investigated in Part II of this thesis.

5. Interference Modelling in Wireless Access Networks

A network set up by a straightforward application of the basic model (4.4) presented in Section 4.1 could suffer from severe problems caused by inter-cell interference. For instance, TNs which are claimed as covered in an optimal solution of the model might not receive a sufficiently large signal power to fulfil the requested bit rates due to other BSs interfering the signal. Hence, to improve the performance of the network, interference has to be taken into account already in the planning phase. There exist several techniques and approaches to address inter-cell interference by influencing the signal transmission, which are especially suitable to decrease/eliminate interference during operation of the wireless network. Examples of such approaches are randomisation, where interference is averaged across the spectrum, signal processing techniques to eliminate interference completely, multi-antenna techniques using several antennae simultaneously for signal transmission and reception, or coordination, where interference is handled via restrictions to the utilisable resources; see Kosta et al. [115] for more details.

In this chapter, we focus on inter-cell interference free planning of wireless networks, where the proposed approaches are based on SINR conditions, which can guarantee a sufficient quality of an established link. We present conventional SINR constraints in Section 5.2, which can, however, lead to numerical instabilities due to huge differences in magnitude of the coefficients. Hence, we propose approximate as well as novel exact approaches to model inter-cell interference via SINR conditions but avoid the explicit incorporation of the corresponding constraints.

Before investigating different approaches and proposing corresponding formulations, we give some preliminaries on the computation of spectral efficiencies and the evaluation of the exactness of the formulations in Section 5.1. Additionally, we present the conventional SINR constraint formulation in Section 5.2. Afterwards, we first state three approximate formulations and then three exact and novel ones. The first approximate approach is the concept of a conflict graph presented in Section 5.3. By means of a conflict graph, we can on the one hand position the transmitting BSs sufficiently far away from each other such that only marginal inter-cell interference occurs, and on the other hand extend the definition of a BS candidate. A more refined but still approximate approach is presented in Section 5.4 where we require that the ratio between the spectral efficiencies corresponding to the serving and to any interfering BS exceeds a certain threshold. The last approximate approach is proposed in Section 5.5 which depicts an iterative algorithm updating the values for the spectral efficiencies according to the actual deployed BSs. Afterwards, we present three novel exact approaches and corresponding formulations in Sections 5.6 to 5.8. The first formulation mitigates interference from a BS point of view, where the

second formulation is a TN oriented model. The approach in Section 5.8, exploits the fact that it is sufficient to consider only a discrete number of different spectral efficiencies.

Finally, we find good values for formulation specific parameters in a computational study in Section 5.9 and compare the different approaches in terms of a SINR-corrected objective value which regards also violated SINR conditions.

5.1. Preliminaries

In most of the approaches presented in the following sections, it is necessary to compute the SINR of a link from a BS to a TN; see equation (3.3) in Section 3.3 for the definition of SINR. The received power at TN t from BS s is generally computed as

$$P_{\mathrm{r}}(s,t) = p_s a_{st},$$

where p_s denotes the transmission power of BS s and a_{st} the *fading coefficient* of the link from BS s to TN t, which can incorporate shadow fading and path loss. A general SINR requirement/condition can be formulated as

$$\gamma_{st} \coloneqq \frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{\sigma \in \mathcal{S}_{st}} P_{\mathrm{r}}(\sigma,t) + \eta} \geq \delta, \tag{5.1}$$

where δ is a predefined threshold and η denotes the background noise, which particularly depends on the bandwidth. The largest possible set of interferers $\mathcal{S}_{st}$ is $\mathcal{S}_t \setminus \{s\}$ as a BS can only interfere a signal to t if it is able to cover this TN.

Whenever it is inevitable to specify the fading coefficient explicitly, for instance in the computational study in Section 5.9, we consider only path loss and disregard shadow fading. Hence, the fading coefficient is then computed as

$$a_{st} = 10^{-\frac{1}{10} P_{\mathrm{L}}^{\mathrm{dB}}(s,t)},$$

where $P_{\mathrm{L}}^{\mathrm{dB}}(s,t)$ denotes the path loss of the signal from s to t given in dB.

Based on the SINR value, the spectral efficiency e_{st} is exactly calculated as

$$e_{st} = \log_2(1 + \gamma_{st}). \tag{5.2}$$

To avoid these numerical difficult values, we approximate the spectral efficiencies by a stepwise function. Hence, we use the look-up Table 5.1 taken from Sesia et al. [162], which maps SINR and thus, also SNR values given in dB to spectral efficiencies for a bandwidth of 10 MHz. More precisely, a range of SINR values is mapped to one discrete spectral efficiency. For example, the interval $[-5.1, -2.9)$ is associated with the spectral efficiency 0.25. In Figure 5.1, we display the function (5.2) and the stepwise approximation.

As the spectral efficiency is an indicator for the quality of a signal or channel, each line in Table 5.1 is labelled by an index called *Channel Quality Indicator* (CQI). If not stated

CQI	spec. eff. (bps/Hz)	SINR (dB)
0	out of range	
1	0.25	-5.1
2	0.40	-2.9
3	0.50	-1.7
4	0.66	-1.0
5	1.00	2.0
6	1.33	4.3
7	1.50	5.5
8	1.60	6.2
9	2.00	7.9
10	2.66	11.3
11	3.00	12.2
12	3.20	12.8
13	4.00	15.3
14	4.50	17.5
15	4.80	18.6

Table 5.1.: SINR requirements for LTE and 10 MHz bandwidth according to [162].

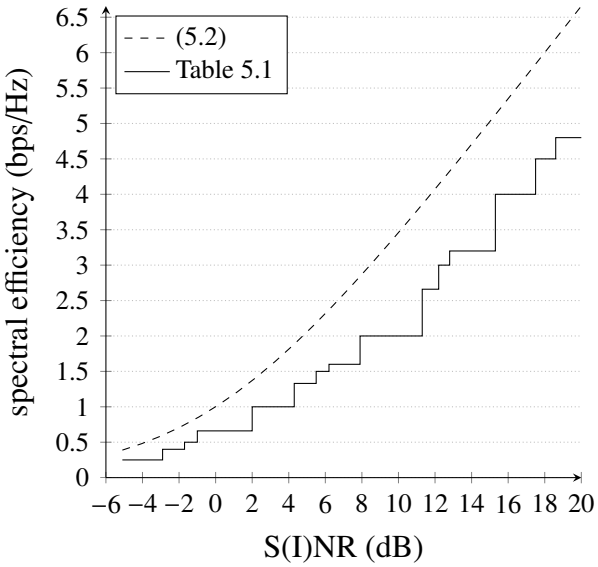


Figure 5.1.: Approximation of the spectral efficiency via look-up Table 5.1.

differently, the values of the spectral efficiencies used in the capacity constraints (4.4c) are based on the SNR values. Since SINR values are usually unkown, it is common practice to use only SNR values as an approximation of the capacity of a wireless link; see Chafekar et al. [41].

For simplicity, we denote by "basic formulation" the ILP (4.4) together with the vub constraints (4.5) in the remainder of this chapter. To evaluate and compare the different approaches proposed in the subsequent sections, we introduce the following notion of exactness.

Definition 5.1. A formulation incorporating interference into the basic formulation is denoted as *exact in terms of SINR* if no violated SINR condition (5.1) exists for an optimal solution.

Note, exactness for any approach is only possible if the minimum spectral efficiency $e_{\min}$ and the SINR threshold δ are chosen accordingly. This means, $e_{\min}$ and δ are set such that δ is the lower bound of the range corresponding to $e_{\min}$ in Table 5.1.

In some approaches proposed in the following sections, we do not restate the capacity constraints (4.4c) which comprise the spectral efficiencies associated to the SNR value. Hence, we cannot guarantee that the BS capacity is not exceeded when the actual SINR values are known for an optimal solution for such formulations. To account for this effect, we introduce the following further notion of exactness.

Section	name	SINR	capacity
5.2	SINR constraint formulation	yes	no
5.3	conflict graph formulation	no	no
5.4	TN coverage requirement form.	no	no
5.5	iterative formulation	no	no
5.6	interference mitigation form.	yes	yes
5.7	TN oriented formulation	yes	yes
5.8	discrete CQIs formulation	yes	yes

Table 5.2.: Summary of exactness in terms of SINR and capacity of the subsequent formulations.

Definition 5.2. A formulation incorporating interference in the basic formulation is denoted as *exact in terms of capacity* if no violated capacity constraint (4.4c), in which the spectral efficiencies are determined via the actual SINR values, exists for an optimal solution.

In Table 5.2, we summarise the exactness in terms of SINR as well as capacity for all formulations discussed in the subsequent sections.

5.2. Conventional SINR Constraints

There exist two most commonly used interference models: graph-based and fading channel or physical models [88]. One example of a graph-based model is discussed in the subsequent Section 5.3 while we propose a physical model in the present section.

The physical model proposed in Gupta and Kumar [91] uses SINR constraints to model interference. Such constraints ensure that the SINR is sufficiently high to establish a physical link from the transmitter (BS) to the receiver (TN); see the SINR requirement (5.1). This model is widely used in the literature to model interference for various kinds of problems occurring in different types of wireless communication networks. We give an overview on several rather recent works in the following. Goussevskaia et al. [87] investigate the problem of scheduling wireless links respecting the SINR model in wireless ad hoc networks, which are decentralised and consist of radio units that can form temporary communication links among each other. Borbash and Ephremides [29] study wireless link scheduling with power control and SINR constraints where the aim is to find a schedule and durations such that all demands are satisfied. Afterwards, power vectors have to be determined such that the SINR requirements are satisfied. In Andrews and Dinitz [9], the authors maximise the number of connections satisfying SINR constraints in arbitrary wireless networks where the transmission powers have to be defined. Furthermore, Ramamurthi et al. [158] study the assignment of channels together with the distribution of capacity and the determination of link flows in wireless mesh networks. They use SINR conditions to determine interfering links for every node. In Johansson and Xiao [101], the

problem of simultaneous routing, resource allocation and scheduling in wireless ad hoc networks is investigated at which the channel capacity is modelled via discrete rate levels with associated SINR targets. Recently, Li et al. [123] investigate two basic approaches of modelling SINR conditions for parallel link transmissions in wireless networks and introduce matching inequalities to improve the optimality gaps.

To include the SINR requirement (5.1) in our basic formulation for wireless network planning, we add the decision variables z_{st} and x_σ denoting the TN assignment and the BS installation and obtain the following constraint.

$$\frac{P_\mathrm{r}(s,t)z_{st}}{\sum\limits_{\sigma\in\mathcal{S}_{st}} P_\mathrm{r}(\sigma,t)x_\sigma + \eta} \geq \delta \;\Leftrightarrow\; P_\mathrm{r}(s,t)z_{st} \geq \delta\left(\sum_{\sigma\in\mathcal{S}_{st}} P_\mathrm{r}(\sigma,t)x_\sigma + \eta\right).$$

The last constraint leads to an undesirable restriction in case that $z_{st} = 0$. Hence, we add a big-M term and obtain the following SINR constraint, which is a big-M constraint.

$$P_\mathrm{r}(s,t)z_{st} + M_{st}(1 - z_{st}) \geq \delta\left(\sum_{\sigma\in\mathcal{S}_{st}} P_\mathrm{r}(\sigma,t)x_\sigma + \eta\right) \quad \forall\,(s,t) \in \mathcal{S} * \mathcal{T}, \tag{5.3}$$

with M_{st} sufficiently large ensuring a restriction only if $z_{st} = 1$. Big-M constraints are in general numerically difficult due to weak LP relaxations and rounding issues. But constraints (5.3) are even worse since the received powers $P_\mathrm{r}(s,t)$ and $P_\mathrm{r}(\sigma,t)$ can vary significantly in magnitude.

Transforming (5.3) such that each variable occurs exactly once leads to

$$(M_{st} - P_\mathrm{r}(s,t))z_{st} + \sum_{\sigma\in\mathcal{S}_{st}} \delta P_\mathrm{r}(\sigma,t)x_\sigma \leq M_{st} - \delta\eta \quad \forall\,(s,t) \in \mathcal{S} * \mathcal{T}. \tag{5.4}$$

In case of $z_{st} = 0$, constraint (5.4) should not present any restriction. If $z_{st} = 0$, (5.4) reduces to

$$\sum_{\sigma\in\mathcal{S}_{st}} \delta P_\mathrm{r}(\sigma,t)x_\sigma \leq M_{st} - \delta\eta \quad \forall\,(s,t) \in \mathcal{S} * \mathcal{T}.$$

Hence, we define

$$M_{st} \coloneqq \sum_{\sigma\in\mathcal{S}_{st}} \delta P_\mathrm{r}(\sigma,t) + \delta\eta \quad \forall\,(s,t) \in \mathcal{S} * \mathcal{T}. \tag{5.5}$$

By means of the SINR constraints (5.3) or (5.4), respectively, a link from a BS to a TN is established if and only if the associated SINR is sufficiently large. Hence, no interference occurs for an optimal solution. However, we cannot guarantee that the capacity constraints (4.4c) are actually satisfied as the spectral efficiencies used for the initialisation are not updated after the actual SINR values are known possibly causing an excess of the BS capacity.

Remark 5.3. The basic formulation together with the conventional SINR constraints (5.3) or (5.4), which we call henceforth the *SINR constraint formulation*, yields an exact formulation in terms of SINR but not in terms of capacity.

Proof. Exactness in terms of SINR is clear by definition. To show a possible violation of a capacity constraint, we consider the following example. We assume a solution with $z_{st} = 1$ for one $(s,t) \in \mathcal{S} * \mathcal{T}$ and $t \in \mathcal{T}_\sigma$ for a BS $\sigma \in \mathcal{S} \setminus \{s\}$ with $x_\sigma = 1$ such that the capacity of BS s is fully used. Moreover, let $P_{\mathrm{L}}^{\mathrm{dB}}(s,t) = 130\,\mathrm{dB}$ and $P_{\mathrm{L}}^{\mathrm{dB}}(\sigma,t) = 140\,\mathrm{dB}$ such that $e_{st} = 2$ based on the SNR value. The spectral efficiency based on the correct SINR value e'_{st} with σ interfering the signal from s to t is only 0.66. Hence, $w_t/e'_{st} > w_t/e_{st}$ causing the violation of the capacity of BS s. □

We investigate the numerical stability of the SINR constraint formulation in a computational study in Section 5.9.3.

Capone et al. [39] derive cover inequalities to replace the conventional SINR constraints (5.3) and to overcome numerical instabilities. We briefly summarise their approach in the following.

As the big-M constraint (5.3) is only restrictive if TN t is assigned to BS s, it reduces to

$$P_{\mathrm{r}}(s,t) \geq \delta \left(\sum_{\sigma \in \mathcal{S}_{st}} P_{\mathrm{r}}(\sigma,t) x_\sigma + \eta \right). \tag{5.6}$$

if $z_{st} = 1$. Defining

$$r_{st} \coloneqq \frac{P_{\mathrm{r}}(s,t)}{\delta} - \eta,$$

constraint (5.6) is equivalent to

$$\sum_{\sigma \in \mathcal{S}_{st}} P_{\mathrm{r}}(\sigma,t) x_\sigma \leq r_{st}, \tag{5.7}$$

which is a knapsack constraint. As described in Section 1.3, the solving of models comprising such constraints can be improved by separating cover inequalities. In the present context, a subset $C_{st} \subseteq \mathcal{S}_{st}$ is a cover if

$$\sum_{\sigma \in C_{st}} P_{\mathrm{r}}(\sigma,t) > r_{st}.$$

The corresponding cover inequality then reads

$$\sum_{\sigma \in C_{st}} x_\sigma \leq |C_{st}| - 1.$$

However, this inequality should only be restrictive if TN t is assigned to BS s, i. e., if $z_{st} = 1$. Hence, we replace the 1 at the right hand side by the assignment variable and obtain the following cover inequality.

$$\sum_{\sigma \in C_{st}} x_\sigma \leq |C_{st}| - z_{st}$$

Note, any integer solution fulfilling every cover inequality is a feasible solution of the original knapsack constraint (5.7).

We do not take this formulation into account in the computational study in Section 5.9 as it has been numerically evaluated in detail in [39].

5.3. Conflict Graph

Our first approach to limit inter-cell interference and avoid SINR constraints is to stipulate that the installed BSs constitute an *independent set* in a predefined *conflict graph* $G = (\mathcal{S}, \mathcal{E})$; see [49]. An independent set in a graph is a subset $\mathcal{S}' \subseteq \mathcal{S}$ of the vertices such that there does not exist an edge $ij \in \mathcal{E}$ for all $i, j \in \mathcal{S}'$. This restriction is reflected in the following constraint, which we add to the basic formulation.

$$x_i + x_j \leq 1 \quad \forall\, ij \in \mathcal{E} \tag{5.8}$$

The concept of conflict or interference graphs is a commonly applied graph-based interference model; see, e. g., Grönkvist and Hansson [89]. It has been employed in the planning of GSM (global system for mobile communications) networks [136], WLANs (wireless local area networks) [159], and LTE networks [66], and in a modified way via complement sets for the deployment of cooperation clusters in general wireless cellular networks [146].

We can strengthen constraints (5.8) by replacing them by *maximal clique inequalities* [150] where a clique is a complete subgraph. Formally speaking, let $\mathcal{U} \subseteq \mathcal{S}$ be a subset of the vertex set. We call $\mathcal{U}$ a clique of conflict graph G if there exists an edge $ij \in \mathcal{E}$ for all $i,\ j \in \mathcal{U}$. A clique is maximal if it is not included in a larger clique.

Constraints (5.8) describe an independent set polytope. Thus, we can replace these constraints by all maximal clique inequalities

$$\sum_{s \in \mathcal{U}} x_s \leq 1 \quad \forall\, \mathcal{U} \subset \mathcal{S},\ \mathcal{U} \text{ is a maximal clique in } G = (\mathcal{S}, \mathcal{E}). \tag{5.9}$$

The problem of finding maximal cliques is NP-complete and maximal cliques can be computed by the Bron-Kerbosch algorithm [31] with complexity $O(3^{n/3})$ [169] which is fast in practice.

The conflict graph is a quite versatile concept as it can on the one hand limit the inter-cell interference and on the other hand it allows for more general candidate sites. For the inter-cell interference limitation, the conflict graph can be defined such that two BSs are adjacent if and only if the distance between them is less than or equal to a minimum required distance $d_{\min}$. For macro cells, $d_{\min}$ is usually set to 500 m; see, e. g., [32, 102, 165]. Note that the resulting conflict graph is a *unit disk graph* which is an intersection graph of equal sized circles in the plane. Gupta et al. [92] present an approximation polynomial in the number of edges m with complexity $O(m\Delta^2)$ (Δ maximal degree) to compute maximal cliques in such graphs.

Remark 5.4. The basic formulation together with constraints (5.8) or the equivalent maximal clique inequalities (5.9) defined via the minimum distance requirement, which we call

henceforth the *conflict graph formulation*, does not yield an exact formulation neither in terms of SINR nor in terms of capacity.

Proof. A minimum distance between deployed BSs cannot prevent the violation of a SINR, e. g., for a TN in the middle of two BSs. The fading coefficient for a signal from a not-serving BS can still be sufficiently high to disturb the signal from the serving BS. Moreover, a violated SINR condition directly implies the violation of a capacity constraint if the capacity is fully used since an unreasonable high bandwidth would be required to serve the TN with insufficient SINR. □

Without the conflict graph concept, a BS candidate consists of a location and a fixed (antenna) configuration. But for each position there can only exist one candidate, i. e., one configuration. By means of the conflict graph, we can introduce several candidates for the same site with different configurations. Since at most one BS can be installed per site, an edge of the graph is then formed by two BS candidates located at the same position with distinct settings.

5.4. TN Coverage Requirement

Based on the paper by Engels et al. [67], we demand that a TN t can be covered by a BS s if the ratio between the spectral efficiencies of the serving BS s and any interfering BS σ exceeds a threshold δ_c, which is related to SINR thresholds. This is modelled by the following *TN coverage requirement*.

$$\frac{e_{st}}{e_{\sigma t}} \geq \delta_c \quad \forall\, \sigma \in \mathcal{S}_{st}, \tag{5.10}$$

$\mathcal{S}_{st}$ denotes the set of BSs interfering the signal from s to t. As explained in Section 5.2, it holds $\mathcal{S}_{st} \subseteq \mathcal{S}_t \setminus \{s\}$. Note that $e_{\sigma t} \geq e_{\min} > 0$ per definition of $\mathcal{S}_t$. A straightforward formulation of the corresponding model constraints is given as follows.

$$\frac{e_{st}}{e_{\sigma t}} \geq (z_{st} + x_\sigma - 1)\delta_c \quad \forall\, t \in \mathcal{T},\ s \in \mathcal{S}_t,\ \sigma \in \mathcal{S}_{st}, \tag{5.11}$$

where s is the serving and σ is an interfering BS. A constraint for $t \in \mathcal{T},\ s \in \mathcal{S}_t,\ \sigma \in \mathcal{S}_{st}$ is restrictive if and only if t is covered by s ($z_{st} = 1$) and σ is installed ($x_\sigma = 1$). However, these constraints can be improved by the following reformulation which shifts the TN coverage requirement to the domain reducing the number of constraints significantly.

$$z_{st} + x_\sigma \leq 1 \quad \forall\, t \in \mathcal{T},\ s \in \mathcal{S}_t,\ \sigma \in \mathcal{S}_{st} \text{ with } \frac{e_{st}}{e_{\sigma t}} < \delta_c \tag{5.12}$$

Note, (5.12) is also a type of a conflict graph with (s, t) and σ forming an edge if $\frac{e_{st}}{e_{\sigma t}} < \delta_c$.

The TN coverage requirement can only guarantee a certain link quality but has no influence on the actual transmission rate. Hence, it depicts a reasonable approach for coverage maximisation, where the number of not covered TNs is minimised. However, violations of SINR requirements (5.1) cannot be eliminated.

Remark 5.5. The basic formulation together with the TN coverage requirement constraints (5.12), which we call henceforth the *TN coverage requirement formulation*, does not yield an exact formulation neither in terms of SINR nor in terms of capacity.

Proof. Requirement (5.12) can only eliminate the worst interference. However, the sum of all interfering signals is included in the SINR. Thus, violated SINR conditions are possible even though every single interfering signal is not too strong.

With the same argumentation as in the proof of Remark 5.4, violated capacity constraints are not excluded. □

In Section 5.9.3, we investigate various values in a computational study to find a good threshold δ_c.

5.5. An Iterative Formulation

In this section, we develop an iterative formulation where the spectral efficiencies are updated in every iteration by computing new SINR values.

For the initial values of the spectral efficiencies e^0_{st}, we compute initial SNR values γ^0_{st} without interference; cf. (3.1).

$$\gamma^0_{st} := \frac{P_r(s,t)}{\eta} \tag{5.13}$$

The corresponding spectral efficiency e^0_{st} is then extracted from the look-up Table 5.1. Moreover, we denote the number of violated SINR constraints at iteration i regarding threshold δ by U^i and initialise this value as

$$U^0 := |\mathcal{T}|.$$

Furthermore, we denote the number of deployed BSs by ς^i and the number of covered TNs by τ^i at iteration i with initial values 0.

The ILP we solve at iteration $i \geq 0$ is the basic formulation but with iteration specific spectral efficiencies:

$$\min \sum_{s\in\mathcal{S}} c_s x_s + \lambda \sum_{t\in\mathcal{T}} u_t \tag{5.14a}$$

$$\text{s.t.} \sum_{s\in\mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{5.14b}$$

$$\sum_{t\in\mathcal{T}_s} \frac{w_t}{e^i_{st}} z_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{5.14c}$$

$$z_{st} \leq x_s \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.14d}$$

$$x_s, z_{st}, u_t \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T},\ t \in \mathcal{T}. \tag{5.14e}$$

For a solution (x^i, z^i, u^i) of ILP (5.14) at iteration i, we compute U^i, ς^i and τ^i as follows.

$$U^i = \left|\left\{t \in \mathcal{T} \,\middle|\, \gamma^{i+1}_{st} < \delta \text{ with } s \in \mathcal{S}_t \text{ and } z^i_{st} = 1\right\}\right|, \text{ with}$$

$$\gamma_{st}^{i+1} = \frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{\sigma \in \mathcal{S}_t \setminus \{s\}} P_{\mathrm{r}}(\sigma,t) x_\sigma^i + \eta}, \tag{5.15}$$

$$\varsigma^i = \left| \left\{ s \in \mathcal{S} \,\middle|\, x_s^i = 1 \right\} \right| = \sum_{s \in \mathcal{S}} x_s^i,$$

$$\tau^i = \left| \left\{ t \in \mathcal{T} \,\middle|\, u_t^i = 0 \right\} \right| = \sum_{t \in \mathcal{T}} u_t^i.$$

We extract the spectral efficiencies e_{st}^{i+1} from Table 5.1 according to the SINR values computed in (5.15) for every pair $(s,t) \in \mathcal{S} * \mathcal{T}$ with $z_{st}^i = 1$. Note, the set $\mathcal{S} * \mathcal{T}$ has to be updated so that (s,t)-pairs with $e_{st}^{i+1} < e_{\min}$ are excluded. The sets $\mathcal{S}_t$ and $\mathcal{T}_s$ are then updated accordingly. Afterwards, we solve ILP (5.14) again with the new spectral efficiencies unless $(U^i, \varsigma^i, \tau^i) = (U^{i-1}, \varsigma^{i-1}, \tau^{i-1}) = (U^{i-2}, \varsigma^{i-2}, \tau^{i-2})$ for $i \geq 2$. This means, a stop criterion catches if the parameters based on the solution have not changed three times in succession. Note, two times do not suffice as can be seen later on in Figure 5.2 on page 49. Without this exit condition, we cannot ensure a termination of the algorithm.

The complete iterative algorithm is summarised in Algorithm 2.

Algorithm 2 Iterative interference modelling

Input: basic parameters of the WNPP and SINR threshold δ

Output: (U, ς, τ) and corresponding solution (x, z, u)

Initialisation:

compute γ_{st}^0 regarding (5.13) and e_{st}^0 according to Table 5.1

define $(U^0, \varsigma^0, \tau^0) := (|\mathcal{T}|, 0, 0)$

$i = 0$, bool CONTINUE=true

while CONTINUE **do**

solve ILP (5.14) with spectral efficiencies e_{st}^i obtaining solution (x^i, z^i, u^i)

compute $(U^i, \varsigma^i, \tau^i)$

if $(U^i, \varsigma^i, \tau^i) = (U^{i-1}, \varsigma^{i-1}, \tau^{i-1}) = (U^{i-2}, \varsigma^{i-2}, \tau^{i-2})$ **then** CONTINUE=false

else

compute γ_{st}^{i+1} according to (5.15) for $(s,t) \in \mathcal{S} * \mathcal{T}$ with $z_{st}^i = 1$

set $\gamma_{st}^{i+1} = \gamma_{st}^i$ for the remaining $(s,t) \in \mathcal{S} * \mathcal{T}$

set the associated values e_{st}^{i+1}

update $\mathcal{S} * \mathcal{T}$, $\mathcal{S}_t$, and $\mathcal{T}_s$

set $i = i + 1$

end if

end while

return $(U^i, \varsigma^i, \tau^i)$ and (x^i, z^i, u^i)

Remark 5.6. Algorithm 2, which we call henceforth the *iterative formulation*, is neither exact in terms of SINR unless $U^i = 0$ nor in terms of capacity.

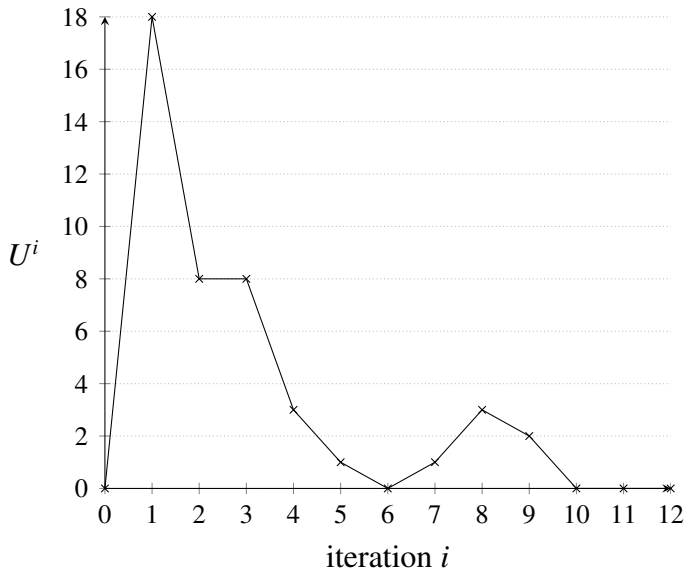


Figure 5.2.: Number of violated SINR conditions U^i per iteration i for a scenario with 40 BS candidates and 450 TNs.

Proof. Algorithm 2 is in general not exact in terms of SINR as $U^i > 0$ is possible for the computed solution (x, z, u). Moreover, the spectral efficiencies used in the computation of a feasible solution do not necessarily correspond to the actual SINR values. Hence, Algorithm 2 is also not exact in terms of capacity. □

Modifications To avoid excessive running times, we apply the following modifications. First, we introduce a limit $\varkappa$ on the integrality gap of ILP (5.14) such that the solving process of the ILP is aborted if the percentage gap between lower and upper bound is below $\varkappa$. Reasonable values for $\varkappa$ are, e. g., 1, 2 or 5 %. Second, as it is possible that an excessive amount of time is needed until the gap limit takes effect, we additionally introduce a time limit for the solving of each ILP. This time limit depends on the global time limit. Since time limits are machine depending, we lose the determinism in case that such a limit takes effect.

The solutions, especially the value of U^i, fluctuate during the iterations. We display fluctuating values of U^i exemplarily for a test scenario with 40 BS candidates and 450 TNs in Figure 5.2 using the best setting determined in Section 5.9.3. Since the solution with the lowest U^i is desired, we compare the current U^i to the best known value U_{best} and update the latter value if necessary at every iteration. Additionally, we save and update ς_{best}, τ_{best} and $(x_{\text{best}}, z_{\text{best}}, u_{\text{best}})$ corresponding to U_{best}. The modified Algorithm 2 then returns $(U_{\text{best}}, \varsigma_{\text{best}}, \tau_{\text{best}})$ and $(x_{\text{best}}, z_{\text{best}}, u_{\text{best}})$.

Finally, the current computation of the values γ_{st}^{i+1} is quite conservative as it keeps (lower) values for (s, t)-pairs for which no link is established in the current ILP solution (x^i, z^i, u^i) but might have been established in a previous solution. Therefore, we present the following variant to compute all γ_{st}^{i+1} values.

$$\gamma_{st}^{i+1} = \begin{cases} \frac{P_r(s,t)}{\sum_{\sigma \in \mathcal{S}_t \setminus \{s\}} P_r(\sigma,t) x_\sigma^i + \eta} & \text{if } z_{st}^i = 1, \\ \frac{P_r(s,t)}{P_r(\sigma,t) + \eta} & \text{if } z_{st}^i = 0 \text{ and } \exists\, \sigma \in \mathcal{S}_t \setminus \{s\} \text{ with } z_{\sigma t}^i = 1 \\ \gamma_{st}^0 & \text{if } u_t^i = 1 \end{cases} \tag{5.16}$$

Here, the first case is the same as before. In the second case, when TN t is assigned to a BS σ in the current solution, we set $\gamma_{st}^{i+1} = \gamma_{st}^i$ for all remaining BSs $s \in \mathcal{S}_t$, $s \neq \sigma$ in the former version of Algorithm 2. But these values are, in general, too low since the link from s to t is currently not established and the former interfering BSs might also not (all) be deployed in the present solution. Instead, we now define a lower bound on the SINR value, where the currently serving BS σ is handled as the only interfering BS of the signal from BS s to t. Eventually, if TN t is currently not covered, the SINR values for every BS $s \in \mathcal{S}_t$ are set back to the initial values resembling only SNR.

Preliminary computational tests have revealed that the SINR update routine (5.16) gives better results with regard to the number of violated SINR conditions U^i than (5.15). On that account, we use only (5.16) to update the SINR values henceforth.

The SINR calculation (5.16) assumes constant signal transmission also from the interfering BSs. However in real world wireless communication networks, BSs are not transmitting constantly. Hence, as a further variant, we incorporate also the percentage load of a BS in the calculation of the SINR values as follows. The percentage load ℓ_s^i in iteration i is defined as

$$\ell_s^i := \frac{1}{b_s} \sum_{t \in \mathcal{T}_s} \frac{w_t}{e_{st}^i} z_{st}^i.$$

The SINR extended by the load of the interfering BSs is computed as

$$\gamma_{st}^{i+1} = \begin{cases} \frac{P_r(s,t)}{\sum_{\sigma \in \mathcal{S}_t \setminus \{s\}} \ell_\sigma^i P_r(\sigma,t) x_\sigma^i + \eta} & \text{if } z_{st}^i = 1, \\ \frac{P_r(s,t)}{\ell_\sigma^i P_r(\sigma,t) + \eta} & \text{if } z_{st}^i = 0 \text{ and } \exists\, \sigma \in \mathcal{S}_t \setminus \{s\} \text{ with } z_{\sigma t}^i = 1, \end{cases} \tag{5.17}$$

where the value for $u_t^i = 1$ remains as before; see (5.16).

By means of a computational study in Section 5.9.3, we intend to find good values for the gap limit $\varkappa$ as well as time limit and to decide if the inclusion of load in the SINR calculation can improve the performance.

5.6. Interference Mitigation

Interference mitigation techniques have recently been developed for LTE networks especially for heterogeneous networks. These techniques include but are not limited to the detection of cells of other types (macro, pico, femto), and power and encoding adjustment. Furthermore, interference can be mitigated by avoiding the worst interferers during the BS placement decision and the assignment decision. We propose a variant of interference

mitigation in this section developed from the concept presented by Kosta et al. [115].

For every BS $s \in \mathcal{S}$, we define $\Im_s \subseteq \mathcal{S} \setminus \{s\}$ to be the set of BSs which interfere signals from s. Furthermore, let $\mathcal{I}_s$ be the power set of $\Im_s$. By $i \in \{1, \dots, |\mathcal{I}_s|\}$, we denote the index of a set in $\mathcal{I}_s$ and by $\mathcal{I}_s^i$ the set of interferers with index i in $\mathcal{I}_s$. Furthermore, let $\mathcal{H}_s$ be the set of BSs which are interfered by s, i. e., $\mathcal{H}_s := \{\sigma \in \mathcal{S} \setminus \{s\} \mid s \in \Im_\sigma\}$. The aim is now to decide which BS should be installed and at the same time which set of BSs should not be installed to avoid the interference with signals transmitted from the deployed BSs. For that purpose, we define new decision variables

$$x_s^i = \begin{cases} 1, & \text{BS } s \text{ is deployed and } \mathcal{I}_s^i \text{ is the maximal set of interfering BSs not deployed} \\ 0, & \text{otherwise,} \end{cases}$$

where "maximal" means with respect to cardinality. If a BS s is not deployed, i. e., $x_s^i = 0$ for all $i \in \{1, \dots, |\mathcal{I}_s|\}$, then the deployment of at least one BS $\sigma \in \mathcal{H}_s$, which is interfered by s, is possible. But then, a set $\mathcal{I}_\sigma^j$ of interfering BSs that contains s with $j \in \{1, \dots, |\mathcal{I}_\sigma|\}$ cannot be deployed at the same time. In the case that a BS s is deployed, i. e., $x_s^i = 1$ for one $i \in \{1, \dots, |\mathcal{I}_s^i|\}$, no interfered BS $\sigma \in \mathcal{H}_s$ can be deployed if its deployment requires that a set $\mathcal{I}_\sigma^j$ with $j \in \{1, \dots, |\mathcal{I}_\sigma|\}$ containing s is not installed. These interrelations and the connection to the former decision variables x_s are summarised in the following model constraints.

$$\sum_{i=1}^{|\mathcal{I}_s|} x_s^i + \sum_{\substack{j \in \{1, \dots, |\mathcal{I}_\sigma|\} : \\ s \in \mathcal{I}_\sigma^j}} x_\sigma^j \leq 1 \qquad \forall\, s \in \mathcal{S},\, \sigma \in \mathcal{H}_s \tag{5.18}$$

$$\sum_{i=1}^{|\mathcal{I}_s|} x_s^i = x_s \qquad \forall\, s \in \mathcal{S} \tag{5.19}$$

Note, equations (5.19) do not have to be added as model constraints and also variables x_s are dispensable. We just quoted these equations for a better understanding here.

For any installed BS s, the selection of $i \in \{1, \dots, |\mathcal{I}_s|\}$ impacts the SINR of all TNs assigned to s. If the interfering BSs in $\mathcal{I}_s^i$ are not deployed and all other interfering BSs are deployed, the SINR for the signal from s to t is denoted by

$$\gamma_{t,s}^i = \frac{P_\mathrm{r}(s,t)}{\sum\limits_{\sigma \in \left(\mathcal{S}_t \setminus \mathcal{I}_s^i\right) \cap \Im_s} P_\mathrm{r}(\sigma,t) + \eta} \quad \forall\, (s,t) \in \mathcal{S} * \mathcal{T},\, i \in \{1, \dots, |\mathcal{I}_s|\}. \tag{5.20}$$

Note, the intersection of $\mathcal{S}_t \setminus \mathcal{I}_s^i$ with $\Im_s$ is necessary in case that $\mathcal{S}_t$ is not completely contained in the set of interfering BSs $\Im_s$. Depending on the definition of $\Im_s$ this case is not unlikely; see the paragraph on the definition of the interfering set at the end of this section. The spectral efficiency e_{st}^i is then taken from Table 5.1 based on the SINR $\gamma_{t,s}^i$.

We now extend the basic formulation to incorporate the presented interference mitigation.

$$\min \sum_{s \in \mathcal{S}} \left(c_s \sum_{i=1}^{|\mathcal{I}_s|} x_s^i \right) + \lambda \sum_{t \in \mathcal{T}} u_t \tag{5.21a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{5.21b}$$

$$\sum_{i=1}^{|\mathcal{I}_s|} x_s^i + \sum_{\substack{j \in \{1,\dots,|\mathcal{I}_\sigma|\}\,:\\ s \in \mathcal{I}_\sigma^j}} x_\sigma^j \leq 1 \qquad \forall\, s \in \mathcal{S},\, \sigma \in \mathcal{H}_s \tag{5.21c}$$

$$\sum_{t \in \mathcal{T}_s : e_{st}^i \geq e_{\min}} \frac{w_t}{e_{st}^i} z_{st} \leq b_s x_s^i + M_s^i (1 - x_s^i) \qquad \forall\, s \in \mathcal{S},\, i \in \{1,\dots,|\mathcal{I}_s|\} \tag{5.21d}$$

$$z_{st} \leq \sum_{i=1}^{|\mathcal{I}_s|} x_s^i \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.21e}$$

$$z_{st} + \sum_{i \in \{1,\dots,|\mathcal{I}_s|\} : e_{st}^i < e_{\min}} x_s^i \leq 1 \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.21f}$$

$$\sum_{i=1}^{|\mathcal{I}_s|} x_s^i \leq 1 \qquad \forall\, s \in \mathcal{S} \tag{5.21g}$$

$$x_s^i,\, z_{st},\, u_t \in \{0,1\} \qquad \begin{array}{l} \forall\, s \in \mathcal{S},\, i \in \{1,\dots,|\mathcal{I}_s|\},\\ (s,t) \in \mathcal{S} * \mathcal{T},\, t \in \mathcal{T} \end{array} \tag{5.21h}$$

The objective function (5.21a) exploits (5.19) and is thus exactly the same as the standard objective function (4.4a), which minimises the cost of installed BSs and the number of not covered TNs. Furthermore, constraints (5.21b) are the same as (4.4b) ensuring that a TN is either covered by one BS or not served at all and (5.21c) are a copy of (5.18). The capacity constraints (4.4c) are reformulated in terms of the new variables x_s^i and corresponding parameters e_{st}^i in (5.21d). We introduce a big-M for every s and i to guarantee that the capacity constraint is non-restrictive in case that $x_s^i = 0$. The smallest possible value for M_s^i is $\sum_{t \in \mathcal{T}_s : e_{st}^i \geq e_{\min}} \frac{w_t}{e_{st}^i}$.

The vub constraints (4.5) are reformulated in (5.21e). Due to the recalculation of the spectral efficiencies based on the SINR values defined in (5.20), $e_{st}^i < e_{\min}$ is possible for a pair $(s,t) \in \mathcal{S} * \mathcal{T}$. To exclude an invalid assignment of a TN t to a BS s if all BSs in $\mathcal{I}_s^i$ are not deployed with $e_{st}^i < e_{\min}$, we add constraints (5.21f). Note, these constraints implicitly include the vub constraints (5.21e). Finally, (5.21g) guarantee that at most one set of interfering BSs $\mathcal{I}_s^i$ is not deployed if BS s is installed.

Remark 5.7. The *interference mitigation formulation* (5.21) is exact in terms of SINR as well as capacity if the sets $\mathcal{I}_s$ comprise every possible subset of interfering BSs for every $s \in \mathcal{S}$.

Proof. If $\mathcal{I}_s$ comprises every possible subset of interfering BSs for every $s \in \mathcal{S}$, then a feasible solution of (5.21) respects the minimum spectral efficiency requirement by means

of constraints (5.21f) and thus, also the SINR requirement. A prerequisite is that $e_{\min}$ and δ are defined accordingly.

Furthermore, the capacity constraints (5.21d) contain the correct spectral efficiencies derived from the actual SINR values and can therefore not be violated by any feasible solution. □

The definition of the interfering sets One degree of freedom in the presented interference mitigation approach is the definition of the sets $\mathcal{I}_s$. To preserve the exactness, $\mathcal{I}_s$ has to be defined as the power set of $\mathcal{H}_s := \mathcal{S} \setminus \{s\}$ for $s \in \mathcal{S}$. But then we have to solve (5.21) with an exponential number of variables and constraints. To avoid this, we propose two alternative definitions of $\mathcal{I}_s$ in the following, which however cannot guarantee the exactness in terms of SINR and thus, also not in terms of capacity.

A definition based on the minimum distance constraint described in Section 5.3 is the following. If the distance between BSs s and σ is less than $d_{\min}$, then $s \in \mathcal{H}_\sigma$ and $\sigma \in \mathcal{H}_s$. The set of subsets of interferers is now defined as $\mathcal{I}_s = 2^{\mathcal{H}_s}$. By this means, we assume the same level of interference between two BSs in both directions.

The interference mitigation approach with the set of interferers specified above is comparable to the conflict graph model since both formulations use the minimum distance requirement. However, the interference mitigation is more capable of modelling inter-cell interference since the SINR values $\gamma^i_{t,s}$ are implicitly taken into account via the spectral efficiencies e^i_{st}.

Another possibility to determine the sets $\mathcal{I}_s$ of interferers is based on the TN coverage requirement introduced in Section 5.4. For that purpose, we define a new parameter $\varrho_{s\sigma}$ for every BS pair (s, σ) with $s \in \mathcal{S}$ and $\sigma \in \mathcal{S} \setminus \{s\}$. $\varrho_{s\sigma}$ gives the percentage of TNs $t \in \mathcal{T}_s \cap \mathcal{T}_\sigma$ for which $e_{st}/e_{\sigma t} < \delta$. If $\varrho_{s\sigma}$ exceeds a predefined threshold, then σ is an interferer of s and hence, $\sigma \in \mathcal{I}^i_s$ for at least one $i \in \{1, \ldots, |\mathcal{I}_s|\}$ and $s \in \mathcal{H}_\sigma$. Note that BS s is not necessarily an interferer to BS σ as the interference is directed.

For a restriction of the size of sets $\mathcal{I}_s$, we define a parameter $\varkappa \in \mathbb{Z}_{>0}$ as a positive integer and assume that only the $\varkappa$ many closest BSs to a BS s can interfere signals from s. Hence, the number of interferers that have to be considered in the two definitions stated above is limited by $\varkappa$.

We investigate the two presented approaches to define the sets $\mathcal{I}_s$ and aim at finding a good value for $\varkappa$ in a computational study in Section 5.9.3.

5.7. A TN Oriented Formulation

The interference mitigation approach proposed in the previous section can be regarded as a BS oriented formulation since interference is defined from a BS point of view. In this section, we develop a TN oriented model where interference is defined on a TN basis.

For every TN $t \in \mathcal{T}$, we define K_t different configurations where a configuration $k \in \{1, \ldots, K_t\}$ comprehends a set $\mathcal{S}^k \subseteq \mathcal{S}_t$ of deployed BSs and an accentuated BS $s^k \in \mathcal{S}^k$ which is the serving BS for t. Furthermore, we define $k = 0$ such that TN t is not served by

any BS. Note, for configuration $k = 0$, we do not make any assumption on the deployment of BSs.

For each TN t and configuration k, we can compute the SINR value as follows.

$$\gamma_t^k := \frac{P_r(s^k, t)}{\sum_{s \in \mathcal{S}^k \setminus \{s^k\}} P_r(s, t) + \eta} \qquad \forall\, t \in \mathcal{T},\ k = 1, \dots, K_t \tag{5.22}$$

Based on these SINR values, we define the associated spectral efficiencies e_t^k according to the look-up Table 5.1. Note that we consider only configurations with $e_t^k \geq e_{\min}$ henceforth.

Finally, we introduce indicator variables y_t^k which are set to 1 if configuration k is chosen for TN t and 0 otherwise. In case $y_t^k = 1$ for $k \geq 1$, the set $\mathcal{S}^k$ of BSs is deployed, all BSs in $\mathcal{S}_t \setminus \mathcal{S}^k$ are not installed and s^k is the serving BS for t. If $y_t^0 = 1$, TN t is not served and we do not assume anything on the deployment of BSs in $\mathcal{S}_t$.

We propose the TN oriented formulation as the following ILP.

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \lambda \sum_{t \in \mathcal{T}} y_t^0 \tag{5.23a}$$

$$\text{s.t.} \sum_{k=0}^{K_t} y_t^k = 1 \qquad \forall\, t \in \mathcal{T} \tag{5.23b}$$

$$\sum_{k \in \{1,\dots,K_t\}: s \in \mathcal{S}^k} y_t^k \leq x_s \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.23c}$$

$$\sum_{k \in \{1,\dots,K_t\}: s \in \mathcal{S}^k} y_t^k \geq x_s - y_t^0 \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.23d}$$

$$\sum_{t \in \mathcal{T}_s} \sum_{k \in \{1,\dots,K_t\}: s = s^k} \frac{w_t}{e_t^k} y_t^k \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{5.23e}$$

$$x_s,\ y_t^k \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\ t \in \mathcal{T}, k = 0, \dots, K_t \tag{5.23f}$$

The objective (5.23a) is the reformulated objective function (4.4a) in terms of the new variables y_t^k. Constraints (5.23b) ensure that exactly one configuration is chosen for every TN while (5.23c) and (5.23d) connect variables y_t^k to the BS decision variables x_s. Note, the latter constraints are stronger than constraints comparable to the vub constraints (4.5). If BS s is the serving BS in a selected configuration for one TN, this BS has to be installed. Finally, the equivalent of the capacity constraints (4.4c) are constraints (5.23e).

Remark 5.8. If we consider, for every TN $t \in \mathcal{T}$, all subsets of $\mathcal{S}_t$ and every possible configuration where each BS $s \in \mathcal{S}_t$ is a potential serving BS, the *TN oriented formulation* (5.23) is exact in terms of SINR as well as capacity.

Proof. Similar to the interference mitigation formulation (5.21), a violation of the minimum spectral efficiency requirement (SINR requirement) by a feasible solution of (5.23) is impossible due to the definition of feasible configurations. Again, $e_{\min}$ and δ have to be defined accordingly.

Moreover, also the capacity constraints (5.23e) incorporate the correct spectral efficiencies and hence, are not violated by any feasible solution. □

Similar to the interference mitigation model in Section 5.6, the size of the sets $\mathcal{S}_t$ has great impact on the number of configurations K_t and hence, on the size of the ILP (5.23). To limit the number of variables and constraints, we define a parameter $\varkappa \in \mathbb{Z}_{>0}$ as a positive integer and assume that only $\varkappa$ many BSs with the strongest signal are included in the sets $\mathcal{S}^k$ for each configuration k per TN t. The strongest signal is defined as the signal with the highest fading coefficient. Note, if we limit the number of configurations by $\varkappa$, we lose the exactness stated in Remark 5.8.

By the parameter $\varkappa$, at most $2^\varkappa$ many different sets have to be considered in the configurations corresponding to one TN. However, also the definition of the serving BS s^k impacts the number of configurations K_t. We propose the following two possibilities. First for a given subset $\mathcal{S}^k \subseteq \mathcal{S}_t$, we select $s^k \in \mathcal{S}^k$ to be the BS with the strongest signal to TN t. This is a quite intuitive method to define a serving BS and is also performed in practice for homogeneous networks; see, e. g. Madan et al. [130], Siomina et al. [166]. For more flexibility, we consider every BS $s \in \mathcal{S}^k$ as the potential serving BS and add configurations accordingly in our second approach. In this approach, we have $K_t = \sum_{i=1}^{\varkappa} \binom{\varkappa}{i} \cdot i$ when using the restricting parameter $\varkappa$.

We investigate the two proposed possibilities of defining the serving BS as well as reasonable values for $\varkappa$ in a computational study in Section 5.9.3.

5.8. Exploiting Discrete Channel Quality Indicators

In this section, we exploit the determination of the spectral efficiencies using discrete values associated to a range of SINR values; see Table 5.1. In total, we have to consider 16 different values for spectral efficiencies each labelled by the corresponding CQI k . We denote the highest possible CQI for a link from s to t, which is associated to the SNR value (no interference), by κ_{st} and the value of the spectral efficiency for any CQI k by e^k.

To incorporate these discrete CQIs in the basic formulation, we add new binary variables z_{st}^k for every $(s,t) \in \mathcal{S} * \mathcal{T}$ and every CQI $k \in \{1, \dots, \kappa_{st}\}$. It holds $z_{st}^k = 1$ if the signal from s to t has the quality k, thus spectral efficiency e^k. If TN t is assigned to BS s, then this link has exactly one specified CQI. Hence, we add the following constraint.

$$\sum_{k=1}^{\kappa_{st}} z_{st}^k = z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.24}$$

The quality of a link is impaired by interfering BSs. If a subset $C \subseteq \mathcal{S}_t \setminus \{s\}$ of BSs is deployed, the SINR of the signal from s to t is calculated as

$$\gamma_C = \frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{\sigma \in C} P_{\mathrm{r}}(\sigma,t) + \eta}. \tag{5.25}$$

We denote the CQI corresponding to γ_C by $\kappa_{st}(C)$. In the case that all BSs in C are deployed, better spectral efficiencies than for the CQI $\kappa_{st}(C)$ cannot occur for this link from s to t. Hence, $\sum_{k=\kappa_{st}(C)+1}^{\kappa_{st}} z_{st}^k = 0$. We can formulate this condition as the following model constraint.

$$\sum_{k=\kappa_{st}(C)+1}^{\kappa_{st}} z_{st}^k \leq |C| - \sum_{\sigma \in C} x_\sigma \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T}, C \subseteq \mathcal{S}_t \setminus \{s\}. \tag{5.26}$$

These inequalities have a similar structure as cover inequalities for the KP; see (1.14). Hence, we use the ambiguous term *cover* to denote C in the present context.

The complete model which exploits the discrete CQIs then reads

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \lambda \sum_{t \in \mathcal{T}} u_t \tag{5.27a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{5.27b}$$

$$\sum_{k=1}^{\kappa_{st}} z_{st}^k = z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.27c}$$

$$\sum_{t \in \mathcal{T}_s} \sum_{k=1}^{\kappa_{st}} \frac{w_t}{e^k} z_{st}^k \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{5.27d}$$

$$\sum_{k=\kappa_{st}(C)+1}^{\kappa_{st}} z_{st}^k \leq |C| - \sum_{\sigma \in C} x_\sigma \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T}, C \subseteq \mathcal{S}_t \setminus \{s\} \tag{5.27e}$$

$$z_{st} \leq x_s \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{5.27f}$$

$$x_s, z_{st}, z_{st}^k, u_t \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T},\ k = 1, \ldots, \kappa_{st},\ t \in \mathcal{T}. \tag{5.27g}$$

This ILP resembles the basic formulation with adapted capacity constraint (4.4c) in terms of the new variables and corresponding spectral efficiencies; see constraints (5.27d). Furthermore, we have added constraints (5.24) and (5.26) as explained before.

Remark 5.9. The *discrete CQIs formulation* (5.27) denotes an exact formulation in terms of SINR and capacity.

Proof. The argumentation is the same as in the proofs of Remarks 5.7 and 5.8. □

The ILP (5.27) is computationally challenging since there can exist exponentially many constraints (5.27e) due to the existence of exponentially many covers C. This is why, we do not add the entire set of these constraints but separate violated inequalities on the fly. In the following, we present two different separation routines.

5.8.1. ILP-based Separation

For any solution of (5.27), either LP or integer solution, denoted by $(\tilde{x}, \tilde{z}, \tilde{z}^k, \tilde{u})$, we determine the most violated and hopefully, the most restrictive inequalities via an exact ILP

based separation. The ILP presented in this section is specific for the case in which the fading coefficients a_{st} are determined only by the path loss.

First, we define the set of partially deployed BSs as $\Sigma := \{s \in \mathcal{S} \mid \tilde{x}_s > 0\}$. Based on Σ, we define $\Sigma_t := \mathcal{S}_t \cap \Sigma$ for all $t \in \mathcal{T}$ with $\tilde{u}_t < 1$. To propose an ILP for the separation of valid inequalities for a fixed pair (s,t), we introduce binary variables y_σ for every $\sigma \in \Sigma_t \setminus \{s\}$ deciding which BS is included in the new cover C and α_k for every $k \in \{0,1,\dots,\kappa_{st}\}$ to determine $\kappa_{st}(C)$ indirectly. Hence,

$$y_\sigma = \begin{cases} 1, & \text{if } \sigma \in C \\ 0, & \text{otherwise,} \end{cases} \qquad \alpha_k = \begin{cases} 1, & \text{if } k \geq \kappa_{st}(C)+1 \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, let γ_k denote the S(I)NR value for CQI k in Table 5.1,

$$\delta_k := \begin{cases} 10^{-\gamma_k/10} & \text{if } k = 1,\dots,15, \\ \mathrm{M} & \text{if } k = 0, \\ 0 & \text{if } k = 16, \end{cases}$$

with M sufficiently large, and let $\eta' := \eta/(p_s a_{st})$.

Given a solution $(\tilde{x}, \tilde{z}, \tilde{z}^k, \tilde{u})$, we propose the following ILP to separate a violated inequality (5.26) for a pair (s,t) with $\tilde{z}_{st} > 0$ or to prove that no violated inequality exists.

$$\max \sum_{\sigma \in \Sigma_t \setminus \{s\}} (\tilde{x}_\sigma - 1) y_\sigma + \sum_{k=1}^{\kappa_{st}} \tilde{z}^k_{st} \alpha_k \tag{5.28a}$$

$$\text{s.t. } \alpha_k \leq \alpha_{k+1} \qquad \forall\, k = 1,\dots,\kappa_{st} \tag{5.28b}$$

$$\sum_{k=1}^{\kappa_{st}+1} (\delta_{k-1} - \delta_k)\alpha_k - \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} y_\sigma \geq \eta' \tag{5.28c}$$

$$\sum_{k=1}^{\kappa_{st}+1} (\delta_k - \delta_{k+1})\alpha_k - \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} y_\sigma \leq \eta' - \varepsilon \tag{5.28d}$$

$$y_\sigma,\ \alpha_k \in \{0,1\} \qquad \begin{array}{l} \forall\, \sigma \in \Sigma_t \setminus \{s\}, \\ \forall\, k = 1,\dots,\kappa_{st}, \end{array} \tag{5.28e}$$

with $\varepsilon > 0$ small. Constraints (5.28b) guarantee that all α variables for the subsequent indices $k+1,\dots,\kappa_{st}$ are also set to 1 as soon as $\alpha_k = 1$. The objective (5.28a) maximises the violation of a potentially violated inequality (5.26). If the optimal objective value is ≤ 0, no violated inequality exists. Otherwise, for an optimal solution $(\tilde{y}, \tilde{\alpha})$ of (5.28), the cover C and the corresponding maximal CQI $\kappa_{st}(C)$ are defined by

$$C := \{\sigma \in \Sigma_t \setminus \{s\} \mid \tilde{y}_\sigma = 1\}, \quad \kappa_{st}(C) := \operatorname*{argmin}_{k=1,\dots,\kappa_{st}} \{\tilde{\alpha}_k = 1\} - 1. \tag{5.29}$$

Lemma 5.10. *Constraints* (5.28c) *and* (5.28d) *determine the CQI* $\kappa_{st}(C)$ *if the cover* C *is determined as described in* (5.29).

Proof. The value $\kappa_{st}(C)$ is taken from the look-up Table 5.1 based on the SINR value γ_C of the signal from BS s to TN t when all BSs in the cover C are deployed. However, the SINR values in Table 5.1 are given in dB. The SINR value γ_C converted to dB is calculated as

$$\gamma_C^{\mathrm{dB}} := 10 \cdot \log_{10}(\gamma_C) = 10 \cdot \log_{10}\left(\frac{p_s a_{st}}{\sum\limits_{\sigma \in \Sigma_t \setminus \{s\}} p_\sigma a_{\sigma t} y_\sigma + \eta}\right).$$

Hence, it holds

$$\gamma_{\kappa_{st}(C)} \leq \gamma_C^{\mathrm{dB}} < \gamma_{\kappa_{st}(C)+1}.$$

To determine the correct lower and upper bounds, we sum over $\alpha_{k+1} - \alpha_k$ as $\alpha_{k+1} - \alpha_k = 1$ if and only if $\alpha_{k+1} = 1$ and $\alpha_k = 0$. Additionally, we define $\alpha_0 := 0$ and γ_0 sufficiently small (negative). Thus,

$$\sum_{k=0}^{\kappa_{st}} \gamma_k(\alpha_{k+1} - \alpha_k) \leq 10 \cdot \log_{10}\left(\frac{p_s a_{st}}{\sum\limits_{\sigma \in \Sigma_t \setminus \{s\}} p_\sigma a_{\sigma t} y_\sigma + \eta}\right) < \sum_{k=0}^{\kappa_{st}} \gamma_{k+1}(\alpha_{k+1} - \alpha_k). \tag{5.30}$$

We linearise these inequalities and reformulate them by exploiting the fact that all coefficients are positive as follows.

$$\begin{aligned}
(5.30) \Leftrightarrow & 10^{\sum_{k=0}^{\kappa_{st}} \frac{\gamma_k}{10}(\alpha_{k+1}-\alpha_k)} \leq \frac{p_s a_{st}}{\sum\limits_{\sigma \in \Sigma_t \setminus \{s\}} p_\sigma a_{\sigma t} y_\sigma + \eta} < 10^{\sum_{k=0}^{\kappa_{st}} \frac{\gamma_{k+1}}{10}(\alpha_{k+1}-\alpha_k)} \\
\Leftrightarrow & 10^{-\sum_{k=0}^{\kappa_{st}} \frac{\gamma_k}{10}(\alpha_{k+1}-\alpha_k)} \geq \frac{\sum\limits_{\sigma \in \Sigma_t \setminus \{s\}} p_\sigma a_{\sigma t} y_\sigma + \eta}{p_s a_{st}} > 10^{-\sum_{k=0}^{\kappa_{st}} \frac{\gamma_{k+1}}{10}(\alpha_{k+1}-\alpha_k)} \\
\Leftrightarrow & \prod_{k=0}^{\kappa_{st}} 10^{\frac{-\gamma_k}{10}(\alpha_{k+1}-\alpha_k)} \geq \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} y_\sigma + \eta' > \prod_{k=0}^{\kappa_{st}} 10^{\frac{-\gamma_{k+1}}{10}(\alpha_{k+1}-\alpha_k)} \\
\Leftrightarrow & \sum_{k=0}^{\kappa_{st}} 10^{\frac{-\gamma_k}{10}}(\alpha_{k+1} - \alpha_k) \geq \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} y_\sigma + \eta' > \sum_{k=0}^{\kappa_{st}} 10^{\frac{-\gamma_{k+1}}{10}}(\alpha_{k+1} - \alpha_k)
\end{aligned}$$

The last equivalence holds as there exists exactly one value $\alpha_{k+1} - \alpha_k = 1$ and all others are equal to 0. Hence, exactly one factor in the product is $\neq 1$.

By the presented derivation, it becomes clear that the values δ_k are the negative of the SINR values γ_k for $k = 1, \ldots, 15$, which are given in dB, converted to W. Using the values δ_k and exploiting $\alpha_0 = 0$, we can reformulate the last inequalities so that each variable α_k occurs only once:

$$\sum_{k=1}^{\kappa_{st}} (\delta_{k-1} - \delta_k)\, \alpha_k + \delta_{\kappa_{st}} \alpha_{\kappa_{st}+1} \geq \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} y_\sigma + \eta' > \sum_{k=1}^{\kappa_{st}} (\delta_k - \delta_{k+1})\, \alpha_k + \delta_{\kappa_{st}+1} \alpha_{\kappa_{st}+1}. \tag{5.31}$$

Note, by setting $\delta_{16} = 0$, the second inequality becomes non-restrictive in case $\kappa_{st} = 15$

and $\alpha_{\kappa_{st}+1} = 1$. Furthermore, M for the definition of δ_0 has to be sufficiently large so that the first inequality is non-restrictive in case $\alpha_1 = 1$, e. g., $M \coloneqq \delta_1 + \sum_{\sigma \in \Sigma_t \setminus \{s\}} \frac{p_\sigma a_{\sigma t}}{p_s a_{st}} + \eta'$.

The inequalities (5.31) are exactly (5.28c) and (5.28d) using the definition of δ_k and η'. □

Improvements To accelerate the solving of the separation ILP (5.28), we apply the following preprocessing. For a BS $\sigma \in \Sigma_t \setminus \{s\}$ with $\tilde{x}_\sigma = 1$, we assume that this BS is definitely contained in the new cover C as it contributes a 0 to the objective function (5.28a). Hence, we consider only the set $\Sigma_t' \coloneqq \Sigma_t \setminus (\{s \in \mathcal{S}_t \,|\, \tilde{x}_s = 1\} \cup \{s\})$ in the ILP (5.28) and define $C \coloneqq \{\sigma \in \mathcal{S}_t \,|\, \tilde{x}_s = 1\} \setminus \{s\} \cup \{\sigma \in \Sigma_t' \,|\, \tilde{y}_\sigma = 1\}$. Thereby, we can reduce the size of the ILP (5.28).

Furthermore, we simplify the test of feasibility of an integer solution. Per (s,t)-pair, we define $C \coloneqq \{\sigma \in \mathcal{S}_t \,|\, \tilde{x}_\sigma = 1\}$, calculate the corresponding value $\kappa_{st}(C)$ based on the SINR γ_C as given in (5.25), and test if the associated model constraint (5.26) is violated. Therefore, we do not have to solve the ILP (5.28) for an integer feasibility test.

Finally, we test the performance of the separation routine based on ILP (5.27) when restricting the total number of separated violated inequalities per separation round to a positive integer $\varkappa \geq 1$ in the computational study presented in Section 5.9.3. More precisely, we stop one separation round at a pair (s', t') as soon as $\varkappa$ many violated inequalities have been found. For the next solution, we start the separation routine at the pair (s', t'). After the last pair, we start again with the first possible BS-TN pair. Such a process is also called a round robin scheduling algorithm.

5.8.2. Combinatorial Separation

Despite the improvements described in the previous section, the solving of the ILP (5.28) for several BS-TN pairs with $\tilde{z}_{st} > 0$ can be quite time consuming. Thus, we describe a combinatorial separation algorithm, which determines all possible covers C and maximal CQIs $\kappa_{st}(C)$ for each pair $(s,t) \in \mathcal{S} * \mathcal{T}$ with $\tilde{z}_{st} > 0$, in this section. If the corresponding inequality (5.26) is violated, we add it to the formulation.

To determine possible covers, we use the definitions of Σ and Σ_t as given at the beginning of Section 5.8.1. We can then compute every subset $C \subseteq \Sigma_t$ and the corresponding value $\kappa_{st}(C)$ and calculate the violation of inequality (5.26). This routine is summarised in Algorithm 3.

If this algorithm does not return a cover, then there does not exist a violated inequality (5.26) for the current solution since all possible covers have been tested. Hence, also the combinatorial separation is exact.

To reduce the number of added inequalities and find the more restrictive ones, we introduce the following concept of dominance. But first recall the definition of dominating valid inequalities.

Definition 5.11. If $\pi x \leq \pi_0$ and $\pi' x \leq \pi_0'$ are two valid inequalities for a polyhedron $P \subseteq \mathbb{R}_{\geq 0}$, $\pi x \leq \pi_0$ *dominates* $\pi' x \leq \pi_0'$ if there exists $\lambda \geq 0$ such that $\pi \geq \lambda \pi'$, $\pi_0 \leq \lambda \pi_0'$, and $(\pi, \pi_0) \neq (\pi', \pi_0')$.

Algorithm 3 Combinatorial separation of inequalities of type (5.26)

Input: solution $(\tilde{x}, \tilde{z}, \tilde{z}^k, \tilde{u})$

Output: pair $(C, \kappa_{st}(C))$ with violated corresponding inequality or proof that none exists

define $\Sigma := \{s \in \mathcal{S} \mid \tilde{x}_s > 0\}$ and $\Sigma_t := \mathcal{S}_t \cap \Sigma$ for all $t \in \mathcal{T}$ with $\tilde{u}_t < 1$

for $(s,t) \in \mathcal{S} * \mathcal{T}$ with $\tilde{z}_{st} > 0$ **do**

for $C \subseteq \Sigma_t$ with $s \notin C$ **do**

compute SINR value $\gamma_C := \frac{P_r(s,t)}{\sum_{\sigma \in C} P_r(\sigma,t) + \eta}$

set value for $\kappa_{st}(C)$ based on γ_C and Table 5.1

end for

if $\sum_{k=\kappa_{st}(C)+1}^{\kappa_{st}} \tilde{z}^k_{st} > |C| - \sum_{\sigma \in C} \tilde{x}_\sigma$ **then return** $(C, \kappa_{st}(C))$

end if

end for

Definition 5.12. (*Dominance*) A cover C_1 dominates a cover C_2 if the corresponding "cover" constraint

$$\sum_{k=\kappa_{st}(C_1)+1}^{\kappa_{st}} z^k_{st} \leq |C_1| - \sum_{\sigma \in C_1} x_\sigma \tag{5.32}$$

dominates the respective constraint

$$\sum_{k=\kappa_{st}(C_2)+1}^{\kappa_{st}} z^k_{st} \leq |C_2| - \sum_{\sigma \in C_2} x_\sigma. \tag{5.33}$$

Lemma 5.13. *A cover C_1 dominates a cover C_2 if*

i) $C_1 \subset C_2$, *and*

ii) $\kappa_{st}(C_1) = \kappa_{st}(C_2)$.

Proof. The left hand sides of the associated inequalities (5.32) and (5.33) are equal. To prove that any (partial) solution $(\tilde{x}, \tilde{z})$ satisfying (5.32) also satisfies (5.33) it remains to show

$$|C_1| - \sum_{\sigma \in C_1} x_\sigma \leq |C_2| - \sum_{\sigma \in C_2} x_\sigma.$$

As $C_1 \subset C_2$, we have $|C_2| - |C_1| = |C_2 \setminus C_1|$ and naturally, $|C_2 \setminus C_1| \geq \sum_{\sigma \in C_2 \setminus C_1} x_\sigma$. Hence,

$$|C_1| - \sum_{\sigma \in C_1} x_\sigma \leq |C_1| - \sum_{\sigma \in C_1} x_\sigma + |C_2 \setminus C_1| - \sum_{\sigma \in C_2 \setminus C_1} x_\sigma = |C_2| - \sum_{\sigma \in C_2} x_\sigma.$$

□

Applying Lemma 5.13, we compute all possible covers C by Algorithm 3 but add only violated inequalities for covers which are not dominated by any other. Moreover, we use

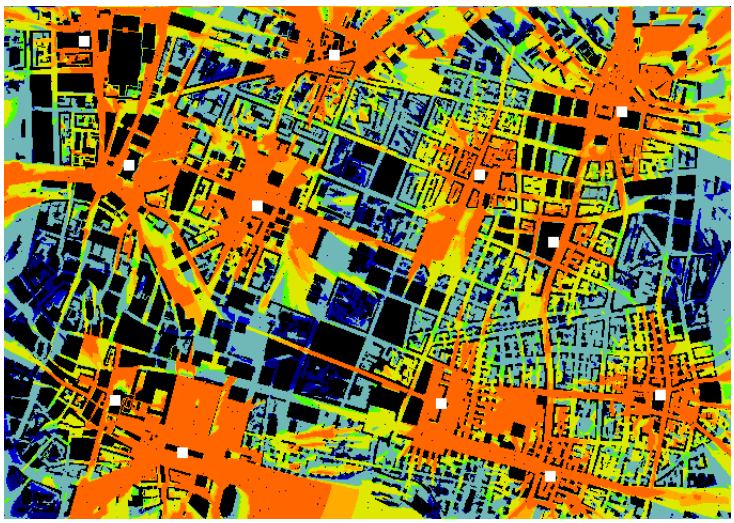


Figure 5.3.: An example of a signal propagation scenario for Munich with dimension 2.5 × 3.5 km. The black rectangles denote houses, the white squares BSs and the colours denote the signal strength. The warmer the colour, the better the signal strength.

the same integer feasibility test as described at the end of Section 5.8.1 instead of Algorithm 3.

5.9. Numerical Comparison of Different Formulations

5.9.1. Creation of Test Instances

To evaluate the performance of the models presented in the previous sections and any modifications or generalisations of the basic formulation, we use different test scenarios which depict a simplification of the real world. The general procedure to create such a test instance and the involved parameters are described in this section.

All planning scenarios are based on signal propagation data for Munich, available at [55]. This data set comprises 60 BS candidate sites from which we randomly choose BSs if we do not intend to regard the total number of available BSs. Unless otherwise stated, the BS candidates are limited to the location of the BS. Furthermore, TNs are distributed randomly in the available area. To compute the path loss from a BS to a TN, we use a cube oriented ray launching algorithm [137] for signal prediction; see also Section 3.1.2. in Engels [65]. An example of the signal propagation for 11 BSs is displayed in Figure 5.3. The black rectangles are houses, which designate also the streets, and the white squares are BSs. The colours denote the signal strength: The warmer the colour, the better the signal, i. e., the lower the path loss. Based on the path loss, we compute the SNR value (3.1) for every $(s,t) \in \mathcal{S} * \mathcal{T}$ and extract the corresponding spectral efficiency e_{st} from Table 5.1. Additionally, we assume a transmission power p_s of 46 dBm for every BS $s \in \mathcal{S}$. This

service	usage (%)	bit rate (kbps)
data	[10,20]	[512,2000]
web	[20,40]	[128,512]
VoIP	remaining	64

Table 5.3.: Traffic profiles for TNs

is a typical value for a macro cell in LTE used together with a bandwidth $b_s = 10\,\text{MHz}$; cf. Chapter 19 in Kahn [102]. Moreover, the noise value in dB is computed as thermal noise (dB) + noise figure for users (dB), where the thermal noise in W is computed as $\eta_t(\text{W}) := k_B \cdot T \cdot B$. Here, k_B is the Boltzmann constant, T the temperature in Kelvin and B is the bandwidth given in Hz. In total, the noise η in dB is computed as

$$\eta = 10 \log_{10}\left(1.3806503 \cdot 10^{-23} \cdot 290.0\,\text{K} \cdot 10000000\,\text{Hz}\right) + 9\,\text{dB},$$

where 9 dB represents the noise figure for users; cf. [67]. Note, for the computation of S(I)NR values, this value has to be converted to W.

Since data of mobile users is not available due to data privacy limitations, we compute the demand values w_t for each $t \in \mathcal{T}$ by randomly generating user profiles from Table 5.3; compare [65] and Chapter 19 in [102]. A percentage for both data (downloading including streaming) and web (browsing) services is uniformly drawn from the "usage (%)" column and multiplied by a bit rate uniformly drawn from the "bit rate (kbps)" column. The remaining percentage is used for Voice-over-IP (VoIP) telephony with a bit rate of 64 kbps. As an example, the minimum required bit rate of any TN is

$$10\% \cdot 512\,\text{kbps} + 20\% \cdot 128\,\text{kbps} + 70\% \cdot 64\,\text{kbps} = 121.6\,\text{kbps}.$$

A fixed parameter which is used in every test instance is the BS cost, $c_s = 4000$, where the value is taken from Deruyck et al. [63] and depicts the (rounded) total power consumed by a BS in W.

Moreover, reasonable values for the scaling parameter λ depend on the cost of a BS. We study values of 1000, 2000, and 4000 in our work Claßen et al. [49], where $\lambda = 1000$ (2000, 4000) denotes that four (two, one) TNs can be lost before it becomes beneficial to deploy an additional BS. Even though the scale of the results presented in [49] depend on the value of λ, the nature of the results is identical. Based on this observation, we choose the least conservative value 1000 for λ if not stated otherwise.

Any further necessary parameters to set up a complete test scenario are listed when needed.

5.9.2. Specifications for the study of interference modelling

To decrease the magnitude of the coefficients occurring in the models, we consider multiples of thousands. Hence, we set the BS cost c_s to 4 and the scaling parameter λ to 1.

name	# BSs	# TNs
20_200a	20	200
20_300b	20	300
20_400c	20	400
30_300a	30	300
30_400b	30	400
30_500c	30	500
40_400a	40	400
40_450a	40	450

Table 5.4.: Names and dimensions of test instances used in the computational study to investigate different approaches for interference modelling.

Furthermore, we round the demand values up to integer kbps and the bandwidth b_s is set to $10,000$ kHz. We define the SINR threshold as $\delta = -5.1$ dB, which is the lowest possible value to establish a link, and the corresponding minimum required spectral efficiency is $e_{\min} = 0.25$ bps/Hz.

In this computational study, we consider eight different test scenarios which are summarised in Table 5.4. For scenarios with the same number of BSs, the characters "a", "b", "c" indicate that the sets of BSs are different.

Evaluating interference To evaluate the quality of a solution in terms of interference, we first count the number of violated SINR requirements. The SINR values are computed as

$$\frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{\sigma \in I_{st}} P_{\mathrm{r}}(\sigma,t) + \eta}, \tag{5.34}$$

where s is the serving BS for TN t and I_{st} is the set of interfering and deployed BSs of the signal from s to t. We denote the number of SINR values below the threshold δ by U.

For a direct comparison of different solutions, we introduce the *SINR-corrected objective value* υ, which exploits the fact that c_s is the same for every BS for the studied scenarios. Given a solution x^* for the BS deployment, we define

$$\upsilon \coloneqq c_s \sum_{s \in \mathcal{S}} x_s^* + \lambda(|\mathcal{T}| - (\tau - U)) = c_s \sum_{s \in \mathcal{S}} x_s^* + \lambda(|\mathcal{T}| - \tau + U), \tag{5.35}$$

with τ denoting the number of TNs marked as covered in the (optimal) solution. Thus, υ denotes the objective value where TNs with violated SINR requirements are regarded as not covered.

The values of the spectral efficiencies used in the different formulations described in the previous sections are not always based on the correct SINR values, which can sometimes be computed only after the complete solution is known. Therefore, a violation of the capacity constraints (4.4c) is possible. To regard such violations in the evaluation of the

solutions, we compute the actual percentage load for all deployed BSs incorporating the correct spectral efficiencies e'_{st} based on the SINR values (5.34) as follows.

$$\ell_s := \frac{1}{b_s} \sum_{t \in \mathcal{T}'_s} \frac{w_t}{e'_{st}},$$

where $\mathcal{T}'_s$ denotes the set of TNs assigned to s with a SINR value computed via (5.34) exceeding δ. Note, $\ell_s > 1$ indicates a capacity violation. The SINR value that includes the percentage load of the interfering BSs is computed as

$$\frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{\sigma \in I_{st}} \ell_s P_{\mathrm{r}}(\sigma,t) + \eta}.$$

Since $\ell_s > 1$ is possible, violated SINR conditions can still occur and we denote the number of such SINR values below δ by U_{l}. The SINR-corrected objective value υ_{l} is computed accordingly; cf. (5.35).

All computations in the following sections are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of two hours if not stated differently. Additionally, we limit the available RAM to 11 GB and use CPLEX 12.4 [98] as (I)LP solver.

5.9.3. Evaluation of Different Settings

In this subsection, we investigate the different parameter settings per formulation to find the best setting for the comparison of all proposed formulations depicted in Section 5.9.4.

SINR constraint formulation For the SINR constraint formulation proposed in Section 5.2, we investigate the numerical stability via different values to scale the fading coefficients. Without a multiplication of the SINR constraints (5.4) with a scaling parameter $10^{\varkappa}$ where $\varkappa \in \mathbb{Z}_{\geq 0}$, the coefficients are too small for the precision of CPLEX. We study $\varkappa \in \{1, \ldots, 10\}$ and compare the solutions to the solutions obtained without scaling ($\varkappa = 0$). We display the values of the SINR-corrected objective value υ in Figure 5.4. Since the number of deployed BSs varies around one BS more or less for every scenario, variations in υ mainly occur due to fluctuating numbers of actually served TNs. By "actually", we designate the number of TNs with $\tilde{u}_t = 0$ in an optimal solution minus the number of violated SINR conditions, i. e., $|\mathcal{T}| - \tau + U$. The values of υ fluctuate for $\varkappa \leq 8$ since violated SINR conditions arise due to numerical difficulties. Only for a scaling exponent $\varkappa \geq 9$, no violated SINR condition exists for the optimal solutions. Hence, it is difficult to judge beforehand without conducting a computational study which scaling exponent has to be chosen to obtain a meaningful/correct solution. This demonstrates why we study alternative formulations to model interference. Nevertheless, for the comparison of the different formulations presented in Section 5.9.4, we set $\varkappa = 10$.

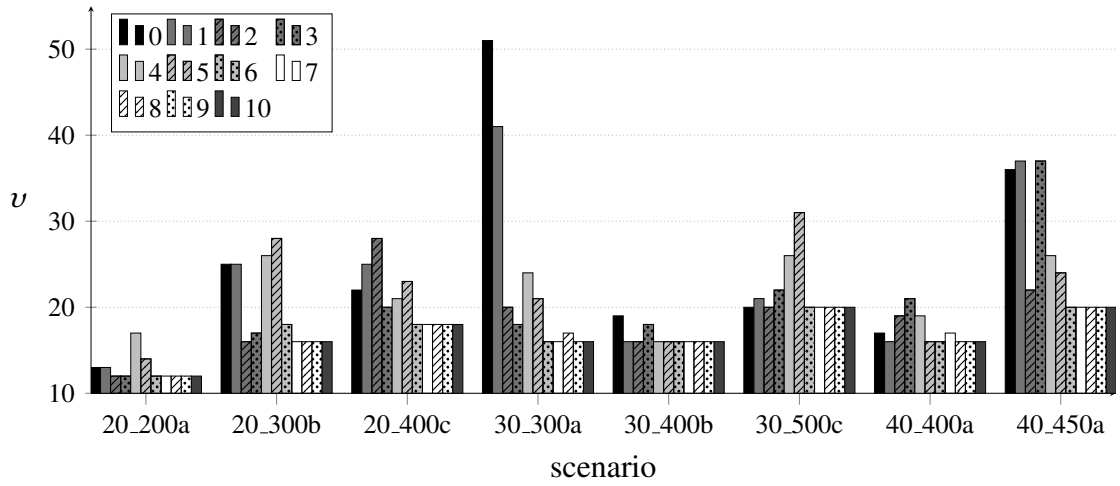


Figure 5.4.: SINR-corrected objective values υ for the SINR constraint formulation and different scaling exponents $\varkappa$.

TN coverage requirement formulation For the TN coverage requirement formulation discussed in Section 5.4, we evaluate different values of the threshold δ_c. Engels et al. [67] set $\delta_c = 0.6$ while we also investigate values greater than 0.6. Note, $\delta_c \leq 0.5$ is not reasonable if we intend to avoid violated SINR conditions. In total, we study $\delta_c \in \{0.6, 0.7, 0.8, 0.9, 1.0, 1.1\}$ and present the SINR-corrected objective values υ in Figure 5.5. The numbers of deployed BSs are rather the same for distinct thresholds. Thus, varying values for υ occur mainly due to different numbers of actually served TNs. The threshold $\delta_c = 1.0$ gives the lowest SINR-corrected objective values most frequently for the studied scenarios. Additionally, since the values of υ increase drastically for $\delta_c = 1.1$, we cannot expect to improve the formulation by setting $\delta_c > 1.1$. In summary, we set $\delta_c = 1.0$ for the comparison in Section 5.9.4.

Iterative formulation For the iterative formulation presented in Section 5.5, we investigate three parameters and their combination. These parameters comprise the gap limit for the ILP (5.14), the time limit for this ILP, and the SINR calculation including the load of a BS (5.17) or not (5.16). As the gap limit for the ILP, we evaluate values of 1, 2 and 5 % and for the time limit, we study 72 s (one hundredth of the global time limit) and 1800 s.

We analyse the three parameters in Figure 5.6, where Figure 5.6(a) displays the SINR-corrected objective values υ for different time and gap limit combinations in case the SINR is calculated via (5.16) without load and Figure 5.6(b) displays υ but with SINR calculation (5.17) including the load. For better readability, we introduce the label "TL-gap" to denote a combination of time and gap limit. We observe that all results are quite similar. However, to determine the best setting for the comparison in Section 5.9.4, we drop all settings which lead to higher values of υ for at least one scenario in case of no load in the SINR calculation. The remaining settings are 72-1 and 1800-1. The same evaluation in case of load inclusion is not possible as every setting gives a higher υ for at least one

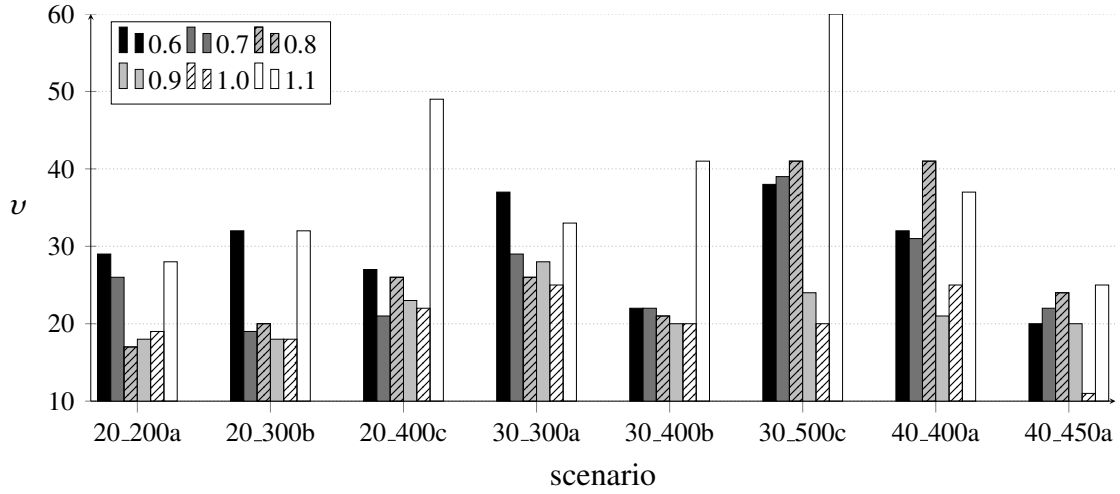


Figure 5.5.: SINR-corrected objective values υ for the TN coverage requirement formulation and different thresholds δ_c.

scenario	20_200a	20_300b	20_400c	30_300a	30_400b	30_500c	40_400a	40_450a
highest	199	300	400	300	400	500	400	450

Table 5.5.: Highest numbers of actually served TNs for the iterative formulation.

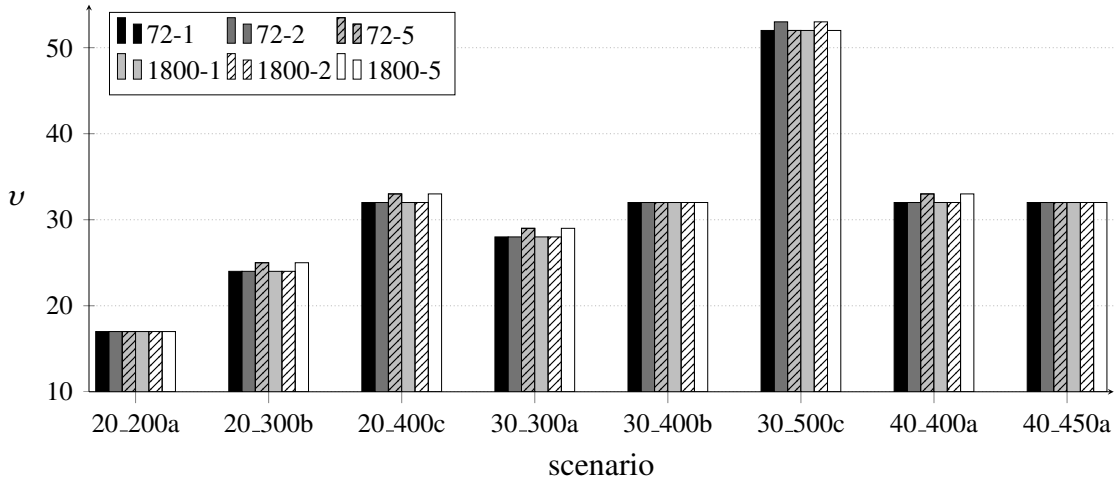
scenario.

Hence, we compare the numbers of actually served TNs for settings 72-1 and 1800-1 without load inclusion and all settings with load inclusion. In Table 5.5, we display the highest achieved values per scenario. These numbers of actually served TNs can be achieved only by the settings 72-1 and 1800-1 via SINR calculation (5.16) without load. For all settings with load inclusion, the optimal solution serves one TN less either for scenario 30_300a or 30_500c. Moreover, the settings 72-1 and 1800-1 with SINR calculation (5.16) perform identically under the studied aspects.

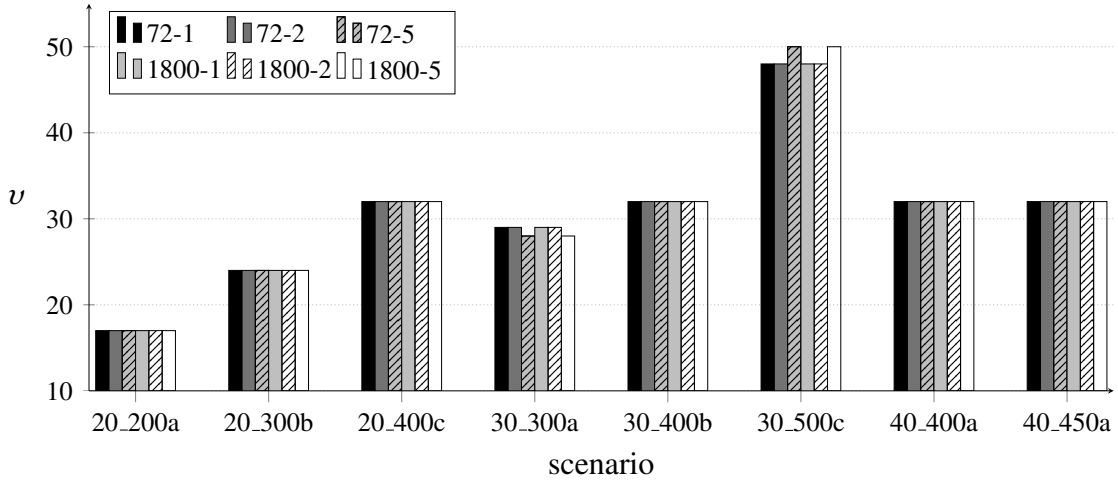
If we compare these two settings in more detail via time consumption and number of performed iterations, we cannot observe any significant difference. Hence, the time limit for the ILP is never reached and the gap limit of 1 % is the only restriction. For the global comparison in Section 5.9.4, we choose the setting 72-1 arbitrarily.

Interference mitigation formulation We cannot solve the exact approach for any scenario due to excessive memory use such that the process is killed by the operating system before the creation of the ILP (5.21) is completed. Thus, the memory limit of 11 GB could not be checked at this state and instead, the complete available memory of 32 GB of the machine is exceeded. Which is the reason why an increase of the memory limit is not an option.

In the following, we compare the minimum distance based definition of the interfering sets to the TN coverage requirement based definition with $\delta_c = 1$, where this value is based



(a) No load in SINR calculation (5.16)



(b) Load in SINR calculation included (5.17)

Figure 5.6.: SINR-corrected objective values υ for the iterative formulation and different values of time and gap limit.

	$\varkappa$			
threshold for $\varrho_{s\sigma}$	1	5	10	15
20 %	7.7	10.1	8.7	–
40 %	9.1	9.1	9.2	–
50 %	7.7	8.1	9.8	–
60 %	9.0	6.1	8.4	7.4
70 %	3.8	4.7	5.3	4.8
80 %	5.3	4.5	4.5	4.5
90 %	5.5	5.9	5.9	5.9
100 %	4.4	4.2	4.2	4.2
dist. based	7.0	12.2	9.5	9.5

Table 5.6.: Average percentage violation of SINR conditions for minimum distance based as well as TN coverage requirement with $\delta_c = 1$ based definition of the interfering sets in the interference mitigation formulation and $\varkappa \in \{1, 5, 10, 15\}$.

on the previously presented results for the TN coverage formulation defined in Section 5.4. For both types of defining the interfering sets, we investigate four values for the number of interferers, $\varkappa = 1, 5, 10$ and 15. Moreover, for the TN coverage based definition, we additionally study values of $20, 40, 50, 60, 70, 80, 90$ and $100\,\%$ for the percentage of TNs $\varrho_{s\sigma}$ violating the TN coverage requirement. To compare the two variants, we compute the percentage number of violated SINR conditions per setting and instance and average over the scenarios. These numbers are depicted in Table 5.6. Note, for 15 interferers and a threshold of 20, 40 or 50 % not every scenario could be solved by the TN coverage requirement based set definition, which is the reason why we do not state average values for these settings. For the remaining settings, the maximum average value is 10.1 % while the lowest value is 3.8 %. In contrast, for the minimum distance based definition, the average percentage SINR violation varies between 7.0 % and 12.2 %. This indicates that the former variant performs better than the letter and thus, we do not investigate the distance based set definition any further.

To find the best setting among the different settings for the TN coverage requirement based set definition, we study those ten settings which give an average percentage SINR violation value of less than 5 % by means of the SINR-corrected objective value υ. For better readability, we introduce the label "$\varkappa$-threshold" to denote a specific setting. We display the SINR-corrected objective values in Figure 5.7. From this figure, it is difficult to judge which setting is best as every setting gives the lowest value υ at least once but also (significantly) worse values for another scenario. Therefore, we count the number of scenarios for which each setting gives the lowest SINR-corrected objective value. For four out of the eight scenarios, the settings 1-70 and 5-70 give the best υ, while all other settings give the lowest numbers only for at most two scenarios. Thus, we rate these two settings as better than the others. As 1-70 also gives the minimum average value of the percentage SINR violation; cf. Table 5.6, we regard it as the best setting. Note,

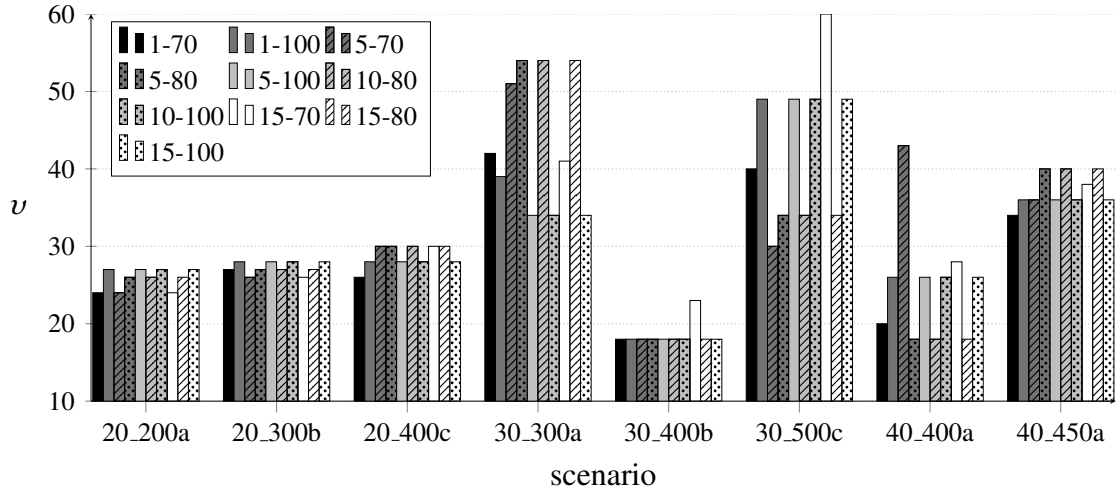


Figure 5.7.: SINR-corrected objective values υ for the interference mitigation formulation with TN coverage requirement based definition of the interfering sets with $\delta_c = 1$ and all settings with less than 5 % average percentage SINR violation.

as the differences between the various parameter combinations are not that pronounced, the determination of the best setting strongly depends on our considered aspects and their order. Nevertheless, we need one designated setting for the comparison of the different formulations in Section 5.9.4. In summary for the interference mitigation approach, the definition of the sets via the TN coverage requirement with a threshold $\delta_c = 1$ taking only the closest interfering BS into account and assuming that the TN coverage requirement fails for 70 % of the TNs, which can be served by one BS, turned out to be the best setting with respect to our evaluation criteria.

TN oriented formulation To find a good setting for the TN oriented formulation proposed in Section 5.7, we investigate $\varkappa \in \{1, 2, \ldots, 10\}$ for the number of potential interfering BSs, and the two possibilities of defining the serving BS. First, we observe that not every scenario can be solved within the time limit for $\varkappa \geq 7$ regardless of the selection of the serving BS. Hence, we only study $\varkappa \in \{1, 2, \ldots, 6\}$ in the following.

To evaluate the two possibilities to select the serving BS, we denote by υ_{all} the SINR-corrected objective value when the serving BS is selected from all potential serving BSs and by $\upsilon_{\text{strongest}}$ the SINR-corrected objective value when the serving BS is the one with the strongest signal. In Figure 5.8, we display the difference between υ_{all} and $\upsilon_{\text{strongest}}$ for $\varkappa \in \{1, 2, 3, 4, 5, 6\}$ and all scenarios, where a negative value signifies that the selection among all potential serving BSs gives a better result than the definition of the serving BS with the strongest signal. We observe that there does not exist an overall best setting. However, υ_{all} is more frequently below $\upsilon_{\text{strongest}}$ than the other way round. Thus, we regard the possibility to select the serving BS from all potential serving BSs as better than to choose the BS with the strongest signal as serving BS and do not investigate the latter case

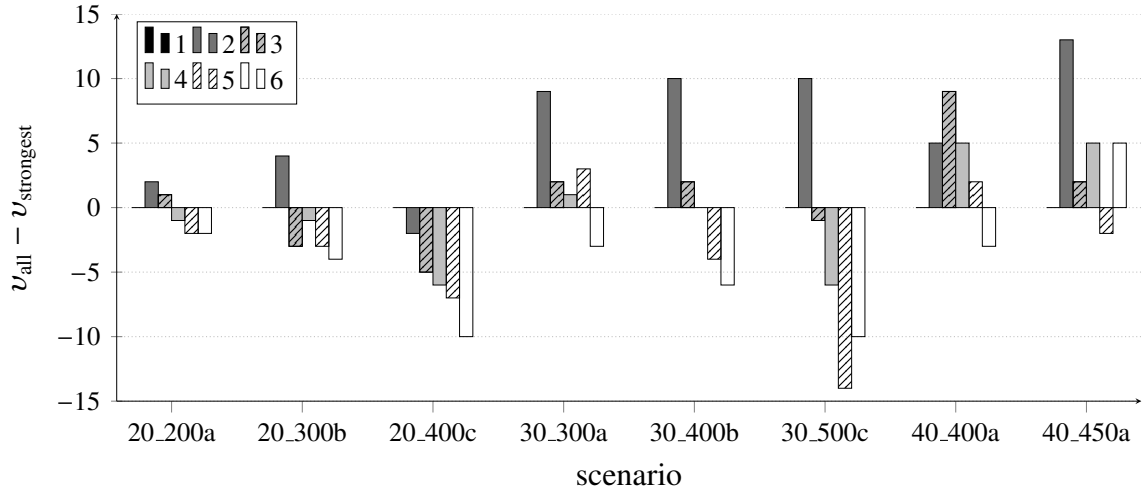


Figure 5.8.: Difference between SINR-corrected objective values v_{all} and $v_{\text{strongest}}$ for $\varkappa \in \{1, 2, 3, 4, 5, 6\}$ in the TN oriented formulation.

any further.

Now, we aim at finding a good value for the number of potential interfering BSs $\varkappa$. To this end, we display the SINR-corrected objective values for $\varkappa \in \{1, 2, 3, 4, 5, 6\}$ in Figure 5.9. Again, there does not exist an overall best value for $\varkappa$. However, $\varkappa \leq 3$ yields in general higher values than another $\varkappa$. Moreover, $\varkappa = 5$ yields the lowest values for the four scenarios 20_300b, 30_400b, 30_500c, 40_450a, while $\varkappa = 6$ ($\varkappa = 4$) yield the lowest values only for three (one) scenarios, 20_200a, 30_300a and 40_400a (20_400c). Thus, we apply $\varkappa = 5$ for the TN oriented formulation in the comparison given in Section 5.9.4.

Discrete CQIs formulation We investigate the two separation routines (via an ILP and combinatorial) for the discrete CQIs formulation (5.27). Irrespective of the separation routine, we have to decide at which node(s) of the B&B the separation should be performed. We study the separation at every node (the frequency of the separation callback is set to 1), and only at the root node and for every integer solution (the frequency is set to 0).

First, we investigate the performance of the separation based on the ILP proposed in Section 5.8.1. To this end, we analyse the restriction of the total number of separated violated inequalities per separation round to $\varkappa = 5$, 10, 15, 20 or not. We observe that even the smallest scenario 20_200a cannot be solved within the time limit of 2 h independent of the value of $\varkappa$ and the separation frequency. To obtain any result, we increase the time limit for the present study to 12 h. But even then only the smallest scenario can be solved to optimality as the larger scenarios either exceed the time limit or the available memory. We display the optimality gaps obtained for scenarios 20_200a, 20_300b, and 30_300a, for $\varkappa = 5$, 10, 15, 20, no restriction of the number of separated inequalities, and for a separation frequency of 0 and 1 in Table 5.7. We indicate settings which reach the memory limit by a $*$. The separation at every node leads in general to an excess of the available

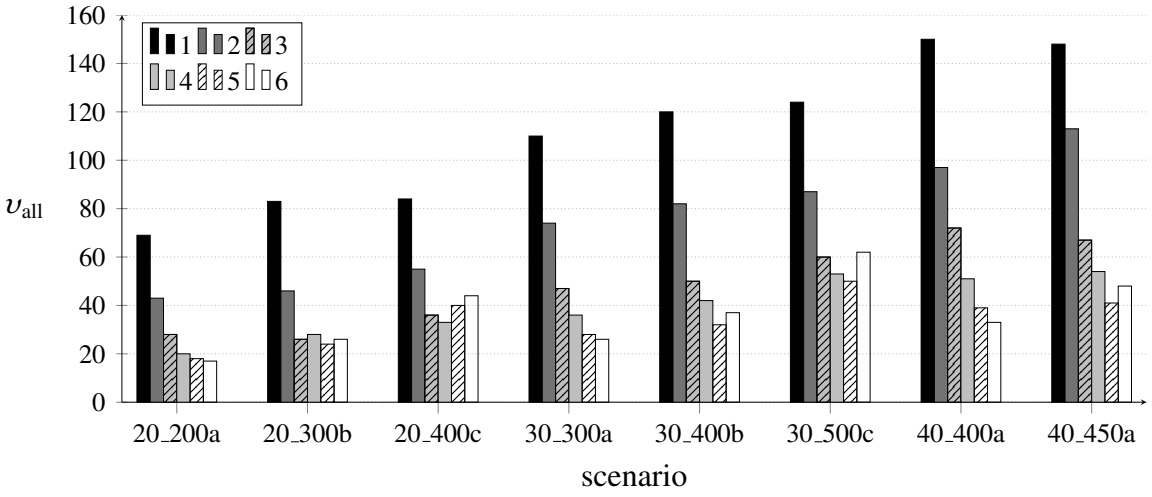


Figure 5.9.: SINR-corrected objective values v_{all} for $\varkappa \in \{1, 2, 3, 4, 5, 6\}$ in the TN oriented formulation.

frequency	0			1		
$\varkappa$	20_200a	20_300b	30_300a	20_200a	20_300b	30_300a
–	96.6*	610*	902.7*	0	220.7	263.1
5	4.8	1031.5*	822	17.6	48	358.7
10	331.3*	397.3*	791.7	0	248.3	737.6
15	382.2*	746.7*	1279.1*	0	96.1	212.1
20	283.7*	742.2*	1314.7*	0	161.9	269.9

Table 5.7.: Optimality gaps in % for selected scenarios, $\varkappa = 5$, 10, 15, 20, no restriction of the number of separated inequalities, and separation frequency 0 and 1 considering the ILP separation routine in the discrete CQI formulation. Settings which reach the memory limit are indicated by a $*$. All other gaps are obtained after 12 h.

frequency	20_200a	20_300b	30_300a
0	0	836.2	567.5
1	0	105.7	254.9

Table 5.8.: Optimality gaps in % for selected scenarios and separation frequency 0 and 1 considering the combinatorial separation routine in the discrete CQI formulation.

memory. Only the smallest scenario 20_200a is solved to optimality if the separation is performed at every node and for $\varkappa = 10$, 15, 20 or if the number of cuts is not restricted. The lowest solving time (23356 s) is observed by $\varkappa = 10$. For the other two scenarios and all settings, the obtained solution are too far from optimality to give meaningful objective values. Based on these results, we cannot expect larger scenarios to be solved to optimality by any setting of the ILP separation.

Now, we study the performance of the combinatorial separation routine presented in Section 5.8.2, where we also set a time limit of 12 h and test a separation frequency of 0 as well as 1. In Table 5.8, we depict the optimality gaps obtained for scenarios 20_200a, 20_300b, and 30_300a. Again, we can solve only scenario 20_200a to optimality and the gaps for the two other scenarios, which are obtained after the available memory is fully utilised, are too large to give meaningful solutions. Thus, we can also not expect a better performance for larger scenarios.

The lowest solving time for scenario 20_200a are 1257 s for a separation frequency of 1. Hence, the combinatorial separation is remarkably faster than the ILP separation for the only solvable scenario.

5.9.4. Comparison of Best Settings

In this section, we first compare all presented formulations with their best setting by means of the SINR-corrected objective value υ. Apart from the discrete CQIs formulation, these results are displayed in Table 5.9. Certainly, the SINR-corrected objective values are lowest for the SINR constraint formulation as no violated SINR conditions exist for an optimal solution. The second lowest values of υ provides the TN coverage requirement formulation while the remaining formulations yield significantly higher values. Since we cannot solve the exact versions of the interference mitigation and the TN oriented formulation, all models regarded in Table 5.9 only approximate the SINR requirements (5.1). Among these approximative methods, the TN coverage requirement formulation performs best.

The SINR-corrected objective value we obtain for scenario 20_200a by the discrete CQIs formulation is 22 and this is the exact optimal value as this formulation is exact. In contrast, the SINR constraint formulation gives a SINR-corrected objective value of 12 which indicates a significant violation of the capacity constraints. Otherwise, the optimal value computed by the discrete CQIs formulation would also be lower. Nevertheless, the corresponding models, which have to be solved for the exact discrete CQIs formulation, are

scenario	SINR	conflict	TN cov. requir.	iterative	mitigation	TN oriented
20_200a	12	27	15	17	24	18
20_300b	16	28	18	24	27	24
20_400c	18	31	22	32	26	40
30_300a	16	59	25	28	42	28
30_400b	16	18	16	32	18	32
30_500c	20	42	24	52	40	50
40_400a	16	30	21	32	20	39
40_450a	20	36	31	32	34	41

Table 5.9.: SINR-corrected objective values υ for different formulations of interference modelling with the best settings.

scenario	SINR	conflict	cov. requir.	iterative	mitigation	TN oriented
20_200a	21	35	16	18	31	18
20_300b	27	58	20	35	56	26
20_400c	55	65	51	46	50	349
30_300a	38	94	42	48	84	230
30_400b	40	44	37	54	43	37
30_500c	83	116	60	154	86	101
40_400a	56	68	52	58	42	309
40_450a	43	69	48	63	67	221

Table 5.10.: SINR-corrected objective values υ_ℓ including the load for different formulations of interference modelling with the best settings.

quite complex as discussed in Section 5.9.3 and hence, we cannot obtain optimal solutions for all other scenarios by means of the available resources.

As stated in Section 5.9.1, the spectral efficiencies in the capacity constraints, which are satisfied by the solutions of the different formulations, do not necessarily correspond to the actual SINR values. This results from the fact that correct SINR values can only be computed after the solution is known for most investigated formulations. As a consequence, even an optimal solution might violate a capacity constraint (4.4c). To account for such violations, we now analyse the SINR-corrected objective value υ_ℓ which incorporates the load of the deployed BSs and has been defined in Section 5.9.1. These values are displayed in Table 5.10 but again not for the discrete CQIs formulation. We observe that the values of υ_ℓ are, in general, considerably higher than υ stated in Table 5.9 implying that capacity constraints are strongly violated; the higher υ_ℓ, the higher the violation. The strongest effect can be seen for the TN oriented formulation, which indicates that the approximative version is far from exactness in terms of capacity. Overall, the TN coverage requirement formulation yields the lowest values of υ_ℓ, which are often even lower than the values obtained by the conventional SINR constraint formulation.

scenario	SINR	conflict	cov. requir.	iterative	mitigation	TN oriented
20_200a	0.8	0.2	0.7	0.6	0.2	18.1
20_300b	7.4	1.6	2.1	6.5	3.3	69.4
20_400c	18.2	6.2	17.8	12.8	8.8	146.3
30_300a	24.8	12.1	17.5	42.1	25.8	45.4
30_400b	3.6	0.9	3.4	3.7	14.0	264.7
30_500c	55.6	18.1	37.2	142.3	86.3	859.0
40_400a	29.9	2.7	18.3	11.4	115.8	16.1
40_450a	429.6	52.0	256.9	386.6	90.5	269.3
average	71.2	11.7	44.3	75.7	43.1	211.0

Table 5.11.: Time consumption in s for different formulations of interference modelling with the best settings.

Regarding the discrete CQIs formulation, we have $\upsilon = \upsilon_\ell = 22$ for 20_200a due to the exactness of this formulation in terms of SINR as well as capacity.

Finally, we compare the solving times for the different approximate formulations in Table 5.11.

For all formulations, the solving times are rather moderate. Certainly, more time is in general consumed to solve larger instances. However, outliers occur especially for the TN oriented formulation. The by far fastest approach is the conflict graph formulation with an average solving time of 11.7 s. This is not surprising as this formulation does not increase the complexity of the models significantly.

5.10. Conclusion

In this chapter, we have adapted and applied three approaches of interference limitation to the WNPP and have proposed four new formulations to model interference. Three of the new methods are exact in terms of SINR as well as capacity. However, for practical reasons, we have investigated only approximative versions for two of these approaches. In a computational study based on eight scenarios of various dimensions, we have identified the best parameter setting for each formulation. Then, we have compared these best settings by means of the SINR-corrected objective value υ, which incorporates the number of deployed BSs, covered TNs, and violated SINR constraints, as well as by means of υ_ℓ, which additionally accounts for violated capacity constraints via the load of a BS. The TN coverage requirement formulation has yielded the lowest values for both and thus, we regard this formulation as the best among the proposed for practical applicability.

However, we would like to point out that the SINR-corrected objective value including the load has also revealed that the solutions of all approximate formulations violate the capacity constraints even though hardly any or no SINR condition is violated. Thus, to obtain a solution that is feasible for SINR constraints as well as capacity constraints,

the proposed formulations have to be extended by a post-processing step which reassigns TNs after the deployment of BSs has been determined. Such a post-processing procedure though increases the complexity and cannot guarantee the optimality of a generated solution.

From a theoretical point of view, the exact methods, interference mitigation, TN oriented and discrete CQIs formulation model SINR requirements best. Among these three, the discrete CQIs method has the greatest potential to compute optimal solutions if the performance of the separation routine can be improved further. For the interference mitigation as well as the TN oriented formulation, an exponentially sized model has to be solved, where this exponential dimension cannot be avoided by a separation routine as in the case of discrete CQIs.

In the following main parts of this thesis, we study various robust optimisation aspects which entail rather complex models. Hence, a fairly simple basic formulation is necessitated to obtain still tractable models. The conflict graph formulation has proven to be the fastest among the various investigated formulations. The reason is that the addition of a conflict graph does not increase the complexity of the formulation significantly. Additionally, the interference limitation obtained by the conflict graph formulation is reasonable even though not as good as in case of the TN coverage requirement formulation. A further reason to choose the conflict graph formulation to limit interference in the remainder of this thesis is the fact that it enables the inclusion of versatile BS candidates.

As interference modelling is not the main focus of this thesis, this chapter is just a digression to complete the modelling of the planning of wireless networks.

Part II.

Chance Constraints

6. General Concept

Chance or probabilistic constraints are a specific model of stochastic optimisation to incorporate random parameters in optimisation problems. They were first mentioned in Charnes et al. [43] in 1958 in the context of scheduling of heating oil production. This work can be regarded as an example of the method for solving chance-constrained programs presented in the subsequent work by Charnes and Cooper [42]. This method splits the problem into two parts. First, a probability distribution that maximises the objective and respects the probability constraints is determined. In the second part, this distribution is approximated by a decision rule which is a concept of stochastic programming describing a rule to calculate decisions as functions of former observations. Further early work on chance-constrained programming can be found in Miller and Wagner [139] and Prékopa [156].

In the context of stochastic optimisation, chance constraints are in particular well suited to model problems involving high uncertainties or requiring reliability [164]. Though, the resulting formulations are typically severely computationally intractable due to the following reasons; see [16, 143]. First, the occurrent probability in the constraints is difficult to evaluate with high accuracy entailing that it can be quite costly to check if a chance constraint is satisfied. Second, the feasible set of the problem is often non-convex impeding the solving of such a problem. Moreover, Luedtke et al. [128] show that the problem

$$\begin{aligned} &\min c^t x \\ &\text{s.t. } \mathbb{P}\{Ax \geq b\} \geq 1 - \varepsilon \\ &\quad x \in \mathbb{R}^n_{\geq 0} \end{aligned}$$

with $A \in \mathbb{R}^{m \times n}$, b a random vector in $\mathbb{R}^m$, $c \in \mathbb{R}^n$ and $\varepsilon \in (0, 1)$ with $\varepsilon << 1$ is NP-hard. For a computationally intractable problem, a common approach is to replace the chance constraint by a tractable approximation which possibly involves new variables [16, 143]. Such an approximation is a sufficient condition for the validity of the chance constraint. Additionally, it can be verified efficiently and should define a convex and computationally tractable set in the space of the x-variables in the original formulation. For a so-called conservative approximation, the projection onto the space of the x-variables is contained in the feasible set of the chance-constrained model. An example is the scenario approximation where samples of realisations of the random vector (or matrix) are generated. All realisations have to fulfil the inner constraint ($Ax \geq b$) inside the probability term.

We distinguish between *joint* and *individual/separate* chance constraints. A joint chance constraint is of the form

$$\mathbb{P}\left\{a_j x \geq b_j \quad \forall\, j \in \{1, \ldots, m\}\right\} \geq 1 - \varepsilon, \tag{6.1}$$

where a_j is the j-th column of matrix $A \in \mathbb{R}^{m \times n}$ and b_j is the j-th entry of vector $b \in \mathbb{R}^m$ while the corresponding individual chance constraints can be written as

$$\mathbb{P}\left\{a_j x \geq b_j\right\} \geq 1 - \varepsilon_j \quad \forall\, j \in \{1, \ldots, m\} \tag{6.2}$$

with $\varepsilon_j \in (0, 1)$ small. The joint chance constraint guarantees a certain reliability for the whole problem, whereas the individual chance constraints can only guarantee a certain reliability for each constraint separately. Every vector $x \in \mathbb{R}^n$ fulfilling the joint chance constraint (6.1) also fulfils all separate chance constraints (6.2) if $\varepsilon_j = \varepsilon \,\forall\, j \in \{1, \ldots, m\}$, i. e., $z_{\text{joint}} \geq z_{\text{sep}}$ with

$$\begin{aligned} z_{\text{joint}} = \min\ & c^t x \\ \text{s.t. } & \mathbb{P}\left\{a_j x \geq b_j \quad \forall\, j\right\} \geq 1 - \varepsilon \\ & x \geq 0, \end{aligned} \qquad \begin{aligned} z_{\text{sep}} = \min\ & c^t x \\ \text{s.t. } & \mathbb{P}\left\{a_j x \geq b_j\right\} \geq 1 - \varepsilon \qquad \forall\, j \\ & x \geq 0 \end{aligned}$$

see [108]. Hence, the separate chance constraint program with objective z_{sep} represents a relaxation of the joint chance constraint program with objective z_{joint} and the probability that all separate constraints are fulfilled is $(1 - \varepsilon)^m < (1 - \varepsilon)$ when assuming independent random variables. In general, $\varepsilon_j \neq \varepsilon$ for separate chance constraints.

There exists a vast literature on chance-constrained programming also applied to a variety of problems such as scheduling in electricity markets [180], optimal power flow [112], or beamforming in cognitive radio networks [129] to name just some recent applications. However, only few papers on combinatorial problems exist. One of the first works on combinatorial problems with chance constraints was published by Hillier [95] in 1967. The author presents an exact solution procedure for 0-1 bounded chance-constrained models based on its deterministic equivalent form and deterministic uniformly tighter and looser constraints. Additionally, a two-stage approach is presented if the values of some random variables are known before some decision variables must be set.

Atamtürk and Narayanan [10] study the submodular knapsack polytope since chance constraints with only binary variables can be modelled as a submodular knapsack set under certain assumptions. A recent paper on chance-constrained combinatorial problems is published by Klopfenstein [108]. The author gives a detailed overview on existing literature and presents B&C algorithms to solve 0-1 chance-constrained IPs to optimality, where the right hand side is deterministic and the coefficient matrix is random. Furthermore, linear inequalities which are valid for the set of vectors fulfilling the individual chance constraints via so-called basic scenarios are proposed. An exponential number of such inequalities describes the set of feasible points completely. The developed solution algorithm separates these valid inequalities in a B&C framework and serves as a basis for an algorithm to solve a joint chance-constrained model.

One of the earliest works on solving general chance-constrained IPs to optimality was published by Tayur et al. [168]. The framework of the presented solution method can be described as follows. First, the "complicating constraints" such as joint chance constraints are relaxed and the resulting reduced IP is then solved to optimality. Afterwards, the

algorithm moves along specified directions (by Gröbner bases) to further feasible solutions of the reduced IP. The feasibility of the new solution is checked for the original problem. The algorithm is tested only for small scheduling problems (with at most 4 machines, 7 jobs).

In the case that only the right hand side is random with discrete distribution Beraldi and Ruszczyński [17] propose a B&B algorithm for probabilistic IPs with a joint chance constraint. The solution method is based on (partial) enumeration of so-called p-efficient points where p is the probability level which should be met by the chance constraint.

7. Application to Fixed Broadband Wireless Networks

In this chapter, we apply separate and joint chance constraints both presented in the previous chapter to the FBWN problem as stated in Section 4.2. The presented theoretical and computational results are based on several publications [46, 47, 51] originated from a collaboration with David Coudert from INRIA, Prof. Arie M. C. A. Koster from RWTH Aachen University and Napoleão Nepomuceno from Universidade de Fortaleza.

7.1. Separate Chance-Constrained Formulation

To model the random factor, radio link configuration, of the FBWN formalised in (4.6), we first introduce a random variable η_{uv}^p with a discrete probability distribution which represents the bandwidth efficiency of the current radio configuration for bandwidth choice p on arc uv. This configuration or modulation, varies over time in response to channel fluctuations. We replace the capacity constraints (4.6c) by the following *separate chance constraints*.

$$\mathbb{P}\left\{\sum_{k\in\mathcal{K}} x_{uv}^k \leq \sum_{p=1}^{P_{uv}} \eta_{uv}^p B_{uv}^p y_{uv}^p\right\} \geq 1-\varepsilon_{uv} \qquad \forall uv \in \mathcal{A}, \tag{7.1}$$

where $\varepsilon_{uv} > 0$ is the *infeasibility tolerance per link* which is typically close to zero and set by the network operator. These constraints ensure that the available capacity on each link, taking bandwidth choice as well as the random modulation into account, supports the total traffic routed through it with (high) probability $1-\varepsilon_{uv}$.

The incorporation of the chance constraints entails that the model is computationally intractable. However, we can reformulate this probabilistic program as a standard ILP since there exists only a finite though huge number of scenarios, i. e., realisations of the random vector η consisting of the random variables η_{uv}^p. To reduce the number of scenarios to be considered, we apply the approach of basic scenarios by [108]. To establish the deterministic counterpart of constraints (7.1), we introduce the parameter $\mathfrak{n}_{uv}^p$ computed from the (reverse) cumulative probability distribution of the random variable η_{uv}^p. It represents the highest radio configuration (number of bits) whose cumulative probability value ρ_{uv}^p is greater than $1-\varepsilon_{uv}$. Hence, $\rho_{uv}^p = \mathbb{P}\{\eta_{uv}^p \geq \mathfrak{n}_{uv}^p\} \geq 1-\varepsilon_{uv}$ where $\mathbb{P}\{\eta_{uv}^p \geq \mathfrak{n}_{uv}^p+1\} < 1-\varepsilon_{uv}$. An example to compute $\mathfrak{n}_{uv}^p$ for an infeasibility tolerance $\varepsilon_{uv} = 0.06$ is displayed in Table 7.1. The cumulative probability value for 32-QAM is the last value of the ascending modulation schemes to exceed $1-\varepsilon_{uv} = 0.94$, hence, 32-QAM is the highest possible modulation for this link uv.

modulation scheme	individual probability	(reverse) cumulative probability ρ^p_{uv}	
QPSK	0.02	1.00	
16-QAM	0.03	0.98	
32-QAM	0.03	0.95	$\rightarrow \mathfrak{n}^p_{uv} = \log_2(32) = 5$
64-QAM	0.02	0.92	
128-QAM	0.10	0.90	
256-QAM	0.80	0.80	

Table 7.1.: Example of (reverse) cumulative probability values ρ^p_{uv} and the parameter $\mathfrak{n}^p_{uv}$ for an infeasibility tolerance of $\varepsilon_{uv} = 0.06$.

The deterministic counterpart of separate chance constraints (7.1) can be stated as

$$\sum_{k \in \mathcal{K}} x^k_{uv} \leq \sum_{p=1}^{P_{uv}} \mathfrak{n}^p_{uv} B^p_{uv} y^p_{uv} \qquad \forall uv \in \mathcal{A}. \tag{7.2}$$

The separate infeasibility tolerance enforced on each link yields a lower bound for the objective when the joint chance constraint is included. This bound can become arbitrarily small if the number of links is large (compare z_{joint} and z_{sep} in Chapter 6). Therefore, the separate chance-constrained formulation is sensible only for smaller networks. In the following section, we present a formulation which is more appropriate to cope with instances of practically relevant sizes for a FBWN.

7.2. Joint Chance-Constrained Formulations

For a joint chance-constrained formulation, let $\varepsilon > 0$ be the *global infeasibility tolerance* which has the same properties as ε_{uv} named in the previous section. Now we replace the capacity constraints (4.6c) by the following *joint chance constraint*.

$$\mathbb{P}\left\{ \sum_{k \in \mathcal{K}} x^k_{uv} \leq \sum_{p=1}^{P_{uv}} \eta^p_{uv} B^p_{uv} y^p_{uv} \quad \forall uv \in \mathcal{A} \right\} \geq 1 - \varepsilon, \tag{7.3}$$

where η^p_{uv} denotes the random variable introduced in Section 7.1. This constraint enforces an infeasibility tolerance on all capacity constraints simultaneously. This means (7.3) guarantees that the assigned bandwidth is sufficient to support the total traffic routed through each arc of the network with probability $1 - \varepsilon$. Just like the separate chance constraints impair the solving of the model, also the incorporation of the joint chance constraint entails that the model is computationally intractable. Hence, the development of reformulations that can be solved with standard ILP solvers such as CPLEX [98] is desirable and studied in the following subsections.

7.2.1. Big-M Reformulation

One possible reformulation of (4.6) combined with (7.3) is a *big-M* ILP (cf. [128, 160]) considering a finite number of realisations $\eta^1, \ldots, \eta^R$ of the random vector η which occur with probabilities $\pi_1, \ldots, \pi_R$ and $\sum_{r=1}^{R} \pi_r = 1$. Let $z_r \in \{0,1\},\ r = 1, \ldots, R$ be decision variables, where $z_r = 0$ guarantees that the capacity constraints are satisfied taking into account realisation η^r. Setting $M := \sum_{k \in \mathcal{K}} d^k$, the chance constraint (7.3) can be rewritten as

$$\sum_{k \in \mathcal{K}} x_{uv}^k - M z_r \leq \sum_{p=1}^{P_{uv}} (\eta^r)_{uv}^p B_{uv}^p y_{uv}^p \qquad \forall uv \in \mathcal{A},\ r = 1, \ldots, R \tag{7.4a}$$

$$\sum_{r=1}^{R} \pi_r z_r \leq \varepsilon \tag{7.4b}$$

$$z_r \in \{0,1\} \qquad \forall\, r = 1, \ldots, R. \tag{7.4c}$$

The knapsack constraint (7.4b) is equivalent to

$$\sum_{r=1}^{R} \pi_r (1 - z_r) \geq 1 - \varepsilon. \tag{7.5}$$

The big-M constraints (7.4a) in association with (7.5) guarantee that the probability of scenarios which satisfy the capacity constraints is greater than or equal to $1 - \varepsilon$, thus enforcing the probabilistic constraint (7.3). Replacing the capacity constraints (4.6c) in model (4.6) by the big-M constraints (7.4), we obtain the *big-M formulation*

$$\min \sum_{uv \in \mathcal{A}} \sum_{p=1}^{P_{uv}} c_{uv}^p y_{uv}^p \tag{7.6a}$$

$$\text{s.t.} \sum_{u \in \delta^-(v)} x_{uv}^k - \sum_{u \in \delta^+(v)} x_{vu}^k = \begin{cases} -d^k, & v = s^k \\ d^k, & v = t^k \\ 0, & \text{otherwise} \end{cases} \qquad \forall\, v \in \mathcal{V},\ k \in \mathcal{K} \tag{7.6b}$$

$$\sum_{k \in \mathcal{K}} x_{uv}^k - M z_r \leq \sum_{p=1}^{P_{uv}} (\eta^r)_{uv}^p B_{uv}^p y_{uv}^p \qquad \forall uv \in \mathcal{A},\ r = 1, \ldots, R \tag{7.6c}$$

$$\sum_{r=1}^{R} \pi_r z_r \leq \varepsilon \tag{7.6d}$$

$$\sum_{p=1}^{P_{uv}} y_{uv}^p \leq 1 \qquad \forall uv \in \mathcal{A} \tag{7.6e}$$

$$x_{uv}^k \geq 0, y_{uv}^p,\ z_r \in \{0,1\} \qquad \begin{array}{l} \forall uv \in \mathcal{A},\ k \in \mathcal{K}, \\ p = 1, \ldots, P_{uv},\ r = 1, \ldots, R. \end{array} \tag{7.6f}$$

Computation of probabilities In general, the (merely unknown) correlation among outage events of different radio links prohibits the computation of the probabilities π_r. Under the assumption that microwave links suffer fades independently, we can define an artificial set of realisations. Note, we assume that the random variables of two links are independent and not that η_{uv}^p, $p = 1, \dots, P_{uv}$ on a single link $uv \in \mathcal{A}$ are independent. However, we will show in the following that the dependency between the bandwidth efficiencies η_{uv}^p, $p = 1, \dots, P_{uv}$ does not have any effects on the calculation of the relevant probabilities.

By the assumed independence between the links, we can limit the discussion to a single link $uv \in \mathcal{A}$. Let Q^p be the number of possible radio configurations for the chosen link uv and bandwidth choice p. Further, let $\mathcal{D}^p$ be the domain, which denotes the possible bandwidth efficiencies, of the random variable η_{uv}^p and define a bijection $f^p \colon \mathcal{D}^p \to \{1, \dots, Q^p\}$ with $f^p(\eta_{uv}^p) = q$ mapping bandwidth efficiency to radio configuration. For the sake of simplicity, we write f instead of f^p. For a fixed bandwidth choice $\tilde{p}$, the probability that link uv runs with radio configuration $\nabla \in \{1, \dots, Q^{\tilde{p}}\}$ is

$$\mathbb{P}\left\{f(\eta_{uv}^{\tilde{p}}) = \nabla\right\} = \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}}) = \nabla}^{R} \pi_r, \tag{7.7}$$

which is the sum over all probabilities of those realisations for which the entry $(\eta^r)_{uv}^{\tilde{p}}$ is also mapped to radio configuration ∇. Now, we define all possible bandwidth-independent realisations with probabilities π_r^* such that the probability $\mathbb{P}\left\{f(\eta_{uv}^{\tilde{p}}) = \nabla\right\}$ can be determined similar to (7.7). More precisely, let R^* be the number of all possible bandwidth and configuration constellations, that is $R^* = \prod_{p=1}^{P_{uv}} Q^p$ and $\pi_r^* := \prod_{p=1}^{P_{uv}} \mathbb{P}\{\eta_{uv}^p = (\eta^r)_{uv}^p\}$ for $r = 1, \dots, R^*$.

Lemma 7.1. *Let $\tilde{p} \in \{1, \dots, P_{uv}\}$ be a bandwidth choice and $\nabla \in \{1, \dots, Q^{\tilde{p}}\}$ a radio configuration. It holds that*

$$\sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}}) = \nabla}^{R^*} \pi_r^* = \mathbb{P}\left\{f(\eta_{uv}^{\tilde{p}}) = \nabla\right\}.$$

Proof. By definition of π_r^*, we obtain

$$\begin{aligned}\sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}}) = \nabla}^{R^*} \pi_r^* &= \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}}) = \nabla}^{R^*} \left(\prod_{p=1}^{P_{uv}} \mathbb{P}\{\eta_{uv}^p = (\eta^r)_{uv}^p\} \right) \\ &= \mathbb{P}\left\{f(\eta_{uv}^{\tilde{p}}) = \nabla\right\} \cdot \underbrace{\sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}}) = \nabla}^{R^*} \left(\prod_{p=1 \mid p \neq \tilde{p}}^{P_{uv}} \mathbb{P}\{\eta_{uv}^p = (\eta^r)_{uv}^p\} \right)}_{(\star)}\end{aligned}$$

It remains to show that $(\star) = 1$. For this purpose, we fix another bandwidth choice $\hat{p} \in \{1, \dots, P_{uv}\} \setminus \{\tilde{p}\}$ and separate all corresponding summands (regarding the radio configuration) as follows.

$$
\begin{aligned}
(\star) &= \sum_{q=1}^{Q^{\hat{p}}} \left[\mathbb{P}\left\{ f(\eta_{uv}^{\hat{p}}) = q \right\} \cdot \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}})=\nabla}^{R^*} \left(\prod_{p=1 \mid p \neq \tilde{p}, \hat{p}}^{P_{uv}} \mathbb{P}\left\{ \eta_{uv}^p = (\eta^r)_{uv}^p \right\} \right) \right] \\
&= \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}})=\nabla}^{R^*} \left(\prod_{p=1 \mid p \neq \tilde{p}, \hat{p}}^{P_{uv}} \mathbb{P}\left\{ \eta_{uv}^p = (\eta^r)_{uv}^p \right\} \right) \cdot \underbrace{\sum_{q=1}^{Q^{\hat{p}}} \mathbb{P}\left\{ f(\eta_{uv}^{\hat{p}}) = q \right\}}_{=1}
\end{aligned}
$$

Separating all bandwidth choices subsequently that way, it follows $(\star) = 1$ and the proof is complete. □

Corollary 7.2. *The probability that a link with a fixed bandwidth runs with a certain radio configuration can be computed either via bandwidth-dependent or bandwidth-independent realisations, i. e.,*

$$
\mathbb{P}\left\{ f(\eta_{uv}^{\tilde{p}}) = \nabla \right\} = \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}})=\nabla}^{R} \pi_r = \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}})=\nabla}^{R} \pi_r^* = \sum_{r=1 \mid f((\eta^r)_{uv}^{\tilde{p}})=\nabla}^{R} \prod_{p=1}^{P_{uv}} \mathbb{P}\left\{ \eta_{uv}^p = (\eta^r)_{uv}^p \right\}.
$$

Despite the presented possibility to compute the probabilities π_r, the big-M formulation (7.6) is intractable due to the very large number of scenarios or realisations of the random vector η, to be considered. Additionally, big-M models are often numerically unstable; cf. Chapter 5. In the sequel, we propose a computationally more tractable ILP model in case of independent link outages.

7.2.2. Independent Link Outages

Henceforth, we assume that the link outages are independent. Hence, also η_{uv}^p, $p = 1, \dots, P_{uv}$ are assumed to be independent for a link $uv \in \mathcal{A}$ and the probability in (7.3) can be replaced by the product of probabilities. To do so, let Q_{uv}^p be the number of possible configurations for arc uv with assigned bandwidth choice p. The index of a configuration is denoted by $q = 1, \dots, Q_{uv}^p$ where a larger q means a higher-level (larger m) m-QAM modulation scheme. Moreover, let ρ_{uv}^{pq} be the probability that arc uv with bandwidth choice p is running at configuration q or higher. Thus, ρ_{uv}^{pq} is the (reverse) cumulative probability value of the random variable regarding modulation q; cf. Section 7.1, and the minimum available capacity of arc uv operating at bandwidth choice p with configuration q is now represented by B_{uv}^{pq}. Moreover, we append the additional index q to the decision variable y and introduce a binary slack variable y_{uv}^0 which is set to 1 if arc uv is not operated. Without this additional variable, a zero-product when rewriting (7.3) as the product of probabilities

could occur. The complete reformulated model then reads

$$\min \sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} c_{uv}^p y_{uv}^{pq} \tag{7.8a}$$

$$\text{s.t.} \sum_{u\in\delta^-(v)} x_{uv}^k - \sum_{u\in\delta^+(v)} x_{vu}^k = \begin{cases} -d^k, & v = s^k \\ d^k, & v = t^k \\ 0, & \text{otherwise} \end{cases} \quad \forall\, v \in \mathcal{V},\, k \in \mathcal{K} \tag{7.8b}$$

$$\sum_{k\in\mathcal{K}} x_{uv}^k \le \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} B_{uv}^{pq} y_{uv}^{pq} \quad \forall uv \in \mathcal{A} \tag{7.8c}$$

$$\prod_{uv\in\mathcal{A}} \left(y_{uv}^0 + \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \rho_{uv}^{pq} y_{uv}^{pq} \right) \ge 1 - \varepsilon \tag{7.8d}$$

$$y_{uv}^0 + \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} y_{uv}^{pq} = 1 \quad \forall uv \in \mathcal{A} \tag{7.8e}$$

$$x_{uv}^k \ge 0, y_{uv}^{pq}, y_{uv}^0 \in \{0,1\} \quad \begin{array}{l} \forall uv \in \mathcal{A},\, k \in \mathcal{K}, \\ p = 1,\dots,P_{uv},\, q = 1,\dots,Q_{uv}^p. \end{array} \tag{7.8f}$$

The objective (7.8a) is the same as (4.6a) just incorporating the choice on the radio configuration. Furthermore, the flow conservation constraints (7.8b) are exactly (4.6b). In the capacity constraints (7.8c), we explicitly assume a hypothesis on the radio configuration: For a given arc and bandwidth, the lower the configuration is, the lower the bandwidth efficiency assumed to this arc will be in time of design and also the higher the probability that the effective capacity on this arc in time of operation supports the total traffic to be routed through it. In other words, more conservative hypotheses on the radio configuration lead to more reliable solutions. Constraint (7.8d) formally denotes this relation. According to the bandwidth assignment and the hypotheses on the radio configuration, it guarantees that the reliability of a solution is at least $1 - \varepsilon$. The configuration constraints (7.8e) are equal to (4.6d) just incorporating the additional slack variable y_{uv}^0.

Theorem 7.3. *The big-M formulation* (7.6) *and model* (7.8) *are equivalent in case of independent link outages.*

Proof. Similar to Section 7.2.1, let $\mathcal{D}_{uv}^p$ be the domain of the random variable η_{uv}^p and we define the bijection $f_{uv}^p \colon \mathcal{D}_{uv}^p \to \{1,\dots,Q_{uv}^p\}$ with $f_{uv}^p(\eta_{uv}^p) = q$ which maps bandwidth efficiency to radio configuration. For the sake of simplicity, we write f instead of f_{uv}^p. We show that for every feasible bandwidth assignment and routing of traffic demands to the big-M formulation, there exists a corresponding feasible solution to formulation (7.8) with the same cost, and vice versa. Let $(\tilde{x}, \tilde{y}, \tilde{z})$ be a feasible solution for the big-M formulation. Recall that for feasible realisations $r = 1,\dots,R$ of η holds

$$\tilde{z}_r = 0 \iff \sum_{k\in\mathcal{K}} \tilde{x}^k_{uv} \le \sum_{p=1}^{P_{uv}} (\eta^r)^p_{uv} B^p_{uv} \tilde{y}^p_{uv} \qquad \forall\, uv \in \mathcal{A}.$$

We can easily obtain a feasible solution $(\tilde{x}, \hat{y})$ to (7.8) with the same cost as shown in the following. Let us define

$$\mathcal{A}_1 := \left\{ uv \in \mathcal{A} \,\middle|\, \sum_{p=1}^{P_{uv}} \tilde{y}^p_{uv} = 1 \right\} \text{ and } \mathcal{A}_0 := \mathcal{A} \setminus \mathcal{A}_1 \tag{7.9}$$

as the sets of installed and non-installed arcs, respectively, and let $\tilde{p}_{uv}$, $uv \in \mathcal{A}_1$, be the selected bandwidth choice for arc uv, i. e., $\tilde{y}^{\tilde{p}_{uv}}_{uv} = 1$. For simplicity, whenever it is understood from the context, we write $\tilde{p}$ instead of $\tilde{p}_{uv}$. Moreover, we define the largest radio configuration necessary to fulfil the capacity constraints (7.4a) for every feasible realisation as

$$\tilde{q}_{uv} := f\left(\min_{r=1,\dots,R} \left\{ (\eta^r)^{\tilde{p}}_{uv} \,\middle|\, \tilde{z}_r = 0 \right\} \right) \qquad \forall\, uv \in \mathcal{A}_1.$$

Again, we write $\tilde{q}$ instead of $\tilde{q}_{uv}$ whenever it is unambiguous. Note, $\tilde{z}_r = 0,\ r \in \{1,\dots,R\}$ if $(\eta^r)^{\tilde{p}}_{uv} \ge f^{-1}(\tilde{q})$ for all $uv \in \mathcal{A}$ simultaneously. Then we set

$$\hat{y}^{pq}_{uv} := \begin{cases} 1, & \text{if } uv \in \mathcal{A}_1, p = \tilde{p} \text{ and } q = \tilde{q} \\ 0, & \text{otherwise.} \end{cases}$$

Besides, we set

$$\hat{y}^0_{uv} := \begin{cases} 1, & \text{if } uv \in \mathcal{A}_0 \\ 0, & \text{if } uv \in \mathcal{A}_1. \end{cases}$$

By definition of $\hat{y}$, we have

$$\sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q^p_{uv}} c^p_{uv} \hat{y}^{pq}_{uv} = \sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} c^p_{uv} \tilde{y}^p_{uv},$$

which demonstrates that the bandwidth cost are equal. Furthermore, constraints (7.8b), (7.8c), and (7.8e) are fulfilled with $B^{pq}_{uv} = f^{-1}(q) \cdot B^p_{uv}$.

To prove the feasibility of the constructed solution $(\tilde{x}, \hat{y})$, it remains to show that the reliability constraint (7.8d) is fulfilled. For this purpose, we introduce the following notation. First, we reduce the space of the random vector η such that we consider only the variables $\eta^{\tilde{p}}_{uv}$, $\forall\, uv \in \mathcal{A}_1$. Let $\tilde{\eta}$ be this reduced random vector. Again we have to deal with a finite number of realisations $\tilde{\eta}^1, \dots, \tilde{\eta}^S$ of the random vector $\tilde{\eta}$. Consider the set

$$\tilde{\mathcal{S}} := \left\{ s \in \{1,\dots,S\} \,\middle|\, f((\tilde{\eta}^s)^{\tilde{p}}_{uv}) \ge \tilde{q}\ \forall\, uv \in \mathcal{A}_1 \right\}$$

of feasible realisations of $\tilde{\eta}$ with respect to solution $\tilde{y}$. Then it holds for the left hand side of (7.8d)

$$
\begin{aligned}
&\prod_{uv\in\mathcal{A}}\left(\hat{y}^0_{uv}+\sum_{p=1}^{P_{uv}}\sum_{q=1}^{Q^p_{uv}}\rho^{pq}_{uv}\hat{y}^{pq}_{uv}\right)\\
&=\underbrace{\prod_{uv\in\mathcal{A}_0}\left(\underbrace{\hat{y}^0_{uv}}_{=1}+\sum_{p=1}^{P_{uv}}\sum_{q=1}^{Q^p_{uv}}\rho^{pq}_{uv}\underbrace{\hat{y}^{pq}_{uv}}_{=0}\right)}_{=1}\cdot\prod_{uv\in\mathcal{A}_1}\left(\underbrace{\hat{y}^0_{uv}}_{=0}+\sum_{p=1}^{P_{uv}}\sum_{q=1}^{Q^p_{uv}}\rho^{pq}_{uv}\hat{y}^{pq}_{uv}\right)\\
&=\prod_{uv\in\mathcal{A}_1}\sum_{p=1}^{P_{uv}}\sum_{q=1}^{Q^p_{uv}}\rho^{pq}_{uv}\hat{y}^{pq}_{uv}=\prod_{uv\in\mathcal{A}_1}\left(\rho^{\tilde{p}\tilde{q}}_{uv}\underbrace{\hat{y}^{\tilde{p}\tilde{q}}_{uv}}_{=1}\right)\\
&\overset{\text{def. of }\rho\text{ and }f}{=}\prod_{uv\in\mathcal{A}_1}\mathbb{P}\left\{f(\eta^{\tilde{p}}_{uv})\geq\tilde{q}_{uv}\right\}=\sum_{s\in\tilde{\mathcal{S}}}\prod_{uv\in\mathcal{A}_1}\mathbb{P}\left\{\eta^{\tilde{p}}_{uv}=(\tilde{\eta}^s)^{\tilde{p}}_{uv}\right\} \qquad (*)\\
&\overset{(+)}{=}\sum_{r=1\,|\,\tilde{z}_r=0}^{R}\prod_{uv\in\mathcal{A}}\prod_{p=1}^{P_{uv}}\mathbb{P}\left\{\eta^p_{uv}=(\eta^r)^p_{uv}\right\}\\
&\overset{\text{Cor. 7.2}}{=}\sum_{r=1\,|\,\tilde{z}_r=0}^{R}\pi_r=\sum_{r=1}^{R}\pi_r(1-\tilde{z}_r)\geq 1-\varepsilon.
\end{aligned}
$$

For (+), we use a similar argumentation of separating bandwidth choices subsequently as in the proof of Lemma 7.1. In the last equality, we apply Corollary 7.2 and exploit the independence of $\eta^p_{uv},\ p=1,\dots,P_{uv}$. Altogether, $(\tilde{x},\hat{y})$ is a feasible solution for (7.8).

Conversely, given a feasible solution $(\tilde{x},\hat{y})$ to formulation (7.8), one can obtain a feasible solution $(\tilde{x},\tilde{y},\tilde{z})$ to the big-M formulation with the same cost. We set

$$\tilde{y}^p_{uv}:=\sum_{q=1}^{Q^p_{uv}}\hat{y}^{pq}_{uv}\qquad\forall uv\in\mathcal{A},p=1,\dots,P_{uv}$$

and define $\hat{p}_{uv}$, $\hat{q}_{uv}$ such that $\hat{y}^{\hat{p}_{uv}\hat{q}_{uv}}_{uv}=1$ for all $uv\in\mathcal{A}_1$, where $\mathcal{A}_1$ is defined in (7.9). For $r=1,\dots,R$, we set

$$\tilde{z}_r:=\begin{cases}0, & \text{if } f((\eta^r)^{\hat{p}_{uv}}_{uv})\geq\hat{q}_{uv}\ \forall uv\in\mathcal{A}_1\\ 1, & \text{otherwise.}\end{cases}$$

Again, the objectives (4.6a) and (7.8a) are equal and constraints (4.6b), (4.6d), and (7.4a) are fulfilled. To show that constraint (7.4b) is satisfied, we follow the same argumentation as before in (∗), just in the reverse direction and by replacing $\tilde{p}$ and $\tilde{q}$ by $\hat{p}_{uv}$ and $\hat{q}_{uv}$.

Consequently, the big-M formulation and model (7.8) are equivalent in case of independent link outages. □

Model (7.8) is hard to solve due to the non-linearity of constraint (7.8d). However, this constraint can be linearised as follows. By employing monotonicity of logarithmic functions and because the logarithm of a product is equal to the sum of the logarithms,

(7.8d) is equivalent to

$$\sum_{uv\in\mathcal{A}} \log\left(1 \cdot y_{uv}^0 + \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \rho_{uv}^{pq} y_{uv}^{pq}\right) \geq \log(1-\varepsilon).$$

By (7.8e), exactly one of the sum elements within the logarithmic function will be non-zero and, hence, the constraint is equivalent to

$$\sum_{uv\in\mathcal{A}} \left(\underbrace{\log(1)}_{=0} y_{uv}^0 + \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \log(\rho_{uv}^{pq}) y_{uv}^{pq}\right) \geq \log(1-\varepsilon).$$

Thus, the slack variables y_{uv}^0 are not required anymore. The complete linearised problem in case of independent link outages is modelled by the following ILP.

$$\min \sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} c_{uv}^p y_{uv}^{pq} \tag{7.10a}$$

$$\text{s.t.} \sum_{u\in\delta^-(v)} x_{uv}^k - \sum_{u\in\delta^+(v)} x_{vu}^k = \begin{cases} -d^k, & v = s^k \\ d^k, & v = t^k \\ 0, & \text{otherwise} \end{cases} \qquad \forall\, v \in \mathcal{V},\, k \in \mathcal{K} \tag{7.10b}$$

$$\sum_{k\in\mathcal{K}} x_{uv}^k \leq \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} B_{uv}^{pq} y_{uv}^{pq} \qquad \forall uv \in \mathcal{A} \tag{7.10c}$$

$$\sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \log(\rho_{uv}^{pq}) y_{uv}^{pq} \geq \log(1-\varepsilon) \tag{7.10d}$$

$$\sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} y_{uv}^{pq} \leq 1 \qquad \forall uv \in \mathcal{A} \tag{7.10e}$$

$$x_{uv}^k \geq 0, y_{uv}^{pq} \in \{0,1\} \qquad \begin{array}{l} \forall uv \in \mathcal{A},\, k \in \mathcal{K}, \\ p = 1, \ldots, P_{uv}, \\ q = 1, \ldots, Q_{uv}^p. \end{array} \tag{7.10f}$$

Note, this formulation is still a large scale ILP, which is, in general, hard to solve.

7.2.3. Budget-Constrained Formulation for Independent Link Outages

The problem formulation (7.10) minimises the cost for the bandwidth allocation while a certain reliability is guaranteed. But there is a trade-off between cost and reliability. Depending on the value of the infeasibility tolerance, many problems may be infeasible since the reliability constraint (7.10d) becomes too restrictive. Investigating the problem

from the opposite perspective, we ask how reliable can the network be while a certain budget $\mathfrak{B}$ is not exceeded? This perception of the problem is formalised in the following ILP.

$$\max \sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \log(\rho_{uv}^{pq}) y_{uv}^{pq} \tag{7.11a}$$

$$s.t.\ (7.10b), (7.10c), (7.10e), (7.10f) \tag{7.11b}$$

$$\sum_{uv\in\mathcal{A}} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} c_{uv}^p y_{uv}^{pq} \leq \mathfrak{B} \tag{7.11c}$$

The formulation (7.11) maximises the reliability of the network while the budget $\mathfrak{B}$ is not exceeded and it is also a special case of a network design problem extended by a knapsack constraint, the budget constraint (7.11c). Thus, (7.11) is strongly NP-hard.

7.2.4. Dependent Random Variables

In real world applications, the random variables η_{uv}^p are usually not independent as, e.g., bad weather conditions influence more than one link at the same time. Nevertheless, we can embed the presented formulation (7.10) in a B&B framework on the basis of [78] via a Benders like decomposition to model the case of dependent random variables as described in the following.

First, we solve model (7.10) where the probabilities ρ_{uv}^{pq} in constraint (7.10d) describe the marginal probabilities on a single link. Every integer solution $(\tilde{x}, \tilde{y})$ found during the B&B process is then tested for feasibility regarding the actual (dependent) random variables. This means, we fix the binary decision variables y to $\tilde{y}$ and determine a corresponding flow x that maximises the probability given in (7.3). If the computed probability is less than $1-\varepsilon$, the configuration given by $\tilde{y}$ is not part of a feasible solution in case of dependent random variables and hence, the current solution has to be prohibited. In such a case, we redefine

$$\mathcal{A}_1 := \left\{ uv \in \mathcal{A} \;\middle|\; \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \tilde{y}_{uv}^{pq} = 1 \right\} \text{ and } \mathcal{A}_0 := \mathcal{A} \setminus \mathcal{A}_1$$

as the sets of installed and non-installed links, respectively. Moreover, define $\tilde{p}_{uv}$ and $\tilde{q}_{uv}$ for every $uv \in \mathcal{A}_1$ such that $\tilde{y}_{uv}^{\tilde{p}_{uv}\tilde{q}_{uv}} = 1$. Based on these values and the sets $\mathcal{A}_1$ and $\mathcal{A}_0$, we add the following constraint as a so-called *lazy constraint* prohibiting the current solution $(\tilde{x}, \tilde{y})$.

$$\sum_{uv\in\mathcal{A}_1} y_{uv}^{\tilde{p}_{uv}\tilde{q}_{uv}} + \sum_{uv\in\mathcal{A}_0} \left(1 - \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} y_{uv}^{pq} \right) \leq |\mathcal{A}| - 1 \tag{7.12}$$

Including the new constraint, we continue the B&B routine solving (7.10). The whole process is depicted in Figure 7.1 and continues as long as the B&B algorithm provides new integer solutions.

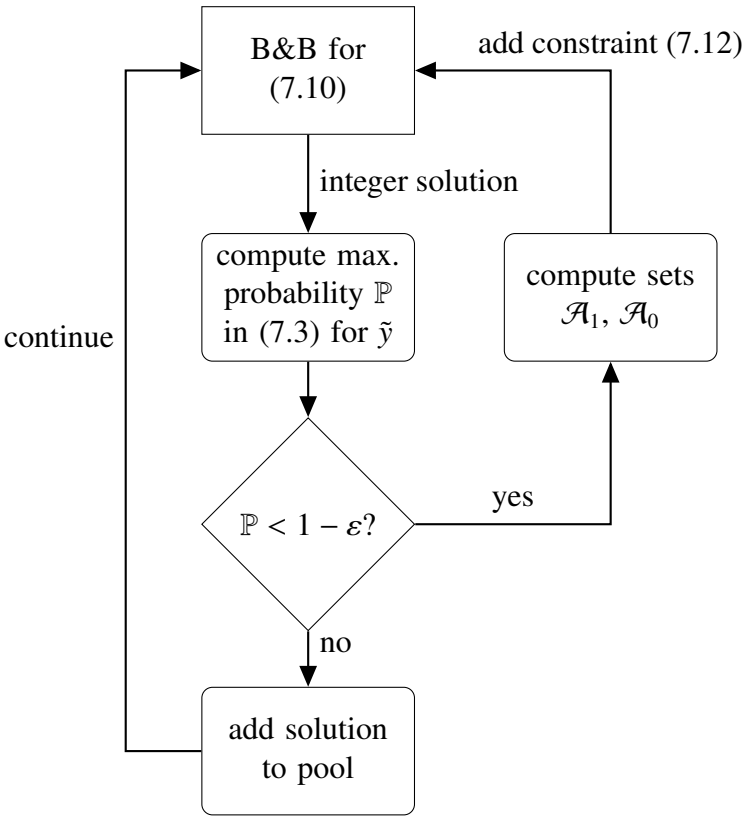


Figure 7.1.: Flowchart of lazy constraint separation during the B&B solution framework for dependent random variables, see [51].

This framework models random variable dependencies. However, we would like to point out that this is just a theoretical model as the computation of the maximum probability for fixed $\tilde{y}$ is typically intractable due to the correlation between the random variables.

7.3. Performance Improvements

Even small instances of the models described in Section 7.2.2 and 7.2.3 cannot be solved within a reasonable time due to the NP-hardness of the problem. To accelerate the B&B solving process, we present cutset inequalities to improve the dual bound. We separate these inequalities on the fly by exact separation ILPs. Finally, to improve the primal bound of the budget-constrained model (7.11) we introduce a primal heuristic.

7.3.1. Cutset Inequalities

Neglecting constraints (7.10d) and (7.10e), formulation (7.10) defines a classical network design problem studied intensively in the literature [24, 25, 131, 132, 157]. To enhance the performance of ILP solvers, several valid inequalities have been introduced for the classical network design problem. A special class are so-called cutset-based inequalities

which exploit knowledge about the required capacity on a cut in the network; see also Section 1.2.2.

Let $\mathcal{S} \subset \mathcal{V}$ be a proper and non-empty subset of the node set $\mathcal{V}$ and $\overline{\mathcal{S}} = \mathcal{V} \setminus \mathcal{S}$ its complement. The set

$$\mathcal{A}(\mathcal{S}, \overline{\mathcal{S}}) := \left\{ uv \in \mathcal{A} \,|\, u \in \mathcal{S}, v \in \overline{\mathcal{S}} \right\},$$

connecting a node in $\mathcal{S}$ to a node in $\overline{\mathcal{S}}$, defines a *cutset*. Similarly, let

$$\mathcal{K}(\mathcal{S}, \overline{\mathcal{S}}) := \left\{ k \in \mathcal{K} \,|\, s^k \in \mathcal{S}, t^k \in \overline{\mathcal{S}} \right\}$$

be the set of commodities originating in $\mathcal{S}$ and terminating in $\overline{\mathcal{S}}$. Finally, define

$$d(\mathcal{S}, \overline{\mathcal{S}}) := \sum_{k \in \mathcal{K}(\mathcal{S}, \overline{\mathcal{S}})} d^k.$$

An appropriate aggregation of constraints (7.10b), (7.10c), and non-negativity of the variables results in the following *base cutset inequalities*.

$$\sum_{uv \in \mathcal{A}(\mathcal{S}, \overline{\mathcal{S}})} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} B_{uv}^{pq} y_{uv}^{pq} \geq d(\mathcal{S}, \overline{\mathcal{S}}) \qquad \forall\, \mathcal{S} \subset \mathcal{V} \tag{7.13}$$

These inequalities ensure that there is enough capacity assigned to the arcs of any cutset in order to satisfy the demands that must be routed through the cutset. Base cutset inequalities are necessary for a capacity vector to be feasible, but it is well-known that they are not sufficient in general [53, 56]. By applying Chvátal-Gomory (CG) rounding to base cutset inequalities (cf. [179]), we obtain the well-known *cutset inequalities*

$$\sum_{uv \in \mathcal{A}(\mathcal{S}, \overline{\mathcal{S}})} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \left\lceil \frac{B_{uv}^{pq}}{a} \right\rceil y_{uv}^{pq} \geq \left\lceil \frac{d(\mathcal{S}, \overline{\mathcal{S}})}{a} \right\rceil \qquad \forall\, \mathcal{S} \subset \mathcal{V}, \tag{7.14}$$

where $a \in \{B_{uv}^{pq} \,|\, uv \in \mathcal{A}(\mathcal{S}, \overline{\mathcal{S}}), p = 1, \ldots, P_{uv}, q = 1, \ldots, Q_{uv}^p\}$. In general, the LP relaxation of (7.10) does not satisfy (7.14). On the contrary, all integer solutions have to satisfy it as explained in Section 1.2.2; see also [157].

A further class of valid inequalities are the *shifted cutset inequalities*. We obtain these inequalities from the base cutset inequalities by shifting the coefficients first before applying CG-rounding. Given a cutset $\mathcal{A}(\mathcal{S}, \overline{\mathcal{S}})$ and a link $uv \in \mathcal{A}(\mathcal{S}, \overline{\mathcal{S}})$, define

$$a_{uv} := \min_{p=1,\ldots,P_{uv}} \min_{q=1,\ldots,Q_{uv}^p} B_{uv}^{pq}, \text{ and}$$
$$a' \in \left\{ B_{uv}^{pq} - a_{uv} \,\middle|\, uv \in \mathcal{A}(\mathcal{S}, \overline{\mathcal{S}}), p = 1, \ldots, P_{uv}, q = 1, \ldots, Q_{uv}^p \right\} \setminus \{0\}.$$

Note, the parameter a_{uv} is strictly greater than 0 as $B_{uv}^{pq} > 0$. Multiplying constraints (7.10e)

by $-a_{uv}$ results in

$$\sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} (-a_{uv}) y_{uv}^{pq} \geq -a_{uv} \qquad \forall\, uv \in \mathcal{A}. \tag{7.15}$$

Adding $\sum_{uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} (-a_{uv}) y_{uv}^{pq}$ to the left hand side of (7.14) and $\sum_{uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})} -a_{uv}$ to the right hand side leads to a valid inequality due to (7.15). Then, we apply CG-rounding using the notation $a(\mathcal{S},\overline{\mathcal{S}}) \coloneqq \sum_{uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})} a_{uv}$ and obtain the shifted cutset inequalities

$$\sum_{uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \left\lceil \frac{B_{uv}^{pq} - a_{uv}}{a'} \right\rceil y_{uv}^{pq} \geq \left\lceil \frac{d(\mathcal{S},\overline{\mathcal{S}}) - a(\mathcal{S},\overline{\mathcal{S}})}{a'} \right\rceil \qquad \forall\, \mathcal{S} \subset \mathcal{V}. \tag{7.16}$$

Note that the presented cutset inequalities are valid for both formulations (7.10) and (7.11).

7.3.2. Separation of Cutset Inequalities

As there exist exponentially many subsets $\mathcal{S} \subset \mathcal{V}$, it is not efficient to add all possible cutset inequalities and shifted cutset inequalities described in the previous section. Hence, we rather generate only violated inequalities on the fly. Since no polynomial algorithm is known, we propose ILPs to separate the most violated (shifted) cutset inequality for the current LP solution exactly; see, e.g., [75] or [117] for cutset separation in the robust network design problem. A cutset inequality (7.14) is violated if

$$\sum_{uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})} \sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \left\lceil \frac{B_{uv}^{pq}}{a} \right\rceil \tilde{y}_{uv}^{pq} - \left\lceil \frac{d(\mathcal{S},\overline{\mathcal{S}})}{a} \right\rceil < 0,$$

where $\tilde{y}_{uv}^{pq}$ is part of the current LP solution.

For the exact separation of cutset inequalities given a, we introduce variables α_v indicating whether node $v \in \mathcal{V}$ is an element of the subset $\mathcal{S}$, and variables β_{uv} deciding whether $uv \in \mathcal{A}(\mathcal{S},\overline{\mathcal{S}})$. For simplicity we further define

$$D \coloneqq \frac{\sum\limits_{k \in \mathcal{K}} d^k \beta_{s^k t^k}}{a}.$$

The exact separation of violated cutset inequalities can be formulated as the following ILP.

$$\min \sum_{uv \in \mathcal{A}} \left(\sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \left\lceil \frac{B_{uv}^{pq}}{a} \right\rceil \tilde{y}_{uv}^{pq} \right) \beta_{uv} - z \tag{7.17a}$$

$$s.t.\ D \leq z \leq D + \frac{a-1}{a} \tag{7.17b}$$

$$\alpha_u - \alpha_v \leq \beta_{uv} \leq \min\{1 - \alpha_v, \alpha_u\} \qquad \forall\, u, v \in \mathcal{V} \tag{7.17c}$$

$$\alpha_v,\, \beta_{uv} \in \{0, 1\},\, z \in \mathbb{N} \qquad \forall\, u, v \in \mathcal{V}. \tag{7.17d}$$

If the optimal objective value is negative, then a violated cutset inequality is found. The variable z together with constraint (7.17b) determines the rounding of the right hand side of the cutset inequality, where $\frac{a-1}{a}$ depicts a number close to but strictly less than 1. Constraints (7.17c) determine the link between variables β_{uv}, α_v, and α_u, which is

$$\beta_{uv} = 1 \iff \alpha_u = 1 \wedge \alpha_v = 0.$$

For the exact separation of shifted cutset inequalities with fixed a', we restate the objective (7.17a) as

$$\min \sum_{uv \in \mathcal{A}} \left(\sum_{p=1}^{P_{uv}} \sum_{q=1}^{Q_{uv}^p} \left\lceil \frac{B_{uv}^{pq} - a_{uv}}{a'} \right\rceil \tilde{y}_{uv}^{pq} \right) \beta_{uv} - z,$$

set

$$D := \frac{\sum\limits_{k \in \mathcal{K}} d^k \beta_{s^k t^k} - \sum\limits_{uv \in \mathcal{A}} a_{uv} \beta_{uv}}{a'},$$

and replace a in constraint (7.17b) by a'.

For an optimal solution $(\tilde{\alpha}, \tilde{\beta}, \tilde{z})$ of either separation problem (7.17) or the separation problem for shifted cutset inequalities, the subset $\mathcal{S}$ is defined by

$$\mathcal{S} := \{v \in \mathcal{V} \mid \tilde{\alpha}_v = 1\}$$

and the cutset

$$\mathcal{A}(\mathcal{S}, \overline{\mathcal{S}}) := \left\{ uv \in \mathcal{A} \,\middle|\, \tilde{\beta}_{uv} = 1 \right\}.$$

The right hand side of the violated inequality (7.14) or (7.16) that has to be added to the cut pool is given by $\tilde{z}$.

7.3.3. A Primal Heuristic for the Budget-Constrained Formulation

To find good solutions of (7.11) faster, we introduce the following heuristic which computes values for the decision variables y based on the current LP solution without modifying the flow variables x, see Algorithm 4.

Based on the flow values given in the LP solution $(\tilde{x}, \tilde{y})$, we compute the best bandwidth-configuration pair for each arc, i. e., the pair for which the flow is satisfied and the cost is as low as possible while the reliability is maximal. If the sum over all costs is lower than or equal to the budget, we have found a feasible solution. However, this cannot be guaranteed.

We experienced that the budget $\mathfrak{B}$ is not always used completely by the constructed solution. Hence, if there is some budget left, we attempt to improve the new solution by replacing bandwidth-configuration pairs with pairs having a higher reliability and still fulfilling the requirements. Note, we assume a non-decreasing ordering of the bandwidths and consider only larger bandwidths in the improvement step.

Algorithm 4 Primal Heuristic for (7.11)

Input: current LP solution $(\tilde{x}, \tilde{y})$
Output: new solution $(\tilde{x}, \hat{y})$ or ABORT

for $uv \in \mathcal{A}$ **do**
Compute left hand side of constraint (7.10c): $\text{lhs}_{uv} \coloneqq \sum_{k \in \mathcal{K}} \tilde{x}^k_{uv}$
Find best bandwidth-configuration pair fulfilling the demands with lowest cost and highest reliability:

$$(\hat{p}, \hat{q})_{uv} \coloneqq \underset{(p,q)}{\operatorname{argmin}} \left\{ c^p_{uv} \;\middle|\; B^{pq}_{uv} \geq \text{lhs}_{uv} \text{ and } \log(\rho^{pq}_{uv}) = \max_{(\tilde{p},\tilde{q})} \left\{ \log(\rho^{\tilde{p}\tilde{q}}_{uv}) \;\middle|\; c^p_{uv} = c^{\tilde{p}}_{uv} \right\} \right\}$$

Define minimum cost and maximum reliability: $\hat{c}_{uv} \coloneqq c^{\hat{p}}_{uv}$, $\hat{\rho}_{uv} \coloneqq \log(\rho^{\hat{p}\hat{q}}_{uv})$
Set new solution: $\hat{y}^{\hat{p}\hat{q}}_{uv} = 1$, $\hat{y}^{pq}_{uv} = 0 \; \forall \, (p,q) \neq (\hat{p}, \hat{q})_{uv}$
end for
if $\mathfrak{B} - \sum_{uv \in \mathcal{A}} \hat{c}_{uv} < 0$ **then** no solution found **return** ABORT
else if $\mathfrak{B} - \sum_{uv \in \mathcal{A}} \hat{c}_{uv} = 0$ **then** new solution found **return** $(\tilde{x}, \bar{y})$
else Try to improve the solution successively for every arc:
for $uv \in \mathcal{A}$ **do**
for $p > \hat{p}$ **do**
if $B^{pq}_{uv} \geq \text{lhs}_{uv}$, $\log(\rho^{pq}_{uv}) > \hat{\rho}_{uv}$ **and** $\sum_{\tilde{u}\tilde{v} \in \mathcal{A}} \hat{c}_{\tilde{u}\tilde{v}} - \hat{c}_{uv} + c^p_{uv} \leq \mathfrak{B}$ for one q **then**
Change new solution: $\hat{y}^{\hat{p}\hat{q}}_{uv} = 0$, $\hat{y}^{pq}_{uv} = 1$, set $\hat{c}_{uv} \coloneqq c^p_{uv}$, $\hat{\rho}_{uv} \coloneqq \log(\rho^{pq}_{uv})$
break for loop over p
end if
end for
end for
return $(\tilde{x}, \bar{y})$
end if

7.4. Computational Study

The number of scenarios to be considered in the big-M formulation (7.6) is at least

$$\#\text{configurations}^{\#\text{arcs}\cdot\#\text{bandwidths}}$$

in case that the bandwidth choices as well as the possible configurations are the same for each link. For the smallest realistic instance investigated in this computational study (the network Polska), the number of scenarios is $6^{36 \cdot 3}$. Hence, the big-M formulation is unmanageable for practical instances which is the reason for focussing on the case of independent link outages in the following computational study.

In this section, we first briefly present some preliminary results for formulation (7.10) concerning the *price of reliability* and the performance of the valid inequalities presented in Section 7.3.1 obtained for a 5×5 grid network. Afterwards, we describe the network

	7 MHz	28 MHz
cost	1000 $	6000 $
default config. (99.9 % availability)	128-QAM	256-QAM
config. for fading (100 % availability)	16-QAM	32-QAM

Table 7.2.: Considered bandwidth choices and configurations for the 5×5 grid network.

topologies and the configurations we used to generate realistic problem instances. For these instances, we present on the one hand results on the achievable reliability of the networks with the budget-constrained formulation compared to models without chance-constraints and on the other hand detailed results on the performance improvements discussed in Section 7.3.

7.4.1. Investigations for a Grid Network

To receive an impression for the reliability that can be achieved by formulation (7.10) and for the effectiveness of the (shifted) cutset inequalities, we perform simplified preliminary computational experiments on a 5×5 grid network based on [120]. The network comprises 25 nodes, 80 links, and 50 commodities. The considered bandwidth choices and configurations are displayed in Table 7.2. We assume that the availability of the lowest configuration is 100 %, thus, $\rho_{uv}^{p1} = 1$, and that the availability of the highest configuration is 99.9 %, hence, $\rho := \rho_{uv}^{p2} = 0.999 \, \forall uv \in \mathcal{A}, p = 1, \ldots, P_{uv}$. Including this assumption in constraint (7.10d) and applying some reformulations leads to

$$\sum_{uv \in \mathcal{A}} \sum_{p=1}^{P_{uv}} y_{uv}^{p2} \leq \left\lfloor \frac{\log(1-\varepsilon)}{\log(\rho)} \right\rfloor =: N.$$

Therefore, N is the maximum number of links which use the highest radio configuration. The higher the infeasibility tolerance ε, the larger is N and the lower the reliability of the network. In this computational study, we consider $N \in \{0, 10, \ldots, 80\}$.

Price of reliability The subsequent results are obtained on a Linux machine with a 3.20 GHz Intel Xeon W5580 CPU and 64 GB RAM, using IBM ILOG CPLEX 12.1 [98] as underlying solver. A time limit of two hours of computation is set and all other solver settings are preserved at their defaults. In Figure 7.2, we display the cost of the network as a function of N. For $N = 0$, the available capacity is not sufficient to route the total flow. Hence, the problem is infeasible. Moreover, the cost for $N = 10$ are 38.6 % higher than for $N = 80$. However, the results for $N = 60, 70, 80$ are the same as the reliability requirement is not restrictive due to the fact that the number of links which have to run at the highest configuration to satisfy the total traffic is less than or equal to 60. The reliability of the network can be improved without large effects on the cost until $N = 20, 30$. Summing up, the results indicate the trade-off between reliability and cost. For

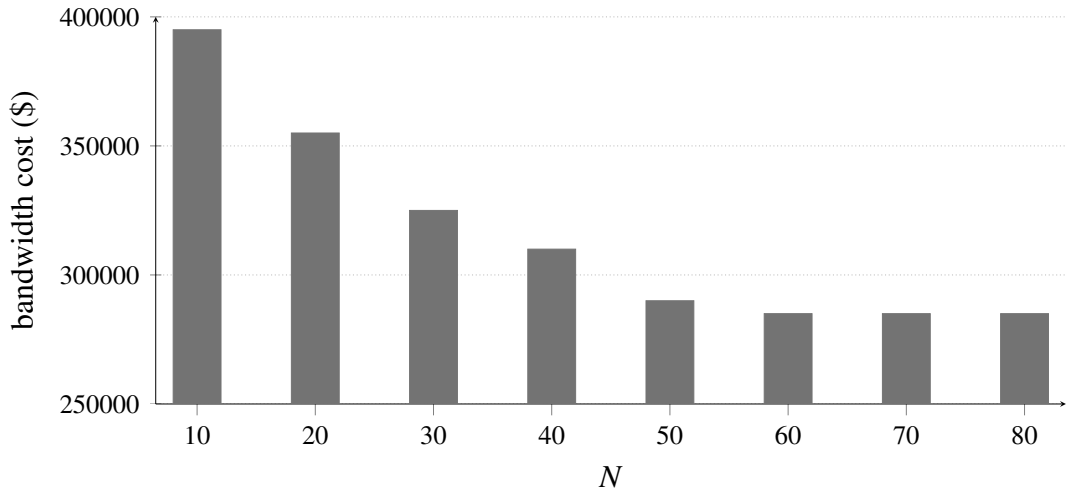


Figure 7.2.: Bandwidth cost as a function of N, the maximum number of links with highest configuration.

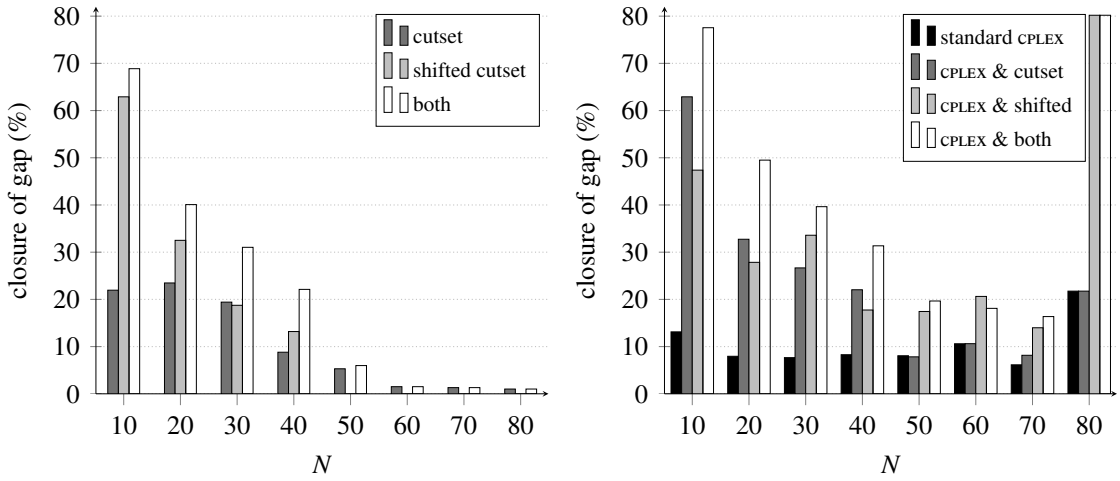
further details, we refer to our work [47].

Performance of (shifted) cutset inequalities The (shifted) cutset inequalities presented in Section 7.3.1 are separated only at the root node as a separation at every B&B node is too time consuming. We evaluate all three possible settings, the separation of i) only cutset inequalities, ii) only shifted cutset inequalities, iii) both types of inequalities. Additionally, we investigate the performance of these settings if the internal CPLEX cuts are also switched on. The results in Figure 7.3 were obtained on a Linux machine with a 2.67 GHz Intel Xeon X5650 CPU and 12 GB RAM, using IBM ILOG CPLEX 12.2 [98] as underlying solver.

For every setting and value of N, we compute the percentage gap closed at the root node defined as

$$\frac{\mathrm{DB}_{\mathrm{cut}} - \mathrm{DB}_{\mathrm{root}}}{\mathrm{PB}_{\mathrm{best}} - \mathrm{DB}_{\mathrm{root}}},$$

where $\mathrm{DB}_{\mathrm{root}}$ denotes the dual bound at the root node obtained by CPLEX, $\mathrm{DB}_{\mathrm{cut}}$ denotes the dual bound for i), ii), or iii), and $\mathrm{PB}_{\mathrm{best}}$ the best known primal bound computed after 12 h. The results are displayed in Figure 7.3 and the actual minimal, maximal, and average values over all N are listed in Table 7.3. Either with or without internal CPLEX cuts switched on, Figure 7.3 shows that the gap reduction decreases in general for increasing N. In most cases, the combination of both types of inequalities performs best where this statement is also supported by the numbers listed in Table 7.3. Additionally, the enabling of the internal CPLEX cuts usually supports the performance improvements gained by the separation of the (shifted) cutset inequalities. However, it is possible that the performance degrades when internal CPLEX cuts are active. Reasons for such a behaviour are given in the discussion in the final remarks on pages 233 to 235. Hence, the performance of the cuts can be better than of the shifted cuts (N = 10, 20, 40) and also the performance of the shifted cuts can



(a) (Shifted) Cutset inequalities without CPLEX cuts (b) (Shifted) cutset inequalities with CPLEX cuts

Figure 7.3.: Bandwidth cost as a function of the infeasibility tolerance.

	without internal CPLEX cuts			with internal CPLEX cuts			
	i)	ii)	iii)	CPLEX cuts	i)	ii)	iii)
min. (%)	1.01	0.0	1.0	6.1	7.8	14.0	16.4
max. (%)	23.5	62.9	68.9	21.8	62.9	80.2	80.2
average (%)	10.4	15.9	21.5	10.5	24.1	32.4	41.5

Table 7.3.: Minimal, maximal and average gap reduction values in % for the three settings either with or without internal CPLEX cuts enabled.

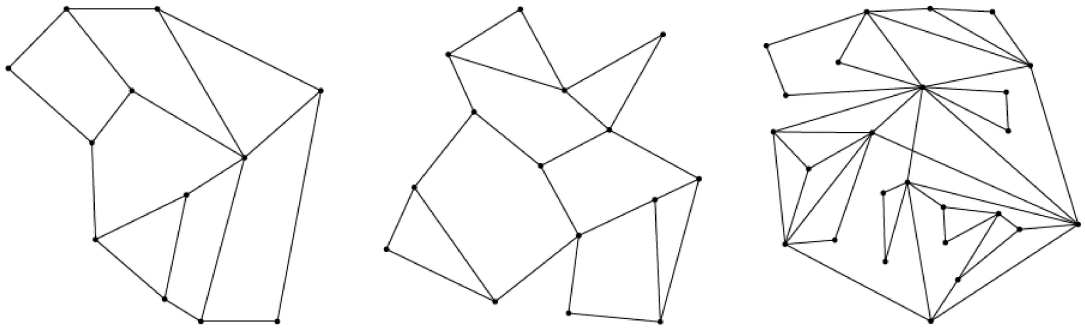
better than any other setting ($N = 60$). For further details on the computational results of this paragraph, we refer to [46].

7.4.2. Realistic Problem Instances

The grid network discussed in the previous section simplifies model (7.10) significantly. Hence, we investigate more realistic test instances in the remainder. Additionally, we focus on the budget-constrained formulation (7.11) henceforth as formulation (7.10) is infeasible for many problems which have insufficient reliability. Furthermore, we benefit from the primal heuristic presented in Section 7.3.3.

Given the absence of benchmark instances available in the literature for the studied problem, we generate test instances as follows. Network topologies and traffic demands are based on instances from a data library for fixed telecommunication network design, the Survivable Network Design Library (SNDlib) [149]. The studied network topologies are shown in Figure 7.4.

Since microwave links present limited capacity compared to optical fibre (as given in



(a) Polska (b) Atlanta (c) France

Figure 7.4.: SNDlib network topologies [149].

Network	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{A}\rvert$	$\lvert\mathcal{K}\rvert$	γ
Polska	12	36	66	0.2252
Atlanta	15	44	210	0.0170
France	25	90	300	0.0372

Table 7.4.: Summary of SNDlib problem instances.

the SNDlib [149]), we rescale the volumes of traffic demands by a factor γ to fit our application scenario. The scaling factor γ is obtained from the optimisation problem (7.18) and represents the maximum value for which there exists a feasible flow over the network assuming a network reliability of 99 % ($\varepsilon = 0.01$).

$$\max \gamma \tag{7.18a}$$

$$s.t. \sum_{u \in \delta^-(v)} x_{uv}^k - \sum_{u \in \delta^+(v)} x_{vu}^k = \begin{cases} -d^k\gamma, & \text{if } v = s^k, \\ d^k\gamma, & \text{if } v = t^k, \\ 0, & \text{otherwise} \end{cases} \qquad \forall\, v \in \mathcal{V},\ \forall\, k \in \mathcal{K} \tag{7.18b}$$

$$(7.10c), (7.10d), (7.10e), (7.10f)$$

$$\gamma \geq 0 \tag{7.18c}$$

Note, the proportionality of the original demands is preserved. The magnitudes of the studied SNDlib instances together with the corresponding scaling parameter are stated in Table 7.4.

We consider three bandwidth choices, 7 MHz, 14 MHz and 28 MHz based on a standard defined by the European Telecommunications Standards Institute [68], and six radio configurations for each link. The corresponding capacities B_{uv}^{pq} which are the product of bandwidth and bandwidth efficiency are presented in Table 7.5, where the values are based on confidential specifications of the WSL500 product by 3Roam [2]. The estimation of

radio configuration	e (bps/Hz)	B_{uv}^{pq} for 7 MHz (Mbps)	B_{uv}^{pq} for 14 MHz (Mbps)	B_{uv}^{pq} for 28 MHz (Mbps)
16-QAM coded	3.6	25.2	50.4	100.8
16-QAM uncoded	4.0	28.0	56.0	112.0
64-QAM coded	5.4	37.8	75.6	151.2
64-QAM uncoded	6.0	42.0	84.0	168.0
256-QAM coded	7.2	50.4	100.8	201.6
256-QAM uncoded	8.0	56.0	112.0	224.0

Table 7.5.: Radio configuration, bandwidth efficiency e, and capacity B_{uv}^{pq} for the considered bandwidth choices.

the probabilities ρ_{uv}^{pq} is obtained from the Vigants-Barnett fading model [13, 175]. More details can be found in [144] and [51].

Since spectrum pricing is usually a linear function of the amount of spectrum with which a license is associated, we set a cost of 1 \$ per 1 MHz of bandwidth and use bandwidth utilisation and cost interchangeably.

For each network, we detect a range of reasonable values for the budget $\mathfrak{B}$. We set the budget interval for Polska to $[644, 840]$, where 644 is the lowest possible value. For a budget $\mathfrak{B} < 644$, the problem is infeasible since we cannot install sufficient capacity on the arcs to serve the total demand. Beyond the budget of 840, the behaviour changes only very marginally; see Section 7.4.3. For Atlanta and France, similar arguments lead to the intervals $[749, 1057]$ and $[1414, 2002]$, respectively. Due to the possible bandwidth values of 7, 14 or 28 MHz, we consider budgets by a step of 7.

7.4.3. Reliability Analysis

In this subsection, based on the reliability of the network topologies, we compare the budget-constrained formulation (7.11) to two formulations without outage probability constraints of the form (7.3). Computations in this and the subsequent section are carried out on a Linux machine with a 3.40 GHz Intel i7-3770 CPU and 32 GB RAM, using IBM ILOG CPLEX 12.4 [98] as underlying solver. A time limit of two hours of computation is set for solving each instance, and all other solver settings are preserved at their defaults.

First, we consider (7.11) without AMC. Thus, only one radio configuration is available. For all three instances and bandwidth choices, the fixed radio configuration must be the highest one, 256-QAM uncoded as otherwise, the problems become infeasible. Thus, the model just selects the bandwidth for all links such that the total demand is fulfilled. We refer to this model as the *restricted budget-constrained formulation*. Note that by selecting a single radio configuration, the solution value is a lower bound on the network reliability of (7.11).

In a post-processing step, for every link uv and the bandwidth p chosen in the solution, we compute the lowest configuration q for which the capacity is sufficient. The actual

	(7.11)		restricted (7.11)		post-processing	
network	min	max	min	max	min	max
Polska	98.9	99.4	98.8	98.9	98.8	99.0
Atlanta	98.5	99.1	97.7	97.9	98.3	98.5
France	98.8*	99.3	97.8	97.9	97.8	98.1

Table 7.6.: The highest and lowest achieved network reliability values in % for the three network topologies and the budget-constrained formulation, the restricted budget-constrained formulation, and the post-processing. The value marked by $*$ is the lowest value that is computed.

network reliability is computed as the product of the corresponding link probabilities ρ_{uv}^{pq}. The post-processing adds AMC to the restricted budget-constrained formulation.

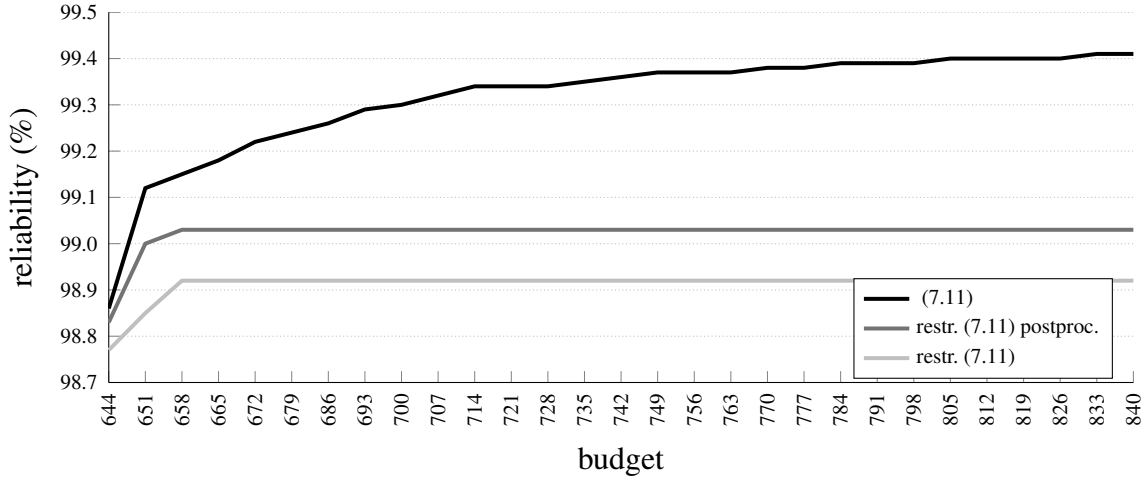
We display the reliabilities realised by the budget-constrained formulation (7.11), by the restricted budget-constrained formulation (lower bound), and by the post-processing for Polska with a budget interval $[644, 840]$, Atlanta ($\mathfrak{B} \in [749, 1057]$), and France ($\mathfrak{B} \in [1414, 2002]$) in Figure 7.5. For the latter two networks, we additionally display the dual bounds since some/all problems are not solved to optimality within the time limit.

The results for Polska and Atlanta are comparable whereas the use of AMC in the post-processing seems to be less restrictive. For networks of the size of France, the budget-constrained formulation (7.11) is harder to solve, in particular, for more restrictive budgets. The first feasible solution we find within the time limit is 98.78 % for a budget of 1470. The next solution is then computed for a budget $\mathfrak{B} = 1498$, which is why the corresponding curve starts at 1498 where the others start at 1414. Furthermore, many problems are not solved to optimality leading to fluctuating curves in Figure 7.5(c). In general, for higher budgets the solutions are very close to optimal.

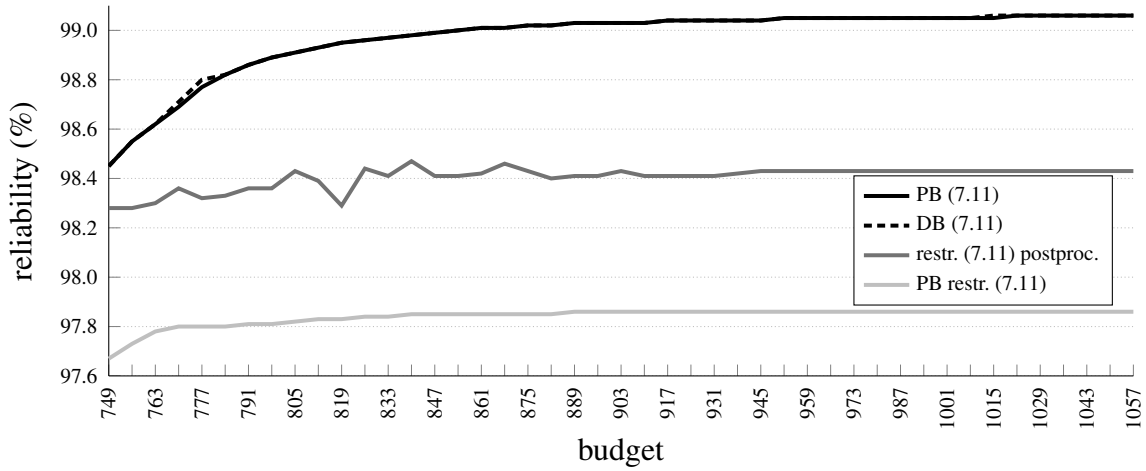
We summarise the results in Table 7.6 by giving the minimum and maximum reliability achieved by the three different approaches for the three network topologies. The minimum value for France and formulation (7.11) is the lowest one we can compute but there might exist a lower value for a problem with a lower budget where no solution is found within the time limit.

Altogether, the presented results illustrate the significant advantage of the chance-constrained model over the restricted (post-processed) budget-constrained formulation: we gain higher network reliabilities with reasonable computational effort for small to medium-sized networks.

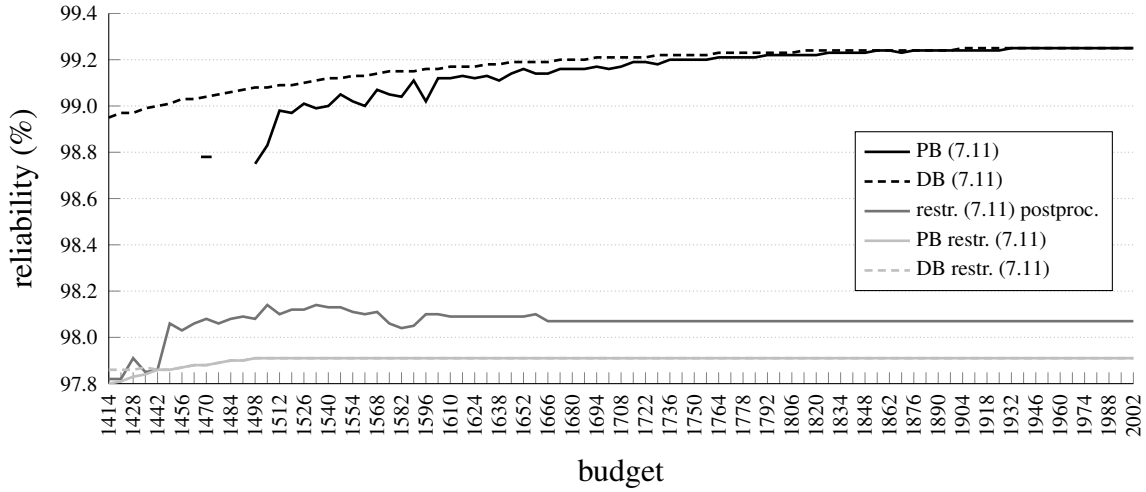
Since the achievable reliabilities when restricting (7.11) to one possible configuration might be too low in practice, engineers prefer a different strategy to configure the network. Instead of fixing a single radio configuration for the whole network, we select a configuration for every arc-bandwidth combination. A reliability of at least, e.g., 99 % can be achieved by requiring an uniform minimum probability for all links: $0.99^{\frac{1}{|A|}}$. Since the lowest configuration has typically a very high cumulative probability, this minimum probability is achievable for every arc and bandwidth choice. We select the highest con-



(a) Polska, $644 \leq \mathfrak{B} \leq 840$



(b) Atlanta, $749 \leq \mathfrak{B} \leq 1057$



(c) France, $1414 \leq \mathfrak{B} \leq 2002$

Figure 7.5.: Reliabilities for the three network topologies considering only one configuration with/without post-processing or (7.11) for different budgets.

figuration satisfying the minimum probability. This approach basically boils down to the model with individual chance constraints for all arcs; see Section 7.1. However, a disadvantage is that if we solve the restricted model (7.11), the problem becomes infeasible for any budget value. Thus, not all traffic can be routed in such a configuration. Similar to the choice of γ, all traffic should be scaled down. If we reduce the traffic requirements from 100 % by steps of 10 %, the first percentage resulting in feasible instances for all budgets is 70 % for Polska, and 60 % for Atlanta and France. The highest reliability for Polska with a network load of 70 % is 99.64 %, for Atlanta with 60 % network load it is 99.70 % and for France 99.69 %. Consequently, we exceed the required reliability of 99 % clearly but for the price of routing less traffic (only 70 % or 60 %, respectively). In contrast, the clear benefit of formulation (7.11) is that 100 % of the traffic can be routed with a higher reliability than the required 99 %.

7.4.4. Analysis of Valid Inequalities and Primal Heuristic

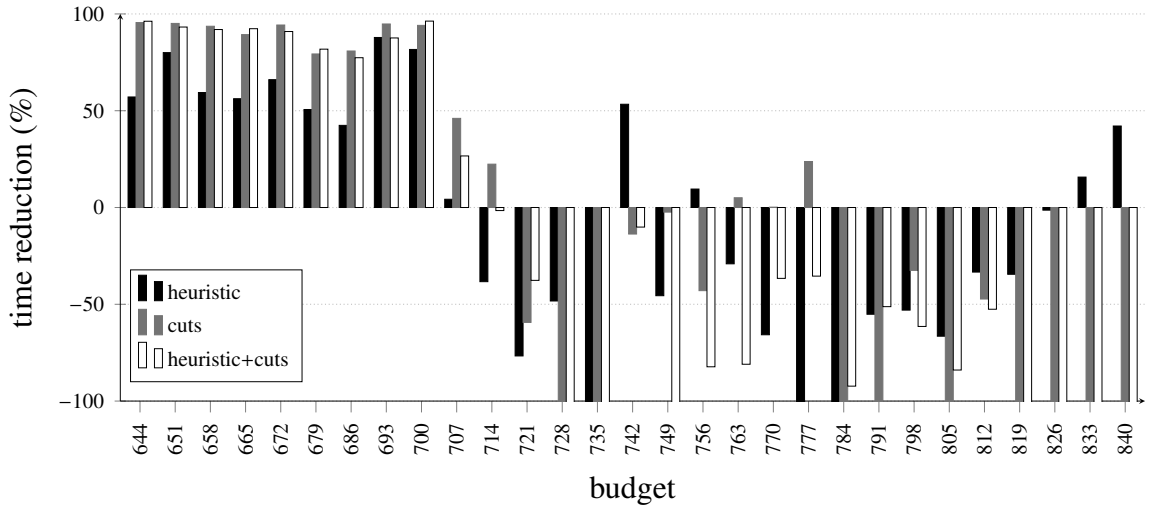
Most results in the previous subsection can only be achieved by applying the valid inequalities and primal heuristic presented in Section 7.3. We have demonstrated the effectiveness of the cutset inequalities at the root node for a 5×5 grid network in Section 7.4.1. Now, we study their performance, the performance of the primal heuristic, and the combination of both for the three realistic network topologies. Therefore, we consider four different settings: standard CPLEX, CPLEX and the primal heuristic, CPLEX and the valid inequalities, and CPLEX, the primal heuristic and the valid inequalities. As before, cutset inequalities are separated only at the root node of the B&B tree via the auxiliary ILPs presented in Section 7.3.1. Additionally, the primal heuristic is applied with a frequency of 20. This means that the heuristic is called in every 20th node of the B&B tree, where the node selection strategy is "best-bound" choosing the next node based on the best objective function.

For Polska and Atlanta, we investigate the time reduction for the different settings per budget as well as the absolute CPU times. We compute the time reduction as

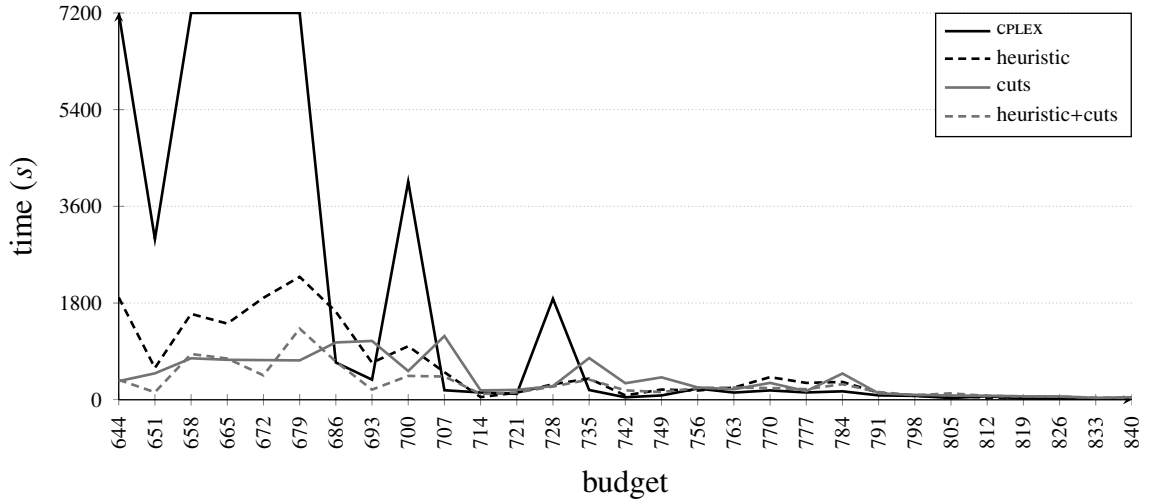
$$\frac{\text{CPLEX time} \;-\; \text{advanced time}}{\text{CPLEX time}},$$

where "advanced time" denotes the time consumed by one of the settings "heuristic", "cuts", or "heuristic+cuts". As an example, a value of 20 % means that we can reduce the solving time by 20 % due to the application of the cuts/the primal heuristic compared to the time needed when using standard CPLEX, while a value of -20 % says that we are 20 % slower than CPLEX. Note, if standard CPLEX exceeds the time limit, the computed time reduction is just a lower bound. Hence, the cuts and the primal heuristic can give a time reduction of at least the computed values if standard CPLEX reaches the time limit.

In Figures 7.6(a) and 7.7(a), we display the time reduction achieved for Polska and Atlanta. For better readability we set the lowest y-axis value to -100 %, where the highest time reduction possible is naturally $+100$ %. Due to this restriction, we cannot display the worst cases which exceed the value -100 %. Hence, we additionally display the absolute solving times for all four settings in Figures 7.6(b) and 7.7(b) demonstrating that a slow

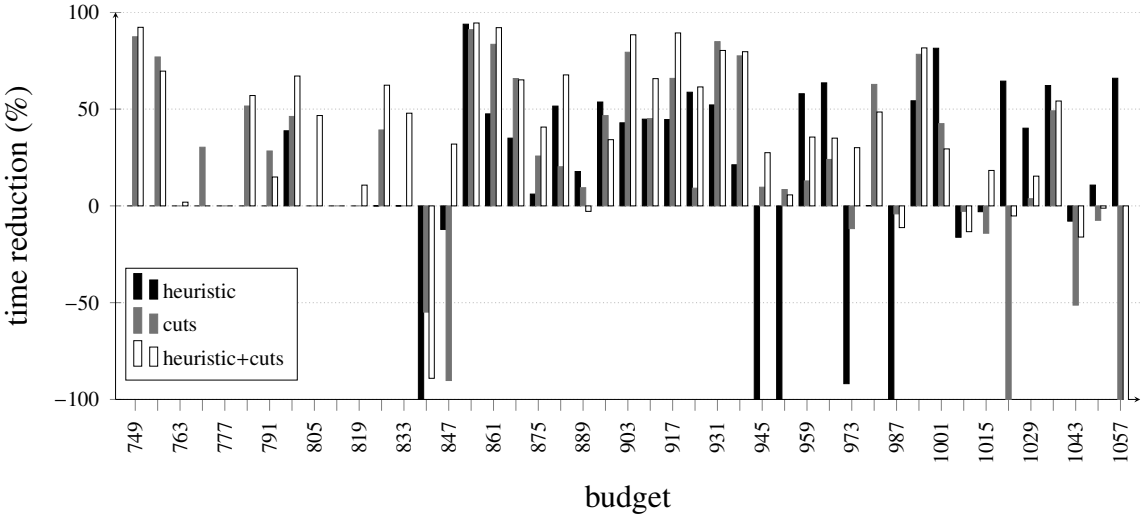


(a) Time reduction in %

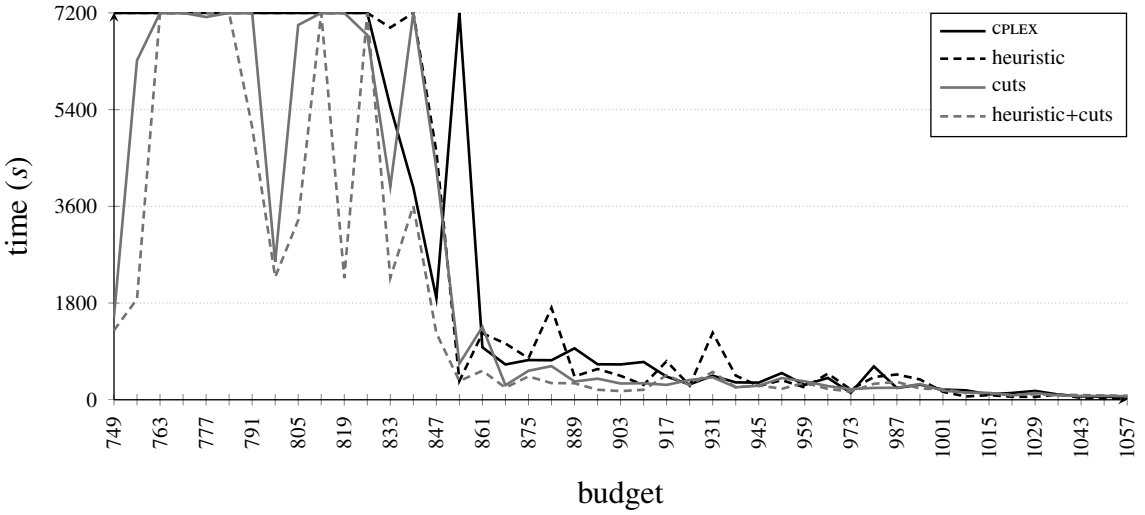


(b) Absolute times in s

Figure 7.6.: Reduction of computation times and absolute times for Polska considering different settings and budget values.



(a) Time reduction in %



(b) Absolute times in s

Figure 7.7.: Reduction of computation times and absolute times for Atlanta considering different settings and budget values.

down of 100 % or higher does not correspond to exorbitant absolute solving times.

The problems for lower budget values are naturally harder to solve. The problems for Polska with a budget $\mathfrak{B}$ between 644 and 700 excluding 651 ($\mathfrak{B} \in [749, 833]$ for Atlanta) are not solved within the time limit by standard CPLEX. For Polska, all of these problems are solved applying either the primal heuristic or the cutset inequalities while eleven out of thirteen problems are solved for Atlanta. The highest time reduction achieved for Polska (Atlanta) is 96.3 % (94.5 %) by means of cutset inequalities and the primal heuristic. For Polska and most budget values $\mathfrak{B} \geq 707$ standard CPLEX consumes significantly less time. This is why the inequalities and the primal heuristic cannot reduce the solving time for these problems in most cases. For a more detailed analysis in the case when all settings exceeded the time limit for Atlanta, we refer to our work [51].

Finally, we evaluate the results for France where we fix the budget $\mathfrak{B}$ to the interval $[1414, 2002]$. As no problem can be solved to optimality for neither setting, we consider the optimality gaps computed as $\frac{|\text{PB}-\text{DB}|}{\text{PB}}$ reached after two hours instead of the times. Hence, we compute the gap reduction as

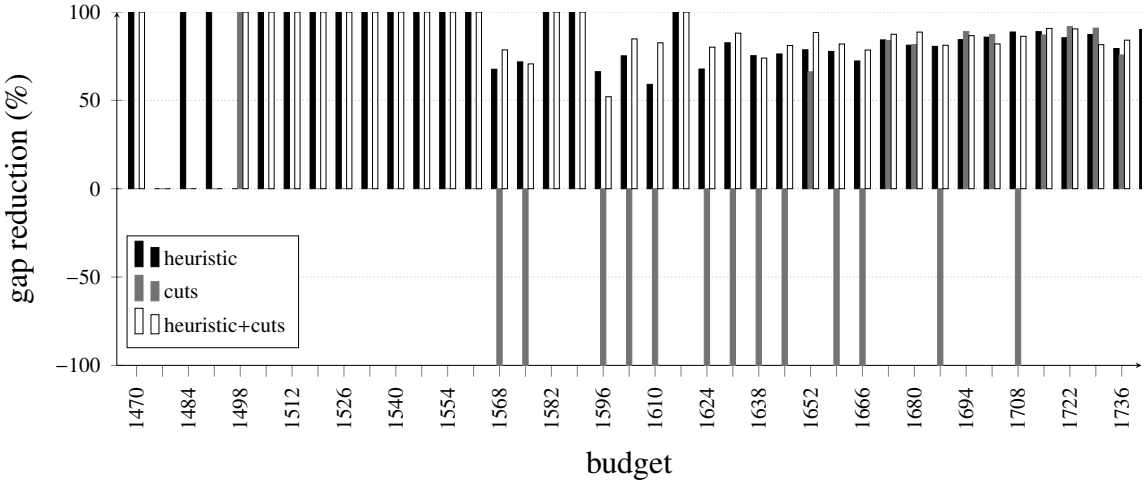
$$\frac{\text{CPLEX gap} \;-\; \text{advanced gap}}{\text{CPLEX gap}},$$

where "advanced gap" denotes the optimality gap reached by one of the settings "heuristic", "cuts", or "heuristic+cuts". Compared to the gap reduction defined in Section 7.4.1, we use the primal bound computed by the current setting (included in the gap) instead of the best known bound. If no primal bound is found, we set the gap to 100 %. Hence, the presented values are again the lower bounds.

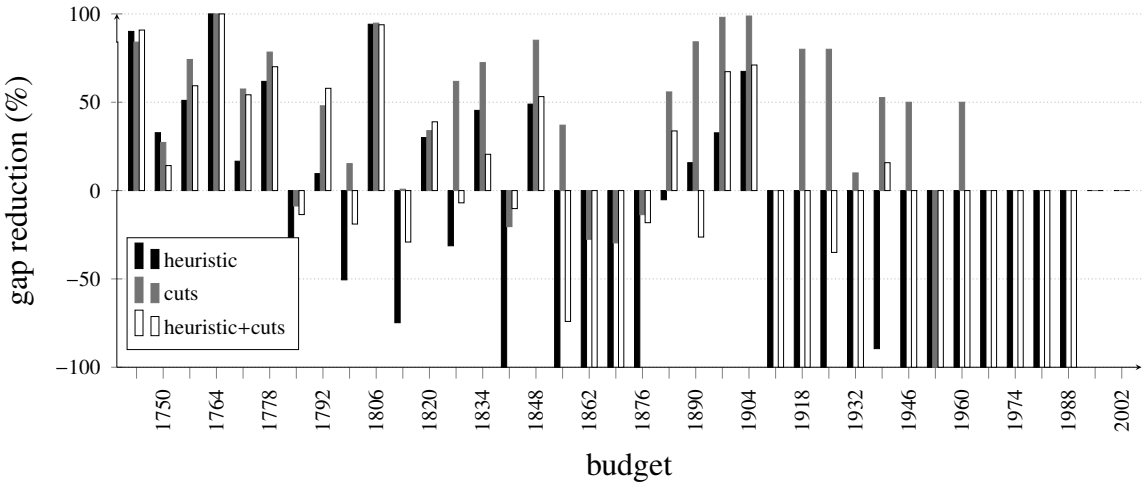
We display the gap reduction for France in Figures 7.8(a) and 7.8(b) for budget values in $[1470, 1736]$ and $[1743, 2002]$, respectively. Since not a single solution is found for $\mathfrak{B} \in [1414, 1463]$, we start with $\mathfrak{B} = 1470$ in Figure 7.8(a). For better readability, we once more scale from -100 % to 100 % although the gap can be increased by more than 100 %. For the first interval, negative gap reductions occur only for the setting "cuts" where no optimality gap can be computed within the time limit. For the second budget interval, we also give the absolute values of the optimality gaps in Figure 7.8(c). This figure demonstrates that an increase in the gap only occurs if the absolute values are quite low (≤ 4 %).

For almost all problems, the optimality gap is reduced significantly by the combination of both improvement techniques. Just for the easier problems with a budget greater than 1848 standard CPLEX computes already quite low gaps, which we cannot decrease. Since the cutset inequalities just improve the dual bound, a primal bound is usually found later. This is the reason why the gaps can be higher when only the valid inequalities are separated.

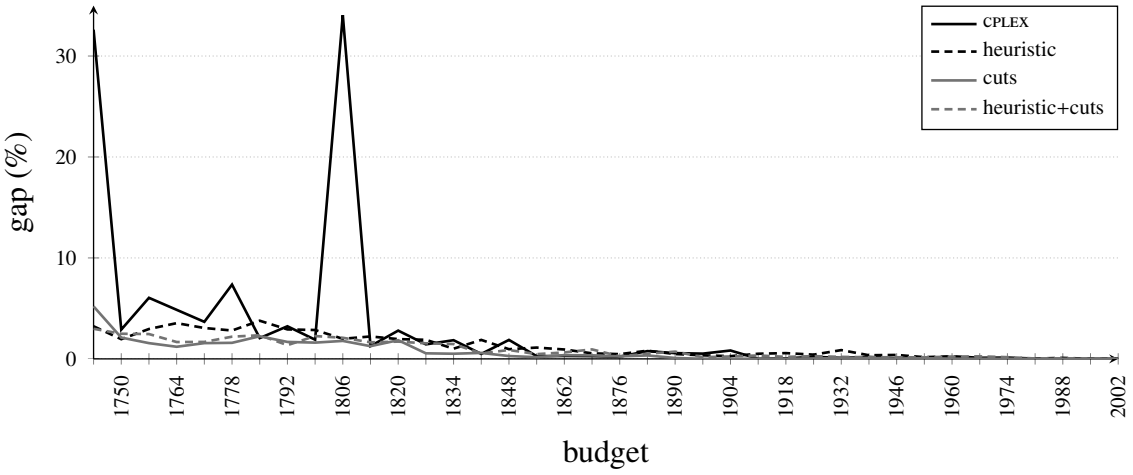
The gap reduction by the heuristic (with or without cutting planes) is dramatic, showing the importance of this relative simple idea. To understand its effectiveness, we revisit the primal heuristic once again briefly but from different perspectives. Table 7.7 displays all considered aspects and the used budget intervals for the three network topologies. Based on the intervals, the number of test instances differs per network. For Polska, the first



(a) Gap reduction in % for $1470 \leq \mathfrak{B} \leq 1736$



(b) Gap reduction in % for $1743 \leq \mathfrak{B} \leq 2002$



(c) Absolute gaps in % for $1743 \leq \mathfrak{B} \leq 2002$

Figure 7.8.: Reduction of optimality gaps and absolute gaps for France considering different settings and budget values in $[1743, 2002]$.

	Polska [644, 840]	Atlanta [749, 1057]	France [1414, 2002]
# instances	29	45	85
# inst. first sol. by heur. (in %)	23 (79.3 %)	42 (93.3 %)	75 (88.2 %)
# inst. no sol. found without heur.	–	11	26
# inst. no sol. found with heur. (in %)	–	2 (18.2 %)	10 (38.5 %)
absolute increase in primal bound	0.2 %	0.6 %	2.8 %
% of time used to find first sol. with heur.	45.0 %	13.6 %	23.7 %

Table 7.7.: Performance of primal heuristic with respect to different aspects.

solution found is computed by the primal heuristic in 23 out of 29 cases, which corresponds to 79.3 % of all instances. For Atlanta and France, this is the case for 42 out of 45 (93.3 %) instances and 75 out of 85 (88.2 %), respectively. Note, for the remaining instances of Atlanta and France, no primal solution is found at all within the time limit. The number of not solved instances is reduced from 11 to 2 for Atlanta and from 26 to 10 for France by the application of the presented primal heuristic. Hence, even for the more complex network topologies, the proposed heuristic is very effective in finding a feasible solution.

Moreover, the absolute values of the first primal solution found by the heuristic (given as a percentage) are usually larger than the values found without the heuristic if comparing only those cases where a primal solution is computed with both settings. On average, the reliability is increased by 0.2 % for Polska, 0.6 % for Atlanta and 2.8 % for France. Thus, the larger the network topology, the more the first primal bound found is improved by the primal heuristic.

Finally, regarding the computation times until the first solution is found, the primal heuristic uses on average only 45.0 %, 13.6 % and 23.7 % of the time spent without the heuristic for Polska, Atlanta and France, respectively. These speed-ups are implicitly also included in Figures 7.6 and 7.7 but not as considerably as in the numbers of Table 7.7 since the speed-up until the first solution is found is not necessarily conveyed to the end of the solution process.

In summary, the results of this section demonstrate the gains of both the valid inequalities and the primal heuristic, and especially of their combination.

7.5. Conclusion

In this chapter, we have applied the chance-constrained programming approach presented in Chapter 6 to tackle the problem of assigning bandwidths for reliable fixed point-to-point wireless networks under uncertain radio configurations. We have stated separated as well as joint chance constraints modelling the capacity requirements under uncertainty. Furthermore, we have introduced ILP formulations for the capacity planning problem in broadband wireless networks including a budget-constrained model. To improve the performance, we have introduced valid inequalities, their exact separation by ILPs, and a

primal heuristic. The computational studies have revealed a reduction of the solving times or optimality gaps for larger instances by means of the valid inequalities and the primal heuristic. Furthermore, we have investigated the reliability of various network topologies for different budget values and have compared the budget-constrained model to two alternative formulations which do not incorporate the joint outage probability constraint. The results show a significant gain in reliability by the joint probability model, though solving times increase.

Part III.

Γ-Robustness

8. General Concept

In this part and chapter, we study one special concept of robust optimisation, the so-called Γ-robustness. This approach limits the number of uncertain coefficients by a robustness parameter Γ, i. e., it uses a budgeted or cardinality constrained uncertainty set. It comprises the advantages of the linear approach by Soyster [167] but provides the possibility to control the level of conservatism. Γ-robustness has attracted a great deal of attention during the last decade due to its wide application spectrum and computational tractability.

In the first section, we give an introduction to this concept stating the Γ-robust uncertainty set, the robust counterpart, and commonly used measurements to evaluate a Γ-robust solution. In the subsequent section, we summarise the known results on probability bounds for the violation of a constraint subject to a Γ-robust uncertainty set. As an example, we briefly summarise known results on the Γ-robust knapsack problem in Section 8.5. We discuss a generalisation of Γ-robustness in detail in the subsequent part IV.

8.1. The Basic Principle

The most popular robust optimisation approach in the last years is Γ-*robustness* introduced by Bertsimas and Sim [18, 19] in 2003. Its popularity mainly arises from two results [19]: i) the robust counterpart of a LP with a Γ-robust uncertainty set remains computationally tractable and a LP (see Section 8.2), and ii) there exist bounds on the probability that a Γ-robust constraint is violated (see Section 8.4). Hence, Γ-robustness has been applied to a variety of optimisation and real-world problems such as resource allocation problems, logistics, telecommunication/network problems, railway planning and revenue management to name just some. In the following, we state the Γ-robustness concept formally.

Let $\Gamma \in [0, n]$ denote the robustness parameter which adjusts the robustness against the level of conservatism of the solution with $n \in \mathbb{N}$ denoting the number of coefficients which are subject to uncertainty. The following assumptions on the vector $u \in \mathbb{R}^n$ of uncertain data are necessary to define the concept of Γ-robustness.

- Each uncertain entry u_j of u can be modelled as an independent and bounded random variable.

- The distribution of each random variable need not to be known but symmetrical with nominal value $\bar{u}_j$. This together with the boundedness means for every realisation of u holds $u_j \in [\bar{u}_j - \hat{u}_j, \bar{u}_j + \hat{u}_j]$, $j \in \{1, \dots, n\}$, where $\hat{u}_j \geq 0$ denotes the largest possible deviation.

- For every realisation of the random vector u, up to $\lfloor \Gamma \rfloor$ many entries u_j can deviate simultaneously either to their minimum value $\bar{u}_j - \hat{u}_j$ or their peak value $\bar{u}_j + \hat{u}_j$. Additionally, one entry u_j is allowed to change by at most $(\Gamma - \lfloor \Gamma \rfloor)\hat{u}_j$.

The last assumption is quite intuitive in the following sense. It is most unlikely that all of the n entries u_j change at the same time. This assumption guarantees the feasibility of the obtained solution if less than Γ many uncertain entries deviate. In case that more than Γ maximal deviations occur, we present probabilistic guarantees that the robust solution is feasible with high probability in Section 8.4.

Note, in this thesis without loss of generality we restrict to $\Gamma \in \mathbb{N}$ for simplicity even though the Γ-robustness approach allows fractional values for Γ. Additionally, we assume that every entry u_j of the vector u is affected by uncertainty, thus $\hat{u}_j > 0$ for all $j \in \{1, \dots, n\}$.

In the following, we give a formal definition of the Γ-robust uncertainty set.

Definition 8.1. For a vector $u \in \mathbb{R}^n$ of uncertain data with $n \in \mathbb{N}$, let $\bar{u}$, $\hat{u} \in \mathbb{R}^n$ with $\hat{u} \geq 0$ denote the nominal and the deviation values, respectively, so that for every realisation holds $u_j \in [\bar{u}_j - \hat{u}_j, \bar{u}_j + \hat{u}_j]$. The Γ-*robust uncertainty set* $\mathcal{U}^\Gamma$ with $\Gamma \in \{0, 1, \dots, n\}$ is defined as

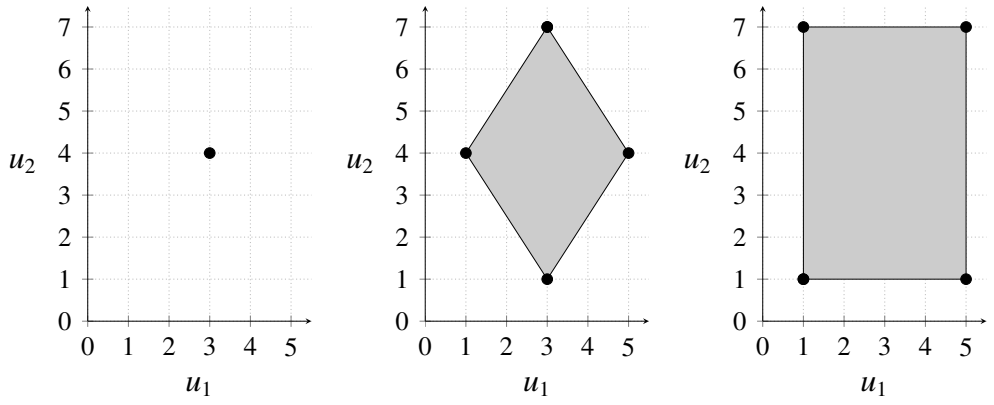
$$\mathcal{U}^\Gamma := \left\{ u \in \mathbb{R}^n \;\middle|\; \exists\, \delta \in [0,1]^n \,:\, \bar{u}_j - \delta_j \hat{u}_j \leq u_j \leq \bar{u}_j + \delta_j \hat{u}_j \;\forall\, j \in \{1, \dots, n\}, \sum_{j=1}^{n} \delta_j \leq \Gamma \right\}. \tag{8.1}$$

This definition allows more than Γ many simultaneous deviations as long as these deviations are not all maximal. Due to the convexity of $\mathcal{U}^\Gamma$, it is sufficient to consider only realisations with $\delta_j \in \{0, 1\}$, $j \in \{1, \dots, n\}$ and $\sum_{j=1}^{n} \delta_j = \Gamma$ ($\Gamma \in [0, n]$) since these points are the extreme points of the polytope. Note, for $\Gamma = n$ the uncertainty set $\mathcal{U}^\Gamma$ is equivalent to the uncertainty set defined by Soyster [167],which is the most conservative. An example of uncertainty sets for different values of Γ is displayed in Figure 8.1.

8.2. The Γ-robust Counterpart

In this section, we investigate the Γ-robust counterpart of a standard LP given in (1.1). Let a_i denote the i-th row of matrix A. The row by row representation of a standard LP is given by

$$\begin{aligned} &\max\ c^t x \\ &\text{s.t. } a_i x \leq b_i && \forall\, i \in \{1, \dots, m\} \\ &\quad x \in \mathbb{R}^n_{\geq 0}. \end{aligned} \tag{8.2}$$



(a) Γ = 0 (b) Γ = 1 (c) Γ = 2

Figure 8.1.: Γ-robust uncertainty sets for $\Gamma = 0, 1, 2$, nominal vector $\bar{u} = (3, 4)$ and deviation vector $\hat{u} = (2, 3)$, where the markers denote the extreme points.

The corresponding Γ-robust counterpart which consists of an exponential number of constraints is given as

$$\begin{aligned} &\max\ c^t x \\ &\text{s.t. } a_i x \le b_i && \forall\, i \in \{1, \dots, m\}, a_i \in \mathcal{U}_i^\Gamma \\ &\quad x \in \mathbb{R}^n_{\ge 0}, \end{aligned} \tag{8.3}$$

where

$$\mathcal{U}_i^\Gamma = \left\{ a_i \in \mathbb{R}^n \;\middle|\; \bar{a}_{ij} - \delta_j \hat{a}_{ij} \le a_{ij} \le \bar{a}_{ij} + \delta_j \hat{a}_{ij}, \sum_{j=1}^{n} \delta_j \le \Gamma, \delta_j \in \{0, 1\}\, \forall\, j \in \{1, \dots, n\} \right\}.$$

In [72], the authors present a dynamic cut generation scheme to solve this exponential sized formulation. In this approach, not all constraints of (8.3) are included in the formulation at the beginning, but violated constraints are separated on the fly as robustness cuts. Computational results show that this approach is promising for LPs but the computational tractability for ILPs cannot be ensured. These difficulties for ILPs are also demonstrated by Koster et al. [117].

A further approach which avoids the exponential nature of formulation (8.3), is to include only the most restrictive constraints as follows.

$$\begin{aligned} &\max\ c^t x \\ &\text{s.t. } \sum_{j=1}^{n} \bar{a}_{ij} x_j + \max_{S \subseteq \{1,\dots,n\}, |S| \le \Gamma} \sum_{j \in S} \hat{a}_{ij} x_j \le b_i && \forall\, i \in \{1, \dots, m\} \\ &\quad x \in \mathbb{R}^n_{\ge 0}. \end{aligned} \tag{8.4}$$

In this model, we consider only the worst, which are the highest deviation values, as they increase the left hand side most and hence, lead to the most restrictive constraints. Thus, negative deviations are discarded. All other constraints in (8.3) are dominated by those stated in (8.4). However, this model is non-linear. One possibility to linearise formulation (8.4) is to replace each constraint for $i \in \{1, \dots, m\}$ by the following potentially exponential many constraints.

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j + \sum_{j \in S} \hat{a}_{ij} x_j \leq b_i \qquad \forall\, S \subseteq \{1, \dots, n\} \text{ with } |S| \leq \Gamma \tag{8.5}$$

An alternative linear reformulation of (8.4) without exponential many constraints can be derived as follows. We consider the max-term for fixed variables x_j and row i separately. I. e.,

$$\max_{S \subseteq \{1,\dots,n\}, |S| \leq \Gamma} \sum_{j \in S} \hat{a}_{ij} x_j = \qquad \max \sum_{j=1}^{n} \hat{a}_{ij} x_j z_{ij} \tag{8.6a}$$

$$\text{s.t.} \sum_{j=1}^{n} z_{ij} \leq \Gamma \tag{8.6b}$$

$$z_{ij} \in \{0, 1\} \qquad \forall\, j \in \{1, \dots, n\}, \tag{8.6c}$$

where $z_{ij} = 1$ if and only if the j-th entry a_{ij} in row i takes its peak value $\bar{a}_{ij} + \hat{a}_{ij}$, which means that its deviation value $\hat{a}_{ij}$ is included in the max-term. Constraint (8.6b) guarantees that at most Γ many entries can deviate simultaneously. The coefficient matrix formed by this constraint and the integrality constraints (8.6c) is totally unimodular. Thus, the polytope of the feasible solutions of the LP relaxation has only integer vertices and we can relax constraints (8.6c) obtaining a LP. By strong duality (Theorem 1.3), the objectives of the LP and its dual coincide, where the dual problem is defined as follows.

$$\min \Gamma \pi_i + \sum_{j=1}^{n} \rho_{ij} \tag{8.7a}$$

$$\text{s.t. } \pi_i + \rho_{ij} \geq \hat{a}_{ij} x_j \qquad \forall\, j \in \{1, \dots, n\} \tag{8.7b}$$

$$\pi_i, \rho_{ij} \geq 0 \qquad \forall\, j \in \{1, \dots, n\} \tag{8.7c}$$

Variable π_i is the dual variable to constraint (8.6b) while variables ρ_{ij} are the duals to the upper bound of the relaxed version of constraints (8.6c), i. e., $z_{ij} \leq 1$. Now, we replace the max-term in (8.4) by (8.7) and relax the minimum. This last relaxation is possible as every feasible solution for the relaxed constraint still fulfils it when including only the lowest values. In summary, the compact Γ-robust counterpart of (8.2) reads

$$\max\ c^t x \tag{8.8a}$$

$$\text{s.t.} \sum_{j=1}^{n} \bar{a}_{ij} x_j + \Gamma \pi_i + \sum_{j=1}^{n} \rho_{ij} \leq b_i \qquad \forall\, i \in \{1,\dots,m\} \tag{8.8b}$$

$$\pi_i + \rho_{ij} \geq \hat{a}_{ij} x_j \qquad \forall\, i \in \{1,\dots,m\},\ j \in \{1,\dots,n\} \tag{8.8c}$$

$$x_j, \pi_i, \rho_{ij} \geq 0 \qquad \forall\, i \in \{1,\dots,m\},\ j \in \{1,\dots,n\}. \tag{8.8d}$$

This *compact* formulation has $m(n+1)$ additional variables and $n \cdot m$ additional constraints. Thus, its size is polynomial in the size of the non-robust LP (8.2). This observation is the crucial characteristic of the Γ-robustness concept introduced in [18, 19]. Applying Γ-robustness to an (I)LP does not increase its theoretical complexity. However, it can impair the computational tractability of the problem. In case of ILPs, Fischetti and Monaci [72] show that the compact formulation typically performs better than their dynamic cut generation scheme for formulation (8.3). Further computational studies in [116, 117] also show that the compact formulation outperforms other approaches. The advantage of the cutting plane approach to solve the non-compact formulation (8.3) is that it can handle also non-compact nominal (non-robust) formulations occurring, e. g., in routing problems, and it is also applicable to other uncertainty sets.

Uncertain Objective Now, we briefly present a result for a special case of a 0-1-optimisation problem when the objective coefficients c are subject to uncertainty. We extend this result for multi-band robustness; see Part IV. The non-linear (and non-exponential) Γ-robust counterpart of such a problem reads

$$z^* := \min \sum_{j=1}^{n} \bar{c}_j x_j + \max_{S \subseteq \{1,\dots,n\}, |S| \leq \Gamma} \sum_{j \in S} \hat{c}_j x_j \tag{8.9a}$$

$$\text{s.t. } x \in X \subseteq \{0,1\}^n. \tag{8.9b}$$

Theorem 8.2. (Bertsimas and Sim [18]) *Problem* (8.9) *can be solved by solving* $n+1$ *nominal (non-robust) problems of the form*

$$G^l = \Gamma \hat{c}_l + \min \left(\sum_{j=1}^{n} \bar{c}_j x_j + \sum_{j=1}^{l} (\hat{c}_j - \hat{c}_l) x_j \right) \tag{8.10}$$

$$s.t.\ x \in X, \tag{8.11}$$

i. e.,

$$z^* = \min_{l=1,\dots,n+1} G^l. \tag{8.12}$$

We assume $\hat{c}_{n+1} = 0$. The main ideas of the proof are as follows. First, formulation (8.9) is linearised via the dualisation approach described before introducing dual variables π

and ρ_j. The equivalent to constraints (8.8c) for this problem then reads

$$\pi + \rho_j \geq \hat{c}_j x_j \quad \forall\, j \in \{1, \ldots, n\}.$$

Due to these constraints, an optimal solution of the compact Γ-robust counterpart satisfies

$$\rho_j = \max\left\{\hat{c}_j x_j - \pi, 0\right\} = \max\left\{\hat{c}_j - \pi, 0\right\} x_j,$$

where the last equality holds since $x_j \in \{0, 1\}$. This property is exploited and $\mathbb{R}_{\geq 0}$ is decomposed into intervals $[0, \hat{c}_n]$, $[\hat{c}_n, \hat{c}_{n-1}], \ldots, [\hat{c}_2, \hat{c}_1]$, $[\hat{c}_1, \infty)$ to find the optimal value of π. Note, we assume that the entries j are sorted non-increasingly regarding $\hat{c}_j$ for the decomposition.

8.3. Evaluation of Robustness

A commonly used measurement to evaluate a Γ-robust formulation is the so-called *price of robustness* which gives the trade-off between the probability of constraint violation and the effect to the objective function of the nominal problem [19]. We use a slightly modified definition based on the percentage change of the objective:

$$\text{PoR} := \frac{|z - z(\Gamma)|}{|z|}, \tag{8.13}$$

with z being the optimal objective value of the nominal problem and $z(\Gamma)$ denotes the optimal objective value for the Γ-robust model. Hence, $|z - z(\Gamma)|$ gives the deterioration of the optimal value required to guarantee robustness. We include the absolute value for a definition of PoR regardless of the objective sense.

A different approach to measure the gain of a Γ-robust solution is the protection of the robust optimal solution; see [34]. The *protection level* can be computed as follows.

$$\text{ProL} := \frac{|\{\sigma \in \Sigma \,|\, \Gamma\text{-robust solution is feasible in } \sigma\}|}{|\Sigma|}, \tag{8.14}$$

where Σ denotes the set of all possible or considered scenarios, where a scenario represents one realisation of the uncertain coefficients. Hence, ProL gives the percentage of protected scenarios for which the Γ-robust solution is feasible. Koster et al. [116] apply a less conservative variant of the (inverse of the) protection level to measure the gain of Γ-robustness, the so-called realised robustness. It gives the average percentage of violated constraints for all considered scenarios.

While the price of robustness is based only on the objective function value, the protection level considers the actual realisations of the uncertain data. The advantage of the Γ-robust approach over the nominal problem can hence be depicted better by the value ProL.

8.4. Probability Bounds for Constraint Violation

The second major achievement of the works [18, 19] is the investigation of the probabilities of constraint violation. In the following, we briefly cite the main results.

Theorem 8.3. (Bertsimas and Sim [19]) *Let $\tilde{x}$ be an optimal solution of the Γ-robust counterpart* (8.8)*. The probability that the i-th constraint is violated satisfies*

$$\mathbb{P}\{a_i\tilde{x} > b_i\} \leq \exp\left(-\frac{\Gamma^2}{2n}\right). \tag{8.15}$$

The probability bound (8.15) is independent of the solution $\tilde{x}$ but can be quite bad if the fraction $\Gamma^2/(2n)$ is small. A better bound is given in the following theorem.

Theorem 8.4. (Bertsimas and Sim [18, 19]) *Let $\tilde{x}$ be an optimal solution of the Γ-robust counterpart* (8.8)*. Then the following holds.*

a)

$$\mathbb{P}\{a_i\tilde{x} > b_i\} \leq B(n,\Gamma) = \frac{1}{2^n}\left((1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu\sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right) \tag{8.16}$$

$$= \frac{1}{2^n}\left((1-\mu)\binom{n}{\lfloor\nu\rfloor} + \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right), \tag{8.17}$$

where $\nu = (\Gamma + n)/2$ and $\mu = \nu - \lfloor\nu\rfloor$.

b) Bound (8.16) *is tight.*

c) The bound (8.16) *satisfies*

$$B(n,\Gamma) \leq (1-\mu)C(n,\lfloor\nu\rfloor) + \sum_{l=\lfloor\nu\rfloor+1}^{n} C(n,l), \tag{8.18}$$

where

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

d) For $\Gamma = \theta\sqrt{n}$,

$$\lim_{n\to\infty} B(n,\Gamma) = 1 - \Phi(\theta), \tag{8.19}$$

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 a standard normal.

noitems	bound (8.15)	bounds (8.16),(8.18)	approximation (8.20)
5	5.0	5.0	5.0
10	9.6	8.2	8.4
100	30.3	24.3	24.3
200	42.9	33.9	33.9
2000	135.7	105.0	105.0

Table 8.1.: (Bertsimas and Sim [19]) Choice of Γ as a function of n so that the probability of constraint violation is less than 1 %.

The best possible bound (8.16) is computational difficult for large n. This is the reason why Bertsimas and Sim also derive bound (8.18) which is simple to compute and also "very tight" [19]. Moreover, they use the De Moivre-Laplace approximation to the binomial distribution to derive the following approximation of bound (8.19).

$$B(n, \Gamma) \approx 1 - \Phi\left(\frac{\Gamma - 1}{\sqrt{n}}\right) \tag{8.20}$$

For a comparison of these bounds, Table 8.1 states the choices of Γ for a selection of values of n such that the probability of a constraint violation is below 1 %. Obviously, using bounds (8.16) and (8.18) gives identical values for Γ. As bound (8.15) is usually worse, it also gives higher values for Γ. Additionally, (8.20) approximates bound (8.16) appropriately. In total, relatively small values for Γ are sufficient to guarantee a constraint satisfaction of 99 %.

8.5. The Γ-Robust Knapsack Problem

In this section, we extend the classical 0-1 KP introduced in Section 1.3 to the Γ-robust KP (Γ-RKP) using the same notation. We assume that the weights w are subject to uncertainty. Diverse aspects of this problem have been investigated by a variety of authors. We name just some basic works relevant for this thesis here. Bertsimas and Sim [18, 19] study the Γ-RKP as an example in their experimental analysis of the price of robustness and the computationally tractability of their newly introduced Γ-robustness approach. The Γ-RKP polyhedron is studied by Klopfenstein and Nace [111] including (extended) Γ-robust cover inequalities. These results are summarised and extended by a stronger class of extended Γ-robust cover inequalities by Kutschka [118]. Furthermore, Monaci et al. [141] present an exact solution algorithm using a dynamic program for the Γ-RKP which is based on the DP presented in Section 1.3.1 and has a running time of $O(\Gamma nB)$.

In the following, we summarise the known results which will be used and extended in the remainder of this part.

The Γ-RKP asks to select a subset of items $i \in N$ having an uncertain (positive) weight w_i with realisation in $[\bar{w}_i - \hat{w}_i, \bar{w}_i + \hat{w}_i]$ and a (positive) profit p_i such that a given capacity B

is not exceeded for any realisation of the weights and the total profit is maximized.

Definition 8.5. *(Γ-robust KP)* The Γ-*robust (0-1) KP* can be formalised as

$$\max \sum_{i \in N} p_i x_i \tag{8.21a}$$

$$\text{s.t.} \sum_{i \in N} \bar{w}_i x_i + \max_{S \subseteq N, |S| \leq \Gamma} \sum_{i \in S} \hat{w}_i x_i \leq B \tag{8.21b}$$

$$x_i \in \{0, 1\} \qquad \forall\, i \in N \tag{8.21c}$$

Introducing dual variables π and ρ_i to constraints (8.21b) and (8.21c), respectively, and following the dualisation technique described in Section 8.2 we obtain the following *compact robust counterpart* of the Γ-RKP (8.21).

$$\max \sum_{i \in N} p_i x_i \tag{8.22a}$$

$$\text{s.t.} \sum_{i \in N} \bar{w}_i x_i + \Gamma \pi + \sum_{i \in N} \rho_i \leq B \tag{8.22b}$$

$$\pi + \rho_i \geq \hat{w}_i x_i \qquad \forall\, i \in N \tag{8.22c}$$

$$\pi,\, \rho_i \geq 0 \qquad \forall\, i \in N \tag{8.22d}$$

$$x_i \in \{0, 1\} \qquad \forall\, i \in N. \tag{8.22e}$$

The Γ-robust knapsack polytope can be defined as follows; see [111].

Definition 8.6. *(Γ-Robust Knapsack Polytope)* The Γ-robust knapsack polytope P^{Γ}_{KP} is defined as the convex hull of the set of feasible solutions of the Γ-RKP (8.21):

$$P^{\Gamma}_{KP} := \operatorname{conv}\left\{ x \in \{0,1\}^n \,\middle|\, \forall\, S \subseteq N \text{ with } |S| \leq \Gamma \text{ holds } \sum_{i \in N} \bar{w}_i x_i + \sum_{i \in S} \hat{w}_i x_i \leq B \right\} \tag{8.23}$$

$$= \operatorname{conv}\{ x \in \{0,1\}^n \,|\, \exists\, \pi,\, \rho \text{ such that } x,\, \pi,\, \rho \text{ satisfy (8.22b) – (8.22d)} \} \tag{8.24}$$

The definition of the Γ-robust knapsack polytope uses the linearised version of constraint (8.21b) which is equal to constraints (8.22b)-(8.22d).

8.5.1. Γ-Robust Cover

The cover inequalities (1.14) introduced in Section 1.3 can be generalised to valid inequalities for the Γ-robust knapsack polytope as follows; see Klopfenstein and Nace [111].

Definition 8.7. *(Γ-Robust Cover)* A set $\tilde{C} \subseteq N$ is a Γ-*robust cover* if

$$\sum_{i \in \tilde{C}} \bar{w}_i + \max_{C' \subseteq \tilde{C} : |C'| \leq \Gamma} \sum_{i \in C'} \hat{w}_i > B.$$

It is *minimal* if for any item $i \in \tilde{C}$, $\tilde{C} \setminus \{i\}$ is not a robust cover.

Just like (non-robust) cover inequalities can be strengthened to extended cover inequalities (1.16) as explained in Section 1.3, Klopfenstein and Nace [111] extend Γ-robust cover inequalities as follows.

Definition 8.8. *(Extended Γ-Robust Cover)* For a Γ-robust cover $\tilde{C}$, the corresponding *extended Γ-robust cover* is defined as

$$E(\tilde{C}) := \tilde{C} \cup \begin{cases} \left\{ i \in N \setminus \tilde{C} \;\middle|\; \bar{w}_i + \hat{w}_i \geq \max\limits_{j \in \tilde{C}} \left(\bar{w}_j + \hat{w}_j \right) \right\}, & \text{if } |\tilde{C}| \leq \Gamma, \\ \left\{ i \in N \setminus \tilde{C} \;\middle|\; \bar{w}_i \geq \max\limits_{j \in \tilde{C}} \bar{w}_j,\ \bar{w}_i + \hat{w}_i \geq \max\limits_{j \in \tilde{C}} \left(\bar{w}_j + \hat{w}_j \right) \right\}, & \text{if } |\tilde{C}| > \Gamma. \end{cases} \tag{8.25}$$

If a Γ-robust cover $\tilde{C}$ can be partitioned into $\tilde{C} = C \cup J$ with $C \cap J = \emptyset$,

$$J = \operatorname*{argmax}_{C' \subseteq \tilde{C} : |C'| \leq \Gamma} \sum_{i \in C'} \hat{w}_i,$$

and $|C| \geq 0$, the extension $E(\tilde{C})$ can be strengthened; see Kutschka [118].

Definition 8.9. *(Strengthened Extended Γ-Robust Cover)* For a Γ-robust cover $\tilde{C} = C \cup J$, we define the corresponding *strengthened extended Γ-robust cover* as

$$E^+(C, J) := (C \cup J) \cup \left\{ i \in N \setminus (C \cup J) \;\middle|\; \bar{w}_i \geq \max_{j \in C} \bar{w}_j,\ \bar{w}_i + \hat{w}_i \geq \max_{j \in J} \left(\bar{w}_j + \hat{w}_j \right) \right\}. \tag{8.26}$$

Lemma 8.10. *(Klopfenstein and Nace [111], Kutschka [118]) Let $\tilde{C} = C \cup J \subseteq N$ be a robust cover with extension $E(\tilde{C})$ and strengthened extension $E^+(C, J)$. Then the* Γ-robust cover inequality

$$\sum_{i \in \tilde{C}} x_i \leq |\tilde{C}| - 1, \tag{8.27}$$

the extended Γ-robust cover inequality

$$\sum_{i \in E(\tilde{C})} x_i \leq |\tilde{C}| - 1, \tag{8.28}$$

and the strengthened extended Γ-robust cover inequality

$$\sum_{i \in E^+(C,J)} x_i \leq |\tilde{C}| - 1 \tag{8.29}$$

are valid for the Γ-robust knapsack polytope P^{Γ}_{KP}.

Moreover, Klopfenstein and Nace [111] state cases for which an extended robust cover inequality defines a facet of P^{Γ}_{KP}.

8.5.2. Separation of Γ-Robust Cover Inequalities

In this section, we briefly summarise an exact ILP based algorithm and a greedy heuristic by Klopfenstein and Nace [111] to separate violated Γ-robust cover inequalities and an exact ILP based algorithm by Kutschka [118] to separate violated strengthened extended Γ-robust cover inequalities.

Let x^{LP} be the solution of the LP relaxation of (8.22) with at least one $x_i^{\mathrm{LP}} \notin \{0, 1\}$. Again, we assume integer weights $\bar{w}_i$, $\hat{w}_i$ and $B \in \mathbb{Z}$ without loss of generality henceforth.

Define binary variables $\bar{y}_i$ and $\hat{y}_i$, where $\bar{y}_i = 1$ if and only if $i \in C \cup J$ and $\hat{y}_i = 1$ if and only if $i \in J$. The separation ILP introduced in [111] to separate a most violated Γ-robust cover inequality (8.27) then reads

$$\min \sum_{i \in N} (1 - x_i^{\mathrm{LP}}) \bar{y}_i \tag{8.30a}$$

$$\text{s.t.} \sum_{i \in N} (\bar{w}_i \bar{y}_i + \hat{w}_i \hat{y}_i) \geq B + 1 \tag{8.30b}$$

$$\sum_{i \in N} \hat{y}_i \leq \Gamma \tag{8.30c}$$

$$\hat{y}_i \leq \bar{y}_i \qquad \forall\, i \in N \tag{8.30d}$$

$$\bar{y}_i,\ \hat{y}_i \in \{0, 1\} \qquad \forall\, i \in N. \tag{8.30e}$$

Similar to the non-robust separation problem (1.15), the objective (8.30a) minimises the negative of the violation. This means, it maximises the violation of a potential robust cover inequality. If the objective value is strictly less than 1, we have found a violated Γ-robust cover inequality. Otherwise, the current LP solution x^{LP} satisfies all Γ-robust cover inequalities. Constraint (8.30b) ensures the cover condition while constraints (8.30c) and (8.30d) guarantee the Γ-robustness of the cover.

For an optimal solution $(\bar{y}^*, \hat{y}^*)$ of (8.30) with objective value < 1, the minimal Γ-robust cover corresponding to a most violated Γ-robust cover inequality is defined as

$$J := \{i \in N \mid \hat{y}_i^* = 1\}, \quad C := \{i \in N \mid \bar{y}_i^* = 1\} \setminus J.$$

The separation problem (8.30) reduces to a minimisation (surrogate) knapsack problem in case of $\Gamma = 0$. Since the separation of cover inequalities is NP-hard [107], (8.30) is also NP-hard. Klopfenstein and Nace [111] present a greedy heuristic which is an adaption of a well-known greedy heuristic for the classical KP; see Martello and Toth [135]. The main idea of the greedy heuristic is to first sort the items with respect to the ratio of profit to peak weight non-decreasingly. Up to Γ many items with the smallest ratios are selected to define the set J. As soon as the knapsack capacity is exceeded, this process is stopped. If $|J| = \Gamma$ and the capacity B is not exceeded, the remaining items are sorted non-decreasingly with respect to the ratio of profit to nominal weight. The set C is then filled with items until the capacity is exceeded; see Klopfenstein and Nace [111] for more details. We state a detailed slightly improved version of this greedy heuristic in Section 9.1.1.

Kutschka [118] presents the following exact ILP based algorithm so separate violated

strengthened extended cover inequalities (8.29). Define binary variables $\bar{y}_i$, $\hat{y}_i$ and α_i where $\bar{y}_i = 1$ if and only if $i \in C$, $\hat{y}_i = 1$ if and only if $i \in J$, and $\alpha_i = 1$ if and only if $i \in E^+(C, J) \setminus (C \cup J)$. Note, the slightly different meaning of $\bar{y}_i = 1$ here in comparison with ILP (8.30). The separation ILP then reads

$$\min \sum_{i \in N} (1 - x_i^{\mathrm{LP}})(\bar{y}_i + \hat{y}_i) - \sum_{i \in N} x_i^{\mathrm{LP}} \alpha_i \tag{8.31a}$$

$$\text{s.t.} \sum_{i \in N} \bar{w}_i \bar{y}_i + \sum_{i \in N} (\bar{w}_i + \hat{w}_i) \hat{y}_i \geq B + 1 \tag{8.31b}$$

$$\sum_{i \in N} \hat{y}_i \leq \Gamma \tag{8.31c}$$

$$\bar{y}_i + \hat{y}_i + \alpha_i \leq 1 \qquad \forall\, i \in N \tag{8.31d}$$

$$\bar{y}_i + \alpha_j \leq 1 \qquad \forall\, i,\ j \in N \text{ with } \bar{w}_j \geq \bar{w}_i \tag{8.31e}$$

$$\hat{y}_i + \alpha_j \leq 1 \qquad \forall\, i,\ j \in N \text{ with } \bar{w}_j + \hat{w}_j \geq \bar{w}_i + \hat{w}_i \tag{8.31f}$$

$$\bar{y}_i,\ \hat{y}_i,\ \alpha_i \in \{0, 1\} \qquad \forall\, i \in N. \tag{8.31g}$$

Again, the objective (8.31a) minimises the negative of the violation while a violated strengthened extended Γ-robust cover inequality is found if the objective value is strictly less than 1. Constraint (8.31b) ensures the cover condition whereas constraint (8.31c) guarantees that at most Γ many items are considered with their peak demand. Any item can either be in C or J or $E^+(C, J) \setminus (C \cup J)$ which is ensured by constraints (8.31d). Finally, constraints (8.31e) and (8.31e) formulate the requirements imposed on an item in the extension.

For an optimal solution $(\bar{y}^*, \hat{y}^*, \alpha^*)$ of (8.31) with objective value < 1 the strengthened extended Γ-robust cover is defined as $E^+(C, J) \coloneqq \{i \in N \mid \bar{y}_i^* + \hat{y}_i^* + \alpha_i^* = 1\}$ with $C \coloneqq \{i \in N \mid \bar{y}_i^* = 1\}$ and $J \coloneqq \{i \in N \mid \hat{y}_i^* = 1\}$.

9. Application to Wireless Networks

The Γ-robust optimisation approach presented in the previous chapter has been widely applied to a variety of problems. However, it is not as prevalently used for wireless communication networks as, e. g., for fixed (line) networks; see Altin et al. [6], Bertsimas et al. [21], Kutschka [118] to name just some publications. For wireless communication networks, Paschalidis and Wu [152] apply Γ-robustness to a routing problem occurring in wireless sensor networks. Such networks consist of autonomous sensors which control different factors such as environmental conditions and transmit the collected data through the network to a gateway. The authors of [152] apply Γ-robustness to uncertain energy consumption for packet transmission and reception. Adasme and Lisser [4] utilise Γ-robustness to model robust resource allocation in wireless networks which are based on OFDMA. In this work, the power consumption is subject to uncertainty but occurs only in the objective function. A further application of Γ-robustness is given by Parsaeefard and Sharafat [151] who study robust power control in cognitive radio networks with uncertain channel gains. Moreover, Zola et al. [181] investigate robust association for multi-radio devices where the download rate is subject to uncertainty due to the possibility to choose between different access networks.

In contrast, we apply the Γ-robustness approach to the planning of wireless cellular networks with uncertain traffic demands or uncertain spectral efficiencies in this chapter. First, we introduce the uncertainty of demands and derive a compact formulation of the robust WNPP based on demand uncertainty (d-RWNPP) in Section 9.1. Moreover, we recap two cutting plane approaches presented in Section 4.1 and introduce a further class of cutting planes, the robust (extended) cover inequalities, to improve the solving performance of the compact formulation of the d-RWNPP. The effectiveness of the cutting planes as well as the robustness regarding different aspects such as price of robustness, level of protection, and in comparison with conventional planning is analysed in a computational study. Subsequently in Section 9.2, we apply Γ-robustness to the WNPP when the spectral efficiencies, modelling interference, are subject to uncertainty (s-RWNPP). Again, we derive a compact formulation and perform a computational study to evaluate different possibilities of defining nominal and deviation values. Finally, we investigate the quality of interference modelling of s-RWNPP in comparison to the various formulations presented in Chapter 5.

Before introducing Γ-robustness to the WNPP, we first briefly restate the nominal formulation.

The Nominal Formulation of the WNPP The basic formulation of the WNPP introduced in Section 4.1 extended by the conflict graph concept discussed in Section 5.3 reads

as follows.

$$\min \sum_{s\in\mathcal{S}} c_s x_s + \lambda \sum_{t\in\mathcal{T}} u_t \tag{9.1a}$$

$$\text{s.t.} \sum_{s\in\mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{9.1b}$$

$$x_i + x_j \leq 1 \qquad ij \in \mathcal{E} \tag{9.1c}$$

$$\sum_{t\in\mathcal{T}_s} \frac{w_t}{e_{st}} z_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{9.1d}$$

$$x_s,\, z_{st},\, u_t \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\, (s,t) \in \mathcal{S} * \mathcal{T},\, t \in \mathcal{T} \tag{9.1e}$$

Note, we include the conflict graph in the basic formulation here to study the most general model which can handle BS candidate sites, which comprise different configurations in addition to the location. That is, by means of the conflict graph we can exclude the deployment of two BSs at the same position but with different configurations and at the same time limit the inter-cell interference; cf. Chapter 5.

9.1. The RWNPP with Demand Uncertainties

In this section, we focus on the uncertainty of TN demands in the WNPP and denote this problem by d-RWNPP. The presented formulations and improvements are based on our joint works [48, 49]. As TNs are aggregated users, uncertainties in the demands occur due to fluctuating bit rate requirements as well as due to the movement of users from one TN to another. Following the Γ-robustness approach discussed in Chapter 8, we model demand values as symmetric and bounded random variables w_t with realisations lying in the interval $[\bar{w}_t - \hat{w}_t, \bar{w}_t + \hat{w}_t]$, where $\bar{w}_t$ denotes a nominal value and $\hat{w}_t$ its highest deviation. We limit the number of simultaneous deviations by a robustness parameter $\Gamma \in \{0, \ldots, |\mathcal{T}|\}$. The (non-linear) robust counterpart of the capacity constraints (9.1d) is then given as follows; compare (8.4).

$$\sum_{t\in\mathcal{T}_s} \frac{\bar{w}_t}{e_{st}} z_{st} + \max_{\mathcal{T}'\subseteq\mathcal{T}_s, |\mathcal{T}'|\leq\Gamma} \sum_{t\in\mathcal{T}'} \frac{\hat{w}_t}{e_{st}} z_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S}. \tag{9.2}$$

As discussed in Section 8.2, the straightforward linearisation via the computation of all possible subsets $\mathcal{T}'$ can lead to an exponential number of constraints; cf. constraints (8.5). In [117], computations for a structurally comparable robust problem show that the exponential-sized formulation is outperformed by the compact reformulation which is derived via the dualisation technique described in detail in Section 8.2 and reads as follows.

min (9.1a)
s.t. (9.1b), (9.1c), (9.1e)

$$\sum_{t\in\mathcal{T}_s} \frac{\bar{w}_t}{e_{st}} z_{st} + \Gamma\pi_s + \sum_{t\in\mathcal{T}_s} \rho_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{9.3a}$$

$$\pi_s + \rho_{st} \geq \frac{\hat{w}_t}{e_{st}} z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{9.3b}$$

$$\pi_s,\, \rho_{st} \geq 0 \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T}. \tag{9.3c}$$

Compared to the nominal problem (9.1), the compact robust counterpart (9.3) has $|\mathcal{S}| + |\mathcal{S} * \mathcal{T}|$ additional variables and $|\mathcal{S} * \mathcal{T}|$ additional constraints.

9.1.1. Performance Improvements

The performance of a B&B algorithm can be improved by adding valid inequalities; cf. Section 1.2.2. Therefore, we study three types of cutting planes in the following whereby the first two have already been presented in Section 4.1 and Section 5.3, respectively.

Variable upper bound constraints One class of valid inequalities for the WNPP which are also valid for the d-RWNPP are the variable upper bound constraints (vub) (4.5)

$$z_{st} \leq x_s \quad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{9.4}$$

introduced in Section 4.1 and applied already in formulations for interference modelling in Chapter 5. We investigate the gain of adding these constraints explicitly in a computational study in Section 9.1.3.

Maximal clique inequalities A second class of cutting planes for the WNPP are the maximal clique inequalities (mci) introduced in Section 5.3 in the context of a conflict graph. This means, we can replace the conflict graph constraints (9.1c) by the mci

$$\sum_{s\in\mathcal{U}} x_s \leq 1 \quad \forall\, \mathcal{U} \subset \mathcal{S},\ \mathcal{U} \text{ is a maximal clique in } G = (\mathcal{S}, \mathcal{E}), \tag{9.5}$$

where all maximal cliques are computed by the Bron-Kerbosch algorithm [31] with complexity $O(3^{n/3})$ [169]. We also evaluate the effectiveness of these inequalities in the computational study in Section 9.1.3.

Robust cover inequalities The capacity constraints (9.1d) are knapsack constraints with a variable right hand side. Hence, the Γ-robust version (9.3a)-(9.3c) resembles Γ-robust knapsack constraints with variable right hand side. Rewriting (9.3a) for a fixed s as a general Γ-robust knapsack constraint with variable right hand side leads to

$$\sum_{i\in N} \bar{w}_i z_i + \max_{J\subseteq N : |J|\leq\Gamma} \sum_{i\in J} \hat{w}_i z_i \leq Bx, \tag{9.6}$$

where N denotes the set of items, $\bar{w}_i$ the nominal weight of item i, $\hat{w}_i$ its deviation, and B the knapsack capacity. For $x = 1$, (9.6) states a Γ-robust knapsack constraint as given

in (8.21b). Due to this fact and since $x = 0$ implies $z_i = 0$ for all $i \in N$, every valid inequality for the Γ-RKP can be adapted to a valid inequality for (9.6) by multiplying its right hand side with x.

In the following, we adapt the strengthened extended Γ-robust cover inequalities introduced in Section 8.5 to valid inequalities for a Γ-RKP with variable right hand side. For conciseness, we drop the word "strengthened" henceforth as we are only considering strengthened extended Γ-robust cover inequalities and do not need the distinction between extended and strengthened extended.

Definition 9.1. *(Γ-Robust Knapsack Polytope with Variable Right Hand Side)* The Γ-robust knapsack polytope Q^{Γ}_{KP} with variable right hand side is defined as the convex hull of the set of feasible solutions of the Γ-RKP where the capacity constraint (9.6) has a variable right hand side.

$$Q^{\Gamma}_{KP} \coloneqq \operatorname{conv}\left\{(z,x) \in \{0,1\}^n \times \{0,1\} \;\middle|\; \forall\, S \subseteq N,\ |S| \leq \Gamma \text{ holds } \sum_{i\in N} \bar{w}_i z_i + \sum_{i\in S} \hat{w}_i z_i \leq Bx\right\}$$

Corollary 9.2. *Let $\tilde{C} = C \cup J \subseteq N$ be a robust cover defined in Definition 8.7 with extension $E^+(C,J)$ as defined in Definition 8.9. Then the* Γ*-robust cover inequality*

$$\sum_{i\in C\cup J} z_i \leq (|C \cup J| - 1)x, \tag{9.7}$$

and the extended Γ*-robust cover inequality*

$$\sum_{i\in E^+(C,J)} z_i \leq (|C \cup J| - 1)x \tag{9.8}$$

are valid for the Γ*-robust knapsack polytope Q^{Γ}_{KP} with variable right hand side.*

This result is an immediate consequence from Lemma 8.10 as $x \in \{0,1\}$.

Separation of extended Γ-robust cover inequalities We do not compute all extended robust cover inequalities, as explained in Section 8.5.2. Instead, only violated inequalities are separated on the fly. For a non-integral LP solution $(z^{\mathrm{LP}}, x^{\mathrm{LP}})$ of (9.3), an exact ILP based separation problem is given as follows; cf. (8.30).

$$\min \sum_{i\in N} \left(x^{\mathrm{LP}} - z_i^{\mathrm{LP}}\right)(\bar{y}_i + \hat{y}_i) - \sum_{i\in N} z_i^{\mathrm{LP}} \alpha_i \tag{9.9a}$$

$$\text{s.t. } (8.31\text{b}) - (8.31\text{g}) \tag{9.9b}$$

The objective (9.9a) minimises the negative of the violation of a potential robust cover inequality where the extended robust cover is defined by the constraints of ILP (8.31). If the optimal objective value is strictly less than x^{LP}, the robust cover $C \cup J$ defines a violated inequality of type (9.8) with $C \coloneqq \{i \in N \mid \bar{y}_i = 1\}$, $J \coloneqq \{i \in N \mid \hat{y}_i = 1\}$, and extension $E^+(C,J) \coloneqq \{i \in N \mid \bar{y}^*_i + \hat{y}^*_i + \alpha^*_i = 1\}$.

Solving the exact separation problem (9.9) consumes a significant amount of time. Instead, we develop the following heuristic. First, we adapt and slightly improve the greedy heuristic by Klopfenstein and Nace [111] separating Γ-robust cover inequalities, which we stated briefly in Section 8.5.2, in the following Algorithm 5.

Algorithm 5 Heuristic for Separating Γ-Robust Cover Inequalities

Input: weights $\bar{w}_i$, $\hat{w}_i$ and current non-integral LP solution $(z^{\mathrm{LP}}, x^{\mathrm{LP}})$
Output: Γ-robust cover $C \cup J$ and violation v

Step 0: Set $v = 0$, $\bar{V} = 0$, $\hat{V} = 0$, $C = \emptyset$, $J = \emptyset$.
For all $i \in N$, let $\alpha_i = \frac{x^{\mathrm{LP}} - z_i^{\mathrm{LP}}}{\bar{w}_i + \hat{w}_i}$ and $\beta_i = \frac{x^{\mathrm{LP}} - z_i^{\mathrm{LP}}}{\bar{w}_i}$.

Step 1: Let $L(k)$ be the index of the k-th smallest coefficient in $\{\alpha_i\}_{i \in N}$.
for $k = 1$ to Γ **do**
$J \leftarrow J \cup \{L(k)\}$
$\bar{V} \leftarrow \bar{V} + \bar{w}_{L(k)}$
$\hat{V} \leftarrow \hat{V} + \hat{w}_{L(k)}$
$v \leftarrow v + \left(x^{\mathrm{LP}} - z^{\mathrm{LP}}_{L(k)}\right)$
if $\bar{V} + \hat{V} > B$ **and** $v < x^{\mathrm{LP}}$ **then**
STOP
end if
end for

Step 2: Let $L'(k)$ be the index of the k-th smallest coefficient in $\{\beta_i\}_{i \in N \setminus J}$.
for $k = 1$ to $n - \Gamma$ **do**
$\bar{V} \leftarrow \bar{V} + \bar{w}_{L'(k)}$
if $\hat{w}_{L'(k)} > \min_{i \in J} \hat{w}_i$ **then**
$j := \operatorname{argmin}_{i \in J} \hat{w}_i$
$C \leftarrow C \cup \{j\}$
$J \leftarrow J \setminus \{j\} \cup \{L'(k)\}$
$\hat{V} \leftarrow \hat{V} - \hat{w}_j + \hat{w}_{L'(k)}$
else
$C \leftarrow C \cup \{L'(k)\}$
end if
$v \leftarrow v + \left(x^{\mathrm{LP}} - z^{\mathrm{LP}}_{L'(k)}\right)$
if $\bar{V} + \hat{V} > B$ **and** $v < x^{\mathrm{LP}}$ **then**
STOP
end if
end for

In Step 0, we initialise the sets C and J and the parameters v, $\bar{V}$ and $\hat{V}$, where v presents the objective value of the separation problem, $\bar{V}$ the sum of the nominal weights $(\bar{w}_i)$ of all items in the Γ-robust cover $C \cup J$, and $\hat{V}$ presents the sum of the deviations $(\hat{w}_i)$ of the items in J. Moreover, we save the ratios of profit $x^{\mathrm{LP}} - z_i^{\mathrm{LP}}$ to peak weight $\bar{w}_i + \hat{w}_i$ in parameters α_i and the ratios of profit to nominal weight $\bar{w}_i$ in parameters β_i. Then the items are sorted

service	normal (%)	high (%)	bit rate (kbps)
data	[10,20]	[30,40]	[512,2000]
web	[20,40]	[40,50]	[128,512]
VoIP	remaining	remaining	64

Table 9.1.: Normal and high traffic profiles for calculation of TN demands

non-decreasingly with respect to α_i.

In Step 1, the set J is filled by at most Γ many items with the lowest α_i values which are considered with their peak weight. If the capacity B is exceeded and the objective value is less than x^{LP} at any point in this step, the algorithm stops and returns J. Otherwise, it continues with Step 2. In this step, the remaining items are first sorted non-decreasingly according to β_i. Afterwards, the set C is filled by the items with the lowest β_i values which are considered with their nominal weight. However, if the deviation value of the current item $L'(k)$, which should be included in C, is larger than the lowest deviation value of any item in J, we reshuffle sets J and C as follows. The item $j \in J$ with the lowest deviation value is removed from J and included in C while the current item $L'(k)$ is added to J. The parameter $\hat{V}$ is adapted accordingly. By this interchange routine, we can strengthen the cover $C \cup J$ as we might need less items to exceed the capacity B. As before, the algorithm stops as soon as the capacity is exceeded and the objective value is less than x^{LP}.

Algorithm 5 returns a Γ-robust cover $C \cup J$ and a corresponding (negative) violation v. If it holds $v < x^{\mathrm{LP}}$ for the final violation value v, the cover $C \cup J$ is strengthened according to (8.26) if possible leading to $E^+(C, J)$ and we add an extended Γ-robust cover inequality (8.26) with violation

$$-v + x^{\mathrm{LP}} + \sum_{i \in E^+(C,J) \setminus (C \cup J)} z_i^{\mathrm{LP}}.$$

9.1.2. Generating Test Instances

To evaluate the performance of the presented cutting planes as well as of the Γ-robustness applied to the WNPP, we use the scenarios 20_200a, 30_300a, 40_400a and 40_450a described in Section 5.9.1 and create three larger scenarios with 50 (60) BSs and 500 (600,1000) TNs analogously for a broad range of magnitudes of test instances. In total, the sizes of the considered scenarios range from 20 to 60 BSs and from 200 to 1000 TNs. The demand values computed regarding the traffic profiles given in Table 5.3 and repeated in Table 9.1 in column "normal" are now assumed to be the nominal demands $\bar{w}_t$. In addition, the peak demands $\bar{w}_t + \hat{w}_t$ are computed analogously but regarding higher traffic profiles as depicted in Table 9.1 in column "high". In case that $\bar{w}_t + \hat{w}_t < \bar{w}_t$, we interchange the two values. Furthermore as before, we set $e_{\min} = 0.25$ for the minimum required spectral efficiency. All remaining parameters are fixed as described in Section 5.9.1. In particular, the conflict graph is established via the minimum distance requirement with $d_{\min} = 500\,\mathrm{m}$; cf. Section 5.3.

scenario	min.	max.	average	Γ_{max}
s_20_200	13	34	25.20	26
s_30_300	19	42	32.97	34
s_40_400	21	48	36.85	38
s_40_450	21	50	39.33	40
s_50_500	20	53	41.16	42
s_60_600	20	57	44.38	46
s_60_1000	30	73	57.02	58

Table 9.2.: Minimum, maximum and average value of Γ_{max} computed via (8.17) with probability 1 % and actually selected values for Γ_{max}.

Due to the different dimensions of the test scenarios, the maximum value for the robustness parameter Γ also varies. To find a reasonable value Γ_{max} per scenario, we use the probability bound (8.17) assuming a probability of constraint violation of 1 %. This bound is computed for every BS s with $n = |\mathcal{T}_s|$. In Table 9.2, we display the minimum, maximum and average Γ_{max} value per scenario. To limit the number of problems to be solved, we consider only even values for Γ. We set Γ_{max} to the next larger even number of the average value depicted in Table 9.2. We select Γ_{max} according to the average value to be secured against constraint violation on average. If we chose the maximum value, the probability of constraint violation would be less than 1 % for every BS, but Γ_{max} would also be unjustifiably large for many BSs. Moreover, we will see in Section 9.1.5 that setting Γ_{max} to the average value is sufficient. In total, we consider $\Gamma \in \{0, 2, \ldots, \Gamma_{max}\}$ for each scenario.

All computations in the following sections are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor using CPLEX 12.4. Additionally, we set a general CPU time limit of two hours and a memory limit of 11 GB.

9.1.3. Analysis of Improvements

Apart from slightly different scenarios, the computational study performed in this section to evaluate the improvements proposed in Section 9.1.1 differs from the computational study presented in our work [49] in the number of threads that CPLEX is allowed to use. In [49], we limited the number of threads to one "to obtain comparable results" as former versions of the solver CPLEX were not able to use multiple threads as soon as user defined inequalities were separated. In contrast, we do not limit the number of threads in the present computational study to be able to evaluate the improvements we actually gain when using the highest available computing power. Thus, the following evaluation is performed from a more practical point of view than in [49].

To analyse the performance of the cutting planes presented in Section 9.1.1, we compute the additional gap closed at the root node for four different settings: vub only ("vub"), mci only ("mci"), extended robust cover inequalities only ("c"), all three types together ("all").

The additional gap closed is given by

$$\frac{\mathrm{DB}_{\mathrm{cut}} - \mathrm{DB}_{\mathrm{root}}}{\mathrm{PB}_{\mathrm{best}} - \mathrm{DB}_{\mathrm{root}}},$$

with $\mathrm{PB}_{\mathrm{best}}$ being the best known primal bound, $\mathrm{DB}_{\mathrm{root}}$ the dual bound at the root node computed without any cuts (also no internal CPLEX cuts), which is the LP relaxation, and $\mathrm{DB}_{\mathrm{cut}}$ being the dual bound at the root node computed when the investigated cutting planes are applied; see also Section 7.4.1.

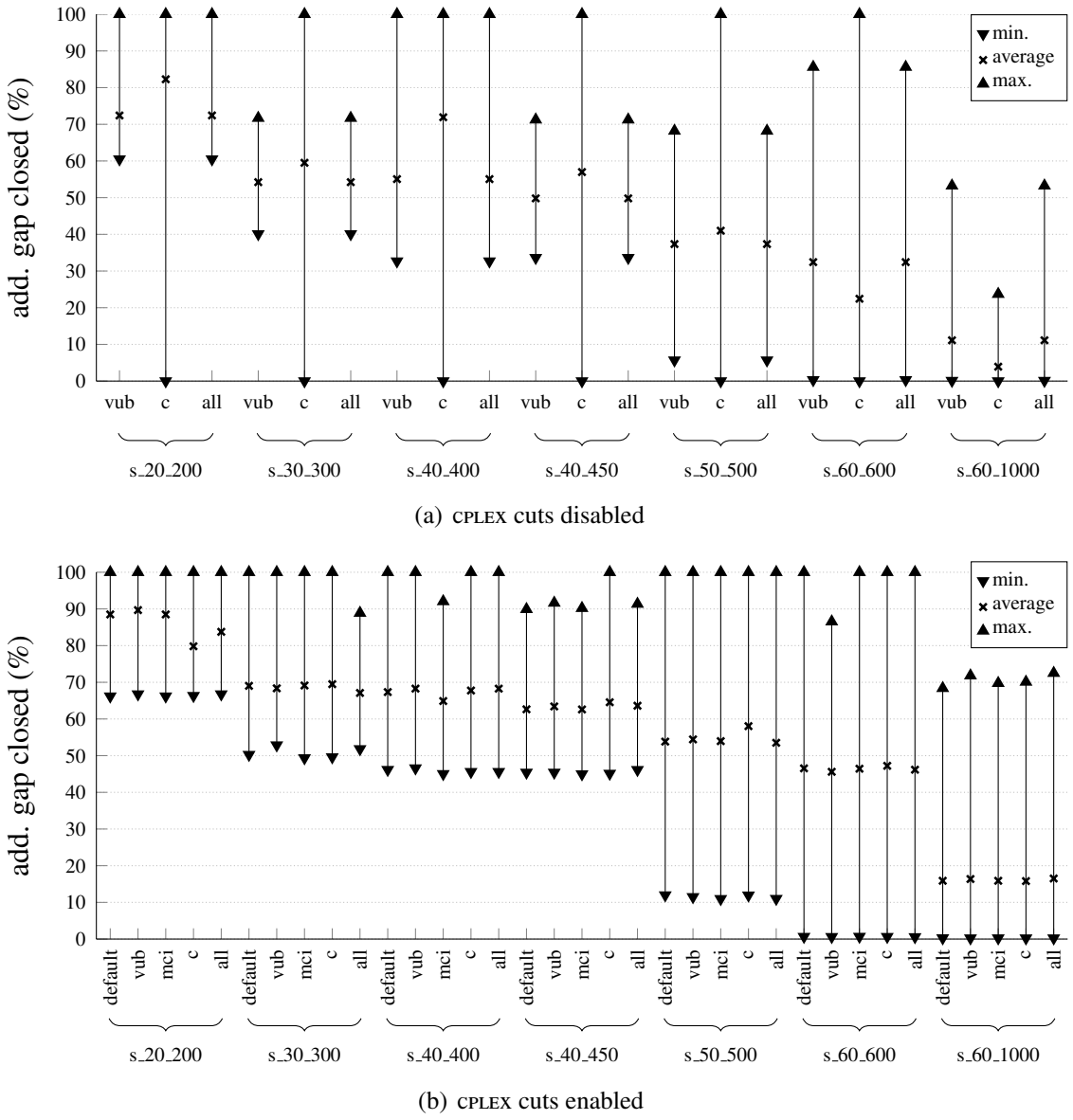
In Figure 9.1, we display the minimum, average (over Γ) and maximum additional gap closed for the studied cutting planes either with internal CPLEX cuts disabled (Figure 9.1(a)) or enabled (Figure 9.1(b)).

For disabled CPLEX cuts, the additional gap closed for the mci is always 0 %, which is why we do not depict these results in Figure 9.1(a). A reason for this behaviour, which differs from earlier results in our work [49], is that the capacity of BSs with a positive LP solution value is used only sparsely in the investigated scenarios. Therefore, TNs can be fully assigned to a BS ($z_{st}^{\mathrm{LP}} = 1$) even though $x_s^{\mathrm{LP}} << 1$. This causes the mci to be less or even non-effective. Note that we only investigated the performance of the mcis in combination with the vub constraints in [49].

For scenario s_40_450 and both scenarios with 60 BSs, some instances with the setting "c" are stopped by the operating system due to exceeding the available memory. This is possible as many cover inequalities are added but the memory limit is not checked at the root node. Hence, the additional gaps closed obtained for these instances are only lower bounds. In general, we observe that the average gap closed achieved for smaller scenarios is higher than for larger when CPLEX cuts are disabled. For the setting "c", the additional gap closed ranges from 0 % to 100 % while the average lies between 41 and 82 % for scenarios with up to 50 BSs. Such a high fluctuation occurs since a large number of cover inequalities is generated but with a heuristic separation routine. Hence, for some instances, strong inequalities are found while this is not the case for others impairing the solving process by a large number of additional constraints. Additionally, the heuristic separation is also the reason why the setting "all" does not always perform as good as the best among the two other settings. When all cutting plane approaches are combined, significantly less cover inequalities are added but there is no guarantee that all of these inequalities have been found.

In summary, the separation of cover inequalities performs best on average for scenarios with up to 50 BSs while "all" and "vub" yield the best average gap closed for the two largest scenarios in case that internal CPLEX cuts are disabled.

By the setting "default", we denote CPLEX with its default settings including its internal cuts. The additional gap closed when these cuts are enabled are quite similar to all considered settings; cf. Figure 9.1(b), indicating that CPLEX closes most of the gap. However, for scenarios 40_450 and s_60_1000 the activation of all cutting planes yields a slightly higher gap closed (minimum, average and maximum) than "default". Again, the setting "mci" has hardly any positive effect on the solving of the root node. Nevertheless, as the mci replace the distance constraints (9.1c) and the conflict graph is sparse for the investigated



(a) CPLEX cuts disabled

(b) CPLEX cuts enabled

Figure 9.1.: Minimum, average and maximum additional gap closed for different settings and scenarios when CPLEX cuts are dis- or enabled.

scenarios, the number of constraints is not increased. Therefore, the solving process is also not deteriorated by these types of cutting planes.

Summing up, the performed computational study indicates that the vub constraints, the extended robust cover inequalities as well as the combination of all cutting planes can strengthen the compact formulation (9.3) while the mci hardly improve the performance. For enabled CPLEX cuts, most of the gap is closed by the internal cuts. Nevertheless, it should not be neglected that the adding of further constraints causes a different structure of the B&B tree in CPLEX. Hence, even if the combination of all valid inequalities could not

close the gap at the root node, the consecutive solving process might still be improved; see also the numerical results in our work [49]. Additionally, we would like to point out that all presented results have to be treated with caution as the performance of CPLEX can be significantly influenced by adding (redundant) constraints; see the discussion in the final remarks on pages 233 to 235.

Based on the results above and those achieved in our work [49], we decide to use the setting "all" in the following section when continuing the solving process after the root node.

9.1.4. The Price of Robustness

In this section, we investigate the solutions (primal and dual bounds) computed within four hours for the considered values of Γ and the different scenarios. In Figure 9.2, we display these primal (solid lines) and dual (dotted lines) bounds. For instances which are solved to optimality, the two lines coincide. For s_60_1000 and $\Gamma \geq 20$, the primal bounds found within the time limit are quite close to the first solution where no BS is installed and no TN is served (objective value 1000). Hence, the dual bound can only give an indication of the magnitude of the optimal objective value. Moreover, for scenario s_60_600 and $\Gamma = 44$, the best primal bound found within the time limit is 498. For a better readability, we do not display these high numbers for s_60_600 and s_60_1000 in Figure 9.2 by setting the maximum y-value to 45.

Due to the nature of the Γ-robust approach, the objective values rise with increasing values of Γ in a stepwise manner, i. e., they are monotonically increasing functions. Thus, we could have for instance replaced the primal value for $\Gamma = 26$ and s_60_600 by the lower value for $\Gamma = 28$ as every objective value for a larger value of Γ is an upper bound for the instance with a smaller Γ value. However, we have decided to display the actual bounds obtained after the time limit is reached.

For (close to) optimal solutions, we observe that the objective values stagnate at some point. For example for s_20_200, $\Gamma \geq 12$ leads to the same objective value and hence, we could have set $\Gamma_{\max} = 12$ obtaining already the most conservative solution. As mentioned in Section 8.3, the increase in the objective is called price of robustness. We display the percentage price of robustness PoR as defined in (8.13) in Figure 9.3. Note, due to the bad primal bounds, which are far from optimality, for scenarios with 60 BSs and some values of Γ we do not display PoR for these instances to obtain a readable figure. The corresponding PoR values exceed 100 extremely. For all scenarios with up to 40 BSs, the highest deterioration, i. e., increase of the objective value is less than or equal to 50 %. In case of $\Gamma \leq 40$ ($\Gamma = 42$), the highest deterioration for s_50_500 is 40 % (120 %).

Analogously to the objective value, also the PoR is a monotonically increasing function, where the same smoothing as described above could be applied in Figure 9.3 in particular for s_60_600.

The question remains, which value of Γ has to be chosen and thus, which deterioration of the objective value is inevitable, to obtain a robust optimal solution that is feasible for a number of realisations of the uncertain demand values.

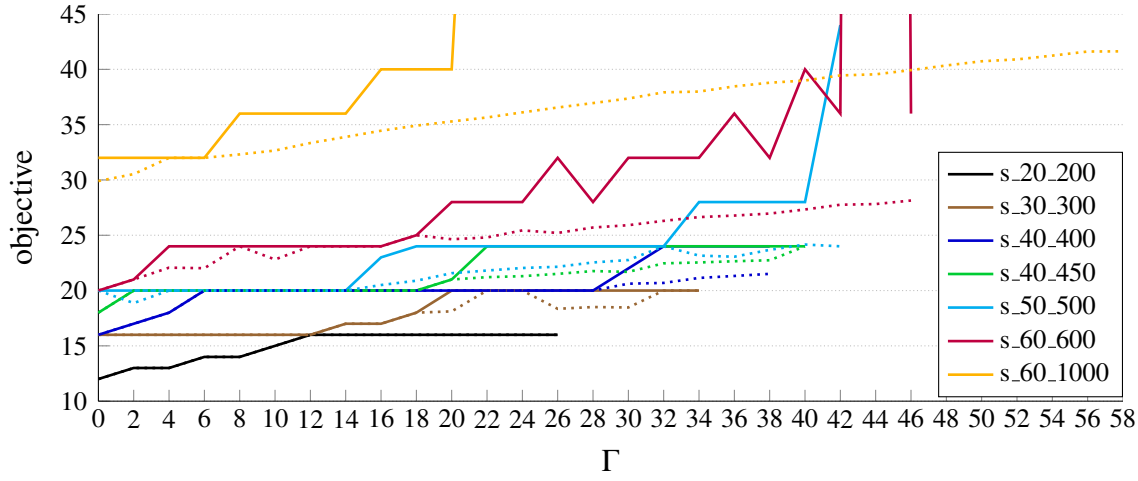


Figure 9.2.: Primal (solid lines) and dual (dotted lines) bounds for the considered values of Γ and the different scenarios.

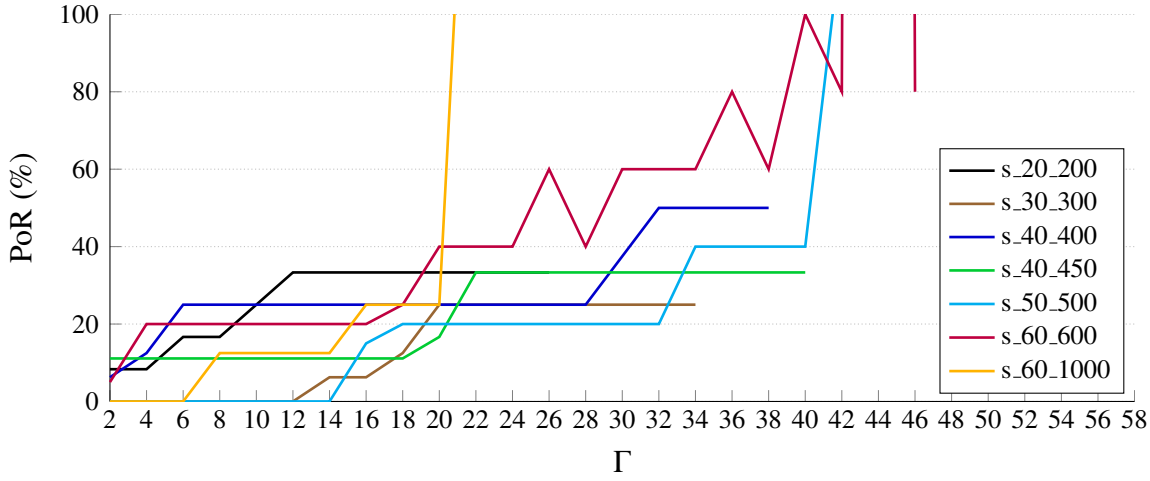


Figure 9.3.: Percentage price of robustness PoR.

9.1.5. Level of Protection

In this section, we study the level of protection ProL defined in (8.14). A robust solution is not feasible for a realisation if the capacity of at least one BS is exceeded. As the probability distribution of demand values is unknown, we investigate two types of realisations of the uncertain demands for which $\bar{w}_t$ actually is the mean and $\bar{w}_t + \hat{w}_t$ the highest deviation. For the first type of realisation, we generate 1000 snapshots based on an (integer) uniform distribution of w_t in the interval $[\bar{w}_t - \hat{w}_t, \bar{w}_t + \hat{w}_t]$. Note, if the lowest value $\bar{w}_t - \hat{w}_t$ is negative, we use 0 as the lower bound. For the second type, we generate 1000 snapshots

	s_20_200	s_30_300	s_40_400	s_40_450	s_50_500	s_60_600	s_60_1000
uniform	12	10	12	16	18	20	20
normal	12	12	12	14	16	16	18
$\Gamma_{\max}$	26	34	38	40	42	46	58

Table 9.3.: Minimum Γ for ProL = 100 % per distribution.

based on an (integer) normal distribution in the same interval with nominal value $\bar{w}_t$ and standard deviation $\hat{w}_t$. More precisely, we compute a value following the described normal distribution. In case this value is negative/less than $\bar{w}_t - \hat{w}_t$/higher than $\bar{w}_t + \hat{w}_t$, we replace it by $0/\bar{w}_t - \hat{w}_t/\bar{w}_t + \hat{w}_t$. Finally, we round the value to obtain only integer demand values.

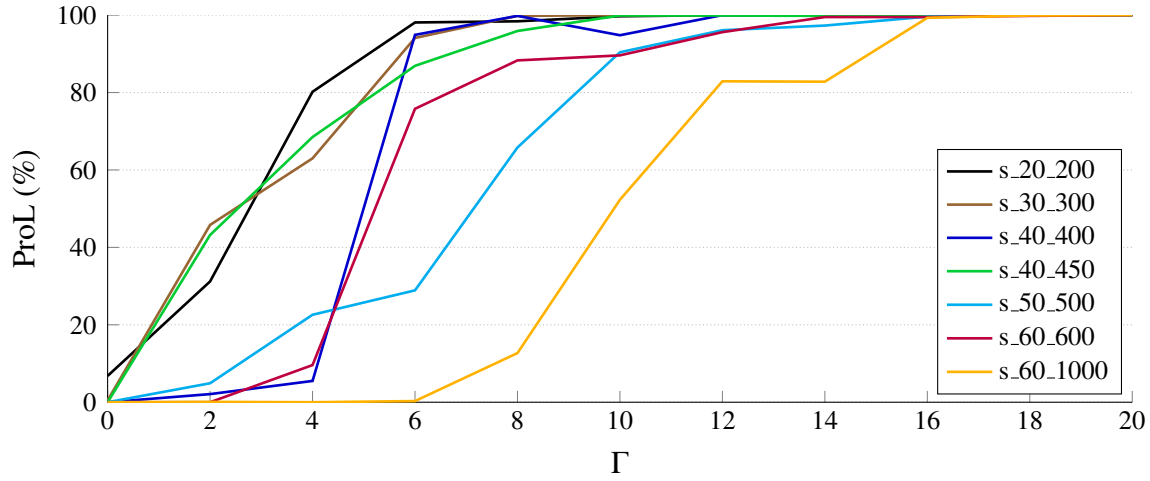
In Figure 9.4, we display ProL for the two distributions, all scenarios and $\Gamma \leq 20$ ($\Gamma \leq 18$). Note, spikes can occur as solutions for two different values of Γ are not necessarily comparable. We observe that the levels of protection for the uniform and the normal distribution are quite close and that the trend of the curves is to increase for larger values of Γ as expected.

In Table 9.3, we depict the lowest values for Γ per scenario for which the robust solution is feasible for all 1000 snapshots, i. e., the probability of a violated constraint is 0 %. Compared to the values for $\Gamma_{\max}$ displayed in Table 9.2 and computed via the probability bound (8.17), where we assumed a probability for constraint violation of at most 1 %, we observe that a value less than half as large as $\Gamma_{\max}$ is sufficient to be secured against fluctuating demand values for all investigated realisations. Note, these good results strongly depend on the chosen probability distributions. In the following section we use the maximum of the two values for Γ in Table 9.3 to compare the robust solution with the solution obtained by conventional planning.

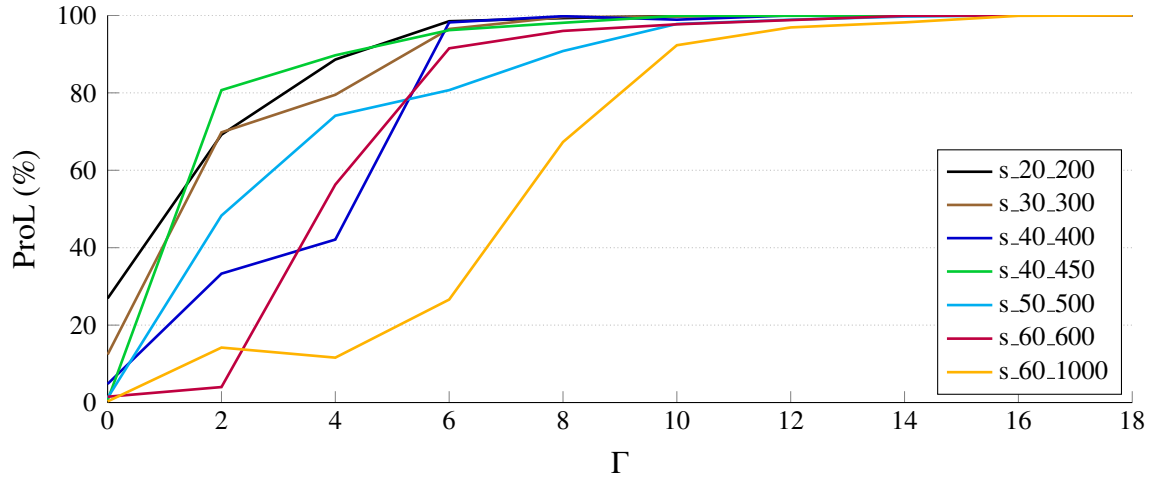
9.1.6. Conventional Planning

In conventional planning of wireless communication networks, which is usually performed in practice [148], uncertain parameters are assumed to be static. To be able to compensate demand fluctuations, the planning should be performed with demand values equal or close to the peak values; cf. Olinick [148]. Therefore, we solve the non-robust formulation (9.1) with demands $\bar{w}_t + \hat{w}_t$ and denote the optimal solution as conventional. Note, a planning that considers only the nominal demands $\bar{w}_t$ gives solutions which are infeasible for at least 93 % (73 %) of the snapshots with a uniform (normal) distribution; see Figure 9.4 and $\Gamma = 0$.

For every scenario, the conventional formulation is solved to optimality within the time limit. We compare these optimal solutions to the robust solutions obtained for the maximum of the two values of Γ depicted in Table 9.3. In both variants, all TNs can be served in all scenarios. Hence, the only difference between the two solutions is the number of deployed BSs. These numbers are given in Figure 9.5. We observe that we can save at least one and up to three BSs for all scenarios with at least 30 BS candidates with our



(a) Level of protection for an uniform distribution of traffic demands.



(b) Level of protection for a normal distribution of traffic demands.

Figure 9.4.: Level of protection for two types of probability distribution of the traffic demands.

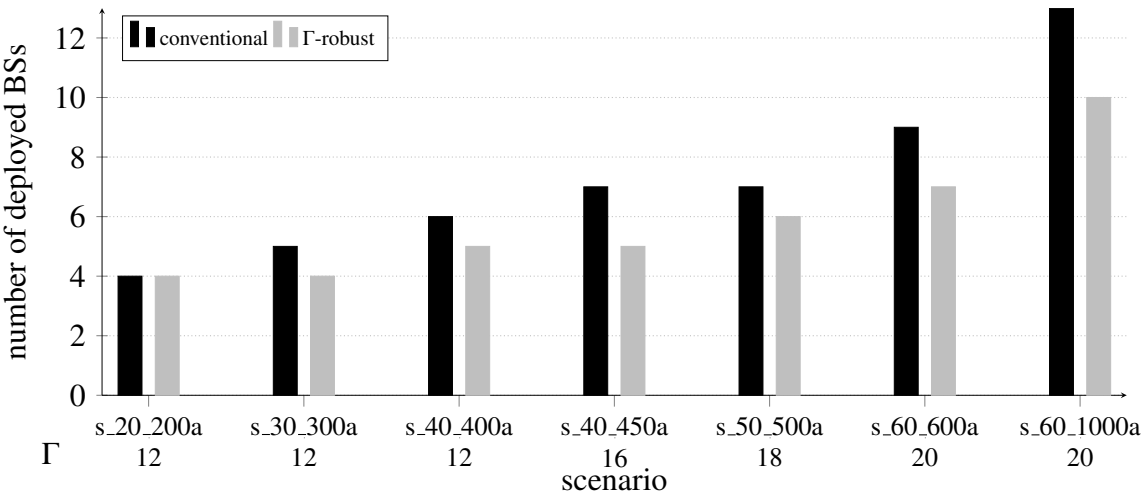


Figure 9.5.: Number of deployed BSs for the conventional planning and the Γ-robust version with Γ chosen according to Section 9.1.5.

robust solution, which is feasible for all considered 1000 snapshots per studied probability distribution, compared to conventional planning. Such a saving has a significant effect on the capital as well as operational expenditures for a network operator.

9.1.7. Conclusion

In this section, we have applied Γ-robustness to the WNPP with uncertain demands and have proposed a compact reformulation. Moreover, we have presented three types of valid inequalities, variable upper bound constraints, maximal clique inequalities, and extended Γ-robust cover inequalities and have developed a heuristic separation algorithm.

We have evaluated the Γ-robustness and the performance of the valid inequalities in a computational study performed on seven test scenarios of various dimensions. Investigations at the root node have demonstrated that vub and the cover inequalities can strengthen the formulation while the setting which activates all types has the greatest potential to improve the subsequent solving process when internal CPLEX cuts are enabled. Furthermore, we have analysed the price of robustness, in particular, for scenarios with up to 40 BSs since we could not obtain reasonable primal bounds for the largest scenarios. The percentage deterioration of the objective value is less than 50 % for all of these scenarios.

Additionally, we have computed the level of protection of the Γ-robust solutions for an uniform as well as normal distribution of demands based on 1000 snapshots per distribution. We have used the smallest Γ values which yield solutions with a level of protection of 100 % to compare our Γ-robust solutions to conventional network planning, where the peak demands have to be fulfilled. This comparison has unveiled the economy of at least one and up to three BSs, which represents savings of 14 to 29 %, for all but the smallest scenario by means of Γ-robustness.

9.2. The RWNPP with Uncertain Spectral Efficiencies

In this section, we apply the Γ-robustness approach to the WNPP (9.1) when the spectral efficiencies are subject to uncertainty and the demands are assumed to be static, which is denoted by s-RWNPP. Investigating this type of uncertainty, we continue the discussion of interference modelling of Chapter 5. For that purpose, we would like to model the spectral efficiencies e_{st} as symmetric and bounded random variables with realisations lying in an interval $[\bar{e}_{st} - \hat{e}_{st}, \bar{e}_{st} + \hat{e}_{st}]$ with nominal value $\bar{e}_{st}$ and deviation value $\hat{e}_{st}$. However, these values are computed on a logarithmic scale based on the SINR value γ_{st}; see Section 5.1. Due to the described nature of the values, the realisations do not lie in a symmetric interval. Hence, we model the SINR values γ_{st} as symmetric bounded random variables instead and derive the corresponding spectral efficiencies from the look-up Table 5.1 on page 41. We assume that realisations of γ_{st} lie in the interval $[\bar{\gamma}_{st} - \hat{\gamma}_{st}, \bar{\gamma}_{st} + \hat{\gamma}_{st}]$ with nominal value $\bar{\gamma}_{st}$ and deviation value $\hat{\gamma}_{st}$. Since the spectral efficiencies occur in the denominator in the capacity constraints (9.1d), the worst case range is $[\bar{\gamma}_{st} - \hat{\gamma}_{st}, \bar{\gamma}_{st}]$. This means, the higher the SINR value, the better. Let

$$e : \mathbb{R} \mapsto \{0, 0.25, 0.4, 0.5, 0.66, 1, 1.33, 1.5, 1.6, 2, 2.66, 3, 3.2, 4, 4.5, 4.8\}$$

be the mapping of SINR values given in dB to spectral efficiencies as defined in Table 5.1. Again, the number of simultaneous deviations is limited by a robustness parameter $\Gamma \in \{0, \dots, |\mathcal{T}|\}$. Then, the non-linear robust counterpart of the capacity constraints (9.1d) can be formulated as follows.

$$\sum_{t \in \mathcal{T}_s} \frac{w_t}{e(\bar{\gamma}_{st})} z_{st} + \max_{\mathcal{T}' \subseteq \mathcal{T}_s, |\mathcal{T}'| \leq \Gamma} \sum_{t \in \mathcal{T}'} \left(\frac{w_t}{e(\bar{\gamma}_{st} - \hat{\gamma}_{st})} - \frac{w_t}{e(\bar{\gamma}_{st})} \right) z_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S}, \tag{9.10}$$

where we replace the fraction $w_t / e(\bar{\gamma}_{st} - \hat{\gamma}_{st})$ by b_s if $e(\bar{\gamma}_{st} - \hat{\gamma}_{st}) = 0$. As explained before, the max-term can be reformulated as an ILP. Introducing dual variables π_s and ρ_{st}, we derive the linear compact robust counterpart for s-RWNPP; compare Sections 8.2 and 9.1.

min (9.1a)
s.t. (9.1b), (9.1e)

$$\sum_{t \in \mathcal{T}_s} \frac{w_t}{e(\bar{\gamma}_{st})} z_{st} + \Gamma \pi_s + \sum_{t \in \mathcal{T}_s} \rho_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{9.11a}$$

$$\pi_s + \rho_{st} \geq \left(\frac{w_t}{e(\bar{\gamma}_{st} - \hat{\gamma}_{st})} - \frac{w_t}{e(\bar{\gamma}_{st})} \right) z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{9.11b}$$

$$\pi_s,\, \rho_{st} \geq 0 \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T}. \tag{9.11c}$$

Remark 9.3. The s-RWNPP (9.11), does not yield an exact formulation neither in terms of SINR nor in terms of capacity.

Proof. Since the nominal and deviation values $\bar{\gamma}_{st}$ and $\hat{\gamma}_{st}$ are predefined, they are not necessarily the correct SINR values when the BS deployment is decided. Thus, violated SINR conditions are possible and hence, also violated capacity constraints. □

9.2.1. Defining the Interval

The crucial part of the compact formulation (9.11) is the determination of the nominal values $\bar{\gamma}_{st}$ and the lowest values $\bar{\gamma}_{st} - \hat{\gamma}_{st}$. In this subsection, we present three possibilities to determine $\bar{\gamma}_{st}$ and three possibilities to determine $\bar{\gamma}_{st} - \hat{\gamma}_{st}$ which can also be combined with each other. We investigate the performance of the various definitions in a computational study in Section 9.2.2.

As before, compare Section 5.1, we assume that the fading coefficient a_{st} for the signal from BS s to TN t is defined only by the path loss, i. e., $a_{st} = 10^{-\frac{1}{10}P_{\mathrm{L}}^{\mathrm{dB}}(s,t)}$. As a reminder, the received power at TN t from BS s is computed as $P_{\mathrm{r}}(s,t) = p_s a_{st}$ with p_s denoting the transmission power of s.

The nominal value for a fixed pair $(s,t) \in \mathcal{S} * \mathcal{T}$.

1. The best possible value for the SINR γ_{st} is the SNR value, in which no interference is present. Hence, we set the nominal value $\bar{\gamma}_{st}$ to the SNR value; see equation (5.1) on page 40.
$$\bar{\gamma}_{st} = 10 \log_{10}\left(\frac{P_{\mathrm{r}}(s,t)}{\eta}\right)$$

2. The nominal value equal to the SNR as assumed in the first approach might be too optimistic since it is unlikely that there occurs no interference at all in a realistic network. Hence, in our second approach, we set the nominal value $\bar{\gamma}_{st}$ to the SINR value which includes the BS candidate with the strongest signal as interference:
$$\bar{\gamma}_{st} = 10 \log_{10}\left(\frac{P_{\mathrm{r}}(s,t)}{P_{\mathrm{r}}(\sigma,t) + \eta}\right),$$
with $\sigma = \operatorname{argmax}_{s' \in \mathcal{S}_t \setminus \{s\}} a_{s't}$.

3. An alternative to incorporate interference in the nominal value is the following. First, we compute the SNR value $10 \log_{10}(P_{\mathrm{r}}(s,t)/\eta)$ and then take the SINR value from Table 5.1 with the next worse CQI as the nominal value. For example, for an SNR value of 3 dB (corresponding to CQI 5), we set $\bar{\gamma}_{st} = -1.0$. We can generalise this approach by taking the second or third worse value. We denote the number of shifts by the parameter $\bar{\varkappa}$.

The lowest possible value for a fixed pair $(s,t) \in \mathcal{S} * \mathcal{T}$.

1. Similar to the interference mitigation and the TN oriented approach proposed in Sections 5.6 and 5.7, we include $\varkappa \in \mathbb{Z}_{\geq 0}$ many BSs with the strongest signals as

interferers in the lowest SINR value:

$$\bar{\gamma}_{st} - \hat{\gamma}_{st} = 10 \log_{10} \left(\frac{P_{\mathrm{r}}(s,t)}{\sum\limits_{s' \in \mathcal{S}'} P_{\mathrm{r}}(s',t) + \eta} \right),$$

with $\mathcal{S}' \subseteq \mathcal{S}_t \setminus \{s\}$, $|\mathcal{S}'| = \varkappa_t$, $\varkappa_t := \min\{\varkappa, |\mathcal{S}_t|\}$ and $a_{s't} \geq a_{\sigma t}$ for every $s' \in \mathcal{S}'$ and $\sigma \notin \mathcal{S}'$.

2. Including $\varkappa$ many worst interferers as proposed in the previous approach might be too conservative since BSs are usually not installed right next to each other in a realistic network. Hence, we now present a slight modification of the first approach. We take $\varkappa$ times the average power of all interfering signals as the interference power included in the SINR value:

$$\bar{\gamma}_{st} - \hat{\gamma}_{st} = 10 \log_{10} \left(\frac{P_{\mathrm{r}}(s,t)}{\varkappa \cdot \frac{\sum_{\sigma \in \mathcal{S}_t \setminus \{s\}} P_{\mathrm{r}}(\sigma,t)}{|\mathcal{S}_t|} + \eta} \right).$$

3. Following the third approach to define the nominal value $\bar{\gamma}_{st}$, we introduce a parameter $\hat{\varkappa} \in \mathbb{Z}_{\geq 0}$ to denote the number of shifts of the SNR value regarding Table 5.1 to define the lowest SINR value. Note, to guarantee that $\bar{\gamma}_{st} - \hat{\gamma}_{st} < \bar{\gamma}_{st}$, we do not combine this approach with a different possibility to define the nominal value than the third approach and we stipulate $\hat{\varkappa} > \bar{\varkappa}$.

9.2.2. Computational Study

In the computational study presented in this section, we analyse the performance of the compact formulation (9.11) for the s-RWNPP with the different approaches to set the nominal and the lowest SINR values proposed in the previous section. The criteria we investigate are the number of deployed BSs and the number of actually served TNs, which is the number of TNs with $u_t = 0$ in the best solution minus the number of violated SINR conditions (5.1). The studied scenarios are the same as described in Section 5.9.1 and we find good values for $\Gamma_{\max}$ via the probability bound (8.17) assuming a probability of constraint violation of 1 %; see Section 9.1.2. In Table 9.4, we display the average value per scenario. As before, we consider only even values for Γ and set $\Gamma_{\max}$ to the next larger even number of these average values.

The different investigated settings of nominal and lowest SINR values are summarised in Table 9.5. For every setting, we also apply the performance improvements presented in Section 9.1.1 where we adapt the computation of extended robust cover inequalities in a straightforward way.

All computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of two hours. Additionally, we set a memory limit of 11 GB and use CPLEX 12.4 [98].

scenario	average	$\Gamma_{\max}$
20_200a	25.20	26
20_300b	32.10	34
20_400c	39.75	40
30_300a	32.97	34
30_400b	35.17	36
30_500c	43.03	44
40_400a	36.85	38
40_450a	39.33	40

Table 9.4.: The average values of $\Gamma_{\max}$ computed via (8.17) with probability 1 % and actually selected values for $\Gamma_{\max}$.

$\bar{\gamma}_{st}$	$\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$
1. SNR	1. $\varkappa$ strongest interferers, $\varkappa \in \{1, 5, 10\}$
1. SNR	2. $\varkappa$ times average interference, $\varkappa \in \{1, 5, 10\}$
2. SINR with strongest interferer	1. $\varkappa$ strongest interferers, $\varkappa \in \{1, 5, 10\}$
2. SINR with strongest interferer	2. $\varkappa$ times average interference, $\varkappa \in \{1, 5, 10\}$
3. $\bar{\varkappa} = 1$	3. $\hat{\varkappa} \in \{2, 3\}$
3. $\bar{\varkappa} = 2$	3. $\hat{\varkappa} = 3$

Table 9.5.: Investigated settings of nominal and lowest SINR values.

$\bar{\gamma}_{st}$ = SNR Foremost, we investigate the first setting for the nominal demands where $\bar{\gamma}_{st}$ is defined as the SNR value, i. e., the first two lines of Table 9.5. In Figure 9.6, we display exemplarily for scenario 20_400c the objective values and the SINR-corrected objective values υ (see (5.35) on page 63) for $\Gamma \in \{2, 4, \ldots, 44\}$, where $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ is defined via the $\varkappa = 5$ strongest interferers. As mentioned in Section 9.1.4, a characteristic of Γ-robustness is that the objective value represents a monotonically increasing function. In contrast, the SINR-corrected objective value υ, which is not explicitly optimised, is not monotonically increasing due to the post-calculation of actually served TNs. Higher values of Γ can still give a lower value for υ. We define the Γ yielding the lowest SINR-corrected objective value as "best". However, to be sure that we have found the best solution, we additionally have to test larger values of Γ which causes significant computational overhead. Note that this behaviour does not depend on the definition of the deviation interval.

In Table 9.6 on page 146, we display the number of actually served TNs and the number of deployed BSs for the non-robust formulation ($\Gamma = 0$) and the best values achieved for $\Gamma > 0$ (without stating the precise value of the corresponding Γ here), where the two possibilities to define $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ as well as all values $\varkappa \in \{1, 5, 10\}$ are considered. For increasing values of $\varkappa$, i. e., more BSs are considered in the interfering set $\mathcal{S}'$, the solution becomes more conservative and hence, the probability of a SINR violation decreases implying an increasing number of actually served TNs. Moreover, if the strongest interferers

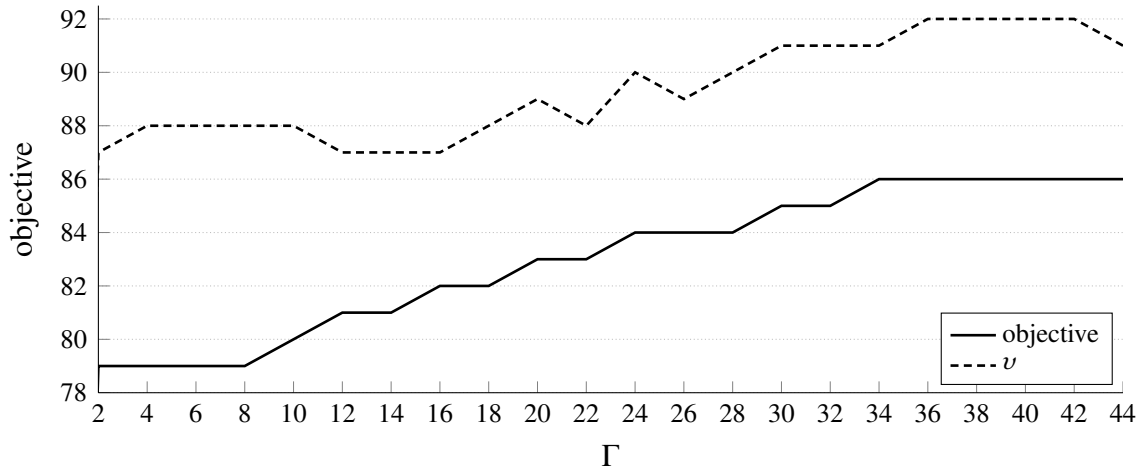


Figure 9.6.: The objective value and the SINR-corrected objective value υ for scenario 20_400c and $\Gamma \in \{2, 4, \ldots, 44\}$ with $\bar{\gamma}_{st}$ being defined as the SNR value and $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ by the $\varkappa = 5$ strongest interferers.

are considered, we can in general also serve more TNs than if considering the averaged interference. However, up to 21 TNs (40_450a) cannot be covered for the best robust setting (strongest interference and $\varkappa = 10$). The non-robust formulation usually gives results which are quite close to those of the best robust solution but deploys significantly less BSs. Thus, the definition of the nominal value via the SNR value does not give satisfying results in terms of robustness or interference modelling and we do not investigate this setting any further.

$\bar{\gamma}_{st}$ = SINR with strongest interferer We now investigate the second setting for the nominal demands where $\bar{\gamma}_{st}$ is set to the SINR value which regards the strongest interferer as the only interference. Thus, we study the third and forth line in Table 9.5. As before, we compute the number of actually served TNs for the non-robust formulation as well as the two possibilities to define $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ with $\varkappa \in \{1, 5, 10\}$. We display these values and the number of deployed BSs for $\Gamma = 0$ and the best value for $\Gamma > 0$ (without stating the precise value of the corresponding Γ here) in Table 9.7.

Again, the number of actually served TNs in general increases for an increasing value of $\varkappa$ and including the strongest interferers instead of the averaged value gives better results. However as before, the non-robust results are quite close to the best robust results but this time also the non-robust solution deploys an unjustified large amount of BSs. Thus, also the definition of the nominal value via the SINR value with the strongest interferer does not give satisfying results in terms of robustness or interference modelling and we also do not investigate this setting any further.

			strongest interference						average interference					
	$\Gamma = 0$		$\varkappa = 1$		$\varkappa = 5$		$\varkappa = 10$		$\varkappa = 1$		$\varkappa = 5$		$\varkappa = 10$	
scenario	T	B	T	B	T	B	T	B	T	B	T	B	T	B
20_200a	185	3	192	13	192	14	192	14	164	9	179	11	194	14
20_300b	284	3	279	14	286	17	287	17	233	10	273	12	280	16
20_400c	388	4	371	17	385	18	390	18	325	12	356	17	376	18
30_300a	270	4	255	19	275	20	280	21	236	12	241	17	260	19
30_400b	398	4	359	22	381	24	385	25	342	16	357	19	364	21
30_500c	459	5	450	26	487	27	490	27	420	16	442	22	448	26
40_400a	380	4	344	28	379	31	384	32	345	15	329	21	341	27
40_450a	430	4	387	29	427	31	429	31	385	15	355	23	388	25

Table 9.6.: Number of actually served TNs and deployed BSs for $\bar{\gamma}_{st}$ =SNR, $\Gamma = 0$ and best values for one $\Gamma > 0$, where $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ is defined via $\varkappa \in \{1, 5, 10\}$ strongest interferers or times the average interference.

			strongest interference						average interference					
	$\Gamma = 0$		$\varkappa = 1$		$\varkappa = 5$		$\varkappa = 10$		$\varkappa = 1$		$\varkappa = 5$		$\varkappa = 10$	
scenario	T	B	T	B	T	B	T	B	T	B	T	B	T	B
20_200a	184	12	184	12	183	12	184	12	176	11	178	12	180	12
20_300b	266	14	266	14	283	17	283	17	262	13	265	15	273	16
20_400c	360	17	360	17	381	18	381	18	350	17	357	18	373	18
30_300a	263	20	263	20	271	21	273	21	252	19	242	19	254	18
30_400b	352	22	352	22	372	23	376	24	344	21	341	23	351	23
30_500c	451	26	451	26	483	26	483	27	440	26	445	26	468	24
40_400a	352	28	352	28	375	31	374	31	351	28	358	29	356	30
40_450a	384	29	384	29	419	31	424	31	387	29	390	27	401	29

Table 9.7.: Number of actually served TNs and deployed BSs for $\bar{\gamma}_{st}$ =SINR with strongest interferer, $\Gamma = 0$ and best values for $\Gamma > 0$, where $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ is defined via $\varkappa \in \{1, 5, 10\}$ strongest interferers or times the average interference.

$\bar{\gamma}_{st}$ via $\bar{\varkappa} = 1$ In this paragraph, we study the third approach to define the nominal values via $\bar{\varkappa} = 1$ many shifts. The number of actually served TNs and deployed BSs for the non-robust formulation as well as the best values, including the respective Γ values, for the robust formulation with $\hat{\varkappa} \in \{2, 3\}$ shifts of $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$ are depicted in Table 9.8. The solution obtained for the setting $\hat{\varkappa} = 2$ typically covers more TNs than the setting $\hat{\varkappa} = 3$ and also than the non-robust formulation while the same number of BSs is deployed. Thus, the setting $\bar{\varkappa} = 1$ and $\hat{\varkappa} = 2$ gives the best results so far.

As a further remark, we would like to point out that the results displayed in Table 9.8 indicate that the Γ-robustness approach does not behave as expected in the present application. For higher numbers of TNs we would expect that a higher value of Γ gives the best solution since the objective value represents a monotonically increasing function. But as the number of *actually* served TNs is not optimised, we cannot expect the same monotonicity for this number yielding fluctuating best Γ values for increasing numbers of TNs; see also Figure 9.6.

	$\Gamma = 0$		$\hat{\varkappa} = 2$			$\hat{\varkappa} = 3$		
scenario	T	B	Γ	T	B	Γ	T	B
20_200a	196	3	2	196	3	4	190	3
20_300b	267	4	2	286	4	28	273	4
20_400c	361	5	8	395	5	12	382	5
30_300a	276	4	12	298	4	14	290	4
30_400b	366	5	16	392	5	26	389	5
30_500c	490	5	26	486	5	2	485	5
40_400a	353	5	24	393	5	26	395	5
40_450a	429	5	10	437	5	16	433	5

Table 9.8.: Number of actually served TNs and deployed BSs for $\bar{\varkappa} = 1$, $\Gamma = 0$, and $\hat{\varkappa} \in \{2, 3\}$ and the best value of $\Gamma > 0$.

	$\Gamma = 0$		$\hat{\varkappa} = 3$		
scenario	T	B	Γ	T	B
20_200a	186	3	2	187	4
20_300b	269	4	30	274	4
20_400c	379	5	26	382	5
30_300a	288	4	16	289	5
30_400b	373	5	12	390	5
30_500c	451	6	12	484	6
40_400a	362	5	8	392	5
40_450a	429	5	26	431	5

Table 9.9.: Number of actually served TNs and deployed BSs for $\bar{\varkappa} = 1$, $\Gamma = 0$, and $\hat{\varkappa} = 3$ and the best value of Γ.

$\bar{\gamma}_{st}$ via $\bar{\varkappa} = 2$ Finally, we investigate $\bar{\varkappa} = 2$ shifts for the nominal value and $\hat{\varkappa} = 3$ shifts for $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$. The number of actually served TNs and deployed BSs for the non-robust formulation as well as the best robust formulation, including the respective Γ values, are displayed in Table 9.9. The robust formulation can serve more TNs than the non-robust formulation by deploying at most one BS more for every scenario. If we compare this setting to $\bar{\varkappa} = 1$ and $\hat{\varkappa} = 2$, we observe that the latter setting serves more TNs than the present one and does not deploy more BSs than the non-robust formulation. Altogether, $\bar{\varkappa} = 1$ and $\hat{\varkappa} = 2$ gives the best results in terms of number of covered TNs and deployed BSs.

In the next paragraph, we compare this setting to the TN coverage requirement formulation, which we regard as the best model among the various interference modelling formulations presented in Chapter 5. Additionally, we present a comparison to the conflict graph formulation which we use in d-RWNPP and subsequent formulations to limit interference.

	s-RWNPP		TN cov. req.		conflict	
scenario	υ	υ_l	υ	υ_l	υ	υ_l
20_200a	16	22	15	16	27	35
20_300b	30	65	18	20	28	58
20_400c	25	67	22	51	31	65
30_300a	18	47	25	42	59	94
30_400b	28	57	16	37	18	44
30_500c	34	96	24	60	42	116
40_400a	27	69	21	52	30	68
40_450a	33	86	31	48	36	69

Table 9.10.: SINR-corrected objective values υ and υ_l for s-RWNPP with the best setting $\bar{\varkappa} = 1$, $\hat{\varkappa} = 2$ and for the TN coverage requirement formulation.

Comparison to other formulations In Table 9.10, we display the SINR-corrected objective values υ and υ_l as defined on pages 63 and 64 for s-RWNPP with $\bar{\varkappa} = 1$ and $\hat{\varkappa} = 2$ and for the TN coverage requirement as well as the conflict graph formulation. Comparing υ with υ_l, we observe a significant increase which is due to (strongly) exceeded capacities. The s-RWNPP formulation (9.11) can neither guarantee that no SINR requirement is violated nor that no capacity is exceeded; see Remark 9.3.

With the exception of scenario 30_300a, the s-RWNPP gives (considerably) higher values for υ as well as for υ_l than the TN coverage requirement formulation. However, these values are in general closer to the magnitude of the values for the conflict graph formulation. In summary, the TN coverage requirement formulation is better suited to incorporate interference modelling approximately in the WNPP but s-RWNPP gives similar results to the conflict graph formulation.

9.2.3. Conclusion

In this section, we have applied the Γ-robustness concept to uncertain spectral efficiencies in the WNPP. The compact formulation (9.11) strongly depends on the definition of the deviation interval, hence, on $\bar{\gamma}_{st}$ and $\bar{\gamma}_{st}$-$\hat{\gamma}_{st}$. Based on a computational study performed for eight test scenarios of various dimensions, we have determined the best setting among the presented possibilities which is to shift the SNR value once to obtain the nominal value and twice to obtain the lowest value ($\bar{\varkappa} = 1$ and $\hat{\varkappa} = 2$). However, based on the best Γ values stated in Tables 9.8 and 9.9, we also observe that the Γ-robustness approach does not behave as expected in the studied application. For higher numbers of TNs we would expect that a higher value of Γ gives the best solution. Moreover, a higher value of Γ does not necessarily yield a solution at least as good as the best solution.

In summary, we cannot predict the number of violated SINR conditions or the number of actually served TNs when varying the value of Γ. Thus, the Γ-robustness concept is not suitable to model interference in the WNPP in the presented way, which can also be seen in the comparison to the TN coverage requirement formulation.

10. A B&P Approach for the d-RWNPP

The compact formulation (9.3) of the RWNPP with uncertain demands discussed in Section 9.1 consists of a huge number of variables. In fact, the number of variables is twice the number of BSs and possible BS-TN pairs. Furthermore, the ILP (9.3) can have a weak LP solution. A prominent procedure to tackle problems containing large numbers of variables and/or complex subproblems is column generation or B&P for ILPs as introduced in Section 1.2.3. Applying a Dantzig-Wolfe decomposition, the solution process starts with a subset of variables (columns) and only variables having the potential to improve the objective are generated on the fly. As demonstrated in an example in the subsequent section as well as in a computational study in Section 10.3, this method can significantly improve the LP solution compared to the compact model.

B&P is a commonly used technique to solve a variety of integer problems, in particular, various wireless network design problems. Bjorklund et al. [27] investigate resource allocation in ad hoc networks, where time slots are assigned to nodes or links, respectively. The authors use a set covering formulation and a B&P algorithm to solve this problem. The LP relaxations can be efficiently solved and "provide very tight bounds to the integer solutions". For TDMA (time division multiple access) based networks, Gomes et al. [85] study the problem of joint routing and scheduling, where time is divided into slots and allocated to users. In case of integer flows, the authors apply a B&P algorithm which solves the problem easily as soon as a critical region (bottleneck) is solved. Moreover, Fu et al. [79] analyse the scheduling of wireless links and the control of the power transmission in STDMA (spatial-reuse TDMA) based networks, in which the same time slot can be used for more than one user if these users are positioned sufficiently far away from each other. Assuming that the allocated time is integer, the authors use a B&P algorithm to solve the problem. They further introduce a new pricing problem (PP) and a smart enumerating algorithm to solve the PPs achieving significant reduction in runtime. In the work of Wilhelm and Gokce [177], surveillance systems for port and waterway security via sensor networks are investigated. The aim is to find a strategic design for the sensors which provides sufficient surveillance. Applying B&P, the authors propose a heuristic for a basic initial solution for the restricted master problem (RMP), an effective method for solving PPs and alternative branching rules. The computational study demonstrates that their B&P algorithm requires less runtime than using CPLEX directly.

In this chapter, we present a complete B&P algorithm for the d-RWNPP and performance improvements whereupon the results are based on our work [50]. First, we develop the master problem (MP) via a Dantzig-Wolfe decomposition, present the RMP, the PP and problem specific branching rules in Section 10.1. Afterwards, in Section 10.2, we propose several techniques to speed up the B&P algorithm. The effectiveness of these performance improvements is analysed in a computational study in Section 10.3. A comparison of the

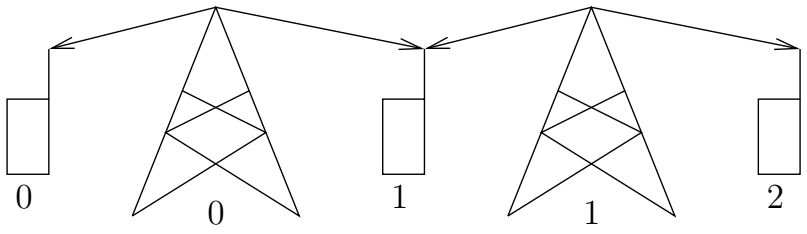


Figure 10.1.: An example consisting of two BSs and three TNs to demonstrate the improvement of the LP solution by means of column generation.

B&P algorithm to the compact formulation (9.3) is conducted in a separate computational study in Section 10.4.

10.1. A B&P formulation

As pointed out in Section 1.2.3, the LP solution computed by a column generation approach can be significantly better than the LP relaxation of the compact model (9.3) if there exist non-integer extreme points of the relaxed PP. To demonstrate this, we give a tiny non-robust example with two BSs and three TNs, see Figure 10.1. Every BS has an available bandwidth of 40 and entails costs of 4000. We assume that the BSs are not interfering with each other so that no conflict graph exists. Every TN has a nominal demand of 30 and no deviation. We choose the spectral efficiencies such that TNs 0 and 1 can be assigned to BS 0 and TNs 1 and 2 to BS 1 represented by arrows in the figure. The scaling parameter λ is set to 10000.

The LP solution of the compact model is 8000 with assignment variables $z_{00} = z_{12} = 1$ and $z_{01} = z_{11} = 0.5$ installing both BSs. The column generation algorithm, which we present in the subsequent sections, gives a LP solution of 18000 with exactly two reformulated assignment variables equal to one such that both BSs are installed, TN 0 is assigned to BS 0, TN 2 to BS 1 and TN 1 remains uncovered. Since we have a minimisation problem, this is a much better lower bound. In fact, for this tiny example the LP solution of the column generation is the optimal integer solution.

Another reason for a B&P algorithm is the decomposition of the compact model into master and pricing problems where constraints, that provide a structure for which the computation of extreme points is a tractable problem, are outsourced to the PPs. For the d-RWNPP, we move the embedded Γ-RKP constraints to the PPs leading to a MP which is independent of Γ-robustness. By this means, we can apply the (extended) Γ-robust cover inequalities introduced in Section 9.1.1 to enhance the solving performance for the PPs.

10.1.1. The Master Problem

To apply a Dantzig-Wolfe decomposition as introduced in Section 1.2.3, we neglect the binary variables x_s for simplicity in the following. We reformulate the compact formulation (9.3) by substituting the assignment variables z_{st} by the convex combination of

the extreme points of conv(Z_s), where Z_s denotes the subproblem which consists of constraints (9.3a) and (9.3b), i. e.,

$$Z_s := \Bigg\{(z,\pi,\rho) \in \{0,1\}^{|\mathcal{T}_s|} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{|\mathcal{T}_s|} \;\Bigg|\; \sum_{t\in\mathcal{T}_s} \frac{\bar{w}_t}{e_{st}} z_t + \Gamma\pi + \sum_{t\in\mathcal{T}_s} \rho_t \leq b_s, \tag{10.1a}$$

$$\pi + \rho_t \geq \frac{\hat{w}_t}{e_{st}} z_t \;\forall\, t \in \mathcal{T}_s\Bigg\}. \tag{10.1b}$$

Therefore, we have one Γ-RKP per BS as subproblem which can be considered separately. An extreme point (z,π,ρ) of conv(Z_s) is represented by a subset $\tau \subseteq \mathcal{T}_s$ of TNs with $\tau = \{t \in \mathcal{T}_s \,|\, z_t = 1\}$ for which (10.1a) and (10.1b) are satisfied. Note that the extreme points of an LP relaxation of a Γ-RKP are not necessarily integer and hence, a strictly better lower bound than the LP relaxation can be possibly obtained by column generation. We introduce a set T_s for every BS $s \in \mathcal{S}$ consisting of all extreme points of conv(Z_s). Then, we denote the new assignment variables by $\zeta_{s\tau}$ for $s \in \mathcal{S}$ and $\tau \in \mathrm{T}_s$ with

$$\zeta_{s\tau} = \begin{cases} 1, & \text{the set of TNs } \tau \subseteq \mathcal{T}_s \text{ is assigned to BS } s \text{ (and } \mathcal{T}_s \setminus \tau \text{ not)} \\ 0, & \text{otherwise.} \end{cases}$$

For a full substitution of z_{st} by the convex combination of the extreme points, we have to add the convexity constraints

$$\sum_{\tau\in\mathrm{T}_s} \zeta_{s\tau} = 1 \quad \forall s \in \mathcal{S}, \tag{10.2}$$

which guarantee that exactly one subset τ is selected. Note that the empty set is also included in T_s. Since an assignment of TNs to a BS s is only possible if s is deployed, we can replace the right hand side of (10.2) by the deployment indicator variable x_s. Moreover, we keep the non-coverage indicator variables u_t as introduced for the compact formulation (9.3). We can now state the MP as the following ILP.

$$\min \sum_{s\in\mathcal{S}} c_s x_s + \lambda \sum_{t\in\mathcal{T}} u_t \tag{10.3a}$$

$$\text{s.t.} \sum_{s\in\mathcal{S}_t} \sum_{\tau\in\mathrm{T}_s : t\in\tau} \zeta_{s\tau} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{10.3b}$$

$$\sum_{s\in\mathcal{U}} -x_s \geq -1 \qquad \forall\, \mathcal{U} \subset \mathcal{S}, \mathcal{U} \text{ max. clique in } G \tag{10.3c}$$

$$x_s - \sum_{\tau\in\mathrm{T}_s} \zeta_{s\tau} \geq 0 \qquad \forall\, s \in \mathcal{S} \tag{10.3d}$$

$$x_s,\ \zeta_{s\tau},\ u_t \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\ \tau \in \mathrm{T}_s,\ t \in \mathcal{T} \tag{10.3e}$$

The objective (10.3a) and the maximal clique inequalities (10.3c) are exactly the same as the compact equivalents (9.1a) and (9.5). Furthermore, constraints (10.3b) present the

reformulated coverage constraints (9.1b) and (10.3d) the relaxed version of the convexity constraints (10.2). Such a relaxation is possible as the deployment of a BS without any assigned TN is not reasonable due to the minimisation of installed BSs and non-covered TNs. Note, these constraints are also reformulated vub constraints (9.4).

In the following, we describe two modifications of the MP which simplify the definition of the RMP and the PPs in the subsequent sections. First, it is sufficient to consider only "≥" in constraints (10.3b) since we minimise the number of u_t set to one. Thus, we replace these constraints by

$$\sum_{s \in \mathcal{S}_t} \sum_{\tau \in \mathrm{T}_s : t \in \tau} \zeta_{s\tau} + u_t \geq 1 \qquad \forall\, t \in \mathcal{T}. \tag{10.4}$$

It is also sufficient to consider only maximal sets τ, to which no further TNs can be added without violating the capacity constraint, since it is beneficial to assign as many TNs as possible to a deployed BS when minimising the number of not covered TNs.

Second, the definition of all variables as binary is not necessary. Instead it is sufficient to have

$$x_s,\ \zeta_{s\tau},\ u_t \in \mathbb{Z}_{\geq 0} \qquad \forall\, s \in \mathcal{S},\ \tau \in \mathrm{T}_s,\ t \in \mathcal{T}. \tag{10.5}$$

The upper bound of 1 for all variables is guaranteed by constraints (10.3c) and (10.3d) as well as the minimisation in the objective (10.3a).

Note, these modifications do not, in general, speed up the solving process of the compact formulation (9.3) studied in Section 9.1. We ran tests for the instances s_20_200–s_60_600 and Γ values described in Table 9.2 and a time limit of two hours. In only 11 cases out of 118 different settings, the compact formulation including similar modifications was solved faster than the compact formulation (9.3). This is the reason why we did not consider these modifications in the computational study in Section 9.1.

The MP (10.3) describes the d-RWNPP completely. However, for each BS there exists a (potentially) exponential number of extreme points/sets $\tau \in \mathrm{T}_s$ resulting in a huge model. Hence, we restrict the MP to subsets $\mathrm{T}'_s \subseteq \mathrm{T}_s$ for each $s \in \mathcal{S}$ obtaining the RMP and compute further necessary columns by the PPs presented in the following section.

10.1.2. The Pricing Problems

As stated before, the RMP does not consider the total amount of assignment variables ζ at the outset. To decide which variable has the potential to improve the objective value but is not yet included in the RMP, we compute the reduced cost of $\zeta_{s\tau}$ for all $s \in \mathcal{S}$ and $\tau \in \mathrm{T}_s \setminus \mathrm{T}'_s$ by means of dual variables. For that purpose, we introduce dual variables α_s for each vub constraint (10.3d) and dual variables β_t for each coverage constraint (10.4). The reduced cost of $\zeta_{s\tau}$ are then computed as the objective value minus the coefficients:

$$0 - (-\alpha_s) - \sum_{t \in \tau} \beta_t.$$

It is beneficial to add the variable $\zeta_{s\tau}$ to the RMP if the reduced cost are negative, hence if

$$\sum_{t\in\tau}\beta_t > \alpha_s.$$

To detect variables with negative reduced cost, we introduce one PP for each BS s as follows. We define a binary decision variable a_t for every TN $t \in \mathcal{T}_s$ with

$$a_t = \begin{cases} 1, & \text{TN } t \text{ is in the newly constructed set } \tau \in \mathrm{T}_s \setminus \mathrm{T}'_s \\ 0, & \text{otherwise.} \end{cases}$$

The PP for a fixed BS s determines an extreme point of $\mathrm{conv}(Z_s)$ where the corresponding variable $\zeta_{s\tau}$ has the most negative reduced cost if such a variable exists. The problem can be formulated as follows.

$$\max \sum_{t\in\mathcal{T}_s}\tilde{\beta}_t a_t \tag{10.6a}$$

$$\text{s.t.} \sum_{t\in\mathcal{T}_s}\frac{\bar{w}_t}{e_{st}}a_t + \Gamma\pi + \sum_{t\in\mathcal{T}_s}\rho_t \le b_s \tag{10.6b}$$

$$\pi + \rho_t \ge \frac{\hat{w}_t}{e_{st}}a_t \qquad \forall\, t \in \mathcal{T}_s \tag{10.6c}$$

$$a_t \in \{0,1\},\ \pi,\ \rho_t \ge 0 \qquad \forall\, t \in \mathcal{T}_s, \tag{10.6d}$$

where $\tilde{\beta}_t$ is the optimal value of the dual β_t of the current RMP and π, ρ_t are the dual variables introduced for the compact Γ-robust counterpart of the capacity constraints; see Section 9.1. Note, the index s is dropped here for these variables as the BS is fixed. Obviously, a PP is a Γ-RKP.

If the objective value (10.6a) for an optimal solution $(\tilde{a}, \tilde{\pi}, \tilde{\rho})$ is greater than the optimal value $\tilde{\alpha}_s$, we have found a variable $\zeta_{s\tau}$ with negative reduced cost, where $\tau := \{t \in \mathcal{T}_s \,|\, \tilde{a}_t = 1\}$. We add this new variable to the RMP and solve it again. The PPs are solved consecutively for every BS $s \in \mathcal{S}$. The whole process of solving the RMP giving new dual values and then solving the PPs to decide if further variables are needed is repeated until no more variables with negative reduced cost exist.

10.1.3. Branching Rules

By means of the previously described decomposition into RMP and PPs, we can solve the LP relaxation of the MP (10.3) by column generation. However, the optimal LP solution is not necessarily integer. Hence, a branching process has to be carried out.

A standard branching procedure is to create two child nodes based on the integrality criterion. We apply this branching rule first to branch on non-integer BS deployment indicator variables x_s creating two child nodes with additional constraint $x_s \le 0$ and $x_s \ge 1$, respectively. If all of these variables are integer, we apply the same branching rule to the non-coverage indicator variables u_t.

When no fractional x_s or u_t variables are left, we have to branch on non-integer assignment variables $\zeta_{s\tau}$. However, the branching rule based on the integrality criterion leads to the following difficulties. Enforcing $\zeta_{s\tau} \leq 0$ in a child node requires to add the additional constraint

$$\sum_{t\in\tau} a_t + \sum_{t\notin\tau}(1 - a_t) \leq |\mathcal{T}_s| - 1$$

to the PP (10.6) corresponding to BS s. This constraint guarantees that we cannot add all TNs in the set τ to a new subset of TNs and at the same time not adding any of the TNs not in τ. Otherwise, we would generate the set τ again. Consequently, the PP is a Γ-RKP with an additional constraint. Future branching decisions lead to further supplementary constraints which destroy the structure of the PPs. On the other hand, if we do not enforce the ζ-decision in the PP corresponding to s, it is possible that we compute exactly the same set τ again which is a problem inherent to the column generation approach. To avoid these difficulties, we develop different problem specific branching rules which are based on the common technique to branch on the corresponding variables of the compact formulation; see, e. g., Barnhart et al. [14]. Hence, we intend to apply the branching rule based on the integrality criterion to the assignment variables z_{st}. Branching on these variables in the B&P approach comprises the generation of two child nodes with $z_{st} \leq 0$ and $z_{st} \geq 1$, respectively, if

$$z_{st}^{\mathrm{LP}} := \sum_{\tau\in\mathrm{T}'_s : t\in\tau} \zeta_{s\tau}^{\mathrm{LP}}$$

is not integer in the current LP solution. Note that we have introduced the notation z_{st}^{LP} just for simplicity. In each branching step, we branch on the most fractional value that is the z_{st}^{LP} closest to 0.5.

For a BS-TN pair (s, t) with a non-integer value z_{st}^{LP}, we generate two child nodes with additional branching constraints

$$z_{st}^{\mathrm{LP}} \leq 0 \tag{10.7}$$

and

$$z_{st}^{\mathrm{LP}} \geq 1, \tag{10.8}$$

respectively. By the first constraint (10.7), every variable $\zeta_{s\tau}$ is implicitly fixed to 0 for all $\tau \in \mathrm{T}'_s$ with $t \in \tau$. Hence, only new subsets $\tau' \setminus \{t\}$ can be beneficial in this subproblem enforcing $a_t = 0$ in the PP corresponding to BS s. Note, when the subset T'_s is complemented during the subsequent solving process, also constraints (10.7) and (10.8) have to be updated accordingly.

On the other hand, constraint (10.8) guarantees that TN t is served by BS s. Including this constraint in the subproblem, we have to consider a new dual variable corresponding to (10.8) for the computation of the reduced cost of a new pricing variable. This new dual variable has to be multiplied by a_t and this product is then added to the objective function of the PP. However, the consideration of further dual variables complicates the whole problem. Hence, instead of including new dual variables in the PPs, we reformulate constraint (10.8) as follows. By the coverage constraints (10.3b), we know that $u_t = 0$ as TN t is assigned to BS s. Further, by (10.8) $\zeta_{\sigma\tau} = 0$ for all BSs $\sigma \neq s$ and subsets τ

containing t. Combining these two observations, constraint (10.8) is equivalent to

$$\sum_{\sigma \in \mathcal{S} \setminus \{s\}} \sum_{\tau \in \mathrm{T}'_\sigma : t \in \tau} \zeta_{\sigma\tau} + u_t \leq 0. \tag{10.9}$$

Replacing (10.8) by (10.9), we do not have to consider the corresponding dual variable in the PP for s. Instead, we set $a_t = 0$ in the PPs corresponding to $\sigma \in \mathcal{S} \setminus \{s\}$. Hence, the dual variable corresponding to (10.9) is multiplied by zero. Additionally, we can further reduce the number of variables in the PP related to s by setting $a_t = 1$.

As mentioned earlier, in each node of the B&B tree it is necessary to know the path to the root node to add all constraints of type (10.7) and (10.9) that have been computed on this path and to adjust the PPs accordingly. Additionally, subsequent computed variables have to be added to the corresponding branching constraints (10.7) and (10.9) respecting s and t.

Proposition 10.1. *The presented branching scheme is complete, i. e., all variables at every leaf of the complete B&B tree are integer.*

Proof. It is directly evident that the variables x and u are integer at the leaves of the B&B tree as we branch on them if and only if they are fractional. On the other hand, the assignment variables ζ are fixed to 0 if they occur in any of the constraints of type (10.7) or (10.9). However, these constraints do not explicitly forbid fractional values of the remaining $\zeta_{s\tau}$ variables.

To exclude this, we assume the original assignment variables z_{st} are integer but there exists a BS s and a set τ_1 with $\zeta_{s\tau_1}$ fractional. By integrality of z_{st}, it holds that $z_{st} = 1$ for all $t \in \tau_1$. Since

$$z_{st} = \sum_{\tau \in \mathrm{T}_s : t \in \tau} \zeta_{s\tau} = 1,$$

there must exist at least one set $\tau_2 \neq \tau_1$ containing t with $\zeta_{s\tau_2}$ fractional. Without loss of generality, $\tau_1 \setminus \tau_2 \neq \emptyset$ and $\zeta_{s\tau_1} + \zeta_{s\tau_2} = 1$ (for $\zeta_{s\tau_1} + \zeta_{s\tau_2} < 1$, we replace $\zeta_{s\tau_2}$ by the sum over all assignment variables $\zeta_{s\tau}$ with $\tau \in \mathrm{T}_s$, $t \in \tau$ and $\tau \neq \tau_1$). For every $t' \in \tau_1 \setminus \tau_2$ there must exist (at least) one set τ_3 containing t' with $\zeta_{s\tau_3}$ fractional but $t \notin \tau_3$. But then

$$\sum_{\tau \in \mathrm{T}_s : t \in \tau} \zeta_{s\tau} + \zeta_{s\tau_3} > 1,$$

which violates constraint (10.3d) for BS s as $x_s \leq 1$, a contradiction. □

10.2. Performance Improvements

A straightforward implementation of the B&P algorithm presented in the previous section does not give satisfying results. Even small test instances cannot be solved to optimality. Hence, we investigate several techniques to improve and to speed up the column generation process for the d-RWNPP in this section.

10.2.1. General Settings

In this subsection, we present general settings which we use for all computations.

Initial solution For the initialisation of the column generation approach, a (dual) feasible initial solution of the MP (10.3) is required (or infeasibility of the MP must be proven). The quality of the initial solution impacts the dual solution of the initial RMP and thus, also the quality of the lower and upper bounds for the optimal solution of the MP.

A promising initial solution can be computed by means of the LP relaxation of the compact formulation (9.3). Denote by $(x^{\mathrm{LP}}, u^{\mathrm{LP}}, z^{\mathrm{LP}})$ the optimal LP solution of the compact formulation. For every BS $s \in \mathcal{S}$ with $x_s^{\mathrm{LP}} \neq 0$, we sort the set of TNs with $z_{st}^{\mathrm{LP}} \neq 0$ such that $z_{s0}^{\mathrm{LP}} \geq z_{s1}^{\mathrm{LP}} \geq \ldots$. For some $j \geq 1$ it holds $z_{st}^{\mathrm{LP}} = x_s^{\mathrm{LP}}$ for the first j TNs (starting from 0) due to the vub constraints (9.4). Hence, let τ denote the set of these TNs:

$$\tau := \left\{ t \in \mathcal{T}_s \,\middle|\, z_{st}^{\mathrm{LP}} = x_s^{\mathrm{LP}} \right\} = \{0, \ldots, j-1\}.$$

It holds $\tau \in \mathrm{T}_s$. Then, we consider the next TN j with $z_{sj}^{\mathrm{LP}} < x_s^{\mathrm{LP}}$. If the Γ-robust capacity constraint

$$\sum_{t \in \tau \cup \{j\}} \frac{\bar{w}_t}{e_{st}} + \max_{\mathcal{T}' \subseteq \tau \cup \{j\}, |\mathcal{T}'| \leq \Gamma} \sum_{t \in \mathcal{T}'} \frac{\hat{w}_t}{e_{st}} \leq b_s \tag{10.10}$$

is still valid, we add this TN to the set τ, i. e., $\tau = \tau \cup \{j\}$. We add the subsequent TNs one by one as long as the BS capacity is not exceeded. By this means, we create one appropriate and as large as possible initial column per BS $s \in \mathcal{S}$. Note that the corresponding solution value in general differs from the LP solution value.

Absolute gap limit The pricing routine of a B&P algorithm stops if no further variables with negative reduced cost exist. However, it is possible that a primal bound found earlier is already an optimal solution. We introduce the following gap limit to stop the solving process before no variables with negative reduced cost exist anymore. Let PB and DB denote the primal and dual bound of the RMP, respectively. If $|\mathrm{PB} - \mathrm{DB}| <$ `abs_gap`, with `abs_gap`:= $\gcd(\min_{s \in \mathcal{S}} c_s, \lambda)$ for integer values of c_s and λ and gcd denotes the greatest common divisor, then there cannot lie another integer solution between PB and DB and we can stop the solving process. This reduces the tailing-off effect; cf. Vanderbeck [174]. Note, this absolute gap limit is automatically known by any solver for the compact formulation (9.3) since all variables of this model are present in every step of the solution process.

Aging Since we compute many columns, it is possible that pricing variables are discarded by cutting off subtrees and hence, are not needed during the subsequent solving process. Therefore, we mark the pricing variables as "removable" so that the corresponding column can be removed from the LP due to aging or cleanup which is automatically performed by the branch-and-price-and-cut framework SCIP [3] which we use in our computational study in Section 10.3.

Cutting planes As mentioned before, the PPs (10.6) are Γ-RKPs and hence, we can apply the (strengthened) extended Γ-robust cover inequalities (8.29). We use the adapted greedy heuristic by Klopfenstein and Nace [111] to separate these inequalities; compare Algorithm 5 without a variable right hand side.

Set extension In the first pricing rounds, many dual variables β_t have a value equal to 0 since the RMP contains only very few columns providing poor dual information. Vanderbeck [174] calls this the *heading-in effect*. A TN t with $\tilde{\beta}_t = 0$ is not considered in the PPs. Hence, the first sets computed in the PPs have a low cardinality. This is why we extend the computed sets of TNs as follows. Assume, a PP for BS s has computed a subset of TNs τ. For all $t \in \mathcal{T}_s \setminus \tau$ with $\beta_t = 0$, we include t in τ if the set $\tau \cup \{t\}$ does not violate the Γ-robust capacity constraint (10.10). Note, we assume an arbitrary ordering of the TNs here. This extension is performed for every computed set of TNs in every B&B node. As a consequence, it is possible that a TN is assigned to more than one BS in an optimal solution. However, we can just drop the redundant assignments afterwards.

All described enhancements are implemented by default and we refer to the B&P algorithm as `simple` henceforth.

10.2.2. The Lagrangian Bound

To evaluate the quality of the current LP solution found by the column generation algorithm, we apply the so-called *Lagrangian bound*; see Desaulniers et al. [64]. Let $Z^{\text{LP}}_{\text{MP}}$ denote the optimal objective value of the LP relaxation of the MP (10.3), $Z^{\text{LP}}_{\text{RMP}}$ of the current LP relaxation of the RMP and let $\tilde{Z}_s$ be the optimal objective value of the PP (10.6) corresponding to BS $s \in \mathcal{S}$. Obviously, every optimal solution of the current RMP yields an upper bound for the MP. Thus,

$$Z^{\text{LP}}_{\text{MP}} \leq Z^{\text{LP}}_{\text{RMP}}.$$

Further, denote by

$$\kappa^\star = \min_{s \in \mathcal{S}} \left(\tilde{\alpha}_s - \tilde{Z}_s \right)$$

the minimum reduced cost in the current pricing round regarding all BSs. If $\kappa^\star \geq 0$, there does not exist a variable $\zeta_{s\tau}$ with negative reduced cost and the optimal solution of the current RMP is also an optimal solution of the MP.

Furthermore, we know that at most $|\mathcal{S}|$ many variables $\zeta_{s\tau}$ are set to one in an optimal solution of the MP since at most one subset of TNs can be assigned to each BS. This leads to the upper bound

$$\sum_{s \in \mathcal{S}} \sum_{\tau \in \mathrm{T}_s} \zeta_{s\tau} \leq |\mathcal{S}|.$$

Based on the previous observations, we can derive a lower bound on $Z^{\text{LP}}_{\text{MP}}$ for $\kappa^\star < 0$.

$$Z^{\text{LP}}_{\text{RMP}} + |\mathcal{S}|\kappa^\star \leq Z^{\text{LP}}_{\text{MP}} \tag{10.11}$$

This bound states that we cannot reduce the optimal objective value of the current RMP by more than $|\mathcal{S}|$ times the minimum reduced cost. This lower bound is called the *Lagrangian bound*.

Let ξ denote the cardinality of a maximum independent set in the conflict graph G. At most ξ many BSs can be deployed at the same time. Consequently, we can enhance the lower bound (10.11) by replacing $|\mathcal{S}|$ with ξ.

The Lagrangian bound is used to speed up computations at B&B nodes, in particular at the root node. Given a value `gap`, we leave the current node if $-\xi\kappa^\star < \texttt{gap}$. To avoid extensive fluctuations, we update the value of the Lagrangian bound only if the current value $-\xi\kappa^\star$ is higher. This results in a stepwise behaviour of the Lagrangian bound. Moreover, in case $Z^{\text{LP}}_{\text{RMP}}$ is a multiple of $\gcd(\min_{s\in\mathcal{S}} c_s, \lambda)$, and $\texttt{gap} \leq \gcd(\min_{s\in\mathcal{S}} c_s, \lambda)$, there do not exist integer solutions having a value less than $Z^{\text{LP}}_{\text{RMP}}$ and this value is a lower bound. Therefore, we set `gap=abs_gap` in our computational study.

We denote the B&P algorithm which includes the settings included in `simple` and the Lagrangian bound as a stop criterion by `LB` henceforth.

10.2.3. Acceleration of the PPs

The time spent solving the PPs has significant influence on the overall performance of the B&P algorithm. For example, it impacts the number of visited B&B nodes. Therefore, we present several techniques to speed up the PPs in this subsection.

Stabilisation As explained, e. g., in Leitner et al. [122] and Vanderbeck [174], column generation suffers from several drawbacks such as slow convergence (*tailing-off effect*), generation of irrelevant columns mainly in the first iterations (*heading-in effect*), and *primal degeneracy* entailing multiple dual optimal solutions. These computational instabilities cause long running times with many iterations. Many stabilisation techniques have been developed to diminish these drawbacks; see Lübbecke and Desrosiers [127] for an overview.

We focus on the primal degeneracy of the RMP and apply stabilisation using alternative dual optimal solutions as described in Leitner et al. [122]. The dual of the LP relaxation of the RMP, the Restricted Dual Problem (RDP), can be stated as the following LP.

$$\max \sum_{t\in\mathcal{T}} \beta_t - \sum_{\mathcal{U}\subseteq\mathcal{S}:\text{ max. clique}} \gamma_{\mathcal{U}} \tag{10.12a}$$

$$\text{s.t.} \sum_{t\in\tau} \beta_t - \alpha_s \leq 0 \qquad \forall\, s\in\mathcal{S}, \tau\in \mathrm{T}'_s \tag{10.12b}$$

$$\beta_t \leq \lambda \qquad \forall\, t\in\mathcal{T} \tag{10.12c}$$

$$\alpha_s - \sum_{\mathcal{U}\subseteq\mathcal{S}:\text{ max. clique},\, s\in\mathcal{U}} \gamma_{\mathcal{U}} \leq c_s \qquad \forall\, s\in\mathcal{S} \tag{10.12d}$$

$$\alpha_s,\ \beta_t,\ \gamma_{\mathcal{U}} \geq 0 \qquad \forall\, s\in\mathcal{S},\ t\in\mathcal{T},\ \mathcal{U}\subseteq\mathcal{S}:\ \text{max. clique} \tag{10.12e}$$

with $\gamma_{\mathcal{U}}$ denoting the dual variables which correspond to the maximal clique inequalities (10.3c). Further, constraints (10.12b) correspond to variables $\zeta_{s\tau}$, (10.12c) to u_t and constraints (10.12d) to x_s. For a dual optimal solution $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma})$, we define the *dual slack* of variable x_s as follows.

$$\Delta_s := c_s - \tilde{\alpha}_s + \sum_{\mathcal{U} \subseteq \mathcal{S}\text{: max. clique, } s \in \mathcal{U}} \tilde{\gamma}_{\mathcal{U}}$$

For every dual optimal solution $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma})$, there exists another optimal solution $(\alpha^\star, \beta^\star, \gamma^\star)$ of the RDP computed as

$$\begin{aligned}
\alpha_s^\star &= \tilde{\alpha}_s + \Delta_s = c_s + \sum_{\mathcal{U} \subseteq \mathcal{S}\text{: max. clique, } s \in \mathcal{U}} \tilde{\gamma_{\mathcal{U}}}, \\
\beta^\star &= \tilde{\beta}, \\
\gamma^\star &= \tilde{\gamma}.
\end{aligned}$$

The solution $(\alpha^\star, \beta^\star, \gamma^\star)$ is optimal for the RDP since the objective value is not changed, constraints (10.12d) are fulfilled with equality and constraints (10.12b) are satisfied more conspicuously. Obviously, $\alpha_s^\star \geq \tilde{\alpha}_s$. Hence, if we compare the objective value of the PP to $\alpha_s^\star$ instead of $\tilde{\alpha}_s$ the comparison becomes more restrictive and the generated columns are more likely to be relevant.

Even though the variables β_t are of greater importance since they occur as cost in the objective function of the PPs, we cannot increase their value in a given dual optimal solution by adding the dual slack of variable u_t. This would on the one hand increase the objective value of the RDP and on the other hand we could not guarantee the compliance of constraints (10.12b) anymore.

Note, the stabilisation is only performed at the root node since the dual values of the branching constraints are typically unknown and cannot be computed easily in the subsequent B&B nodes.

We denote the B&P algorithm which includes the settings of algorithm `LB` and the stabilisation approach by `PP_stab` henceforth.

Suboptimal solving of PPs The proof that a primal solution of a PP is optimal can be rather time consuming for larger test instances. Hence, we stop the solving process of a PP if the optimality gap is less than 1 %. Furthermore, we restrict the number of B&B nodes per PP to 500. If the gap limit is not reached within the first 500 B&B nodes, the gain of solving this PP any further is not sufficient to justify the additional time consumption.

In the case that we have not found any new variable with negative reduced cost at the current B&B node, we solve all PPs again without a gap limit and without the restriction on the number of B&B nodes. By this means, we can guarantee that we have not missed to compute a necessary variable at any node.

We denote the algorithm based on `PP_stab` with the suboptimal solving of the PPs by `PP_subopt`.

10.2.4. Limited Number of Added Columns

Per pricing round, we can add up to $|\mathcal{S}|$ many variables which can lead to a high number of total variables. Therefore, it is reasonable to investigate the restriction of the number of added variables per pricing round. We study the restriction to 1, 5, or 10 variables per round. Even though we restrict this number, we solve all (necessary) PPs and sort the computed variables by their reduced cost in ascending order. We then add the variable or the five or ten variables with the most negative reduced cost.

The setting which limits the number of added columns per pricing round is denoted by `added_cols` and includes the setting `PP_subopt`.

10.2.5. A Primal Heuristic

The primal bounds computed during a B&P procedure are in general rather poor. To overcome this drawback we could solve the current RMP to optimality once in a while. However, this would take some time if the solving process has progressed due to the number of columns in the current subset. This is why we develop a primal heuristic which can be called at the end of each B&B node.

We intend to compute a feasible solution $(\hat{x}, \hat{u}, \hat{\zeta})$ from the current LP solution denoted by $(x^{\mathrm{LP}}, u^{\mathrm{LP}}, \zeta^{\mathrm{LP}})$. For this purpose, we define a new set $\mathcal{S}^\star$ of BSs based on x_s^{LP} as follows.

$$\mathcal{S}^\star \coloneqq \left\{ s \in \mathcal{S} \,\middle|\, x_s^{\mathrm{LP}} > 0.5 \right\}.$$

For all BSs not in this set, we fix $\hat{x}_s = 0$ and $\hat{\zeta}_{s\tau} = 0 \; \forall \, s \in \mathcal{S} \setminus \mathcal{S}^\star, \tau \in \mathrm{T}'_s$. Furthermore, we keep all already decided assignments by setting $\hat{\zeta}_{s\tau} = 1 \; \forall \, (s, \tau)$ if $\zeta_{s\tau}^{\mathrm{LP}} = 1$.

To determine the remaining values, we solve the following sub-ILP which is based on the current RMP.

$$\min \sum_{s \in \mathcal{S}^\star} c_s x_s + \lambda \sum_{t \in \mathcal{T}} u_t \tag{10.13a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t^\star} \sum_{\tau \in \mathrm{T}'_s : t \in \tau} \zeta_{s\tau} + u_t = 1 \qquad \forall \, t \in \mathcal{T} \tag{10.13b}$$

$$\sum_{s \in \mathcal{U}} -x_s \geq -1 \qquad \forall \, \mathcal{U} \subset \mathcal{S}^\star, \mathcal{U} \text{ max. clique} \tag{10.13c}$$

$$x_s - \sum_{\tau \in \mathrm{T}'_s} \zeta_{s\tau} \geq 0 \qquad \forall \, s \in \mathcal{S}^\star \tag{10.13d}$$

$$x_s, \; \zeta_{s\tau}, \; u_t \in \{0, 1\} \qquad \forall \, s \in \mathcal{S}^\star, \; \tau \in \mathrm{T}'_s, \; t \in \mathcal{T} \tag{10.13e}$$

To speed up the primal heuristic, we set a limit of one on the number of solutions. Therefore, as soon as an integer solution with a value better than the known primal solution of the RMP is found, the solving process of the sub-ILP is stopped. Based on this solution, we set the remaining values for $(\hat{x}, \hat{u}, \hat{\zeta})$ to the values of (10.13) and add $(\hat{x}, \hat{u}, \hat{\zeta})$ as a new primal solution to the RMP. We denote the B&P algorithm which applies this primal

identifier	# BSs	# TNs	ξ
BP_20_200a	20	200	12
BP_20_200b	20	200	11
BP_20_200c	20	200	12
BP_30_300a	30	300	14
BP_30_300b	30	300	14
BP_30_300c	30	300	15

Table 10.1.: Number of BSs and TNs, and the maximum independent set number of the six test scenarios.

heuristic and includes the setting `added_cols` with the best parameter (to be determined) by `heuristic`.

10.3. Computational Study

In this section, we present a summary of the comprehensive computational study presented in our work [50] which investigates the performance of the techniques illustrated in Section 10.2. First, we briefly describe the considered scenarios and give some information on the general settings for the performed computations. Afterwards, we analyse the gains of the different settings achieved at the root node and for the complete solving process.

10.3.1. The Scenarios

We consider six different scenarios of two dimensions which are generated as described in Section 5.9.1 as well as Section 9.1.2 and displayed in Table 10.1. The conflict graph is established via the minimum distance requirement with $d_{\min} = 500\,\text{m}$. Additionally, we set $e_{\min} = 0.5\,\text{bps/Hz}$. In this computational study, we consider demand values in bit per second (not kbps) and do not round. Additionally, we set $c_s = 4000$ and $\lambda = 1000$. Furthermore, we fix `abs_gap` $= \gcd(4000, 1000) = 1000$. We consider $\Gamma \in \{0, 1, \ldots, 20\}$ for all scenarios where the maximum value is set to 20 since initial computations performed with the compact model (9.3) showed a constant solution for $\Gamma \geq 20$ for all scenarios.

All computations are performed on a Linux machine with 3.40GHz Intel Core i7-2600 processor, a memory limit of 11 GB RAM and a general CPU time limit of two hours. We use SCIP 3.0.0 [3] with CPLEX 12.4 [98] as underlying LP solver. Furthermore, the PPs are directly solved using CPLEX.

The various investigated algorithms are summarised in Table 10.2 based on the descriptions given in Section 10.2.

In the following subsection, we analyse the quality of the LP relaxation and the behaviour of the Lagrangian bound exemplarily for scenario BP_30_300b. We chose this scenario randomly from the group of the three larger scenarios since the root node of BP_20_200a-c is mostly solved too fast to reveal the information we would like to present.

Identifier	Description
`simple`	straightforward B&P, initial solution via compact LP, absolute gap limit, aging of pricing variables, (strengthened) extended cover inequalities for PPs, extended sets of TNs
`LB`	as `simple` and uses the Lagrangian bound as stop criterion at each B&B node
`PP_stab`	as `LB` plus stabilisation at the root node
`PP_subopt`	as `PP_stab` plus solving the PPs suboptimally: gap limit of 1 %, at most 500 B&B nodes
`added_cols`	as `PP_subopt` plus the number of added variables per pricing round is limited
`heuristic`	as `added_cols` plus deploying a primal heuristic

Table 10.2.: Summary of all settings described in Section 10.2.

10.3.2. LP Relaxation and Lagrangian Bound

In Figure 10.2, we display the values of the LP relaxation of the compact model (9.3) and the values of the LP solutions computed at the root node via the column generation algorithm for $\Gamma \in \{0, 1, \ldots, 50\}$. Note, we consider values for Γ of up to 50 here to illustrate the complete behaviour of the LP solutions for this specific scenario even though, we set $\Gamma \leq 20$ for all other computations. For the more complex problems with $6 \leq \Gamma \leq 32$,

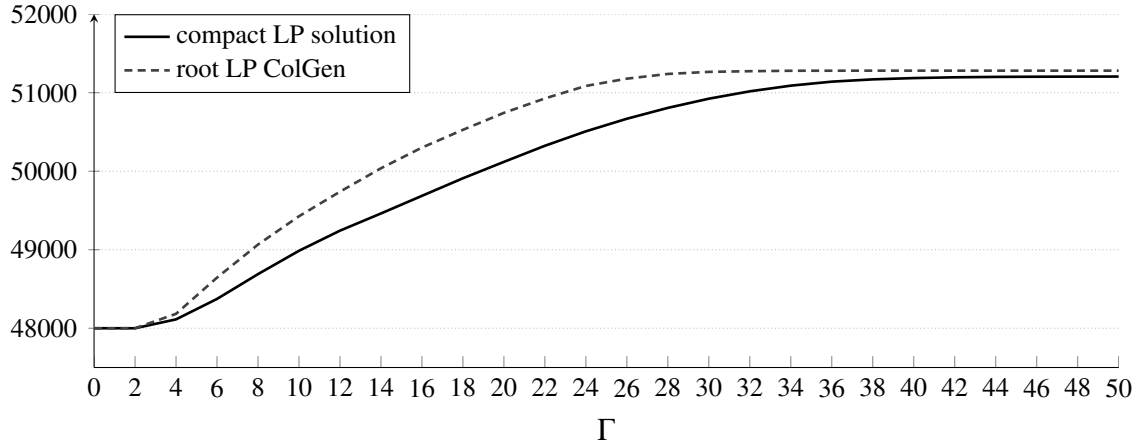
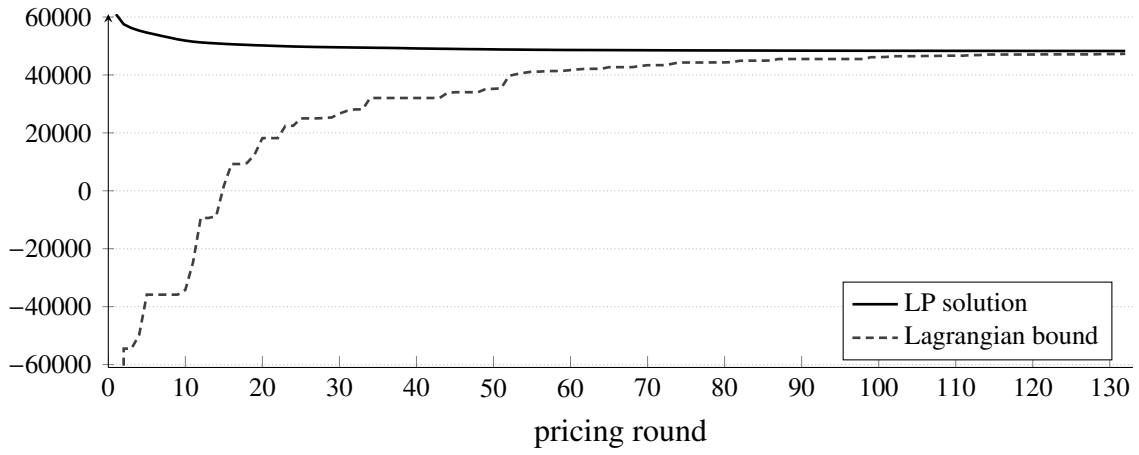
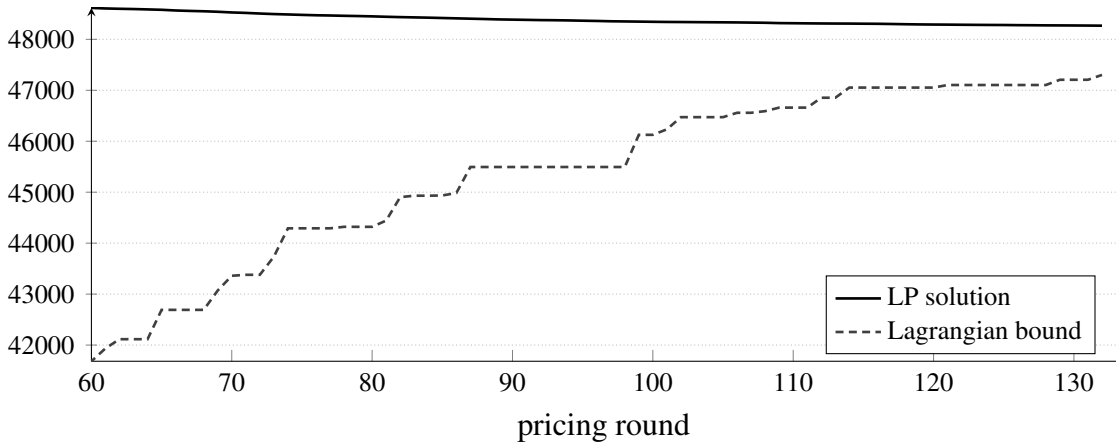


Figure 10.2.: Comparison of the LP solution of the compact model and the LP solution at the root node of the column generation algorithm for scenario BP_30_300b and $\Gamma \in \{0, 1, \ldots, 50\}$.



(a) Complete range.



(b) Zoomed in to rounds 60 to 133.

Figure 10.3.: LP solution and Lagrangian bound per pricing round at the root node for scenario BP_30_300b and $\Gamma = 4$.

the value of the LP solution computed via column generation is significantly better than the value of the LP relaxation of the compact model. By "significantly" we signify the actual values rather than the percentage value since, for example, 50001 is preferred to 49999 as a lower bound due to the fact that the parameters in the objective are multiples of thousand. Concluding, Figure 10.2 demonstrates one reason to apply a B&P algorithm to the d-RWNPP.

Sticking to scenario BP_30_300b, we analyse the behaviour of the Lagrangian bound introduced in Section 10.2.2 in Figure 10.3 exemplarily for $\Gamma = 4$. We display the Lagrangian bound and the value of the current LP solution per pricing round at the root node. For a better readability, we omit the value of the Lagrangian bound at the first round (-611000). As explained in Section 10.2.2, the Lagrangian bound has a stepwise behaviour which becomes apparent in the figure. Furthermore, in the first 60 rounds the

		simple	LB	PP_stab	PP_subopt
BP_20_200a	time	0.0%	0.0%	9.5%	90.5%
	rounds	19.0%	57.1%	19.0%	28.6%
	vars	4.8%	23.8%	38.1%	42.9%
BP_20_200b	time	0.0%	4.8%	14.3%	81.0%
	rounds	28.6%	71.4%	19.0%	14.3%
	vars	23.8%	23.8%	42.9%	38.1%
BP_20_200c	time	0.0%	4.8%	23.8%	71.4%
	rounds	0.0%	71.4%	28.6%	9.5%
	vars	9.5%	28.6%	66.7%	23.8%
total absolute values	time	0	2	10	51
	rounds	10	42	14	11
	vars	8	16	31	22
	total	18	60	55	84

Table 10.3.: Percentage of instances for which each of the four settings simple, LB, PP_stab and PP_subopt gives the best result per scenario and the total absolute values for BP_20_200a–c.

gap between the bound and the value of the LP solution is quite large but it decreases fast in the subsequent pricing rounds; see Figure 10.3(b). As we stop the solving of the root node if the gap between the value of the LP solution and the Lagrangian bound is less than abs_gap = 1000, there is still a gap between the two curves at the last pricing round. The application of the Lagrangian bound as a stop criterion reduces the number of pricing rounds needed to solve the root node from 692 to 133 rounds for the present scenario and Γ value. Therefore, the Lagrangian bound has also the potential to reduce the time spent solving the root node.

10.3.3. Performance of Column Generation at the Root Node

In this subsection, we analyse the performance of the different settings summarised in Table 10.2 for the column generation algorithm at the root node. The aspects we investigate are time consumption, number of pricing rounds and number of computed variables. Detailed results for scenario BP_30_300b can be found in our work [50]. We give a general overview here.

First, we compare the four settings simple, LB, PP_stab, and PP_subopt by counting the number of cases in which each setting gives the best result per scenario. "Best" regards either the lowest time, number of rounds, or number of variables. If two settings give the same best result, the counters for both settings are increased. The results for scenarios BP_20_200a–c are given in Table 10.3 and for BP_30_300a–c in Table 10.4.

The last four lines in each table give the total absolute values regardless of the scenarios

		simple	LB	PP_stab	PP_subopt
BP_30_300a	time	0.0%	9.5%	19.0%	71.4%
	rounds	4.8%	66.7%	19.0%	14.3%
	vars	0.0%	4.8%	81.0%	14.3%
BP_30_300b	time	0.0%	38.1%	38.1%	23.8%
	rounds	0.0%	66.7%	28.6%	4.8%
	vars	0.0%	28.6%	66.7%	4.8%
BP_30_300c	time	0.0%	23.8%	23.8%	52.4%
	rounds	0.0%	85.7%	19.0%	0.0%
	vars	0.0%	4.8%	90.5%	4.8%
total absolute values	time	0	15	17	31
	rounds	1	46	14	4
	vars	0	8	50	5
	total	1	69	81	40

Table 10.4.: Percentage of instances for which each of the four settings simple, LB, PP_stab and PP_subopt gives the best result per scenario and the total absolute values for BP_30_300a–c.

in each group. We do not sum over scenarios with different dimensions since they are too diverse in their solving behaviour. Considering the overall total number, the setting PP_subopt performs best for the smaller instances whereas PP_stab is the best setting for the larger scenarios. Regarding the different aspects separately, the setting LB performs the fewest pricing rounds for both types of scenarios while PP_stab adds the fewest number of variables and PP_subopt consumes the least time. These observations are quite intuitive. First, the Lagrangian bound stops the solving process earlier which implies a reduction in the number of pricing rounds. Second, by means of the stabilisation we compute other variables which can be more effective. Hence, fewer variables are necessary. Finally, the suboptimal solving of the PPs saves time at the expense of increasing the number of pricing rounds and computed variables. However, as the time is usually the most restrictive resource in real world applications, we regard PP_subopt as the most appropriate setting. This setting most frequently gives the best result concerning the time independent of the dimension of the scenarios.

The number of computed pricing variables is in general quite high. This leads to the question if it is beneficial to restrict the number of added variables per pricing round. We test the effect of adding at most 1, 5, or 10 variables per round. Again, we count the cases in which each setting gives the best result regarding solving time, number of rounds and number of variables, see Table 10.5 for BP_20_200a–c and Table 10.6 for BP_30_300a–c.

The highest percentage of instances with the lowest number of computed variables is obtained by added_cols with a limit of 1 independent of the dimension of the scenarios. However, the restriction on the number of variables in general deteriorates the solving

		PP_subopt	added_cols 1	added_cols 5	added_cols 10
BP_20_200a	time	57.1%	19.0%	4.8%	19.0%
	rounds	57.1%	14.3%	9.5%	28.6%
	vars	9.5%	81.0%	4.8%	4.8%
BP_20_200b	time	52.4%	4.8%	23.8%	23.8%
	rounds	47.6%	4.8%	23.8%	28.6%
	vars	14.3%	52.4%	33.3%	0.0%
BP_20_200c	time	52.4%	0.0%	4.8%	42.9%
	rounds	52.4%	0.0%	14.3%	52.4%
	vars	0.0%	81.0%	9.5%	9.5%
total absolute values	time	34	5	7	18
	rounds	33	4	10	23
	vars	5	45	10	3
	total	72	54	27	44

Table 10.5.: Percentage of instances for which each of the four settings PP_subopt, added_cols 1, added_cols 5 and added_cols 10 gives the best result per scenario and the total absolute values for BP_20_200a–c.

		PP_subopt	added_cols 1	added_cols 5	added_cols 10
BP_30_300a	time	76.2%	0.0%	0.0%	23.8%
	rounds	76.2%	0.0%	0.0%	23.8%
	vars	9.5%	81.0%	0.0%	9.5%
BP_30_300b	time	57.1%	0.0%	4.8%	38.1%
	rounds	76.2%	0.0%	4.8%	23.8%
	vars	0.0%	85.7%	14.3%	0.0%
BP_30_300c	time	47.6%	0.0%	9.5%	42.9%
	rounds	66.7%	0.0%	0.0%	33.3%
	vars	0.0%	61.9%	14.3%	23.8%
total absolute values	time	38	0	3	22
	rounds	46	0	1	17
	vars	2	48	6	7
	total	86	48	10	46

Table 10.6.: Percentage of instances for which each of the four settings PP_subopt, added_cols 1, added_cols 5 and added_cols 10 gives the best result per scenario and in the total absolute values.

process at the root node compared to PP_subopt; see the overall total absolute numbers. Hence, we no longer consider the setting added_cols and the setting heuristic is the setting PP_subopt combined with the primal heuristic proposed in Section 10.2.5.

10.3.4. Performance of the B&P Algorithm

For each group of scenarios, where the groups are defined with respect to the dimension, we have considered $63 = 3 \cdot 21$ different settings as $\Gamma \in \{0, 1, 2, \ldots, 20\}$. To give an overview on the performance of the five different settings simple, LB, PP_stab, PP_subopt, and heuristic regarding the complete solving process, we compute the percentage of these 63 instances per group which have at most a certain gap after the time limit of two hours is reached. The results are displayed in Figure 10.4 for the scenarios BP_20_200a–c as well as BP_30_300a–c.

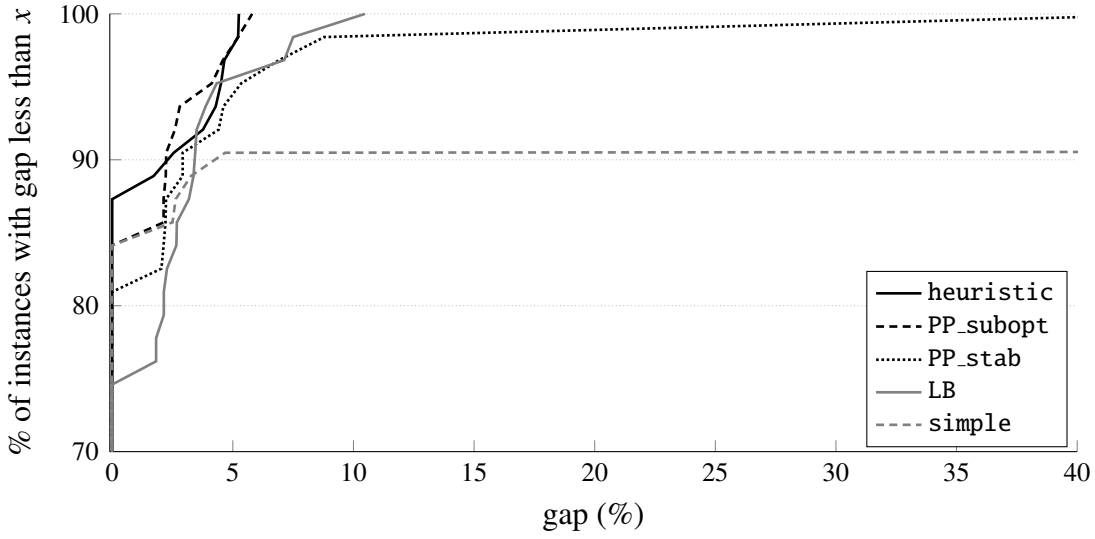
Scenarios BP_20_200b and BP_20_200c can be solved for most values of Γ within a few seconds regardless of which setting. This is the reason why around 75 % of the smaller instances are solved to optimality with any setting; see Figure 10.4(a). Moreover, all instances of the first group solved by the setting heuristic have an optimality gap of less than 5.5 % while all instances of the group of the larger scenarios have an optimality gap of less than 12 % when solved by heuristic. These results are the best we can achieve by the improved B&P algorithm. However, already by means of PP_subopt which is included in heuristic more than 60 % (95 %) of the larger (smaller) instances are solved with a gap less than or equal to 4 %. Concerning the straightforward approach simple, around 43 % (85 %) of the larger (smaller) instances are solved to optimality, whereas an optimality gap could not be computed at all for 30 % (10 %) due to missing dual bounds. Overall, the results demonstrate a clear outperformance of the settings PP_subopt and heuristic compared to simple for computing small gaps within two hours.

In the following section, we investigate the performance of the B&P algorithm with the best settings PP_subopt and heuristic directly compared to the compact formulation (9.3) in a continuative computational study. We perform computations on a selection of scenarios taken from Section 9.1 and on BP_20_200a–c and BP_30_300a–c.

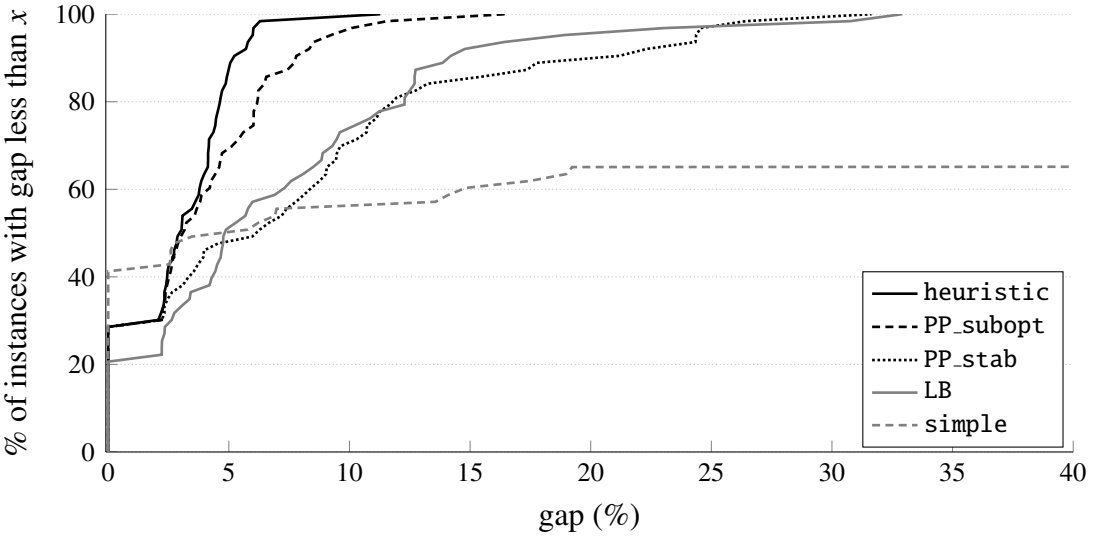
10.4. Numerical Comparison of Compact Formulation and B&P

To compare the compact formulation (9.3) of the d-RWNPP including the improvements presented in Section 9.1.1 to the B&P algorithm with the best settings PP_subopt and heuristic, we run tests for the scenarios with at most 40 BSs presented in Section 9.1 and for all scenarios presented in Section 10.3. We choose the values of Γ regarding the magnitude of the scenarios, i. e., $\Gamma \in \{0, 2, \ldots, 40\}$ and $\Gamma \in \{0, 1, 2, \ldots, 20\}$, respectively. Thus in total, we solve 21 instances for ten scenarios.

All computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of two hours. Additionally, we set a memory limit



(a) BP_20_200a–c



(b) BP_30_300a–c

Figure 10.4.: Percentage of the 63 instances per group of scenarios (BP_20_200a–c or BP_30_300a–c and $\Gamma \in \{0, 1, 2, \ldots, 20\}$) with a gap less than an optimality gap given at the x-axis for a time limit of two hours and the different settings.

of 11 GB. Note, we also use SCIP here to solve the compact formulation (9.3) to obtain comparable results.

In Table 10.7 we display the number of instances per scenario that are solved to optimality with the different algorithms. The compact formulation solves all instances of the scenarios consisting of 20 BSs and half (one third) of the instances with 30 (40) BSs. In contrast, `PP_subopt` and `heuristic` cannot solve any of the instances for the scenarios

identifier	compact	PP_subopt	heuristic
BP_20_200a	21	13	14
BP_20_200b	21	20	20
BP_20_200c	21	21	21
BP_30_300a	11	0	9
BP_30_300b	18	4	6
BP_30_300c	10	8	8
s_20_200	21	0	0
s_30_300	15	0	0
s_40_400	9	0	0
s_40_450	9	0	0

Table 10.7.: Number of instances out of 21 per scenario solved to optimality.

instance	compact	heuristic
BP_20_200a	47.57	904.61
BP_20_200b	1.09	409.10
BP_20_200c	1.33	134.40
BP_30_300a	302.87	2436.15
BP_30_300b	238.61	2656.15
BP_30_300c	447.77	2858.09
s_20_200	79.07	–
s_30_300	988.24	–
s_40_400	1706.31	–
s_40_450	1239.87	–

Table 10.8.: Average solving time for solved instances.

presented in Section 9.1 (identifier "s" and values rounded to kbps). However, the number of solved instances for scenarios presented in Section 10.3 (identifier "BP" and values given in bps) and 20 BSs is reasonable. For these scenarios and 30 BSs, heuristic performs better than PP_subopt but still solves less than half of the instances that can be solved by the compact formulation. In summary, the B&P algorithm with the best setting heuristic is clearly outperformed by the compact formulation in terms of numbers of solved instances.

In Table, 10.8 we display the solving time per scenario averaged over all instances solved to optimality by either the compact formulation or heuristic. The compact formulation is considerably faster than heuristic. This algorithm needs at most 16 % of the average time needed by heuristic and can, in general, solve more instances. The numbers displayed in Table 10.8 further demonstrate that the scenarios taken from Section 9.1 are harder to solve than the scenarios presented in Section 10.3. However, these more difficult scenarios can still be solved in a reasonable amount of time by the compact

formulation if they can be solved to optimality at all within the time limit.

10.5. Conclusion

In this chapter, we have presented a complete B&P approach for the d-RWNPP based on our work [50]. Additionally, we have introduced seven different settings in total to accelerate the solving process and analysed the performance in a computational study executed on six test instances of two dimensions. We have evaluated five settings at the root node and for the complete solving process.

The limitation on the number of added variables per pricing round crystallised to have a negative effect on the solving performance. Though, all other enhancements have in general a positive effect on the solving time, the number of processed pricing rounds and the number of added variables. The suboptimal solving of the PPs performs best at the root node while the primal heuristic which also includes the suboptimal solving of the PPs emerged to be the best setting regarding the complete solving process. Certainly, the judgement of the best performance depends on the focus of the evaluation.

Moreover, we have performed a computational study to compare the best settings for the B&P algorithm to the compact ILP which has demonstrated that even the best setting of the B&P algorithm cannot compete with the (improved) compact formulation (9.3). Nevertheless, for more sophisticated robustness models such as the multi-band robustness investigated in Part IV or the recoverable robustness studied in Part V, the compact model becomes considerably more difficult due to higher numbers of variables and constraints. For such concepts, the presented B&P algorithms might give better results compared to the blown-up compact formulation since the applied robustness approach just affects the PPs and not the complete problem.

Furthermore, an aspect for future investigations on accelerating the B&P algorithm is the solving of the PPs by a dynamic programming algorithm. We present such an algorithm for multi-band RKPs in the subsequent part.

Part IV.

Multi-Band Robustness

11. General Concept

The Γ-robustness concept studied in Part III might estimate the deviations just roughly if the probability of deviation varies noticeably within the deviation interval, which comprises only the extreme values. In such a case, Γ-robustness can lead to a too conservative uncertainty set. To overcome this limitation, the deviation interval can be partitioned into multiple intervals, so-called bands, and the total number of realisations in each band is bounded by one or two parameters per band. Especially in case that historical data on the uncertain coefficients is available, a so-called multi-band uncertainty set can more effectively approximate the distribution of the deviations than a Γ-robust uncertainty set.

A first attempt to use multiple deviation bands was made by Bienstock [22] by means of the so-called histogram model for portfolio optimisation problems in finance. This concept has also been applied to wireless network design problems by Bienstock and D'Andreagiovanni [23] and D'Andreagiovanni [58]. The first theoretical framework of the multi-band robustness concept was presented in Büsing and D'Andreagiovanni [34], [35]. These works comprise fundamental investigations of the properties of this concept in case of a lower as well as upper bound on the number of realisations per band, multi-band robust counterparts and first studies on probabilistic bounds for feasibility guarantees. A good overview on the history of the multi-band concept with lower and upper bounds can be found at the website given in [59]. Shortly after the works [34, 35], Mattia [138] published a multiple interval concept in a technical report. This concept is quite similar to the multi-band robustness but more restricted as the probability distribution of the random variables is assumed to be symmetric and no lower bounds on the number of realisations per interval are assumed. Kutschka [118] studies polyhedral properties of the multi-band RKP and shows that the multi-band RKP polytope is full-dimensional if and only if the highest possible weight of every item does not exceed the knapsack capacity. Additionally, the author states trivial facets of the polytope and derives (extended) cover inequalities, which are valid.

In the following section, we briefly present the basic concept of multi-band robustness based on [34].

11.1. The Basic Principle

The following assumptions on the vector $u \in \mathbb{R}^n$ of uncertain data with $n \in \mathbb{N}$ are necessary to define the concept of multi-band robustness.

- Each uncertain entry u_j of u can be modelled as an independent and bounded random variable.

- The distribution of each random variable is unknown but a realisation is equal to the summation of a nominal value $\bar{u}_j$ and a deviation lying in the range $[\hat{u}_j^{K^-}, \hat{u}_j^{K^+}]$, where $\hat{u}_j^{K^-}, \hat{u}_j^{K^+} \in \mathbb{R}$ represent the maximum negative and positive deviations with $K^-,\ K^+ \in \mathbb{Z}$.

In the multi-band robustness, the deviation interval $[\hat{u}_j^{K^-}, \hat{u}_j^{K^+}]$ of each uncertain coefficient u_j is partitioned into $K = -K^- + K^+ + 1$ disjoint bands based on K deviation values:

$$-\infty < \hat{u}_j^{K^-} < \ldots < \hat{u}_j^{-1} < \hat{u}_j^0 = 0 < \hat{u}_j^1 < \ldots < \hat{u}_j^{K^+} < \infty.$$

A band $k \in \{K^-, \ldots, -1\}$ of the set of negative deviation bands corresponds to a range $\left[\hat{u}_j^k, \hat{u}_j^{k+1}\right)$ while a band $k \in \{1, \ldots, K^+\}$ of the set of positive deviation bands corresponds to a range $\left(\hat{u}_j^k, \hat{u}_j^{k+1}\right]$. Additionally, $k = 0$ denotes the zero-band which is represented by the single value $\hat{u}_j^0$. Thus, a realisation u_j of an uncertain coefficient u_j lies in band $k \in \{K^-, \ldots, K^+\}$ if and only if

$$u_j \in \left[\bar{u}_j + \hat{u}_j^k, \bar{u}_j + \hat{u}_j^{k+1}\right) \qquad \text{for } k \in \{K^-, \ldots, -1\} \tag{11.1a}$$

$$u_j = \bar{u}_j \qquad \text{for } k = 0 \tag{11.1b}$$

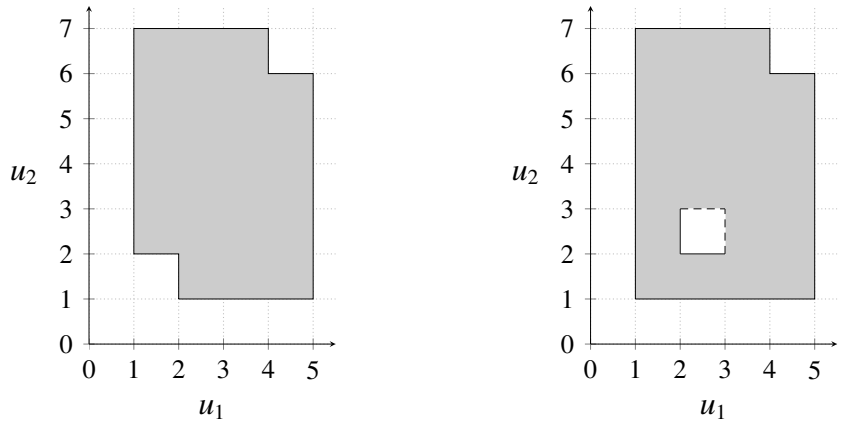
$$u_j \in \left(\bar{u}_j + \hat{u}_j^k, \bar{u}_j + \hat{u}_j^{k+1}\right] \qquad \text{for } k \in \{1, \ldots, K^+\}. \tag{11.1c}$$

For each band k, we define a lower bound γ_k and an upper bound Γ_k on the number of deviations falling in this band with $0 \le \gamma_k \le \Gamma_k \le n$. We assume $\sum_{k=K^-}^{K^+} \gamma_k \le n$ to guarantee the existence of a feasible solution and $\gamma_0 = 0$, $\Gamma_0 = n$ to allow that all values or none deviate.

The multi-band robust uncertainty set can be defined as follows, which represents a slightly modified and corrected version of the uncertainty set given by Kutschka [118].

Definition 11.1. For a vector $u \in \mathbb{R}^n$ of uncertain data with $n \in \mathbb{N}$, define K bands with $\gamma_k, \Gamma_k \in \{0, 1, \ldots, n\}$ for all $k \in \{K^-, \ldots, K^+\}$ and let $\bar{u}_j,\ \hat{u}_j^k \in \mathbb{R}$ such that for every realisation holds (11.1). The *multi-band robust uncertainty set* $\mathcal{U}^{mb}$ is defined as

$$\begin{aligned}\mathcal{U}^{\text{mb}} := \Bigg\{ u \in \mathbb{R}^n \;\Bigg|\; \exists\, \delta_j^k \in \{0,1\}^{Kn} \;:\; & \bar{u}_j + \sum_{k=K^-}^{-1} \delta_j^k \hat{u}_j^k \le u_j < \bar{u}_j + \sum_{k=K^-}^{-1} \delta_j^k \hat{u}_j^{k+1} \\ & \text{or } u_j = \bar{u}_j + \delta_j^0 \hat{u}_j^0, \\ & \text{or } u_j + \sum_{k=1}^{K^+} \delta_j^k \hat{u}_j^k < u_j \le \bar{u}_j + \sum_{k=1}^{K^+} \delta_j^k \hat{u}_j^{k+1}, \\ & \gamma_k \le \sum_{j=1}^{n} \delta_j^k \le \Gamma_k \quad \forall\, k \in \{K^-, \ldots, K^+\}, \\ & \sum_{k=K^-}^{K^+} \delta_j^k = 1 \quad \forall\, j \in \{1, \ldots, n\} \Bigg\}.\end{aligned} \tag{11.2}$$



(a) $\Gamma_{-2} = \Gamma_2 = 1$, $\Gamma_{-1} = \Gamma_1 = 2$ (b) $\Gamma_{-2} = \Gamma_1 = 2$, $\Gamma_{-1} = \Gamma_2 = 1$

Figure 11.1.: Five-band robust uncertainty sets for two combinations of Γ_{-2}, Γ_{-1}, Γ_1, Γ_2. Dashed lines are not included in the set.

We give an example of a multi-band robust uncertainty set in case of two uncertain coefficients which extends the Γ-robust example given in Figure 8.1 on page 117.

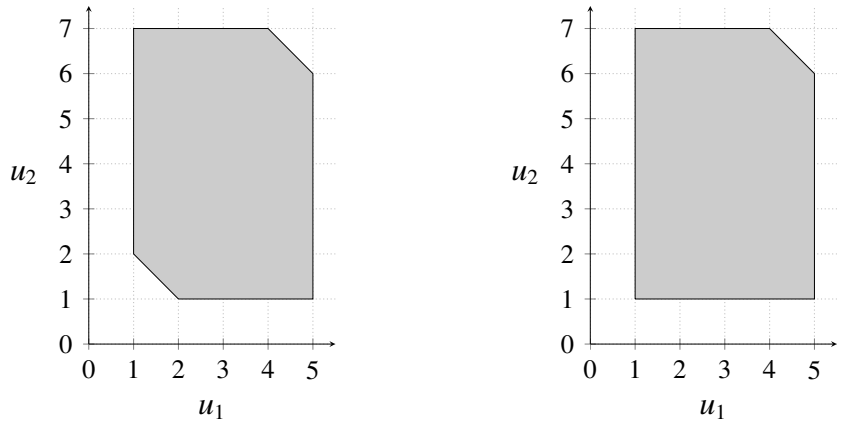
Example 11.2. For a problem with two uncertain coefficients u_1, u_2 and nominal vector $\bar{u} = (3, 4)$, we model five deviation bands with coefficients $\hat{u}^{-2} = (-2, -3)$, $\hat{u}^{-1} = (-1, -2)$, $\hat{u}^0 = (0, 0)$, $\hat{u}^1 = -\hat{u}^{-1}$, and $\hat{u}^2 = -\hat{u}^{-2}$. We assume $\gamma_k = 0$ for all $k \in \{-2, -1, 0, 1, 2\}$ and $\Gamma_0 = 2$. We depict two five-band robust uncertainty sets in Figure 11.1 where the first corresponds to $\Gamma_{-2} = \Gamma_2 = 1$ and $\Gamma_{-1} = \Gamma_1 = 2$ and the second to $\Gamma_{-2} = \Gamma_1 = 2$, $\Gamma_{-1} = \Gamma_2 = 1$.

Note, the five-band uncertainty sets presented in Figure 11.1 are not convex. However, by the multi-band robustness approach, we secure solutions against the convex hull of the uncertainty set. The convex hulls of the uncertainty sets given in Example 11.2 are displayed in Figure 11.2.

We assume henceforth that the left hand side of a linear constraint $ux \leq b$ is affected by uncertainty, where $x \geq 0$. For this type of inequality, positive deviations contribute most to a feasible worst-case realisation of the uncertain data u. Thus, we only focus on positive deviation bands in the remainder of this part. For worst-case realisations which also include negative bands, Büsing and D'Andreagiovanni [35] introduce frequency profiles which determine the number of coefficient realisations u_j per band of a worst-case realisation vector $u \in \mathcal{U}^{\text{mb}}$.

Moreover, we drop the lower bound γ_k on the number of realisations in band k; compare the multiple interval concept by Mattia [138]. The multi-band uncertainty set including only positive bands with upper bounds can be stated as follows.

Definition 11.3. For a vector $u \in \mathbb{R}^n$ of uncertain data with $n \in \mathbb{N}$, define K bands with $\Gamma_k \in \{0, 1, \ldots, n\}$ for all $k \in \{0, \ldots, K\}$ and let $\bar{u}_j$, $\hat{u}_j^k \in \mathbb{R}$ such that for every realisation holds $u_j \in \left(\bar{u}_j + \hat{u}_j^{k-1}, \bar{u}_j + \hat{u}_j^k\right]$ for one band $k \in \{1, \ldots, K\}$ or $u_j = \bar{u}_j + \hat{u}_j^0$. The



(a) $\Gamma_{-2} = \Gamma_2 = 1$, $\Gamma_{-1} = \Gamma_1 = 2$ (b) $\Gamma_{-2} = \Gamma_1 = 2$, $\Gamma_{-1} = \Gamma_2 = 1$

Figure 11.2.: Convex hulls of five-band robust uncertainty sets given in Figure 11.1.

multi-band robust uncertainty set $\mathcal{U}^{mb+}$ with only positive bands and upper bounds is defined as

$$\mathcal{U}^{\text{mb+}} := \left\{ u \in \mathbb{R}^n \,\middle|\, u_j = \bar{u}_j + \delta_j \;\forall\, j \in \{1, \dots, n\},\, \delta \in \mathcal{S}^{\text{mb+}} \right\}, \tag{11.3}$$

with scenarios set

$$\mathcal{S}^{\text{mb+}} := \left\{ \delta \in \mathbb{R}^n \,\middle|\, \delta_j \in \left[\hat{u}_j^0, \hat{u}_j^K\right],\, \left|\left\{ j \in \{1, \dots, n\} \,:\, \delta_j = \hat{u}_j^0 \right\}\right| \leq \Gamma_0, \right.$$
$$\left. \left|\left\{ j \in \{1, \dots, n\} \,:\, \delta_j \in \left(\hat{u}_j^{k-1}, \hat{u}_j^k\right] \right\}\right| \leq \Gamma_k \;\forall\, k \in \{1, \dots, K\} \right\}. \tag{11.4}$$

We call $\delta \in \mathcal{S}^{\text{mb+}}$ a *feasible scenario*.

Note, we use an alternative representation for $\mathcal{U}^{\text{mb+}}$ for a simple introduction of the following dominance relation. Let $\delta, \delta' \in \mathcal{S}^{\text{mb+}}$ be two feasible scenarios. We say δ *dominates* δ' if $\delta_j \geq \delta'_j \;\forall\, j \in \{1, \dots, n\}$. Hence, a solution $x \geq 0$ which is feasible under scenario δ, i. e., $(\bar{u}_j + \delta_j)x_j \leq b_j \;\forall\, j \in \{1, \dots, n\}$, remains feasible for δ'.

The following lemma is a direct consequence of Lemma 1 given in [35].

Lemma 11.4. *Let $\delta \in \mathcal{S}^{mb+}$ be a feasible scenario. If $\delta_j \neq \hat{u}_j^k$ for any $j \in \{1, \dots, n\}$ and $k \in \{0, \dots, K\}$, there exists a feasible scenario $\delta' \in \mathcal{S}^{mb+}$ dominating δ.*

By means of this lemma, it is sufficient to consider only vectors $u \in \mathcal{U}^{\text{mb+}}$ with $u_j \in \{\bar{u}_j + \hat{u}_j^0, \bar{u}_j + \hat{u}_j^1, \dots, \bar{u}_j + \hat{u}_j^K\}$ for the determination of a worst-case realisation.

11.2. The Robust Counterpart

In this section, we investigate the multi-band robust counterpart of a standard LP in row by row representation given in (8.2) assuming only positive deviation bands; cf. [35, 118]

for the general multi-band robust counterpart. The corresponding multi-band robust counterpart which consists of an exponential number of constraints is given as

$$\max\ c^t x \tag{11.5a}$$
$$\text{s.t. } a_i x \le b_i \qquad \forall\, i \in \{1,\dots,m\}, a_i \in \mathcal{U}_i^{\text{mb+}} \tag{11.5b}$$
$$x \in \mathbb{R}^n_{\ge 0}, \tag{11.5c}$$

where $\mathcal{U}_i^{\text{mb+}}$ is the multi-band robust uncertainty set of row i. Similar to the Γ-robustness concept, (11.5) can be reformulated to a formulation with a polynomial number of constraints, which are the most restrictive.

$$\begin{aligned} &\max\ c^t x \\ &\text{s.t. } \max_{a_i \in \mathcal{U}_i^{\text{mb+}}} a_i x \le b_i \qquad \forall\, i \in \{1,\dots,m\} \\ &\quad x \in \mathbb{R}^n_{\ge 0}, \end{aligned} \tag{11.6}$$

This model is non-linear due to the inner maximisation. However, we can linearise it as follows. For row i and fixed variables x, the inner maximisation is equal to

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j + \max \sum_{j=1}^{n} \sum_{k=0}^{K} \hat{a}_{ij}^k x_j z_{ij}^k \tag{11.7a}$$
$$\text{s.t. } \sum_{j=1}^{n} z_{ij}^k \le \Gamma_k \qquad \forall\, k \in \{0,\dots,K\} \tag{11.7b}$$
$$\sum_{k=0}^{K} z_{ij}^k = 1 \qquad \forall\, j \in \{1,\dots,n\} \tag{11.7c}$$
$$z_{ij}^k \in \{0,1\} \qquad \forall\, j \in \{1,\dots,n\},\ k \in \{0,\dots,K\}, \tag{11.7d}$$

where $z_{ij}^k = 1$ if and only if coefficient a_{ij} deviates in band k. The underlying matrix of the maximisation problem in (11.7) is totally unimodular; see [35, 118]. Hence, we can relax the binary constraints (11.7d) to $0 \le z_{ij}^k \le 1$, where the upper bound is dominated by constraints (11.7c) and is thus redundant. For the resulting LP, we introduce dual variables π and ρ for constraints (11.7b) and (11.7c), respectively, and apply strong duality. The dual of the maximisation problem in (11.7) can then be stated as

$$\min \sum_{k=0}^{K} \Gamma_k \pi_i^k + \sum_{j=1}^{n} \rho_{ij} \tag{11.8a}$$
$$\text{s.t. } \pi_i^k + \rho_{ij} \ge \hat{a}_{ij}^k x_j \qquad \forall\, j \in \{1,\dots,n\},\ k \in \{0,\dots,K\} \tag{11.8b}$$
$$\pi_i^k \ge 0 \qquad \forall\, j \in \{1,\dots,n\},\ k \in \{0,\dots,K\} \tag{11.8c}$$
$$\rho_{ij} \text{ free.} \tag{11.8d}$$

Analogously to the compact Γ-robust counterpart developed in Section 8.2, we rewrite the non-linear formulation (11.6) using the dual (11.8) and obtain the *compact multi-band robust counterpart*

$$\max\ c^t x \tag{11.9a}$$

$$\text{s.t.}\ \sum_{j=1}^{n} \bar{a}_{ij} x_j + \sum_{k=0}^{K} \Gamma_k \pi_i^k + \sum_{j=1}^{n} \rho_{ij} \leq b \qquad \forall\, i \in \{1,\dots,m\},\ k \in \{0,\dots,K\} \tag{11.9b}$$

$$\pi_i^k + \rho_{ij} \geq \hat{a}_{ij}^k x_j \qquad \forall\, i \in \{1,\dots,m\},\ j \in \{1,\dots,n\},\ k \in \{0,\dots,K\} \tag{11.9c}$$

$$x,\ \pi \geq 0,\ \rho \text{ free.} \tag{11.9d}$$

This compact formulation possesses $m(n + K + 1)$ additional variables and $nm(K + 1)$ extra constraints. Thus, its size is polynomial in the size of the non-robust LP (8.2). As a consequence, applying multi-band robustness to an (I)LP does not increase its theoretical complexity.

Similar to the separation of robustness cuts for the Γ-robustness approach presented in [72], an alternative formulation to the compact multi-band robust model (11.9) is to solve the non-robust formulation (8.2) and separate violated model constraints as included in (11.5b); see Büsing and D'Andreagiovanni [34].

12. The Multi-Band RKP

The classical KP introduced in Section 1.3 is NP-hard but can be solved in pseudo-polynomial time via a DP with complexity $O(nB)$ as stated in Section 1.3.1. Further algorithms usually first order the items according to non-increasing profit-weight ratios and then derive upper and lower bounds which can be used to fix variables before applying B&B techniques; cf. Pisinger [153] and the references therein.

The Γ-RKP with uncertain weights introduced in Section 8.5 is a generalisation of the KP and hence, also NP-hard. Nevertheless, there exist DPs which solve fairly large instances within reasonable time. Klopfenstein and Nace [110] present a DP which is a modification of the DP for the KP and has complexity $O(nB^2)$ while Monaci et al. [141] develop an algorithm explicitly for the Γ-RKP improving the complexity to $O(\Gamma nB)$. The crucial part of the latter algorithm is the ordering of items according to non-increasing deviation values.

In this chapter, we investigate a further generalisation of the Γ-RKP, the *multi-band RKP*, where multi-band robustness has been introduced in Chapter 11. The presented results are based on our joint work [52]. The multi-band RKP has not been studied intensively. Kutschka [118] proves the full-dimensionality of the associated polytope under certain prerequisites and derives (extended) cover inequalities. In [35, 138], the authors consider a special case of multi-band robustness where uncertainties occur only in the objective coefficients. Mattia [138] shows that the resulting problem can be solved by solving at most $(n + 1)^K$ nominal problems, where K is the number of bands. Since uncertain objective coefficients can be transformed to uncertain weights for the KP, the latter algorithm would yield a $O(n^{K+1}B)$ algorithm. However, as we show by a counterexample in Section 12.2, the stated algorithm is incorrect since the domain for the dual variables in the robust counterpart is not correct. Büsing and D'Andreagiovanni [35] present a corrected version requiring the solving of $O(n^{K^2})$ nominal problems, yielding a $O(n^{K^2}B)$ algorithm for the multi-band RKP.

We develop two DPs. The first algorithm proposed in Section 12.2 has a complexity linear in the number of items but not in the knapsack capacity ($O(nB^{K+1})$). As the capacity can be significantly higher than the number of items, we propose a second DP in Section 12.3, which has a complexity linear in the capacity ($O(K!n^{K+1}B)$). As a side effect, we generalise a result of Bertsimas and Sim [19] on combinatorial optimisation problems with uncertain objective. Additionally in Section 12.4, we develop practical improvements to speed-up the solving process. The performance of the different algorithms is analysed in a computational study in Section 12.5.

In the following section, we briefly present a compact ILP formulation of the multi-band RKP.

12.1. Compact Formulation of the Multi-Band RKP

We first extend the notation used to describe the KP in Section 1.3 and the Γ-RKP in Section 8.5 as follows.

We assume that there exist $K \geq 2$ deviation bands and a realisation $\tilde{w}_i$ of the random variable w_i, which models an uncertain weight, has a deviation value lying in band $k \in \{1, \dots, K\}$ if and only if $\tilde{w}_i \in (\bar{w}_i + \hat{w}_i^{k-1}, \bar{w}_i + \hat{w}_i^k]$, with nominal weight $\bar{w}_i \in \mathbb{Z}_{\geq 0}$ and deviation value $\hat{w}_i^k \in \mathbb{Z}_{\geq 0}$ in band k. Moreover, we presume an increasing ordering among the deviation values for each item $i \in N$, i. e., $\hat{w}_i^k < \hat{w}_i^{k+1}$ for every $k \in \{1, \dots, K-1\}$. The number of possible deviation values lying in band k is limited by a parameter $\Gamma_k > 0$ and $B \in \mathbb{Z}_{\geq 0}$ is presupposed. Note, we consider here a special case of the multi-band robustness concept presented in Chapter 11 as we neglect negative bands. Moreover, we drop the notation of the zero-band but still allow no deviation.

Now, we can formulate the K-band RKP as the following IP with binary decision variables x_i indicating whether item i is included in the knapsack.

$$\max \sum_{i \in N} p_i x_i \tag{12.1a}$$

$$\text{s.t.} \sum_{i \in N} \bar{w}_i x_i + \text{DEV}(x, \hat{w}) \leq B \tag{12.1b}$$

$$x_i \in \{0, 1\} \qquad \forall\, i \in N \tag{12.1c}$$

We maximise the profit in (12.1a) while the knapsack capacity should not be exceeded which is guaranteed by constraint (12.1b). The deviation term DEV is computed via the following ILP for an allocation of the decision variables x, where the binary variables y_i^k define the selection of the items in band k.

$$\text{DEV}(x, \hat{w}) = \max \sum_{i \in N} \sum_{k=1}^{K} \hat{w}_i^k x_i y_i^k \tag{12.2a}$$

$$\text{s.t.} \sum_{i \in N} y_i^k \leq \Gamma_k \qquad \forall\, k \in \{1, \dots, K\} \tag{12.2b}$$

$$\sum_{k=1}^{K} y_i^k \leq 1 \qquad \forall\, i \in N \tag{12.2c}$$

$$y_i^k \in \{0, 1\} \qquad \forall\, i \in N,\ \forall\, k \in \{1, \dots, K\} \tag{12.2d}$$

This formulation maximises the sum of the deviation values on condition that at most Γ_k many deviations can lie in band k; see constraints (12.2b), and the weight of each item can deviate at most in one band; see constraints (12.2c). Note, the possibility of no deviation is implicitly included in the inequality sign of the latter constraints.

Since the polytope described by (12.2) is integral [34], which implies that the underlying matrix is totally unimodular, by strong duality, we can replace $\text{DEV}(x, \hat{w})$ in (12.1) by its

dual problem

$$\min \sum_{k=1}^{K} \Gamma_k \pi_k + \sum_{i \in N} \rho_i \tag{12.3a}$$

$$\text{s.t. } \pi_k + \rho_i \geq \hat{w}_i^k x_i \qquad \forall\, k \in \{1, \ldots, K\},\ \forall\, i \in N \tag{12.3b}$$

$$\pi_k,\ \rho_i \geq 0 \qquad \forall\, i \in N,\ \forall\, k \in \{1, \ldots, K\}, \tag{12.3c}$$

with π_k denoting the dual variables which correspond to constraints (12.2b) and ρ_i corresponding to constraints (12.2c). Note that $\rho_i > 0$ and not free since we have an inequality in constraints (12.2c) in contrary to constraints (11.7c). Including (12.3) in (12.1), we derive the *compact reformulation of the K-band RKP* as follows.

$$\max \sum_{i \in N} p_i x_i \tag{12.4a}$$

$$\text{s.t. } \sum_{i \in N} \bar{w}_i x_i + \sum_{k=1}^{K} \Gamma_k \pi_k + \sum_{i \in N} \rho_i \leq B \tag{12.4b}$$

$$\pi_k + \rho_i \geq \hat{w}_i^k x_i \qquad \forall\, i \in N, k \in \{1, \ldots, K\} \tag{12.4c}$$

$$x_i \in \{0, 1\},\ \pi_k,\ \rho_i \geq 0 \qquad \forall\, i \in N, k \in \{1, \ldots, K\} \tag{12.4d}$$

Using the notation $(\alpha)^+$ to denote $\max\{\alpha, 0\}$, for fixed values of the vectors $x \in \{0,1\}^n$ and $\pi \in \mathbb{R}^K_{\geq 0}$ the optimal value of vector ρ in (12.4) can be computed according to constraints (12.4c)-(12.4d) as

$$\rho_i = \left(\max_{k \in \{1,\ldots,K\}} \left\{ \hat{w}_i^k x_i - \pi_k \right\} \right)^+ = \left(\max_{k \in \{1,\ldots,K\}} \left\{ \hat{w}_i^k - \pi_k \right\} \right)^+ x_i \quad \forall\, i \in \{1, \ldots, n\} \tag{12.5}$$

since $x \in \{0,1\}^n$; see [18]. Furthermore for integer deviation values, optimal dual variables π_k and ρ_i are also integer and obviously, $\pi_k \leq \max_{i \in N} \left\{ \hat{w}_i^k \right\}$ for all bands $k \in \{1, \ldots, K\}$ in an optimal solution.

12.2. A DP for the Multi-Band RKP

In this section, we present an exact DP for the K-band RKP which depicts a straightforward generalisation of the DP for the KP [106, 135] and its generalisation for the Γ-RKP [110].

Applying (12.5), let $f(j, b, \pi)$ denote the highest profit for a given vector π and a feasible solution of ILP (12.4) with weight b, in which only items $\{1, \ldots, j\} \subseteq N$ with $j \in N$ are incorporated assuming an arbitrary ordering. The value, $f(j, b, \pi)$ can be formulated as a classical equality KP

$$f(j, b, \pi) = \max \left\{ \sum_{i=1}^{j} p_i x_i \;\middle|\; \sum_{i=1}^{j} \left(\bar{w}_i + \mathrm{DEV}_\pi(i) \right) x_i = b,\ x_i \in \{0, 1\},\ i \leq j \right\}$$

with total capacity b and

$$\mathrm{DEV}_\pi(i) \coloneqq \left(\max_{k \in \{1,\dots,K\}} \left\{ \hat{w}_i^k - \pi_k \right\} \right)^+ .$$

It must hold $b \in \left\{0, 1, \dots, B - \sum_{k=1}^{K} \Gamma_k \pi_k\right\}$ to fulfil constraint (12.4b). Furthermore, for the dual variables π_k applies

$$\pi_k \in \Pi_k \coloneqq \left\{0, 1, \dots, \max_{i \in N} \left\{\hat{w}_i^k\right\}\right\} \qquad \forall\, k \in \{1, \dots, K\}. \tag{12.6}$$

Similar to the DP described in Section 1.3.1, the DP then consists of the computation of all values of f by the recursive equation

$$f(j, b, \pi) = \max \left\{ f(j-1, b, \pi), f(j-1, b - \left(\bar{w}_j + \mathrm{DEV}_\pi(j)\right), \pi) + p_j \right\} \tag{12.7}$$

with initial values

$$f(1, b, \pi) = \begin{cases} 0, & \text{if } b = 0 \\ p_1, & \text{if } b = \bar{w}_1 + \mathrm{DEV}_\pi(1) \\ -\infty, & \text{otherwise} \end{cases}$$

and $\pi \in \Pi_1 \times \ldots \times \Pi_K$, $b \in \left\{0, 1, \dots, B - \sum_{k=1}^{K} \Gamma_k \pi_k\right\}$. More precisely, $f(j, b, \pi)$ is either equal to the profit when item j is not included or to the profit when j is included. In the second case, the optimal solution value is the sum of the profit p_j of the item and the profit of ILP (12.4) with weight b minus the weight of item j. This means, we can add item j only if there is enough capacity available for its weight.

Formally, the optimal solution value of (12.1) is determined by

$$\max \left\{ f(n, b, \pi) \,\middle|\, \pi \in \Pi_1 \times \ldots \times \Pi_K, b \in \left\{0, 1, \dots, B - \sum_{k=1}^{K} \Gamma_k \pi_k\right\} \right\}. \tag{12.8}$$

Lemma 12.1. *The complexity of the DP described by* (12.7) *and* (12.8)*, which uses the sets* Π_k *for* $k \in \{1, \dots, K\}$ *given in* (12.6)*, is* $O(nB^{K+1})$.

Proof. The computation of

$$\max \left\{ f(n, b, \pi) \,\middle|\, b \in \left\{0, 1, \dots, B - \sum_{k=1}^{K} \Gamma_k \pi_k\right\} \right\}$$

for a fixed vector $\pi \in \Pi_1 \times \ldots \times \Pi_K$ corresponds to solving a (non-robust) KP with a capacity of $B - \sum_{k=1}^{K} \Gamma_k \pi_k$. By (12.7), this can be done in $O(nB)$; see also Algorithm 1 on page 19. Moreover, $|\Pi_k| \leq B + 1$ as $\hat{w}_i^k \leq B$ for all $i \in N$ and $k \in \{1, \dots, K\}$. Hence, the complexity of the presented DP is $O(nB^{K+1})$. □

item	$\hat{w}_i^1 x_i$	$\hat{w}_i^2 x_i$
1	4	8
2	3	7
3	1	3

Table 12.1.: Deviation values for the counterexample.

Note, the complexity $O(nB^{K+1})$ is linear in the number of items n. However, in many applications, the capacity B is larger than the number of items. Hence, an algorithm with complexity linear in the capacity is desirable. For the one-band RKP (Γ-RKP), Monaci et al. [141] derive such a DP. Its crucial assumption is an ordering of the items according to non-increasing deviation values. As no comparable ordering of items with more than one deviation value exists, an extension of this DP to a DP for the problem studied would result in an algorithm with complexity $O(n\Gamma_\ell B^K)$ for one $\ell \in \{1, \dots, K\}$.

Furthermore, conveying the algorithm stated in Mattia [138] for a combinatorial optimisation problem with uncertain objective coefficients to the multi-band RKP, the sets Π_k would be defined as

$$\Pi_k = \tilde{\Pi}_k \coloneqq \left\{ \hat{w}_i^k \middle| i \in N \right\} \cup \{0\} \quad \forall\, k \in \{1, \dots, K\} \tag{12.9}$$

leading to an algorithm with complexity $O((n+1)^{(K+1)}B)$. However, the sets (12.9) are not correct. By means of the following counterexample of a two-band RKP with three items, we show that for an optimal solution of (12.3) $\pi_1 \in \tilde{\Pi}_1$ and $\pi_2 \in \tilde{\Pi}_2$ is not satisfied and hence, these sets cannot be used to compute an optimal solution of the two-band RKP.

The considered deviation values are given in Table 12.1. Further, we set $\Gamma_1 = 2$ and $\Gamma_2 = 1$. Then, the lowest objective value for the dual problem (12.3), which we can compute by using the sets in (12.9), is 13 with corresponding solution $\pi_1 = 0$, $\pi_2 = 3$, $\rho_1 = 5, \rho_2 = 4$, $\rho_3 = 1$. However, the optimal solution value is 12 with $\pi_1 = 0$, $\pi_2 = 4$, $\rho_1 = 4, \rho_2 = 3$, $\rho_3 = 1$ and thus, $\pi_2 \notin \tilde{\Pi}_2$.

In the following section, we present a DP with complexity linear in the capacity which uses modified versions of the sets $\tilde{\Pi}_k$ in (12.9) and which are smaller in magnitude compared to the sets defined in (12.6).

12.3. An Improved DP

In the first part of this section, we derive a DP with complexity linear in the capacity in case the coefficient matrix is subject to uncertainty. Based on the achieved results, we propose a polynomial time algorithm if uncertainty occurs only in the objective coefficients in the subsequent part.

12.3.1. Uncertainty in the Coefficient Matrix

By (12.5), we can restate the dual problem (12.3) as minimising the function χ of π which is given as

$$\chi : \mathbb{R}^K_{\geq 0} \to \mathbb{R} : \chi(\pi) = \sum_{k=1}^{K} \Gamma_k \pi_k + \sum_{i \in N} \mathrm{DEV}_\pi(i). \tag{12.10}$$

Note, we neglect x_i here for simplicity. The inclusion would just cause the usage of a new set $I \coloneqq \{i \in N \mid x_i = 1\}$ instead of N.

Remark 12.2. The function χ is piecewise linear.

Proof. Every term in the maximum of χ is linear. Additionally, χ is convex since linear functions are convex, the maximum of convex functions is convex and the summation of convex functions is also convex. □

Therefore, we know that any local minimum of function χ is also a global minimum. To derive properties of a global minimum for $K > 1$, we use a well-known result for the Γ-RKP, the one-band RKP. Assuming that items are sorted by non-increasing weight deviations $\hat{w}^1_i$, the following lemma holds.

Lemma 12.3. (Monaci et al. [141]) *A subset $I \subseteq N$ is feasible for the one-band RKP if and only if*

$$\sum_{i \in I : i \leq i_{\Gamma_1}} \left(\bar{w}_i + \hat{w}^1_i\right) + \sum_{i \in I : i > i_{\Gamma_1}} \bar{w}_i \leq B$$

where i_{Γ_1} denotes the Γ_1-th item in I if $|I| \geq \Gamma_1$, otherwise i_{Γ_1} is the index of the last item in I.

Corollary 12.4. *For $K = 1$, χ takes a global minimum at $\pi_1 = \hat{w}^1_{\Gamma_1+1}$, where $\hat{w}^1_{\Gamma_1+1}$ denotes the $(\Gamma_1 + 1)$-largest deviation value.*

Proof. We show that $\pi_1 = \hat{w}^1_{\Gamma_1+1}$ is an optimal solution of (12.3) when only items in $I \coloneqq \{i \in N \mid x_i = 1\}$ are considered. By Lemma 12.3, we know that

$$\begin{aligned} &\sum_{i \in I : i \leq i_{\Gamma_1}} \left(\bar{w}_i + \hat{w}^1_i\right) + \sum_{i \in I : i > i_{\Gamma_1}} \bar{w}_i \leq B \\ \Leftrightarrow &\sum_{i \in I} \bar{w}_i + \sum_{i \in I : i \leq i_{\Gamma_1}} \hat{w}^1_i \leq B. \end{aligned}$$

Based on this result, we construct an optimal solution y of the ILP (12.2) with $y_i = 1 \Leftrightarrow i \leq i_{\Gamma_1}$ and objective value $\sum_{i \in I : i \leq i_{\Gamma_1}} \hat{w}^1_i$. Setting $\pi_1 = \hat{w}^1_{\Gamma_1+1}$ it holds for the objective value

of (12.3):

$$\begin{aligned}\Gamma_1\pi_1 + \sum_{i\in I}\left(\hat{w}_i^1 - \pi_1\right)^+ &= \Gamma_1\hat{w}_{\Gamma_1+1}^1 + \sum_{i\in I: i\le i_{\Gamma_1+1}}\left(\hat{w}_i^1 - \hat{w}_{\Gamma_1+1}^1\right)\\ &= \Gamma_1\hat{w}_{\Gamma_1+1}^1 + \sum_{i\in I: i\le i_{\Gamma_1}}\left(\hat{w}_i^1 - \hat{w}_{\Gamma_1+1}^1\right)\\ &= \Gamma_1\hat{w}_{\Gamma_1+1}^1 + \sum_{i\in I: i\le i_{\Gamma_1}}\hat{w}_i^1 - \Gamma_1\hat{w}_{\Gamma_1+1}^1\\ &= \sum_{i\in I: i\le i_{\Gamma_1}}\hat{w}_i^1.\end{aligned}$$

By strong duality, $\pi_1 = \hat{w}_{\Gamma_1+1}^1$ is optimal. Thus, χ takes a global minimum at $\pi_1 = \hat{w}_{\Gamma_1+1}^1$, where $\hat{w}_{\Gamma_1+1}^1$ denotes the (Γ_1+1)-largest deviation value. □

For $K > 1$,we consider the case that all entries of the vector π but one are fixed. Then we limit the domain for the remaining π_ℓ in such a way that $\chi(\pi)$ is as low as possible regarding the fixed entries of π.

For simplicity, we write $k \neq \ell$ instead of $k \in \{1,\dots,K\}\backslash\{\ell\}$ for $\ell \in \{1,\dots,K\}$ henceforth.

Lemma 12.5. *If π_k is fixed for all $k \neq \ell$ with $\ell \in \{1,\dots,K\}$, then the optimal π_ℓ which minimises* (12.10) *is contained in*

$$\left\{\left(\hat{w}_i^\ell - \max_{k\neq\ell}\left\{\left(\hat{w}_i^k - \pi_k\right)^+\right\}\right)^+ \,\middle|\, i \in N\right\}.$$

In fact, the minimum is taken by the $(\Gamma_\ell+1)$-largest of these values.

Proof. If π_k is fixed for all $k \neq \ell$, then

$$\begin{aligned}\chi(\pi) &= \sum_{k\neq\ell}\Gamma_k\pi_k + \Gamma_\ell\pi_\ell + \sum_{i\in N}\left(\max_{k\in\{1,\dots,K\}}\left\{\hat{w}_i^k - \pi_k\right\}\right)^+\\ &= \sum_{k\neq\ell}\Gamma_k\pi_k + \Gamma_\ell\pi_\ell + \sum_{i\in N}\max\left\{\hat{w}_i^\ell - \pi_\ell, \max_{k\neq\ell}\left\{\hat{w}_i^k - \pi_k\right\}, 0\right\}\\ &= \sum_{k\neq\ell}\Gamma_k\pi_k + \sum_{i\in N}\max_{k\neq\ell}\left\{\left(\hat{w}_i^k - \pi_k\right)^+\right\} + \tilde{\chi}(\pi_\ell)\end{aligned}$$

with

$$\begin{aligned}\tilde{\chi}(\pi_\ell) &:= \Gamma_\ell\pi_\ell + \sum_{i\in N}\left(\hat{w}_i^\ell - \pi_\ell - \max_{k\neq\ell}\left\{\left(\hat{w}_i^k - \pi_k\right)^+\right\}\right)^+\\ &= \Gamma_\ell\pi_\ell + \sum_{i\in N}\left(\hat{w}_i^\ell - \max_{k\neq\ell}\left\{\left(\hat{w}_i^k - \pi_k\right)^+\right\} - \pi_\ell\right)^+.\end{aligned}$$

Defining

$$\hat{d}_i^\ell := \left(\hat{w}_i^\ell - \max_{k\neq\ell}\left\{\left(\hat{w}_i^k - \pi_k\right)^+\right\}\right)^+,$$

the function $\tilde{\chi}(\pi_\ell)$ is equal to $\Gamma_\ell \pi_\ell + \sum_{i \in N} \left(\hat{d}_i^\ell - \pi_\ell\right)^+$. This function is the objective of the dual problem (12.3) of a Γ_ℓ-RKP with deviation values $\hat{d}_i^\ell$. Applying Corollary 12.4, we know that an optimal solution is given by setting the π_ℓ to the $(\Gamma_\ell + 1)$-largest deviation value. □

Since every local minimum of the convex function χ as defined in (12.10) is a global minimum, we have to find a point with a gradient equal to zero. If a linear segment with a gradient equalling zero exists, also the corresponding corners have the same property and still constitute global minima of the function χ. At such a corner, χ is not differentiable in one direction, i. e., the directional derivatives differ.

For an index $\ell \in \{1, \dots, K\}$ (an index set $\mathcal{K} \subseteq \{1, \dots, K\}$), we denote by e_ℓ ($e_\mathcal{K}$) the unit vector with all entries set to zero apart from position(s) ℓ ($\mathcal{K}$) which is (are) set to 1.

Theorem 12.6. *If $\pi \in \mathbb{R}^K_{\geq 0}$ is a point of non-differentiability of the function χ with*

$$\nabla_{e_\ell} \chi(\pi) \neq -\nabla_{-e_\ell} \chi(\pi)$$

for an $\ell \in \{1, \dots, K\}$, then there exists at least one $i \in N$ with $\pi_\ell = \hat{w}_i^\ell$ or $\hat{w}_i^\ell - \pi_\ell = \hat{w}_i^l - \pi_l$ for one $l \neq \ell$.

Proof. First, we recall the definition of the directional derivative in the direction e_ℓ and of the negative directional derivative in the direction $-e_\ell$ which are used in this proof.

$$\begin{aligned}
\nabla_{e_\ell} \chi(\pi) &= \lim_{h \searrow 0} \frac{\chi(\pi + h e_\ell) - \chi(\pi)}{h} = \lim_{h \searrow 0} \frac{\chi(\pi_1, \dots, \pi_{\ell-1}, \pi_\ell + h, \pi_{\ell+1}, \dots, \pi_K) - \chi(\pi)}{h} \\
-\nabla_{-e_\ell} \chi(\pi) &= \lim_{h \searrow 0} \frac{\chi(\pi - h e_\ell) - \chi(\pi)}{-h} = \lim_{h \searrow 0} \frac{\chi(\pi_1, \dots, \pi_{\ell-1}, \pi_\ell - h, \pi_{\ell+1}, \dots, \pi_K) - \chi(\pi)}{-h}
\end{aligned}$$

Hence, $\nabla_{e_\ell} \chi(\pi) \neq -\nabla_{-e_\ell} \chi(\pi)$ implies $\chi(\pi + h e_\ell) - \chi(\pi) \neq -\chi(\pi - h e_\ell) + \chi(\pi)$. For $h \to 0$, a difference between $\chi(\pi + h e_\ell)$ and $\chi(\pi)$ can only occur due to differences in the term $\mathrm{DEV}_\pi(i)$ in the definition of χ in (12.10). We now define a partition of the set of items N for point π based on the difference values $\hat{w}_i^k - \pi_k$ which determine $\mathrm{DEV}_\pi(i)$.

$$\begin{aligned}
I_0 &:= \left\{ i \in N \,\middle|\, \hat{w}_i^k - \pi_k < 0 \;\forall\, k \in \{1, \dots, K\} \right\}, \\
I_k &:= \Big\{ i \in N \,\Big|\, \hat{w}_i^k - \pi_k > \hat{w}_i^{k'} - \pi_{k'} \;\forall\, k' \leq k-1,\ \hat{w}_i^k - \pi_k \geq \hat{w}_i^{k'} - \pi_{k'} \;\forall\, k' \geq k+1, \\
&\qquad \text{and } \hat{w}_i^k - \pi_k \geq 0 \Big\} \quad \forall\, k \in \{1, \dots, K\}
\end{aligned}$$

The set I_0 is the set of items for which all difference values (deviation minus dual) are negative. For band k, the set I_k is the set of items for which k is the smallest index whereby $\mathrm{DEV}_\pi(i)$ is defined by the corresponding difference $\hat{w}_i^k - \pi_k$. By definition, the sets I_k for all bands $k \in \{0, 1, \dots, K\}$ are pairwise disjoint defining a partition of $N = I_0 \cup \bigcup_{k=1}^K I_k$. Analogously for the point $\pi + h e_\ell$, we define sets I_0^h and I_k^h by replacing $\hat{w}_i^\ell - \pi_\ell$ by $\hat{w}_i^\ell - (\pi_\ell + h)$ and for point $\pi - h e_\ell$, sets I_0^{-h} and I_k^{-h} by replacing $\hat{w}_i^\ell - \pi_\ell$ by $\hat{w}_i^\ell - (\pi_\ell - h)$.

These sets are also pairwise disjoint and thus,

$$N = I_0 \cup \bigcup_{k=1}^{K} I_k = I_0^h \cup \bigcup_{k=1}^{K} I_k^h = I_0^{-h} \cup \bigcup_{k=1}^{K} I_k^{-h}.$$

The transition from point $\pi + he_\ell$ to point π and then to $\pi - he_\ell$ leads to the following subset relations.

(a) $I_0^{-h} \subseteq I_0 \subseteq I_0^h$,

(b) $I_\ell^h \subseteq I_\ell \subseteq I_\ell^{-h}$,

(c) $I_k^{-h} \subseteq I_k \subseteq I_k^h \quad \forall\, k \neq \ell$.

This means, the changes in the sets by transition, e. g., from $\pi + he_\ell$ to π can only be caused by items shifting from I_ℓ^h to $I_0 \cup \bigcup_{k\neq\ell} I_k$ since $\hat{w}_i^\ell - (\pi_\ell + h) \leq \hat{w}_i^\ell - \pi_\ell$ and all other values remain the same. Hence, the following subset relations are also immediate.

(d) $I_\ell \setminus I_\ell^h \subseteq \left(I_0^h \setminus I_0 \cup \bigcup\limits_{k\neq\ell} (I_k^h \setminus I_k) \right)$,

(e) $I_k^h \setminus I_k \subseteq I_\ell \setminus I_\ell^h \quad \forall\, k \neq \ell$, in particular $\bigcup\limits_{k\neq\ell} \left(I_k^h \setminus I_k \right) \subseteq I_\ell \setminus I_\ell^h$,

(f) $I_\ell^{-h} \setminus I_\ell \subseteq \left(I_0 \setminus I_0^{-h} \cup \bigcup\limits_{k\neq\ell} (I_k \setminus I_k^{-h}) \right)$, and

(g) $\bigcup\limits_{k\neq\ell} \left(I_k \setminus I_k^{-h} \right) \subseteq I_\ell^{-h} \setminus I_\ell$.

For $i \in I_\ell \setminus I_\ell^h$ by (d), either

i) $i \in I_0^h \setminus I_0 \Rightarrow \hat{w}_i^k - \pi_k < 0\ \forall\, k \neq \ell$ but $\hat{w}_i^\ell - \pi_\ell \geq 0$ and $\hat{w}_i^\ell - (\pi_\ell + h) < 0$ hence, for h small $\pi_\ell = \hat{w}_i^\ell$, or

ii) $i \in \bigcup\limits_{k\neq\ell} (I_k^h \setminus I_k) \Rightarrow \hat{w}_i^k - \pi_k \leq \hat{w}_i^\ell - \pi_\ell$ and $\hat{w}_i^k - \pi_k > \hat{w}_i^\ell - (\pi_\ell + h)$ for at least one $k \neq \ell$. Hence, for h small $\hat{w}_i^k - \pi_k = \hat{w}_i^\ell - \pi_\ell$.

Based on these relations, we can now rephrase the numerators of the directional derivatives as follows.

$$\begin{aligned}
& \chi(\pi + he_\ell) - \chi(\pi) \\
= \quad & \sum_{k\neq\ell} \Gamma_k \pi_k + \Gamma_\ell(\pi_\ell + h) + \sum_{i \in \bigcup\limits_{k\neq\ell} I_k^h} \max_{k\neq\ell} \left\{ \hat{w}_i^k - \pi_k \right\} + \sum_{i\in I_\ell^h} (\hat{w}_i^\ell - (\pi_\ell + h)) \\
& - \sum_{k=1}^{K} \Gamma_k \pi_k - \sum_{i \in \bigcup\limits_{k\neq\ell} I_k} \max_{k\neq\ell} \left\{ \hat{w}_i^k - \pi_k \right\} - \sum_{i\in I_\ell} (\hat{w}_i^\ell - \pi_\ell) \\
\overset{\text{(b),(c)}}{=} \quad & \Gamma_\ell h + \sum_{i \in \bigcup\limits_{k\neq\ell} (I_k^h \setminus I_k)} \max_{k\neq\ell} \left\{ \hat{w}_i^k - \pi_k \right\} - |I_\ell^h| h - \sum_{i\in I_\ell \setminus I_\ell^h} (\hat{w}_i^\ell - \pi_\ell) \\
\overset{\text{(e),}i\text{),}ii\text{)}}{=} \quad & (\Gamma_\ell - |I_\ell^h|) h
\end{aligned}$$

By means of an analogue argumentation applying (b),(c),(f), and (g), we obtain

$$\chi(\pi - he_\ell) - \chi(\pi) = (|I_\ell| - \Gamma_\ell)h.$$

In total, we have

$$\begin{aligned}
&\nabla_{e_\ell}\chi(\pi) \neq -\nabla_{-e_\ell}\chi(\pi)\\
\Leftrightarrow &\lim_{h\searrow 0}\frac{\chi(\pi + he_\ell) - \chi(\pi)}{h} \neq \lim_{h\searrow 0}\frac{\chi(\pi - he_\ell) - \chi(\pi)}{-h}\\
\Leftrightarrow &\lim_{h\searrow 0}(\Gamma_\ell - |I_\ell^h|) \neq \lim_{h\searrow 0}(\Gamma_\ell - |I_\ell|)\\
\Leftrightarrow &\lim_{h\searrow 0}|I_\ell^h| \neq |I_\ell|\\
\Leftrightarrow &\exists\, i \in I_\ell \setminus I_\ell^h \text{ for } h \searrow 0\\
\overset{i),ii)}{\Rightarrow} &\pi_\ell = \hat{w}_i^\ell \text{ or } \hat{w}_i^\ell - \pi_\ell = \hat{w}_i^k - \pi_k \text{ for at least one } i \in N \text{ and } k \neq \ell.
\end{aligned}$$

This concludes the proof. □

We have shown now that in an optimal solution of (12.3) either a dual variable π_ℓ is equal to a deviation value $\hat{w}_i^\ell$ or two difference values of the form $\hat{w}_i^k - \pi_k$ for two distinct bands coincide. In the following, we intend to exclude the latter case from the space of optimal solutions but $\pi_k = 0$ remains possible. To this end, we first prove a result for the directional derivatives at a point with the undesirable property in the following lemma.

Lemma 12.7. *Let $\pi^\star \in \mathbb{R}^K_{\geq 0}$ be a point with $\pi_k^\star \notin \left\{\hat{w}_i^k \,\middle|\, i \in N\right\} \cup \{0\}\ \forall\, k \in \{1,\dots,K\}$ and $\hat{w}_j^\ell - \pi_\ell^\star = \hat{w}_j^\varkappa - \pi_\varkappa^\star > 0$, for one $j \in N$ and $\ell, \varkappa \in \{1,\dots,K\}$, $\ell \neq \varkappa$. Furthermore, let $\alpha := \hat{w}_j^\ell - \pi_\ell^\star$ and $\mathcal{K} := \{k \in \{1,\dots,K\} \,|\, \hat{w}_j^k - \pi_k^\star = \alpha\}$ whereas we write $k \notin \mathcal{K}$ instead of $k \in \{1,\dots,K\} \setminus \mathcal{K}$ for simplicity. Similar to the proof of Theorem 12.6, we define a partition of the item set N into two sets, one consisting of all items for which* $\operatorname{argmax}_{k\in K}\left\{\hat{w}_i^k - \pi_k^\star\right\} \in \mathcal{K}$ *and the other one containing the remaining items:*

$$I_\mathcal{K}^\star := \left\{ i \in N \,\middle|\, \max\left\{\max_{k\in\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star\right\}, \max_{k\notin\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star\right\}, 0\right\} = \max_{k\in\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star\right\}\right\}, \quad I^\star := N \setminus I_\mathcal{K}^\star.$$

Then, for the directional derivatives in the direction $e_\mathcal{K}$ at $\pi^\star$ holds

$$\nabla_{e_\mathcal{K}}\chi(\pi^\star) = -\nabla_{-e_\mathcal{K}}\chi(\pi^\star) = \sum_{k\in\mathcal{K}}\Gamma_k - |I_\mathcal{K}^\star|.$$

Proof. In order to define the numerators of the directional derivatives and analog to $N = I_\mathcal{K}^\star \cup I^\star$, we define a partition of the set of items at points $\pi^\star + he_\mathcal{K}$ and $\pi^\star - he_\mathcal{K}$ as

$$\begin{aligned}
I_\mathcal{K}^h &:= \left\{ i \in N \,\middle|\, \max\left\{\max_{k\in\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star - h\right\}, \max_{k\notin\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star\right\}, 0\right\} = \max_{k\in\mathcal{K}}\left\{\hat{w}_i^k - \pi_k^\star - h\right\}\right\},\\
I^h &:= N \setminus I_\mathcal{K}^h,
\end{aligned}$$

and analogously, $I_{\mathcal{K}}^{-h}$ and $I^{-h} := N \setminus I_{\mathcal{K}}^{-h}$. If h is sufficiently small, $I_{\mathcal{K}}^{\star} = I_{\mathcal{K}}^{h} = I_{\mathcal{K}}^{-h}$ and $I^{\star} = I^{h} = I^{-h}$. Then, the numerators can be computed as follows.

$$
\begin{aligned}
&\chi(\pi^{\star} + he_{\mathcal{K}}) - \chi(\pi^{\star}) \\
&= \sum_{k\in\mathcal{K}} \Gamma_k (\pi_k^{\star} + h) + \sum_{k\notin\mathcal{K}} \Gamma_k \pi_k^{\star} + \sum_{i\in I^h} \max_{k\notin\mathcal{K}} \left\{(\hat{w}_i^k - \pi_k^{\star})^+\right\} + \sum_{i\in I_{\mathcal{K}}^h} \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^{\star} - h\right\} \\
&\quad - \sum_{k=1}^{K} \Gamma_k \pi_k^{\star} - \sum_{i\in I^{\star}} \max_{k\notin\mathcal{K}} \left\{(\hat{w}_i^k - \pi_k^{\star})^+\right\} - \sum_{i\in I_{\mathcal{K}}^{\star}} \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^{\star}\right\} \\
&= h \sum_{k\in\mathcal{K}} \Gamma_k + \sum_{k\in\mathcal{K}} \Gamma_k \pi_k^{\star} + \sum_{k\notin\mathcal{K}} \Gamma_k \pi_k^{\star} - \sum_{k=1}^{K} \Gamma_k \pi_k^{\star} + \sum_{i\in I^h} \max_{k\notin\mathcal{K}} \left\{(\hat{w}_i^k - \pi_k^{\star})^+\right\} + \sum_{i\in I_{\mathcal{K}}^h} \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^{\star}\right\} \\
&\quad - \sum_{i\in I_{\mathcal{K}}^h} h - \sum_{i\in I^{\star}} \max_{k\notin\mathcal{K}} \left\{(\hat{w}_i^k - \pi_k^{\star})^+\right\} - \sum_{i\in I_{\mathcal{K}}^{\star}} \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^{\star}\right\} \\
&= h \sum_{k\in\mathcal{K}} \Gamma_k - |I_{\mathcal{K}}^h| h,
\end{aligned}
$$

and

$$
\chi(\pi^{\star} - he_{\mathcal{K}}) - \chi(\pi^{\star}) = |I_{\mathcal{K}}^{\star}| h - h \sum_{k\in\mathcal{K}} \Gamma_k.
$$

Hence, for the directional derivatives holds

$$
\begin{aligned}
\lim_{h\searrow 0} \frac{\chi(\pi^{\star} + he_{\mathcal{K}}) - \chi(\pi^{\star})}{h} &= \lim_{h\searrow 0} \frac{h \sum\limits_{k\in\mathcal{K}} \Gamma_k - |I_{\mathcal{K}}^h| h}{h} = \lim_{h\searrow 0} \left(\sum_{k\in\mathcal{K}} \Gamma_k - |I_{\mathcal{K}}^h| \right) \\
= \lim_{h\searrow 0} \left(\sum_{k\in\mathcal{K}} \Gamma_k - |I_{\mathcal{K}}^{\star}| \right) &= \lim_{h\searrow 0} \frac{|I_{\mathcal{K}}^{\star}| h - h \sum\limits_{k\in\mathcal{K}} \Gamma_k}{-h} = \lim_{h\searrow 0} \frac{\chi(\pi^{\star} - he_{\mathcal{K}}) - \chi(\pi^{\star})}{h}.
\end{aligned}
$$

□

Using this result, we can now show that there exists an optimal solution of (12.3) in which no two difference values for two distinct bands coincide.

Lemma 12.8. *For a point $\pi^{\star} \in \mathbb{R}_{\geq 0}^K$ with $\pi_k^{\star} \notin \left\{\hat{w}_i^k \,\middle|\, i \in N\right\} \cup \{0\}$ $\forall\, k \in \{1, \ldots, K\}$ and $\hat{w}_j^{\ell} - \pi_{\ell}^{\star} = \hat{w}_j^{\varkappa} - \pi_{\varkappa}^{\star} > 0$, for one $j \in N$ and $\ell, \varkappa \in \{1, \ldots, K\}$, $\ell \neq \varkappa$, there exists a point $\pi \in \mathbb{R}_{\geq 0}^K$ with $\pi_k \in \left\{\hat{w}_i^k \,\middle|\, i \in N\right\} \cup \{0\}$ for at least one $k \in \{1, \ldots, K\}$ and $\chi(\pi) \leq \chi(\pi^{\star})$.*

Proof. We use α, $\mathcal{K}$ and the partitioning sets $I_{\mathcal{K}}^{\star}$ and $I^{\star}$ as defined in Lemma 12.7, whereby we know that

$$
\nabla_{e_{\mathcal{K}}} \chi(\pi^{\star}) = -\nabla_{-e_{\mathcal{K}}} \chi(\pi^{\star}) = \sum_{k\in\mathcal{K}} \Gamma_k - |I_{\mathcal{K}}^{\star}|
$$

holds for the directional derivatives in the direction $e_{\mathcal{K}}$ at $\pi^{\star}$.

If the derivative is non-positive, we now increase all $\pi_k^{\star}$, $k \in \mathcal{K}$ simultaneously until we reach a point $\pi \in \mathbb{R}_{\geq 0}^K$ with $\pi_{\ell} \in \{\hat{w}_i^{\ell} \,|\, i \in N\} \cup \{0\}$ for (at least) one $\ell \in \{1, \ldots, K\}$. On the

other hand, if the derivative is (strictly) positive, we decrease all $\pi_k^\star$, $k \in \mathcal{K}$ simultaneously until we reach a point $\pi \in \mathbb{R}^K_{\geq 0}$ with $\pi_\ell \in \{\hat{w}_i^\ell \mid i \in N\}$ or $\pi_\ell = 0$ for (at least) one $\ell \in \{1, \dots, K\}$.

In case of a non-positive derivative, i. e, $\sum_{k\in\mathcal{K}} \Gamma_k - |I^\star_{\mathcal{K}}| \leq 0$, we define the minimum value possible to add to $\pi_k^\star$ without changing the partition of the item set as

$$\varepsilon^+ := \min \left\{ \min_{\substack{k \in \mathcal{K}, i \in N: \\ \hat{w}_i^k - \pi_k^\star > 0}} \left\{ \hat{w}_i^k - \pi_k^\star \right\}, \min_{\substack{k \in \mathcal{K}, k' \notin \mathcal{K}, i \in N: \\ \hat{w}_i^k - \pi_k^\star > \hat{w}_i^{k'} - \pi_{k'}^\star}} \left\{ \hat{w}_i^k - \pi_k^\star - \hat{w}_i^{k'} + \pi_{k'}^\star \right\} \right\}.$$

ε^+ describes the minimum difference between a dual variable $\pi_k^\star$ and a deviation value or between two difference values. Note, this minimum exists since $\hat{w}_j^k - \pi_k^\star = \alpha > 0$. Now, we define

$$\pi_k := \begin{cases} \pi_k^\star + \varepsilon^+ & k \in \mathcal{K} \\ \pi_k^\star & k \notin \mathcal{K}. \end{cases}$$

If ε^+ is defined by the first minimum, the point π has the required property that $\pi_k \in \{\hat{w}_i^k \mid i \in N\} \cup \{0\}$. Otherwise, $\varepsilon^+ = \hat{w}_i^\ell - \pi_\ell^\star - \hat{w}_i^{k'} + \pi_{k'}^\star$ for one $i \in N$, $\ell \in \mathcal{K}$, $k' \notin \mathcal{K}$ and we define $\alpha := \hat{w}_i^\ell - \pi_\ell$ and a new set $\mathcal{K}' := \left\{k \in K \mid \hat{w}_i^k - \pi_k = \alpha\right\}$. Then we start again from the beginning repeating the same steps. We continue this procedure until one π_k has the desired property.

What remains to show is that the objective is not increased when increasing some $\pi_k^\star$ by ε^+. For that purpose, we first show that the set $I^\star_{\mathcal{K}}$ remains the same. So, we define a set $I_{\mathcal{K}}$ analogously to $I^\star_{\mathcal{K}}$ by replacing every $\pi_k^\star$ by π_k. By the definition of these two sets, it holds $I_{\mathcal{K}} \subseteq I^\star_{\mathcal{K}}$. We now show $I^\star_{\mathcal{K}} \subseteq I_{\mathcal{K}}$, i. e., these sets are equal. To this end, we consider an item $i \in I^\star_{\mathcal{K}}$ and show $i \in I_{\mathcal{K}}$. We have

- $\varepsilon^+ \leq \hat{w}_i^k - \pi_k^\star - \hat{w}_i^{k'} + \pi_{k'}^\star$ for all $i \in I^\star_{\mathcal{K}}$, $k \in \mathcal{K}$ and $k' \notin \mathcal{K}$:

$$\begin{aligned} \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k\right\} &= \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^\star - \varepsilon^+\right\} \\ &\geq \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^\star - \hat{w}_i^k + \pi_k^\star + \hat{w}_i^{k'} - \pi_{k'}^\star\right\} \quad \forall\, k' \notin \mathcal{K} \\ &= \hat{w}_i^{k'} - \pi_{k'}^\star \quad \forall\, k' \notin \mathcal{K}. \end{aligned}$$

  Hence, $\max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k\right\} \geq \max_{k'\notin\mathcal{K}} \left\{\hat{w}_i^{k'} - \pi_{k'}\right\}$.

- $\varepsilon^+ \leq \hat{w}_i^k - \pi_k^\star$ for all $i \in I^\star_{\mathcal{K}}$ and $k \in \mathcal{K}$:

$$\max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k\right\} = \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^\star - \varepsilon^+\right\} \geq \max_{k\in\mathcal{K}} \left\{\hat{w}_i^k - \pi_k^\star - \hat{w}_i^k + \pi_k^\star\right\} = 0.$$

Hence, $i \in I_{\mathcal{K}}$ and thus, $I^\star_{\mathcal{K}} \subseteq I_{\mathcal{K}}$. Then, for the objective function holds

$$\begin{aligned}
\chi(\pi) &= \sum_{k\in\mathcal{K}} \Gamma_k \left(\pi^\star_k + \varepsilon^+\right) + \sum_{k\notin\mathcal{K}} \Gamma_k \pi^\star_k + \sum_{i\in I_{\mathcal{K}}} \max_{k\in\mathcal{K}} \left\{\hat{w}^k_i - \pi^\star_k - \varepsilon^+\right\} + \sum_{i\in N\setminus I_{\mathcal{K}}} \max_{k\notin\mathcal{K}} \left\{\left(\hat{w}^k_i - \pi^\star_k\right)^+\right\} \\
&= \sum_{k=1}^{K} \Gamma_k \pi^\star_k + \varepsilon^+ \sum_{k\in\mathcal{K}} \Gamma_k + \sum_{i\in I^\star_{\mathcal{K}}} \max_{k\in\mathcal{K}} \left\{\hat{w}^k_i - \pi^\star_k\right\} - \varepsilon^+ |I^\star_{\mathcal{K}}| + \sum_{i\in N\setminus I^\star_{\mathcal{K}}} \max_{k\notin\mathcal{K}} \left\{\left(\hat{w}^k_i - \pi^\star_k\right)^+\right\} \\
&= \chi(\pi^\star) + \varepsilon^+ \left(\sum_{k\in\mathcal{K}} \Gamma_k - |I^\star_{\mathcal{K}}|\right) \\
&\leq \chi(\pi^\star).
\end{aligned}$$

The second case, when the directional derivative $\sum_{k\in\mathcal{K}} \Gamma_k - |I^\star_{\mathcal{K}}|$ is strictly positive, can be handled quite analogously to the previous case.

We define the minimum value which is possible to subtract from $\pi^\star_k$ without any change of the partition as follows.

$$\varepsilon^- := \min \left\{ \min_{\substack{k\in\mathcal{K}, i\in N: \\ \pi^\star_k - \hat{w}^k_i > 0}} \left\{\pi^\star_k - \hat{w}^k_i\right\}, \min_{\substack{k\in\mathcal{K}, k'\notin\mathcal{K}, i\in N \\ \hat{w}^{k'}_i - \pi^\star_{k'} > \hat{w}^k_i - \pi^\star_k}} \left\{\hat{w}^{k'}_i - \pi^\star_{k'} - \hat{w}^k_i + \pi^\star_k\right\} \right\}$$

Moreover, we define

$$\pi_k := \begin{cases} \pi^\star_k - \varepsilon^- & k \in \mathcal{K} \\ \pi^\star_k & k \notin \mathcal{K}. \end{cases}$$

Again, if ε^- is defined by the first minimum, we stop the decrease of $\pi^\star$. Otherwise, we compute the new value of α, define a new set $\mathcal{K}'$ and start again from the beginning as described in the first case.

Once more we have to show that the objective value does not increase when decreasing $\pi^\star$ as described. To this end, let $I_{\mathcal{K}}$ be the set of items obtained by replacing $\pi^\star$ by π in $I^\star_{\mathcal{K}}$. Obviously, $I^\star_{\mathcal{K}} \subseteq I_{\mathcal{K}}$. To prove also $I_{\mathcal{K}} \subseteq I^\star_{\mathcal{K}}$, we show $N \setminus I^\star_{\mathcal{K}} \subseteq N \setminus I_{\mathcal{K}}$.

For that purpose we consider an item $i \in N \setminus I^\star_{\mathcal{K}}$, i. e.,

$$\max\left\{\max_{k\in\mathcal{K}}\left\{\hat{w}^k_i - \pi^\star_k\right\}, \max_{k'\notin\mathcal{K}}\left\{\hat{w}^{k'}_i - \pi^\star_{k'}\right\}, 0\right\} = \begin{cases} 0 & \text{(a)} \\ \max\limits_{k'\notin\mathcal{K}}\left\{\hat{w}^{k'}_i - \pi^\star_{k'}\right\}. & \text{(b)} \end{cases}$$

(a) is equivalent to

$$\max_{k\in\mathcal{K}}\left\{\hat{w}^k_i - \pi^\star_k\right\} < 0 \Leftrightarrow \hat{w}^k_i - \pi^\star_k < 0 \quad \forall\, k \in \mathcal{K}.$$

By the definition of ε^-,

$$\hat{w}^k_i - \pi^\star_k + \varepsilon^- = \hat{w}^k_i - \pi_k \leq 0 \quad \forall\, k \in \mathcal{K} \;\Rightarrow i \notin I_{\mathcal{K}}.$$

(b) is equivalent to

$$\max_{k\in\mathcal{K}}\left\{\hat{w}_i^k-\pi_k^\star\right\}<\max_{k'\notin\mathcal{K}}\left\{\hat{w}_i^{k'}-\pi_{k'}^\star\right\}\Leftrightarrow\hat{w}_i^k-\pi_k^\star<\max_{k'\notin\mathcal{K}}\left\{\hat{w}_i^{k'}-\pi_{k'}^\star\right\}\quad\forall\,k\in\mathcal{K}.$$

Again, by the definition of ε^-,

$$\hat{w}_i^k-\pi_k^\star+\varepsilon^-=\hat{w}_i^k-\pi_k<\max_{k'\notin\mathcal{K}}\left\{\hat{w}_i^{k'}-\pi_{k'}\right\}\quad\forall\,k\in\mathcal{K}\;\Rightarrow i\notin I_{\mathcal{K}}.$$

Altogether, we have $I_{\mathcal{K}}=I_{\mathcal{K}}^\star$. For the objective function (12.10) then holds

$$\begin{aligned}\chi(\pi)&=\sum_{k\in\mathcal{K}}\Gamma_k(\pi^\star-\varepsilon^-)+\sum_{k\notin\mathcal{K}}\Gamma_k\pi_k^\star+\sum_{i\in I_{\mathcal{K}}}\max_{k\in\mathcal{K}}\left\{\hat{w}_i^k-\pi_k^\star+\varepsilon^-\right\}+\sum_{i\in N\setminus I_{\mathcal{K}}}\max_{k\notin\mathcal{K}}\left\{\hat{w}_i^k-\pi_k^\star\right\}\\&=\sum_{k=1}^{K}\Gamma_k\pi_k^\star-\varepsilon^-\sum_{k\in\mathcal{K}}\Gamma_k+\sum_{i\in I_{\mathcal{K}}^\star}\max_{k\in\mathcal{K}}\left\{\hat{w}_i^k-\pi_k^\star\right\}+\varepsilon^-|I_{\mathcal{K}}^\star|+\sum_{i\in N\setminus I_{\mathcal{K}}^\star}\max_{k\notin\mathcal{K}}\left\{\hat{w}_i^k-\pi_k^\star\right\}\\&=\chi(\pi^\star)-\varepsilon^-\left(\sum_{k\in\mathcal{K}}\Gamma_k-|I_{\mathcal{K}}^\star|\right)\\&<\chi(\pi^\star).\end{aligned}$$

□

We summarise the previous achievements in the following complexity result.

Theorem 12.9. *There exists a DP which solves* (12.4) *in* $O(K!n^{K+1}B)$.

Proof. By Theorem 12.6 and Lemma 12.8, we know that for an optimal $\pi\in\mathbb{R}^K_{\geq 0}$ holds $\pi_\ell\in\{\hat{w}_i^\ell\,|\,i\in N\}\cup\{0\}$ for at least one $\ell\in\{1,\dots,K\}$. We prove the claim by mathematical induction for $K\geq 2$.

For $K=2$, we can either first choose $\pi_1\in\{\hat{w}_i^1\,|\,i\in N\}\cup\{0\}$ or $\pi_2\in\{\hat{w}_i^2\,|\,i\in N\}\cup\{0\}$ having $n+1$ possibilities each and two options to choose the starting $\ell\in\{1,2\}$. For a fixed π_ℓ with $\ell\in\{1,2\}$, the remaining optimal π_l, $l\neq\ell$, lies in $\{\hat{w}_i^l-(\hat{w}_i^\ell-\pi_\ell)^+\,|\,i\in N\}\cup\{0\}$ as shown in Lemma 12.5. Again, we have $n+1$ possibilities for π_l. For a fixed π, the optimal ρ is fixed; cf. Lemma 12.5, and the problem reduces to a knapsack problem which can be solved in $O(nB)$ by the DP (12.7) and (12.8). In total, we have a runtime of $O(2n^3B)$ for $K=2$.

Now, we assume that the claim holds for $K-1$. Thus, the $(K-1)$-band RKP can be solved in $O((K-1)!n^KB)$. For the K-band RKP, we choose a $\ell\in\{1,\dots,K\}$ and one π_ℓ from $\{\hat{w}_i^\ell\,|\,i\in N\}\cup\{0\}$ having $K(n+1)$ many possibilities to fix the first π_ℓ. After the fixation, the problem reduces to a $(K-1)$-band RKP. In total, we have a runtime of $O(K!n^{K+1}B)$. □

The set $\Pi_1\times\ldots\times\Pi_K$ is implicitly described in this proof. For $K=2$, we state a closed form in Section 12.5.

12.3.2. Uncertain Objective Coefficients

Based on the presented results, we give a polynomial time algorithm to solve a binary optimisation problem with uncertain objective coefficients, correcting the result in [138] and extending the result in Theorem 8.2 for the Γ-RKP.

The problem with uncertain objective coefficients is defined as

$$\min \sum_{i \in N} c_i x_i \tag{12.11a}$$

$$\text{s.t. } x \in X, \tag{12.11b}$$

where the feasible region is $X \subseteq \{0,1\}^n$ and c_i are subject to uncertainty. The uncertainty is modelled by a multi-band robust uncertainty set with $\bar{c}_i$ denoting the nominal value and $\hat{c}_i^k$ the deviation in band k. The multi-band robust counterpart of (12.11) can be formalised as

$$\min \sum_{i \in N} \bar{c}_i x_i + \max_{\pi \in \Pi_1 \times \ldots \times \Pi_K} \mathrm{DEV}(x, \pi) \tag{12.12a}$$

$$\text{s.t. } x \in X, \tag{12.12b}$$

with

$$\mathrm{DEV}(x, \pi) \coloneqq \sum_{k \in \{1,\ldots,K\}} \Gamma_k \pi_k + \sum_{i \in N} \left(\max_{k \in \{1,\ldots,K\}} \{\hat{c}_i^k - \pi_k\} \right)^+ x_i. \tag{12.13}$$

Rewriting (12.12) by using (12.13), we obtain

$$\min \left\{ \sum_{k \in \{1,\ldots,K\}} \Gamma_k \pi_k + \min_{x \in X} \left\{ \sum_{i \in N} \left(\bar{c}_i + \left(\max_{k \in \{1,\ldots,K\}} \{\hat{c}_i^k - \pi_k\} \right)^+ \right) x_i \right\} \middle| \pi \in \Pi_1 \times \ldots \times \Pi_K \right\} \tag{12.14}$$

The set $\Pi_1 \times \ldots \times \Pi_K$ is again implicitly defined in the proof of Theorem 12.9 just replacing $\hat{w}$ by $\hat{c}$.

Corollary 12.10. *The problem* (12.11) *with uncertain objective coefficients defined in a multi-band robust uncertainty set can be solved in* $O(K! n^K)$.

Proof. Solve the equivalent formulation (12.14) and use the same argumentation as in the proof of Theorem 12.9. □

12.4. Practical Improvements

By Theorem 12.6 and Lemma 12.8, we can now define alternatives to the sets Π_k defined in (12.6). However, a straightforward application of the former results leads to sets which are still quite large. In this section, we present some improvements to reduce the size of the sets and to speed up the solving process in general.

12.4.1. Reducing the Size of Π_k

For a reduction of the number of π-vectors to be considered in the DP, we define a mapping per band to sort the corresponding deviation values non-increasingly:

$$h_k : N \to N \text{ with } \hat{w}^k_{h_k(i)} \geq \hat{w}^k_{h_k(i+1)}, \forall i \in \{1, \dots, n-1\},\ k \in \{1, \dots, K\}$$

Lemma 12.11. *There exists a point π defining a minimum of function χ with $\pi_k \in \{\hat{w}^k_i \mid i \in N\} \cup \{0\}$ for at least one $k \in \{1, \dots, K\}$ (cf. Lemma 12.8) and*

$$\pi_k \leq \hat{w}^k_{h_k(\Gamma_k+1)} \quad \forall k \in \{1, \dots, K\}.$$

Proof. Let $\pi^\star$ be a point defining a minimum of χ with $\pi^\star_\ell > \hat{w}^\ell_{h_\ell(\Gamma_\ell+1)} =: \alpha$ for one band $\ell \in \{1, \dots, K\}$. Construct a point π with $\chi(\pi) \leq \chi(\pi^\star)$ as follows.

$$\pi_\ell := \alpha,\ \pi_k := \pi^\star_k \quad \forall k \neq \ell.$$

Note that π still satisfies the property of Lemma 12.8. To show $\chi(\pi) \leq \chi(\pi^\star)$, let $z := \pi^\star_\ell - \pi_\ell > 0$. We define a partition of the sets of items as $N = I_\leq \cup I_>$ with $I_\leq := \{i \in N \mid h^{-1}_\ell(i) \leq \Gamma_\ell\}$, $I_> := \{i \in N \mid h^{-1}_\ell(i) > \Gamma_\ell\}$. For $i \in I_\leq$ holds

$$\max_{k\in\{1,\dots,K\}} \left\{\hat{w}^k_i - \pi_k\right\} = \max\left\{\max_{k\neq\ell}\left\{\hat{w}^k_i - \pi^\star_k\right\}, \hat{w}^\ell_i - \pi^\star_\ell + z\right\} \leq \max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi^\star_k\right\} + z$$

and

$$\max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi_k\right\}, 0\right\} \leq \max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi^\star_k\right\}, 0\right\} + z.$$

On the other hand, for $i \in I_>$ holds in particular $\hat{w}^\ell_i \leq \alpha$ and thus,

$$\begin{aligned}\max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi_k\right\}, 0\right\} &= \max\left\{\max_{k\neq\ell}\left\{\hat{w}^k_i - \pi^\star_k\right\}, \hat{w}^\ell_i - \alpha, 0\right\} \\ &\leq \max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi^\star_k\right\}, 0\right\}\end{aligned}$$

Hence, regarding the objective value $\chi(\pi)$ we have

$$\begin{aligned}\chi(\pi) &= \sum_{k=1}^K \Gamma_k\pi_k + \sum_{i\in N}\max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi_k\right\}, 0\right\} \\ &\leq \sum_{k=1}^K \Gamma_k\pi^\star_k - \Gamma_\ell z + \sum_{i\in I_\leq}\left(\max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi^\star_k\right\}, 0\right\} + z\right) \\ &\quad + \sum_{i\in I_>}\max\left\{\max_{k\in\{1,\dots,K\}}\left\{\hat{w}^k_i - \pi^\star_k\right\}, 0\right\} = \chi(\pi^\star),\end{aligned}$$

a contradiction. □

For a further reduction of the necessary values of π_k, recall that $\hat{w}_i^k < \hat{w}_i^{k+1}$ for all bands $k \in \{1, \dots, K-1\}$ and all items $i \in N$.

Lemma 12.12. *There exists a point π defining a minimum of function χ satisfying the properties of Lemma 12.11, i. e.,*

a) $\pi_k \in \left\{\hat{w}_i^k \,\middle|\, i \in N\right\} \cup \{0\}$ *for at least one* $k \in \{1, \dots, K\}$,

b) $\pi_k \leq \hat{w}^k_{h_k(\Gamma_k+1)}$ *for all* $k \in \{1, \dots, K\}$,

and

c) $\pi_{k+1} \geq \pi_k \geq 0$ *for all* $k \in \{1, \dots, K-1\}$.

Proof. We assume $\pi^\star$ defines a minimum of χ and $\pi_\ell^\star > \pi_{\ell+1}^\star \geq 0$ for one $\ell \in \{1, \dots, K-1\}$ but properties a) and b) are fulfilled. We construct a point π with $\chi(\pi) \leq \chi(\pi^\star)$ that has the properties a) and b) and satisfies $\pi_\ell \leq \pi_{\ell+1}$. We distinguish the following two cases.

I) There exists at least one band $k' \neq \ell$ with $\pi_{k'}^\star \in \left\{\hat{w}_i^{k'} \,\middle|\, i \in N\right\} \cup \{0\}$.

II) Band ℓ is the only band satisfying $\pi_\ell^\star \in \left\{\hat{w}_i^\ell \,\middle|\, i \in N\right\} \cup \{0\}$.

Case I): Define $z := \pi_\ell^\star - \pi_{\ell+1}^\star > 0$ and the new point π

$$\begin{aligned}\pi_\ell &:= \pi_\ell^\star - z = \pi_{\ell+1}^\star \\ \pi_k &:= \pi_k^\star \quad \forall k \neq \ell,\end{aligned}$$

which still satisfies properties a) and b). If follows for all items $i \in N$

$$\hat{w}_i^\ell - \pi_\ell = \hat{w}_i^\ell - \pi_{\ell+1}^\star < \hat{w}_i^{\ell+1} - \pi_{\ell+1}^\star$$

and hence,

$$\begin{aligned}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} &= \max\left\{\max_{k\neq\ell}\{\hat{w}_i^k - \pi_k^\star\}, \hat{w}_i^\ell - \pi_\ell, 0\right\} \\ &\leq \max\left\{\max_{k\neq\ell}\{\hat{w}_i^k - \pi_k^\star\}, \hat{w}_i^{\ell+1} - \pi_{\ell+1}^\star, 0\right\} \\ &\leq \max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\}.\end{aligned}$$

For the objective function χ we have

$$\begin{aligned}\chi(\pi) &= \sum_{k=1}^{K} \Gamma_k \pi_k + \sum_{i\in N} \max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} \\ &\leq \sum_{k\neq\ell} \Gamma_k \pi_k^\star + \Gamma_\ell(\pi_\ell^\star - z) + \sum_{i\in N} \max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\} \\ &< \chi(\pi^\star).\end{aligned}$$

Case II): If we just used the same argumentation as in case I), we would obtain a point π which does not necessarily satisfy property a) anymore. Hence, we use a modified approach in the following. Due to the assumption that band ℓ is the only band satisfying $\pi_k^\star \in \left\{\hat{w}_i^k \,\middle|\, i \in N\right\} \cup \{0\}$ and by Lemma 12.5, we know that $\pi_{\ell+1}^\star \in \{(\hat{w}_i^{\ell+1} - \max_{k\neq\ell+1}\{\hat{w}_i^k - \pi_k^\star\})^+ \,|\, i \in N\}$. Again, we distinguish two cases:

i) $\pi_{\ell+1}^\star = 0$,

ii) $\pi_{\ell+1}^\star = \hat{w}_i^{\ell+1} - \hat{w}_j^k + \pi_k^\star$ for at least one $k \neq \ell+1$ and $j \in N$.

For case i), we construct a new point π with $\chi(\pi) < \chi(\pi^\star)$ as follows.

$$\pi_\ell \coloneqq 0, \quad \pi_k \coloneqq \pi_k^\star \quad \forall k \neq \ell.$$

Note that π still fulfils the properties a) and b). It holds

$$\hat{w}_i^\ell - \pi_\ell = \hat{w}_i^\ell < \hat{w}_i^{\ell+1} = \hat{w}_i^{\ell+1} - \pi_{\ell+1}^\star$$

and with the same argumentation as in case I),

$$\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} \leq \max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\}.$$

Therefore, for the objective function χ we have

$$\begin{aligned}\chi(\pi) &= \sum_{k=1}^{K}\Gamma_k\pi_k + \sum_{i\in N}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} \\ &\leq \sum_{k\neq\ell}\Gamma_k\pi_k^\star + \Gamma_\ell\cdot 0 + \sum_{i\in N}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\} < \chi(\pi^\star).\end{aligned}$$

For case ii), we construct a new point π in such a way that we will be able to use the same argumentation as in the proof of Lemma 12.8 to construct a further point $\tilde{\pi}$ which than satisfies all required properties. Therefore, we define $\alpha \coloneqq \pi_{\ell+1}^\star - \hat{w}_j^{\ell+1}$ and $\mathcal{K} \coloneqq \{k \in \{1,\dots,K\} \,|\, \hat{w}_j^k - \pi_k^\star = \alpha\}$ analogously to the proof of Lemma 12.8. Note that $\ell+1 \in \mathcal{K}$ but $\ell \notin \mathcal{K}$. Moreover, we set

$$\varepsilon \coloneqq \min_{\substack{k\in\mathcal{K}, i\in N:\\ \hat{w}_i^k - \pi_k^\star > 0}} \{\hat{w}_i^k - \pi_k^\star\}$$

and $z \coloneqq \pi_\ell^\star - \pi_{\ell+1}^\star > 0$. Then the new point π is constructed as

$$\begin{aligned}\pi_\ell &\coloneqq \pi_\ell^\star - z - \varepsilon = \pi_{\ell+1}^\star - \varepsilon \\ \pi_k &\coloneqq \pi_k^\star \quad \forall k \neq \ell,\end{aligned}$$

which still satisfies property b). If follows for all items $i \in N$

$$\hat{w}_i^\ell - \pi_\ell = \hat{w}_i^\ell - \pi_{\ell+1}^\star + \varepsilon \leq \hat{w}_i^\ell - \pi_{\ell+1}^\star + \pi_{\ell+1}^\star - \hat{w}_i^{\ell-1} < 0$$

and hence,

$$\begin{aligned}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} &= \max\left\{\max_{k\neq\ell}\{\hat{w}_i^k - \pi_k^\star\}, \hat{w}_i^\ell - \pi_\ell, 0\right\} \\ &\leq \max\left\{\max_{k\neq\ell}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\} \\ &\leq \max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\}.\end{aligned}$$

For the objective function χ we have

$$\begin{aligned}\chi(\pi) &= \sum_{k=1}^{K}\Gamma_k\pi_k + \sum_{i\in N}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k\}, 0\right\} \\ &\leq \sum_{k\neq k'}\Gamma_k\pi_k^\star + \Gamma_\ell(\pi_\ell^\star - z - \varepsilon) + \sum_{i\in N}\max\left\{\max_{k\in\{1,\dots,K\}}\{\hat{w}_i^k - \pi_k^\star\}, 0\right\} \\ &< \chi(\pi^\star).\end{aligned}$$

Although π does not satisfy property a) but $\hat{w}_j^{\ell+1} - \pi_{\ell+1}^\star = \hat{w}_j^k - \pi_k^\star$ for one $j \in N$ and $k \neq \ell+1$, we can apply the same argumentation as in the proof of Lemma 12.8 to construct a new point $\tilde{\pi}$ satisfying property a). By the definition of π_ℓ, it is not possible that $\pi_{\ell+1}$ can become smaller again than π_ℓ. In case of a negative considered derivative, $\pi_{\ell+1}$ is increased by ε^+ and π_ℓ not. Note that property b) is preserved by the definition of ε^+. If the considered derivative is positive, $\pi_{\ell+1}$ is decreased by $\varepsilon^- \leq \varepsilon$. □

By means of this lemma, we have to consider only vectors π in the DP with $0 \leq \pi_k \leq \pi_{k+1}$ for all $k \in \{1, \dots, K-1\}$ and $\pi_k \leq \hat{w}_{h_k(\Gamma_k+1)}^k$ for all $k \in \{1, \dots, K\}$.

Corollary 12.13. *For $K = 2$, there exists a point π defining a minimum of function χ satisfying the properties of Lemma 12.12, where property c) is replaced by $\pi_2 > \pi_1 \geq 0$ or $\pi_2 = \pi_1 = 0$.*

12.4.2. Acceleration of the Solving Process

The decrease of the size of the set $\Pi \coloneqq \Pi_1 \times \ldots \times \Pi_K$ achieved in the previous section is just of practical relevance as the worst-case size remains the same. However, even though the size can be reduced in practice, solving a classical KP for every possible $\pi \in \Pi$ still consumes a large amount of time. To decrease the solving time, we order the elements in Π such that the most relevant are considered first and the least relevant are most likely not taken into account at all.

For a fixed $\pi \in \Pi$, the following KP has to be solved.

$$\begin{aligned} \max \quad & \sum_{i \in N} p_i x_i \\ \text{s.t.} \quad & \sum_{i \in N} (\bar{w}_i + \mathrm{DEV}_\pi(i)) x_i \leq B - \sum_{k=1}^{K} \Gamma_k \pi_k \\ & x_i \in \{0,1\} \qquad \forall\, i \in N \end{aligned}$$

We use an upper bound for this problem, which we derive in the following, to define an order of the elements in Π. A straightforward upper bound is an optimal solution of the LP relaxation which is computed by a greedy algorithm. This algorithm sorts the items by their profit-to-weight ratio $p_i/(\bar{w}_i + \mathrm{DEV}_\pi(i))$ non-increasingly and includes the items with the largest ratio in the knapsack as long as the capacity is not exceeded. To use the capacity completely, just a fraction of the last element, the so-called critical item, might be put into the knapsack. We improve this bound as described in Martello and Toth [135] by interchanging items closely before or after the critical item. We save this upper bound for every $\pi \in \Pi$ and sort the elements non-increasingly regarding the bound. A vector π with a large upper bound has a higher potential to increase the objective. Furthermore, if the upper bound of π is less than or equal to the current best known solution, an optimal solution of the corresponding KP cannot improve the best known solution and thus, we disregard the optimal solving of the KP. Additionally, the whole solving process terminates as the upper bound of the next element is not higher than the current best solution.

12.5. Computational Study

In this section, we present an extensive computational study to evaluate the performance of the derived algorithms in practice. To obtain computationally tractable formulations and algorithms, we focus on two bands in this study.

We investigate the following four algorithms. The DP presented in Section 12.2 utilising the sets Π_k as defined in (12.6) is the basic algorithm and denoted by SIMPLE. When replacing the Cartesian product of the sets Π_k by Π_{REDUCED} which is based on the results from Section 12.4.1 and will be defined in (12.15), we obtain the algorithm REDUCED. If we order the (π_1, π_2)-pairs as described in Section 12.4.2 in the algorithm REDUCED, we call the resulting algorithm IMPROVED. Finally, to evaluate the performance of the best DP, we additionally solve the ILP (12.4) for all instances and call this algorithm ILP. In all three DPs, we solve the underlying KPs for fixed vectors π via the algorithm "MinKnap" developed by Pisinger [153]. The considered algorithms are summarised in Table 12.2.

In the following, we derive a closed form expression of the set Π_{REDUCED} for $K = 2$. Assuming $\pi_1 \in \{\hat{w}^1_1, \ldots, \hat{w}^1_n\}$ in an optimal solution, we first define

$$\Pi_1 := \{0\} \cup \left\{ \hat{w}^1_{h_1(i)} \,\middle|\, i \leq \Gamma_1 + 1 \text{ and } \hat{w}^1_{h_1(i)} \leq \left\lfloor \frac{B}{\Gamma_1} \right\rfloor \right\},$$

Algorithm	Description
SIMPLE	The DP using the simple sets Π_k as defined in (12.6) with complexity $O(nB^3)$.
REDUCED	The DP using the set Π_{REDUCED} as defined in (12.15) with complexity $O(2!n^3B)$.
IMPROVED	As algorithm REDUCED but the elements in Π_{REDUCED} are ordered non-increasingly regarding the upper bounds as described in Section 12.4.2 and the solving process stops as soon as the current upper bound is less than or equal to the best solution.
ILP	The ILP (12.4) is solved with CPLEX.

Table 12.2.: Summary of algorithms considered in the computational study.

and then for every $\pi_1 \in \Pi_1$, we define

$$\Pi_2(\pi_1) \coloneqq \{0\} \cup \left\{ \left(\hat{w}_i^2 - \left(\hat{w}_i^1 - \pi_1 \right)^+ \right)^+ \;\middle|\; i \in N \text{ and } \pi_1 < \left(\hat{w}_i^2 - \left(\hat{w}_i^1 - \pi_1 \right)^+ \right)^+ \leq \min \left\{ \hat{w}^2_{h_2(\Gamma_2+1)}, \left\lfloor \frac{B}{\Gamma_2} \right\rfloor \right\} \right\}.$$

Furthermore when assuming $\pi_2 \in \{\hat{w}_1^2, \ldots, \hat{w}_n^2\}$ in an optimal solution, we define

$$\Pi_2' \coloneqq \{0\} \cup \left\{ \hat{w}^2_{h_2(i)} \;\middle|\; i \leq \Gamma_2 + 1 \text{ and } \hat{w}^2_{h_2(i)} \leq \left\lfloor \frac{B}{\Gamma_2} \right\rfloor \right\},$$

and for every $\pi_2 \in \Pi_2'$

$$\Pi_1'(\pi_2) \coloneqq \{0\} \cup \left\{ \left(\hat{w}_i^1 - \left(\hat{w}_i^2 - \pi_2 \right)^+ \right)^+ \;\middle|\; i \in N \text{ and } \left(\hat{w}_i^1 - \left(\hat{w}_i^2 - \pi_2 \right)^+ \right)^+ \leq \min \left\{ \hat{w}^1_{h_1(\Gamma_1+1)}, \left\lfloor \frac{B}{\Gamma_1} \right\rfloor, \pi_2 - 1 \right\} \right\}.$$

Combining these sets, we get the set of all (π_1, π_2)-pairs which have to be considered in the DP as

$$\Pi_{\text{REDUCED}} \coloneqq \left(\bigcup_{\pi_1 \in \Pi_1} \{\pi_1\} \times \Pi_2(\pi_1) \right) \cup \left(\bigcup_{\pi_2 \in \Pi_2'} \Pi_1'(\pi_2) \times \{\pi_2\} \right). \tag{12.15}$$

All computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of two hours. To solve the ILP, we use the standard version of CPLEX 12.4 [98].

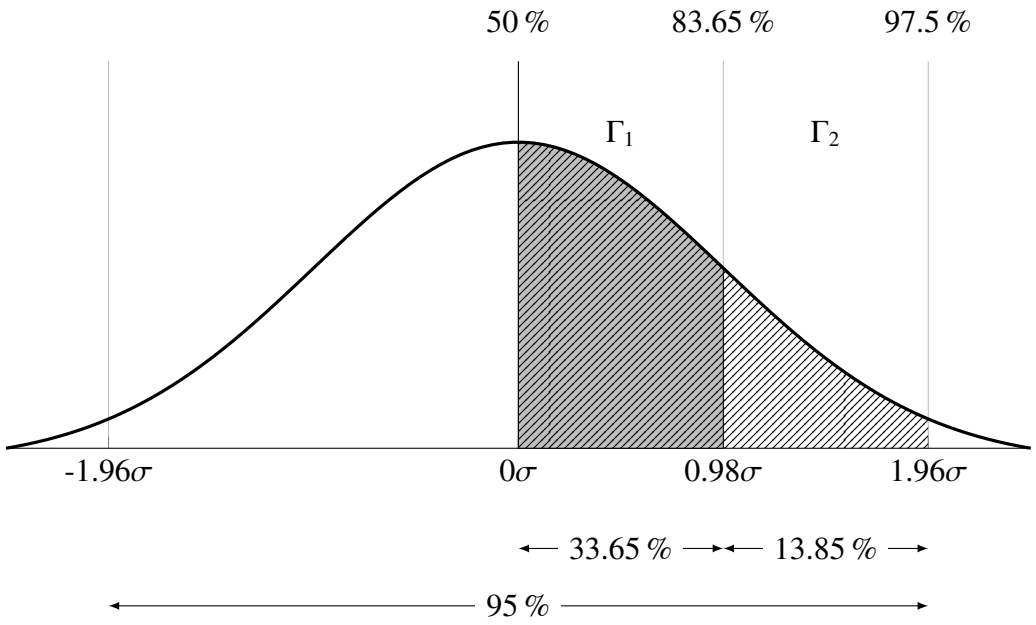


Figure 12.1.: Sketch to explain the computation of the robustness parameters Γ_1 and Γ_2.

12.5.1. Generation of Test Instances

Based on Klopfenstein and Nace [109] and Pisinger [154], we randomly generate hard instances to study the performance of the different algorithms.

The nominal weights are uniformly distributed in a given interval $[1, R]$, where $R = 100$ and $R = 1000$ defines the range, and the profits are proportional to the nominal weights plus a fixed charge, i. e., $p_i = \bar{w}_i + 10$. Hence, the instances are strongly correlated and hard to solve; see [154]. Furthermore, the knapsack capacity B is randomly chosen from $\left[\frac{1}{3}\sum_{i\in N}\bar{w}_i, \frac{2}{3}\sum_{i\in N}\bar{w}_i\right] \cap \mathbb{Z}$. To create two-band robust instances, we additionally have to compute two deviation values $\hat{w}_i^1$ and $\hat{w}_i^2$ for each $i \in N$ with $\hat{w}_i^1 < \hat{w}_i^2$. For that purpose, we define a maximum deviation $\delta \in \{0.2, 0.5, 1.0\}$ relative to the nominal weight, where, e. g., $\delta = 0.2$ signifies a maximum deviation of at most 20 % of the nominal weight. Hence, $\hat{w}_i^2 = \delta\bar{w}_i$. Assuming an (almost) equal bandwidth in each band, we set $\hat{w}_i^1 = \left\lceil \frac{\hat{w}_i^2}{2} \right\rceil$. Finally, we consider three different numbers of items $n \in \{200, 500, 1000\}$. For each setting, we generate a series of ten instances.

The number of possible values for the robustness parameters Γ_1 and Γ_2 is quite large. To focus on the most meaningful values, we assume a normal distribution of the deviation values; see Figure 12.1. 95 % of the deviation values lie between -1.96σ and 1.96σ, where σ denotes the standard deviation; cf. the cumulative distribution function of the standard normal distribution. In particular, 97.5 % of the values are less than 1.96σ. This point is reflected by Γ_2. As we are interested only in positive deviations, we consider $0.98\sigma = 1.96\sigma/2$ for Γ_1. At this point, we have a probability of 83.65 %. Furthermore, we have a probability of 33.65 % = 83.65 % − 50 % between 0σ and 0.98σ and a

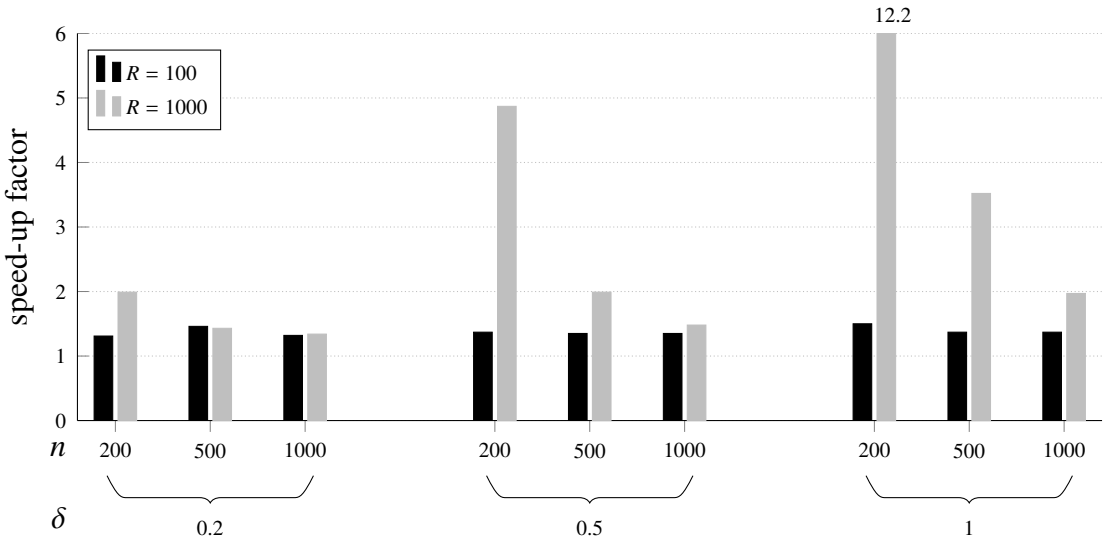


Figure 12.2.: Speed-up factors of algorithm REDUCED normalised to SIMPLE averaged over ten instances and five Γ_2-values for different settings and sizes.

probability of $13.85\,\% = 97.5\,\% - 83.65\,\%$ between 0.98σ and 1.96σ. This, means we have the following ratio.

$$\frac{\Gamma_1}{\Gamma_2} = \frac{33.65\,\%}{13.85\,\%} = 2.43$$

Thus, we assume $\Gamma_2 \in \{\lceil 0.01n \rceil, \lceil 0.02n \rceil, \lceil 0.03n \rceil, \lceil 0.04n \rceil, \lceil 0.05n \rceil\}$, where $\Gamma_2 = \lceil 0.05n \rceil$ means that at most 5 % of all weights deviate, and $\Gamma_1 = \lceil 2.43 \cdot \Gamma_2 \rceil$. Note, we round $2.43 \cdot \Gamma_2$ to receive only integer values for Γ_1 since the DP (12.8) requires integer parameters.

12.5.2. Comparison of different sets Π

First, we compare the performance of algorithm REDUCED to SIMPLE. For that purpose, we examine the solving times of the two algorithms and compute a speed-up factor for REDUCED. For example, a solving time of 100 s for SIMPLE and a solving time of 10 s for REDUCED gives a speed-up factor of 10. For every setting consisting of a range R, a number of items n, and a maximum deviation δ, we take the mean over the ten instances. Furthermore, since differences in the solving time for different values of Γ_2 are marginal, we also average over Γ_2 and display the results in Figure 12.2. The average speed-up factor ranges from 1.31 to 12.20. In general, for higher range R and higher δ, we can also achieve higher speed-up factors. In SIMPLE, the size of the sets Π_k, which is defined by the actual deviation values, has the strongest impact on the computing time, whereas the number of items plays a minor role. In contrast, the number of items influences the computing time of REDUCED most considerably while the deviation values are less important. Hence, the high speed-up factors achieved by REDUCED for $n = 200$ cannot be retained for higher numbers of items.

Without averaging, the individual minimum speed-up factor achieved by REDUCED ranges from 1 to 14.09.

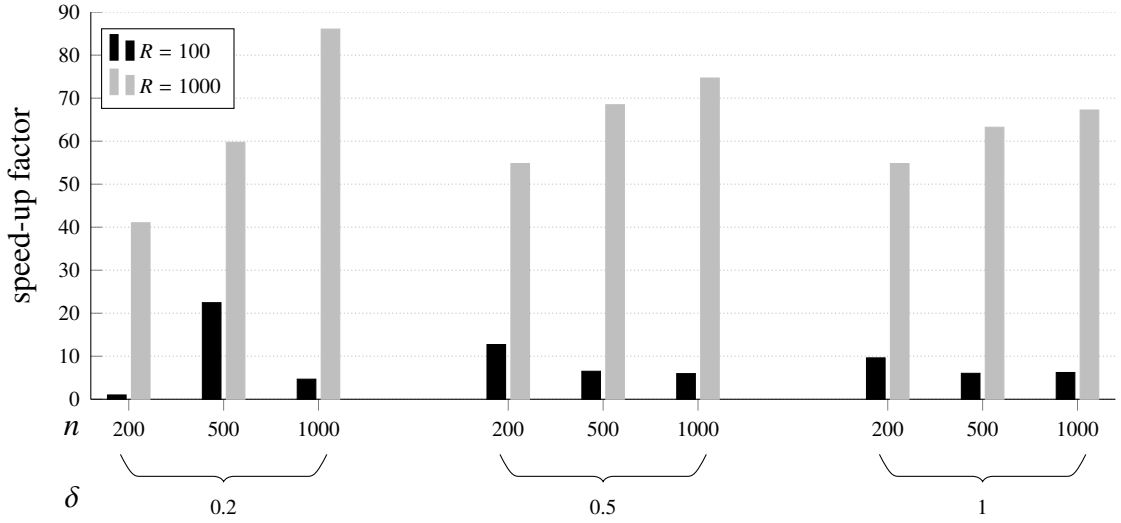


Figure 12.3.: Speed-up factors of algorithm IMPROVED normalised to REDUCED averaged over ten instances and five Γ_2-values for different settings and sizes.

12.5.3. Evaluation of Practical Improvements

We now compare IMPROVED to REDUCED by means of speed-up factors, where we once more average over instances and Γ_2-values. A graphical evaluation of the results is depicted in Figure 12.3. For $R = 100$, $\delta = 0.2$ and $n = 200$, no speed-up factor is computed since the solving times of IMPROVED are strictly below 10^{-2} s. For better readability, we nevertheless draw a bar of height 1. For the remaining settings, the average speed-up factor achieved by IMPROVED ranges from 4.67 to 86.05. For an increasing number of items and $R = 1000$, the speed-up factor also increases because the stop criterion if the upper bound of a (π_1, π_2)-pair is below the current best known solution gains more significance for a higher number of items. For example, the stop criterion catches after 157 (π_1, π_2)-pairs have been considered for an instance with $n = R = 1000$, $\delta = 1$ and $\Gamma_2 = \lceil 0.01n \rceil$ while 240382 pairs are considered by REDUCED. On the contrary, for one of the smallest instances ($n = 200$, $R = 1000$, $\delta = 0.2$ and $\Gamma_2 = \lceil 0.01n \rceil$), the number of considered (π_1, π_2)-pairs is only reduced from 155 to 8. Hence, the reduction for the highest number of items is the largest. For $R = 100$, these effects cannot be seen that clearly since the problems are quite small and are solved in less than 1.5 s with both algorithms.

Without averaging, the minimum speed-up factor that can be achieved by IMPROVED is 2 and the maximum is 500.

12.5.4. Comparison to ILP Formulation

The absolute solving times averaged over ten instances and five Γ_2-values for the algorithm ILP and for IMPROVED are displayed in Table 12.3. Note, some of the 50 problems per average value are not solved within the time limit by ILP. These instances are not considered in the average. However, we give the number of the remaining instances, which are

	δ	0.2		0.5		1.0	
	n \ R	100	1000	100	1000	100	1000
ILP	200	671 (44)	540 (44)	578 (37)	1388 (37)	626 (29)	885 (25)
	500	1666 (36)	490 (13)	1501 (23)	571 (12)	1718 (25)	882 (14)
	1000	1966 (36)	83 (4)	1161 (17)	103 (6)	362 (7)	105 (4)
IMPROVED	200	0.000	0.213	0.007	0.360	0.025	0.383
	500	0.002	0.613	0.028	2.075	0.103	4.030
	1000	0.012	1.073	0.059	5.705	0.219	16.116

Table 12.3.: Absolute solving times in seconds of ILP and IMPROVED for different settings averaged over ten instances and five Γ_2-values. The number of instances (out of 50) not exceeding the time limit by ILP are given in parenthesis.

solved to optimality, in parenthesis. The lowest average time consumption for ILP is 105 s ($R = 100$, $\delta = 1.0$, $n = 1000$) but only 4 out of 50 instances are solved within the time limit. In contrast, IMPROVED requires at most 16 s for these largest instances. Hence, for the 2-band RKP, the presented DP IMPROVED clearly outperforms ILP.

As mentioned before, the runtime of algorithm IMPROVED strongly depends on the number of items. Which means, if the number of items is doubled, also the runtime is increased by a factor of 2. Such a factor cannot be computed for algorithm ILP as more quantities than the number of items influence its running time.

12.5.5. Larger Instances

The solving times of the DP IMPROVED are quite low (at most 18.01 s) for all instances. Therefore, we now briefly study the performance of IMPROVED for larger instances. We generate instances with $R \in \{1000, 10000\}$, $n \in \{500, 1000, 2000, 5000, 10000\}$, and δ as before. Again, we generate ten instances for each setting. Note, we create only entities for a combination of R and n which has not been considered before. The absolute solving times are displayed in Figure 12.4. For $R = 10000$, $n = 10000$ and $\delta = 0.5$, only two instances are solved within the time limit whereas no instance for $\delta = 1.0$ is solved. For all other settings, the instances are solved within a reasonable time of at most 5811.76 s on average.

12.6. Conclusion

In this chapter, we have considered a generalisation of the RKP, the multi-band RKP, where the weights of the items have several deviation values lying in different bands. Based on the compact formulation, we have developed a DP with a complexity linear in the number of items n. However, the complexity also depends on the capacity which is raised to the

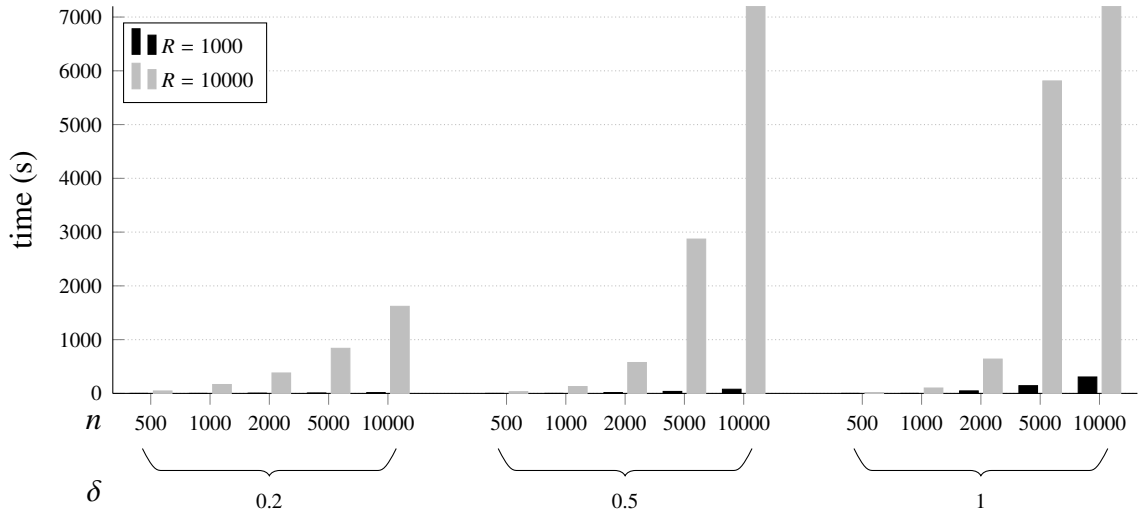


Figure 12.4.: Absolute solving times in seconds of algorithm IMPROVED averaged over ten instances and five Γ_2-values for different settings and sizes.

power of the number of bands K plus one. Since the capacity is usually higher than the number of items, we have additionally developed a DP with a complexity linear in the capacity B, i. e., $O(K!n^{K+1}B)$. From this, it can be concluded that a binary combinatorial optimisation problem with uncertain objective can be solved by solving $O(K!n^K)$ similar problems with certain objective. We have improved the DP in practice and compared the performance of the resulting algorithm to the former two DPs by means of solving times in an extensive computational study with randomly generated instances of various sizes. On the one hand, the results have demonstrated a clear benefit of the DP with a complexity linear in the capacity compared to the first DP. On the other hand, the results have shown the effectiveness of the presented improvements in practice. Furthermore, a comparison of the improved DP and CPLEX solving the compact ILP has illustrated that the improved DP clearly outperforms the ILP. Finally, we have tested the performance of the improved DP for large instances with up to 10000 items whereupon most instances are solved within a time limit of two hours.

Due to the good performance of the improved DP, we apply this algorithm as subroutine to a wireless network planning problem in Section 13.2.

13. Application to Wireless Networks

13.1. The Two-Band RWNPP

As stated earlier, multi-band robustness introduced in Chapter 11 generalises Γ-robustness formalised in Chapter 8 via a refinement of the deviation range. But the question remains, what is the actual gain achieved by multi-band robustness? To investigate this issue, we apply the multi-band robustness approach to the WNPP proposed in Section 4.1 assuming uncertain traffic demands and evaluate the obtained solutions in comparison to the Γ-robust solutions achieved in Section 9.1. Thus, we compare the number of deployed BSs for solutions with a protection level of 100 % revealing the gain of multi-band robustness.

To obtain a computationally tractable multi-band robust formulation, we consider only two positive deviation bands. Thus, we set $K^- = 0$ and $K^+ = K = 2$ treating the nominal band implicitly again. In the following sections, we first extend the notation of the uncertainty of demands introduced for the Γ-robust approach in Section 9.1 and state the compact formulation of the two-band robust WNPP. Afterwards, we evaluate the two-band robustness in comparison to the Γ-robust solutions computed in Section 9.1.

13.1.1. The Compact Formulation

For the two-band robust approach, we assume that the interval $[\bar{w}_t, \bar{w}_t + \hat{w}_t]$ of realisations of the random variables $\tilde{w}_t$ modelling the traffic demand is partitioned into the following two sub-bands.

$$\left[\bar{w}_t, \bar{w}_t + \hat{w}_t^1\right], \left(\bar{w}_t + \hat{w}_t^1, \bar{w}_t + \hat{w}_t^2\right],$$

where $\hat{w}_t^1$ denotes the first deviation value and $\hat{w}_t^2 > \hat{w}_t^1$ the second. Note, the first interval also includes the nominal value. We introduce two robustness parameters Γ_1 and Γ_2 to limit the number of realisations falling in the first band and in the second, respectively.

Following Section 12.1 and applying the standard LP duality manipulation, the compact formulation for the two-band RWNPP can be stated as follows.

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \lambda \sum_{t \in \mathcal{T}} u_t \tag{13.1a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t} z_{st} + u_t = 1 \qquad \forall\, t \in \mathcal{T} \tag{13.1b}$$

$$x_i + x_j \leq 1 \qquad \forall\, ij \in \mathcal{E} \tag{13.1c}$$

$$\sum_{t \in \mathcal{T}_s} \frac{\bar{w}_t}{e_{st}} z_{st} + \Gamma_1 \pi_s^1 + \Gamma_2 \pi_s^2 + \sum_{t \in \mathcal{T}_s} \rho_{st} \leq b_s x_s \qquad \forall\, s \in \mathcal{S} \tag{13.1d}$$

$$\pi_s^1 + \rho_{st} \geq \frac{\hat{w}_t^1}{e_{st}} z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{13.1e}$$

$$\pi_s^2 + \rho_{st} \geq \frac{\hat{w}_t^2}{e_{st}} z_{st} \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T} \tag{13.1f}$$

$$x_s,\, z_{st},\, u_t \in \{0,1\},\, \pi_s^1,\, \pi_s^2,\, \rho_{st} \geq 0 \qquad \forall\, s \in \mathcal{S},\, (s,t) \in \mathcal{S} * \mathcal{T},\, t \in \mathcal{T}_s. \tag{13.1g}$$

Additionally, we apply the vub constraints (4.5) and the maximal clique inequalities (5.9), which are also valid cutting planes for model (13.1).

13.1.2. The Gain of Multi-Band Robustness

In this section, we evaluate the gain we can achieve by the two-band robustness applied to the WNPP in comparison with the Γ-robustness in a proof-of-concept implementation. For that purpose, we use the scenarios with up to 50 BSs investigated in the computational study performed for the d-RWNPP in Section 9.1. All computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of four hours. Moreover, we use the standard version of CPLEX 12.4 [98] as ILP solver.

Following the setting to define test instances for the two-band RKP presented in Section 12.5.1, we assume (almost) equidistant bands and define the two deviation values per TN as follows.

$$\hat{w}_t^2 := \hat{w}_t,\ \hat{w}_t^1 = \left\lceil \frac{\hat{w}_t^2}{2} \right\rceil,$$

where $\hat{w}_t$ is the highest deviation value as generated in Section 9.1.2.

In Section 9.1.5, we have determined the maximum value $\Gamma_{\text{ProL}=100\,\%}$ of the robustness parameter Γ such that no snapshot of 1000 generated for a uniform as well as normal distribution of demand values is violated (ProL = 100 %); cf. Table 9.3. Setting $\Gamma_2 = \Gamma_{\text{ProL}=100\,\%}$ and $\Gamma_1 = 0$ would yield the Γ-robust solution. Thus, we know that any solution for $\Gamma_2 > \Gamma_{\text{ProL}=100\,\%}$ has the same quality as the Γ-robust solution, i. e., ProL = 100 %. Therefore, we investigate only $\Gamma_2 \in \{0, 1, \ldots, \Gamma_{\text{ProL}=100\,\%}\}$.

To determine the range of values to be considered for Γ_1, we extract the number of deployed BSs Σ of the $\Gamma_{\text{ProL}=100\,\%}$-robust solution; cf. Table 9.3. Then we set the maximum value for Γ_1 to

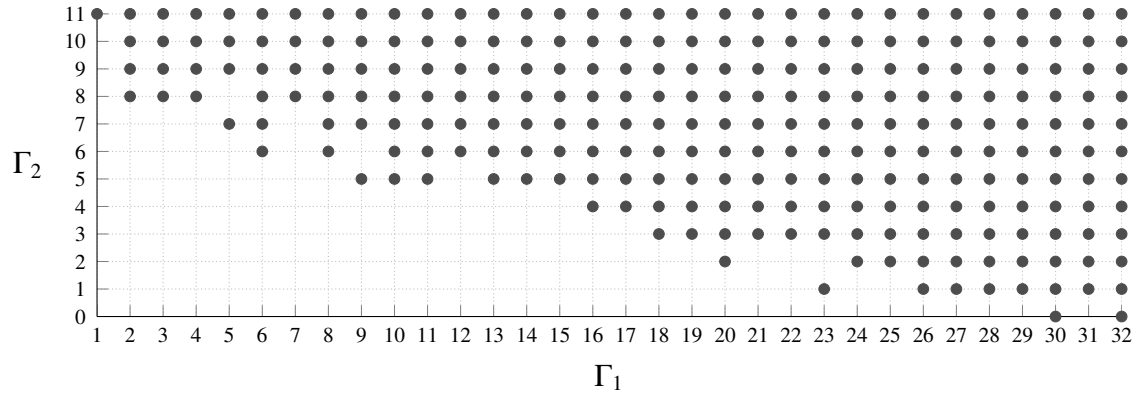
$$\left\lceil \frac{|\mathcal{T}|}{\Sigma} \right\rceil,$$

which is the average number of TNs assigned to a deployed BS if all TNs are covered. As an example, for scenario s_20_200 the Γ-robust solution with $\Gamma_{\text{ProL}=100\,\%} = 12$ deploys four BSs. Hence, we investigate $\Gamma_1 \in \{1, \ldots, 50\}$ and $\Gamma_2 \in \{0, 1, \ldots, 12\}$. Note, we do not consider $\Gamma_1 = 0$ as this setting would just give the Γ-robust solutions. For all five studied scenarios with up to 50 BSs, we depict the maximum considered values for Γ_1 and Γ_2 in Table 13.1.

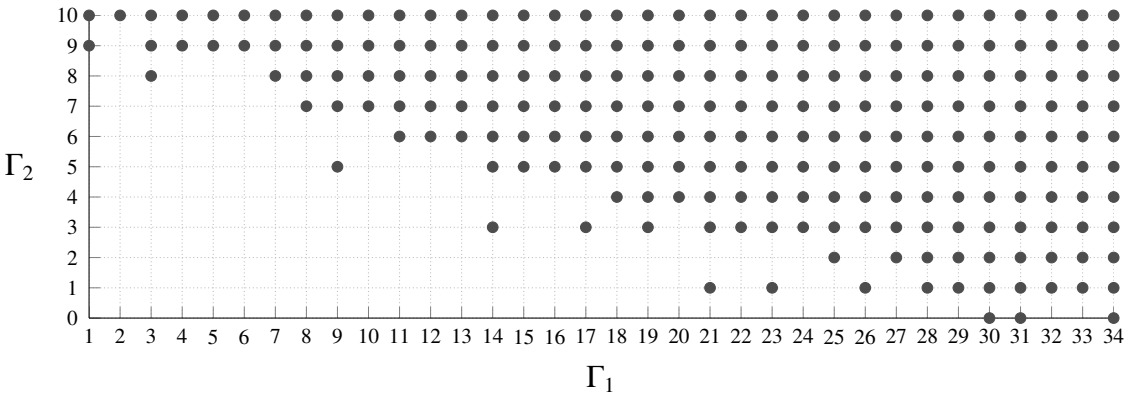
Before evaluating the gain achieved by multi-band robustness, we study the correlation between Γ_1 and Γ_2 exemplarily for s_20_200. To this end, we display those (Γ_1, Γ_2)-pairs, for which the obtained solution yields a protection level of 100 % in Figure 13.1. The

scenario	max. Γ_1	max. Γ_2
s_20_200	50	12
s_30_300	76	12
s_40_400	80	12
s_40_450	90	16
s_50_500	84	18

Table 13.1.: Maximum values of Γ_1 and Γ_2 for all investigated scenarios.



(a) uniform distribution



(b) normal distribution

Figure 13.1.: (Γ_1, Γ_2)-pairs yielding solutions with ProL = 100 %, where the snapshots follow either a uniform or a normal distribution, for scenario s_20_200.

snapshots used to compute ProL are the same as used in the computational study in Section 9.1.5 and obey either an uniform or a normal distribution. Note, the solutions for larger Γ_1 and Γ_2 values, which are not displayed in the figure, are also feasible for all snapshots. The diagrams are just condensed for better readability.

Apart from some outliers for which the solution is feasible for all snapshots by chance, a

scenario	Γ-obj.	obj.	uniform	normal
s_20_200	16	15	(2, 8), (5, 7), (6, 6), (9, 5), (16, 4), (18, 3), (20, 2), (23, 1), (30, 0)	(1, 9), (3, 8), (9, 5), (14, 3), (21, 1), (30, 0)
s_30_300	16	16	(2, 9), (3, 7), (8, 6), (12, 5), (14, 4),(18, 3), (21, 2), (25, 1), (36, 0)	(1, 11), (2, 10), (4, 9), (6, 8), (8, 6), (14, 4), (16, 3), (21, 1), (36, 0)
s_40_400	20	20	(1, 9), (4, 3), (7, 2), (16, 1), (31, 0)	(1, 9), (3, 6), (4, 5), (6, 4), (7, 2), (19, 1), (35, 0)
s_40_450	20	20	(2, 13), (3, 6), (12, 5), (19, 3), (22, 2), (34, 1), (38, 0)	(1, 9), (7, 6), (12, 5), (17, 4), (19, 3), (20, 2), (29, 0)
s_50_500	24*	24*/21	(1, 16), (3, 15), (6, 14), (8, 13), (10, 12), (12, 11), (14, 10), (17, 9), (19, 8), (21, 7), (25, 5), (28, 4), (32, 3), (34, 2), (37, 1), (43, 0)	(8, 11), (20, 6), (28, 3), (38, 0)

Table 13.2.: Non-dominated (Γ_1, Γ_2)-pairs with lowest objective value and ProL = 100 % for a uniform as well as normal distribution in the snapshots for different scenarios and the corresponding Γ-robust objective, where $*$ denotes a non-optimal value.

linear correlation for Γ_1 and Γ_2 becomes apparent. Moreover, for the computational study performed for the two-band RKP in Section 12.5, we had defined $\Gamma_1 = \lceil 2.43 \cdot \Gamma_2 \rceil$. The results in Figure 13.1 indicate that this is, in general, a reasonable choice in case of a normal as well as uniform distribution of demand values. If we decrease the value of Γ_2 by one, we have to increase the value of Γ_1 by three to obtain a solution still feasible for all snapshots.

For the subsequent comparison of two-band robust solutions to Γ-robust solutions, we focus on non-dominated (Γ_1, Γ_2)-pairs with ProL = 100 % and the lowest objective value. By non-dominated, we mean that a decrease of Γ_1 or Γ_2 yields a solution that is not feasible for at least one snapshot. In Table 13.2, we state all non-dominated pairs yielding the lowest objective value for an uniform as well as a normal distribution. Additionally, we state the objective value of the corresponding Γ-robust solution. Only for the smallest scenario s_20_200 and both demand distributions, there exists a two-band robust solution with a lower objective value than the Γ-robust objective. The Γ-robust solution deploys four BSs and serves all 200 TNs, thus yielding an objective value of 16. In case of a uniform (normal) distribution of demand values, 9 (6) non-dominated (Γ_1, Γ_2)-pairs deploy only three BSs while serving 197 TNs yielding an objective value of 15. Moreover, for s_50_500 and a normal distribution, we can also reduce the objective value by 12.5 % by

means of multi-band robustness. However, the time limit is reached for all pairs listed in the column "uniform" and for $(8, 11)$ in the column "normal". Therefore, we cannot be sure if there also exists a (Γ_1, Γ_2)-pair which yields an optimal solution that is feasible for all snapshots generated with an uniform distribution.

The presented results reveal a potential to save cost, i. e., to decrease the objective value of the WNPP, by means of multi-band robustness compared to the Γ-robust solution. However, this gain strongly depends on the test scenario and the Γ-robust solution. On the one hand, if the load of an installed BS is close to the available bandwidth, then the deployed BSs in the Γ-robust solution are more or less the critical number of necessary BSs and some changes in the deviation values do not have a significant impact on the number of installed BSs. On the other hand, if significantly less than the available bandwidth is used in the Γ-robust solution, it is easier to assign all TNs of one BS to others if some deviation values are halved. For the investigated scenarios, the assignment of TNs in the Γ-robust formulation seems not to be a complex task once the BS decision is made. Thus, if it was more challenging, e.g., when the available bandwidth is halved to 5 MHz, there would exist more possibilities to economise one or more BSs. Additionally, an increase of the number of deviation bands, i. e., a further refinement of the deviation interval might also have the potential to save cost.

13.1.3. Conclusion

In this section, we have proposed a compact formulation of the two-band RWNPP and have investigated the gain achieved by this model in comparison to the Γ-robust solutions taken from Section 9.1. By a computational study of a small scenario, we have illustrated a linear correlation between the two robustness parameters Γ_1 and Γ_2. Furthermore, numerical results obtained from a proof-of-concept implementation for five scenarios in total have revealed a potential to save cost via the deployment of less BSs, by means of multi-band robustness in comparison with the Γ-robust solution. However, we have also seen that this gain strongly depends on the test scenario and the Γ-robust solution.

13.2. A Lagrangian Relaxation Approach for a Subproblem of the RWNPP

In Chapter 12, we have developed an improved DP for the multi-band RKP which outperforms CPLEX solving the compact ILP in case of two deviation bands. We study the performance of this DP applied to a wireless network planning problem, the traffic node assignment problem (TNAP), in this chapter. The TNAP is a subproblem of the WNPP introduced in Section 4.1 and aims at assigning TNs to already deployed BSs. We choose the simplified TNAP for the application of the DP for the multi-band RKP as this problem comprehends only one additional type of constraint to the multi-band RKP. We relax these constraints and determine an upper bound by means of Lagrangian relaxation. Similar to a column generation approach, also Lagrangian relaxation can give a better bound than the

LP solution; cf. Table 13.6. In fact, we obtain (close to) optimal solutions.

13.2.1. The Two-Band Robust Traffic Node Assignment Problem

The *two-band robust TNAP* (two-band RTNAP) determines an assignment of TNs to already deployed BSs while maximising the total nominal throughput. To describe this problem formally, we use the same notation as for the two-band RWNPP; cf. Table 4.1 and Section 13.1.1. For a fixed subset $\mathcal{S}' \subseteq \mathcal{S}$ of deployed BSs and $\mathcal{S}'_t := \mathcal{S}_t \cap \mathcal{S}'$ for all $t \in \mathcal{T}$, we propose the following compact formulation of the two-band RTNAP.

$$\max \sum_{s \in \mathcal{S}'} \sum_{t \in \mathcal{T}_s} \bar{w}_t z_{st} \tag{13.2a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}'_t} z_{st} \leq 1 \qquad \forall\, t \in \mathcal{T} \tag{13.2b}$$

$$\sum_{t \in \mathcal{T}_s} \frac{\bar{w}_t}{e_{st}} z_{st} + \Gamma_1 \pi^1_s + \Gamma_2 \pi^2_s + \sum_{t \in \mathcal{T}_s} \rho_{st} \leq b_s \qquad \forall\, s \in \mathcal{S}' \tag{13.2c}$$

$$\pi^1_s + \rho_{st} \geq \frac{\hat{w}^1_t}{e_{st}} z_{st} \qquad \forall\, s \in \mathcal{S}', t \in \mathcal{T}_s \tag{13.2d}$$

$$\pi^2_s + \rho_{st} \geq \frac{\hat{w}^2_t}{e_{st}} z_{st} \qquad \forall\, s \in \mathcal{S}', t \in \mathcal{T}_s \tag{13.2e}$$

$$z_{st} \in \{0, 1\},\ \pi^1_s,\ \pi^2_s,\ \rho_{st} \geq 0 \qquad \forall\, s \in \mathcal{S}', t \in \mathcal{T}_s. \tag{13.2f}$$

The objective (13.2a) maximises the total nominal throughput which is the sum of the satisfied nominal demands $\bar{w}_t$. As the number of served TNs is maximised only indirectly, we omit the variables u_t of the WNPP for simplicity here. Hence, constraints (13.2b) constitute the equivalent to constraints (13.1b) ensuring that a TN is covered by at most one BS. Moreover, constraints (13.2c) to (13.2e) depict the compact reformulation of the two-band robust capacity constraints and are equal to constraints (13.1d) to (13.1f) of the two-band RWNPP, where the variable x_s is dropped.

Without the coverage constraints (13.2b), the MIP (13.2) would just contain several non-correlated two-band RKPs to which we could apply the DP IMPROVED proposed in Chapter 12 directly. However, to allow for the application of this DP, we first have to relax the additional constraints. This is done by Lagrangian relaxation, which has been introduced in Section 1.2.4.

13.2.2. Lagrangian Relaxation for the Two-Band RTNAP

In this section, we apply the Lagrangian relaxation to the two-band RTNAP (13.2) by relaxing the coverage constraints (13.2b). This relaxation results in a two-band RKP with a modified objective function. As there exist extreme points of the LP relaxation of the KP and hence, also of the two-band RKP, which are not integral, there exists the potential that the Lagrangian relaxation gives a strictly better bound than the LP relaxation; cf. Proposition 1.7.

We introduce Lagrange multipliers $\mu_t \geq 0 \; \forall \, t \in \mathcal{T}$ to penalise a constraint violation. The Lagrangian relaxation of the two-band RTNAP can then be formalised as follows.

$$\chi_{\text{LR}}(\mu) \coloneqq \max \sum_{s \in \mathcal{S}'} \sum_{t \in \mathcal{T}_s} \bar{w}_t z_{st} + \sum_{t \in \mathcal{T}} \mu_t \left(1 - \sum_{s \in \mathcal{S}'_t} z_{st} \right) \tag{13.3a}$$

$$(13.2c), (13.2d), (13.2e), (13.2f) \tag{13.3b}$$

As stated before, the Lagrangian relaxation is an upper bound for model (13.2). The best upper bound is defined by the optimal solution of the Lagrangian dual problem

$$\chi_{\text{LD}} \coloneqq \min_{\mu \geq 0} \chi_{\text{LR}}(\mu).$$

To find the optimal solution of this problem, we apply the subgradient algorithm described in Section 1.2.4.

For given multipliers μ_t, $\chi_{\text{LR}}(\mu)$ corresponds to $|\mathcal{S}'|$ two-band RKPs which can be solved separately by the DP IMPROVED developed in Chapter 12. The complete objective is then obtained by adding the separate optimal values of the DPs per BS and the sum over the Lagrange multipliers.

A good lower bound For the definition of the step lengths α_i at iteration i used in the subgradient algorithm, a primal lower bound χ^i, which is the objective value of a feasible solution of (13.2), is required. A straightforward definition of χ^i is the following.

$$\chi^i \coloneqq \sum_{t \in \tau} \bar{w}_t \quad \text{with } \tau \coloneqq \left\{ t \in \mathcal{T} \,\middle|\, \sum_{s \in \mathcal{S}'_t} \tilde{z}^i_{st} \geq 1 \right\}, \tag{13.4}$$

where $\tilde{z}^i$ is a combination of optimal solutions of the two-band RKPs. Since we solve the two-band RKPs for each BS separately, it is most likely that a TN is assigned to more than one BS occupying more capacity than necessary. Hence, we present a two-step heuristic to improve the lower bound in the following.

In the first step, we determine TNs which are currently assigned to more than one BS and define a serving BS, where this decision is based on the potential of the BS to serve further TNs. If no BS fulfils this decision criterion, we select the BS with the lowest remaining capacity as the serving BS. By the first step, we have gained capacity at some BSs. To exploit the capacity completely, we try to assign not yet covered TNs by running a DP consecutively for every BS with available capacity in the second step of the heuristic. We now give a detailed description of the two steps.

Step 1 First, we compute the subgradient or slack ξ^i_t for each TN t at iteration i, where $\tilde{z}^i_{st}$ is the current optimal solution:

$$\xi^i_t = 1 - \sum_{s \in \mathcal{S}'_t} \tilde{z}^i_{st}.$$

If $\xi_t^i \geq 0$, the relaxed constraint (13.2b) is satisfied for TN t. For all remaining TNs, we define the set of BSs to which the TN is currently assigned $\Sigma_t \coloneqq \{s \in \mathcal{S}'_t \,|\, \tilde{z}_{st}^i = 1\}$. Among these BSs, we have to specify the serving BS. To this end, we define the set of free TNs per BS at iteration i as

$$\tau_s^i \coloneqq \left\{t \in \mathcal{T}_s \,\middle|\, \tilde{z}_{st}^i = 0\right\} \quad \forall\, s \in \mathcal{S}',$$

which is the set of TNs potentially served by s but not in the current solution. If there exists a BS $\sigma \in \Sigma_t$ with $|\tau_\sigma^i| = 0$ for a TN t with $\xi_t^i \leq -1$, then σ is assumed to be the serving BS. Such a BS has no free TNs left and hence, no potential to serve further TNs. If no such BS exists, we compute the load ℓ_s of each $s \in \Sigma_t$ as

$$\ell_s = \mathrm{rhs}_s + \Gamma_1 \tilde{\pi}_s^{1,i} + \Gamma_2 \tilde{\pi}_s^{2,i},$$

where rhs_s denotes the optimal right hand side value, which is found among the $b \leq B - \Gamma_1\pi_1 - \Gamma_2\pi_2$ in the notation of Chapter 12; cf. the maximisation problem (12.8). Moreover, $\tilde{\pi}_s^{1,1}$, $\tilde{\pi}_s^{2,1}$ denote the optimal values of the dual variables computed via the DP IMPROVED. Then, we define the serving BS $\sigma \in \Sigma_t$ for TN t as the BS with the lowest remaining capacity, i. e.,

$$\sigma = \operatorname*{argmin}_{s \in \Sigma_t} \{b_s - \ell_s\}. \tag{13.5}$$

We summarise the first step of the heuristic by constructing a new solution $\hat{z}^i$, which is feasible for the two-band RTNAP (13.2).

$$\hat{z}_{st}^i = \begin{cases} 1 & \text{if } (\xi_t^i = 0 \wedge \tilde{z}_{st}^i = 1) \text{ or } s \text{ is the serving BS defined in (13.5)}, \\ 0 & \text{otherwise.} \end{cases}$$
$$\forall\, (s,t) \in \mathcal{S}' * \mathcal{T}$$

Step 2 After the serving BSs are fixed for all TNs with a negative slack, we remove the multiple assignments and thus, gain capacity at some BSs. To exploit the total capacity, we now attempt to assign each TN t with $\sum_{s \in \mathcal{S}'_t} \hat{z}_{st}^i = 0$ to a BS s with $\ell_s < b_s$. Since the realisation of the uncertain traffic demands is not known, we cannot just add TNs to BSs. Instead, we run a DP IMPROVED for each of the potential BSs respecting the fixed assignments and with objective

$$\max \sum_{t \in \tau_s^i \colon \sum_{s \in \mathcal{S}'_t} \hat{z}_{st}^i = 0} \bar{w}_t x_t.$$

As soon as a TN t is assigned to a BS σ in such a post-processing DP, we set $\hat{z}_{\sigma t}^i = 1$ and remove the TN from τ_s^i for all $s \in \mathcal{S}'_t \setminus \{\sigma\}$.

In total, we set

$$\chi^i = \sum_{t \in \mathcal{T}} \bar{w}_t \left(\sum_{s \in \mathcal{S}'_t} \hat{z}^i_{st} \right), \tag{13.6}$$

which is clearly at least as good as the bound defined in (13.4).

Implementation details and parameter settings In the introduction to Lagrangian relaxation in Section 1.2.4, the initial values of the Lagrange multipliers for the subgradient algorithm are set to zero. However, we can define better initial multipliers μ^0_t following the approach by Fisher [77] and using the subsequent reformulated objective. The objective (13.3a) can equivalently be written as

$$\max \sum_{s \in \mathcal{S}'} \sum_{t \in \mathcal{T}_s} (\bar{w}_t - \mu_t) z_{st} + \sum_{t \in \mathcal{T}} \mu_t. \tag{13.7}$$

For Lagrange multipliers μ with $\mu_t \geq \bar{w}_t \, \forall \, t \in \mathcal{T}$, the optimal solution of (13.3) is $z = 0$ since $\bar{w}_t \geq 0$. Hence, setting $\mu^0_t = \bar{w}_t$ minimises the upper bound $\chi_{\text{LR}}(\mu)$ for all μ for which $z = 0$ is an optimal solution of (13.3). Clearly, this yields better initial values than $\mu^0_t = 0$.

To apply the stop criteria specified for the subgradient algorithm, we have to define the necessary parameters. We set $\varepsilon_1 = 0.001$ ensuring that the algorithm stops as soon as the optimality gap is below 0.1 %. Furthermore, as we are only interested in integer solutions, we round the upper bound in the computation of the gap, i. e., $\lfloor \chi_{\text{LR}}(\mu^i) \rfloor$. To improve the optimality gap further, we use the overall best lower bound χ, which is updated only if the current lower bound χ^i is higher.

Moreover, we define $\varepsilon_3 = 0.001$ for the precision of the step lengths. Instead of using ε_2 directly to control the improvement of the upper bound, we exploit the precision set for the underlying algorithm "MinKnap" in the DP IMPROVED; cf. Pisinger [153]. This algorithm is developed only for integer objective coefficients. Due to the non-integer Lagrange multipliers, we have to scale all coefficients by a parameter ε_2 and decide to set $\varepsilon_2 = 1000$. Hence, if $|\mu^{i+1}_t - \mu^i_t| < 1/\varepsilon_2$, the precision used in "MinKnap" is not sufficient and we stop the subgradient algorithm.

Finally, we set

$$\delta^i = \begin{cases} 2 & i = 0 \\ \frac{\delta^{i-1}}{2} & \text{if } \delta^{i-1} = \delta^{i-2} = \delta^{i-3} = \delta^{i-4}, i \geq 3 \\ \delta^{i-1} & \text{otherwise,} \end{cases}$$

i. e., we halve the scaling parameter δ^i if the upper bound was consecutively not improved three times.

13.2.3. Numerical Evaluation

In this section, we investigate the quality of the upper bound for the two-band RTNAP obtained by the Lagrangian relaxation described in the previous section in a proof-of-

scenario	(Γ_1, Γ_2)	reason for selection
s_20_200	$(9, 5)$	non-dom. for uniform and normal
s_30_300	$(8, 6)$	non-dom. for uniform and normal
s_40_400	$(1, 9)$	non-dom. for uniform and normal
s_40_450	$(2, 13)$	lowest time for uniform and dom. for normal
s_50_500	$(20, 8)$	lowest time for normal: $(20, 6)$, take next feasible for uniform

Table 13.3.: Selected (Γ_1, Γ_2)-pairs and reason for selection per scenario.

concept implementation. To this end, we select one (Γ_1, Γ_2)-pair per scenario, for which we have obtained a feasible solution for both demand distributions in Chapter 13.1 and which yields the lowest objective value. Among several possibilities, we select the pair with the lowest solving time. More precisely, if one of the pairs depicted in Table 13.2, is non-dominated for the uniform as well as normal demand distribution, we select this (Γ_1, Γ_2)-pair here. This is the case, e. g., for s_20_200 where we select $(\Gamma_1, \Gamma_2) = (9, 5)$. If such a pair does not exist, we check if a pair with the lowest solving time for the uniform (normal) distribution is also feasible (but dominated) for the normal (uniform) distribution. For instance, the pair $(2, 13)$ consumes the lowest time among the non-dominated pairs for an uniform distribution for s_40_450 and is dominated by $(1, 9)$, which is a non-dominated pair for a normal distribution. Only for scenario s_50_500 we vary this selection routine as there does not exist a feasible pair with the lowest time consumption that is also feasible for the other distribution. Hence, we choose a pair feasible for both distributions and that is as close as possible to a pair with the lowest solving time. The selected pairs per scenario and the reason for their selection are displayed in Table 13.3. Based on these (Γ_1, Γ_2)-pairs, we use the corresponding optimal solution $(\tilde{x}, \tilde{z}, \tilde{u}, \tilde{\pi}^1, \tilde{\pi}^2)$ of the compact formulation (13.1) of the two-band RWNPP to define the set of deployed BSs $\mathcal{S}' := \{s \in \mathcal{S} \mid \tilde{x}_s = 1\}$. Additionally, we round all ratios of demand over spectral efficiency up to obtain only integer values which are required by the DP.

All computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a memory limit of 11 GB. The objective value of the Lagrangian dual problem χ_{LD} is obtained by $\lfloor \chi_{\text{LR}} \rfloor$ where χ_{LR} is the best upper bound obtained by the Lagrangian relaxation. Furthermore, we denote by χ the best lower bound of the Lagrangian relaxation solving process.

We first investigate the effect of the heuristic to obtain an improved lower bound. In Table 13.4, we display the rounded upper bounds $\lfloor \chi_{\text{LR}} \rfloor$ and the lower bounds obtained by either dis- or enabling the heuristic. The lower bound χ obtained when the heuristic is enabled is always better than the lower bound obtained without the heuristic. For scenarios s_40_400 and s_40_450, the upper bound $\lfloor \chi_{\text{LR}} \rfloor$ is slightly worse when the heuristic is enabled. This is the case as the Lagrangian relaxation stops since the gap limit of 0.1 % is reached due to the good lower bound. This is also the reason for the high time savings of around 50 % for these scenarios. The presented results demonstrate that the application of the proposed heuristic improves the solving process of the Lagrangian relaxation while

scenario	no heuristic		with heuristic		% time decrease
	χ_{LD}	$\underline{\chi}$	χ_{LD}	$\underline{\chi}$	
s_20_200	51674	51522	51674	51613	-9.63
s_30_300	80548	80283	80548	80366	-9.27
s_40_400	109295	109173	109304	109200	55.18
s_40_450	120469	120250	120474	120424	47.41
s_50_500	139484	138953	139484	139289	-3.62

Table 13.4.: Upper and lower bounds $\lfloor \chi_{\mathrm{LR}} \rfloor$ and $\underline{\chi}$ when the heuristic is disabled and enabled and the percentage time decrease by the application of the heuristic.

scenario	χ_{LD}	$\underline{\chi}$	gap (%)	time (s)	# iter.	ILP time (s)
s_20_200	51674	51613	0.12	23.8	83	2.4
s_30_300	80548	80366	0.23	77.1	90	24.9
s_40_400	109304	109200	0.10	89.2	38	71.6
s_40_450	120474	120424	0.04	157.9	44	10.5
s_50_500	139484	139289	0.14	630.0	101	3442.3

Table 13.5.: Upper and lower bounds $\lfloor \chi_{\mathrm{LR}}(\mu^i) \rfloor$ and $\underline{\chi}$, integrality gap, solving time, and number of iterations for the Lagrangian relaxation and the solving time of the ILP (13.2) with precision 0.001 per scenario.

the time consumption is increased just moderately. Thus, we enable the heuristic in the subsequent investigations.

In Table 13.5, we display the best upper bound χ_{LD}, the best lower bound $\underline{\chi}$, the integrality gap, the time consumption and the number of iterations of the Lagrangian relaxation. Note, the gap is the percentage value of $(\chi_{\mathrm{LD}} - \underline{\chi})/\chi_{\mathrm{LD}}$. As seen before, by the definition of ε_1 scenarios s_40_400 and s_40_450 are solved to optimality with precision 0.001 while the solutions for s_20_200, s_30_300 and s_50_500 are close to optimal. Moreover, the time consumption is reasonable with less than 630 s for all investigated scenarios. Comparing these times to the time consumed when solving the ILP (13.2) with the same precision, the Lagrangian relaxation is (slightly) slower for scenarios with up to 40 BSs but significantly faster for the largest scenario. In summary, the Lagrangian relaxation gives (close to) optimal upper bounds in an acceptable time for all investigated scenarios.

Finally, we compare the upper bound obtained from the LP relaxation χ_{LP} to the best bound of the Lagrangian relaxation χ_{LD} in terms of gap closed. Similarly to former computational studies the percentage gap closed is defined as

$$\text{gap closed} := \frac{\chi_{\mathrm{LD}} - \chi_{\mathrm{LP}}}{\mathrm{ILP} - \chi_{\mathrm{LP}}},$$

with ILP denoting the best known primal bound obtained from solving the ILP (13.2). In Table 13.6, we present the best known primal bound of the ILP, the LP relaxation χ_{LP}, the

scenario	ILP	χ_{LP}	χ_{LD}	gap closed (%)
s_20_200	51647	52705.67	51674	97.5
s_30_300	80504	82429.67	80548	97.7
s_40_400	109295	112349.86	109304	99.7
s_40_450	120457	123431.62	120474	99.4
s_50_500	139397	143646.40	139484	98.0

Table 13.6.: Best known primal bound of the ILP, LP relaxation value χ_{LP}, upper bound of the Lagrangian relaxation χ_{LD} and the percentage gap closed per scenario.

Lagrangian relaxation χ_{LD} and the percentage gap closed. Note that all ILP values are optimal apart from scenario s_50_500 which is not solved to optimality due to exceeding the available memory. The upper bound obtained by the Lagrangian relaxation is significantly better than the bound of the LP relaxation. In fact, by the Lagrangian relaxation we can close 97.5 to 99.7 % of the optimality gap given by the LP relaxation for all scenarios. These good results support our choice to relax the coverage constraints (13.2b) and apply the Lagrangian relaxation approach with a subgradient algorithm.

13.2.4. Conclusion and Outlook

In this chapter, we have proposed a compact formulation of the two-band RTNAP whose aim is to assign all TNs to already deployed BSs. By means of Lagrangian relaxation, we have obtained several uncorrelated two-band RKPs which are solved separately by the DP IMPROVED proposed in Chapter 12. Moreover, we have proposed a two-step heuristic to improve the lower bound. Numerical results performed on five scenarios have demonstrated that the heuristic improves the lower bound such that the optimal tolerance is reached for two scenarios. Additionally, the Lagrangian relaxation gives (close to) optimal solutions within an acceptable time, where it is even significantly faster than the compact ILP for the largest scenario. Finally, the computational results have corroborated the theory that the Lagrangian dual problem can give a (significantly) better bound than the LP relaxation as we have obtained a percentage gap closed of at least 97.5 %.

These good results entail the question of the performance of the Lagrangian relaxation approach when applied to the complete two-band RWNPP. For future investigations, the conflict graph constraints in the two-band RWNPP can also be relaxed such that the resulting subproblem again only consists of separate two-band RKPs. However, the Lagrangian relaxation would consume significantly more time as more BSs would have to be considered than in the study presented in this section. Here, we have only considered those BSs which are deployed in a pre-computed solution. This number is at most 6 while it would be 50 for the RWNPP.

Moreover, the upper bound computed by Lagrangian relaxation is not optimal for every investigated scenario. Hence, we now briefly describe a branching routine which remains as further future work.

As the upper bound $\chi_{\mathrm{LR}}(\mu^i)$ is calculated via the independent solving of uncorrelated two-band RKPs, it is likely that TNs are assigned to more than one BS. To find an optimal solution for the two-band RTNAP, we have to branch on the multiple assigned TNs. For instance, a TN t is assigned to BS s as well as σ. Then, we create three child nodes containing the following restrictions.

1. $z_{st} = 1 \wedge z_{\sigma t} = 0$; TN t is assigned to BS s but not to σ,
2. $z_{st} = 0 \wedge z_{\sigma t} = 1$; TN t is not assigned to BS σ but to s,
3. $z_{st} = 0 \wedge z_{\sigma t} = 0$; TN t is neither assigned to BS s nor to σ.

In every child node, the subgradient algorithm is restarted and by means of these branching decisions, the size of the DPs can be reduced especially when further descending the branching tree. Moreover, this process can be extended to a full B&B routine by introducing and updating global bounds. The crucial part of the described B&B routine is, on the one hand, to decide on which TN should be branched next and on the other hand, to decide the order in which open branching nodes are processed.

Part V.

Recoverable Robustness

14. General Concept

Recoverable robustness is a quite recent two-stage robustness approach, which was first introduced in a technical report by Liebchen et al. [124] in 2007 and applied to railway optimisation problems with formalisation in [125] two years later. In the classical robust optimisation, a solution is required to be immune against small deviations in the data. A less strict requirement is the possibility to recover solutions after a change in the data, which is included in recoverable robustness. Such an adjustment is regulated by predefined rules.

More precisely, a first-stage solution can be adapted (recovered) in order to make it feasible in a second stage once a realisation of uncertain data, which is called scenario, is detected. Such a recovery action is algorithmically limited, i. e., possible modifications of the solution are predefined, and implies cost which have to be taken into account in the recoverable robust solution. Typical scenarios are, e. g., discrete [37], interval [37], or Γ-robust [36], while Büsing [33] studies further scenario sets.

Recoverable robustness is a quite general framework, hence, widely applied and studied for different problems. We give a brief overview on the literature studying (aspects of) recoverable robustness for problems related to this thesis without claiming exhaustiveness.

Bouman et al. [30] study the knapsack problem with uncertain capacity and the demand robust shortest path problem, where the location of the sink and the cost of the edges may change over time. The authors apply column generation to solve the recoverable robust versions of these problems. Another decomposition approach, Benders decomposition, is used by Cacchiani et al. [38] to assess the Price of Recoverability for railway rolling stock planning. Alvarez-Miranda et al. [7] study a recoverable robust two-level network design problem with two available technologies, where the network topology is determined in the first stage. Uncertainty is modelled by means of discrete scenarios. The KP has also been studied in terms of recoverable robustness. In Büsing et al. [37], the authors investigate the recoverable RKP with discrete scenarios while they study Γ-scenarios in [36]. Moreover, Kutschka [118] presents a polyhedral study of the recoverable RKP with a discrete as well as Γ-robust scenario set and determines valid and facet defining inequalities such as recoverable robust (strengthened/extended) cover inequalities.

In the following section, we briefly present the basic concept of recoverable robustness based on [118, 125]. Afterwards, we state the recoverable robust counterpart and possibilities to evaluate the robustness.

The Basic Principle The aim is to find a solution to an optimisation problem which can be made feasible by a limited effort for a limited set of scenarios [125]. Thus, a recoverable robust problem consists of an (original) optimisation problem, realisations of

uncertain data (scenarios) and limited recovery possibilities.

Let K define the *set of scenarios*, where each $k \in K$ represents a realisation of the uncertain data. These realisations are summarised in the *recoverable robust uncertainty set* $\mathcal{U}^K$. Then, the *recoverable robust optimisation problem* consists of the following two stages; cf. [118].

1. A first-stage solution x^0 of the original optimisation problem before the realisation of uncertain data is known.

2. A second-stage solution x^k, which modifies the first-stage solution x^0 according to a predefined recovery rule to recover feasibility of x^0 according to the realisation k of uncertain data.

By $\mathcal{R}(x^0, k)$, we denote the *set of feasible recovered solutions* of the first-stage solution x^0 for scenario k. Hence, $x^k \in \mathcal{R}(x^0, k)$.

The Recoverable Robust Counterpart For a standard minimisation LP in row by row representation

$$\begin{aligned} &\min c^t x \\ &\text{s.t. } a_i x \geq b_i && \forall\, i \in \{1, \dots, m\} \\ &\quad x \in \mathbb{R}^n_{\geq 0}, \end{aligned} \tag{14.1}$$

we extend the notation for its recoverable robust counterpart as follows. The first stage is defined as $c^0 = c$, $a_i^0 = a_i$, and $b_i^0 = b$. Moreover, for each scenario $k \in K$, the second stage is defined as $c^k \in \mathbb{R}^n$, $a_i^k \in \mathbb{R}^n$, and $b_i^k \in \mathbb{R}^m$. The *recoverable robust counterpart* of (14.1) can then be formalised as

$$\min (c^0)^t x^0 + \max_{k \in K} \left\{ (c^k)^t x^k \right\} \tag{14.2a}$$

$$\text{s.t. } a_i^0 x^0 \geq b_i^0 \qquad \forall\, i \in \{1, \dots, m\} \tag{14.2b}$$

$$a_i^k x^k \geq b_i^k \qquad \forall\, i \in \{1, \dots, m\},\ k \in K \tag{14.2c}$$

$$x^k \in \mathcal{R}(x^0, k) \qquad \forall\, k \in K \tag{14.2d}$$

$$x^0 \in \mathbb{R}_{\geq 0}. \tag{14.2e}$$

The objective (14.2a) minimises the first-stage cost c^0 plus the worst-case second-stage cost c^k. Constraints (14.2b) are the first-stage constraints, while (14.2c) are the second-stage constraints. Furthermore, the recovery action is implemented by constraints (14.2d) which also include the non-negativity requirement of the second-stage solution x^k. The worst-case second-stage cost are called the *robust recovery cost*.

Evaluation of Recovery The price of robustness defined in (8.13) gives the percentage deterioration of the objective value required to guarantee feasibility by means of the Γ-robustness approach. For the definition, the robust objective value is compared to the

nominal problem which considers only nominal values. However, a less conservative view is to consider the objective value of the nominal problem which respects the worst-case setting. Then, the price of robustness rather defines the *gain of robustness* as it states the percentage cost savings achieved by the robust solution, which is less conservative and more resource efficient.

Similar to the gain of robustness, we define the *gain of recovery* (GoR) for a recoverable robust solution with objective value $z^{\mathcal{R}}$:

$$\text{GoR} := \frac{|z^K - z^{\mathcal{R}}|}{|z^K|}, \tag{14.3}$$

with z^K denoting the objective value of the nominal problem where no recovery is allowed but all scenarios have to be satisfied simultaneously. This optimisation problem is given by LP (14.2) and setting $x^k = x^0 \; \forall \, k \in K$. Therefore, GoR gives the percentage improvement of the objective value by allowing recovery; cf. [36, 118].

15. The Recoverable RWNPP

In this chapter, we apply the recoverable robustness approach presented in the previous chapter to the WNPP defined in Section 4.1. Our aim is to save cost (or energy) by the deactivation of BSs during low traffic times such as nights. Since the network design is proposed for the maximum throughput and BSs consume a large amount of their total power consumption when they are not serving any TN, there exists a large potential to increase the energy efficiency of the communication network by switching BSs off.

The improvement of the energy efficiency of wireless communication networks has gained a great deal of attention during the last several years. Investigated levels of the communication system in terms of energy savings are the BS components, the links and the network itself; cf. Correia et al. [54]. Self-organising networks (SONs) provide support for several energy saving approaches currently developed. The switching off requires adaptions and compensating mechanisms for the neighbouring cells as pointed out by Blume et al. [28], where cooperative communication and power control among BSs is necessary [94]. A survey on green strategies, the architecture of BSs, operational cost and energy consumption of BSs in terms of energy efficiency is presented in [133]. Recent works on strategies to turn off BSs during low traffic times have been proposed in, e. g., [28, 44, 54, 97, 147, 161] and the references therein.

While the most recent works on energy aware radio networks make use of the advanced (communication) techniques of 4G networks, our approach based on recoverable robustness can be applied to any cellular access network. Moreover, it gives a benchmark of the achievable savings by means of deactivating BSs.

15.1. Formulation

To limit the number of variables necessary for the recoverable robust formulation, we replace the non-coverage variable u_t by $1 - \sum_{s \in \mathcal{S}_t} z^0_{st}$ for all $t \in \mathcal{T}$, where z^0_{st} denotes the first stage solution variables. Such a replacement is possible due to the coverage constraint (4.1). Hence, the objective (4.3) is reformulated to

$$\min \sum_{s \in \mathcal{S}} c_s x^0_s + \lambda \sum_{t \in \mathcal{T}} \left(1 - \sum_{s \in \mathcal{S}_t} z^0_{st} \right) = \lambda |\mathcal{T}| + \min \sum_{s \in \mathcal{S}} c_s x^0_s - \lambda \sum_{(s,t) \in \mathcal{S} * \mathcal{T}} z^0_{st}, \tag{15.1}$$

with the first stage deployment indicator variables x^0_s.

To apply the recoverable robustness approach, we introduce a finite set of *discrete* scenarios K such that for each scenario $k \in K$ a TN t requires demand w^k_t and we save

cost $c_s^k > 0$ if BS s is not operating for this scenario. Moreover, a TN that is not covered in scenario k is penalised by $\lambda^k \geq 0$.

A feasible first stage solution $(x^0, z^0) \in \{0,1\}^{|\mathcal{S}|} \times \{0,1\}^{|\mathcal{S}*\mathcal{T}|}$ satisfies the coverage constraints

$$\sum_{s \in \mathcal{S}_t} z_{st}^0 \leq 1 \quad \forall\, t \in \mathcal{T}, \tag{15.2}$$

the first-stage capacity constraints

$$\sum_{t \in \mathcal{T}_s} \frac{w_t^0}{e_{st}} z_{st}^0 \leq b_s x_s^0 \quad \forall\, s \in \mathcal{S}, \tag{15.3}$$

with first-stage demands w_t^0, and the maximal clique inequalities

$$\sum_{s \in \mathcal{U}} x_s^0 \leq 1 \quad \forall\, \mathcal{U} \subset \mathcal{S},\ \mathcal{U} \text{ is a max. clique in } G. \tag{15.4}$$

For each scenario $k \in K$, we define the second-stage variables as follows.

$$x_s^k = \begin{cases} 1, & \text{BS } s \text{ is deployed in scenario } k \\ 0, & \text{otherwise} \end{cases}$$

$$z_{st}^k = \begin{cases} 1, & \text{TN } t \text{ is assigned to BS } s \text{ in scenario } k \\ 0, & \text{otherwise} \end{cases}$$

In a feasible second-stage solution (x^k, z^k), a BS can only be installed if it has already been deployed in the first-stage solution. This means, we only permit the deactivation of BSs and not the deployment of new BSs in the recovery action. Furthermore, a TN can be reassigned in the second stage but only if the serving BS of the first stage is switched off for the considered scenario. These requirements are formalised in the following constraints.

$$x_s^k \leq x_s^0 \qquad \forall\, s \in \mathcal{S}, k \in K \tag{15.5}$$

$$z_{st}^k \geq z_{st}^0 + x_s^k - 1 \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T}, k \in K \tag{15.6}$$

Finally, the second-stage equivalents of the coverage constraints (15.2)

$$\sum_{t \in \mathcal{T}} z_{st}^k \leq 1 \quad \forall\, t \in \mathcal{T}, k \in K \tag{15.7}$$

and the capacity constraints

$$\sum_{t \in \mathcal{T}_s} \frac{w_t^k}{e_{st}} z_{st}^k \leq b_s x_s^k \quad \forall\, s \in \mathcal{S}, k \in K \tag{15.8}$$

have to be satisfied by a second-stage solution (x^k, z^k). Note, since we do not deploy more BSs in the second stage than in the first stage, there is no need for an equivalent of the

maximal clique inequalities per scenario.

For each scenario $k \in K$ and a given first-stage solution (x^0, z^0), the set of all feasible recoveries is defined as

$$\mathcal{R}(x^0, z^0, k) := \left\{(x^k, z^k) \in \{0,1\}^{|\mathcal{S}|} \times \{0,1\}^{|\mathcal{S}*\mathcal{T}|} \,\middle|\, (15.5) - (15.8) \text{ are satisfied}\right\}, \tag{15.9}$$

while the robust recovery cost of a first-stage solution (x^0, z^0) is given by

$$C(x^0, z^0) := \max_{k \in K} \min_{(x^k, z^k) \in \mathcal{R}(x^0, z^0, k)} \left\{ \sum_{s \in \mathcal{S}(x^0 - x^k)} -c_s^k + \lambda^k |\mathcal{T}| - \sum_{(s,t) \in \mathcal{S}*\mathcal{T}(z^k)} \lambda^k \right\}. \tag{15.10}$$

Here, $\mathcal{S}(x^0 - x^k)$ denotes the set of BS indices associated to $x^0 - x^k$, i. e., $x_s^0 = 1$ and $x_s^k = 0$ if and only if $s \in \mathcal{S}(x^0 - x^k)$, and $\mathcal{S} * \mathcal{T}(z^k)$ denotes the set of BS-TN pairs associated to z^k, i. e., $z_{st}^k = 1 \Leftrightarrow (s,t) \in \mathcal{S} * \mathcal{T}(z^k)$.

Thus, the complete model of the recoverable RWNPP with a discrete uncertainty set can be stated as follows.

$$\lambda|\mathcal{T}| + \min \sum_{s \in \mathcal{S}} c_s x_s^0 - \lambda \sum_{(s,t) \in \mathcal{S}*\mathcal{T}} z_{st}^0 - \omega \tag{15.11a}$$

$$\text{s.t.} \sum_{s \in \mathcal{S}_t} z_{st}^0 \leq 1 \qquad \forall\, t \in \mathcal{T} \tag{15.11b}$$

$$\sum_{t \in \mathcal{T}_s} \frac{w_t^0}{e_{st}} z_{st}^0 \leq b_s x_s^0 \qquad \forall\, s \in \mathcal{S} \tag{15.11c}$$

$$\sum_{s \in \mathcal{U}} x_s^0 \leq 1 \qquad \forall\, \mathcal{U} \subset \mathcal{S},\ \mathcal{U} \text{ is a max. clique in } G \tag{15.11d}$$

$$\sum_{s \in \mathcal{S}_t} z_{st}^k \leq 1 \qquad \forall\, t \in \mathcal{T}, k \in K \tag{15.11e}$$

$$x_s^k \leq x_s^0 \qquad \forall\, s \in \mathcal{S}, k \in K \tag{15.11f}$$

$$z_{st}^k \geq z_{st}^0 + x_s^k - 1 \qquad \forall\, (s,t) \in \mathcal{S} * \mathcal{T},\ s \in \mathcal{S},\ k \in K \tag{15.11g}$$

$$\sum_{t \in \mathcal{T}_s} \frac{w_t^k}{e_{st}} z_{st}^k \leq b_s x_s^k \qquad \forall\, s \in \mathcal{S}, k \in K \tag{15.11h}$$

$$\sum_{s \in \mathcal{S}} c_s^k (x_s^k - x_s^0) - \lambda^k \sum_{(s,t) \in \mathcal{S}*\mathcal{T}} z_{st}^k + \omega \leq -\lambda^k |\mathcal{T}| \qquad \forall\, k \in K \tag{15.11i}$$

$$x_s^0,\ z_{st}^0,\ x_s^k,\ z_{st}^k \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\ (s,t) \in \mathcal{S} * \mathcal{T},\ k \in K \tag{15.11j}$$

$$\omega \in \mathbb{R} \tag{15.11k}$$

Variable ω denotes the recovery cost defined in (15.10) while we include the maximisation by means of constraints (15.11i). Compared to the basic model (4.4) together with the vub constraints (4.5), the mci (5.9) and the replacement of the non-coverage variables u_t by $1 - \sum_{s \in \mathcal{S}_t} z_{st}$, the recoverable RWNPP formulation (15.11) has $|K| \cdot (|\mathcal{S}| + |\mathcal{S} * \mathcal{T}|) + 1$ additional variables and $|K| \cdot (|\mathcal{T}| + 2|\mathcal{S}| + |\mathcal{S} * \mathcal{T}| + 1)$ extra constraints.

Gain of Recovery To evaluate the benefit obtained by a recovery action, we use the measure of the gain of recovery introduced in (14.3). Thus, we now define the standard robust formulation of the WNPP where no recovery is allowed but all scenarios have to be fulfilled simultaneously.

$$\lambda|\mathcal{T}| + \min \sum_{s\in\mathcal{S}} c_s x_s^0 - \lambda \sum_{(s,t)\in\mathcal{S}*\mathcal{T}} z_{st}^0 - \omega \tag{15.12a}$$

$$\text{s.t. (15.2), (15.3), (15.4)} \tag{15.12b}$$

$$\sum_{t\in\mathcal{T}_s} \frac{w_t^k}{e_{st}} z_{st}^0 \leq b_s x_s^0 \qquad \forall\, s \in \mathcal{S}, k \in K \tag{15.12c}$$

$$-\lambda^k \sum_{(s,t)\in\mathcal{S}*\mathcal{T}} z_{st}^0 + \omega \leq -\lambda^k|\mathcal{T}| \qquad \forall\, k \in K \tag{15.12d}$$

$$x_s^0,\, z_{st}^0, \in \{0,1\} \qquad \forall\, s \in \mathcal{S},\, (s,t) \in \mathcal{S} * \mathcal{T},\, k \in K \tag{15.12e}$$

$$\omega \in \mathbb{R}. \tag{15.12f}$$

In the notation of the definition of GoR in (14.3), formulation (15.11) yields objective value $z^{\mathcal{R}}$ and (15.12) yields objective value z^K.

15.2. Numerical Evaluation

In this section, we evaluate the GoR numerically for five test instances with up to 50 BSs and 500 TNs defined in the computational study of the d-RWNPP in Section 9.1.2 with $e_{\min} = 0.25$. To this end, we define two different scenario sets and appropriate values of the four parameters, BS cost c_s, cost savings c_s^k and the scaling or penalty parameters λ and λ^k.

Based on Table 9.1 stated for the creation of test instances for the evaluation of the d-RWNPP in Section 9.1.2, we extend the set of possible traffic profiles by the profile "mix" which is defined as 2/3 normal + 1/3 high. As an example, the lowest possible value following the profile "mix" is

$$\frac{2}{3}(10\,\% \cdot 512 + 20\,\% \cdot 128 + 70\,\% \cdot 64) + \frac{1}{3}(30\,\% \cdot 512 + 40\,\% \cdot 128 + 30\,\% \cdot 64) = 156\,\text{kbps}.$$

We define two scenario sets representing one week each as follows. The planning demands w_t^0, which denote the daily demand, are computed according to the traffic profile "high". Moreover, each night (eight hours) per week is defined as one scenario k with demand w_t^k computed via traffic profile "normal" or "mix" and we denote the respective scenario set accordingly. Thus, $|K| = 7$ for scenario set "normal" as well as "mix".

In previous studies of the WNPP, we have assumed $c_s = 4\,\forall\, s \in \mathcal{S}$ and $\lambda = 1$; see, for instance, Section 5.9.1. For the recoverable RWNPP, we need a wider scope for these parameters to find a good trade-off between minimising the number of deployed BSs and maximising the number of served TNs. Hence, we set $c_s = 40$ and $\lambda = 10$ keeping the

	"normal"			"mix"		
instance	min.	av.	max.	min.	av.	max.
s_20_200	0.0%	0.0%	0.0%	0.0%	0.1%	0.3%
s_30_300	0.1%	0.4%	1.6%	0.0%	0.5%	0.8%
s_40_400	0.1%	1.1%	1.9%	0.2%	0.9%	1.0%
s_40_450	0.7%	1.4%	2.0%	0.5%	0.8%	1.3%
s_50_500	0.8%	1.3%	2.8%	0.3%	1.0%	1.8%

Table 15.1.: Minimum, average and maximum optimality gaps of MILP (15.11) after 12 h for all five test instances with $\lambda^k \in \{1, \dots, 10\}$ and the two scenario sets.

same ratio. Approximately 10 % of the total power consumption of a BS is consumed during idle periods [54]. Since we intend to switch off or set to idle mode, respectively, BSs during one third of a day, we set $c_s^k = (40 - 40 \cdot 10\,\%) \cdot 1/3 = 12$. This means, we assume 4 (representing 400 W) is consumed by the standby of a BS and the remaining 36 (representing 3600 W) are equally consumed during eight hours per day. Moreover, reasonable values for the scaling parameter λ^k are $1, \dots, \lambda$ for all $k \in K$. Based on the chosen values of the parameters, we can require $\omega \in \mathbb{Z}$ to speed up the solving of the MILPs (15.11) and (15.12).

All subsequent computations are performed on a Linux machine with 3.40GHz Intel Core i7-3770 processor and a general CPU time limit of 12 h. Moreover, we use the standard version of CPLEX 12.4 [98] to solve the MILPs.

We cannot solve the recoverable RWNPP for all instances to optimality within the TL of 12 h. In Table 15.1, we state the minimum, average and maximum optimality gaps for all test instances and both scenario sets. The average is taken over $\lambda^k \in \{1, \dots, 10\}$. All settings for the smallest test instance s_20_200 and the scenario set "normal" are solved to optimality while the higher demands during the night in the scenario set "mix" have a negative effect on the solving of MILP (15.11). However, the maximum optimality gaps for all test instances are at most 2.8 % and at most 1.4 % on average. Therefore, also non-optimal solutions are reasonable for the subsequent investigations.

For the value of λ^k that yields the lowest objective value per test instance and scenario set, we present the number of deployed BSs and covered TNs during the day and the number of operating BSs and the minimum, average, and maximum number of deployed TNs during the night in Table 15.2, where the average is taken over all scenarios (nights). Apart from test instance s_50_500 which does not serve one particular TN due to the non-optimality of the solution, all TNs are served during the day. For all test instances and both scenario sets, the recoverable robust solution economises at least one BS during the night while at most 2 % of the TNs but in most cases 0 % are lost. In six out of seven (all) nights, two BSs are economised for test instance s_50_500 (s_40_400) and three BSs are economised in all nights for s_40_450 in case of the scenario set "normal". In case of the scenario set "mix", two BSs are economised for test instance s_40_450.

Concerning the value of λ^k that gives the results in Table 15.2, lower objective values

	instance	λ^k	# BSs d.	# TNs d.	# BSs n.	min.	av.	max.
"normal"	s_20_200	1	4	200	3	200	200	200
	s_30_300	2	5	300	4	300	300	300
	s_40_400	4	6	400	4	400	400	400
	s_40_450	1	7	450	5	441	441	441
	s_50_500	5	7	499	5/6	498	498	500
"mix"	s_20_200	1	4	200	3	197	197	197
	s_30_300	2	5	300	4	300	300	300
	s_40_400	10	6	400	5	400	400	400
	s_40_450	2	7	450	4	450	450	450
	s_50_500	3	7	500	6	500	500	500

Table 15.2.: Number of deployed BSs and covered TNs during the day and the number of BSs and the minimum, average (over scenarios), and maximum number of covered TNs during the night for λ^k with the lowest objective value per test instance and scenario set.

instance	"normal"	"mix"
s_20_200	7.5%	5.6%
s_30_300	6.0%	6.0%
s_40_400	10.0%	5.0%
s_40_450	9.6%	8.6%
s_50_500	0.7%	4.3%

Table 15.3.: Gain of recovery in % of the recoverable RWNPP (15.11) compared to the robust problem (15.12) per test instance and scenario set.

are naturally obtained for lower values of λ^k. However, if too many TNs are not served during the night as the penalty to leave TNs uncovered is too low, the objective value will rise. Hence, for larger test instances, the value of λ^k yielding the lowest objective value is in general higher than for test instances with a lower number of BSs and TNs. Even though, this behaviour is not surprising, we cannot predict the best value for the scaling parameter λ^k.

Finally, we evaluate the GoR as defined in (14.3) which is achieved by the recoverable RWNPP (15.11) in comparison to the nominal problem (15.12). The GoR for all five test instances and both scenario sets is stated in Table 15.3. Formulation (15.12) without recovery, which is solved optimally for all instances, serves all TNs and deploys the same number of BSs as the recoverable robust solution during the day. As we have seen before, we economise at least one BS during the night by the recovery action. This economisation yields a GoR of at least 0.7 % and up to 10.0 %. Additionally, we would like to point out that these numbers are only lower bounds for the GoR since the recoverable robust solutions are not (all) optimal and hence, there exists the potential for further improvements.

15.3. Conclusion and Outlook

In this chapter, we have proposed a compact formulation with a discrete uncertainty set for the recoverable RWNPP to model the deactivation of BSs during low traffic times such as night. A computational study performed on five test instances with 20 to 50 BSs and 200 to 500 TNs and for two scenario sets has revealed the potential of the recoverable robustness approach. For a good choice of the scaling parameter λ^k, which represents the penalty for not serving TNs during one night, we can economise at least one and up to three BSs for all test instances. However, a good value of λ^k is hard to predict and hence, has been found by solving the recoverable RWNPP for different values and comparing the objective values. The achieved savings measured by the gain of recovery range from 0.7 % to 10.0 %. Certainly, this gain depends on the selected values for the parameters c_s, c_s^k, λ and λ^k.

As the compact formulation for the larger test instances is not solved during twelve hours, most investigated recoverable robust solutions are not optimal. Thus, the improvement of the solving of the recoverable RWNPP, e. g., by means of valid inequalities such as recoverable robust cover inequalities, remains as future work.

Furthermore, the adaption of the recovery action to model diverse aspects of the WNPP and the incorporation of further uncertainty sets such as interval or Γ-robust scenario sets remain as aspects of future research.

Remarks and Conclusions

Interpreting Computational Studies

We have performed computational studies to evaluate the effectiveness of valid inequalities, for instance, in Section 7.4 for a chance-constrained model and in Section 9.1.3 for a Γ-robust optimisation model. Based on recent studies in the field of unexpected behaviour in the performance of MIP solvers, we critically discuss the interpretation of such types of computational studies in this section.

Performance variability In 2008, Danna [60] was the first to point out large differences in solving times and number of B&B nodes for solving the same instance with the identical CPLEX version but on different machines (Linux, AIX) in a presentation. The author named such unexpected changes in performance in the MIP context *performance variability* (PV). Apart from diverse machines, Danna [60] generated PV also by permuting rows and/or columns of the original model and by changing the random seed leading to the common opinion that PV is intrinsic to MIP; see also [73, 126].

A further way to cause PV was discussed in a presentation by Fischetti and Monaci [71] in 2012. They added a redundant constraint which entailed a non-negligible speed-up in eight of nine cases. The technical explanation is on the one hand, that the addition of constraints changes the order of the variables in which they are loaded by the solver causing an implicit permutation of columns. On the other hand, the authors have selected heavily biased test beds for which the default solver does not perform well. As pointed out by Lodi and Tramontani [126], the adding of a redundant constraint just adds noise and hence, it is of great importance to distinguish between the impact of a scientific idea and the noise it adds to the solving process.

The first work discussing a source of PV was presented by Koch et al. [113] in 2011 when introducing the latest release of the benchmark library for MIPs, MIPLIB 2010. The authors identified imperfect tie-breaking as a reason for PV. Decisions such as the selection of branching candidates, which are taken by MIP solvers, are based on scores defined for certain selection criteria. If there is a tie among the best scored candidates and the tie-breaking is imperfect, selection is made arbitrarily. Then the ordering of the candidates and rounding errors gain importance. Methods to solve MIPs such as B&C are prone to PV since the entire subsequent solution process is affected once the path in a B&B tree diverges. Also Lodi and Tramontani [126] list floating point operations, which differ in different computing environments, as a possible source for PV. In this context, the authors further name the heuristic nature of MIP solver decisions, which can be caused if the perfect score is not known or too hard to compute, as another source of PV. Moreover, a

computational study in [113] reveals that also the number of threads influences the performance. A variability score introduced in this work to measure PV goes up and down quite arbitrary for an increasing number of threads. Additionally, numerical investigations of model permutations unveil that "the probability of improving performance by permuting the instances is about as high as the probability to deteriorate performance, but the average improvement is smaller than the average deterioration" [113].

Fischetti et al. [76] investigate and exploit the variability caused by the simplex algorithm which solves LPs. The LP relaxation of a MIP is solved at the root node of a B&B tree and forms the foundation for the subsequent computations. However, due to dual degeneracy of formulations, there usually exists more than one optimal basis which is returned by the simplex algorithm. So far, the returned basis is selected arbitrarily among the optimal while alternative bases may have significant and rather unpredictable impact on the solving of the MIP. Thus, the selection of the first basis is crucial.

Exploiting PV Apart from the problems caused by PV and specified above, it is also possible to exploit the variability to improve the performance. Two recent approaches are the following. Fischetti et al. [76] have developed a sampling scheme which is based on different optimal bases generated by the simplex algorithm. Thus, for a number of optimal bases, the default cutting plane loop and the default primal heuristics of the MIP solver are executed (in parallel). In a second step, the generated cutting planes and feasible solutions of the different samples are used as input for the final run. By a large number of samples, the sensitivity of the solving process to initial conditions can be reduced. Furthermore, also the random seed parameter of CPLEX version 12.5.0 and higher can be applied to generate different root node solutions to be exploited in a sample algorithm.

Another algorithm that makes use of PV is a so-called bet-and-run and has been recently proposed by Fischetti and Monaci [73]; see also [126] for a summary. The main idea of this method is to make a number of short runs for randomised initial conditions and then bet on the most promising run. Only this run is then completed. Bet-and-run adds just a small overhead to any sequential tree search method and numerical results performed on a proof-of-concept implementation have demonstrated its potential to improve the performance for medium to hard instances. The quality of the bet-and-run depends on the following two aspects. First, a cost- or resource-efficient way to generate diversified runs without deteriorating the average performance has to be determined. Second, rules how to choose the best run have to be defined. Fischetti and Monaci [73] specify a number of indicators such as number of open nodes or best bounds with a predefined priority.

Revisiting our computational results For all theoretical achieved results in this thesis, we have performed a computational study to investigate their performance. Especially the studies investigating the gain of valid inequalities in Sections 7.4 and 9.1.3 can be affected by PV.

In Chapter 7, we have applied chance constraints to the FBWN problem and have developed valid inequalities. In the computational study performed on a grid network in Section 7.4.1, the cuts perform better than the shifted cuts for some settings but only if the

internal CPLEX cuts are enabled. PV is a plausible explanation for this unexpected behaviour since the adding of internal cuts affects the complete B&B tree.

Similarly, the computational study in Section 9.1.3, which investigates the performance of three cutting plane approaches and their combinations for the d-RWNPP, has to be treated with caution. The results have indicated that the vub constraints, which are implicitly included in the capacity constraints, are one of the most effective types of cutting planes among the analysed. However, as discussed in [71, 126], the addition of (redundant) constraints adds noise and hence, impacts the MIP solver. Thus, we cannot be totally sure that the cutting planes improve the performance themselves and not just the added noise by chance. To limit this effect, we have followed the measure of selecting an unbiased test bed as recommended in [113, 126] by performing the study in Section 9.1.3 on several test scenarios of various dimensions. A large number of instances is advantageous, especially if the expected performance difference is small.

Based on these remarks, the practical improvements achieved by the addition of valid inequalities should not be overestimated. There can be no assurance that the results are similar for other solvers or even for future versions of the used solver CPLEX. Nevertheless, it is still common practice to perform such computational studies and all presented computational studies of this thesis should rather be seen as a proof of concept implementation.

Moreover, some classes of valid inequalities will become less important for the practical performance of MIPs since solvers become more and more advanced, inter alia, by providing a growing variety of valid inequality classes. As a consequence, users can deteriorate instead of improve the solving process by intervening. However, new problems with a so far unknown structure require new types of valid inequalities which have to be developed before they can be incorporated in the solvers. Therefore, valid, especially facet-defining, inequalities are of interest from a theoretical point of view as well as from a practical perspective if the structure of the corresponding problem is not well-studied.

Conclusions

In this thesis, we have studied diverse robust optimisation approaches as well as one variant of stochastic optimisation to model data uncertainty in the framework of different types of wireless communication networks. The investigated mathematical concepts comprise chance constraints representing a stochastic methodology, the nowadays prevalent Γ-robustness, its recent generalisation multi-band robustness, and the two-staged recoverable robustness.

For a complete description of cellular or mobile communication networks, we have additionally studied various approaches to model interference in a digression. To this end, we have on the one hand adapted three known approaches and on the other hand developed four novel. Three approaches model interference exactly, while the others are approximate. In theory, the exact methods model SINR requirements best. For practical purposes, we have studied approximate versions of the exact formulations, which are in contrast to the exact counterpart computationally tractable, and have first determined good parameter settings by means of a computational study. Afterwards, we have compared

the formulations via a SINR-corrected objective value, which incorporates the number of deployed BSs, covered TNs, and violated SINR constraints. Overall, the TN coverage requirement formulation has yielded the best value but the conflict graph formulation is most suitable for the subsequent investigations in this thesis due to its good results and low complexity. One major disadvantage of all approximate formulations is the violation of capacity constraints which reveals future research prospects to obtain solutions feasible for SINR and also for the capacity constraints.

To model uncertain radio configurations and to obtain reliable fixed broadband wireless networks, we have formalised capacity requirements under uncertainty by separated as well as joint chance constraints. For the latter approach, we have proposed ILP formulations, cutset inequalities, and a primal heuristic applicable to a budget-constrained model. Computational studies performed on generated grid network instances as well as realistic network topologies have revealed a positive effect of the valid inequalities and of the primal heuristic in terms of solving time or optimality gap. A comparison of the reliability obtained for different budget values and two alternative formulations without outage probability constraints has demonstrated a significant gain in reliability by our approach.

As one of the most well-established robust optimisation methodologies, we have studied the Γ-robustness intensively according to various aspects for the WNPP. On the one hand, we have incorporated uncertain demands via two alternative modelling approaches, a compact formulation and a B&P framework, and on the other hand, we have integrated fluctuating channel conditions by means of uncertain spectral efficiencies.
For the compact ILP with uncertain demands, we have presented three types of valid inequalities and a heuristic separation algorithm where a computational study has revealed their potential to improve the complete solving process due to the enhancement of the dual bound. Additionally, we have investigated the price of robustness and the level of protection for this formulation. The percentage deterioration of the objective value is acceptable, at most 50 % for all scenarios for which we have been able to compute a reasonable primal bound, while Γ-robust solutions with a protection level of 100 % economise one BS compared to conventional solutions for all but one test scenario. The latter result particularly demonstrates the potential of Γ-robust optimisation applied to wireless network planning problems.
Since we could not solve the compact ILP with uncertain demands to optimality for every studied test scenario and the obtained LP solutions are quite poor, we have developed a B&P approach as an alternative formulation, which has proven to be efficient for various wireless network problems. To accelerate the solving process, we have additionally proposed various settings where the suboptimal solving of the pricing problems has performed best at the root node and the primal heuristic setting has performed best regarding the complete solving process. Moreover, we have compared the performance of the B&P approach directly to the ILP formulation in a further computational study, which has demonstrated the expected improvement of the LP solution. However, even the best setting of the B&P algorithm cannot compete with the (improved) compact formulation when considering the complete solving process.

The incorporation of uncertain spectral efficiencies in the WNPP by means of Γ-robustness models interference approximately. The crucial aspect of the compact reformulation

is the definition of the deviation intervals. Based on a computational study, we have determined the best setting among the various presented possibilities which is the shifting of the SNR value (once for the nominal value, twice for the lowest value). We have then studied the resulting formulation in terms of validity for interference modelling. It has turned out that we cannot predict the number of violated SINR conditions or the number of actually served TNs when varying the value of Γ and additionally, the TN coverage requirement has given better results. Hence, we do not judge the Γ-robustness suitable to model interference in the WNPP.

A generalisation of Γ-robust optimisation is multi-band robustness where the deviation interval can be partitioned into multiple bands and the total number of realisations in each band is bounded by one parameter per band. We have incorporated this robust optimisation methodology in the WNPP problem and have proposed a compact formulation in case of two bands. We have illustrated a linear correlation between the two robustness parameters by means of a computational study performed on a small scenario. Furthermore, we have investigated the gain achieved by multi-band robustness in comparison to the Γ-robust WNPP. The numerical results have revealed a potential to save cost via the economisation of one BS, where this potential strongly depends on the test scenario and the Γ-robust solution.

Besides this practical application, we have studied multi-band robustness also theoretically for the RKP. To this end, we have developed two alternative DPs for the K-band RKP. The first DP has a complexity linear in the number of items n; $O(nB^{K+1})$. As the capacity B is usually higher than the number of items, we have devised the second DP with a complexity linear in the capacity; $O(K!n^{K+1}B)$. Additionally, we have concluded from this DP that a binary combinatorial optimisation problem with uncertain objective can be solved by solving $O(K!n^K)$ similar problems with certain objective, which generalises a known result for Γ-robust optimisation problems. Moreover, we have improved the DP with complexity linear in the capacity in practice and have depicted the effectiveness of the improvements in a computational study performed for two bands on randomly generated instances of various sizes. The improved DP also outperforms the compact ILP of the multi-band RKP in terms of solving time. This last achievement has accounted for the application of the improved DP to the TNAP, a subproblem of the WNPP. For this purpose, we have proposed a Lagrangian relaxation for this problem, which consists of several uncorrelated two-band RKPs and gives a better bound than the LP relaxation. The two-band RKPs are solved separately by the improved DP. Numerical results have shown that the Lagrangian relaxation gives a (close-to) optimal upper bound in a reasonable time for all investigated scenarios.

Finally, we have investigated the two-stage concept of recoverable robustness in terms of the WNPP. To this end, we have proposed a compact formulation incorporating a discrete uncertainty set. We have studied the recoverable robust solutions for test instances of various dimensions and different values for the scaling parameter which penalises the non-coverage of a TN during one night. The solutions with the best objective values economise at least one BS during the night while at most 2 % of the TNs are lost for all investigated test instances. Comparing these results to a robust solution without recovery, the gain of recovery ranges between 0.7 and 10.0 %. Thus, the recoverable robust optimisation applied

to the WNPP can give a good benchmark of savings that can be achieved by deactivating BSs during low traffic times.

All in all, we have studied various stochastic and robust optimisation methodologies in the context of wireless communication networks and have presented their potentials. The chance-constrained approach has proven to be appropriate to obtain a reliable fixed broadband wireless network. For cellular wireless communication networks, the question remains which robustness approach should be applied. The apparent answer is, it depends. While Γ-robustness is well suited if historical data is known and can reduce cost in comparison to conventional planning, it estimates deviations just roughly. Multi-band robustness can overcome this limitation but also increases the complexity of the formulation polynomially. If a reformulation of the planning problem containing only separate multi-band RKPs is possible, there exists an efficient solving method. Furthermore, the recoverable robustness offers more flexibility and a diversified adaptability especially if the modelling of different time segments is required. Based on the potentials and limitations revealed in our work, the concept to be applied should be chosen according to the problem and the desired optimisation purposes.

Future research topics This thesis provides several aspects for future research. First of all, the proposed formulations for interference modelling should be improved for practical applicability. If the solving of (one of) the exact models is sufficiently fast, it is possible to combine a robustness concept with interference modelling obtaining a more accurate formalisation of wireless communication networks.

Furthermore, the B&P approach, which could not compete with the improved compact formulation for the Γ-robust WNPP, might be beneficial if adopted to a multi-band or recoverable robust model. In particular, for the multi-band RWNPP the solving of the pricing problems can benefit from the improved DP for the multi-band RKP.

In the context of multi-band robust wireless communication networks, the development of a branching routine for the Lagrangian relaxation applied to the TNAP remains as future work. Additionally, the extension of the Lagrangian relaxation for the TNAP to the complete WNPP is an aspect for future investigations.

For the recoverable robust WNPP, the solving of the compact formulation with discrete uncertainty sets has to be improved to obtain optimal solutions. Moreover, further uncertainty sets such as Γ-robust and different recovery actions, i. e., conditions subject to which TNs are allowed to be reassigned in the recovery step, should be investigated and developed to model diverse aspects of the WNPP.

Finally, from a practical point of view, future wireless communication networks such as those of the currently developed fifth generation will apply novel techniques and thus, will pose new challenges for the existing optimisation models.

Bibliography

[1] 3rd Generation Partnership Project (3GPP), 2014. URL http://www.3gpp.org. accessed 30.08.2014. (Cited on pages 27 and 32.)

[2] 3Roam. 3Roam website, 2014. URL http://www.3roam.com. accessed 30.08.2014. (Cited on page 101.)

[3] T. Achterberg. SCIP: Solving constraint integer programs. *Mathematical Programming Computation*, 1(1):1–41, July 2009. URL http://scip.zib.de. (Cited on pages 156 and 161.)

[4] P. Adasme and A. Lisser. Semidefinite and Conic Programming for Robust wireless {OFDMA} networks. *Electronic Notes in Discrete Mathematics*, 36(0):1225–1232, 2010. (Cited on page 127.)

[5] J.H. Ahrens and G. Finke. Merging and Sorting Applied to the Zero-One Knapsack Problem. *Operations Research*, 23(6):1099, 1975. (Cited on page 16.)

[6] A. Altin, E. Amaldi, P. Belotti, and M.Ç. Pinar. Provisioning virtual private networks under traffic uncertainty. *Networks*, 49(1):100–115, 2007. (Cited on page 127.)

[7] E. Alvarez-Miranda, I. Ljubic, S. Raghavan, and P. Toth. The Recoverable Robust Two-Level Network Design Problem. *INFORMS Journal on Computing*, 2014. URL http://homepage.univie.ac.at/ivana.ljubic/research/publications/RRTLND.pdf. online first. (Cited on page 221.)

[8] H. Anderson. *Fixed Broadband Wireless System Design*. John Wiley & Sons, 2003. (Cited on page 36.)

[9] M. Andrews and M. Dinitz. Maximizing Capacity in Arbitrary Wireless Networks in the SINR Model: Complexity and Game Theory. In *IEEE INFOCOM 2009*, pages 1332–1340, April 2009. (Cited on page 42.)

[10] A. Atamtürk and V. Narayanan. The submodular knapsack polytope. *Discrete Optimization*, 6(4):333–344, 2009. (Cited on page 80.)

[11] E. Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8(1):146–164, 1975. (Cited on pages 16, 17, and 18.)

[12] E. Balas and R. Jeroslow. Canonical Cuts on the Unit Hypercube. *SIAM Journal on Applied Mathematics*, 23(1):pp. 61–69, 1972. (Cited on page 18.)

[13] W.T. Barnett. Multipath propagation at 4, 6 and 11 GHz. *Bell System Technical Journal*, 51(2):311–361, 1972. (Cited on page 102.)

[14] C. Barnhart, E.L. Johnson, G.L. Nemhauser, M.W.P. Savelsbergh, and P.H. Vance. Branch-and-Price: Column Generation for Solving Huge Integer Programs. *Operations Research*, 46(3):316–329, 1998. (Cited on pages 11 and 154.)

[15] A. Ben-Tal and A. Nemirovski. Robust solutions of Linear Programming problems contaminated with uncertain data. *Mathematical Programming*, 88(3):411–424, 2000. (Cited on page 21.)

[16] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009. (Cited on pages 21, 24, 25, and 79.)

[17] P. Beraldi and A. Ruszczyński. A branch and bound method for stochastic integer problems under probabilistic constraints. *Optimization Methods and Software*, 17 (3):359–382, 2002. (Cited on page 81.)

[18] D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98(1-3):49–71, 2003. (Cited on pages 115, 119, 121, 122, and 181.)

[19] D. Bertsimas and M. Sim. The Price of Robustness. *Operations Research*, 52(1): 35–53, 2004. (Cited on pages 115, 119, 120, 121, 122, and 179.)

[20] D. Bertsimas, D. Brown, and C. Caramanis. Theory and Applications of Robust Optimization. *SIAM Review*, 53(3):464–501, 2011. (Cited on pages 24 and 25.)

[21] D. Bertsimas, E. Nasrabadi, and S. Stiller. Robust and Adaptive Network Flows. *Operations Research*, 61(5):1218–1242, 2013. (Cited on page 127.)

[22] D. Bienstock. Histogram models for robust portfolio optimization. *Journal on Computational Finance*, 11:1–64, 2007. (Cited on page 173.)

[23] D. Bienstock and F. D'Andreagiovanni. Robust Wireless Network Planning. In *Proceedings of AIRO 2009, the 40th Annual Conference of the Italian Operational Research Society*, pages 131–132, 2009. (Cited on page 173.)

[24] D. Bienstock and O. Günlük. Capacitated Network Design – Polyhedral Structure and Computation. *INFORMS Journal on Computing*, 8(3):243–259, 1996. (Cited on page 93.)

[25] D. Bienstock, S. Chopra, O. Günlük, and C.Y. Tsai. Minimum Cost Capacity Installation for Multicommodity Network Flows. *Mathematical Programming*, 81(2): 177–199, 1998. (Cited on page 93.)

[26] J.R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer, 2011. (Cited on page 22.)

[27] P. Bjorklund, P. Varbrand, and D. Yuan. Resource optimization of spatial TDMA in ad hoc radio networks: a column generation approach. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (IEEE INFOCOM 2003)*, volume 2, pages 818–824, March 2003. (Cited on page 149.)

[28] O. Blume, H. Eckhardt, S. Klein, E. Kuehn, and W. M. Wajda. Energy savings in mobile networks based on adaptation to traffic statistics. *Bell Labs Technical Journal*, 15(2):77–94, Sept 2010. (Cited on page 225.)

[29] S.A. Borbash and A. Ephremides. Wireless Link Scheduling With Power Control and SINR Constraints. *IEEE Transactions on Information Theory*, 52(11):5106–5111, November 2006. (Cited on page 42.)

[30] P.C. Bouman, J.M. van den Akker, and J.A. Hoogeveen. Recoverable Robustness by Column Generation. In C. Demetrescu and M.M. Halldórsson, editors, *Algorithms – ESA 2011*, volume 6942 of *Lecture Notes in Computer Science*, pages 215–226. Springer Berlin Heidelberg, 2011. (Cited on page 221.)

[31] C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM*, 16(9):575–577, September 1973. (Cited on pages 45 and 129.)

[32] O. Bulakci, S. Redana, B. Raaf, and J. Hämäläinen. System Optimization in Relay Enhanced LTE-Advanced Networks via Uplink Power Control. In *IEEE 71st Vehicular Technology Conference (VTC 2010-Spring)*, pages 1–5, 2010. (Cited on page 45.)

[33] C. Büsing. *Recoverable Robustness in Combinatorial Optimization*. PhD thesis, Technische Universität Berlin, Cuvillier Verlag Göttingen, 2011. (Cited on page 221.)

[34] C. Büsing and F. D'Andreagiovanni. New Results about Multi-band Uncertainty in Robust Optimization. In R. Klasing, editor, *Experimental Algorithms*, volume 7276 of *Lecture Notes in Computer Science*, pages 63–74. Springer Berlin Heidelberg, 2012. (Cited on pages 120, 173, 178, and 180.)

[35] C. Büsing and F. D'Andreagiovanni. Robust Optimization under Multi-band Uncertainty — Part I: Theory. accessed 30.08.2014, January 2013. URL `http://arxiv.org/abs/1301.2734`. (Cited on pages 173, 175, 176, 177, and 179.)

[36] C. Büsing, A.M.C.A. Koster, and M. Kutschka. Recoverable Robust Knapsacks: Γ-Scenarios. In J. Pahl, T. Reiners, and S. Voß, editors, *Network Optimization*, volume 6701 of *Lecture Notes in Computer Science*, pages 583–588. Springer Berlin Heidelberg, 2011. (Cited on pages 221 and 223.)

[37] C. Büsing, A.M.C.A. Koster, and M. Kutschka. Recoverable Robust Knapsacks: the Discrete Scenario Case. *Optimization Letters*, 5(3):379–392, 2011. (Cited on page 221.)

[38] V. Cacchiani, A. Caprara, L. Galli, L. Kroon, and G. Maróti. Recoverable Robustness for Railway Rolling Stock Planning. In M. Fischetti and P. Widmayer, editors, *8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08)*, volume 9 of *OpenAccess Series in Informatics (OASIcs)*, Dagstuhl, Germany, 2008. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. (Cited on page 221.)

[39] A. Capone, L. Chen, S. Gualandi, and D. Yuan. A New Computational Approach for Maximum Link Activation in Wireless Networks under the SINR Model. *Wireless Communications, IEEE Transactions on*, 10(5):1368–1372, 2011. (Cited on pages 44 and 45.)

[40] Central Intelligence Agency (CIA). The World Factbook, 2013. URL `https://www.cia.gov/library/publications/the-world-factbook`. accessed 30.08.2014. (Cited on page 1.)

[41] D. Chafekar, V.S.A. Kumart, M.V. Marathe, S. Parthasarathy, and A. Srinivasan. Approximation Algorithms for Computing Capacity of Wireless Networks with SINR Constraints. In *IEEE INFOCOM 2008. The 27th Conference on Computer Communications.*, pages –, April 2008. (Cited on page 41.)

[42] A. Charnes and W.W. Cooper. Chance-Constrained Programming. *Management Science*, 6(1):73–79, 1959. (Cited on page 79.)

[43] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. *Management Science*, 4:235–263, 1958. (Cited on page 79.)

[44] A. Chatzipapas, S. Alouf, and V. Mancuso. On the minimization of power consumption in base stations using on/off power amplifiers. In *IEEE Online Conference on Green Communications (GreenCom 2011)*, pages 18–23, September 2011. (Cited on page 225.)

[45] S. Chopra, I. Gilboa, and S.T. Sastry. Source sink flows with capacity installation in batches. *Discrete Applied Mathematics*, 85(3):165–192, 1998. (Cited on page 37.)

[46] G. Claßen, D. Coudert, A.M.C.A. Koster, and N. Nepomuceno. A Chance-Constrained Model & Cutting Planes for Fixed Broadband Wireless Networks. In J. Pahl, T. Reiners, and S. Voß, editors, *Network Optimization*, volume 6701 of *Lecture Notes in Computer Science*, pages 37–42. Springer Berlin Heidelberg, 2011. (Cited on pages 2, 83, and 100.)

[47] G. Claßen, D. Coudert, A.M.C.A. Koster, and N. Nepomuceno. Bandwidth assignment for reliable fixed broadband wireless networks. In *12th IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM 2011)*, pages 1–6, 2011. (Cited on pages 2, 32, 83, and 99.)

[48] G. Claßen, A.M.C.A. Koster, and A. Schmeink. Robust Planning of Green Wireless Networks. In *5th International Conference on Network Games, Control and Optimization (NetGCooP)*, pages 1–5, October 2011. (Cited on pages 2 and 128.)

[49] G. Claßen, A.M.C.A. Koster, and A. Schmeink. A Robust Optimisation Model and Cutting Planes for the Planning of Energy-Efficient Wireless Networks. *Computers and Operations Research*, 40(1):80–90, January 2013. (Cited on pages 2, 33, 45, 62, 128, 133, 134, and 136.)

[50] G. Claßen, A.M.C.A. Koster, and A. Schmeink. Speeding up column generation for robust wireless network planning. *EURO Journal on Computational Optimization*, 1(3-4):253–281, 2013. (Cited on pages 2, 149, 161, 164, and 170.)

[51] G. Claßen, D. Coudert, A.M.C.A. Koster, and N. Nepomuceno. Chance-Constrained Optimization of Reliable Fixed Broadband Wireless Networks. *INFORMS Journal on Computing*, 26(4):893–909, 2014. (Cited on pages 2, 83, 93, 102, and 108.)

[52] G. Claßen, A.M.C.A. Koster, and A. Schmeink. The Multi-Band Robust Knapsack Problem – A Dynamic Programming Approach –. *Discrete Optimization*, 2014. URL `http://www.optimization-online.org/DB_FILE/2014/02/4246.pdf`. under review. (Cited on pages 2 and 179.)

[53] G. Claßen, A.M.C.A. Koster, M. Kutschka, and I. Tahiri. Robust Metric Inequalities for the Γ-Robust Network Loading Problem. *Asia Pacific Journal of Operational Research*, 2015. accepted. (Cited on pages 2 and 94.)

[54] L.M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. Van der Perre. Challenges and enabling technologies for energy aware mobile radio networks. *Communications Magazine, IEEE*, 48(11):66–72, November 2010. (Cited on pages 225 and 229.)

[55] COST 231. Urban micro cell measurements and building data, 1996. URL `http://www2.ihe.uni-karlsruhe.de/forschung/cost231/cost231.en.html`. accessed 30.08.2014. (Cited on page 61.)

[56] A.M. Costa, J-F. Cordeau, and B. Gendron. Benders, metric and cutset inequalities for multicommodity capacitated network design. *Computational Optimization and Applications*, 42(3):371–392, 2009. (Cited on page 94.)

[57] D. Coudert, N. Nepomuceno, and H. Rivano. Power-efficient radio configuration in fixed broadband wireless networks. *Computer Communications*, 33(8):898–906, 2010. (Cited on pages 31 and 36.)

[58] F. 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 page 173.)

[59] F. D'Andreagiovanni. Multiband Robust Optimization, 2014. URL http://www.dis.uniroma1.it/~fdag/multiband.html. (Cited on page 173.)

[60] E. Danna. Performance variability in mixed integer prgoramming. Presentation, Workshop on Mixed Integer Programming (MIP 2008). accessed 30.08.2014, 2008. URL http://coral.ie.lehigh.edu/~jeff/mip-2008/talks/danna.pdf. (Cited on page 233.)

[61] G.B. Dantzig. Linear Programming under Uncertainty. *Management Science*, 1 (3-4):197–206, 1955. (Cited on page 21.)

[62] G.B. Dantzig and P. Wolfe. Decomposition Principle for Linear Programs. *Operations Research*, 8(1):101–111, 1960. (Cited on page 11.)

[63] M. Deruyck, W. Vereecken, E. Tanghe, W. Joseph, M. Pickavet, L. Martens, and P. Demeester. Comparison of power consumption of mobile WiMAX, HSPA and LTE access networks. In *9th Conference on Telecommunications Internet and Media Techno Economics (CTTE)*, pages 1–7, 2010. (Cited on page 62.)

[64] G. Desaulniers, J. Desrosiers, and M.M. Solomon, editors. *Column Generation*. Springer, 2005. Chapters 1, 12. (Cited on pages 11, 12, and 157.)

[65] A. Engels. *Dimensioning, Cell Site Planning, and Self-Organization of 4G Radio Networks*. PhD thesis, RWTH Aachen University, Shaker Verlag, 2013. (Cited on pages 61 and 62.)

[66] A. Engels, M. Neunerdt, R. Mathar, and H.M. Abdullah. Acceptance as a Success Factor for Planning Wireless Network Infrastructure. In *International Symposium on Wireless Communication Systems 2011 (ISWCS'11)*, pages 889–893, Aachen, Germany, November 2011. (Cited on page 45.)

[67] A. Engels, M. Reyer, X. Xu, R. Mathar, J. Zhang, and H. Zhuang. Autonomous Self-Optimization of Coverage and Capacity in LTE Cellular Networks. *IEEE Transactions on Vehicular Technology*, 62(5):1989–2004, June 2013. (Cited on pages 46, 62, and 65.)

[68] European Telecommunications Standards Institute (ETSI). Harmonized Standard EN 302 217-3, 2014. URL http://www.etsi.org. accessed 30.08.2014. (Cited on page 101.)

[69] B. Faaland. Solution of the Value-Independent Knapsack Problem by Partitioning. *Operations Research*, 21(1):332–337, 1973. (Cited on page 16.)

[70] M. Fischetti and M. Monaci. Light Robustness. In R.K. Ahuja, R.H. Möhring, and C.D. Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 61–84. Springer Berlin Heidelberg, 2009. (Cited on page 25.)

[71] M. Fischetti and M. Monaci. On the role of randomness in exact tree search methods. Presentation, Matheuristics 2012. accessed 30.08.2014, 2012. URL `http://mat.tepper.cmu.edu/blog/wp-content/uploads/2012/09/2012-Matheuristic-Fischetti`. (Cited on pages 233 and 235.)

[72] M. Fischetti and M. Monaci. Cutting plane versus compact formulations for uncertain (integer) linear programs. *Mathematical Programming Computation*, 4(3): 239–273, 2012. (Cited on pages 117, 119, and 178.)

[73] M. Fischetti and M. Monaci. Exploiting Erraticism in Search. *Operations Research*, 62(1):114–122, 2014. (Cited on pages 233 and 234.)

[74] M. Fischetti, D. Salvagnin, and A. Zanette. Fast Approaches to Improve the Robustness of a Railway Timetable. *Transportation Science*, 43(3):321–335, 2009. (Cited on page 25.)

[75] M. Fischetti, A. Lodi, and D. Salvagnin. Just MIP it! In *Matheuristics*, volume 10 of *Annals of Information Systems*, pages 39–70. Springer, 2010. (Cited on page 95.)

[76] M. Fischetti, A. Lodi, M. Monaci, D. Salvagnin, and A. Tramontani. Tree search stabilization by random sampling. Technical report, University of Bologna, Italy, 2013. URL `http://www.dei.unipd.it/~salvagni/pdf/randomness.pdf`. accessed 30.08.2014. (Cited on page 234.)

[77] M.L. Fisher. The Lagrangian Relaxation Method for Solving Integer Programming Problems. *Management Science*, 50(12_supplement):1861–1871, 2004. (Cited on page 213.)

[78] B. Fortz and M. Poss. An improved Benders decomposition applied to a multi-layer network design problem. *Operations Research Letters*, 37(5):359–364, 2009. (Cited on page 92.)

[79] L. Fu, S. Liew, and J. Huang. Fast algorithms for joint power control and scheduling in wireless networks. *IEEE Transactions on Wireless Communications*, 9(3):1186–1197, March 2010. (Cited on page 149.)

[80] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, 1979. (Cited on page 7.)

[81] A.M. Geoffrion. Lagrangean relaxation for integer programming. In M.L. Balinski, editor, *Approaches to Integer Programming*, volume 2 of *Mathematical Programming Studies*, pages 82–114. Springer Berlin Heidelberg, 1974. (Cited on page 13.)

[82] M. Goerigk, M. Schmidt, A. Schöbel, M. Knoth, and M. Müller-Hannemann. The Price of Strict and Light Robustness in Timetable Information. *Transportation Science*, 48(2):225–242, 2014. online first. (Cited on page 25.)

[83] A.J. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005. (Cited on page 27.)

[84] A.J. Goldsmith and S-G. Chua. Adaptive coded modulation for fading channels. *IEEE Transactions on Communications*, 46(5):595–602, 1998. (Cited on page 37.)

[85] C. Gomes, S. Pérennes, and H. Rivano. Bottleneck Analysis for Routing and Call Scheduling in Multi-Hop Wireless Networks. In *IEEE GLOBECOM Workshops*, pages 1–6, November 2008. (Cited on page 149.)

[86] R.E. Gomory. Outline of an Algorithm for Integer Solutions to Linear Programs. *Bulletin of the American Society*, 64:275–278, 1958. (Cited on page 11.)

[87] O. Goussevskaia, Y.A. Oswald, and R. Wattenhofer. Complexity in Geometric SINR. In *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, MobiHoc '07, pages 100–109, New York, NY, USA, 2007. ACM. (Cited on page 42.)

[88] O. Goussevskaia, R. Wattenhofer, M.M. Halldorsson, and E. Welzl. Capacity of Arbitrary Wireless Networks. In *IEEE INFOCOM 2009*, pages 1872–1880, April 2009. (Cited on page 42.)

[89] J. Grönkvist and A. Hansson. Comparison Between Graph-based and Interference-based STDMA Scheduling. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing*, MobiHoc '01, pages 255–258, New York, NY, USA, 2001. ACM. (Cited on page 45.)

[90] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981. (Cited on page 11.)

[91] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000. (Cited on page 42.)

[92] R. Gupta, J. Walrand, and O. Goldschmidt. Maximal Cliques in Unit Disk Graphs: Polynomial Approximation. In *IN PROCEEDINGS INOC 2005*, 2005. (Cited on page 45.)

[93] P.L. Hammer, E.L. Johnson, and U.N. Peled. Facet of regular 0–1 polytopes. *Mathematical Programming*, 8(1):179–206, 1975. (Cited on pages 16 and 18.)

[94] F. Han, Z. Safar, W.S. Lin, Y. Chen, and K.J.R. Liu. Energy-efficient cellular network operation via base station cooperation. In *IEEE International Conference on Communications (ICC 2012)*, pages 4374–4378, June 2012. (Cited on page 225.)

[95] F.S. Hillier. Chance-constrained programming with 0-1 or bounded continuous decision variables. *Management Science*, 14(1):34–57, 1967. (Cited on page 80.)

[96] E. Horowitz and S. Sahni. Computing Partitions with Applications to the Knapsack Problem. *Journal of ACM*, 21(2):277–292, 1974. (Cited on page 16.)

[97] M. Hunukumbure, R. Agarwal, and S. Vadgama. Handover Mechanisms for Planned Cell Outage in Twin State Green Wireless Networks. In *IEEE 73rd Vehicular Technology Conference (VTC 2011-Spring)*, pages 1–5, May 2011. (Cited on page 225.)

[98] IBM – ILOG. CPLEX Optimization Studio 12.4, 2012. URL `http://www.ilog.com/products/cplex`. accessed 30.08.2014. (Cited on pages 35, 64, 84, 98, 99, 102, 143, 161, 199, 206, and 229.)

[99] IEEE Computer Society and the IEEE Microwave Theory and Techniques Society. IEEE Standard for Air Interface for Broadband Wireless Access Systems, 2012. URL `http://standards.ieee.org/getieee802/802.16.html`. IEEE Std 802.16-2012, accessed 30.08.2014. (Cited on page 32.)

[100] International Telecommunications Union (ITU), 2014. URL `http://www.itu.int`. accessed 30.08.2014. (Cited on pages 27 and 28.)

[101] M. Johansson and L. Xiao. Scheduling, routing and power allocation for fairness in wireless networks. In *IEEE 59th Vehicular Technology Conference (VTC 2004-Spring)*, volume 3, pages 1355–1360, May 2004. (Cited on page 42.)

[102] F. Kahn. *LTE for 4G Mobile Broadband*. Cambridge University Press, 2009. (Cited on pages 29, 31, 45, and 62.)

[103] R. Kannan and C.L. Monma. On the Computational Complexity of Integer Programming Problems. In Rudolf Henn, Bernhard Korte, and Werner Oettli, editors, *Optimization and Operations Research*, volume 157 of *Lecture Notes in Economics and Mathematical Systems*, pages 161–172. Springer Berlin Heidelberg, 1978. (Cited on page 9.)

[104] K. Kaparis and A.N. Letchford. Separation algorithms for 0-1 knapsack polytopes. *Mathematical Programming*, 124(1-2):69–91, 2010. (Cited on page 17.)

[105] R.M. Karp. Reducibility among Combinatorial Problems. In R.E. Miller, J.W. Thatcher, and J.D. Bohlinger, editors, *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Springer US, 1972. (Cited on page 15.)

[106] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer, 2004. (Cited on pages 15, 18, 19, and 181.)

[107] D Klabjan, G.L Nemhauser, and C Tovey. The complexity of cover inequality separation. *Operations Research Letters*, 23(1–2):35–40, 1998. (Cited on pages 17 and 125.)

[108] O. Klopfenstein. Solving chance-constrained combinatorial problems to optimality. *Computational Optimization and Applications*, 45(3):607–638, 2010. (Cited on pages 80 and 83.)

[109] O. Klopfenstein and D. Nace. A note on polyhedral aspects of a robust knapsack problem. Optimization Online, 2007. URL `http://www.optimization-online.org/DB_HTML/2006/04/1369.html`. accessed 30.08.2014. (Cited on page 200.)

[110] O. Klopfenstein and D. Nace. A robust approach to the chance-constrained knapsack problem. *Operations Research Letters*, 36(5):628–632, 2008. (Cited on pages 179 and 181.)

[111] O. Klopfenstein and D. Nace. Cover inequalities for robust knapsack sets—Application to the robust bandwidth packing problem. *Networks*, 59(1):59–72, 2012. (Cited on pages 122, 123, 124, 125, 131, and 157.)

[112] M. Kloppel, A. Gabash, A. Geletu, and Pu Li. Chance constrained optimal power flow with non-Gaussian distributed uncertain wind power generation. In *12th International Conference on Environment and Electrical Engineering (EEEIC)*, pages 265–270, 2013. (Cited on page 80.)

[113] T. Koch, T. Achterberg, E. Andersen, O. Bastert, T. Berthold, R.E. Bixby, E. Danna, G. Gamrath, A.M. Gleixner, S. Heinz, A. Lodi, H. Mittelmann, T. Ralphs, D. Salvagnin, D.E. Steffy, and K. Wolter. MIPLIB 2010. *Mathematical Programming Computation*, 3(2):103–163, 2011. (Cited on pages 233, 234, and 235.)

[114] P.J. Kolesar. A Branch and Bound Algorithm for the Knapsack Problem. *Management Science*, 13(9):723–735, 1967. (Cited on page 15.)

[115] C. Kosta, B. Hunt, A.U. Quddus, and R. Tafazolli. On Interference Avoidance Through Inter-Cell Interference Coordination (ICIC) Based on OFDMA Mobile Systems. *IEEE Communications Surveys Tutorials*, 15(3):973–995, 2013. (Cited on pages 39 and 51.)

[116] A.M.C.A. Koster, M. Kutschka, and C. Raack. Towards Robust Network Design using Integer Linear Programming Techniques. In *6th EURO-NF Conference on Next Generation Internet (NGI)*, pages 1–8, June 2010. (Cited on pages 119 and 120.)

[117] A.M.C.A. Koster, M. Kutschka, and C. Raack. Robust network design: Formulations, valid inequalities, and computations. *Networks*, 61(2):128–149, 2013. (Cited on pages 95, 117, 119, and 128.)

[118] M. Kutschka. *Robustness Concepts for Knapsack and Network Design Problems under Data Uncertainty*. PhD thesis, RWTH Aachen University, Cuvillier Verlag Göttingen, 2013. (Cited on pages 122, 124, 125, 127, 173, 174, 176, 177, 179, 221, 222, and 223.)

[119] A.H. Land and A.G. Doig. An Automatic Method of Solving Discrete Programming Problems. *Econometrica*, 28(3):497–520, 1960. (Cited on page 10.)

[120] T. Larsson and D. Yuan. An Augmented Lagrangian Algorithm for Large Scale Multicommodity Routing. *Computational Optimization and Applications*, 27(2): 187–215, February 2004. (Cited on page 98.)

[121] E.L. Lawler and D.E. Wood. Branch-and-Bound Methods: A Survey. *Operations Research*, 14(4):699–719, 1966. (Cited on page 10.)

[122] M. Leitner, M. Ruthmair, and G.R. Raidl. On Stabilized Branch-and-Price for Constrained Tree Problems. Technical report, Vienna University of Technology, Austria, 2011. (Cited on page 158.)

[123] Y. Li, M. Pióro, D. Yuan, and J. Su. Optimizing compatible sets in wireless networks through integer programming. *EURO Journal on Computational Optimization*, 2(1-2):1–15, 2014. (Cited on page 43.)

[124] C. Liebchen, M.E. Lübbecke, R. Möhring, and S. Stiller. Recoverable Robustness. Technical report, TU Berlin, 2007. ARRIVAL - TR - 0066. (Cited on pages 25 and 221.)

[125] C. Liebchen, M.E. Lübbecke, R. Möhring, and S. Stiller. The Concept of Recoverable Robustness, Linear Programming Recovery, and Railway Applications. In R. Ahuja, R. Möhring, and C. Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 1–27. Springer Berlin Heidelberg, 2009. (Cited on page 221.)

[126] A. Lodi and A. Tramontani. Performance variability in mixed-integer programming. In H. Topaloglu, editor, *Tutorials in Operations Research*, Vol. 10, pages 1–12. INFORMS, 2013. (Cited on pages 233, 234, and 235.)

[127] M.E. Lübbecke and J. Desrosiers. Selected Topics in Column Generation. *Operations Research*, 53(6):1007–1023, November-December 2005. (Cited on page 158.)

[128] J. Luedtke, S. Ahmed, and G.L. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2): 247–272, 2010. (Cited on pages 79 and 85.)

[129] S. Ma and D. Sun. Chance Constrained Robust Beamforming in Cognitive Radio Networks. *IEEE Communications Letters*, 17(1):67–70, 2013. (Cited on page 80.)

[130] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji. Cell Association and Interference Coordination in Heterogeneous LTE-A Cellular Networks. *IEEE Journal on Selected Areas in Communications*, 28(9):1479–1489, December 2010. (Cited on page 55.)

[131] T.L. Magnanti, P. Mirchandani, and R. Vachani. The convex hull of two core capacitated network design problems. *Mathematical Programming*, 60(2):233–250, 1993. (Cited on page 93.)

[132] T.L. Magnanti, P. Mirchandani, and R. Vachani. Modelling and Solving the Two-Facility Capacitated Network Loading Problem. *Operations Research*, 43(1):142–157, 1995. (Cited on page 93.)

[133] V. Mancuso and S. Alouf. Reducing costs and pollution in cellular networks. *Communications Magazine, IEEE*, 49(8):63–71, August 2011. (Cited on page 225.)

[134] T. Manning. *Microwave Radio Transmission Design Guide*. Artech House, 2009. (Cited on page 36.)

[135] S. Martello and P. Toth. *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & Sons, 1990. (Cited on pages 15, 18, 125, 181, and 198.)

[136] R. Mathar and T. Niessen. Optimum positioning of base stations for cellular radio networks. *Wireless Networks*, 6:421–428, 2000. (Cited on page 45.)

[137] R. Mathar, M. Reyer, and M. Schmeink. A Cube Oriented Ray Launching Algorithm for 3D Urban Field Strength Prediction. In *IEEE International Conference on Communications (ICC '07)*, pages 5034–5039, June 2007. (Cited on page 61.)

[138] S. 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 173, 175, 179, 183, and 193.)

[139] B.L. Miller and H.M. Wagner. Chance Constrained Programming with Joint Constraints. *Operations Research*, 13(6):930–945, 1965. (Cited on page 79.)

[140] M. Minoux. Multicommodity Network Flow Models and Algorithms in Telecommunications. In M.G.C. Resende and P.M. Pardalos, editors, *Handbook of Optimization in Telecommunications*, pages 163–184. Springer US, 2006. (Cited on page 37.)

[141] M. Monaci, U. Pferschy, and P. Serafini. Exact solution of the robust knapsack problem. *Computers & Operations Research*, 40(11):2625–2631, 2013. (Cited on pages 122, 179, 183, and 184.)

[142] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience, New York, NY, USA, 1999. (Cited on pages 13 and 14.)

[143] A. Nemirovski and A. Shapiro. Convex Approximations of Chance Constrained Programs. *SIAM Journal on Optimization*, 17(4):969–996, 2007. (Cited on page 79.)

[144] N. Nepomuceno. *Network optimization for wireless microwave backhaul*. PhD thesis, INRIA Sophia Antipolis, M ASCOTTE Team, 2010. (Cited on pages 27 and 102.)

[145] Next Generation Mobile Networks (NGMN), 2015. URL `http://www.ngmn.org`. accessed 30.03.2015. (Cited on page 27.)

[146] Z. Niu, S. Zhou, Y. Hua, Q. Zhang, and D. Cao. Energy-Aware Network Planning for Wireless Cellular System with Inter-Cell Cooperation. *IEEE Transactions on Wireless Communications*, 11(4):1412–1423, April 2012. (Cited on page 45.)

[147] E. Oh and B. Krishnamachari. Energy Savings through Dynamic Base Station Switching in Cellular Wireless Access Networks. In *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pages 1–5, December 2010. (Cited on page 225.)

[148] E. Olinick. Mathematical Programming Models for Third Generation Wireless Network Design. In J. Kennington, E. Olinick, and D. Rajan, editors, *Wireless Network Design: Optimization Models and Solution Procedures*. Springer, 2011. (Cited on page 138.)

[149] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly. SNDlib 1.0–Survivable Network Design Library. *Networks*, 55(3):276–286, 2010. URL `http://sndlib.zib.de`. (Cited on pages 100 and 101.)

[150] M. Padberg. On the Facial Structure of Set Packing Polyhedra. *Mathematical Programming*, 5(1):199–215, 1973. (Cited on page 45.)

[151] S. Parsaeefard and A.R. Sharafat. Robust Distributed Power Control in Cognitive Radio Networks. *IEEE Transactions on Mobile Computing*, 12(4):609–620, April 2013. (Cited on page 127.)

[152] I.C. Paschalidis and R. Wu. On robust maximum lifetime routing in wireless sensor networks. In *47th IEEE Conference on Decision and Control (CDC 2008)*, pages 1684–1689, December 2008. (Cited on page 127.)

[153] D. Pisinger. A minimal algorithm for the 0-1 Knapsack Problem. *Operations Research*, 45(5):758–767, 1997. (Cited on pages 179, 198, and 213.)

[154] D. Pisinger. Where are the hard knapsack problems? *Computers & Operations Research*, 32(9):2271–2284, 2005. (Cited on pages 16 and 200.)

[155] M. Poss. Robust combinatorial optimization with variable budgeted uncertainty. *4OR*, 11(1):75–92, March 2013. (Cited on page 25.)

[156] A. Prékopa. On probabilistic constrained programming. In *Proceedings of the Princeton Symposium on Mathematical Programming*, pages 113–138. Princeton University Press, 1970. (Cited on page 79.)

[157] C. Raack, A.M.C.A. Koster, S. Orlowski, and R. Wessäly. On cut-based inequalities for capacitated network design polyhedra. *Networks*, 57(2):141–156, 2011. (Cited on pages 93 and 94.)

[158] V. Ramamurthi, A. Reaz, D. Ghosal, S. Dixit, and B. Mukherjee. Channel, capacity, and flow assignment in wireless mesh networks. *Computer Networks*, 55(9):2241–2258, 2011. (Cited on page 42.)

[159] J. Riihijarvi, M. Petrova, and P. Mahonen. Frequency allocation for WLANs using graph colouring techniques. In *Second Annual Conference on Wireless On-demand Network Systems and Services (WONS 2005)*, pages 216–222, January 2005. (Cited on page 45.)

[160] A. Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93(2):195–215, 2002. (Cited on page 85.)

[161] D. Sabella, M. Caretti, W. Tomaselli, V. Palestini, B. Cendón, J. Valino, A. Medela, Y. Fernández, and L. Sanchez. Evaluation of ON-OFF Schemes and Linear Prediction Methods for Increasing Energy Efficiency in Mobile Broadband Networks. In Z. Becvar, R. Bestak, and L. Kencl, editors, *NETWORKING 2012 Workshops*, volume 7291 of *Lecture Notes in Computer Science*, pages 27–34. Springer Berlin Heidelberg, 2012. (Cited on page 225.)

[162] S. Sesia, I. Toufik, and M. Baker. *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley, 2009. (Cited on pages 40 and 41.)

[163] C.E. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27, 1948. (Cited on page 30.)

[164] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming*. MPS-SIAM Series on Optimization, 2009. (Cited on pages 22, 23, and 79.)

[165] I. Siomina and D. Yuan. Analysis of Cell Load Coupling for LTE Network Planning and Optimization. *IEEE Transactions on Wireless Communications*, 11(6):2287–2297, 2012. (Cited on page 45.)

[166] I. Siomina, P. Varbrand, and D. Yuan. An Effective Optimization Algorithm for Configuring Radio Base Station Antennas in UMTS Networks. In *IEEE 64th Vehicular Technology Conference (VTC 2006-Fall)*, pages 1–5, 2006. (Cited on page 55.)

[167] A.L. Soyster. Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21(5):1154–1157, 1973. (Cited on pages 24, 115, and 116.)

[168] S.R. Tayur, R.R. Thomas, and N.R. Natraj. An algebraic geometry algorithm for scheduling in presence of setups and correlated demands. *Mathematical Programming*, 69(1-3):369–401, 1995. (Cited on page 80.)

[169] E. Tomita, A. Tanaka, and H. Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theoretical Computer Science*, 363(1):28–42, October 2006. Computing and Combinatorics 10th Annual International Conference on Computing and Combinatorics (COCOON 2004). (Cited on pages 45 and 129.)

[170] P. Toth. Dynamic programming algorithms for the Zero-One Knapsack Problem. *Computing*, 25(1):29–45, 1980. (Cited on page 18.)

[171] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005. (Cited on page 27.)

[172] K. Tutschku, N. Gerlich, and P. Tran-Gia. An Integrated Approach To Cellular Network Planning. In *Proceedings of the 7th International Network Planning Symposium Networks 96*, pages 185–190, 1996. (Cited on page 33.)

[173] T. J. van Roy. A Cross Decomposition Algorithm for Capacitated Facility Location. *Operations Research*, 34(1):145–163, 1986. (Cited on page 35.)

[174] F. Vanderbeck. Implementing mixed integer column generation. In G. Desaulniers, J. Desrosiers, and M. Solomon, editors, *Column Generation*, chapter 12, pages 331–358. Springer, 2005. (Cited on pages 12, 156, 157, and 158.)

[175] A. Vigants. Space-diversity engineering. *Bell System Technical Journal*, 54(1): 103–142, 1975. (Cited on page 102.)

[176] R. Weismantel. On the 0/1 knapsack polytope. *Mathematical Programming*, 77(3): 49–68, 1997. (Cited on page 18.)

[177] W.E. Wilhelm and E.I. Gokce. Branch-and-Price Decomposition to Design a Surveillance System for Port and Waterway Security. *IEEE Transactions on Automation Science and Engineering*, 7(2):316–325, April 2010. (Cited on page 149.)

[178] L.A. Wolsey. Faces for a linear inequality in 0–1 variables. *Mathematical Programming*, 8(1):165–178, 1975. (Cited on pages 16 and 18.)

[179] L.A. Wolsey. *Integer Programming*. John Wiley & Sons, 1998. (Cited on pages 13 and 94.)

[180] H. Wu, M. Shahidehpour, Z. Li, and W. Tian. Chance-Constrained Day-Ahead Scheduling in Stochastic Power System Operation. *IEEE Transactions on Power Systems*, 29(4):1583–1591, 2014. (Cited on page 80.)

[181] E. Zola, P. Dely, A.J. Kassler, and F. Barcelo-Arroyo. Robust Association for Multi-radio Devices under Coverage of Multiple Networks. In V. Tsaoussidis, A.J. Kassler, Y. Koucheryavy, and A. Mellouk, editors, *Wired/Wireless Internet Communication*, volume 7889 of *Lecture Notes in Computer Science*, pages 70–82. Springer Berlin Heidelberg, 2013. (Cited on page 127.)

Index