SAPIENZA
UNIVERSITÀ DI ROMA

Sapienza University of Rome

Department of Computer, Control and Management Engineering
PhD in Automatic Control, Bioengineering and Operations Research

Doctoral Thesis in Operations Research

Location problems with covering constraints: models and solution approaches for the telecommunications

Supervisor
Laura Palagi

Candidate
Alice Calamita
1705585

Academic Year MMXXIII-MMXXIV (XXXVI cycle)

Abstract

Driven by the growing influence of telecommunications in contemporary society, this doctoral thesis offers novel contributions to the modeling and solution of location problems involving covering constraints for the telecommunications. We address two distinct network design problems: the first pertains specifically to the telecommunications sector, while the second has broader applicability to service and communications networks.

Specifically, the first contribution focuses on the location of the transmitters, i.e. the facilities enabling wireless connection, to meet service coverage requirements. In the modern context of increasing traffic, establishing suitable locations and power emissions for the transmitters is a relevant but challenging task due to heavy radio spectrum congestion, leading to signal interference and subsequent service degradation. Traditional network design formulations are very ill-conditioned and suffer from numerical inaccuracies and limited applicability to large-scale practical scenarios. Our contribution consists of speeding up the solution of the problem under consideration by addressing its drawbacks from a modeling point of view. We discuss the modeling of the technological constraints concerning the quality of service, and propose valid cutting plans and constraints aggregation, along with various presolve operations to reduce the problem size and strengthen existing formulations. Our proposals prove effective, allowing us to achieve optimality on large-scale scenarios in solution times aligning well with planning windows.

The second contribution concerns the introduction of mixed-integer quadratic formulations of a novel problem related to the design of service and communications networks. The problem is a location problem with a covering constraint allowing for partial coverage that takes into consideration both the minimization of the congestion and the protection from the uncertainty in customer demand. In particular, motivated by the contemporary society's growing demand for high service quality, we penalize congestion responsible for degrading the service and account for uncertainties in a robust framework. To solve this problem, we propose several Benders decomposition approaches and introduce a cut-strengthening technique to efficiently deal with the degeneracy of the Benders subproblem. Our tailored approach clearly outperforms the state-of-the-art solver Gurobi on adapted instances from the literature.

Keywords: *Mixed-integer nonlinear programming, 0-1 Linear programming, Location problems, Covering constraints, Partial set covering location, Wireless network design, Benders decomposition, Robust optimization, Perspective reformulation, Valid inequalities*

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

0-1 LP 0-1 (or Binary) Linear Programming

4G 4^{th} Generation

5G 5^{th} Generation

B&B Branch-and-Bound

B&C Branch-and-Cut

BD Benders Decomposition

CFLP Congested Facility Location Problem

CPSCLP Congested Partial Set Covering Location Problem

EMF ElectroMagnetic Field

FLP Facility Location Problem

FUB Fondazione Ugo Bordoni

GUB Generalized Upper Bound

ILP Integer Linear Programming

LB Lower Bound

LP Linear Programming

LTE Long Term Evolution

MCLP Maximal Covering Location Problem

MILP Mixed-Integer Linear Programming

MINLP Mixed-Integer Non-Linear Programming

MIP Mixed-Integer Programming

MIQCP Mixed-Integer Quadratically Constrained Programming

MIQP Mixed-Integer Quadratic Programming

MISOCP Mixed-Integer Second-Order Cone Programming

NLP Non-Linear Programming

PSCLP Partial Set Covering Location Problem

RCF Reduced Cost Fixing

SCLP Set Covering Location Problem

SINR Signal-to-Interference-plus-Noise Ratio

SOC Second-Order Cone

UB Upper Bound

VUB Variable Upper Bound

WND Wireless Network Design

Wireless Network Design: Sets & Parameters

δ SINR threshold

ϵ Penetration factor of the technology

$\mathcal{B}$ Set of potential transmitters

$\mathcal{L}$ Set of feasible power indices

$\mathcal{P}$ Set of feasible power values

$\mathcal{T}$ Set of receivers

μ System noise

$\tilde{p}_t^{add}$ Additional power density in testpoint $t \in \mathcal{T}$ produced by the newly installed transmitters

$\tilde{p}_t^{base}$ Baseline power density in testpoint $t \in \mathcal{T}$ produced by the existing transmitters

$\tilde{a}_{tb}$ Fading coefficient applied to the signal from transmitter $b \in \mathcal{B}$ to testpoint $t \in \mathcal{T}$

c_l Cost of using power level $l \in \mathcal{L}$

I_{tb}^{zone} 0-1 Parameter defining if testpoint $t \in \mathcal{T}$ is inside the exclusion zone of transmitter $b \in \mathcal{B}$

l_t Power density limit in testpoint $t \in \mathcal{T}$

M_{tb} Big-M associated with the SINR(t, b) constraint

P_{max} Maximum value of feasible power

r Coverage level

r_t Weight of testpoint $t \in \mathcal{T}$ in the coverage constraint

u Maximum number of users

a_{tb} Power received in testpoint $t \in \mathcal{T}$ of the signal emitted by transmitter $b \in \mathcal{B}$

d_t Number of users in testpoint $t \in \mathcal{T}$

Wireless Network Design: Variables

z_{bl} 0-1 Variable representing if transmitter $b \in \mathcal{B}$ is emitting at power level $l \in \mathcal{L}$

z_b 0-1 Variable representing if transmitter $b \in \mathcal{B}$ is activated

w_t 0-1 Variable defining if testpoint $t \in \mathcal{T}$ falls inside at least one exclusion zone

x_{tb} 0-1 Variable representing if testpoint $t \in \mathcal{T}$ is served by transmitter $b \in \mathcal{B}$

y_{tb} 0-1 Variable defining if the power density in testpoint $t \in \mathcal{T}$ given by the activation of transmitter $b \in \mathcal{B}$ should be considered

Congested Partial Set Covering Location: Sets & Parameters

Γ Maximum number of demand deviations taken into account

$\hat{d}_j$ Deviation from the nominal demand of customer $j \in J$

a Weight of the quadratic component of congestion cost

b Weight of the linear component of congestion cost

D Target demand

d_j Nominal demand of customer $j \in J$

f_i Opening cost of facility $i \in I$

I Set of potential facilities

J Set of customers

Congested Partial Set Covering Location: Variables

ρ, σ Variable accounting for the uncertainty in the congestion

τ, π Variables accounting for the uncertainty in the coverage

u_i Variables representing the quadratic load of facility $i \in I$

v_i Variable representing the load of facility $i \in I$

x_{ij} Variable representing the fraction of demand of customer $j \in J$ served by facility $i \in I$

y_i 0-1 Variable representing if facility $i \in I$ is open

Part I

Motivation & Background

Chapter 1

Introduction

1.1 Research motivation

Facility location problems pose a fundamental question that captures the interest of both academics and practitioners: "Where should facilities be optimally placed to serve a set of customers?". This question gives rise to numerous subclasses of problems, each catering to different aspects, constraints, and objectives derived from real-world applications. One of these variants encompasses *location problems* including a *covering constraint*, which seek to identify the optimal locations for facilities or service providers in such a way that they satisfy coverage for a targeted set of customers.

Location problems with covering objectives or constraints are commonplace in several sectors (see e.g., [58, 67, 71, 81, 132]): in the location of service providers (such as schools, hospitals, libraries, restaurants, retail outlets, banks) or emergency facilities and vehicles (such as fire stations, ambulances, oil spill equipment), or base stations in the telecommunications context. While many applications involve a relatively small number of demand points and potential facility locations and can therefore be solved in a satisfactory way by existing heuristics or by general-purpose solvers, there are also cases where the number of demand points can run in the thousands or even millions. This is the case of *telecommunications systems*, which is the main application of this thesis. Specifically, this thesis addresses two network design problems, one specifically for the telecommunications context, and the other that can be generally applied to service and communications networks (including, e.g., telecommunication networks, healthcare networks, supply chain networks, and many others). In the sphere of telecommunications, network design consists of determining the optimal locations for the *transmitters* (sometimes we will indicate them with the terms *antennas* or *base stations*) to guarantee that every user or device of the

infrastructure gets satisfactory signal strength or data at the minimum possible cost.

The motivation behind this study can be summarized in the following quotations: from the Preface of the book [122]

> Telecommunications has had a major impact in all aspects of life in the last century. There is little doubt that the transformation from the industrial age to the information age has been fundamentally influenced by advances in telecommunications. [...] Optimization problems are abundant in the telecommunications industry. The successful solution of these problems has played an important role in the development of telecommunications and its widespread use. Optimization problems arise in the design of telecommunication systems and in their operation.[...] The spectrum of topics covered includes design of telecommunication networks, routing, network protection, grooming, restoration, wireless communications, network location and assignment problems, internet protocol, world wide web, and stochastic issues in telecommunications.

from the section dedicated to optimization for telecommunications industry [90] in the website of Gurobi, a leading general purpose mixed-integer programming (MIP) solver

> As we enter into the era of 5G, the telecommunications industry is undergoing profound and rapid change. Companies and organizations across the telecommunications value chain – from telecom services providers to telecom equipment manufacturers to government regulators and other key players such as vendors and consultants – must be able to transform their businesses to cope with the changes, overcome the challenges, and capitalize on the opportunities created by the emergence of 5G.

and from the Gurobi collection of challenges for telecommunications companies in 2021 [89]

> Major telecom organizations across the industry – from operators and service providers to equipment manufacturers and government regulators – use mathematical optimization to: address a wide (and ever-expanding) array of strategic, tactical, and operational problems; make optimal plans and decisions; and achieve their business objectives. Here are five key areas where mathematical optimization can deliver immense business benefits for telecom companies. [Then, about the first key area, which is Network Planning we read:] It is essential that telecom operators plan and deploy their 5G networks to optimize coverage, service levels, demand, and time to market. [...]

> Leading global telecom players, like Vodafone, use mathematical optimization to plan, configure, and operate their networks as efficiently and profitably as possible. Mathematical optimization is extremely effective in automating and optimizing numerous network planning processes including fiber optic network planning, facility location planning, coverage and frequency planning, radio planning, and demand planning.[...] Other planning tools, such as machine learning and meta-heuristics, are simply not capable of handling the complexity of telecommunications network planning problems and delivering optimal solutions.

These quotations stress the following *motivating factors* for this thesis: (i) the ubiquitous influence of telecommunications on modern life, (ii) the crucial role of mathematical optimization in the design of telecommunications networks and its employment by major organizations operating in this industry, (iii) the advent of 5G posing even more stringent network quality requirements and making the study of planning problems in the telecommunications sector still relevant, (iv) the fact that, in this constantly evolving landscape of telecommunications, at the forefront of challenges faced by network providers and policymakers there is still the need to ensure optimal facility location planning, coverage and connectivity.

1.2 Thesis overview

In this doctoral thesis entitled "Location problems with covering constraints: models and solution approaches for the telecommunications", we address the design of telecommunications networks, essential components shaping the current and future functionality of global connectivity. Specifically, we focus on the location of facilities enabling connection to meet coverage requirements, using rigorous mathematical modeling and proposing efficient algorithmic approaches to drive optimal planning.

To a greater extent, this thesis addresses two distinct network design problems: the first pertains specifically to the telecommunications sector, while the second has broader applicability to service and communications networks.

The first research topic addressed in this thesis concerns the *optimal design of wireless networks*. In the current era of pervasive connectivity, the design of wireless networks still plays a pivotal role in shaping modern societal infrastructure. These networks, which enable various services, from mobile communication to internet services, have become

the foundation of the digital age. The importance of studying wireless network design lies in its profound implications for enhancing communication and enabling technological advancements, and its influence on the evolving dynamics of global connectivity in daily lives and professional interactions.

A *wireless network* is a telecommunications network that uses radio waves, or other wireless communication technologies, to transmit and receive data without the need for physical cables, allowing for the wireless connectivity of devices. From a design point of view, the basic elements of wireless networks are *transmitters* and *receivers*. Hence, the design of these networks consists of identifying the proper *locations* for the transmitters and setting their *operational parameters* – like frequency and power emission – in such a way to *cover* with service the receivers located in the area of interest.

Wireless networks are the basis of several services, such as television, radio, mobile communication, and the internet. In this thesis, we will address the design of wireless networks for 4G LTE/5G technology. Even if wireless networks rely on different technologies and standards based on the service they are meant to provide, they all share a common feature: the need to reach users scattered over a vast area with a radio signal that must be strong enough to prevail against other unwanted interfering signals. The quality of service (and hence the coverage) indeed depends on the interplay of numerous signals, wanted and unwanted, generated from a large number of transmitting devices. The increasing traffic and the densification of the base stations along the territory have led to an increase in *interfering* signals. Consequently, establishing suitable locations and power emissions for all the transmitters, coexisting within a heavily congested radio spectrum, has become a challenging and relevant task.

This research was carried out under the supervision of Professor Pasquale Avella and in collaboration with Fondazione Ugo Bordoni (FUB) [76], a higher education and research institution in telecommunications under the supervision of Italian Ministry of Enterprises and Made in Italy (MISE), providing innovative services for government entities. The collaboration with the FUB was specifically intended to identify a tool, manageable by practitioners, that does the optimal planning of LTE and 5G radio base stations for the Municipality of Bologna in Italy.

The second research topic concerns the investigation of the *optimal design of service and communications networks* with the operational objectives of safeguarding against both network *congestion* and the *unpredictability* of service demand. Again, the goal is to locate facilities or service providers along the territory, in order to satisfy the demand of the

customers distributed throughout the territory. This time, service *coverage requirements* are not expressed through signal strength and hence we do not need to deal with signal interference. Coverage is instead defined as meeting the demand, where the type of demand varies, depending on the application, and may represent, for example, a specific amount of data within a telecommunications system or a specific number of product units within a manufacturing system.

In this context, we address the cases in which (i) congestion arising at the service providers affects the quality of the service or induces extra costs, and (ii) customer demand cannot be precisely estimated. This not only enables planning decisions that ensure delivery schedules and uphold rigorous quality standards, but also aligns with the modern society's escalating demand for high-quality service. Consider, for instance, the design of a telecommunications network in hard scenarios like dense urban environments. Neglecting uncertainty in demand for service in such a situation can lead to unpleasant surprises in coverage plans: an unexpectedly higher quantity of users in a specific area means highly congested transmitters. Due to limited capacity at the transmitter, congestion can result in poor or even completely unavailable service. We also note that planning a new communications service with a time horizon of several years, implies that the demand from the customers is naturally subject to uncertainty and pretty hard to predict due to a lack of historical data or lack of trust in data from other services. We therefore employ robust optimization to protect network planning from demand uncertainty and pursue the minimization of congestion to prevent facility overload.

This topic was developed in collaboration with Professor Ivana Ljubić during a visiting period at her institute.

1.3 Research approach and contributions

For what concerns the first application, addressing wireless network design in a modern context of increasing traffic and consequent densification of base stations, is a challenge that requires the application of sophisticated optimization techniques. The first contributions on the subject appeared already at the beginning of the 1980s. Several studies (see [49, 53]) suggested that employing optimization-oriented planning techniques can heavily reduce infrastructure costs and enhance coverage quality and resource efficiency significantly. Traditional solution methods, employing (mixed-)integer linear programs with (very) ill-conditioned coefficient matrices, suffer from numerical inaccuracies and limited applicability to large-scale practical scenarios. Indeed, the traditional modeling choice typically includes

big-M coefficients to model coverage conditions, leading to very weak linear relaxations and solutions – returned by state-of-the-art MIP solvers – typically far from the optimum [48].

Our contribution consists of speeding up the solution of the problem under consideration, by addressing its drawbacks, from a modeling point of view. We worked on multiple fronts to reduce memory and numerical issues so as to allow a rapid and accurate solution to these problems, in line with the times and precision required in the planning phase.

In the second application of this thesis, we define a novel problem that integrates features from two combinatorial problems: partial set covering location and congested facility location. Partial set covering location is a special case of set covering location problem, which seeks to optimally place a minimum-cost set of facilities to satisfy a partial target demand. The introduction of partial covering provides more cost-efficient and practical solutions. Congested facility location problem instead incorporates in facility location the complexity of possible congestions at open facilities to express the diseconomies of scale or penalization in the quality of service produced by the congestion.

Our modeling contribution in this context is the introduction of the congested version of the partial set covering location problem. We also formulate its robust counterpart, taking into consideration the case where up to a certain (given) number of customers deviate from the nominal demand and we want to protect from the worst-case realization that meets this hypothesis. Finally, we give a methodological contribution by presenting some approaches relying on Benders decomposition to solve the robust problem efficiently.

1.4 Thesis structure

The thesis is structured in two parts. Part I – including chapters from 1 to 4 – introduces the thesis project, motivation and objectives and gives the theoretical background of the problems and solution methods that will be discussed in the second part of the thesis. Part II – including chapters from 5 to 7 – presents our contributions to the modeling and solution of location problems with covering constraints.

Chapter 2 presents the underlying theoretical framework and the methodological tools necessary to interpret and understand our contributions. We introduce and formulate generic mixed-integer programming problems. The basics of two well-known solution algorithms for this class, namely the branch-and-bound algorithm and the Benders decomposition method, are given. Finally, a robust framework to deal with parameter uncertainty is presented.

In Chapter 3, the wireless network design problem is discussed in detail. We provide an overview of the current scientific literature on the topic, along with background information on wireless networks. A discussion on the negative sides of the formulations typically used in the field is provided. Our contribution to the wireless network design literature is given in Chapter 5, where we investigate the strengthening of existing formulations and propose several presolve operations. Then, we report the computational experience testing our proposals on realistic instances provided by the FUB.

Chapter 4 focuses on two variants of facility location problems: the congested facility location and the partial set covering location. We introduce these two classes of problems – which inspired the problem of congested partial set covering location – and give an overview of the existing literature. Our contribution to the field is given in Chapter 6, which focuses on the congested partial set covering location problem. We motivate and formulate the deterministic problem and do the same for its robust counterpart. We discuss how to apply Benders decomposition on the robust counterpart of the problem and report all the implementation ingredients that have been instrumental in the success of the method. We then show the results of our proposals on adapted instances from the existing literature.

Finally, Chapter 7 summarizes the main findings and contributions of this thesis.

1.5 Dissemination

Our contributions to the wireless network design literature were presented at the International Conference on Optimization and Decision Sciences held in Rome (Italy) in September 2021, the 6$^{\text{th}}$ AIROYoung Workshop held in Rome (Italy) in March 2022 and the 32$^{\text{nd}}$ Conference of the Association of European Operational Research Societies held in Espoo (Finland) in July 2022. The following paper on the topic has already been published

P. Avella, A. Calamita, and L. Palagi (2023). A compact formulation for the base station deployment problem in wireless networks. Networks, 82(1), 52-67. https://doi.org/10.1002/net.22146

and the following

P. Avella, A. Calamita, and L. Palagi. Speeding up the solution of the site and power assignment problem in wireless networks. Under review. arXiv preprint arXiv:2210.04022

has been recently submitted for publication to an international journal and is currently undergoing the peer-reviewing process.

Our contributions to the field of set covering location were exposed at the 3rd EUROYoung Workshop held in Cercy (France) in June 2023, and the International Conference on Optimization and Decision Sciences held in Ischia (Italy) in September 2023. The following paper on the topic

> *A. Calamita, I. Ljubić, and L. Palagi. Benders decomposition for congested partial set covering location with uncertain demand. Under review. arXiv preprint arXiv:2401.12625*

has been recently submitted for publication to an international journal and is currently undergoing the peer-reviewing process.

Chapter 2

Methodological tools

This chapter is devoted to the exposure of the theoretical framework and methodological tools necessary to understand the research contributions shown in the present thesis. Specifically, we introduce mixed-integer programming (MIP) problems and give some basic notions to assess the quality of formulations. Then we discuss two exact methods for MIP problems that we will apply in Part II. Finally, we introduce a robust optimization framework to address data uncertainty, that will be useful to understand the second half of Part II.

2.1 Mixed-integer optimization problems and solution algorithms

An *optimization problem* is a problem of choosing the best alternative among a set S of distinct alternatives. For this reason, optimization lies at the heart of many *decision-making* processes. The choice of the best option is made according to a quantitative criterion, namely according to a given *objective function* $f(\cdot) : S \to \mathbb{R}$ that has to be minimized or maximized, and that assigns a value to each alternative.

A generic optimization problem can be defined as

$$\min\{f(x) : x \in S\}$$

where x is the vector of the *variables* to be optimized, and S is the *feasible region* including all possible choices. The problem can be

- *infeasible*, when $S = \emptyset$ and no solution exists;
- or *unbounded*, when for any given value $M > 0$ there always exists a vector $\bar{x} \in S$

such that $f(\bar{x}) < -M$;

- or *bounded below*, when there exists a value M such that $f(x) \geq M$ for all $x \in S$ but the problem does not admit an optimal solution;
- or admits an *optimal* solution, when there exists a vector $x^\star \in S$ such that $f(x^\star) \leq f(x)$ for all $x \in S$; in this case, $x^\star$ is said optimal solution and $f(x^\star)$ optimal value.

In general, the characteristics of the set S and the objective function f determine the class of the problem. We therefore introduce some of these classes that will be useful in the reading of this thesis.

LP The first class mentioned is *linear programming* (LP) and includes problems having a linear objective function and a feasible region that can be expressed through a finite set of linear constraints. An LP can be modeled as follows

$$\begin{aligned} \min \quad & c^T x \\ & Ax \leq b \\ & x \in \mathbb{R}^n \end{aligned}$$

where A is a matrix in $\mathbb{R}^{m \times n}$, b is a vector in $\mathbb{R}^m$ and c is a vector in $\mathbb{R}^n$.

ILP We discussed problems modeled using continuous variables. Sometimes, however, fractional solutions are not realistic. We therefore introduce discrete modeling, based on the use of integer or binary variables. Many optimization problems of significant importance, such as assignment problems or capital budgeting, can be formulated using *integer linear programming* (ILP) models.

We define (pure) ILP the class of problems having a linear objective function and a feasible region that can be expressed through the intersection of a finite set of linear constraints and a set of integer points. An ILP can be modeled as follows

$$\begin{aligned} \min \quad & c^T x \\ & Ax \leq b \\ & x \in \mathbb{Z}^n \end{aligned}$$

where A is a matrix in $\mathbb{R}^{m \times n}$, b is a vector in $\mathbb{R}^m$ and c is a vector in $\mathbb{R}^n$. We talk about

pure *binary* (or *0-1*) *linear programming* (0-1 LP) when we have all binary variables, i.e. when $x \in \{0,1\}^n \subseteq \mathbb{Z}^n$.

MILP Several applications involve both discrete decisions and decisions that are continuous in nature. Hence, we introduce the broader class of *mixed-integer linear programming* (MILP) problems, characterized by a linear objective function, and a feasible region that can be expressed through a finite set of linear constraints, some variables are discrete and others are continuous.

MILPs can be modeled as follows

$$\begin{aligned} \min \quad & c^T x \\ & Ax \leq b \\ & x_C \in \mathbb{R}^{|C|} \\ & x_I \in \mathbb{Z}^{|I|} \end{aligned} \tag{2.1}$$

where A is a matrix in $\mathbb{R}^{m \times n}$, b is a vector in $\mathbb{R}^m$ and c is a vector in $\mathbb{R}^n$. The set I is the index set of discrete variables and C is the index set of continuous variables, such that $I \cup C = \{1, \ldots, n\}$. MILPs are considered NP-hard problems and therefore recognized as hard to solve. We also observe that they include ILPs and LPs as special cases.

MINLP *Mixed-integer nonlinear programming* (MINLP) is one of the most general modelling paradigms in optimization as it combines the combinatorial challenge of optimizing over discrete variable sets with the difficulties of handling nonlinear functions. The past 30 years have seen a dramatic increase in new MINLP models in several real-world applications, which has motivated the development of a broad range of new techniques to tackle this challenging class of problems [18]. For comprehensive reviews on models, applications and algorithmic advances for MINLPs see [18] and [103].

Formally, we define MINLP as the class of problems having a possibly nonlinear objective function, and/or a feasible region that can be expressed through a finite set of possibly nonlinear constraints, some variables are discrete and others are continuous. MINLPs can

be modeled as follows

$$\begin{aligned} \min \quad & f(x) \\ & g(x) \leq 0 \\ & x \in X \\ & x_C \in \mathbb{R}^{|C|} \\ & x_I \in \mathbb{Z}^{|I|} \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable functions, X is a bounded polyhedral set. The set I is the index set of discrete variables and C is the index set of continuous variables, such that $I \cup C = \{1, \ldots, n\}$. An MINLP is said convex if the functions f and g are convex. It is said non-convex if either f or any g_i (with $i = 1, \ldots, m$) are non-convex functions. MINLPs include MILPs as a special case and are therefore considered NP-hard problems.

Special cases of this class also include

- mixed-integer quadratically constrained programming (MIQCP) problems, having a possibly quadratic objective function, and a feasible region that can be expressed through a finite set of quadratic constraints, some variables are discrete and others are continuous;
- mixed-integer quadratic programming (MIQP) problems, having a quadratic objective function, and a feasible region that can be expressed through a finite set of linear constraints, some variables are discrete and others are continuous.

In this thesis, we mainly focus on ILPs and convex MINLPs (more specifically MIQCPs). Therefore, in this section we introduce the following exact methods to address these classes of problems:

- methods based on *branch-and-bound* for ILPs/MILPs, that can be extended to convex MINLPs;
- *Benders decomposition* method for MILPs and convex MINLPs that meet specific characteristics.

Building blocks of solution methods The two main ingredients on which methods for ILP/MILP/MINLP problems are based are relaxation and constraint enforcement (see [18]).

A *relaxation* of the problem can be obtained by enlarging the feasible set, e.g., by ignoring some constraints of the problem. See an example of relaxation in Figure 2.1, where integrality constraints are relaxed. Relaxations are used to provide lower bounds to the

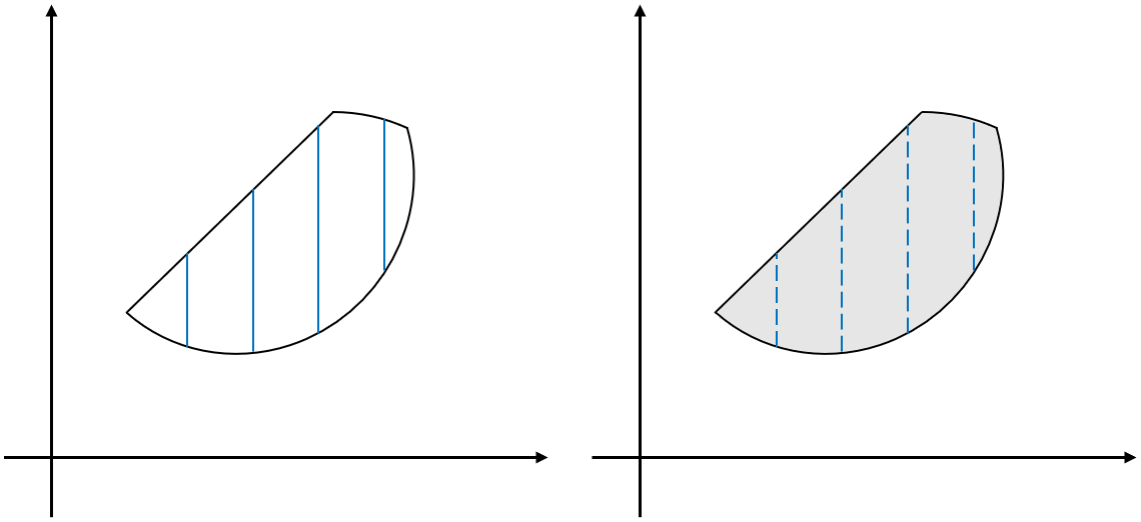

Figure 2.1: On the left is the feasible set of a MINLP problem, given by the vertical blue lines. On the right is the formulation we get by relaxing integrality, given by the grey region.

MIP problem. If the solution to a relaxation is feasible for the MIP, then it also solves the MIP. In general, however, the solution is not feasible for the MIP, and we must somehow exclude this solution from the relaxation.

Constraint enforcement is a procedure used to exclude solutions that are feasible for the relaxation but not to the original problem. Constraint enforcement may be done by refining or tightening the relaxation, or by branching. One way to tighten the relaxation is by adding valid inequalities to the relaxation as in Figure 2.2. A *valid inequality* is an inequality that is satisfied by all feasible solutions for the MIP. When it successfully excludes a given infeasible solution, it is called a *cut* or a *cutting plane*. This type of separation is used, e.g., in the cutting plane algorithm – that recursively solves a relaxation of the problem and produces a cutting plane to exclude the infeasible relaxation solution until a feasible relaxation solution is found – or in Benders decomposition (see Section 2.1.3).

Branching consists of dividing the feasible region into subsets such that every solution to MIP is feasible in one of the subsets. When integrality is relaxed as in Figure 2.3, branching on an integer variable that takes a fractional value (the black point) yields two separate relaxations, such that all solutions of the MIP lie in one of these two. This approach is the basis of the branch-and-bound discussed in Section 2.1.2.

In the following sections we discuss the importance of having a good formulation, and provide two well-known algorithms used to solve MILPs and some convex MINLPs, the

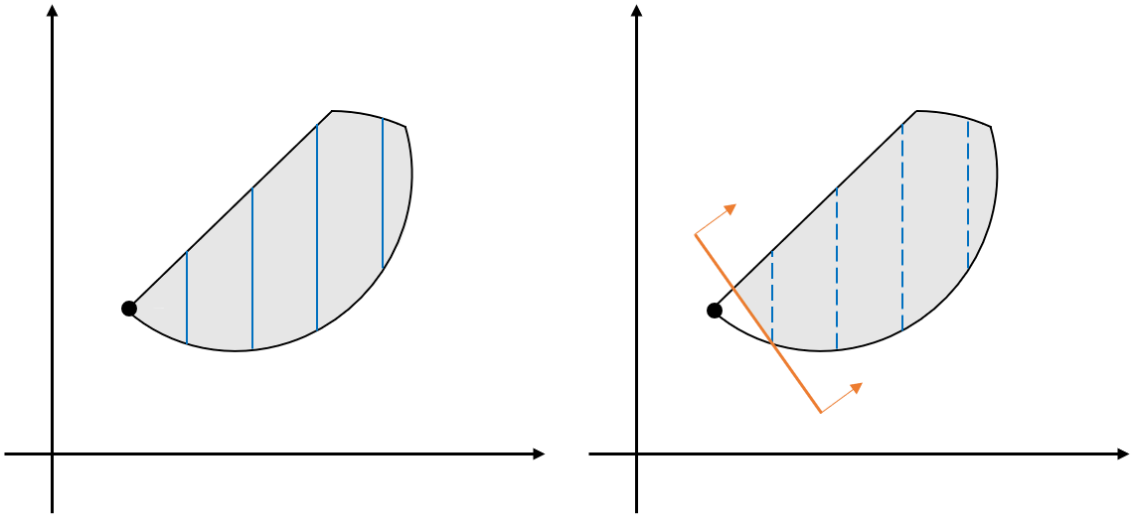

Figure 2.2: How to refine the relaxation using a valid cut discarding the black point.

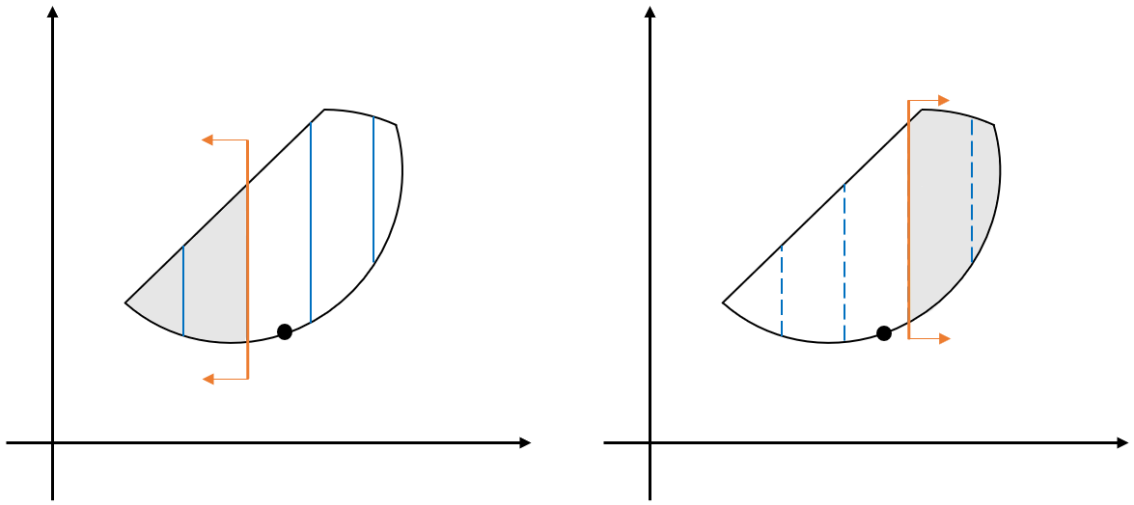

Figure 2.3: The two sets we get by branching on the black point.

branch-and-bound method and the Benders decomposition method.

2.1.1 Quality of formulations

Throughout the thesis, we see how not all formulations are equal: in Chapter 5, we discuss how to strengthen the natural formulation of an ILP problem, while in Chapter 6, we exploit perspective reformulation to speed up the solution of a MIQP problem. It is, therefore, important to introduce the concept of formulation and determine how to compare different formulations to identify the best one (if possible) between them.

In particular, consider the ILP problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \mathbb{Z}^n \end{aligned}$$

with A a matrix in $\mathbb{R}^{m \times n}$, b vector in $\mathbb{R}^m$, and c vector in $\mathbb{R}^n$. The set of linear constraints of the problem, given by $Ax \leq b$, forms a polyhedron and is called *formulation* as it coincides with the LP relaxation of the problem. The *LP relaxation* of the problem is obtained by relaxing the integrality constraints on x.

To solve a problem with this structure, the methods used iteratively attempt to reduce the difference between a *lower bound* (LB) and an *upper bound* (UB) of the problem. This difference is called *gap*. The lower bound refers to a value less than or equal to the optimal value, while the upper bound refers to a value greater than or equal to the optimal value. For minimum problems, an upper bound is given by the value of a feasible solution, while the optimal value of a relaxation of the problem gives a lower bound. One of the most common techniques to identify a lower bound is the LP relaxation of the problem. The closer the bounds are to the optimal value, the smaller the gap (and therefore the "distance" from the optimal value), and the higher the quality of these bounds. When the gap is zero, the feasible solution is optimal and the solution of the LP relaxation is integral.

The quality of the gap is influenced by the quality of the formulation. Very often the same problem admits multiple formulations, and there might be a hierarchy in the quality of these formulations. In particular, the following criteria can be stated.

Criterion 2.1.1 *Given a set $X \subseteq \mathbb{R}^n$ and formulations P and P' for X, P' is stronger (better) than P if and only if $LB(P') \leq LB(P)$, regardless of the objective function of the problem.*

Criterion 2.1.2 *Given a set $X \subseteq \mathbb{R}^n$ and formulations P and P' for X, P' is stronger (better) than P if and only if $P' \subseteq P$ and $\exists x \in P : x \notin P'$.*

Criterion 2.1.3 *Given a set $X \subseteq \mathbb{R}^n$ and formulation P_I for X, P_I is said to be ideal if and only if it coincides with the convex hull of the integer points of X, i.e., if and only if $P_I = Conv(X)$.*

We observe that Criteria 2.1.1 and 2.1.2 are equivalent. The ideal formulation P_I is therefore contained in every other formulation of the problem, i.e., $P_I \subseteq P$ for every formulation P of the problem; we also have that $LB(P_I)$ is the optimal value.

If the ideal formulation is available, it will clearly be the best possible. However, often finding the convex hull of an ILP is just as hard as solving the ILP. In Figure 2.4, there is an example of three formulations that are comparable in terms of quality. Moreover, it is not always possible to compare the quality of two formulations of the same problem, i.e., when the following criterion applies.

Criterion 2.1.4 *Given a set $X \subseteq \mathbb{R}^n$ and formulations P and P' for X, P and P' are incomparable if and only if $\exists x \in P : x \notin P'$ and $\exists x \in P' : x \notin P$.*

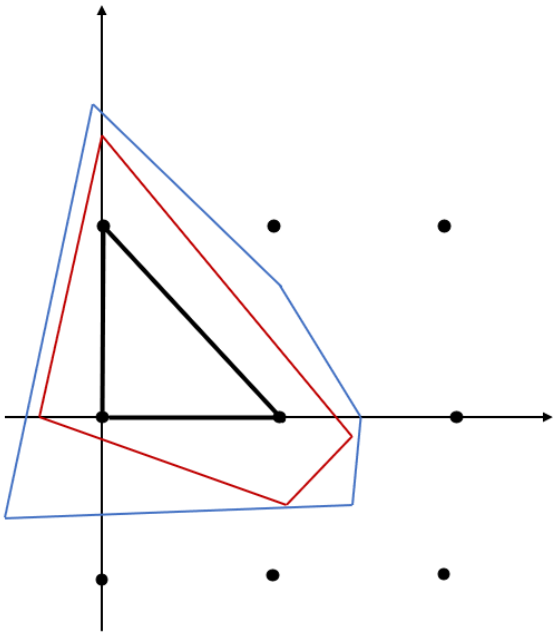

Figure 2.4: Three different formulations of the same problem. The red one is better than the blue one. The best is the ideal formulation, which is the one in black.

Perspective reformulation We report the results of [86] about perspective reformulation. Consider a set of three variables in which we can find "on-off" type decisions. Specifically, let y be a binary indicator variable that controls the continuous variable v. The set is defined as follows

$$S = \left\{(v, u, y) \in \mathbb{R}^2 \times \{0, 1\} : u \geq v^2,\, by \geq v \geq ay,\, u, v \geq 0\right\}$$

where a and b are constants in $\mathbb{R}_+$. We can observe that $S = S^0 \cup S^1$ where

$$S^0 = \left\{(0, u, 0) \in \mathbb{R}^3 : u \geq 0\right\} \quad \text{and} \quad S^1 = \left\{(v, u, 1) \in \mathbb{R}^3 : u \geq v^2,\, b \geq v \geq a,\, u, v \geq 0\right\}.$$

We now define what we mean by perspective function and state one property of perspective functions.

Definition 1 *The perspective of a given function $f : \mathbb{R}^n \to \mathbb{R}$ is the function $\tilde{f} : \mathbb{R}^{n+1} \to \mathbb{R}$ defined as follows*

$$\tilde{f}(y, v) = \begin{cases} yf(v/y) & \text{if } y > 0 \\ 0 & \text{if } y = 0 \\ \infty & \text{otherwise.} \end{cases}$$

Proposition 2.1.5 *Let $\tilde{f}$ be the perspective of function f. We have that $\tilde{f}$ is convex provided that f is convex.*

In our case, we have $f = v^2$ and $y \in \{0, 1\}$, which means the perspective of function f is given by

$$\tilde{f}(v, y) = \begin{cases} y(v^2/y^2) & \text{if } y = 1 \\ 0 & \text{if } y = 0. \end{cases}$$

where $y(v^2/y^2)$ can be simplified into v^2/y for $y = 1$.

Hence, if we replace the quadratic function of set S with its perspective, we get the following set

$$S^{persp} = \left\{(v, u, y) \in \mathbb{R}^2 \times \{0, 1\} : uy \geq v^2,\, by \geq v \geq ay,\, u, v \geq 0\right\}$$

and by relaxing the integrality on y we get its continuous relaxation

$$R^{persp} = \left\{(v, u, y) \in \mathbb{R}^3 : uy \geq v^2,\, by \geq v \geq ay,\, 1 \geq y \geq 0,\, u, v \geq 0\right\}.$$

From [38], the following result is valid.

Lemma 2.1.6 *$Conv(S) = R^{persp}$, i.e., the continuous relaxation of set S^{persp} is the convex hull of the integer points of set S.*

We now give the definition of *rotated second-order cones.*

Definition 2 *A rotated second-order cone in $\mathbb{R}^{n+2}$ is the set*

$$\{(v, u, y) \in \mathbb{R}^{n+2} : uy \geq v^T v, y \geq 0, u \geq 0\}.$$

We observe that the set S^{persp} contains rotated second-order cones. An example of a rotated second-order cone is $\mathbb{R}^3$ is given in Figure 2.5.

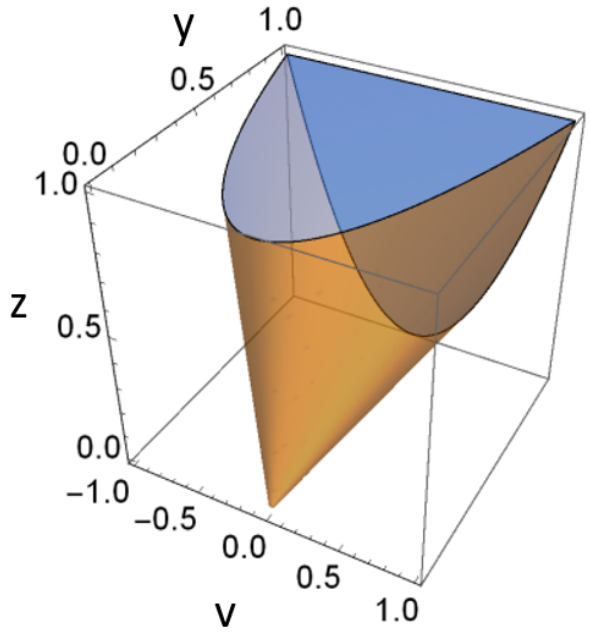

Figure 2.5: An example of a rotated second-order cone is $\mathbb{R}^3$. Picture borrowed from [101].

MIQCP problems with second-order cone constraints are called mixed-integer second-order cone programming (MISOCP) problems.

2.1.2 Branch-and-bound

In 1960, Land and Doig [99] introduced an exact method which is now the paradigm of all state-of-the-art solvers for MIP problems: the *branch-and-bound* (B&B) algorithm. B&B is

an efficient method that allows an *implicit* enumeration of the solution space, by avoiding the exploration of some regions of the feasible set that are considered somehow unpromising. Indeed, an exhaustive search of the solution space based on *explicit* enumeration of all the feasible solutions of an integer program is doable in principle, however it is computationally intractable for most of the cases.

There are two operations underlying the B&B method: *branching* and *bounding*. The branching operation is done to derive two smaller subproblems starting from a bigger one, and it is executed several times during a B&B procedure. The decomposition of each problem into two subproblems is represented through a tree structure, which is referred to as the *branch-and-bound tree* (or simply *branching tree*).
Bounding, instead, consists of evaluating the quality of each subproblem by solving its relaxation. As bounds are estimations of the optimal value, this technique is used to avoid the exploration of subproblems that are guaranteed not to lead to a better solution, making a more efficient exploration of the solution space.

We now describe the procedure in detail. Given a MILP minimization problem (e.g., modeled as in (2.1)), we denote by P_0 the initial problem and by R_0 its feasible set when we relax integrality constraints. We also denote by LB_0 the optimal value of the relaxation R_0, which is a lower bound to the problem P_0, given by the (possibly fractional) solution $\underline{x}_0$ of R_0. Finally, we denote by UB the best upper bound to the problem, corresponding to the feasible solution $\bar{x}_i$ (referred to as an *incumbent*). Starting from the initial problem at node 0 (also known as *root node*) of the branching tree, the B&B algorithm performs a bounding operation to determine the optimal value of the relaxation (LB_0), then decides if that region is interesting or not. If it is, it branches on the relaxation solution by imposing branching constraints and creating two more disjoint subproblems to analyze, otherwise it prunes the node, avoiding the chance to derive (and then analyze) subproblems from that node. The union of the feasible regions of the two subproblems contains all the feasible integer solutions of the problem from which the subproblems are derived. The subproblems not yet analyzed are inserted in a list denoted as $\mathcal{L}$.
Then, the B&B algorithm proceeds by solving subproblems from list $\mathcal{L}$ and for every subproblem P_i (corresponding to a node of the branching tree) the algorithm relaxes integrality constraints on discrete variables to get R_i, gets the optimal solution $\underline{x}_i$ of R_i, whose value is LB_i, and decides if pruning node i or imposing branching constraints on P_i to derive two more subproblems to be inserted in the list $\mathcal{L}$.

It remains to explain how we decide that an area of the feasible region is not interesting. Generally, a node i is pruned if one of the following conditions is met:

- R_i is infeasible, i.e. $R_i = \emptyset$, and accordingly the corresponding mixed-integer problem P_i is infeasible. In this case we *prune by infeasibility*;

- R_i is feasible and its optimal value LB_i is not less than the best upper bound found so far. Thus, P_i cannot contain any feasible solution whose value is better than the best upper bound and the node is pruned. In this case we *prune by bound*;

- R_i is feasible, LB_i is less than the best upper bound and the optimal solution $\underline{x}_i$ of R_i is feasible for the MILP. In this case, we update the best upper bound and the best incumbent respectively to LB_i and $\underline{x}$, and the node is not further explored since the associated MILP problem has been solved to optimality. In this case we *prune by integrality*.

We observe that we say that one solution is feasible for the MILP problem when all the components of the solution that belongs to the set I of the integer components are integral.

A scheme of how the B&B algorithm works for a minimization problem is depicted in Figure 2.6.

As discussed in [18], the B&B algorithm can be extended to solve convex MINLP problems at global optimality. This variant takes the name of *nonlinear branch-and-bound* and was introduced in [47, 87]. Specifically, MINLPs can be solved using (nonlinear) B&B in two manners:

- by replacing all convex nonlinear constraints with a set of linear constraints obtained from first-order Taylor series approximation; in this way the MINLP becomes an MILP and traditional B&B can be applied. It should be noticed that the linear approximations undergo dynamic refinement during the process to produce an optimal solution that meets all of the nonlinear constraints of the model;

- by using a nonlinear programming (NLP) solver, instead of a linear one, to solve each node relaxation R_i. Indeed, if the problem is an MILP, R_i is an LP problem; if the problem is an MINLP, R_i is an NLP problem.

We also observe that there are significant choices that we did not discuss affecting the B&B process, such as the selection of the next branching node to be explored or the selection of the variable we want to branch on. More details about these strategies can be found in [3] for MILPs, and in [33] for MINLPs.

Furthermore, several variants of the B&B method have been provided after its introduction. One of the most famous is the *branch-and-cut* (B&C) algorithm, which is currently at

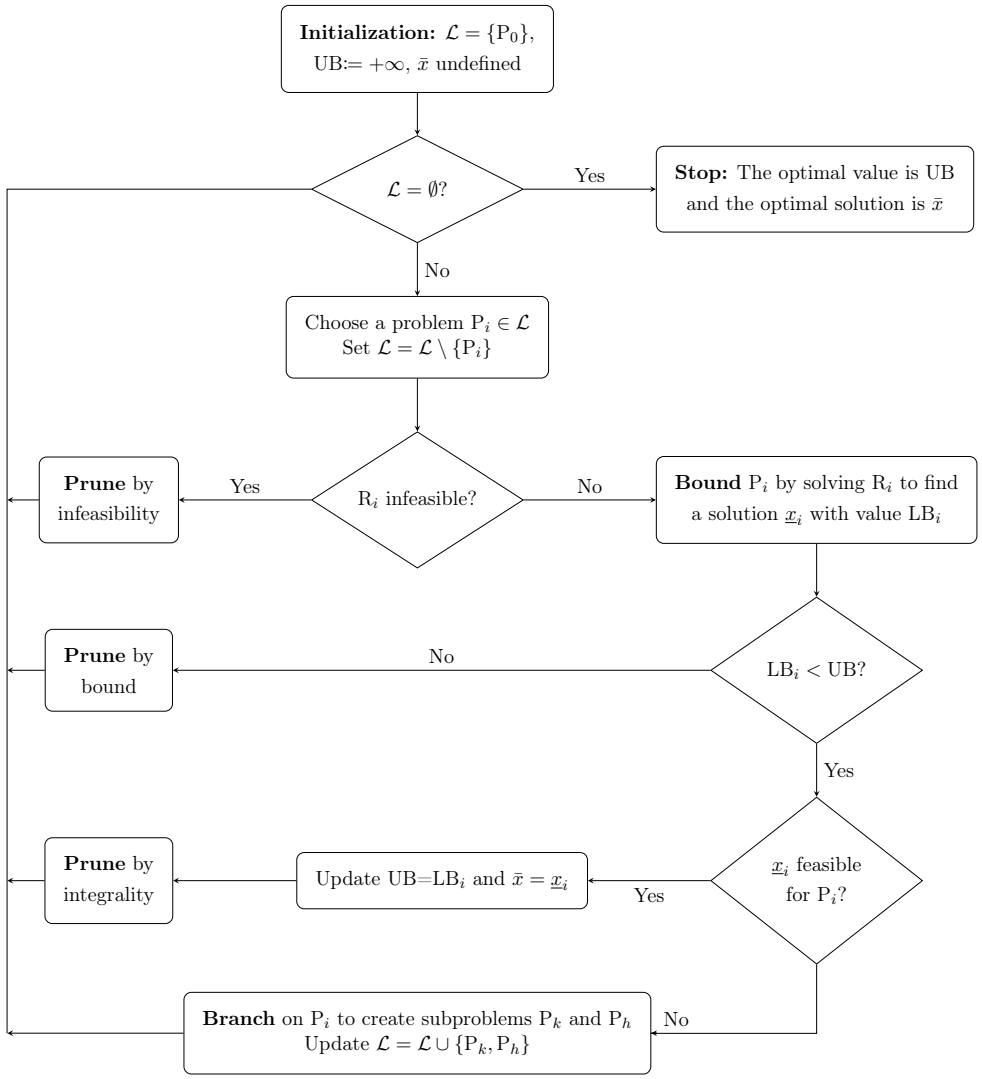

Figure 2.6: Main steps of the branch-and-bound algorithm for a minimization problem.

the basis of every state-of-art solver for MIP problems. Essentially, at each node i of the B&B tree, before the branching step, there is a cutting plane step in which one or more cutting planes may be dynamically generated to improve the optimal value of the relaxation R_i by cutting off its fractional optimal solution $\underline{x}_i$. Node i is then branched only if its relaxed optimal solution remains fractional even after a specified number of cut rounds or if no additional cuts can be generated. The decision to introduce new cuts is often made empirically, usually by assessing the effectiveness of previously introduced cuts. A strong cut strategy is typically applied at the root node, involving multiple rounds of cuts, while fewer or no cuts may be introduced at deeper levels within the branch-and-bound tree. Cuts added at the root node remain valid throughout the entire branching tree, whereas cuts generated within specific tree nodes are usually valid only for that particular sub-tree.

The B&C approach aims to reduce the tree size and decrease the need for branching. The underlying rationale for the B&C method comes from the observation that tightening

the lower bound is crucial for pruning the branching tree and, consequently, for accelerating the optimization process. A reference to B&C for MILPs is [17], and a reference to B&C for MINLPs is [134].

2.1.3 Benders decomposition

Benders decomposition (BD) was proposed in 1962 by Benders in [22] and can be considered a classic paradigm of mixed-integer programming. For surveys on this method, we refer the reader to [46, 120].

Consider the following MIP problem in the continuous variables x and the discrete variables y

$$\begin{aligned} \min \quad & c^T x + d^T y \\ \text{s.t.} \quad & Ax + By \geq b \\ & x \in \mathbb{R}^n_+ \\ & y \in Y \subset \mathbb{Z}^q \end{aligned} \tag{2.2}$$

where b is a vector in $\mathbb{R}^m$, c is a vector in $\mathbb{R}^n_+$, d is a vector in $\mathbb{R}^q$, A is a matrix in $\mathbb{R}^{m\times n}$, B is a matrix in $\mathbb{R}^{m\times q}$, and Y is a finite set describing restrictions on the discrete variables y. We assume that Problem (2.2) is feasible.

We observe that the objective function of the problem is *separable* in x and y and, once we fix the discrete y variables – considered the *complicating* variables as they are discrete – the problem simplifies to a *linear programming* problem in x. These properties are necessary to apply the BD and can be exploited to completely get rid of the continuous x variables, which is the idea behind the Benders reformulation (giving rise to the BD).

Indeed, the problem at hand can be reformulated as the bilevel problem

$$\min_{y\in Y}\{d^T y + \min_{x\geq 0}\{c^T x : Ax \geq b - By\}\}$$

where the outer optimization problem is called *master problem*, whereas the inner optimization problem is called *Benders subproblem*. We can reformulate again the master problem by replacing the inner problem with its value function $\phi(y)$

$$\min_{y\in Y}\{d^T y + \phi(y)\}$$

where $\phi(y)$ is defined as follows

$$\phi(y) = \min_{x \geq 0} \{c^T x : Ax \geq b - By\}.$$

Finally, by introducing the auxiliary variable w, the problem becomes

$$\begin{aligned} \min \quad & d^T y + w \\ \text{s.t.} \quad & w \geq \phi(y) \\ & w \geq 0 \\ & y \in Y. \end{aligned}$$

At this point it should be clearer that, since we can express the part belonging to the x variables as a function of the y values, and since we have a finite number of choices for y (Y is a finite set), we can reformulate the problem eliminating the x variables.

Now, let us focus on the Benders subproblem. We assume we are in the case where the subproblem is feasible and bounded (i.e. well-posed), in the sense that any given $\bar{y} \in Y$ is a feasible decision, and we derive its dual formulation

$$\phi(\bar{y}) = \max_{u \geq 0} \{(b - B\bar{y})^T u : A^T u \leq c\}.$$

We can derive the dual formulation of the subproblem as it is an LP, hence strong duality holds. Given the extreme points of the dual polyhedron $u_1, \ldots, u_P$, we have that the optimal dual value is attained at one of these extreme points

$$\phi(\bar{y}) = \max_{p=1,\ldots,P} \{(b - B\bar{y})^T u^p\} \tag{2.3}$$

as depicted in Figure 2.7. We observe that the polyhedron in the dual space ($A^T u \leq c$) does not depend on the choice of the y variables. This means that the solution (extreme point) does not depend on the y variables values, but its optimality still depends on the y variables values which appear in the dual objective.

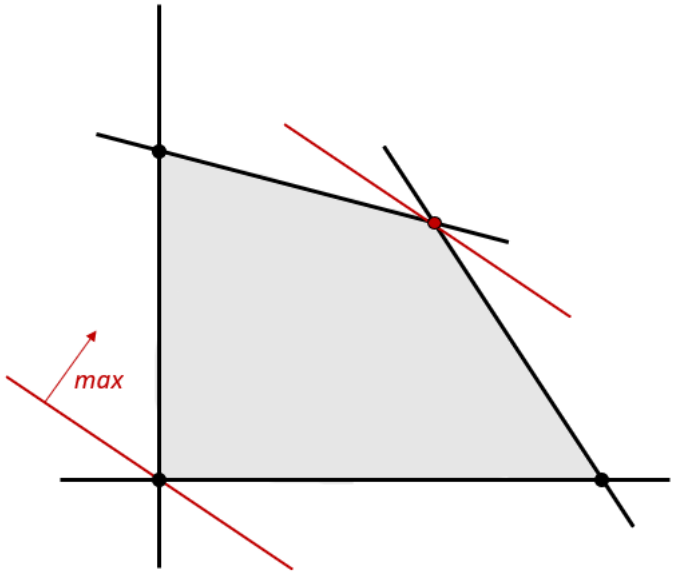

Figure 2.7: In a polyhedron the optimal solution is reached in (at least) one extreme point of the polyhedron. In this case, it is reached at the red point.

Using (2.3), we can reformulate the master problem as

$$\begin{aligned} \min \quad & d^T y + w \\ \text{s.t.} \quad & w \geq \max_{p=1,\dots,P} \{(b - By)^T u^p\} \\ & w \geq 0 \\ & y \in Y \end{aligned}$$

and since we impose w to be greater than or equal to the maximum (optimal) dual objective value, we also have that w is greater or equal to each dual objective value obtained at one extreme point

$$\begin{aligned} \min \quad & d^T y + w \\ \text{s.t.} \quad & w \geq (b - By)^T u^p \qquad p = 1, \dots, P \\ & w \geq 0 \\ & y \in Y. \end{aligned} \tag{2.4}$$

where constraints (2.4) are known as *Benders optimality cuts*.

If the subproblem is infeasible at a point $\bar{y}$, it means that in the dual space we are optimizing over an unbounded cone, and there exists an extreme ray that goes to infinity for the Farkas' Lemma. This extreme ray u^r gives the unbounded direction such that

$(b - B\bar{y})^T u^r > 0$ and hence the dual objective goes to infinity. To discard the point $\bar{y}$ at which the Benders subproblem is infeasible, we introduce the *Benders feasibility cuts* given by $(b - By)^T u^r \leq 0$ for each extreme ray u^r of the recession cone of the dual feasible region $\{u \geq 0 : A^T u \leq c\}$.

Finally, the so-called *Benders Reformulation* of problem (2.2) is given by

$$\begin{aligned} \min \quad & d^T y + w \\ \text{s.t.} \quad & w \geq (b - By)^T u^p \quad \forall u^p \in U_p \quad \text{(OptCut)} \\ & 0 \geq (b - By)^T u^r \quad \forall u^r \in U_r \quad \text{(FeasCut)} \\ & w \geq 0 \\ & y \in Y. \end{aligned} \tag{2.5}$$

where U_p set of all extreme points of the dual feasible region, and U_r set of all extreme rays of the dual feasible region. By means of this reformulation, we get rid of the x variables, but this comes at the expense of introducing a (possibly huge) number of additional constraints, i.e., the *Benders cuts*. In particular, the (OptCut) are the Benders *optimality* cuts, used to underestimate the value of the objective value of the subproblem, and (FeasCut) are the Benders *feasibility* cuts, eliminating master solutions that are infeasible for the (primal) subproblem.

The scheme in Figure 2.8 summarizes the Benders reformulation. Basically, we start from a problem in the variables x (continuous) and y (discrete), and we end up having a reformulation of this problem in y and w (continuous). This reformulation contains the Benders optimality and feasibility cuts generated from the solutions of the subproblem, obtained by fixing one by one the y variables.

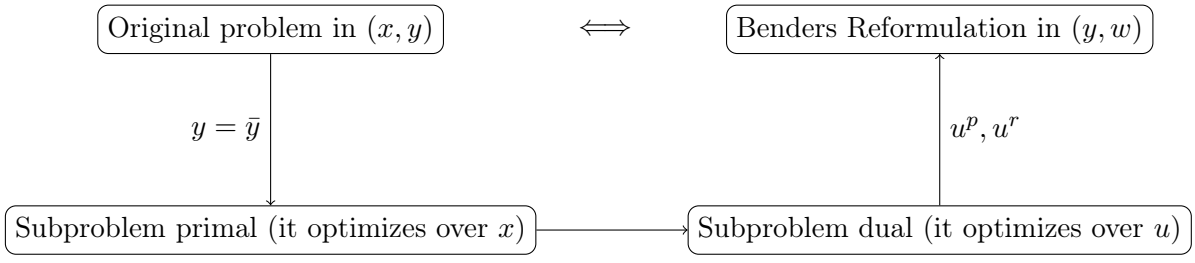

Figure 2.8: A scheme of the Benders Reformulation.

Since the number of Benders cuts is exponential (one for each extreme point and one for each extreme ray), instead of using the full Benders reformulation (2.5), it may be beneficial to use a relaxation of this problem containing only a subset of the Benders cuts that are separated dynamically through a cutting plane algorithm. This procedure is known as

Benders decomposition or *cutting plane Benders*, see a scheme in Figure 2.9.

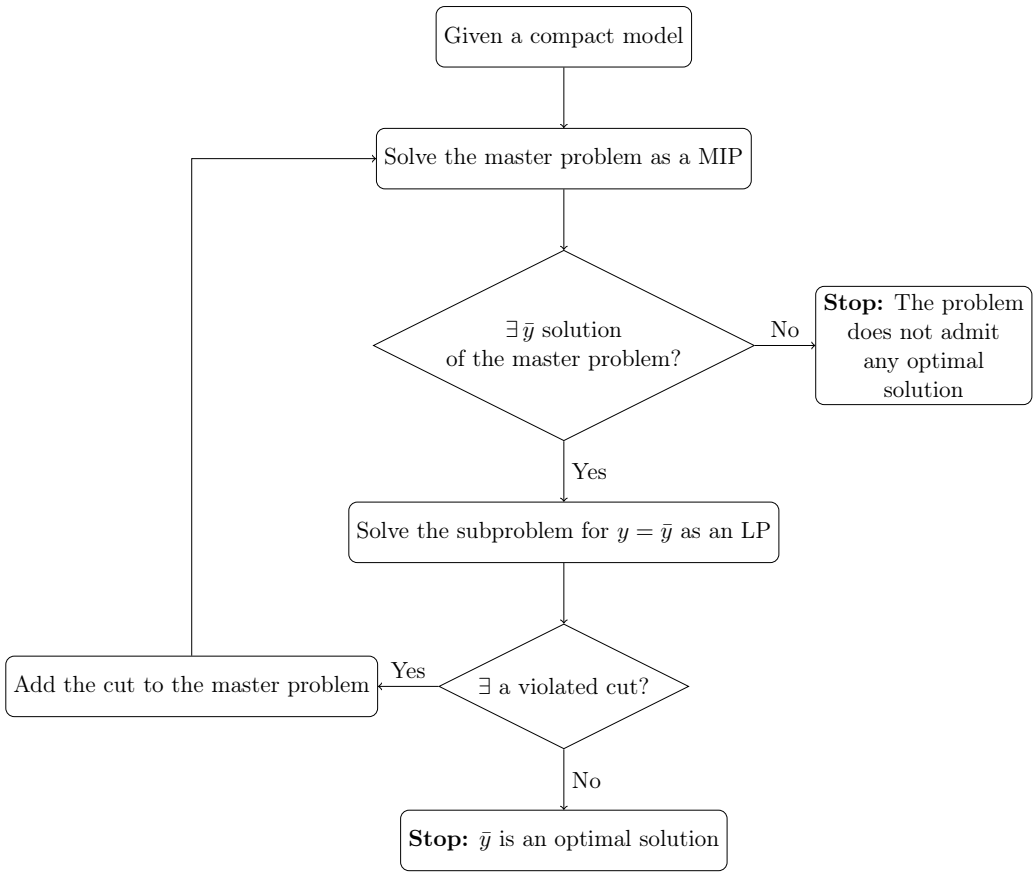

Figure 2.9: Scheme of cutting plane (or multi-tree) Benders.

Basically, this algorithm alternates between solving a relaxation of the master problem obtained by considering only a subset of the Benders cuts, and solving the LP subproblem for the given master solution to generate violated Benders optimality or feasibility cuts. One drawback of this "old school" approach is that each new cut requires the solution of a B&B tree as the relaxed master problem is an MIP, and for this reason it is also known as *multi-tree Benders*.

Modern approaches of Benders, instead, require the solution of only one branching tree as the violated Benders cuts are generated on the fly inside a branch-and-cut procedure, as schematized in Figure 2.10. For this reason, this approach is known as *branch-and-Benders-cut* or *single-tree Benders*.

We observe that this method can be applied both to MILPs and convex MINLPs in which the parameterized subproblem is an LP, as in the case we will discuss in Chapter 6. An extension of the Benders decomposition method to a broader class of MINLP problems in which the parameterized subproblem needs no longer be an LP, was provided by Geoffrion [82] in 1972, and it is known as *Generalized Benders decomposition*.

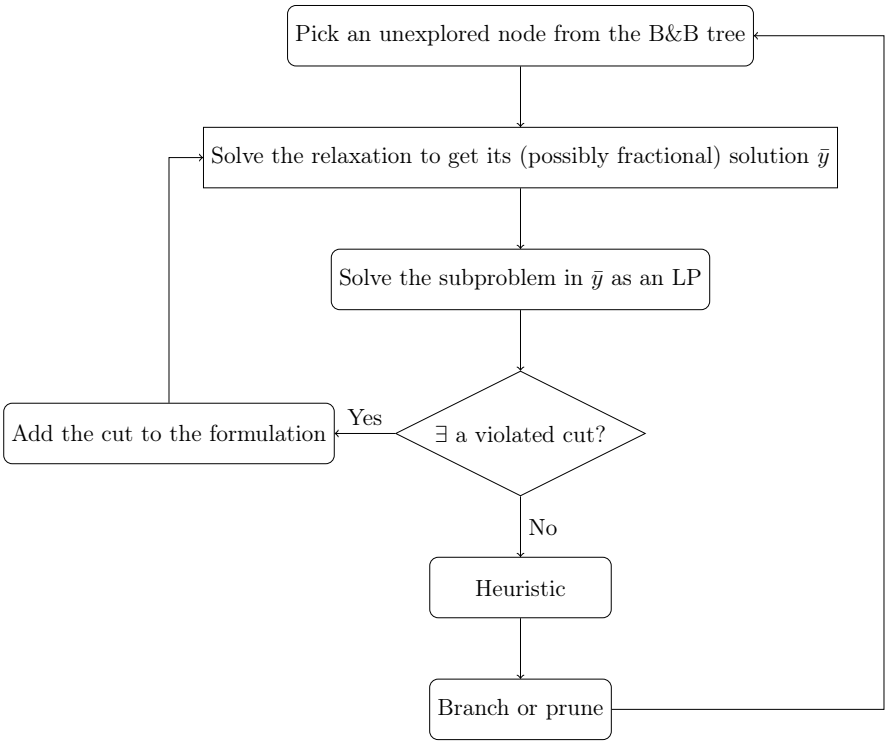

Figure 2.10: Scheme of branch-and-Benders-cut (or single-tree) Benders.

2.2 Optimization under uncertainty

The classical paradigm in mathematical programming is to assume that all the input data to the mathematical model, such as demands and costs, are assumed to be known precisely and equal to some nominal values. However, this assumption does not hold for most real world applications. In such applications, the classical deterministic optimization approach is not useful because the solution found through such an approach is possibly sensitive to even slight changes in the problem parameters. Indeed, deterministic optimization does not take into account the influence of *data uncertainties* on the quality and feasibility of the model: as the data take values different than the nominal ones, several constraints may be violated, and the optimal solution found using the nominal data may no longer be optimal or even feasible.

To illustrate the importance of robustness in practical applications, we quote from the case study by Ben-Tal and Nemirovski (2000) on linear optimization problems from the Net Lib library:

> In real-world applications of Linear Programming, one cannot ignore the possibility that a small uncertainty in the data can make the usual optimal solution completely meaningless from a practical viewpoint.

This observation raises the natural question of designing solution approaches that are

immune to data uncertainty. Both robust optimization and stochastic programming have been extensively studied to address this problem of uncertainty [27, 28, 31, 50, 92, 127], with uncertainty being modeled as explicit scenarios or implicitly specified via a probability distribution over the uncertain parameters.

Dantzig [50] introduced *stochastic* programming in the mid-1950s as an approach to model data uncertainty by assuming scenarios for the data occurring with different probabilities. The two main difficulties with such an approach are: (i) knowing the exact distribution for the data, and thus enumerating scenarios that capture this distribution is rarely satisfied in practice, and (ii) the size of the resulting optimization model increases drastically as a function of the number of scenarios, which poses substantial computational challenges.

A body of literature focused on *robust* approaches, in which we optimize against the worst instances that might arise by using a min-max objective. The first robust approaches proposed in [115, 133], before the the two major contributions by Bertsimas and Sim [27, 28], were considered over-conservative, as highlighted by [19–21, 64, 65]. Bertsimas and Sim in [27, 28] were the first to address over-conservatism of the robust optimization, proposing an approach leading to a *linear* optimization model. Their approach is known as Γ-*robustness*: it allows the control of the degree of conservatism of the solution, and it is computationally tractable both practically and theoretically, unlike the previously proposed non-linear approaches.

2.2.1 Γ-Robustness

We now discuss in detail the approach introduced by Bertsimas and Sim in 2003. The content of this section comes from their seminal papers [27, 28].

The idea at the basis of these papers is that when the data in the constraints of a linear programming problem are subject to uncertainty, we can model a robust optimization problem that optimizes against the worst-case scenario in which up to Γ data are going to deviate from their nominal values (where Γ is an input parameter). This robust problem can be modeled as a *linear programming* of moderately larger size than the size of the nominal problem. Moreover, the robust problem allows for control of the degree of conservatism of the solution in terms of probabilistic bounds of constraint violations.

Let c, l, u be vectors in $\mathbb{R}^n$, let A be an matrix in $\mathbb{R}^{m\times n}$, and b be a vector in $\mathbb{R}^m$. We consider the following nominal MIP on a set of n variables, the first k of which are integral

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & l \leq x \leq u \\ & x_i \in \mathbb{Z} \qquad i = 1, \ldots, k. \end{aligned} \tag{2.6}$$

We assume, without loss of generality, that data uncertainty affects only the elements of the matrix A, and not the vectors b and c. Indeed, if b is uncertain, then we can introduce a new variable x_{n+1}, and write $Ax - bx_{n+1} \leq 0$, $l \leq x \leq u$, $1 \leq x_{n+1} \leq 1$, thus augmenting A to include b; if, instead, c is uncertain, we can use the objective maximize z, add the constraint $z - c^T x \leq 0$, and thus include this constraint into $Ax \leq b$.

In typical applications, we have reasonable estimates for the mean value of the coefficients a_{ij} and its range $\hat{a}_{ij}$. We feel that it is unlikely that we know the exact distribution of these coefficients. Specifically, the model of data uncertainty we consider is as follows. Let $N = \{1, 2, \ldots, n\}$. Each entry a_{ij}, $j \in N$ is modeled as an independent, symmetric and bounded random variable (but with unknown distribution) $\tilde{a}_{ij}$, $j \in N$ that takes values in $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Note that we allow the possibility that $\hat{a}_{ij} = 0$. Note also that the only assumption that we place on the distribution of the coefficients a_{ij} is that it is symmetric.

For robustness purposes, for every i, we introduce a parameter Γ_i, not necessarily integer, that takes values in the interval $[0, |J_i|]$, where $J_i = \{j | \hat{a}_{ij} > 0\}$. The role of the parameter Γ_i in the constraints is to adjust the robustness of the proposed method against the level of conservatism of the solution. Consider the i-th constraint of the nominal problem $a_i^T x \leq b_i$. Let J_i be the set of coefficients a_{ij}, $j \in J_i$ that are subject to parameter uncertainty, i.e., $\tilde{a}_{ij}$, $j \in J_i$ independently takes values according to a symmetric distribution with mean equal to the nominal value a_{ij} in the interval $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Speaking intuitively, it is unlikely that all of the a_{ij}, $j \in J_i$ will change. Our goal is to be protected against all cases in which up to $\lfloor \Gamma_i \rfloor$ of these coefficients are allowed to change, and one coefficient a_{it} changes by at most $(\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it}$.

In other words, we stipulate that nature will be restricted in its behavior, in that only a subset of the coefficients will change in order to adversely affect the solution. We will then guarantee that if nature behaves like this then the robust solution will be feasible deterministically. We will also show that, essentially because the distributions we allow are symmetric, even if more than $\lfloor \Gamma_i \rfloor$ change, then the robust solution will be feasible with very high probability. Hence, we call Γ_i the protection level for the i-th constraint.

Specifically, we consider the following (still nonlinear) formulation of the *robust counterpart* of Problem (2.6)

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & \sum_j a_{ij}x_j + \max_{\{S_i \cup \{t_i\} |\, S_i \subseteq J_i, |S_i| \leq \lfloor \Gamma_i \rfloor\}, t_i \in J_i \setminus S_i\}} \{\sum_{j \in S_i} \hat{a}_{ij} y_j + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it_i} y_{t_i}\} \leq b_i \quad \forall i \\
& -y_j \leq x_j \leq y_j \qquad \forall j \\
& y_j \geq 0 \qquad \forall j \\
& l_j \leq x_j \leq u_j \qquad \forall j \\
& x_i \in \mathbb{Z} \qquad i = 1, \ldots, k.
\end{aligned}
\tag{2.7}
$$

Theorem 2.2.1 *Problem* (2.7) *has an equivalent MILP formulation as follows*

$$
\begin{aligned}
\min \quad & c^T x \\
s.t. \quad & \sum_j a_{ij}x_j + z_i\Gamma_i + \sum_{j \in J_i} p_{ij} \leq b_i \quad \forall i \\
& z_i + p_{ij} \geq \hat{a}_{ij} y_j && \forall i, j \in J_i \\
& -y_j \leq x_j \leq y_j && \forall j \\
& l_j \leq x_j \leq u_j && \forall j \\
& p_{ij} \geq 0 && \forall i, j \in J_i \\
& y_j \geq 0 && \forall j \\
& z_i \geq 0 && \forall i \\
& x_i \in \mathbb{Z} && i = 1, \ldots, k.
\end{aligned}
\tag{2.8}
$$

Proof We show how to model the constraints in (2.7) as linear constraints. Given a vector $x^\star$, we define the protection function:

$$
\beta_i(x^\star, \Gamma_i) = \max_{\{S_i \cup \{t_i\} |\, S_i \subseteq J_i, |S_i| \leq \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} |x_j^\star| + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it_i} |x_{t_i}^\star| \right\}.
$$

This equals to

$$\begin{aligned} \beta_i(x^\star, \Gamma_i) = \max \quad & \sum_{j \in J_i} \hat{a}_{ij} |x_j^\star| z_{ij} \\ \text{s.t.} \quad & \sum_{j \in J_i} z_{ij} \leq \Gamma_i \\ & 0 \leq z_{ij} \leq 1 \qquad \forall j \in J_i. \end{aligned} \tag{2.9}$$

Clearly the optimal solution value of Problem (2.9) consists of $\lfloor \Gamma_i \rfloor$ variables at 1 and one variable at $\Gamma_i - \lfloor \Gamma_i \rfloor$. This is equivalent to the selection of subset $\{S_i \cup \{t_i\} | \, S_i \subseteq J_i, |S_i| \leq \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}$ with corresponding cost function $\sum_{j \in S_i} \hat{a}_{ij} |x_j^\star| + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} |x_{t_i}^\star|$.

We next consider the dual of Problem (2.9):

$$\begin{aligned} \min \quad & \sum_{j \in J_i} p_{ij} + \Gamma_i z_i \\ \text{s.t.} \quad & z_i + p_{ij} \geq \hat{a}_{ij} |x_j^\star| \quad \forall i, j \in J_i \\ & p_{ij} \geq 0 \qquad\qquad\qquad \forall j \in J_i \\ & z_i \geq 0 \qquad\qquad\qquad\qquad \forall i. \end{aligned} \tag{2.10}$$

By strong duality, since Problem (2.9) is feasible and bounded for all $\Gamma_i \in [0, |J_i|]$, then the dual problem (2.10) is also feasible and bounded and their objective values coincide. We have that $\beta_i(x^\star, \Gamma_i)$ is equal to the objective function value of Problem (2.10). Substituting to Problem (2.7), we obtain that Problem (2.7) is equivalent to the linear optimization problem (2.8).

□

In other words, Bertsimas and Sim propose an MINLP approach addressing over-conservatism of the robust optimization, and then show how to reformulate their robust counterpart of the problem as a MILP, leading to a much more computationally tractable approach.

Chapter 3

The wireless network design problem

3.1 Introduction

Wireless network design (WND) is the problem of configuring a set of transmitters to provide service in a target area. The term configuring refers both to the optimal identification of the *locations* and some parameters of the transmitters, such as transmission *power* and/or *frequency*.

When designing a telecommunications network, the goal is to provide connection service to a portion of the territory, namely the *target area*. To test the connection on this area, the area is usually partitioned into elementary areas, in line with the recommendations of the telecommunications regulatory bodies. Each elementary area is called *testpoint*, and it is assumed to represent all users within the corresponding elementary area. We can imagine a testpoint as a sort of *superuser* or as a representative *receiver*. A simplified illustration of a target area and its testpoints is shown in Figure 3.1.

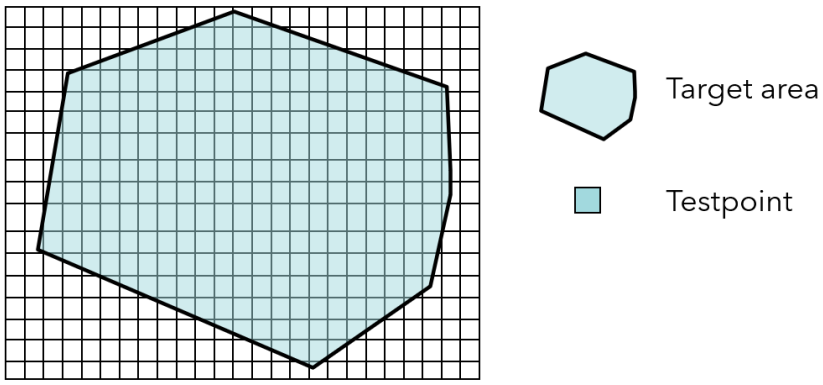

Figure 3.1: Target area with testpoints.

In each testpoint, we can measure the signals coming from all the transmitters that have been activated. The power received by each receiver is classified as *serving power* if

it relates to the power emitted by the transmitter serving that receiver; otherwise, it is classified as *interfering power*, according to the physical layer specifications of the LTE/5G standard (for LTE standard see [128]). This situation is depicted in Figure 3.2. The received power, both serving and interfering, can be considered proportional to the power emitted by a *fading coefficient* which models the reduction in power during the transmission, the so-called *path loss*. The fading coefficient can be evaluated through simulation tools that consider the orography and morphology of the territory, the receiver-transmitter distance, the frequency channel, the shape of the antenna, and other parameters (e.g., the average height of buildings, and the average width of streets).

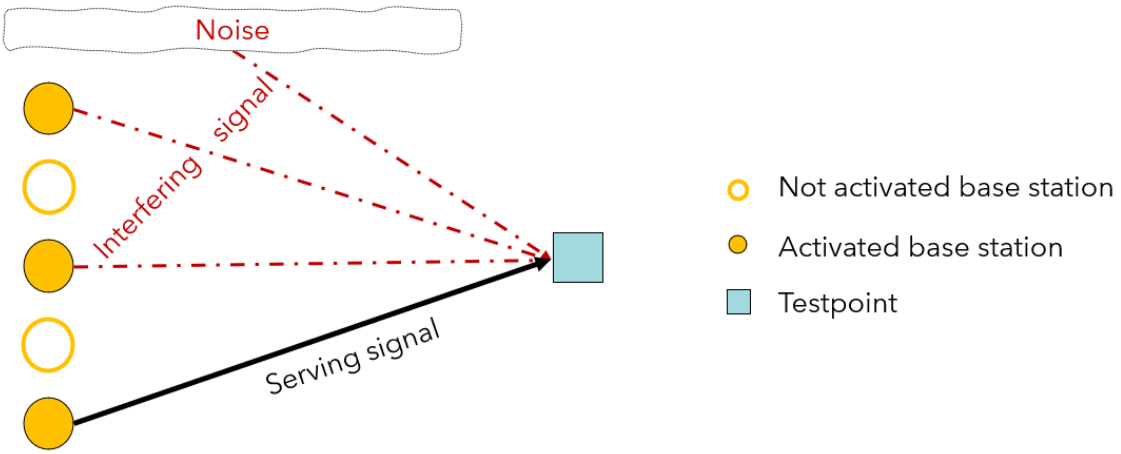

Figure 3.2: Signals received in a testpoint. The wanted signal is in black, the unwanted signals are in red.

As we previously mentioned, the aim is to guarantee service coverage in the designated area, and consequently, on the representative receivers of this area. A receiver is regarded as served by a base station if the ratio of the serving power to the sum of the interfering powers and noise power (*Signal-to-Interference-plus-Noise Ratio* or SINR) is above a threshold, whose value depends on the desired quality of service.

The more general wireless network design problem can take different names depending on the variables we aim to optimize when building the network. This thesis considers (i) the *site and the power assignment problem*, assuming the frequency as fixed whereas the emitted power as variable, and (ii) the *base station deployment problem*, assuming both emitted power and frequency as fixed. We will discuss the two cases separately.

For the site and power assignment problem we assume that frequency channel is given and equal for all transmitters. Having a fixed frequency is a straightforward assumption: for different frequency channels the problem decomposes as there is no interference among non-co-channel signals. For the base station deployment problem, we also assume that the

power emissions are given and equal for all transmitters.

3.2 Literature overview

As wireless networks are becoming denser, due to technological advancements and increased traffic [95], practitioners' traditional design approach based on trial-and-error supported by simulation has exhibited many limitations. The inefficiency of this approach leads to the need for *optimization approaches*, which are critical for lowering costs and meeting user-demanded service quality standards.

Many optimization models for WND have been investigated over the years; we recommend [97] for a thorough overview of the optimization challenges in modern wireless network design. However, the *natural formulation* on which most models are based presents *severe limitations* since it involves numerical problems in the problem-solving phase, which emerge even in small instances. Indeed, the constraint matrices of these models contain coefficients that range in a huge interval, as well as large big-M leading to weak bounds.

Several approaches have been proposed to solve the WND problem, both exact MIP-based [5, 6, 34, 37, 48, 49, 56, 60, 62, 117] and heuristic [41, 59, 61, 63, 80, 109, 110].
The *exact approaches* that have been proposed in the literature are mainly oriented towards non-compact formulations and row generation methods. The authors of [117] tackle the problem of optimal base station locations and power emissions, maximizing the service provider's profit. Their MILP model is solved via an exact solution method that combines combinatorial Benders decomposition, classical Benders decomposition, and valid cuts in a nested way. In [34], the authors aim to maximize mobile operator profits by maximizing coverage and minimizing costs. This paper takes its basis on the formulations presented in previous works [5, 6, 56] related to LTE-RAN applications. The formulations proposed in [5, 6, 34, 56] are solved using standard exact solvers, and tests have been done on randomly generated instances.
In [110], a two-stage heuristic algorithm is proposed to solve large instances of MILP network planning. In the first stage emission powers are assigned, while in the second stage frequencies are optimized. In [80], the minimization of the total number of deployed base stations is pursued. The model uses binary variables and is solved with a meta-heuristic approach based on swarm intelligence. Another heuristic contribution is [41], which considers the use of constraints on electromagnetic emissions and the distance from sensitive areas. The pursued objective is a weighted function of installation costs and

service coverage. The authors propose a heuristic based on a decomposition step, followed by a partial enumeration of the solutions. In [63], a planning methodology is introduced that minimizes interference while maintaining the coverage. The resulting non-convex model is tackled by a two-stage heuristic, consisting of a local search followed by an integer programming heuristic.

Some contributions investigate the aspect of *numerical problems* [37, 48, 49, 59–62, 109] that arise in the natural formulations of WND due to a constraints matrix with coefficients that greatly vary in their order of magnitude. The numerical issues make the problem intractable for real-world instances, which are typically large. Consequently, most exact MIP-based approaches that aim to overcome numerical drawbacks have been oriented towards non-compact formulations and row generation methods. In particular, a non-compact binary formulation is proposed in [49], in which both locations and power emissions of the antennas are optimized to maximize population coverage. The authors discretize the power emissions getting knapsack inequalities with generalized upper bounds (GUBs, defined in [140]), which are reformulated using GUB covers. They also investigate an exact algorithm consisting of row generation embedded in a branch-and-cut framework. In [62], the same authors of [49] introduce a 0-1 model for a variant of WND related to the feasible server assignment problem. In [48], numerically stable 0-1 formulations and a robust optimization model managing signal propagation uncertainty are investigated. In [37], the maximum link activation problem is considered, and a non-compact formulation is proposed that uses cover inequalities to replace the source of numerical instability. In [60], the source of numerical issues in WND is deeply investigated, and the use of numerically safe LP solvers is suggested to make the solutions reliable.
As for the *heuristic approaches* proposed to overcome the numerical instabilities, in [59], a MILP problem for the optimal allocation of power emissions and frequency channels is tackled through a genetic algorithm suitable for large-scale problems. The model maximizes the service coverage. In [109], a MILP model for the power assignment problem in wireless networks is employed, and the signal orientation of the antenna is also considered. A constructive heuristic followed by an improving local search is proposed. In [61], a matheuristic is investigated, which combines a genetic algorithm exploiting a suitable linear relaxation of the problem and an integer linear programming heuristic, improving the solutions found with the relaxation.

A summary classification of the cited works can be found in Table 3.1, where NI stands for numerical issues, indicating the contributions that address them. We refer the reader

to [97] for a complete view of optimization and numerical challenges in modern wireless network design.

Table 3.1: Scheme of contributions to wireless network design literature.

Paper	Decision Variables	NI	Method	Model	Technology
[117]	location, power emission	✗	exact	MILP	UMTS, W-CDMA
[34]	location, power emission, frequency channel	✗	exact	MILP	5G
[5, 6, 56]	location, power emission, frequency channel, service (except for [6]), period	✗	exact	MILP	LTE-RAN
[110]	power emission, frequency channel	✗	heuristic	MILP	DVB-T
[80]	location	✗	heuristic	0-1 LP	5G
[41]	location, frequency channel	✗	heuristic	MILP	5G
[63]	interference	✗	heuristic	MINLP	UMTS
[49]	location, power emission	✓	exact	0-1 LP	WiMAX, DVB-T
[48, 62]	location, power emission	✓	exact	0-1 LP	WiMAX
[37]	location	✓	exact	0-1 LP	-
[60]	location, power emission	✓	exact	MILP	WiMAX
[59]	power emission, frequency channel	✓	heuristic	MILP	WiMAX
[109]	power emission, horizontal orientation	✓	heuristic	MILP	DVB
[61]	power emission, frequency channel, transmission scheme	✓	heuristic	0-1 LP	DVB-T

DVB-T, Digital Video Broadcasting-Terrestrial; WiMAX, Worldwide Interoperability for Microwave Access; UMTS, Universal Mobile Telecommunications System; W-CDMA, Wideband Code Division Multiple Access; LTE-RAN, Long Term Evolution-Radio Access Network; 5G, 5^{th} Generation.

3.3 Mathematical formulation

In this section, we discuss the mathematical modeling of the two wireless network design problems we address. In particular, we present the mathematical formulation of the site and power assignment problem in Section 3.3.1, and the mathematical formulation of the base station deployment problem in Section 3.3.2.

3.3.1 The variable-power case

The site and power assignment problem considers the optimization of both locations and power emissions of the base stations. Locations are typically modeled through discrete choices, whereas power emissions have been initially modeled through continuous variables, leading to classical MIP models. Actually, the use of continuous decision variables contrasts with the standard network planning practice of considering a small number of discrete power values. Indeed, the design specifications of real-life antennas are always expressed as rational numbers with bounded precision and, consequently, assume a finite number of values. Motivated by this telecommunications practice, the authors of [49], introduced

the practice of *power discretization* for modeling purposes. We will therefore refer to the formulation introduced in [49] – which considers the discretization of the power – as the *natural formulation* of the problem.

Let $\mathcal{B}$ be the finite set of potential transmitters and $\mathcal{T}$ be the finite set of receivers located at the testpoints. Let $\mathcal{P} = \{P_1, \ldots, P_{|\mathcal{P}|}\}$ be the finite set of feasible power values assumed by the activated transmitters, with $P_1 > 0$ and $P_{|\mathcal{P}|} = P_{max}$. Hence $\mathcal{L} = \{1, \ldots, |\mathcal{P}|\}$ is the finite set of power value indices (or simply power levels). We introduce the variables

$$z_{bl} = \begin{cases} 1 & \text{if transmitter } b \text{ is emitting at power } P_l \\ 0 & \text{otherwise} \end{cases} \qquad b \in \mathcal{B},\, l \in \mathcal{L}$$

and

$$x_{tb} = \begin{cases} 1 & \text{if testpoint } t \text{ is served by transmitter } b \\ 0 & \text{otherwise.} \end{cases} \qquad b \in \mathcal{B},\, t \in \mathcal{T}$$

To enforce the choice of only one (strictly positive) power level for each activated transmitter we use

$$\sum_{l \in \mathcal{L}} z_{bl} \leq 1 \qquad b \in \mathcal{B}. \tag{3.1}$$

The mathematical formulation of the WND problem contains the so-called *SINR inequalities* used to assess service coverage conditions. We recall that a testpoint $t \in \mathcal{T}$ is regarded as *covered* by a base station $\beta \in \mathcal{B}$ if the $\text{SINR}_{t\beta}$ is above a threshold. The $\text{SINR}_{t\beta}$ is given by the ratio of the serving power coming from β and the sum of the interfering powers and noise power measured in t. The serving power is given by the product of the power emitted by the serving base station β and the fading coefficient, modeling the reduction of power in the signal from β to t. The interfering power is given by the sum of all interfering contributions that are measured at the testpoint t. These interfering contributions are modeled as the serving power, with the unique difference that the signal is emitted by all the base stations that are activated, except for the serving one (i.e., except for β).

The SINR is a good interference indicator since it can be used to define the effect of interference on communication. Measuring the interference is important as it affects the amount of information that can be transmitted [11]. Indeed, the Shannon capacity theorem, which is at the basis of telecommunication systems, states that the maximum data rate at which information can be reliably transmitted over a communication channel is subject to

the amount of noise and interference present in the channel as follows

$$C_{t\beta} = B\, log(1 + \text{SINR}_{t\beta})$$

where $C_{t\beta}$ is the channel capacity (or maximum data rate) and B is the channel bandwidth. Therefore, guaranteeing minimum SINR levels also means guaranteeing minimum data rate levels.

To formulate the SINR inequalities, we refer to the discrete big-M formulation reported in [49], which considers a discretization of the power range. Let $\tilde{a}_{tb} > 0$ be the fading coefficient applied to the signal received in $t \in \mathcal{T}$ and emitted by $b \in \mathcal{B}$. Let $\mu > 0$ be the system noise. Then a receiver t is served by a base station $\beta \in \mathcal{B}$ if the $\text{SINR}_{t\beta}$ is above a given SINR threshold $\delta > 0$, namely

$$\frac{\tilde{a}_{t\beta} \sum_{l \in \mathcal{L}} P_l z_{\beta l}}{\mu + \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb} \sum_{l \in \mathcal{L}} P_l z_{bl}} \geq \delta \qquad t \in \mathcal{T}, \beta \in \mathcal{B} : \; x_{t\beta} = 1 \tag{3.2}$$

where the numerator represents the serving signal in t (coming from β), and the denominator is the sum of the noise and the interfering signals in t (coming from all $b \neq \beta$). Following [49], we can rewrite the SINR condition (3.2) through the big-M constraints

$$\tilde{a}_{t\beta} \sum_{l \in \mathcal{L}} P_l z_{\beta l} - \delta \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb} \sum_{l \in \mathcal{L}} P_l z_{bl} \geq \delta\mu - M_{t\beta}(1 - x_{t\beta}) \qquad t \in \mathcal{T}, \beta \in \mathcal{B} \tag{3.3}$$

where $M_{t\beta}$ is a large (strictly) positive constant. When $x_{t\beta} = 1$, (3.3) reduces to (3.2); when $x_{t\beta} = 0$ and $M_{t\beta}$ is sufficiently large, (3.3) becomes redundant. We can set, e.g.,

$$M_{t\beta} = \delta\mu + \delta P_{max} \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{a}_{tb}. \tag{3.4}$$

Note that we can claim that a testpoint $t \in \mathcal{T}$ is covered if and only if there exists at least one (t, β) with $\beta \in \mathcal{B}$ that can satisfy (3.3) with $x_{t\beta} = 1$.

A constraint to express a minimum service coverage of the territory is included. We assume that each testpoint weights $r_t \in \mathbb{R}$ to account for the fact that testpoints can represent a different number of users or be more or less crucial in service coverage. A minimum coverage level $r \in \mathbb{R}$ of the testpoint is enforced by the constraint:

$$\sum_{t \in \mathcal{T}} r_t \sum_{b \in \mathcal{B}} x_{tb} \geq r. \tag{3.5}$$

When r_t are all equal to 1, r represents the minimum number of testpoint to be covered, and the constraint above corresponds to a territorial coverage. When $r_t \in [0,1]$ can be interpreted as the fractions of the population living in the elementary area $t \in \mathcal{T}$, then $r \in [0,1]$ represents the fraction of the population to be covered by the service and the constraint represents a population coverage.

Each testpoint must be covered by at most one serving base station, namely

$$\sum_{b \in \mathcal{B}} x_{tb} \leq 1 \qquad t \in \mathcal{T}. \tag{3.6}$$

The objective functions proposed in the literature of WND are several, going from the maximization of the coverage to the maximization of the quality of service. According to the FUB, a considerable goal to pursue nowadays is the citizens' welfare; therefore, the model we refer to aims at identifying solutions with low environmental impact in terms of electromagnetic pollution and/or power consumption. Reducing electromagnetic pollution indeed involves reducing the power emitted by the transmitters [40]. Hence, we aim to minimize the total number of activated base stations with a penalization on the use of stronger power levels: the cost associated with the use of a power level equal to $l \in \mathcal{L}$, namely c_l, is greater the greater is P_l.

Thus a natural formulation of the WND is the following 0-1 LP model:

$$\begin{aligned} \min_{x,z} \quad & \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} c_l z_{bl} \\ & (x,z) \in S \end{aligned} \tag{3.7}$$

where the feasible region S is defined as

$$S = \left\{ (x,z) \in \{0,1\}^{n+m} : \text{ satisfying } (3.1), (3.3), (3.5), (3.6) \right\}$$

with $x = (x_{tb})_{t \in \mathcal{T}, b \in \mathcal{B}}$, $z = (z_{bl})_{b \in \mathcal{B}, l \in \mathcal{L}}$ and $n = |\mathcal{T}| \times |\mathcal{B}|, m = |\mathcal{B}| \times |\mathcal{L}|$.

3.3.2 The fixed-power case

The base station deployment problem considers the optimization of base station locations. Let $\mathcal{B}$ be the finite set of potential transmitters and $\mathcal{T}$ be the finite set of receivers located

at the testpoints. We introduce the variables

$$z_b = \begin{cases} 1 & \text{if transmitter } b \text{ is activated} \\ 0 & \text{otherwise} \end{cases} \qquad b \in \mathcal{B}$$

and

$$x_{tb} = \begin{cases} 1 & \text{if testpoint } t \text{ is served by transmitter } b \\ 0 & \text{otherwise} \end{cases} \qquad b \in \mathcal{B},\ t \in \mathcal{T}.$$

We denote by a_{tb} the power value measured in $t \in \mathcal{T}$ of the signal emitted by $b \in \mathcal{B}$. Given a transmitter b and a receiver t, the received power in t, denoted as a_{tb}, is given by the product of the power emitted by b and the fading coefficient $\tilde{a}_{tb}$.
The service conditions can be modeled through the SINR conditions

$$\frac{a_{t\beta} z_\beta}{\mu + \sum_{b \in \mathcal{B}\setminus\{\beta\}} a_{tb} z_b} \geq \delta \qquad t \in \mathcal{T}, \beta \in \mathcal{B} :\ x_{t\beta} = 1$$

that can be rewritten as the following big-M constraints

$$a_{t\beta} z_\beta - \delta \sum_{b \in \mathcal{B}\setminus\{\beta\}} a_{tb} z_b \geq \delta\mu - M_{t\beta}(1 - x_{t\beta}) \qquad t \in \mathcal{T}, \beta \in \mathcal{B} \tag{3.8}$$

in which we can set, e.g., $M_{t\beta} = \delta\mu + \delta \sum_{b \in \mathcal{B}\setminus\{\beta\}} a_{tb}$.

Under the service and coverage constraints, we can minimize the number of activated base stations. Thus a natural formulation is the following 0-1 LP model:

$$\begin{aligned} \min_{x,z} \quad & \sum_{b \in \mathcal{B}} z_b \\ & (x, z) \in S \end{aligned} \tag{3.9}$$

where S is the feasible region defined as

$$S = \left\{(x, z) \in \{0,1\}^{n+m} :\ \text{satisfying } (3.5), (3.6), (3.8)\right\}$$

with $x = (x_{tb})_{t \in \mathcal{T}, b \in \mathcal{B}}$, $z = (z_b)_{b \in \mathcal{B}}$ and $n = |\mathcal{T}| \times |\mathcal{B}|, m = |\mathcal{B}|$.

3.4 Practical issues

In principle, MIP solvers can solve models (3.7) or (3.9). However, it is well-known that numerical and memory issues arise when solving practical instances, as widely described in [49, 60]. The main issues encountered are:

- the power received in each testpoint ranges in a large interval, from small values (order of 10^{-7}) to huge (10^5), which makes the range of coefficients $\tilde{a}_{tb}$ in the constraints matrix large and the solution process numerically unstable and possibly affected by error;
- the big-M coefficients lead to poor quality bounds that impact the effectiveness of standard solution procedures;
- real-life instances of these problems lead to models with a large number of variables and constraints.

As a result, practical WND problems are hard to solve using standard optimal procedures and/or natural formulations. Therefore, the objective of our contribution, given in Section 5, aims to render the optimal solution of realistic instances of the WND problem more practicable.

Chapter 4

Variants of the facility location problem

This chapter delves into two specific variants of the facility location problem: congested facility location and partial set covering location. These combinatorial problems serve as the foundation for the problem discussed in Chapter 6.

4.1 The congested facility location problem

Over the last century, *facility location* has emerged as a critical area of study, attracting the interest of academics and industry professionals because of its wide-ranging applications. For an introduction to location models, the reader can refer to [125], for a survey on this problem to [138], whereas for a comprehensive bibliography of modern research in the main location problem types to [126].

In this chapter, we introduce a variant of facility location problems (FLP), which is the quadratic multiple allocation facility location problem, commonly referred to as the *congested facility location problem* (CFLP). In this problem, we are given a set of potential facility locations and a set of customers. The goal is to meet customer demand through open facilities. The demand can be fractionally met by multiple facilities, allowing for multiple allocation. Multiple allocations are particularly relevant in contexts where the customer represents the population of a specific area, and individuals within the same area may be served by different facilities, such as in a telecommunications context. In the standard FLP, having a linear objective function, the cost to be minimized comprises the sum of facility opening cost and customer allocation cost. Instead, the aim of the congested variant of FLP is to attain a more balanced solution, also taking into account the minimization of

congestion to prevent facility *overload*. The CFLP incorporates the complexity of possible congestions at open facilities to express the *diseconomies* of scale or *penalization* in the quality of service produced by the congestion. Congestion costs can arise due to a range of factors, such as the employment of additional overtime workers, the use of more expensive materials, or neglecting or postponing equipment maintenance schedules [93]. The quality of service is instead influenced by congestion for what concerns, e.g., waiting times [10, 66] or user data rates, depending on the application at hand. Consequently, the goal of CFLP is to minimize the sum of facility opening cost, customer allocation cost, and *congestion cost*. The latter is usually characterized in terms of service or production costs at facilities or customer waiting times. Among the different ways of modeling congestion cost, we refer to the one using *convex quadratic functions*.

This problem can be seen as a telecommunication problem in which there is the need to place transmitters (the facilities), minimizing the cost of the infrastructure and the congestion at the transmitters. Congestion in telecommunications networks may result from a sudden increase in the demand of users allocated to a specific transmitter. Addressing congestion is crucial, as it directly impacts the performance and efficiency of the networks. Indeed, elevated demand for a single transmitter can compromise the quality of service, leading to reduced data rates for the users connected to that particular transmitter. Moreover, taking into account the minimization of the congestion has an effect which is similar to imposing a limited *capacity* on the transmitters. Indeed, as stated in [54], CFLP could be considered a reformulation of the facility location problem with limited capacities on the facilities, as the congestion function in the objective function can be seen as a *penalty function* associated with capacity constraints, that penalizes every additional unit being served by a given facility.

4.1.1 Literature overview

Despite its practical applicability and the theoretical challenges it offers due to its mixed-integer nonlinear structure, the literature gave CFLP relatively little attention. The problem was introduced in 1995 by Desrochers et al. [54], inspired by [98], in which a similar formulation was proposed to illustrate a brief example involving skiers waiting for chairlifts. The seminal paper [54] provides a column generation method embedded in a branch-and-bound scheme. A comparison between different MIP formulations for the case of convex and piece-wise linear production cost functions is made in [93]. Other contributions include the two master theses [107] and [129] investigating the problem, and the article [68],

where the authors propose a Benders decomposition method that is effective even though the subproblem is not separable. A recent study on how off-the-shelf MIP solvers behave on CFLP instances modeled throw mixed-integer second-order cone programs can be found in [16]. Finally, [42] suggests a branch-and-bound algorithm based on Lagrangian relaxation and subgradient optimization.

The CFLP we refer to in this chapter relies on convex quadratic functions to model congestion costs. However, there are several other studies modeling facility congestion using queuing theory, or taking into account different objective functions, that we do not intend to revise in this thesis. We refer the reader to the survey [32] and the literature overviews provided in [1, 141] for further information.

4.1.2 Mathematical formulation

We are given a finite set J of customers, each characterized by a demand $d_j \geq 0$ for $j \in J$, and a finite set I of potential facility locations. Opening a facility $i \in I$ has a cost $f_i \geq 0$, whereas serving a unit of demand of customer $j \in J$ by facility $i \in I$ has a cost $c_{ij} \geq 0$. Congestion at a facility $i \in I$ is measured by the load of i, namely the overall amount of demand served by facility i, considering all the customers served by i.

The CFLP consists of determining the facility locations to open and the fraction of demand served by the selected facilities for each customer in order to minimize the overall cost obtained as the sum of facility opening, customer allocation, and congestion costs. Let $y_i \in \{0,1\}$ denote the closing/opening of the facility location $i \in I$ and $x_{ij} \geq 0$ for each $i \in I, j \in J$ the fraction of the demand of customer j served by facility i when i is open. Given the penalty function $F(\cdot)$, which is assumed to be non-negative, continuous and convex for non-negative arguments, congestion cost are expressed through function $F(\cdot)$ and penalize the load of the facilities.

Then, the CFLP can be modeled as the following MINLP

$$\min \quad \sum_{i\in I} f_i y_i + \sum_{i\in I}\sum_{j\in J} d_j c_{ij} x_{ij} + \sum_{i\in I} F\left(\sum_{k\in J} d_k x_{ik}\right) \sum_{j\in J} d_j x_{ij} \tag{4.1}$$

$$\sum_{i\in I} x_{ij} = 1 \qquad j \in J \tag{4.2}$$

$$x_{ij} \leq y_i \qquad i \in I, j \in J \tag{4.3}$$

$$x_{ij} \geq 0 \qquad i \in I, j \in J$$

$$y_i \in \{0,1\} \qquad i \in I$$

where constraints (4.2) impose full customer demand satisfaction, whereas constraints (4.3) guarantee that allocation to a facility i is only possible if it is open.

We introduce the aggregated variables v_i to model the load of facility i, namely $v_i = \sum_{j \in J} d_j x_{ij}$. Then, the congestion cost at a facility $i \in I$ can be expressed through $F(v_i)v_i$. If we assume that $F(t)$ is a linear function, i.e., $F(t) = at + b$, with a and b non-negative input coefficients, we can write the objective function (4.1) as

$$\sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} d_j c_{ij} x_{ij} + b \sum_{i \in I} v_i + a \sum_{i \in I} v_i^2.$$

Under the assumptions made on $F(\cdot)$, the objective function remains convex in (x, v, y) for all non-negative values of x, as shown in [54], hence the model is a convex MINLP, and in particular it is a mixed-integer convex quadratic program with indicator variables. Consequently, we can derive its *perspective reformulation* (see Chapter 2.1.1 for details) in which the quadratic term in the objective function representing the congestion, v_i^2, is replaced by non-negative (additional) variables u_i so as to move the non-linearity from the objective to the constraints. The perspective reformulation of the CFLP is the following mixed-integer conic programming

$$\begin{aligned}
\min_{x,y,v,z} \quad & \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} d_j c_{ij} x_{ij} + b \sum_{i \in I} v_i + a \sum_{i \in I} u_i & & \\
& v_i^2 \leq u_i y_i & i \in I & \quad (4.4) \\
& v_i = \sum_{j \in j} d_j x_{ij} & i \in I & \\
& \sum_{i \in I} x_{ij} = 1 & j \in J & \\
& x_{ij} \leq y_i & i \in I, j \in J & \\
& x_{ij} \geq 0 & i \in I, j \in J & \\
& z_i \geq 0 & i \in I & \\
& y_i \in \{0, 1\} & i \in I &
\end{aligned}$$

where the rotated second-order cone constraints (4.4), together with the objective function's minimization, guarantee the quadratic load v_i^2 of a facility i is zero if the facility is closed, and it is u_i if the facility i is open. The use of perspective reformulation on such problems has been demonstrated to provide a considerably stronger continuous relaxation than the natural problem formulation (see [85, 86] for details).

4.2 The partial set covering location problem

Covering location problems constitute an important family of facility location problems, incorporating a notion of coverage. In these problems, *coverage* is defined by a proximity measure – often distance or travel time – establishing whether a facility can serve a demand point. Indeed, a demand point is said to be covered by a facility if it lies within the coverage radius of this facility.

Applications of covering location problems can be found in several sectors, e.g., in the service sectors and emergency facilities for the ambulances [4], the healthcare facilities [131], and the fire detection [55]; but also in wireless telecommunication [36], or in monitoring equipment [102], and bank branch [9].

Among the notable problems of this class, there is the *set covering location problem* (SCLP), which aims to locate a minimum-cost set of facilities such that all demand points are covered at least once. However, the limitation of SCLP lies in the fact that it arbitrarily values each demand point, without taking into account factors like its position and size. This often leads to costly or impractical solutions. To mitigate this, the *partial set covering location problem* (PSCLP) and the *maximal covering location problem* (MCLP) were introduced. As opposed to the traditional SCLP, the PSCLP aims to minimize the cost of establishing facilities while ensuring a specified (*partial*) amount of the total demand is covered, whereas MCLP looks for the subset of facilities maximizing the coverage while respecting a budget constraint.

The following example is designed to illustrate why PSCLP is a perfect fit for telecommunications network design. Suppose there is a telecommunications company that needs to design a network of cellular transmitters to provide coverage for an area. The company has identified various candidate locations for installing these transmitters, each having different installation costs. The area can be partitioned into smaller blocks, each representing a demand point that requires cellular coverage.
In a perfect world, each of these demand points would be fully covered by at least one transmitter. However, doing so might incur prohibitive costs and could be unnecessary for regions with low population density or less strategic importance. In such cases, it may be acceptable to only partially cover the area. The goal of PSCLP is then to choose the best locations for the transmitters in a way that minimizes the total cost while ensuring that a target percentage of the total demand is covered. A graphical representation of the PSCLP applied to telecommunications is shown in Figure 4.1, in which we see that only a subset of transmitter locations has been chosen, the customer demand is only partially covered, there

exists a coverage radius for each transmitter and the customers served by the transmitter are inside this radius.

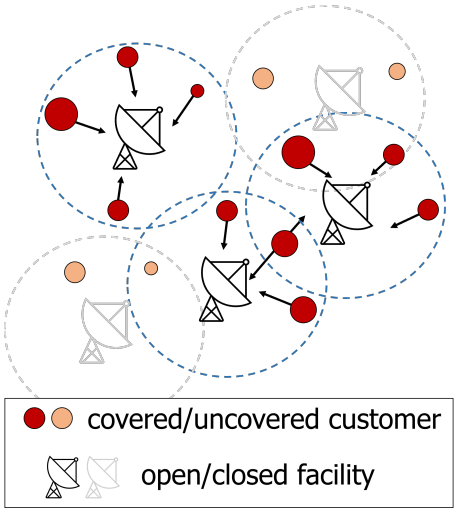

Figure 4.1: Representation of the partial set covering location problem.

4.2.1 Literature overview

In this section, we survey the literature on PSCLP and MCLP. Minimum cost covering problems trace back to Hakimi's work [91] in 1965, where he introduced the problem of locating the minimum number of police officers so that everyone is within a given distance from an officer; Hakimi suggests a solution procedure based on enumeration. The first integer programming formulation of the problem was proposed in [136] to solve the problem of locating emergency service facilities in a discrete space. Then, in [139], a heuristic is proposed to assign ladder trucks to fire stations, and the problem is formulated as a minimum cost covering.

Many authors [123, 132, 136] observed that the solution of the LP relaxation of MIP formulations of the SCLP often has few fractional variables and provides a very good lower bound. Despite this, there are few exact algorithms for solving MCLP and PSCLP.

PSCLP specifically has not received extensive study after its introduction in 1999 by Daskin and Owen [51], who employed a Lagrangian heuristic in their approach. In 2019, an exact approach based on Benders decomposition has been proposed in [45] (for both PSCLP and MCLP), and several large-scale instances were made available by the authors as a benchmark for future works; [45] also makes a complete description of the problem and list the contributions made in the literature, on which this chapter is mainly based. More recently in [39], five customized presolving methods have been discussed to enhance the capability of employing MIP solvers in solving PSCLP.

Other studies consider partial set covering problems in different contexts than location, such as a mining application in [30], and a related problem in which customers whose distance falls between a lower and an upper bound from their nearest facility are only partially covered [26].

MCLP received more attention than PSCLP, especially for what concerns heuristic and metaheuristic algorithms. An overview of the use, application, solution, and extension of MCLP can be found in [116]. The MCLP was formulated for the first time by Church and ReVelle [43] in 1974 as a 0-1 linear programming and in [114] was proved to be NP-hard. There are two versions of this problem in the literature, one imposing an upper bound on the number of open facilities – whose LP relaxation could be integral for relatively small size instances (see [132])– and the other using a budget constraint – whose LP relaxation usually leads to more fractional solutions.

For what concerns exact algorithms, Lagrangian relaxations are employed in [57] to dualize the covering constraints. The Lagrangian dual is then solved using subgradient optimization. This approach is integrated into a branch-and-bound framework.

Several heuristics have been proposed for the MCLP. Two greedy procedures that iteratively choose the facility that increases the coverage the most, one allowing substitution, are discussed in [43]. Another greedy procedure based on an adaptive search strategy is provided in [121]. Several heuristic approaches based on the Lagrangian relaxation of the constraints are suggested in [57, 78, 79, 106]. A decomposition heuristic employing partial relaxation of the covering constraints, resulting in stronger bounds than LP relaxation, is presented in [130]. Among the metaheuristics, we have [124] investigating an approach based on heuristic concentration which reduces the problem size by discarding potential facility locations that are unlikely to appear in an optimal solution. A genetic algorithm to solve even large instances of MCLP is proposed in [13]. In [113], the authors discuss an approach based on an adaptive memory to guide the search towards promising regions. An improved version of the algorithm proposed in [113] is given in [112]. More recently, a metaheuristic paradigm based on an adaptive iterated local search has been provided in [111] to address large instances of the MCLP. Another metaheuristic based on the harmony search algorithm is presented in [12].

For studies investigating the MCLP under data uncertainty, see e.g., [25, 44, 52, 104].

Finally, we refer the interested reader to the survey chapters [81] and [132] offering reviews on covering problems, and to [67] for a classification of many variants of covering location problems. The book [100] instead gives a general overview of location problems.

4.2.2 Mathematical formulation

We are given a set of potential facility locations I with opening costs $f_i \geq 0$, $i \in I$, and a set of customer locations J such that each customer $j \in J$ is associated with a demand $d_j \geq 0$. For each customer j, we are also given a subset $I(j) \subseteq I$ of facility locations that can cover j, i.e., that can fully serve the demand d_j. Similarly, for a subset of facilities $N \subseteq I$, let $J(N) \subseteq J$ be the subset of customers that can be covered by N, and let $J(i) = J(\{i\})$, for $i \in I$. In Figure 4.2, we provide an example of an instance of the problem having 4 potential facilities (represented by the white vertices) and 7 customers (represented by the grey vertices). The incident edges to a customer vertex j are linked to the facilities covering this customer, i.e., to the set $I(j)$; while the incident edges to a facility vertex i represent the set of customers covered by the facility, i.e., the set $J(i)$.

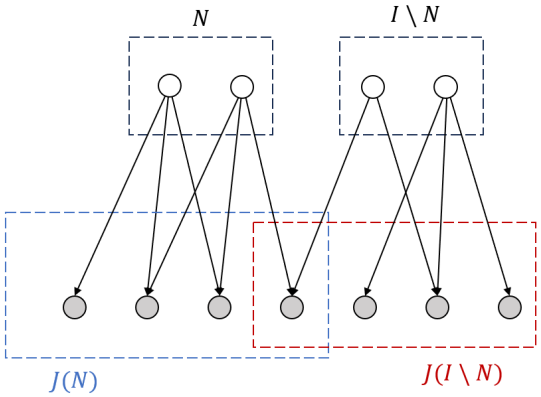

Figure 4.2: Example of an instance of PSCLP with 4 potential facilities (the white vertices) and 7 customers (the grey vertices).

Given a parameter $D > 0$, the PSCLP consists of choosing a subset of facilities to open so as to make sure that the covered customer demand is at least D and the opening cost of the facilities is minimized.

To formulate the problem, we introduce the binary variables y_i for every potential facility location $i \in I$, such that y_i takes value 1 if facility i is open. We also introduce the binary variables z_j for every customer $j \in J$, such that z_j takes value 1 if customer j is covered by at least one open facility.

The problem can then be formulated as the following integer linear programming model

$$\min \quad \sum_{i \in I} f_i y_i \tag{4.5}$$

$$s.t. \quad \sum_{i \in I(j)} y_i \geq z_j \qquad j \in J \tag{4.6}$$

$$\sum_{j \in J} d_j z_j \geq D \tag{4.7}$$

$$y_i \in \{0,1\} \qquad i \in I$$

$$z_j \in \{0,1\} \qquad j \in J.$$

The objective function (4.5) of this model minimizes the opening cost of the facilities. Constraints (4.6) guarantee that whenever a customer j is covered, at least one of the facilities from its neighbourhood is open. Constraint (4.7) ensures that the covered customer demand is at least D.

Part II

Contributions

Chapter 5

The wireless network design problem

We mentioned that practical WND problems are hard to solve using optimal procedures as numerical and memory issues arise even in small instances of the problem. In this chapter, we analyze separately the site and power assignment problem and the base stations deployment problem. We discuss how to strengthen their natural formulations by means of valid cutting planes, coefficient tightening and other presolve operations. All these operations will speed up the solution of the problem and reduce its size, as shown in the computational results of each case.

The content of this chapter is part of a paper [15] by myself and Professors Pasquale Avella from Università del Sannio and Laura Palagi from Sapienza University of Rome published in *Networks*, and a work [14] with the same authors recently submitted for publication in a journal.

5.1 The variable-power case

In this section, we discuss the site and power assignment problem, considering the power emissions as variables.

5.1.1 Presolve operations

Reducing the model size is crucial as real-life instances typically involve many variables and constraints. In this section we describe several operations that allow us to reduce the size of the problem by eliminating some x_{tb} and z_{bl} variables a priori. Note that the elimination of the x_{tb} variables also leads to the a priori elimination of the corresponding SINR_{tb} constraint (3.3).

Reducing the number of servers

Usually, for reasons related to signal quality, only a certain number of transmitters (the ones emitting the strongest signals received from the testpoint) are generally considered as possible servers for a given testpoint. Therefore, we establish a number of possible servers for each testpoint a priori. Namely, for each testpoint t, we select a subset of servers S_t, corresponding to the transmitters emitting the strongest signals received in t.

Then, to further reduce the size of the problem, we delete all the variables modeling transmitter-receiver pairs which any feasible solution would exclude. Specifically, suppose we exclude the possibility of serving the full target area with a single transmitter, indeed this case is trivial and would not require the use of optimization to be solved. In that case, the best-case scenario is the one with only two transmitters deployed, where one transmitter works as a server and the other as an interferer.

Hence, for each testpoint t, we fix $x_{ts} = 0$ for all those transmitters $s \in S_t$ such that the SINR_{ts} is below the threshold δ for each possible single interferer $b \in \mathcal{B} : b \neq s$. Namely, we eliminate all x_{ts} such that

$$\text{SINR}_{ts} = \frac{\tilde{a}_{ts} P_i}{\mu + \tilde{a}_{tb} P_j} < \delta \qquad \forall b \in \mathcal{B} \setminus \{s\}, \forall P_i, P_j \in \mathcal{P}$$

which can be easily verified directly with

$$\text{SINR}_{ts}^{max} = \frac{\tilde{a}_{ts} P_{max}}{\mu + \tilde{a}_{th} P_{min}} < \delta \qquad \text{with} \quad \tilde{a}_{th} \coloneqq \min_{b \in \mathcal{B} \setminus \{s\}} \{\tilde{a}_{tb}\}.$$

Reducing power levels

To reduce the size of the problem, we also select a subset of transmitter power levels which any feasible solution would exclude. Indeed, excluding again the possibility of serving the full target area with a single transmitter, we can consider the best-case scenario as the one with only two transmitters deployed, one server and one interferer. Hence, for each transmitter $b \in \mathcal{B}$ and level of power $l \in \mathcal{L}$, we fix $z_{bl} = 0$ for all those (b, l) such that the SINR is below the threshold δ for each possible interferer $\beta \in \mathcal{B} : \beta \neq b$ and for each testpoint $t \in \mathcal{T}$. Namely, we eliminate all z_{bl} such that

$$\text{SINR}_{tb} = \frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{t\beta} P_j} < \delta \qquad \forall t \in \mathcal{T}, \forall \beta \in \mathcal{B} \setminus \{b\}, \forall P_j \in \mathcal{P}$$

which can be easily verified directly with

$$\text{SINR}_{tb}^{max} = \frac{\tilde{a}_{tb}P_l}{\mu + \tilde{a}_{th}P_{min}} < \delta \qquad \forall t \in \mathcal{T}, \text{ with } \quad \tilde{a}_{th} \coloneqq \min_{\beta \in \mathcal{B}\setminus\{b\}} \{\tilde{a}_{t\beta}\}.$$

Heuristic sparsification

We perform a heuristic sparsification to deal with the numerical issues arising from the coefficients of the SINR inequalities. Specifically, we set a minimum threshold ε on the received power below which the received power can be considered null. Namely, for each $t \in \mathcal{T}$, $b \in \mathcal{B}$, $P_i \in \mathcal{P}$ we set

$$\tilde{a}_{tb}P_i = \begin{cases} \tilde{a}_{tb}P_i & \text{if } \tilde{a}_{tb}P_i \geq \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

This allows us to reduce the size of the problem by eliminating some x_{tb} variables a priori.

5.1.2 Cutting planes

A standard procedure for solving 0-1 LPs is the branch-and-bound algorithm, which can significantly be improved by cutting planes, i.e., inequalities that are valid for all integer solutions but not for some solutions of the linear relaxation. By means of such inequalities, fractional linear relaxation solutions can be cut off. Valid inequalities are internally generated by state-of-the-art MIP solvers. However, MIP solvers cannot take advantage of the particular problem structure known to the user. For the problem at hand, we identify some problem-specific cutting planes, including variable upper bounds and families of clique inequalities, that we provide in this section.

Families of clique inequalities

Suppose again to exclude the possibility of serving the full target area with a single transmitter and consider the best-case scenario as the one with only two transmitters deployed, one server and one interferer. We observe that we can exclude:

- potential levels of power for a certain transmitter $b \in \mathcal{B}$;
- potential serving signals for a certain testpoint $t \in \mathcal{T}$;

simply considering the minimum SINR required in t.

Hence, for each testpoint $t \in \mathcal{T}$, interferer $b \in \mathcal{B}$ with power $P_l \in \mathcal{P}$ and server $\beta \in \mathcal{B}$ such that the SINR measured in t and that considers as the only interferer b is always below the threshold δ, i.e.

$$\frac{\tilde{a}_{t\beta}P_i}{\mu + \tilde{a}_{tb}P_l} < \delta \quad \forall P_i \in \mathcal{P}$$

which can be easily verified directly with

$$\frac{\tilde{a}_{t\beta}P_{max}}{\mu + \tilde{a}_{tb}P_l} < \delta \tag{5.1}$$

we can exclude the possibility that b is activated at a power level equal to l and simultaneously β serves t using

$$z_{bl} + x_{t\beta} \leq 1. \tag{5.2}$$

Using (3.1) and (3.6), we can strengthen the cliques (5.2).

Theorem 5.1.1 *Given $(t, b, \beta, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{B}, \mathcal{L}\}$ such that (5.1) is satisfied, with l minimum power level satisfying (5.1) and $b \neq \beta$, the following cliques are valid inequalities*

$$\sum_{\lambda \in \mathcal{L}:\, \lambda \geq l} z_{b\lambda} + x_{t\beta} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}, \beta \in \mathcal{B} \setminus \{b\}. \tag{5.3}$$

Proof If $x_{t\beta} = 1$, then for each $\lambda \geq l$ we have $z_{b\lambda} = 0$ otherwise the $\text{SINR}_{t\beta}$ constraint (3.3) is violated. If instead $z_{b\lambda} = 1$ for one $\lambda \geq l$, then $x_{t\beta} = 0$ ($\beta \neq b$) since the $\text{SINR}_{t\beta}$ constraint (3.3) is violated. Hence, we cannot have simultaneously that $\sum_{\lambda \in \mathcal{L}:\, \lambda \geq l} z_{b\lambda} = 1$ and $x_{t\beta} = 1$.

Moreover, inequalities (5.2) are implied by (5.3) as

$$z_{bl} + x_{t\beta} \leq \sum_{\lambda \in \mathcal{L}:\, \lambda \geq l} z_{b\lambda} + x_{t\beta} \leq 1.$$

□

Theorem 5.1.2 *Given $(t, b, \beta, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{B}, \mathcal{L}\}$ such that (5.1) is satisfied for all $\beta \neq b$, the following cliques are valid inequalities*

$$z_{bl} + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}, l \in \mathcal{L}. \tag{5.4}$$

Proof If $z_{bl} = 1$, then each $x_{t\beta} = 0$ ($\beta \neq b$) since the $\text{SINR}_{t\beta}$ constraint (3.3) is violated. If instead $x_{t\beta} = 1$ for one $\beta \neq b$, then $z_{bl} = 0$ otherwise the $\text{SINR}_{t\beta}$ constraint (3.3) is

violated. Hence, we cannot have simultaneously that $z_{bl} = 1$ and $\sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} = 1$.

Moreover, inequalities (5.2) are implied by (5.4) as

$$z_{bl} + x_{t\beta} \leq z_{bl} + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1.$$

□

Moreover, given the testpoint $t \in \mathcal{T}$, the server $b \in \mathcal{B}$ with power $P_l \in \mathcal{P}$ and the interferer $\beta \in \mathcal{B}$ such that the SINR measured in t and that considers as the only interferer β is always below the threshold δ for each possible $\beta \neq b$, i.e.

$$\frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{t\beta} P_j} < \delta \quad \forall \beta \in \mathcal{B} \setminus \{b\}, \forall P_j \in \mathcal{P}$$

which can be easily verified directly with

$$\frac{\tilde{a}_{tb} P_l}{\mu + \tilde{a}_{th} P_{min}} < \delta, \quad \tilde{a}_{th} \coloneqq \min_{\beta \in \mathcal{B} \setminus \{b\}} \{\tilde{a}_{t\beta}\} \tag{5.5}$$

we can exclude the possibility that b is activated at a power level equal to l and simultaneously b serves t using

$$z_{bl} + x_{tb} \leq 1. \tag{5.6}$$

Using (3.1), we can strengthen the cliques (5.6).

Theorem 5.1.3 *Given $(t, b, l) \in \{\mathcal{T}, \mathcal{B}, \mathcal{L}\}$ such that (5.5) is satisfied and l corresponds to the maximum power level such that (5.5) is satisfied, the following cliques are valid inequalities*

$$\sum_{\lambda \in \mathcal{L}:\, \lambda \leq l} z_{b\lambda} + x_{tb} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}. \tag{5.7}$$

Proof If $x_{tb} = 1$, then for each $\lambda \leq l$ we have $z_{b\lambda} = 0$, otherwise the SINR_{tb} constraint (3.3) is violated. If instead $z_{b\lambda} = 1$ for one $\lambda \leq l$, then $x_{tb} = 0$ since the SINR_{tb} constraint (3.3) is violated. Hence, we cannot have simultaneously that $\sum_{\lambda \in \mathcal{L}:\, \lambda \leq l} z_{b\lambda} = 1$ and $x_{tb} = 1$.

Moreover, inequalities (5.6) are implied by (5.7) as

$$z_{bl} + x_{tb} \leq \sum_{\lambda \in \mathcal{L}:\, \lambda \leq l} z_{b\lambda} + x_{tb} \leq 1.$$

□

Variable upper bounds

Variable upper bound constraints (VUBs)

$$x_{tb} \leq \sum_{l \in \mathcal{L}} z_{bl} \qquad t \in \mathcal{T}, b \in \mathcal{B} \tag{5.8}$$

enforce that a testpoint $t \in \mathcal{T}$ can be assigned to a transmitter $b \in \mathcal{B}$ only if b is activated, i.e., only if b is using one strictly positive power level. They are known to strengthen the quality of linear relaxation significantly.

Let us denote by $\mathcal{L}_{tb} \subseteq \mathcal{L}$ the subset of power levels l that satisfy (5.5) for a given transmitter-receiver pair $(b, t) \in \{\mathcal{B}, \mathcal{T}\}$, meaning that we can exclude that b is activated at a power level $l \in \mathcal{L}_{tb}$ and simultaneously b serves t. The VUBs (5.8) can be tightened to

$$x_{tb} \leq \sum_{l \in \mathcal{L} \setminus \mathcal{L}_{tb}} z_{bl} \qquad t \in \mathcal{T}, b \in \mathcal{B} \tag{5.9}$$

5.1.3 Reduced cost fixing to tighten the big-M

To further reduce the model size, we propose using a reduced cost fixing method, in short RCF (see [2] for a survey on presolve techniques). Although this procedure is well-known and widespread, no one has ever tried to see its effects on this type of problem (based on our knowledge).

By solving the linear relaxation of the problem, we can get the lower bound lb and the corresponding reduced costs $\bar{c}_{bl}$ associated with the sole variables z_{bl} in the optimal solution of the linear relaxation. Then, given an upper bound $ub > lb$, if $\bar{c}_{bl} \geq ub - lb$ for some $b \in \mathcal{B}, l \in \mathcal{L}$, the corresponding z_{bl} must be at its lower bound in every optimal solution; hence we can fix $z_{bl} = 0$.

Whenever the fixing of a variable z_{bl} occurs at a given $l \in \mathcal{L}$ such that $P_l = P_{max}$, we can recompute and reduce the big-M, resulting in a tightening of the formulation. Indeed, after the RCF, we may have some transmitters b that cannot emit at the maximum power level l such that $P_l = P_{max}$, since the corresponding z_{bl} variables have been fixed to zero. In such cases, the value by which a_{tb} is weighted in the big-M (see (3.4)) can be reduced to the highest power value that b can assume, which is strictly less than P_{max}.

To formalize it, let us define the set $\mathcal{B}^R \subseteq \mathcal{B}$ of base stations affected by RCF, i.e., such that $b \in \mathcal{B}^R$ if the variable z_{bl} has been fixed to zero for at least one $l \in \mathcal{L}$. Then, for a given $b \in \mathcal{B}^R$, we can define the set $\mathcal{L}_b^R \subset \mathcal{L}$ of power levels that b can assume after the RCF. We denote by $P_{b,max}^R$ the power value corresponding to the maximum power level

that $b \in \mathcal{B}^R$ can assume. Since $\mathcal{L}_b^R \subset \mathcal{L}$, we have that $P_{b,max}^R \leq P_{max}$. Using this notation, we can write down the new value of the big-M

$$M'_{t\beta} = \delta\mu + \delta \left(P_{max} \sum_{b \in \mathcal{B} \setminus \{\beta, \mathcal{B}^R\}} \tilde{a}_{tb} + \sum_{b \in \mathcal{B}^R} P_{b,max}^R \tilde{a}_{tb} \right) = \delta\mu + \delta \sum_{b \in \mathcal{B} \setminus \{\beta\}} \tilde{P}_b \tilde{a}_{tb}$$

which satisfies $M_{t\beta} \geq M'_{t\beta}$ since $P_{max} \geq \tilde{P}_b = \begin{cases} P_{max} & \text{if } b \in \mathcal{B} \setminus \mathcal{B}^R \\ P_{b,max}^R & \text{if } b \in \mathcal{B}^R. \end{cases}$

The smaller the optimality gap given by the estimated lower and upper bounds, the greater the number of z_{bl} variables that can be fixed to zero, and the smaller the big-M coefficients. Hence, applying a standard algorithm – as implemented in MIP commercial solvers – to the tightened formulation (i.e., the formulation got after the RCF) produces stronger bounds and a faster solution.

We can apply a further reduction. Since we have information on the maximum number of transmitters that can be installed from the *ub*, we can further reduce the value of the big-M by replacing the sum of all the interfering signals in the testpoint t with the sum of the *strongest* interfering signals in t. In particular, only the strongest γ interferers are considered, where γ is the maximum number of transmitters that can be activated. Given γ, the big-M can be computed as

$$M''_{t\beta} = \delta\mu + \delta \sum_{b \in \mathcal{A}_t \setminus \{\beta\}} \tilde{P}_b \tilde{a}_{tb} \leq M'_{t\beta} \leq M_{t\beta} \tag{5.10}$$

where $\mathcal{A}_t \subset \mathcal{B}$ is the set of the γ base stations emitting the strongest signals received in t, i.e $|\mathcal{A}_t| = \gamma$. The smaller is γ, and the smaller is the big-M; therefore, the estimate of γ should be as accurate as possible.

We observe that getting a good lower bound is straightforward since constraints (5.8) naturally lead to a good linear relaxation value. Conversely, finding a good upper bound is a more daunting task. Although commercial MIP solvers can be used to derive a feasible solution, it may be time-consuming. Consequently, we derived a fixing heuristic based on the observation that the fractional values of the variables in the LP relaxation are often good predictors of zero/non-zero variables in an optimal ILP solution. This especially occurs when the LP relaxation is extremely tight. The scheme of our fixing heuristic is reported in Figure 5.1. The variables rounded to zero correspond to the power levels that are not needed by the transmitters to cover the target area.

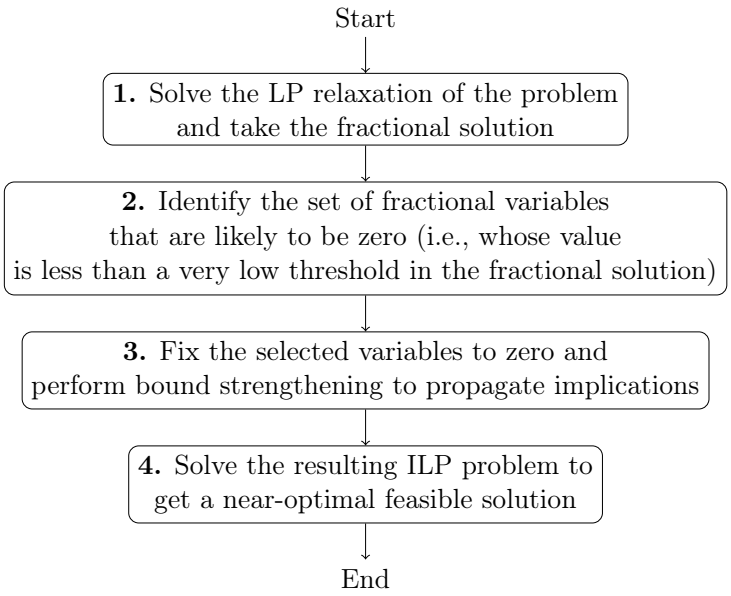

Figure 5.1: Block diagram representing the fixing heuristic.

5.1.4 The final formulation

The final formulation (F) we propose differs from the initial formulation (3.7) since:

(i) it is the result of a reduced cost fixing procedure;

(ii) it includes the addition of the VUBs (5.9) and the cliques (5.3), (5.4), (5.7);

(iii) the big-M appearing in the SINR is formulated as in (5.10);

(iv) the number of servers and the number of levels for the power has been reduced according to the three operations described in Section 5.1.1.

5.1.5 Computational experiments

We compare the results obtained using formulation (3.7) – that we denote as the basic formulation, or B – and the final setting in Section 5.1.4 – that we denote as F – to solve the site and power assignment problem. The code has been implemented in Python, and the experiments have been carried out on a Ubuntu server with an Intel(R) Xeon(R) Gold 5218 CPU running at 2.30 GHz, with 96 GB of RAM and 8 cores. Gurobi Optimizer 10.0.1 [88] with default settings has been employed as an MIP solver. We set a time limit of four hours for computation time.

The testbed

The FUB provided us with the testbed of an existing LTE network, in which the set of transmitters consists of the 135 authentic potential locations for the transmitters in the

Municipality of Bologna (Italy). Since the technology we refer to is 4G LTE 800 MHz, the elementary areas should have a square shape with a side of 100 m following the Italian Resilience Plan [94], which is based on the guidelines given by the Body of European Regulators for Electronic Communications [23, 24]. Thus, the set of the testpoints provided by the FUB corresponds to the centres of the 14 078 squared elementary areas that can be obtained on Bologna.

The system noise and the power received in each receiver by each possible transmitter have been simulated by the FUB using the *Cost Hata* model [7, 77, 128]. The power values are in W and scaled by a factor of 10^{10} to avoid numerical issues and obtain better accuracy on optimal solutions, as suggested in [60]. Accordingly, the threshold on the quality of service δ and the system noise μ are expressed in W; μ is also scaled by 10^{10}. The emitted power values considered in this study are three: $20W$, $40W$, $80W$. The threshold ε has been set around -110 dBmW [135, 137]. In (3.5), the parameter r_t has been set to the fraction of the population living in the elementary area $t \in \mathcal{T}$, and r to the minimum fraction of the population to be covered by service. We used $|S_t| = 10$, namely we selected the strongest ten signals as possible servers for each testpoint t.

Since the number of the testpoints, which is strictly related to the size of the elementary areas, significantly impacts the size of the problem, we decided to aggregate the testpoints and consequently reduce the number of variables and constraints of the model. It is crucial to note that this step was indispensable, since without it the size of the problem exceeded the memory limits of our computational resources, preventing us from even creating a mathematical model of the problem. Specifically, we used *k-means clustering* to smartly partition testpoints into groups (or clusters) such that intra-cluster data points are as similar as possible while keeping the clusters as different (far) as possible. We considered the sum of the squared distance between the cluster centroid and the vectors describing the data points as a measure of similarity. For this analysis, we considered the medium level of power only. In particular, we assume that each testpoint $t \in \mathcal{T}$ is characterized by a power vector $p_t = [\tilde{a}_{tb}P_2]_{b\in\mathcal{B}} \in R^{|\mathcal{B}|}$ of the received power values from each transmitter $b \in \mathcal{B}$ emitting at the medium power level $l = 2$. Intuitively, we assume that two points with similar power vectors have similar behavior with respect to the SINR, and hence they can be aggregated in the same cluster. This behaviour is independent of the power level used for this analysis.

We fixed the number of clusters to K and used the k-means++ algorithm implemented in the Scikit-learn library [119] with the default setting values for the maximum number of iterations and the relative tolerance, enabling a multistart procedure to improve the

local solution. The solution returned by the algorithm is made up of the clusters $\mathcal{C}_i$ and their centroids p_i^c with $i \in 1, \ldots, K$, that generally do not correspond to any point p_t. We selected as the representative for each cluster $\mathcal{C}_i$, the receiver $t \in \mathcal{C}_i$ having the closest p_t to p_i^c.

Depending on the urban and territorial morphology and transmitter positions, clusters may have different cardinality. Specifically, it is possible to identify areas split into many small clusters as the variation of the power received differs significantly even in neighboring points (e.g., in dense urban areas). In other areas, a coarse cluster is sufficient to represent the behavior of the receivers within it (e.g., in rural areas). This procedure resembles the widespread coarsening and refinement approach of the METIS algorithm [96], which suggests a first selection and subsequent refinement where necessary.

The territorial distribution of the testpoints of the Municipality of Bologna is depicted in Figure 5.2. The blue dots are the 14 078 original testpoints of the Municipality of Bologna identified by the FUB according to [23, 24, 94]. The black dots are the 4 693 testpoints selected with the clustering procedure when the number of clusters is fixed to $K = 4\,693$ (about one-third of the original testpoints).

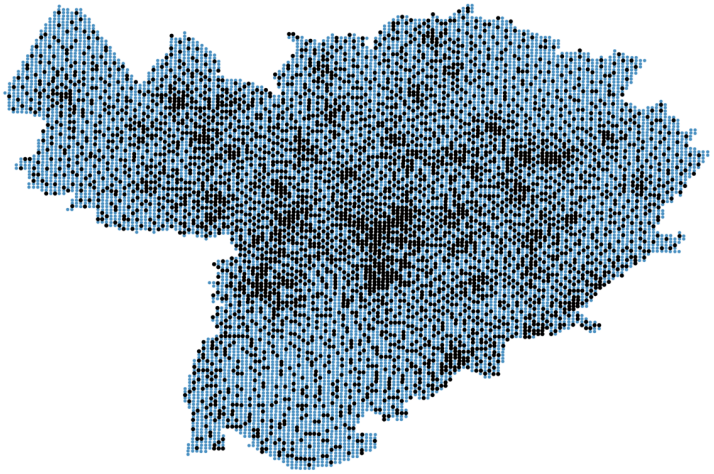

Figure 5.2: In black are the testpoints selected with the clustering procedure, and in blue are the original testpoints of the Municipality of Bologna.

By reducing the number of testpoints and the number of base stations serving each testpoint, we got a reduced network. From this network, we derived eleven instances (BOV1 to BOV11), each of them differing in the quality of service required (increasing with the number) in the receivers and in the fraction of the population to be served, as reported in Table 5.1. The estimate of the upper bound has then been obtained on each instance using the fixing heuristic.

Compared to the instances we derived for the fixed power case (see Section 5.2.5), we

observe that the instances used for the variable power case were obtained considering a more restricted range of the SINR threshold δ: this is because unfortunately both Gurobi and our fixing heuristic are not in able to find good feasible solutions to be used in the RCF within the time limit for instances that require high service values ($\delta > -5$ dBW).

Table 5.1: Characteristics of the instances: SINR threshold (δ) and fraction of population to be served (r).

Instance	**BOV1**	**BOV2**	**BOV3**	**BOV4**	**BOV5**	**BOV6**	**BOV7**	**BOV8**	**BOV9**	**BOV10**	**BOV11**
δ[dBW]	-10	-9,5	-9	-8.5	-8	-7.5	-7	-6.5	-6	- 5.5	- 5
r	1	1	1	1	1	1	1	1	1	0.99	0.99

Results

Here we show the impact of the operations discussed in the previous Sections 5.1.1, 5.1.2, 5.1.3. The tested formulations are described in Table 5.2. In particular, we focus on three evaluation criteria, namely size, sparsity, quality of the bounds at the root node and the end of the optimization.

Table 5.2: Characteristics of the tested formulations.

Formulation	**Characteristics**
B	Basic formulation (3.7)
B+CPs	Formulation (3.7) plus the addition of cutting planes (i.e., VUBs (5.9) and cliques (5.3), (5.4), (5.7))
F	Final setting reported in Section 5.1.4

Figures 5.3-5.5 show the average number of variables, constraints and non-zeros of each formulation respectively. Figures 5.6-5.7 show three box plots, one for each formulation, displaying the distribution of the data concerning the relative gap at the root node and the end of the optimization, based on a summary of five numbers (minimum value, first quartile, median value, third quartile, and maximum value).

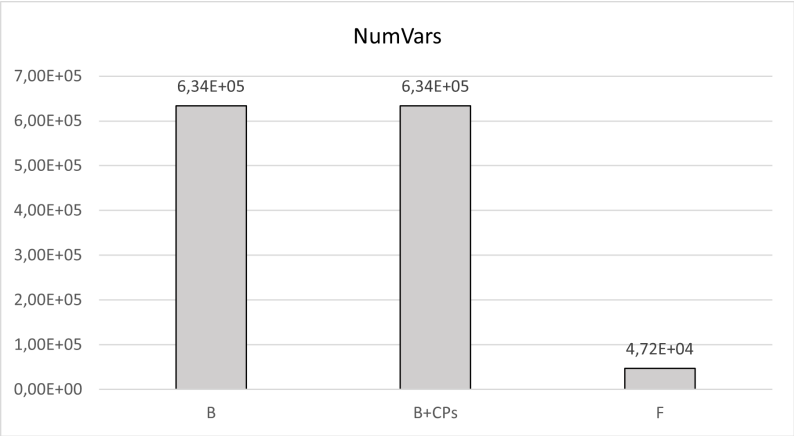

Figure 5.3: Average number of variables in the tested formulations (see Table 5.2).

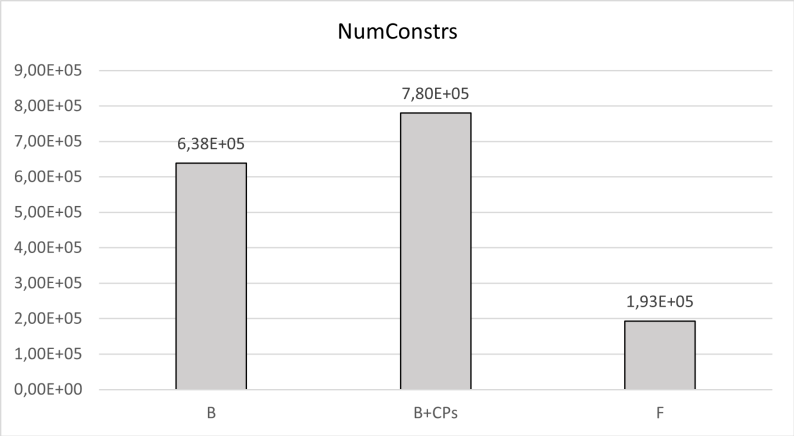

Figure 5.4: Average number of constraints in the tested formulations (see Table 5.2).

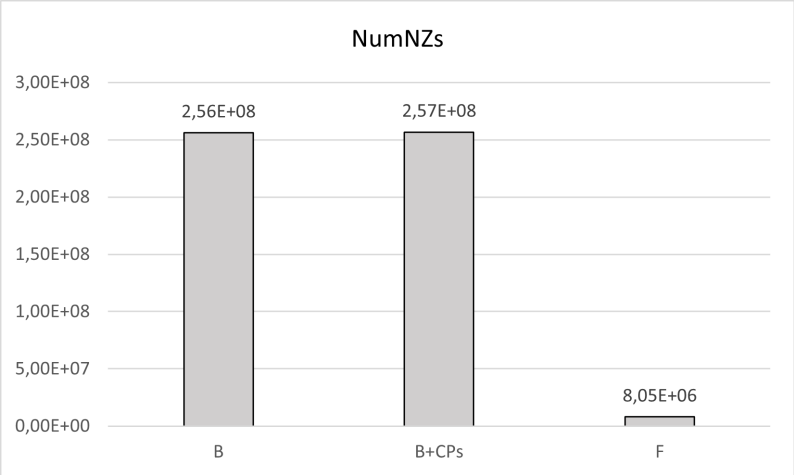

Figure 5.5: Average number of non-zeros in the tested formulations (see Table 5.2).

Graphs in Figures from 5.3 to 5.7 show that:

- formulation B+CPs, leads to a significant improvement in the root relaxation value, revealing the decisive effect of adding the cutting planes;
- the aggressive reduction scheme in F, also including the RCF procedure, produces

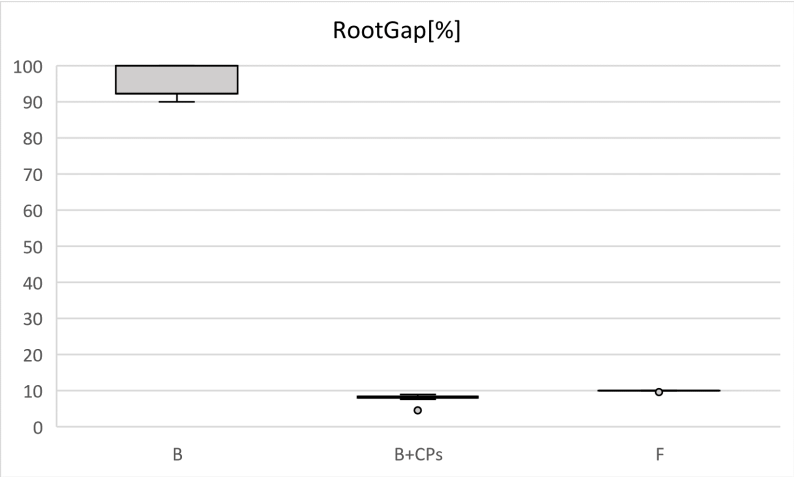

Figure 5.6: Box plots of the relative gap at the root node in the tested formulations (see Table 5.2).

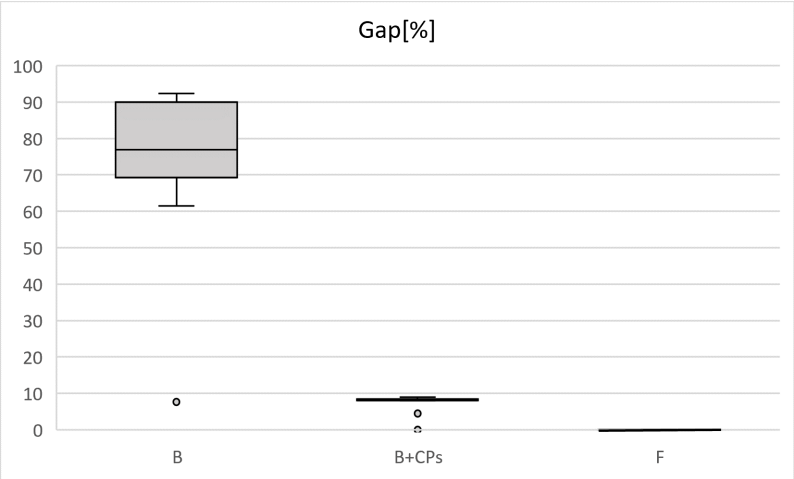

Figure 5.7: Box plots of the relative gap at the end of the optimization phase in the tested formulations (see Table 5.2).

a slight worsening of the root bound but leads to a definitely reduced and sparser formulation;

- we can solve the instances to optimality within the time limit only using formulation F; however, a good final gap can be reached already using B+CPs.

A comparison between formulation B and formulation F is reported in Table 5.3. The evaluation metrics considered are: number of variables (NumVars), constraints (NumConstrs) and non-zeros (NumNZs); maximum value of the big-M (MaxBigM); relative optimality gap at the root node (RootGap), final relative optimality gap (Gap); number of explored nodes (Nodes); total solution time (Time). Denoting the optimal value with Opt, the value of the lower bound at the root node with RootLB, and the value of the best lower bound with BestLB, the RootGap is computed as 100(Opt-RootLB)/Opt, and the Gap

as 100(Opt-BestLB)/Opt. We reported in bold the gap of instances solved to optimality within the time limit.

Table 5.3: Optimization results for the instances BOV1–BOV11 (see Table 5.1): a comparison between the basic (B) and the final (F) formulations. Time is expressed in seconds.

Instance	Formulation	NumVars	NumConstrs	NumNZs	MaxBigM	RootGap[%]	Gap[%]	Nodes	Time[s]
BOV1	**B**	633960	638383	256097415	91958,73	100,00	92,31	1484	TL
	F	47201	164365	8170212	45979,28	9,97	**0,00**	83	429,96
BOV2	**B**	633960	638383	256097415	103179,39	100,00	7,69	1503	TL
	F	47201	169027	8185122	103179,20	9,97	**0,00**	110	534,67
BOV3	**B**	633960	638383	256097415	115769,18	100,00	69,23	14472	TL
	F	47201	174015	8201068	115768,96	9,97	**0,00**	122	518,33
BOV4	**B**	633960	638383	256097415	129895,16	92,31	76,92	14505	TL
	F	47201	179348	8217919	64947,46	9,97	**0,00**	129	471,71
BOV5	**B**	633960	638383	256097415	145744,77	100,00	76,92	12182	TL
	F	47201	185052	8265317	145744,49	9,97	**0,00**	105	541,17
BOV6	**B**	633960	638383	256097415	161284,63	92,31	76,92	15032	TL
	F	47200	190259	8241827	80642,16	9,97	**0,00**	151	680,76
BOV7	**B**	633960	638383	256097415	183481,79	92,31	61,54	14708	TL
	F	47200	197513	8265097	183481,44	9,97	**0,00**	293	586,75
BOV8	**B**	633960	638383	256097415	205869,96	92,31	84,62	14353	TL
	F	47201	204542	8322856	205869,56	9,97	**0,00**	127	598,99
BOV9	**B**	633960	638383	256097415	230989,89	92,31	69,23	14793	TL
	F	47202	212046	8394185	115494,79	9,97	**0,00**	86	537,43
BOV10	**B**	633960	638384	256598940	259174,92	90,00	90,00	14756	TL
	F	47158	220226	7108664	259174,08	9,63	**0,00**	37	1922,65
BOV11	**B**	633960	638384	256598940	290799,04	90,00	90,00	14753	TL
	F	47157	229013	7156663	142654,51	9,63	**0,00**	37	2311,70

TL, time limit reached

Results show that formulation F is definitely reduced in size and sparser than B. This is mainly due to the presolve operations and the reduced cost fixing procedure. The value of the largest big-M used in F is half that used in B in half the instances. The quality of the bounds at the root node is better in F, mainly due to the addition of cutting planes. Thanks to a good root bound, the number of the explored nodes is heavily reduced in F, and consequently, the time spent on the branch-and-cut tree search. Overall, solution times are significantly reduced if we compare F to B. Indeed, none of the basic formulations have been successfully solved within the time limit of 4 hours, whereas each final formulation has been solved to optimality in less than 40 minutes.
In the end, formulation F turns out to be much more competitive than the basic formulation B.

Finally, in Table 5.4 we reported the computational time of every phase of the final setting F. In particular, by LBTime we denoted the time needed to get the lower bound

Table 5.4: Solution times on the instances BOV1–BOV11 (see Table 5.1) of the final formulation. Time is expressed in seconds.

Instance	LBTime[s]	UBTime[s]	SolTime[s]	Time[s]
BOV1	34,85	31,53	363,58	429,96
BOV2	33,61	34,56	466,50	534,67
BOV3	32,79	34,88	450,66	518,33
BOV4	37,57	41,40	392,74	471,71
BOV5	33,25	42,21	465,71	541,17
BOV6	39,08	47,49	594,19	680,76
BOV7	34,14	52,14	500,47	586,75
BOV8	33,75	54,54	510,70	598,99
BOV9	36,43	64,36	436,64	537,43
BOV10	47,49	28,61	1846,55	1922,65
BOV11	52,87	29,39	2229,44	2311,70

given by the linear relaxation of the problem, by UBTime the time needed to get an upper bound using our fixing heuristic, by SolTime the time needed to solve the problem after the RCF procedure, and by Time the overall solution time given by the sum of the previous components. The speed with which the instances can be solved using the scheme based on RCF is due to the speed of calculation of the lower and upper bounds. Both depend on the good quality of the linear relaxation of the problem, achieved thanks to the inclusion of the cuts we introduced. Note that for higher quality values, unfortunately the lower bound gets slightly worse and it is no longer possible to exploit our fixing heuristic. We also tried to exploit Gurobi as a heuristic to obtain an upper bound on high service instances, but unfortunately the bounds produced by Gurobi in times compatible with the use of this procedure were not good enough to actually apply the RCF.

5.2 The fixed-power case

In this section, we focus on the base station deployment problem, considering the power level as fixed. We discuss how to strengthen the natural formulation of the problem by means of valid cutting planes and coefficient tightening operations. We also propose several operations and an aggregation of constraints in order to reduce the problem size.

5.2.1 Cutting planes

In the fixed-power case, the VUBs become:

$$x_{tb} \leq z_b \quad t \in \mathcal{T}, b \in \mathcal{B}. \tag{5.11}$$

Moreover, for each testpoint $t \in \mathcal{T}$, interferer $b \in \mathcal{B}$ and server $\beta \in \mathcal{B}$ such that the SINR measured in t and that considers as the only interferer b is below the threshold δ, i.e.

$$\frac{a_{t\beta}}{\mu + a_{tb}} < \delta \tag{5.12}$$

we can exclude the possibility that b is activated and simultaneously β serves t using the clique

$$z_b + x_{t\beta} \leq 1. \tag{5.13}$$

Using (3.6), we can strengthen the cliques (5.13).

Theorem 5.2.1 *Given* $(t, b, \beta) \in \{\mathcal{T}, \mathcal{B}, \mathcal{B}\}$ *such that* (5.12) *is satisfied for all* $\beta \neq b$*, the following cliques are valid inequalities*

$$z_b + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1 \qquad b \in \mathcal{B}, t \in \mathcal{T}. \tag{5.14}$$

Proof If $z_b = 1$, then each $x_{t\beta} = 0$ ($\beta \neq b$) since the $\text{SINR}_{t\beta}$ constraint (3.8) is violated. If instead $x_{t\beta} = 1$ for one $\beta \neq b$, then $z_b = 0$ otherwise the $\text{SINR}_{t\beta}$ constraint (3.8) is violated. Hence, we cannot have simultaneously that $z_b = 1$ and $\sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} = 1$.

Moreover, inequalities (5.13) are implied by (5.14) as

$$z_b + x_{t\beta} \leq z_b + \sum_{\beta \in \mathcal{B} \setminus \{b\}} x_{t\beta} \leq 1.$$

□

5.2.2 Aggregation of the SINR constraints

To address the numerical criticalities outlined in Section 3.4, we present a reformulation of the SINR constraints.

Each SINR constraint (3.8) can be strengthen by replacing the z_β with the $x_{t\beta}$ as in

the following:

$$a_{t\beta}x_{t\beta} - \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu - M_{t\beta}(1 - x_{t\beta}) \qquad t \in \mathcal{T}, \beta \in \mathcal{B}. \tag{5.15}$$

Theorem 5.2.2 *Inequalities* (5.15) *are valid for* S.

Proof When $x_{t\beta} = 1$, the corresponding inequality of (5.15) becomes $a_{t\beta} - \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu$ which is the required SINR inequality. When $x_{t\beta} = 0$, we get $-\delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b \geq \delta\mu - M_{t\beta}$ that is trivially satisfied.

□

Proposition 5.2.3 *Inequalities* (3.8) *are implied by inequalities* (5.11) *and* (5.15).

Proof For each $t \in \mathcal{T}$ and $\beta \in \mathcal{B}$, inequalities (3.8) and (5.15) can be written respectively as $z_\beta \geq \frac{h}{a_{t\beta}}$ and $x_{t\beta} \geq \frac{h}{a_{t\beta}}$ where $h = \delta \sum_{b\in\mathcal{B}\setminus\{\beta\}} a_{tb}z_b + \delta\mu - M_{t\beta}(1 - x_{t\beta})$ and using $a_{t\beta} > 0$. Inequality (5.11) can be written as $z_\beta - x_{t\beta} \geq 0$. Hence we get that inequality (3.8) can be obtained summing up inequalities (5.11) and (5.15), meaning that it is implied by them.

□

In order to reduce the size of the problem, we aggregate constraints (5.15) using (3.6), producing a new form of SINR constraints that we name *aggregate SINR constraints*

$$(1 + \delta)\sum_{b\in\mathcal{B}} a_{tb}x_{tb} - \delta\sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t(1 - \sum_{b\in\mathcal{B}} x_{tb}) \qquad t \in \mathcal{T} \tag{5.16}$$

where the big-M term depends only on the testpoint t and could be set to

$$M_t = \delta\mu + \delta\sum_{b\in\mathcal{B}} a_{tb} \geq M_{t\beta}.$$

Theorem 5.2.4 *Inequalities* (5.16) *are valid for* S.

Proof Given $t \in \mathcal{T}$, we can have the following cases according to (3.6): t is covered by exactly one base station β, or t is not covered by any base station. In the former case, thanks to (3.6), the sum $\sum_{b\in\mathcal{B}} a_{tb}x_{tb}$ reduces at most to a single element $a_{t\beta}x_{t\beta}$, being β

the unique transmitter serving testpoint t, and the same is for $\sum_{b\in\mathcal{B}} x_{tb} = x_{t\beta}$. Thus (5.16) reduces to

$$(1+\delta)a_{t\beta}x_{t\beta} - \delta\sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t(1-x_{t\beta}).$$

Since $x_{t\beta} = 1$ implies $z_\beta = 1$, we get that (5.15) is satisfied . In the latter case, no transmitter is serving testpoint t and we get $\sum_{b\in\mathcal{B}} a_{tb}x_{tb} = 0$ and $\sum_{b\in\mathcal{B}} x_{tb} = 0$. Thus (5.16) reduces to

$$-\delta\sum_{b\in\mathcal{B}} a_{tb}z_b \geq \delta\mu - M_t$$

which satisfy (5.15) when $x_{t\beta} = 0$.

□

We observe that Theorems 5.2.2 and 5.2.4 and Proposition 5.2.3 allow using the aggregate constraints (5.16) for replacing all the SINR constraints (3.8) in problems (3.9). Indeed, for each receiver t the aggregate SINR constraint is only one, whereas the SINR constraints are $m = |\mathcal{B}|$, meaning that we can significantly reduce the number of constraints by using the aggregate constraints only.

5.2.3 Presolve operations

We describe how to reduce the model size by eliminating some x_{tb} variables a priori and how to tighten the formulation by reducing the value of the big-M.

Reducing the number of servers

For each testpoint t, we select a subset of servers S_t corresponding to the transmitters emitting the strongest signals received in t. Excluding the possibility of serving the full target area with a single transmitter, we can consider the best-case scenario as the one with only two transmitters deployed, where one transmitter works as a server and the other as an interferer. Hence, for each testpoint t, we fix $x_{ts} = 0$ for all those transmitters $s \in S_t$ such that the SINR is below the threshold δ for each possible interferer $b \in \mathcal{B} : b \neq s$. Namely, we eliminate all x_{ts} such that

$$\text{SINR}_{ts} = \frac{a_{ts}}{\mu + a_{tb}} < \delta \qquad \forall b \in \mathcal{B} \setminus \{s\}$$

which can be easily verified directly with

$$\text{SINR}_{ts}^{max} = \frac{a_{ts}}{\mu + a_{th}} < \delta, \quad \text{with} \quad a_{th} \coloneqq \min_{b \in \mathcal{B} \setminus \{s\}} \{a_{tb}\}.$$

Heuristic sparsification

We set a minimum threshold ε on the received power below which the received power can be considered null. Namely, we set

$$a_{tb} = \begin{cases} a_{tb} & \text{if } a_{tb} \geq \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

This allows us to reduce the size of the problem by eliminating some x_{tb} variables a priori.

Coefficient tightening by big-M reduction

We can tighten the formulation by computing the smallest possible big-M. Consider the big-M suitable for the aggregate SINR constraints, namely

$$M_t = \delta\mu + \delta \sum_{b \in \mathcal{B}} a_{tb} \quad \text{for } t \in \mathcal{T}.$$

To decrease its value, we replace the sum of the signal power received from all the transmitters in testpoint t with the sum of the strongest interferers. In particular, only the strongest α interferers might be considered, where α is an upper bound of the optimal number of activated base stations, i.e. $\alpha \geq \sum_{b \in \mathcal{B}} z_b^*$ with (x^*, z^*) optimal solution. Given α, the big-M can be computed as

$$M_t' = \delta\mu + \delta \sum_{b \in \mathcal{A}_t} a_{tb} \quad \text{for } t \in \mathcal{T} \tag{5.17}$$

where $\mathcal{A}_t$ is the set of the α base stations emitting the strongest signals received in t, i.e. $|\mathcal{A}_t| = \alpha$. The better is the bound α, the smaller is the big-M. Hence, the estimate of α must be as accurate as possible.

5.2.4 The final formulation

The final formulation (F) we propose differs from the initial formulation (3.9) since:

(i) it includes the addition of the VUBs (5.11) and the cliques (5.14);

(ii) it contains the aggregate SINR constraints (5.16) instead of the SINR constraints (3.8);

(iii) the big-M appearing in the aggregate SINR is formulated as in (5.17);

(iv) the number of servers has been reduced according to the first two operations described in Section 5.2.3.

5.2.5 Computational experiments

We compare the results obtained using formulation (3.9) – that we denote as the basic formulation, or B – and the final setting presented in 5.2.4 – that we denote as F – to solve the base station deployment problem. The code has been implemented in Python, and the experiments have been carried out on a Ubuntu server with an Intel(R) Xeon(R) Gold 5218 CPU running at 2.30 GHz, with 96 GB of RAM and 8 cores. Gurobi Optimizer 10.0.1 [88] with default settings has been employed as an MIP solver. We set a time limit of four hours for computation time.

The testbed

We obtained the instances from the same testbed previously described in Section 5.1.5. We here underline a few differences between the variable-power case and the fixed-power case. To generate the power data, the transmitter power emission was set at the mean value of 40 W.

From the reduced network, we derived ten instances (BOF1 to BOF10), each of them differing in the quality of service required (increasing with the number) in the receivers and in the fraction of the population to be served, as reported in Table 5.5. Not to affect the solution time, we fixed a priori the upper bounds α to small percentages of activated transmitters (15%, about 10% and 5%); a posteriori, we verified that the optimal number of activated transmitters was lower than α. An estimate of α can also be obtained using a heuristic. We observe that the instances used for the fixed-power case have been obtained considering a bigger range of the SINR threshold δ and are indeed more representative than the instances used in the variable-power case.

Results

Here we show the impact of the operations discussed in the previous Sections 5.2.1, 5.2.2, 5.2.3. The tested formulations are described in Table 5.6. In particular, we focus on four evaluation criteria, namely size, sparsity, quality of root bounds, and times.

Table 5.5: Characteristics of the instances: SINR threshold (δ) and fraction of population to be served (r).

Instance	BOF1	BOF2	BOF3	BOF4	BOF5	BOF6	BOF7	BOF8	BOF9	BOF10
δ[dBW]	-10	-8,5	-7,5	-6	-5	-3	0	+3	+5	+7
r	1	1	1	1	0.99	0.95	0.85	0.75	0.70	0.65

Table 5.6: Characteristics of the tested formulations.

Formulation	Characteristics
B	Basic formulation (3.9)
B+CPs	Formulation (3.9) plus the addition of cutting planes (i.e., VUBs (5.11) and cliques (5.14))
B+CPs+Agg	Formulation B+CPs, plus the replacement of the SINRs (3.8) with the aggregate SINRs (5.16)
F	Final setting reported in Section 5.2.4

Figures 5.8-5.10 show the average number of variables, constraints and non-zeros of each formulation respectively. Figures 5.11-5.12 shows four box plots, one for each formulation, displaying the distribution of the data concerning the relative gap at the root node and the computational time based on a summary of five numbers (minimum value, first quartile, median value, third quartile, and maximum value).

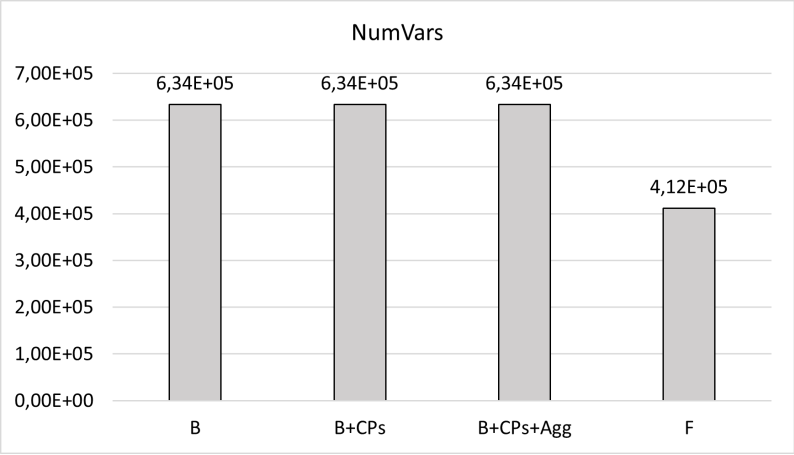

Figure 5.8: Average number of variables in the tested formulations (see Table 5.6).

Graphs in Figures from 5.8 to 5.12 show that:

- formulation B+CPs, leads to a significant improvement in the root relaxation value, revealing the decisive effect of adding the cutting planes;
- the aggregation of the SINR constraints in formulation B+CPs+Agg involves, in some instances, a slight worsening of the bounds but leads to a definitely reduced and sparser formulation;
- sparsity continues to increase after carrying out coefficient tightening operations:

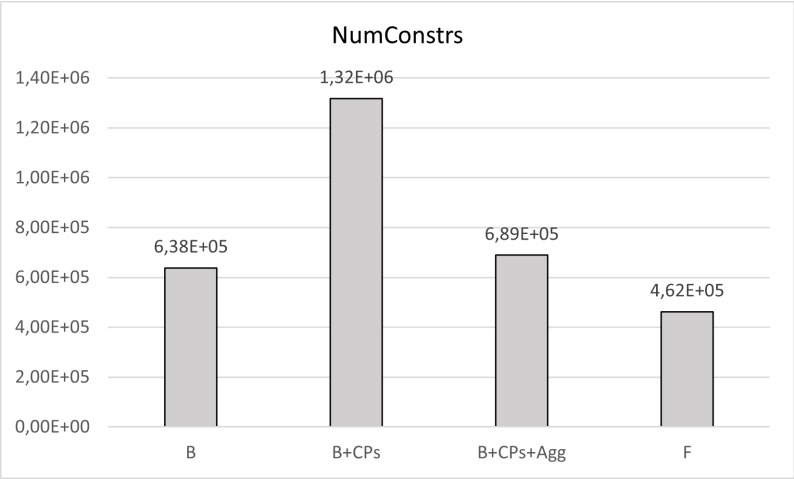

Figure 5.9: Average number of constraints in the tested formulations (see Table 5.6).

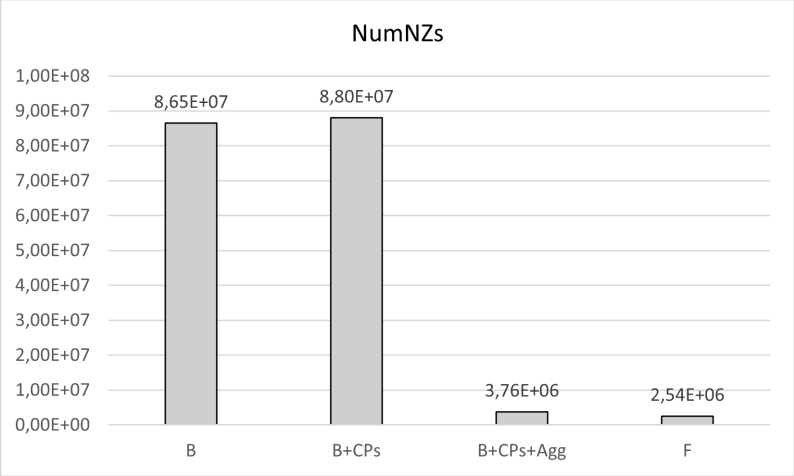

Figure 5.10: Average number of non-zeros in the tested formulations (see Table 5.6).

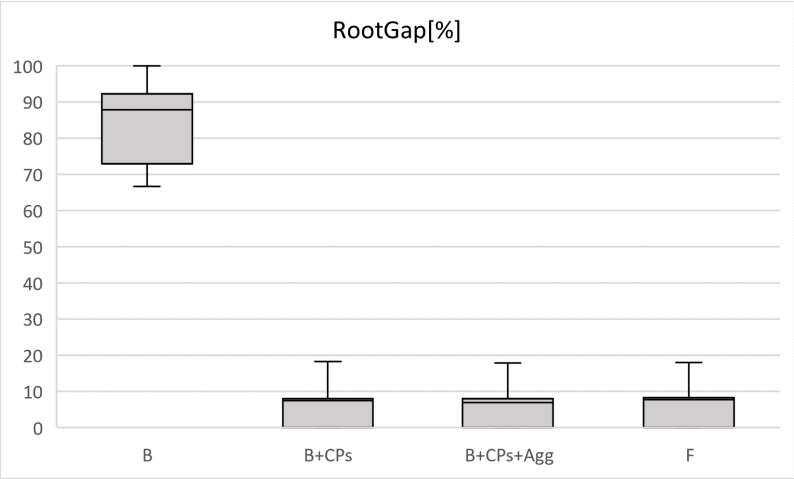

Figure 5.11: Box plots of the relative gap at the root node in the tested formulations (see Table 5.6).

formulation F is characterized by fewer non-zeros and also by fewer variables and constraints.

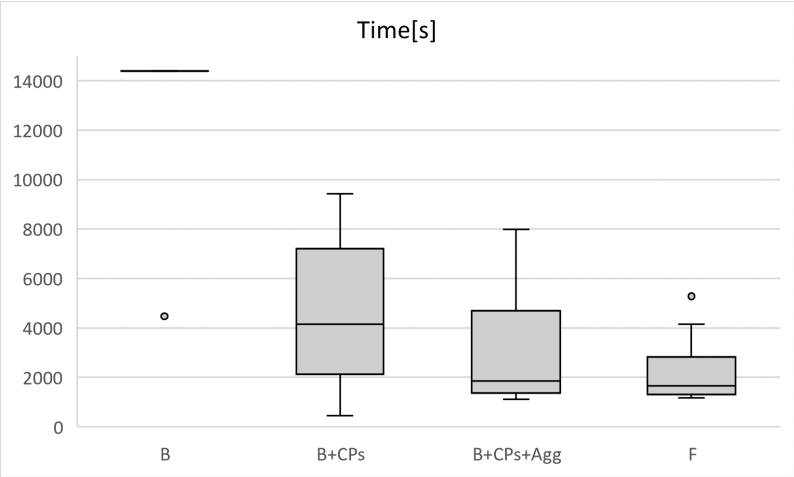

Figure 5.12: Box plots of the computational times in the tested formulations (see Table 5.6).

A comparison between formulation B and formulation F is reported in Table 5.7. The evaluation metrics considered are: number of variables (NumVars), constraints (NumConstrs) and non-zeros (NumNZs); relative optimality gap at the root node (RootGap), final relative optimality gap (Gap); number of explored nodes (Nodes); total solution time (Time). Denoting the optimal value with Opt, the value of the lower bound at the root node with RootLB, and the value of the best lower bound with BestLB, the RootGap is computed as 100(Opt-RootLB)/Opt, and the Gap as 100(Opt-BestLB)/Opt. We reported in bold the gap of instances solved to optimality within the time limit.

Results show that formulation F is definitely reduced in size and sparser than B. This is mainly due to the aggregation of the constraints and the coefficient tightening, as previously discussed. The quality of the bounds at the root node is better in F, mainly due to the addition of cutting planes. In some instances (e.g., BOF6, BOF8–BOF10), the problem formulated with F is even solved at the root node. Thanks to a good root bound, the number of the explored nodes is heavily reduced in F, and consequently, the time spent on the branch-and-cut tree search. Overall, solution times are significantly reduced if we compare F to B. Indeed, only BOF4 has been successfully solved within the time limit of 4 hours using formulation B, whereas each instance has been solved to optimality in less than 1 hour and 30 minutes using formulation F.
In the end, formulation F turns out to be much more competitive than the basic formulation B.

We refer the reader to Appendix A for the figures that show the assignment of testpoints to each activated base station according to the optimal solutions found in the instances.

Table 5.7: Optimization results for the instances BOF1–BOF10 (see Table 5.5): a comparison between the basic (B) and the final (F) formulations. Time is expressed in seconds.

Instance	Formulation	NumVars	NumConstrs	NumNZs	RootGap[%]	Gap[%]	Nodes	Time[s]
BOF1	**B**	633690	638248	86210410	100.00	61.54	14825	TL
	F	409434	427433	2102620	7.55	**0.00**	35	5283.24
BOF2	**B**	633690	638248	86210410	92.31	76.92	29755	TL
	F	409434	429434	2115448	7.92	**0.00**	79	1246.83
BOF3	**B**	633690	638248	86210410	92.31	84.62	14964	TL
	F	409434	430889	2124711	8.01	**0.00**	80	1682.48
BOF4	**B**	633690	638248	86210410	92.31	**0.00**	4114	4470.41
	F	433907	409434	2143364	7.88	**0.00**	124	1164.48
BOF5	**B**	633690	638249	86711935	90.00	90.00	29208	TL
	F	409434	436297	2481646	8.87	**0.00**	41	1907.63
BOF6	**B**	633690	638249	86711935	85.71	85.71	30235	TL
	F	409434	443191	2520896	0.00	**0.00**	1	1618.57
BOF7	**B**	633690	638249	86711935	80.00	80.00	29993	TL
	F	409434	462855	2630988	17.97	**0.00**	148	2399.62
BOF8	**B**	633690	638249	86711935	75.00	75.00	30075	TL
	F	409434	498128	2859026	0.00	**0.00**	1	1324.79
BOF9	**B**	633690	638249	86711935	66.67	66.67	44649	TL
	F	409434	526184	3066771	0.00	**0.00**	1	1594.35
BOF10	**B**	633690	638249	86711935	66.67	66.67	44492	TL
	F	409434	558474	3315367	0.00	**0.00**	1	4144.37

TL, time limit reached

5.3 Critical analysis of the model

In the previous sections, we discussed how to accelerate the solution of WND problems requiring service and coverage requirements. In this chapter, we take under consideration the modeling of an additional requirement which is usually not taken into account in the literature, but could have a huge impact on the optimal solution of the problem: the (finite) capacity of the transmitters. We also report the modeling of another phenomenon, the electromagnetic field produced by the activation of the antennas.

We only discuss the modeling of these requirements, without presenting any computational experiments for lack of data. However, we believe that the modeling is itself interesting for the reader, and represents a first step towards a future study that would take under consideration these aspects.

Capacity constraints

Capacity constraints can be used to enforce a maximum number of users per transmitter and are based on a target user data rate. In particular, by assuming that the users have a target data rate (DR) that they aim to achieve in downlink, we can obtain the maximum number of users u that any transmitter can serve simultaneously as (see [83])

$$u = \lfloor \frac{B \times SE}{DR} \rfloor$$

in which the numerator is the cell capacity, given by the product between the bandwidth B and the downlink spectral efficiency SE.

Capacity constraints can be formulated as

$$\epsilon \sum_{t \in \mathcal{T}} d_t x_{tb} \leq u \qquad b \in \mathcal{B}$$

where the number of users d_t inside testpoints $t \in \mathcal{T}$ that are served by transmitter $b \in \mathcal{B}$ must be not greater than the number of users u that b can serve at the same time. The parameter d_t can be extracted from population/traffic data in t, whereas $\epsilon \in [0, 1]$ is a penetration factor of the technology[1].

Capacity constraints can be tightened by adding the variable z on the right hand side, namely

$$\epsilon \sum_{t \in \mathcal{T}} d_t x_{tb} \leq u \sum_{l \in \mathcal{L}} z_{bl} \qquad b \in \mathcal{B}.$$

For the fixed-power case, they become

$$\epsilon \sum_{t \in \mathcal{T}} d_t x_{tb} \leq u z_b \qquad b \in \mathcal{B}.$$

Electromagnetic field constraints

Another aspect we did not consider, that was recently studied in the literature, concerns the electromagnetic field emissions (EMF). A measure used to evaluate the EMF emissions is the power density (see [41]). To guarantee that the overall power density measured in a testpoint $t \in \mathcal{T}$ is below the limit l_t, we can use the so-called *EMF constraints*, introduced in [41]. These constraints are defined over all the testpoints that are not falling in the exclusion zones of the installed base stations: indeed these testpoints are not subject to the EMF limits defined for the general public as they are too close to the activated transmitters.

[1] E.g., $\epsilon = 0.3$ means that we estimate that only 30% of the population makes use of that technology.

We focus on the fixed-power case first. To identify the testpoints that must be excluded, the binary variables w_t are introduced. In particular, w_t is defined as follows

$$w_t \geq I_{tb}^{zone} z_b \qquad t \in \mathcal{T}, b \in \mathcal{B}$$

$$w_t \leq \sum_{b \in \mathcal{B}} I_{tb}^{zone} z_b \qquad t \in \mathcal{T}$$

where I_{tb}^{zone} is a binary input parameter taking value 1 if testpoint t is inside the exclusion zone of transmitter b. This means that w_t is activated if testpoint t is inside at least one exclusion zone and it is deactivated if t is outside the exclusion zones of all installed transmitters. The EMF constraints are then defined as follows

$$\tilde{p}_t^{base}(1 - w_t) + \sum_{b \in \mathcal{B}} \tilde{p}_{tb}^{add}(1 - w_t) z_b \leq l_t \qquad t \in \mathcal{T}. \tag{5.18}$$

In (5.18), the power density is modeled through the coefficients $\tilde{p}_{tb}^{add}$ and $\tilde{p}_t^{base}$. In particular, the total power density is given by the contributions from the newly installed base stations ($\sum_{b \in \mathcal{B}} \tilde{p}_{tb}^{add} z_b$) as well as the ones from the already installed base stations ($\tilde{p}_t^{base}$). The limit l_t established by the National law could depend on the characteristics of the testpoint t: e.g., the limits on residential areas are different from those on general public areas. The term $(1 - w_t)$ is used in such a way that a testpoint t falling inside the exclusion zone of an installed base station is not considered when the power density is evaluated against the limits.

Note that constraints (5.18) are nonlinear, therefore the authors of [41] linearized them through the McCormick envelopes. Namely, they introduce the binary variables y_{tb} defined as

$$y_{tb} \leq 1 - w_t \qquad t \in \mathcal{T}, b \in \mathcal{B}$$

$$y_{tb} \leq z_b \qquad t \in \mathcal{T}, b \in \mathcal{B}$$

$$y_{tb} \geq z_b - w_t \qquad t \in \mathcal{T}, b \in \mathcal{B}$$

to replace the nonlinear product $(1 - w_t) z_b$ in (5.18). Hence, the EMF constraints become

$$\tilde{p}_t^{base}(1 - w_t) + \sum_{b \in \mathcal{B}} \tilde{p}_{tb}^{add} y_{tb} \leq l_t \qquad t \in \mathcal{T}.$$

For the variable-power case, we can simply replace z_b with $\sum_{l \in L} z_{bl}$.

5.4 Conclusions

The design of wireless networks is a typical problem in the telecommunications sector, with relevant practical applications. Due to the increasing size of the new generation networks, co-existing in an extremely congested radio spectrum and subject to local and international constraints, determining suitable transmitter locations and power emissions has become an increasingly challenging task. Traditional design methods employed by practitioners, relying on trial-and-error supported by simulation, have demonstrated various limitations. Therefore, optimization approaches have become indispensable for cost reduction and meeting user-demanded service quality standards.
Since the early 1980s, several optimization models have been developed for designing wireless networks. However, the natural formulation on which most models are based presents severe limitations since it involves numerical issues in the problem-solving phase, which emerge even in small instances, preventing the solution of instances of practical interest. Indeed, the constraint matrices of these models contain coefficients that range in a huge interval, as well as large big-M leading to weak bounds.

This chapter addresses two primary wireless network design problems: (i) the site and the power assignment problem, where power is allocated among a set of discrete values, and (ii) the base station deployment problem, assuming fixed-power emissions.

For the variable-power case, we worked on improving the natural formulation proposed in the literature by introducing several presolve operations to reduce the number of problem variables and overall problem size. Additionally, we implemented cliques and variable upper bounds to cut off fractional solutions, accelerating solution times. Furthermore, we proposed an aggressive reduction scheme based on a reduced cost fixing procedure that leads to the reduction of the big-M values, strengthening the formulation and reducing the problem size.

For the fixed-power case, we intervened in the natural formulation of the problem by carrying out both strengthening operations (i.e., additions of variable upper bounds and cliques, variable replacement, big-M reduction) and size reduction operations (i.e., constraints aggregation, heuristic coefficient tightening).

The reformulation and operations we introduced enabled us to efficiently solve large instances of the problem to optimality, in solution times consistent with planning operations, for both the fixed and the variable-power case. All the tests were conducted on realistic LTE data from the Municipality of Bologna in Italy.

We remark that the reduced cost fixing procedure – employed in the variable-power

case – should only be used when proper upper bounds can be computed quickly: in this regard, our fixing heuristic played a crucial role in the success of this approach. However, finding an upper bound is not always trivial, especially when high-quality service is required. Indeed, we could test our reduced cost fixing approach only on instances requiring low to medium service.

To conclude, this chapter has provided a comprehensive exploration of the wireless network design problem from a modeling perspective. Emphasis was placed on improving the natural formulation of the problem with the final scope of tackling real-world scenarios. It is noteworthy to mention that the scope of this research did not delve deeply into ad hoc methods for solving the problem, with the exception of reduced cost fixing – which was found to be an efficient methodology only in a limited application area. This approach was taken in line with the specifications from the FUB, which sought solutions that would be straightforward for practitioners (namely, telecommunication operators) to implement. While this constraint offered a clear direction for our current research, future work could certainly expand upon this by investigating specialized methods for solving the wireless network design problem.

In future work, we also plan to investigate how the inclusion of capacity constraints in the problem formulation affects the optimal solution of the problem. In fact, even if capacity constraints are usually not taken into consideration in the literature, we believe they can actually influence the topology of the solution, and consequently the coverage achieved. We also intend to consider the modeling of the electromagnetic field, produced by the activation of the base stations, to determine if the solution complies with National law limits and, if not, to assess how these limits impact solution topology.

Chapter 6

The congested partial set covering location problem

6.1 Introduction

In Chapter 4 we discussed the congested facility location problem and the partial set covering location problem. The former was introduced in 1995 by Desrochers et al. [54] to prevent the overloading effect of the facilities and induce better resource allocation. The latter was introduced in 1999 by Daskin and Owen [51] to mitigate the drawbacks of the well-known set covering location problem, yielding to costly and impractical solutions due to its full coverage requirement. Despite their practical relevance, both the PSCLP and the CFLP have received very little attention in the scientific literature. However, recent works [45, 68, 69] showed promising results given by exact methods based on Benders decomposition for the deterministic PSCLP and CFLP. Moreover, the literature on PSCLP gave no consideration to the inherent volatility of the parameters used in the mathematical model and the, consequently, unreliable deterministic solution. Indeed, uncertainty has been investigated in another variant of set covering location, that is the maximal covering location problem, but not specifically on the PSCLP. For all these reasons, we address the PSCLP in a new quadratic and robust variant, which considers the minimization of the congestion at the facilities and the protection against the changes in demand.

Hence, the problem we introduce in this chapter is a novel problem that we denote as the *congested partial set covering location problem* (CPSCLP), a variant of the facility location problem that may arise in the context of designing telecommunication networks, or (more in general) in service and communications networks. The problem consists of choosing where to locate the facilities among the candidate sites that can satisfy a (partial)

target demand, in such a way as to minimize the cost of facility opening and congestion.

This problem can be seen as a telecommunication problem in which there is the need to place transmitters (the facilities), minimizing the cost of the infrastructure and the congestion at the transmitters. In fact, the more the demand associated with the transmitter, the less the quality of service (i.e., the data rate) given to the users.

We study the CPSCLP under the assumption of uncertain customer demand. In order to create a reliable and efficient network architecture that is robust against demand changes, we propose to deal with data uncertainty using the approach known as Γ-robustness (see [27]). For an introduction to the Γ-robustness and optimization under uncertainty, we refer the reader to Chapter 2.2. The basis of this approach is that we assume that nature is restricted in its behaviour, in the sense that at most Γ customers are supposed to deviate from their expected demand and adversely affect the feasibility and the optimality of the deterministic solution. Uncertainty in customer demand is quite realistic as in general the demand comes out from an estimate due to its inner volatility or a lack of historical data.

In the following sections, we formulate the deterministic problem and derive its robust counterpart, which consists of optimizing against the worst-case scenario. We then reformulate the robust counterpart as a mixed-integer second-order cone program using the perspective reformulation and finally address its solution with Benders decomposition-based approaches.

The content of this chapter is part of a paper [35] with Professors Ivana Ljubic from ESSEC Business School of Paris and Laura Palagi from Sapienza University of Rome. The paper has been recently submitted for publication in a journal.

6.2 The deterministic problem

We are given a set I of potential facility locations with opening cost $f_i \geq 0$ for $i \in I$, and a set J of customer locations such that each customer $j \in J$ is associated with a demand $d_j \geq 0$. For each customer j, we are also given a subset $I(j) \subseteq I$ of facility locations that can "cover" j, i.e., that can fully serve the demand d_j. Similarly, let $J(N) \subseteq J$ for a subset of facilities $N \subseteq I$, be the subset of customers that can be covered by N, and let $J(i) = J(\{i\})$, for $i \in I$.

Given a parameter $0 < D \leq \sum_{j \in J} d_j$, CPSCLP aims at identifying a subset of facilities to open in order to ensure that the total served customer demand is at least D, while minimizing the overall costs, given by facility opening expenses and congestion cost.

For each $i \in I$, the binary variable y_i is set to one if facility i is open, and to zero

otherwise. For each $i \in I$ and $j \in J$, the continuous allocation variable $x_{ij} \geq 0$ denotes the fraction of the demand of customer $j \in J$ served by facility $i \in I$.

We observe that we chose to use fractional x_{ij} motivated by the telecommunication context. Indeed, if we consider the customer as a representative user – namely a sort of superuser that represents multiple users, just like in wireless network design – it could be useful to allow for fractional demand satisfaction, i.e. for fractional x_{ij}. This is because if we only allow for full demand satisfaction or nothing ($x_{ij} \in \{0, 1\}$), we may end up with certain superusers that are not covered at all because their demand cannot be satisfied in full. Instead, allowing fractional x_{ij}, part of a superuser demand can be served too. Furthermore, the demand from a superuser can be split among multiple transmitters (every transmitter may serve only some users inside the superuser), which is amenable when we minimize congestion costs.

To model congestion cost, we follow what is done in the literature of CFLP [54, 68], in which to penalize each additional unit of demand served by a given facility, convex quadratic cost functions are used. Specifically, we introduce the auxiliary aggregated variables $v_i \geq 0$ for each $i \in I$, defined as

$$v_i = \sum_{j \in J} d_j x_{ij}$$

denoting the total demand served by facility i, also known as *facility load*. Then, given a function $F(\cdot)$, which is assumed to be non-negative, continuous and convex for non-negative arguments, the congestion cost at a facility $i \in I$ is given by $F(v_i)v_i$. We can assume that $F(t)$ is a linear function, i.e., given the non-negative input coefficients a and b, we have that $F(t) = at + b$.

The problem at hand can be modeled as follows

$$\min \quad \sum_{i \in I} f_i y_i + b \sum_{i \in I} v_i + a \sum_{i \in I} v_i^2 \tag{6.1}$$

$$s.t. \quad \sum_{i \in I} \sum_{j \in J} d_j x_{ij} \geq D \tag{6.2}$$

$$v_i = \sum_{j \in J} d_j x_{ij} \qquad i \in I \tag{6.3}$$

$$\sum_{i \in I(j)} x_{ij} \leq 1 \qquad j \in J \tag{6.4}$$

$$0 \leq x_{ij} \leq y_i \qquad j \in J, i \in I(j) \tag{6.5}$$

$$y_i \in \{0, 1\} \qquad i \in I.$$

The objective (6.1) aims to minimize the sum of the facility opening cost and the congestion cost. Constraint (6.2) ensures the total customer demand served is at least D. Constraints (6.3) define the facility load. Assignment constraints (6.4) make sure that the fraction of covered demand of a customer does not exceed the unit. Finally, constraints (6.5) ensure that allocation to a facility is only possible if it is open.

We can observe that: (i) the problem is a mixed-integer non-linear program, (ii) the constraints are all linear, (iii) the objective function is separable, and is made of a linear component in y and a convex quadratic component in v, (iv) the problem is NP-hard since it is a generalization of the traditional set covering location problem, which is NP-hard [51].

6.3 The robust counterpart of the problem

In real-world scenarios, customer demand often varies or is difficult to estimate. To capture this uncertainty, we assume that each entry d_j, $j \in J$ of the vector of demand d is modeled as an independent, symmetric and bounded random variable (with unknown distribution) $\tilde{d}_j$, $j \in J$ that takes values in $[d_j - \hat{d}_j, d_j + \hat{d}_j]$. We allow the possibility that the deviations $\hat{d}_j$ from the nominal coefficients d_j could also be zeros, i.e. that $\hat{d}_j = 0$ for some $j \in J$. We adopt the notion of protection introduced by Bertsimas and Sim in [27] known as Γ-*robustness*, which assumes that only a subset of the coefficients of d will deviate from their nominal values, adversely affecting the solution. Hence, we introduce an integer number Γ, taking values in the interval $[0, |J|]$, that limits the number of demand deviations. Parameter Γ controls the level of robustness against the solution: if $\Gamma = 0$, we completely ignore the uncertainty (deterministic setting), while if $\Gamma = |J|$, we are considering all possible demand deviations (which is the most conservative strategy).

We note that the vector of demand d is involved in congestion and coverage affecting both optimality and feasibility. By protecting against the uncertainty, we mean that we are interested in finding an optimal solution that:

1. optimizes against all scenarios under which up to Γ demand coefficients can vary in such a way as to maximally influence the objective; the worst-case scenario is given by increasing demand;

2. is protected against all cases in which up to Γ demand coefficients change affecting the feasibility; the worst-case scenario is given by decreasing demand.

We observe that the two worst-case realizations play against each other.

We now introduce a *robust counterpart of the problem*, that optimizes against the worst-case realizations under demand uncertainty and reads

$$\min \quad \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} v_i^2 \tag{6.6}$$

$$s.t. \quad v_i = \sum_{j\in J} d_j x_{ij} + \max_{\{S|S\subseteq J,|S|\leq\Gamma\}} \left\{\sum_{j\in S} \hat{d}_j x_{ij}\right\} \qquad i\in I \tag{6.7}$$

$$\sum_{i\in I}\sum_{j\in J} d_j x_{ij} - \max_{\{S|S\subseteq J,|S|\leq\Gamma\}} \left\{\sum_{i\in I}\sum_{j\in S} \hat{d}_j x_{ij}\right\} \geq D \tag{6.8}$$

$$\sum_{i\in I(j)} x_{ij} \leq 1 \qquad j\in J \tag{6.9}$$

$$0 \leq x_{ij} \leq y_i \qquad j\in J, i\in I(j) \tag{6.10}$$

$$y_i \in \{0,1\} \qquad i\in I \tag{6.11}$$

where, for a given Γ and allocation choice x_{ij}, the load v_i (and consequently the objective function) is now taking into account also the sum of the Γ largest deviations in case the demand is increasing from the nominal value, whereas in the covering constraint we are considering the sum of the Γ largest deviations in case the demand is decreasing from the nominal value. We observe that we can relax constraints (6.7) replacing the equalities with inequalities of type $\geq$. By applying strong duality to the inner maximization terms in (6.7) and (6.8), we can derive the following equivalent *convex* MIQP model of the robust counterpart

$$\min_{(\tau,\rho,\pi,\sigma)\geq 0,\, y\in\{0,1\}^{|I|}} \quad \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} v_i^2 \tag{6.12}$$

$$s.t. \quad v_i - \sum_{j\in J} d_j x_{ij} - \left(\Gamma\rho_i + \sum_{j\in J}\sigma_{ij}\right) \geq 0 \qquad i\in I \tag{6.13}$$

$$\sum_{i\in I}\sum_{j\in J} d_j x_{ij} - \left(\Gamma\tau + \sum_{j\in J}\pi_j\right) \geq D \tag{6.14}$$

$$\tau + \pi_j \geq \sum_{i\in I} \hat{d}_j x_{ij} \qquad j\in J \tag{6.15}$$

$$\rho_i + \sigma_{ij} \geq \hat{d}_j x_{ij} \qquad j\in J, i\in I \tag{6.16}$$

$$\sum_{i\in I} x_{ij} \leq 1 \qquad j\in J \tag{6.17}$$

$$0 \leq x_{ij} \leq y_i \qquad j\in J, i\in I(j) \tag{6.18}$$

which is a much more computationally tractable formulation, having a (separable) convex quadratic objective function and all linear constraints. We denote formulation (6.13)-(6.18) as the *extended formulation*.

Theorem 6.3.1 *The extended formulation* (6.13)-(6.18) *is equivalent to* (6.7)-(6.11).

Proof We show how to get formulation (6.13)-(6.18), following what is done in Theorem 1 of [27]. Given $i \in I$ and a vector $x_i^\star$, we define the protection function α representing the sum of the Γ largest deviations as

$$\alpha(x_i^\star) = \max_{\{S|S\subseteq J, |S|\leq\Gamma\}} \left\{ \sum_{j\in S} \hat{d}_j x_{ij}^\star \right\}.$$

By introducing the binary variables w_{ij}, the protection function α can be equivalently formulated as

$$\begin{aligned} \alpha(x_i^\star) = \max \quad & \sum_{j\in J} \hat{d}_j x_{ij}^\star w_{ij} \\ s.t. \quad & \sum_{j\in J} w_{ij} \leq \Gamma \\ & w_{ij} \in \{0,1\} \qquad j \in J \end{aligned}$$

which equals to

$$\begin{aligned} \alpha(x_i^\star) = \max \quad & \sum_{j\in J} \hat{d}_j x_{ij}^\star w_{ij} \\ s.t. \quad & \sum_{j\in J} w_{ij} \leq \Gamma & (\rho_i) \\ & 0 \leq w_{ij} \leq 1 \qquad j \in J \qquad (\sigma_{ij}) \end{aligned} \tag{6.19}$$

where we relaxed the integrality of variables w_{ij} since the matrix of the constraints is totally unimodular and Γ is integral (linear programs of this form have integral optima). Clearly the optimal solution of Problem (6.19) consists of Γ variables at 1. We next consider the dual of Problem (6.19)

$$\begin{aligned} \min \quad & \Gamma\rho_i + \sum_{j\in J} \sigma_{ij} \\ s.t. \quad & \rho_i + \sigma_{ij} \geq \hat{d}_j x_{ij}^\star \qquad j \in J \\ & \sigma_{ij} \geq 0 \qquad\qquad\qquad j \in J \\ & \rho_i \geq 0. \end{aligned} \tag{6.20}$$

By strong duality, since Problem (6.19) is feasible and bounded for all $\Gamma \in [0, |J|]$, then the dual problem (6.20) is also feasible and bounded and their objective functions assume the same value in an optimal solution. Hence, $\alpha(x_i^\star)$ is equal to the optimal value of Problem (6.20) for each $i \in I$.

Then, given a vector $x^\star$, we define the protection function β representing the sum of the Γ largest deviations as:

$$\beta(x^\star) = \max_{\{S|S\subseteq J, |S|\leq\Gamma\}} \left\{ \sum_{i\in I} \sum_{j\in S} \hat{d}_j x_{ij}^\star \right\}.$$

By introducing the binary variables w_{ij}, the protection function β can be equivalently formulated as

$$\begin{aligned} \beta(x^\star) = \max \quad & \sum_{i\in I} \sum_{j\in J} \hat{d}_j x_{ij}^\star w_j \\ s.t. \quad & \sum_{j\in J} w_j \leq \Gamma \\ & w_j \in \{0,1\} \qquad\qquad j \in J \end{aligned}$$

which equals to

$$\begin{aligned} \beta(x^\star) = \max \quad & \sum_{i\in I} \sum_{j\in J} \hat{d}_j x_{ij}^\star w_j \\ s.t. \quad & \sum_{j\in J} w_j \leq \Gamma \qquad\qquad (\tau) \\ & 0 \leq w_j \leq 1 \qquad j \in J \qquad (\pi_j) \end{aligned} \tag{6.21}$$

where we relaxed the integrality of variables w_{ij} since the matrix of the constraints is totally unimodular and Γ is integral (linear programs of this form have integral optima). Clearly the optimal solution of Problem (6.21) consists of Γ variables at 1. We next consider the

dual of Problem (6.21)

$$\begin{aligned} \min \quad & \Gamma\tau + \sum_{j\in J} \pi_j \\ s.t. \quad & \tau + \pi_j \geq \sum_{i\in I} \hat{d}_j x^\star_{ij} \qquad j \in J \\ & \pi_j \geq 0 \qquad\qquad\qquad\quad j \in J \\ & \tau \geq 0. \end{aligned} \tag{6.22}$$

By strong duality, since Problem (6.21) is feasible and bounded for all $\Gamma \in [0, |J|]$, then the dual problem (6.22) is also feasible and bounded and their objective functions assume the same value in an optimal solution. Hence, $\beta(x^\star)$ is equal to the optimal value of Problem (6.22).

□

Interpretation of the variables ρ, σ, τ, π

We can give the following interpretation to the dual variables ρ, σ, τ, π of the convex reformulation of the robust counterpart of the problem. For instance, let us consider the protection function β used in the covering constraint. When we go to the dual, we replace the primal objective representing the sum of the Γ largest deviations with the sum of the components of the dual variables, namely with $\Gamma\tau + \sum_{j\in J} \pi_j$.

Suppose to sort all the deviations $\hat{d}_j \sum_{i\in I} x^\star_{ij}$ for a given allocation $x^\star_{ij}$ from the largest to the smallest. We then consider only the Γ largest deviations, hence the first Γ. Now, we fix τ to the Γ-th (sorted) deviation, and π_j to what remains of the j-th deviation if we subtract τ, as depicted in Figure 6.1. Then, the sum of the components of the dual variables represents exactly the sum of the Γ largest deviations and the meaning of each dual variable is explained.

Similarly, we can derive the meaning of the other dual variables ρ representing the Γ-th deviation and σ representing what remains of the deviation once we subtract ρ.

An example

Suppose we are given a problem having five possible locations for the facilities ($|I| = 5$) and six customers ($|J| = 6$). In Figure 6.2 we represent this problem using the white vertices for the facilities and the grey vertices for the customers. The dotted lines indicate

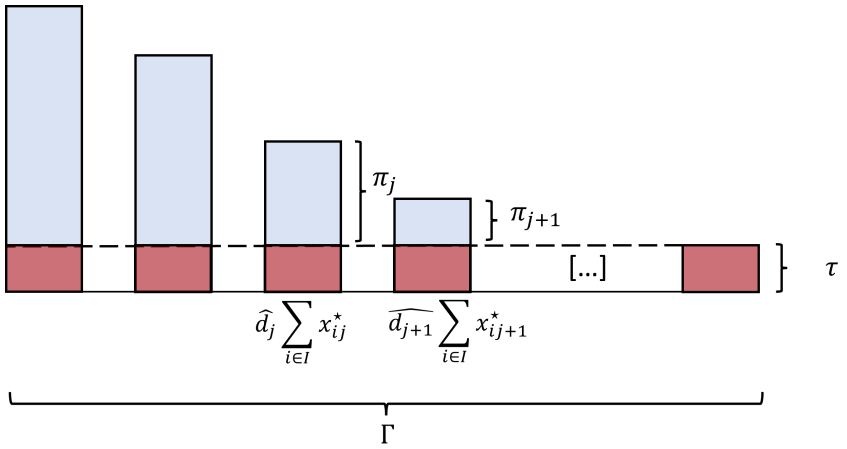

Figure 6.1: Meaning of dual variables in the covering constraint of the robust congested partial set covering location problem.

potential links between facilities and customers; the absence of a dotted line between a facility i and a customer j implies that facility i cannot serve customer j (for example because the distance between them is greater than the given radius of coverage). We set

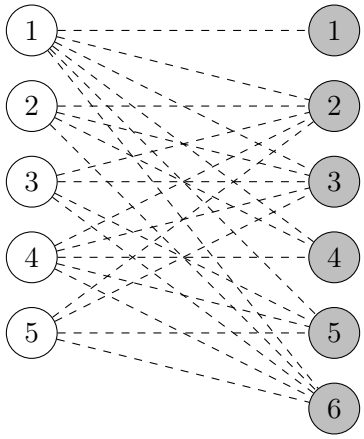

Figure 6.2: Possible links between facilities (white vertices) and customers (grey vertices).

parameter $\Gamma = 4$, implying that at most four customers (out of six) will deviate from their nominal demand, representing a pretty conservative scenario. The costs for opening the facilities are given by $f = [27 \quad 34 \quad 50 \quad 80 \quad 79]^T$. The nominal demands are given by $d = [92 \quad 15 \quad 71 \quad 51 \quad 59 \quad 50]^T$, and the deviations from the nominal demands are $\hat{d} = [17 \quad 3 \quad 0 \quad 6 \quad 13 \quad 15]^T$. The target coverage is set at $D = 70\% \sum_{j \in J} d_j$.

Figures 6.3-6.4 present the optimal solutions under deterministic and robust settings for the given value of Γ. The robust solutions are obtained considering three different cases: protection against uncertainty only in congestion, only in coverage, and in both congestion and coverage.

In the case we consider an increasing demand affecting the load, what happens is that we face more congestion cost and the optimal value ($ObjVal$) increases. If we instead consider

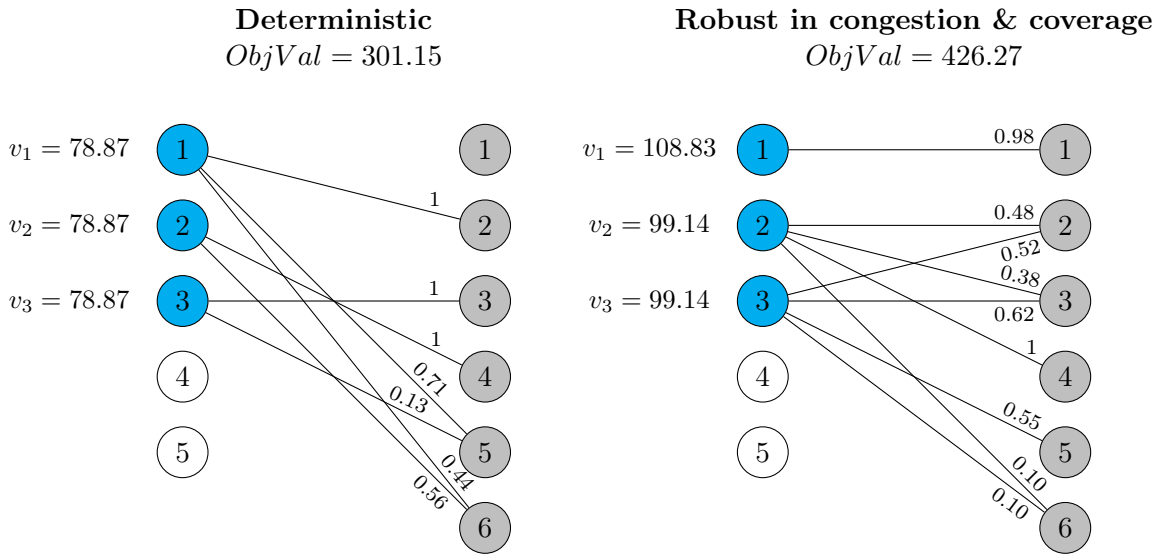

Figure 6.3: Optimal solution and value ($ObjVal$) of the deterministic case on the left. Optimal solution and value ($ObjVal$) of the robust case in which we protect against uncertainty in both congestion and coverage on the right. Open (closed) facilities are represented by the light-blue (white) vertices, customers by the grey vertices. The value v_i next to facility i is its load, whereas the value on the line between i and j is the fraction of demand x_{ij} of costumer j served by facility i.

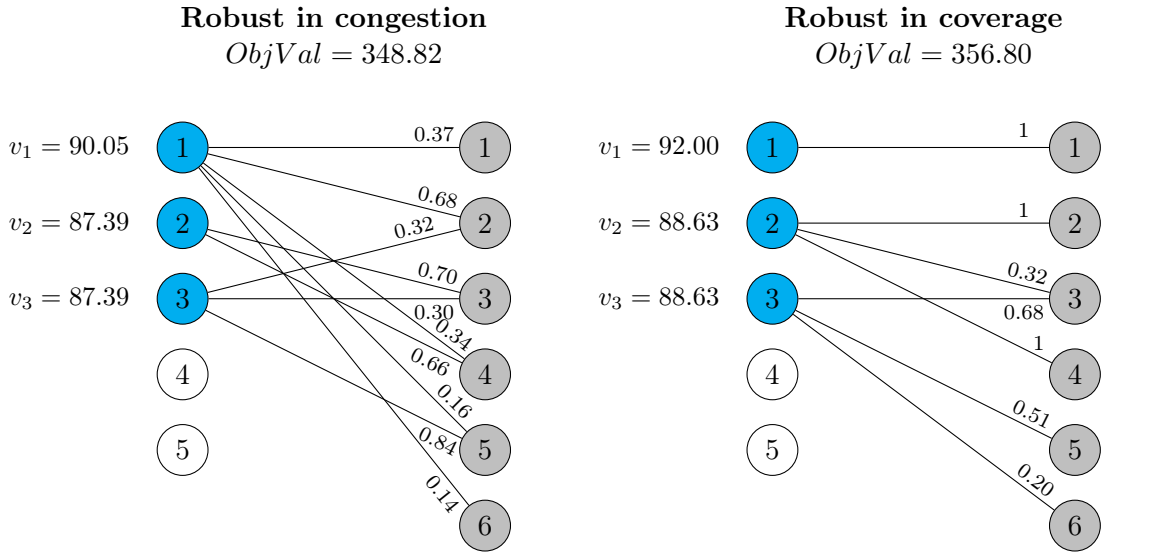

Figure 6.4: Optimal solution and value ($ObjVal$) of the robust case in which we protect against uncertainty in congestion on the left. Optimal solution and value ($ObjVal$) of the robust case in which we protect against uncertainty in coverage on the right. Open (closed) facilities are represented by the light-blue (white) vertices, customers by the grey vertices. The value v_i next to facility i is its load, whereas the value on the line between i and j is the fraction of demand x_{ij} of costumer j served by facility i.

a decreasing demand affecting coverage, solving the robust counterpart of the problem leads to covering a demand higher than the target demand, increasing the facility loads and again congestion cost. When we consider both uncertainties, we end up covering more than D and considering more load, with a double impact on congestion cost. Additionally, the robust model differentiates between customers whose deviation from the nominal demand

could highly affect the coverage and/or the congestion and customers who are not affecting the deterministic solution at all. Hence, there is this extra information in the model, which may influence the distribution of the demand among the facilities and the coverage among the customers. Remember that we are looking for partial set covering, and not full coverage.

6.3.1 Perspective reformulation of the robust counterpart

The extended formulation of the robust counterpart of the problem belongs to the the class of convex MINLPs having "off/on" decisions forcing the continuous variables either to be 0 or to be in a convex set. A very effective reformulation technique that can be applied to such situations is the so-called *perspective reformulation*, described in [74, 86] and in Chapter 2.1.1, which is based on the property that the perspective of a convex function is itself convex; this property can be used to construct a *tighter* reformulation.

To derive the perspective reformulation, we introduce the non-negative variables u_i for $i \in I$, and use them to replace the quadratic term v_i^2 in the objective function. We thus impose that $v_i^2 \leq u_i$, and replace the convex function v_i^2 with its perspective defined by $y_i(v_i/y_i)^2$ if $y_i > 0$, and zero if $y_i = 0$. Then, the *perspective reformulation* of the robust counterpart of the problem reads as follows

$$\begin{aligned} \min_{(u,v,\tau,\rho,\pi,\sigma)\geq 0,\, y\in\{0,1\}^{|I|}} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} u_i & \\ s.t. \quad & v_i^2 \leq u_i y_i & i \in I \qquad (6.23)\\ & (6.13) - (6.18) & \end{aligned}$$

where constraints (6.23) are rotated second-order cone (SOC) constraints, imposing that the quadratic load v_i^2 is 0 when facility i is closed ($y_i = 0$), and it is equal to u_i when facility i is open ($y_i = 1$). The perspective reformulation of the problem is still convex as the right-hand side of the SOC is the product of non-negative variables.

It is well-known (see e.g., [8, 72, 73, 75, 84, 86]) that the continuous relaxation of a perspective reformulation produces stronger bounds than the bounds given by the continuous relaxation of the original formulation. Consequently, we can say that the original formulation (6.13)-(6.18) lies in an extended space compared to the perspective reformulation, and for this reason, we called it the "extended" formulation.

6.4 A Benders decomposition approach

The size of the formulation of the robust counterpart depends on the number of customers and facility locations. In real world context, the number of customers can be quite large, affecting the size and so the solution time of this type of problem. In this regard, we propose to use *Benders decomposition* (for an introduction to this solution method, see Chapter 2.1.3). Our idea is to reduce as much as possible the number of variables in the problem, by projecting out of the master problem (at least) all the variables that depend on the number of customers, i.e., x and σ. This will lead to a boosting in the solution time, as we will see in the computational results Section 6.6.

In this section, we investigate the use of two different separations, both leading to a convex MISOCP master problem containing SOC constraints and the complete objective, and an LP subproblem. The advantage of keeping SOC constraints in the master formulation is that the master problem has a very tight relaxation, which is amenable when dealing with Benders decomposition as it reduces solution times incredibly. We also observe that the subproblem – which is solved several times to generate valid cuts for the master – is an LP that can be therefore efficiently solved by any state-of-the-art solver.

First choice of separation

The first Benders implementation considers the separation of the continuous variables (x, π, σ). The master problem is then defined on the variables (y, v, u, ρ, τ) and is given by

$$\begin{aligned} \min_{(u,\rho,\tau)\geq 0,\, y\in\{0,1\}^{|I|}} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} u_i & & \\ s.t. \quad & \phi(y, v, \rho, \tau) \geq 0 & & \\ & v_i \geq \Gamma\rho_i & i \in I & \qquad (6.24) \\ & v_i^2 - u_i y_i \leq 0 & i \in I & \end{aligned}$$

where $\phi(y, v, \rho, \tau)$ refers to the convex function expressing the optimal solution value of the subproblem as a function of (y, v, ρ, τ), and it can be approximated by linear cuts to be generated on the fly, known as Benders cuts, that are valid for any given vector (y, v, ρ, τ). Note that there are two types of Benders cuts: the optimality cuts and the feasibility cuts. We will generate only Benders feasibility cuts. Indeed, there is no need for Benders optimality cuts in our approach, as the only variables appearing in the objective function (y, v, u) belong to the master problem. Hence, we will only need Benders feasibility cuts to

discard infeasible points.

Constraints (6.24) are lower bounds to the variable v used to strengthen the formulation and derived from constraint (6.13), which appears in the subproblem.

The number of Benders feasibility cuts is exponential, which makes it impractical to enumerate all of them in advance. Since only some of them are necessary to find an optimal solution, they will be dynamically separated by the decomposition approach. Consequently, the master formulation will contain only a subset of Benders cuts, belonging to a so-called relaxed master problem.

Our master problem is a MISOCP, thus a convex MINLP that can be solved as an MILP by a branch-and-cut approach where the integrality requirement on y is relaxed and linear outer-approximations of the SOC constraints are generated on the fly, or it can be solved as an MINLP by a branch-and-cut approach relying on an NLP solver for the solution of the continuous problem at every node.

Benders cuts can be computed by solving a convex subproblem (an LP in our approach). Given a master solution $(\bar{y}, \bar{v}, \bar{u}, \bar{\rho}, \bar{\tau})$, the subproblem is given by

$$
\begin{aligned}
\phi(\bar{y}, \bar{v}, \bar{\rho}, \bar{\tau}) = \min_{(x,\pi,\sigma)\geq 0} \quad & 0 & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \sum_{j\in J} \sigma_{ij} \leq \bar{v}_i - \Gamma\bar{\rho}_i & i \in I \\
& \sum_{i\in I}\sum_{j\in J} d_j x_{ij} - \sum_{j\in J} \pi_j \geq D + \Gamma\bar{\tau} & \\
& \sum_{i\in I} \hat{d}_j x_{ij} - \pi_j \leq \bar{\tau} & j \in J \\
& \hat{d}_j x_{ij} - \sigma_{ij} \leq \bar{\rho}_i & j \in J, i \in I(j) \\
& \sum_{i\in I} x_{ij} \leq 1 & j \in J \\
& x_{ij} \leq \bar{y}_i & j \in J, i \in I(j).
\end{aligned}
$$

The solution of the subproblem leads to the generation of a Benders feasibility cut, i.e., a cutting plane that discards the infeasible master solution $(\bar{y}, \bar{v}, \bar{u}, \bar{\rho}, \bar{\tau})$. Solving a subproblem that could potentially be infeasible may lead to computational issues. An infeasible primal solution means that in the dual space we are optimizing over an unbounded cone. Among the successful strategies to overcome these difficulties are the normalization techniques that consist of solving the dual LP over a bounded polyhedron. There is abundant literature on different normalization techniques for Benders feasibility cuts, see, e.g., [45, 70, 105, 108, 118]. We use a very natural approach exploiting the fact that the

solution $(\bar{y}, \bar{v}, \bar{u}, \bar{\rho}, \bar{\tau})$ of the relaxed master problem is infeasible for the subproblem if and only if the demand covered by $\bar{y}$ is strictly less than $D + \Gamma\bar{\tau}$. Indeed, in our case, the covering constraint forms the irreducible infeasible subsystem, which is the minimal subset of constraints whose removal makes the problem feasible. Hence, instead of solving a feasibility LP given by the formulation above, we search for the maximum coverage. The resulting subproblem is the following LP

$$
\begin{aligned}
\phi'(\bar{y}, \bar{v}, \bar{\rho}, \bar{\tau}) = \max_{(x,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \sum_{j\in J} \pi_j & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \sum_{j\in J} \sigma_{ij} \leq \bar{v}_i - \Gamma\bar{\rho}_i & i \in I \\
& \sum_{i\in I} \hat{d}_j x_{ij} - \pi_j \leq \bar{\tau} & j \in J \\
& \hat{d}_j x_{ij} - \sigma_{ij} \leq \bar{\rho}_i & j \in J, i \in I(j) \\
& \sum_{i\in I} x_{ij} \leq 1 & j \in J \\
& x_{ij} \leq \bar{y}_i & j \in J, i \in I(j)
\end{aligned}
$$

in which we observe that for every given $(\bar{y}, \bar{v}, \bar{u}, \bar{\rho}, \bar{\tau})$, $(x, \pi, \sigma) = (0, 0, 0)$ is a feasible solution, i.e., this problem is always feasible and bounded as both x and π are bounded. If the optimal value of the subproblem $\phi'(\bar{y}, \bar{v}, \bar{\rho}, \bar{\tau})$ is greater or equal to $D + \Gamma\bar{\tau}$, the master solution is feasible and we do not need to generate any Benders cut. If, instead, $\phi'(\bar{y}, \bar{v}, \bar{\rho}, \bar{\tau})$ is strictly less than $D + \Gamma\bar{\tau}$, the master solution is infeasible and we do generate a Benders feasibility cut. The Benders cut is then given by $\phi'(y, v, \rho, \tau) \geq D + \Gamma\tau$ (and no more by $\phi(y, v, \rho, \tau) \geq 0$) for every feasible (y, v, ρ, τ).

We introduce the auxiliary variable $\delta \geq 0$ defined as $v - \Gamma\rho$. Because of concavity, $\phi'(\cdot)$ can be overestimated by a supporting hyperplane at $(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$, so the following linear cut is valid

$$\phi'(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau}) + \xi_y(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})^T(y - \bar{y}) + \xi_\rho(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})^T(\rho - \bar{\rho}) + \xi_\delta(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})^T(\delta - \bar{\delta}) + \xi_\tau(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})(\tau - \bar{\tau}) \geq \phi'(y, \rho, \delta, \tau) \geq D + \Gamma\tau$$

where $\xi_y(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$, $\xi_\rho(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$, $\xi_\delta(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$, $\xi_\tau(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$ denote subgradients of ϕ' with respect to y, ρ, δ, τ in $(\bar{y}, \bar{\rho}, \bar{\delta}, \bar{\tau})$. Depending on the problem, the computation of the subgradients could be heavy.

Therefore, we introduce a simple reformulation of the problem that makes their calculation straightforward, following what was done in [68]. The reformulation of the subproblem reads

$$
\begin{aligned}
\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \max_{(x,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \sum_{j\in J}\pi_j & & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \sum_{j\in J}\sigma_{ij} - \delta_i \leq 0 & i\in I & \\
& \sum_{i\in I} \hat{d}_j x_{ij} - \pi_j - \tau \leq 0 & j\in J & \\
& \hat{d}_j x_{ij} - \sigma_{ij} - \rho_i \leq 0 & j\in J, i\in I(j) & \\
& \sum_{i\in I} x_{ij} \leq 1 & j\in J & \\
& x_{ij} - y_i \leq 0 & j\in J, i\in I(j) & \\
& y_i = \bar{y}_i & i\in I & (6.25)\\
& \rho_i = \bar{\rho}_i & i\in I & (6.26)\\
& \delta_i = \bar{\delta}_i & i\in I & (6.27)\\
& \tau = \bar{\tau} & & (6.28)
\end{aligned}
$$

where we keep the master variables as variables of the subproblem as well, and we apply variable fixing through (6.25)–(6.28).

Note that y is defined as a continuous variable. Each variable-fixing equation is meant to be imposed by modifying the lower and upper bounds on the corresponding variable, e.g., for the master variable y we can impose $\bar{y} \leq y \leq \bar{y}$, which can be handled very efficiently by the solver in a preprocessing phase when the $\bar{y}$ is given.

Note that by construction, y only appears in the trivial constraints (6.25), hence the subgradient is simply $\xi_y(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \bar{r}_y$, where $\bar{r}_y$ is the vector of reduced costs associated with y. Accordingly, we have $\xi_\rho(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \bar{r}_\rho$, $\xi_\delta(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \bar{r}_\delta$, $\xi_\tau(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \bar{r}_\tau$, with r_ρ, r_δ, r_τ vectors of reduced cost associated with the variables ρ, δ, τ. This leads to the Benders cut

$$
\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) + \sum_{i\in I}\bar{r}_{y_i}(y_i-\bar{y}_i) + \sum_{i\in I}\bar{r}_{\rho_i}(\rho_i-\bar{\rho}_i) + \sum_{i\in I}\bar{r}_{\delta_i}(\delta_i-\bar{\delta}_i) + \bar{r}_\tau(\tau-\bar{\tau}) \geq D + \Gamma\tau \quad (6.29)
$$

where each component of the reduced cost defines an upper bound on the increase of the objective function $\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau})$ when one component of $(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau})$ increases.

Second choice of separation

We now discuss an alternative choice for the separation. We consider having (y, u, v) as master variables and we separate $(x, \rho, \tau, \pi, \sigma)$. The master problem is given by the MISOCP

$$\begin{aligned} \min_{(u,v)\geq 0,\, y\in\{0,1\}^{|I|}} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} u_i \\ s.t. \quad & \phi(y,v) \geq 0 \\ & v_i^2 - u_i y_i \leq 0 \qquad\qquad i \in I \end{aligned} \tag{6.30}$$

where $\phi(y, v)$ refers to the convex function expressing the optimal solution value of the subproblem as a function of y and v, and it can be approximated by linear Benders feasibility cuts, that are valid for any given vector (y, v).

Note that, also in this case, there is no need for Benders optimality cuts, as the only variables appearing in the objective function are y, v, and u that are not separated.

Given a master solution $(\bar{y}, \bar{v}, \bar{u})$, the subproblem is the following LP

$$\begin{aligned} \phi(\bar{y},\bar{v}) = \min_{(x,\rho,\tau,\pi,\sigma)\geq 0} \quad & 0 \\ s.t. \quad & \sum_{j\in J} d_j x_{ij} + \Gamma\rho_i + \sum_{j\in J}\sigma_{ij} \leq \bar{v}_i && i \in \\ & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\tau - \sum_{j\in J}\pi_j \geq D \\ & \tau + \pi_j - \sum_{i\in I}\hat{d}_j x_{ij} \geq 0 && j \in J \\ & \sigma_{ij} + \rho_i - \hat{d}_j x_{ij} \geq 0 && j \in J, i \in I(j) \\ & \sum_{i\in I} x_{ij} \leq 1 && j \in J \\ & x_{ij} \leq \bar{y}_i && i \in I(j), j \in J. \end{aligned}$$

If we apply the same normalization technique of the Benders cut described in the previous paragraph, the subproblem becomes

$$
\begin{aligned}
\phi'(\bar{y},\bar{v}) = \max_{(x,\rho,\tau,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\tau - \sum_{j\in J}\pi_j & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \Gamma\rho_i + \sum_{j\in J}\sigma_{ij} \leq \bar{v}_i & i\in I \\
& \tau + \pi_j - \sum_{i\in I}\hat{d}_j x_{ij} \geq 0 & j\in J \\
& \sigma_{ij} + \rho_i - \hat{d}_j x_{ij} \geq 0 & j\in J, i\in I(j) \\
& \sum_{i\in I} x_{ij} \leq 1 & j\in J \\
& x_{ij} \leq \bar{y}_i & j\in J, i\in I(j).
\end{aligned}
$$

which is always feasible and bounded.

We generate a Benders cut if and only if $\phi'(\bar{y},\bar{v}) < D$, i.e., if $(\bar{y},\bar{v})$ does not guarantee the required coverage. The Benders cut is then given by $\phi'(y,v) \geq D$ (and no more by $\phi(y,v) \geq 0$) for every feasible (y,v).

Because of concavity, $\phi'(\cdot)$ can be overestimated by a supporting hyperplane at $(\bar{y},\bar{v})$, so the following linear cut is valid

$$
\phi'(\bar{y},\bar{v}) + \xi_y(\bar{y},\bar{v})^T(y-\bar{y}) + \xi_v(\bar{y},\bar{v})^T(v-\bar{v}) \geq \phi'(y,v) \geq D
$$

where $\xi_y(\bar{y},\bar{v})$, $\xi_v(\bar{y},\bar{v})$ denote subgradients of ϕ' with respect to y,v in $(\bar{y},\bar{v})$. We can reformulate the subproblem as follows

$$
\begin{aligned}
\phi'(\bar{y},\bar{v}) = \max_{(x,\rho,\tau,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\tau - \sum_{j\in J}\pi_j & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \Gamma\rho_i + \sum_{j\in J}\sigma_{ij} - v_i \leq 0 & i\in I \\
& \tau + \pi_j - \sum_{i\in I}\hat{d}_j x_{ij} \geq 0 & j\in J \\
& \sigma_{ij} + \rho_i - \hat{d}_j x_{ij} \geq 0 & j\in J, i\in I(j) \\
& \sum_{i\in I} x_{ij} \leq 1 & j\in J \\
& x_{ij} - y_i \leq 0 & j\in J, i\in I(j) \\
& y_i = \bar{y}_i & i\in I \qquad (6.31) \\
& v_i = \bar{v}_i & i\in I \qquad (6.32)
\end{aligned}
$$

where we keep the master variables as variables of the subproblem as well, and we apply variable fixing through (6.31)–(6.32). By construction, y and v only appear in the trivial constraints (6.31) and (6.32), hence the subgradients are simply $\xi_y(\bar{y}, \bar{v}) = \bar{r}_y$ and $\xi_v(\bar{y}, \bar{v}) = \bar{r}_v$ where $\bar{r}_y$ is the vector of reduced costs associated with y and $\bar{r}_v$ is the vector of reduced costs associated with v. Note that y must be defined as a continuous variable as reduced cost can only be computed for continuous problems. The Benders cut reads

$$\phi'(\bar{y}, \bar{v}) + \sum_{i \in I} \bar{r}_{y_i}(y_i - \bar{y}_i) + \sum_{i \in I} \bar{r}_{v_i}(v_i - \bar{v}_i) \geq D \tag{6.33}$$

where each component of the reduced cost $\bar{r}_{y_i}$ (or $\bar{r}_{v_i}$) defines an upper bound on the increase of the objective function $\phi'(\bar{y}, \bar{v})$ when $\bar{y}_i$ (or $\bar{v}_i$) increases.

6.5 Embedding Benders decomposition within an MIP solver

There are two ways of implementing Benders decomposition:

1. using a *cutting-plane procedure*, also known as *multi-tree approach*, where each time a Benders cut is generated, the cut is included in the master problem and the latter is solved again until optimality;

2. using a *branch-and-Benders-cut algorithm*, also known as *single-tree approach*, in which a single enumeration tree is created and Benders cuts are separated at the nodes as in a classical branch-and-cut procedure.

The *traditional* approach is the multi-tree and it requires the solution of possibly several master problems to optimality. The single-tree approach instead requires the solution of only one master problem and has become very popular in the *recent* literature (see, e.g., [68, 69, 105]); however, its implementation was mainly applied in an MILP (rather than MISOCP) context. For more information about these two approaches and flowchart diagrams, see Section 2.1.3.

In this section we describe the steps for the design of a Benders decomposition approach to be embedded in a modern MIP solver, revealing all the implementation ingredients that play an important role in the design of an effective code. The discussion is based on the MIP solver we used, namely Gurobi; however, our approach can be easily extended to other solvers.

Note that we do not use any specialized solution method to solve the master problem or the subproblem as well. Indeed, using a general-purpose MIP solver benefits many

advantages, such as a better warm-start mechanism and effective heuristics. Moreover, this choice simplifies the implementation process considerably.

Implementation of single-tree approach

For the single-tree approach, the procedure is to define a customized callback function to be called at every node of the branch-and-cut applied to the master problem.

The callback function, described in Algorithm 1, defines how the separation is done. We restricted the separation to integer solutions only, thereby solving the subproblem to possibly generate Benders cuts exclusively when the master relaxation solution $y \in [0,1]^{|I|}$. If at the subproblem solution the target coverage is not reached, this incumbent must be discarded. Therefore, we generate a normalized Benders cut, inject this cut into the master problem and proceed through the branching tree until an optimal solution is found or the maximum time is exceeded.

Algorithm 1 Single-tree Benders implementation

```
if an MIP incumbent of the master problem is found then
    extract the MIP incumbent solution of relaxed the master problem
    update the lower and upper bounds of the master variables appearing in the feasible
subproblem
    solve the updated subproblem
    if the target coverage is not reached then
        get the reduced costs associated with the master variables in the subproblem
        add the normalized feasibility Benders cut to the master problem
    end if
end if
```

In order to use the custom callback function, we set the optimization parameter *LazyConstraints* to 1 for the master problem. We also set the parameter *timeLimit* to stop the solution process when the maximum time is reached.

Implementation of multi-tree approach

The code of the multi-tree Benders decomposition method is reported in Algorithm 2. It involves solving a master problem iteratively, updating the subproblem at every optimal solution of the master, and generating Benders cuts until the master solution does not violate the required coverage (we denote it as the termination condition in the algorithm) or the process terminates due to exceeding of the time limit.

Algorithm 2 Multi-tree Benders implementation

while termination condition is not met or a maximum time is not exceeded **do**
solve the relaxed master problem
extract the optimal solution of the master problem
update the lower and upper bounds of the master variables appearing in the feasible subproblem
solve the updated subproblem
if the target coverage is not reached **then**
get the reduced costs associated with the master variables in the subproblem
add the normalized feasibility Benders cut to the master problem
else
termination condition is met
end if
end while

6.5.1 Addressing degeneracy in Benders decomposition

When applying Benders decomposition, a highly degenerate subproblem admitting several optimal solutions can result in the generation of shallow Benders cuts. This slows down the convergence of the Benders decomposition, requiring the addition of many cuts, which do not improve the bound that much. There are many techniques proposed in the literature to address degeneracy (see, e.g., [120]). Inspired by the perturbation technique proposed in [69] to accelerate the convergence of a cut loop (i.e., the cut separation at the root node of the branching tree), we adapt this technique, that we will refer to as ϵ-*technique*, leading to stronger cuts. The result is an accelerated convergence and a fewer number of generated cuts, as shown from the computational experience provided later in Section 6.6.

The ϵ-technique can be summarized as follows. Consider the bound constraints that fix the master variables to their values in the subproblem and replace the bounds equal to zero with a sufficiently small $\epsilon > 0$. We denote the problem obtained as a perturbed subproblem. We show that we are forcing the perturbed problem to produce smaller reduced costs associated with the variables we fix, leading to stronger Benders cuts. Finally, to calculate the cut on the non-perturbed subproblem, we can derive ϕ' as a function of the optimal value of the perturbed problem and use the reduced costs of the perturbed problem as well.

We now analyze in detail this perturbation technique for each separation choice.

First choice of separation

Consider the *non-perturbed* subproblem formulated as follows

$$
\begin{aligned}
\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \max_{(x,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \sum_{j\in J}\pi_j & & & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \sum_{j\in J}\sigma_{ij} - \delta_i \leq 0 & i\in I & & \\
& \sum_{i\in I}\hat{d}_j x_{ij} - \pi_j - \tau \leq 0 & j\in J & & \\
& \hat{d}_j x_{ij} - \sigma_{ij} - \rho_i \leq 0 & j\in J, i\in I(j) & & \\
& \sum_{i\in I} x_{ij} \leq 1 & j\in J & (\alpha_j) & (6.34)\\
& x_{ij} - y_i \leq 0 & j\in J, i\in I(j) & & \\
& y_i = \bar{y}_i & i\in I & (\beta_i) & \\
& \rho_i = \bar{\rho}_i & i\in I & (\gamma_i) & \\
& \delta_i = \bar{\delta}_i & i\in I & (\Delta_i) & \\
& \tau = \bar{\tau} & & (\mu) &
\end{aligned}
$$

where by α,β,γ,Δ and μ we denote the dual variables. Given the optimal dual solution $(\bar{\alpha},\bar{\beta},\bar{\gamma},\bar{\Delta},\bar{\mu})$, the optimal value of the dual of (6.34) is

$$
\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \sum_{j\in J}\bar{\alpha}_j + \sum_{i\in I}\bar{y}_i\bar{\beta}_i + \sum_{i\in I}\bar{\rho}_i\bar{\gamma}_i + \sum_{i\in I}\bar{\delta}_i\bar{\Delta}_i + \bar{\tau}\bar{\mu}.
$$

We can get the expression of the Benders cut from the LP dual by imposing that the dual objective at the optimal solution is at least the target coverage, namely

$$
\sum_{j\in J}\bar{\alpha}_j + \sum_{i\in I}\bar{\beta}_i y_i + \sum_{i\in I}\bar{\gamma}_i\rho_i + \sum_{i\in I}\bar{\Delta}_i\delta_i + \bar{\mu}\tau \geq D + \Gamma\tau.
$$

The dual problem is highly degenerate as many objective coefficients, i.e., many components of $\bar{y}$, $\bar{\rho}$, $\bar{\delta}$ and $\bar{\tau}$ can be zero, resulting in several free components of the dual variables β, γ_i, Δ, μ and many equivalent optimal solutions. To get the strongest Benders cut, we need the dual multipliers appearing in the cut to take the smallest values possible. To get this, we can apply the ϵ-technique that consists of replacing the zero objective coefficients of $\bar{y}$, $\bar{\rho}$, $\bar{\delta}$, $\bar{\tau}$ with a sufficiently small $\epsilon > 0$ and solving the resulting problem. This induces the dual model to minimize also the components of the dual variables β, γ, Δ, μ associated with the originally zero objective coefficients.

Unfortunately, solving the dual of the subproblem often leads to more numerical issues than solving its primal, hence we explain how to apply this technique to the primal subproblem. Specifically, given a sufficiently small $\epsilon > 0$, applying the ϵ-technique to the primal problem consists of replacing $\bar{y}$, $\bar{\rho}$, $\bar{\delta}$ and $\bar{\tau}$ with

$$\bar{y}_i^\epsilon = \begin{cases} \bar{y}_i & \text{if } \bar{y}_i > 0 \\ \epsilon & \text{if } \bar{y}_i = 0, \end{cases} \qquad \bar{\rho}_i^\epsilon = \begin{cases} \bar{\rho}_i & \text{if } \bar{\rho}_i > 0 \\ \epsilon & \text{if } \bar{\rho}_i = 0, \end{cases} \qquad \bar{\delta}_i^\epsilon = \begin{cases} \bar{\delta}_i & \text{if } \bar{\delta}_i > 0 \\ \epsilon & \text{if } \bar{\delta}_i = 0, \end{cases} \qquad \bar{\tau}^\epsilon = \begin{cases} \bar{\tau} & \text{if } \bar{\tau} > 0 \\ \epsilon & \text{if } \bar{\tau} = 0 \end{cases}$$

for each $i \in I$, simply replacing each zero component of $\bar{y}$, $\bar{\rho}$, $\bar{\delta}$ and $\bar{\tau}$ with a sufficiently small ϵ. Then, the *perturbed* subproblem reads

$$\begin{aligned} \phi'(\bar{y}^\epsilon, \bar{\rho}^\epsilon, \bar{\delta}^\epsilon, \bar{\tau}^\epsilon) = \max_{(x,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \sum_{j\in J} \pi_j & & \\ s.t. \quad & \sum_{j\in J} d_j x_{ij} + \sum_{j\in J} \sigma_{ij} - \delta_i \leq 0 & i \in I & \\ & \sum_{i\in I} \hat{d}_j x_{ij} - \pi_j - \tau \leq 0 & j \in J & \\ & \hat{d}_j x_{ij} - \sigma_{ij} - \rho_i \leq 0 & j \in J, i \in I(j) & \\ & \sum_{i\in I} x_{ij} \leq 1 & j \in J & (\alpha_j^\epsilon) \quad (6.35) \\ & x_{ij} - y_i \leq 0 & j \in J, i \in I(j) & \\ & y_i = \bar{y}_i^\epsilon & i \in I & (\beta_i^\epsilon) \\ & \rho_i = \bar{\rho}_i^\epsilon & i \in I & (\gamma_i^\epsilon) \\ & \delta_i = \bar{\delta}_i^\epsilon & i \in I & (\Delta_i^\epsilon) \\ & \tau = \bar{\tau}^\epsilon & & (\mu^\epsilon) \end{aligned}$$

where by α^ϵ,β^ϵ,γ^ϵ,Δ^ϵ and μ^ϵ we denote the dual multipliers. The Benders cut obtained from the Lagrangian dual of (6.35) is

$$\phi'(\bar{y}^\epsilon, \bar{\rho}^\epsilon, \bar{\delta}^\epsilon, \bar{\tau}^\epsilon) + \sum_{i\in I} \bar{r}_{y_i}^\epsilon (y_i - \bar{y}_i^\epsilon) + \sum_{i\in I} \bar{r}_{\rho_i}^\epsilon (\rho_i - \bar{\rho}_i^\epsilon) + \sum_{i\in I} \bar{r}_{\delta_i}^\epsilon (\delta_i - \bar{\delta}_i^\epsilon) + \bar{r}_\tau^\epsilon (\tau - \bar{\tau}^\epsilon) \geq D + \Gamma\tau$$

where $\bar{r}_y^\epsilon$, $\bar{r}_\rho^\epsilon$, $\bar{r}_\delta^\epsilon$, $\bar{r}_\tau^\epsilon$ are the vectors of reduced costs associated with y, ρ, δ, τ. We now state some properties of the perturbed problem and clarify the relationship between the perturbed and the non-perturbed problem.

Lemma 6.5.1 *The following equations are valid*

$$\bar{r}_y = \bar{\beta}, \quad \bar{r}_\rho = \bar{\gamma}, \quad \bar{r}_\delta = \bar{\Delta}, \quad \bar{r}_\tau = \bar{\mu}, \quad \bar{r}^\epsilon_y = \bar{\beta}^\epsilon, \quad \bar{r}^\epsilon_\rho = \bar{\gamma}^\epsilon, \quad \bar{r}^\epsilon_\delta = \bar{\Delta}^\epsilon, \quad \bar{r}^\epsilon_\tau = \bar{\mu}^\epsilon.$$

Proof Consider the non-perturbed problem (6.34). Since (6.34) is an LP, from the equivalence between LP duality and Lagrangian duality for LP problems we have

$$\sum_{j\in J}\bar{\alpha}_j+\sum_{i\in I}\bar{\beta}_i y_i+\sum_{i\in I}\bar{\gamma}_i\rho_i+\sum_{i\in I}\bar{\Delta}_i\delta_i+\bar{\mu}\tau = \phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau})+\sum_{i\in I}\bar{r}_{y_i}(y_i-\bar{y}_i)+\sum_{i\in I}\bar{r}_{\rho_i}(\rho_i-\bar{\rho}_i)+\sum_{i\in I}\bar{r}_{\delta_i}(\delta_i-\bar{\delta}_i)+\bar{r}_\tau(\tau-\bar{\tau})$$

which implies

$$\sum_{i\in I}\bar{\beta}_i y_i = \sum_{i\in I}\bar{r}_{y_i} y_i, \quad \sum_{i\in I}\bar{\gamma}_i\rho_i = \sum_{i\in I}\bar{r}_{\rho_i}\rho_i, \quad \sum_{i\in I}\bar{\Delta}_i\delta_i = \sum_{i\in I}\bar{r}_{\delta_i}\delta_i, \quad \bar{r}_\tau\tau = \bar{\mu}\tau,$$

$$\sum_{j\in J}\bar{\alpha}_j = \phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) - \sum_{i\in I}\bar{r}_{y_i}\bar{y}_i - \sum_{i\in I}\bar{r}_{\rho_i}\bar{\rho}_i - \sum_{i\in I}\bar{r}_{\delta_i}\bar{\delta}_i - \bar{r}_\tau\bar{\tau},$$

meaning that $\bar{r}_{y_i} = \bar{\beta}_i$, $\bar{r}_{\rho_i} = \bar{\gamma}_i$, $\bar{r}_{\delta_i} = \bar{\Delta}_i$, $\bar{r}_\tau = \bar{\mu}$ for each $i \in I$, namely the dual variables associated with the fixing constraints are the reduced costs associated with the fixed variables. By applying the same procedure to the perturbed problem (6.35), we get $\bar{r}^\epsilon_{y_i} = \bar{\beta}^\epsilon_i$, $\bar{r}^\epsilon_{\rho_i} = \bar{\gamma}^\epsilon_i$, $\bar{r}^\epsilon_{\delta_i} = \bar{\Delta}^\epsilon_i$, $\bar{r}^\epsilon_\tau = \bar{\mu}^\epsilon$ for each $i \in I$.

□

Proposition 6.5.2 *If there exists a sufficiently small $\epsilon > 0$ such that the dual solution $(\alpha^\epsilon, \beta^\epsilon, \gamma^\epsilon, \Delta^\epsilon, \mu^\epsilon)$ of the perturbed problem* (6.35) *is an optimal solution of the dual of the non-perturbed problem* (6.34)*, then*

(1)

$$\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) = \begin{cases} \phi'(\bar{y}^\epsilon,\bar{\rho}^\epsilon,\bar{\delta}^\epsilon,\bar{\tau}^\epsilon) - \epsilon\sum_{i\in I:\,\bar{y}_i=0}\bar{r}^\epsilon_{y_i} - \epsilon\sum_{i\in I:\,\bar{\rho}_i=0}\bar{r}^\epsilon_{\rho_i} - \epsilon\sum_{i\in I:\,\bar{\delta}_i=0}\bar{r}^\epsilon_{\delta_i} & \text{if } \tau > 0 \\ \phi'(\bar{y}^\epsilon,\bar{\rho}^\epsilon,\bar{\delta}^\epsilon,\bar{\tau}^\epsilon) - \epsilon\sum_{i\in I:\,\bar{y}_i=0}\bar{r}^\epsilon_{y_i} - \epsilon\sum_{i\in I:\,\bar{\rho}_i=0}\bar{r}^\epsilon_{\rho_i} - \epsilon\sum_{i\in I:\,\bar{\delta}_i=0}\bar{r}^\epsilon_{\delta_i} - \epsilon\bar{r}^\epsilon_\tau & \text{if } \tau = 0 \end{cases}$$

(2) the following inequality

$$\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau}) + \sum_{i\in I}\bar{r}^\epsilon_{y_i}(y_i-\bar{y}_i) + \sum_{i\in I}\bar{r}^\epsilon_{\rho_i}(\rho_i-\bar{\rho}_i) + \sum_{i\in I}\bar{r}^\epsilon_{\delta_i}(\delta_i-\bar{\delta}_i) + \bar{r}^\epsilon_\tau(\tau-\bar{\tau}) \geq D + \Gamma\tau \tag{6.36}$$

is a valid Benders cut for the non-perturbed problem (6.34)*.*

Proof From the definition of y^ϵ,ρ^ϵ,δ^ϵ and τ^ϵ we have

$$\sum_{i\in I}\bar{y}_i^\epsilon = \sum_{i\in I}\bar{y}_i + \epsilon \sum_{i\in I:\,\bar{y}_i=0} 1 \tag{6.37}$$

$$\sum_{i\in I}\bar{\rho}_i^\epsilon = \sum_{i\in I}\bar{\rho}_i + \epsilon \sum_{i\in I:\,\bar{\rho}_i=0} 1 \tag{6.38}$$

$$\sum_{i\in I}\bar{\delta}_i^\epsilon = \sum_{i\in I}\bar{\delta}_i + \epsilon \sum_{i\in I:\,\bar{\delta}_i=0} 1 \tag{6.39}$$

$$\bar{\tau}^\epsilon = \begin{cases} \tau & \text{if } \tau > 0 \\ \\ \epsilon & \text{if } \tau = 0. \end{cases} \tag{6.40}$$

Consider the expression of the dual objective of the perturbed problem (6.42) at optimal value

$$\phi'(\bar{y}^\epsilon, \bar{\rho}^\epsilon, \bar{\delta}^\epsilon, \bar{\tau}^\epsilon) = \sum_{j\in J}\bar{\alpha}_j^\epsilon + \sum_{i\in I}\bar{y}_i^\epsilon\bar{\beta}_i^\epsilon + \sum_{i\in I}\bar{\rho}_i^\epsilon\bar{\gamma}_i^\epsilon + \sum_{i\in I}\bar{\delta}_i^\epsilon\bar{\Delta}_i^\epsilon + \bar{\tau}^\epsilon\bar{\mu}^\epsilon.$$

By using (6.37)-(6.40) we get

$$\phi'(\bar{y}^\epsilon, \bar{\rho}^\epsilon, \bar{\delta}^\epsilon, \bar{\tau}^\epsilon) = \begin{cases} \underbrace{\sum_{j\in J}\bar{\alpha}_j^\epsilon + \sum_{i\in I}\bar{y}_i\bar{\beta}_i^\epsilon + \sum_{i\in I}\bar{\rho}_i\bar{\gamma}_i^\epsilon + \sum_{i\in I}\bar{\delta}_i\bar{\Delta}_i^\epsilon + \bar{\tau}\bar{\mu}^\epsilon}_{\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau})} + \epsilon \sum_{i\in I:\,\bar{y}_i=0} \overbrace{\bar{\beta}_i^\epsilon}^{\bar{r}_{y_i}^\epsilon} + \epsilon \sum_{i\in I:\,\bar{\rho}_i=0} \overbrace{\bar{\gamma}_i^\epsilon}^{\bar{r}_{\rho_i}^\epsilon} + \epsilon \sum_{i\in I:\,\bar{\delta}_i=0} \overbrace{\bar{\Delta}_i^\epsilon}^{\bar{r}_{\delta_i}^\epsilon} & \text{if } \tau > 0 \\ \\ \underbrace{\sum_{j\in J}\bar{\alpha}_j^\epsilon + \sum_{i\in I}\bar{y}_i\bar{\beta}_i^\epsilon + \sum_{i\in I}\bar{\rho}_i\bar{\gamma}_i^\epsilon + \sum_{i\in I}\bar{\delta}_i\bar{\Delta}_i^\epsilon + \bar{\tau}\bar{\mu}^\epsilon}_{\phi'(\bar{y},\bar{\rho},\bar{\delta},\bar{\tau})} + \epsilon \sum_{i\in I:\,\bar{y}_i=0} \overbrace{\bar{\beta}_i^\epsilon}^{\bar{r}_{y_i}^\epsilon} + \epsilon \sum_{i\in I:\,\bar{\rho}_i=0} \overbrace{\bar{\gamma}_i^\epsilon}^{\bar{r}_{\rho_i}^\epsilon} + \epsilon \sum_{i\in I:\,\bar{\delta}_i=0} \overbrace{\bar{\Delta}_i^\epsilon}^{\bar{r}_{\delta_i}^\epsilon} + \epsilon \overbrace{\bar{\mu}^\epsilon}^{\bar{r}_\tau^\epsilon} & \text{if } \tau = 0 \end{cases}$$

which proves *(1)* for the assumption we made on ϵ and using Lemma 6.5.1.

If ϵ is sufficiently small that $(\alpha^\epsilon, \beta^\epsilon, \gamma^\epsilon, \Delta^\epsilon, \mu^\epsilon)$ is an optimal solution of the dual of the non-perturbed problem (6.34), then for Lemma 6.5.1, we can use $\bar{r}_{y_i}^\epsilon$, $\bar{r}_{\rho_i}^\epsilon$, $\bar{r}_{\delta_i}^\epsilon$ and $\bar{r}_\tau^\epsilon$ in (6.29) and *(2)* is proved.

□

Second choice of separation

Consider the *non-perturbed* subproblem formulated as follows

$$\begin{aligned}
\phi'(\bar{y},\bar{v}) = \max_{(x,\rho,\tau,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\tau - \sum_{j\in J}\pi_j & & \\
s.t. \quad & \sum_{j\in J} d_j x_{ij} + \Gamma\rho_i + \sum_{j\in J}\sigma_{ij} - v_i \leq 0 & i\in I & \\
& \tau + \pi_j - \sum_{i\in I}\hat{d}_j x_{ij} \geq 0 & j\in J & \\
& \sigma_{ij} + \rho_i - \hat{d}_j x_{ij} \geq 0 & j\in J, i\in I(j) & \\
& \sum_{i\in I} x_{ij} \leq 1 & j\in J & (\alpha_j) \\
& x_{ij} - y_i \leq 0 & j\in J, i\in I(j) & \\
& y_i = \bar{y}_i & i\in I & (\beta_i) \\
& v_i = \bar{v}_i & i\in I. & (\gamma_i)
\end{aligned} \tag{6.41}$$

where by α, β and γ we denote the dual variables. Given the optimal dual solution $(\bar{\alpha},\bar{\beta},\bar{\gamma})$, the optimal value of the dual of (6.41) is

$$\phi'(\bar{y},\bar{v}) = \sum_{j\in J}\bar{\alpha}_j + \sum_{i\in I}\bar{y}_i\bar{\beta}_i + \sum_{i\in I}\bar{v}_i\bar{\gamma}_i.$$

We can get the expression of the Benders cut from the LP dual by imposing that the dual objective at the optimal solution is at least the target coverage, namely

$$\sum_{j\in J}\bar{\alpha}_j + \sum_{i\in I}\bar{\beta}_i y_i + \sum_{i\in I}\bar{\gamma}_i v_i \geq D.$$

The dual problem is highly degenerate as many objective coefficients, i.e., many components of $\bar{y}$ and $\bar{v}$ can be zero, resulting in several free components of the dual variables β and γ and many equivalent optimal solutions. To get the strongest Benders cut, we need the dual multipliers appearing in the cut to take the smallest values possible. To get this, we can apply the ϵ-technique that consists of replacing the zero objective coefficients of $\bar{y}$ and $\bar{v}$ with a sufficiently small $\epsilon > 0$ and solving the resulting problem. This induces the dual model to minimize also the components of the dual variables β and γ associated with the originally zero objective coefficients.

Unfortunately, solving the dual of the subproblem often leads to more numerical issues than solving its primal, hence we explain how to apply this technique to the primal

subproblem. Specifically, given a sufficiently small $\epsilon > 0$, applying the ϵ-technique to the primal problem consists of replacing $\bar{y}$ and $\bar{v}$ with

$$\bar{y}_i^\epsilon = \begin{cases} \bar{y}_i & \text{if } \bar{y}_i > 0 \\ \epsilon & \text{if } \bar{y}_i = 0 \end{cases} \qquad \text{and} \qquad \bar{v}_i^\epsilon = \begin{cases} \bar{v}_i & \text{if } \bar{v}_i > 0 \\ \epsilon & \text{if } \bar{v}_i = 0 \end{cases}$$

for each $i \in I$, simply replacing each zero component of $\bar{y}$ and $\bar{v}$ with a sufficiently small ϵ. Then, the *perturbed* subproblem reads

$$\begin{aligned} \phi'(\bar{y}^\epsilon, \bar{v}^\epsilon) = \max_{(x,\rho,\tau,\pi,\sigma)\geq 0} \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\tau - \sum_{j\in J}\pi_j & & \\ s.t. \quad & \sum_{j\in J} d_j x_{ij} + \Gamma\rho_i + \sum_{j\in J}\sigma_{ij} - v_i \leq 0 & i \in I & \\ & \tau + \pi_j - \sum_{i\in I}\hat{d}_j x_{ij} \geq 0 & j \in J & \\ & \sigma_{ij} + \rho_i - \hat{d}_j x_{ij} \geq 0 & j \in J, i \in I(j) & \\ & \sum_{i\in I} x_{ij} \leq 1 & j \in J & (\alpha_j^\epsilon) \\ & x_{ij} - y_i \leq 0 & j \in J, i \in I(j) & \\ & y_i = \bar{y}_i^\epsilon & i \in I & (\beta_i^\epsilon) \\ & v_i = \bar{v}_i^\epsilon & i \in I. & (\gamma_i^\epsilon) \end{aligned} \tag{6.42}$$

where by $\alpha^\epsilon, \beta^\epsilon$ and γ^ϵ we denote the dual multipliers. The Benders cut obtained from the Lagrangian dual of (6.42) is

$$\phi'(\bar{y}^\epsilon, \bar{v}^\epsilon) + \sum_{i\in I} \bar{r}_{y_i}^\epsilon (y_i - \bar{y}_i^\epsilon) + \sum_{i\in I} \bar{r}_{v_i}^\epsilon (v_i - \bar{v}_i^\epsilon) \geq D$$

where $\bar{r}_y^\epsilon$ and $\bar{r}_v^\epsilon$ are the vectors of reduced costs associated with y and v. We now state some properties of the perturbed problem and clarify the relationship between the perturbed and the non-perturbed problem.

Lemma 6.5.3 *The following equations are valid*

$$\bar{r}_y = \bar{\beta}, \quad \bar{r}_v = \bar{\gamma}, \quad \bar{r}_y^\epsilon = \bar{\beta}^\epsilon, \quad \bar{r}_v^\epsilon = \bar{\gamma}^\epsilon.$$

Proof Consider the non-perturbed problem (6.41). Since (6.41) is an LP, from the equiva-

lence between LP duality and Lagrangian duality for LP problems we have

$$\sum_{j\in J}\bar{\alpha}_j + \sum_{i\in I}\bar{\beta}_i y_i + \sum_{i\in I}\bar{\gamma}_i v_i = \phi'(\bar{y},\bar{v}) + \sum_{i\in I}\bar{r}_{y_i}(y_i-\bar{y}_i) + \sum_{i\in I}\bar{r}_{v_i}(v_i-\bar{v}_i)$$

which implies

$$\sum_{i\in I}\bar{\beta}_i y_i = \sum_{i\in I}\bar{r}_{y_i} y_i, \quad \sum_{i\in I}\bar{\gamma}_i v_i = \sum_{i\in I}\bar{r}_{v_i} v_i, \quad \sum_{j\in J}\bar{\alpha}_j = \phi'(\bar{y},\bar{v}) - \sum_{i\in I}\bar{r}_{y_i}\bar{y}_i - \sum_{i\in I}\bar{r}_{v_i}\bar{v}_i,$$

meaning that $\bar{r}_{y_i} = \bar{\beta}_i$ and $\bar{r}_{v_i} = \bar{\gamma}_i$ for each $i \in I$, namely the dual variables associated with the fixing constraints are the reduced costs associated with the fixed variables. By applying the same procedure to the perturbed problem (6.42), we get $\bar{r}^{\epsilon}_{y_i} = \bar{\beta}^{\epsilon}_i$ and $\bar{r}^{\epsilon}_{v_i} = \bar{\gamma}^{\epsilon}_i$ for each $i \in I$.

□

Proposition 6.5.4 *If there exists a sufficiently small $\epsilon > 0$ such that the dual solution $(\alpha^{\epsilon}, \beta^{\epsilon}, \gamma^{\epsilon})$ of the perturbed problem* (6.42) *is an optimal solution of the dual of the non-perturbed problem* (6.41)*, then*

(1) $\phi'(\bar{y},\bar{v}) = \phi'(\bar{y}^{\epsilon},\bar{v}^{\epsilon}) - \epsilon \sum_{i\in I:\,\bar{y}_i=0} \bar{r}^{\epsilon}_{y_i} - \epsilon \sum_{i\in I:\,\bar{v}_i=0} \bar{r}^{\epsilon}_{v_i}$

(2) the following inequality

$$\phi'(\bar{y},\bar{v}) + \sum_{i\in I}\bar{r}^{\epsilon}_{y_i}(y_i-\bar{y}_i) + \sum_{i\in I}\bar{r}^{\epsilon}_{v_i}(v_i-\bar{v}_i) \geq D \tag{6.43}$$

is a valid Benders cut for the non-perturbed problem (6.41).

Proof From the definition of y^{ϵ} and v^{ϵ} we have

$$\sum_{i\in I}\bar{y}^{\epsilon}_i = \sum_{i\in I}\bar{y}_i + \epsilon \sum_{i\in I:\,\bar{y}_i=0} 1 \tag{6.44}$$

$$\sum_{i\in I}\bar{v}^{\epsilon}_i = \sum_{i\in I}\bar{v}_i + \epsilon \sum_{i\in I:\,\bar{v}_i=0} 1. \tag{6.45}$$

Consider the expression of the dual objective of the perturbed problem (6.42) at optimal value

$$\phi'(\bar{y}^{\epsilon},\bar{v}^{\epsilon}) = \sum_{j\in J}\bar{\alpha}^{\epsilon}_j + \sum_{i\in I}\bar{y}^{\epsilon}_i\bar{\beta}^{\epsilon}_i + \sum_{i\in I}\bar{v}^{\epsilon}_i\bar{\gamma}^{\epsilon}_i.$$

By using (6.44)-(6.45) we get

$$\phi'(\bar{y}^\epsilon, \bar{v}^\epsilon) = \underbrace{\sum_{j\in J} \bar{\alpha}_j^\epsilon + \sum_{i\in I} \bar{y}_i \bar{\beta}_i^\epsilon + \sum_{i\in I} \bar{v}_i \bar{\gamma}_i^\epsilon}_{\phi'(\bar{y},\bar{v})} + \epsilon \sum_{i\in I:\, \bar{y}_i=0} \overbrace{\bar{\beta}_i^\epsilon}^{\bar{r}_{y_i}^\epsilon} + \epsilon \sum_{i\in I:\, \bar{v}_i=0} \overbrace{\bar{\gamma}_i^\epsilon}^{\bar{r}_{v_i}^\epsilon}$$

which proves *(1)* for the assumption we made on ϵ and using Lemma 6.5.3.

If ϵ is sufficiently small that $(\alpha^\epsilon, \beta^\epsilon, \gamma^\epsilon)$ is an optimal solution of the dual of the non-perturbed problem (6.41), then for Lemma 6.5.3, we can use $\bar{r}_{y_i}^\epsilon$ and $\bar{r}_{v_i}^\epsilon$ in (6.33) and *(2)* is proved.

□

6.5.2 Other implementation details of the Benders decomposition

In this section, we reveal all the rounding operations we made and the thresholds we used to mitigate numerical issues.

When we get a master solution (an MIP incumbent for the single-tree, an optimal solution for the multi-tree approach), we round to 0-1 the binary variables belonging to the master solution. This is done even if the values should already be 0-1, as the solver may lack precision. Note that we do this rounding as the separation is only on integer master solutions.

When applying the ϵ-procedure, instead of replacing zero coefficients with ϵ, we replace the coefficients below a given small tolerance $\text{tol}_\beta > 0$ with ϵ. This is because the solver is supposed to treat values below some internal threshold as zero.

Once the subproblem at a given master solution is solved, the condition we should theoretically check to decide on the generation of the Benders cut is if $\phi'(\bar{y}, \bar{v}, \bar{\tau}, \bar{\rho}) < D + \Gamma\bar{\tau}$ for the first separation choice, and if $\phi'(\bar{y}, \bar{v}) < D$ for the second separation choice. To get a more precise check of this condition from the numerical point of view, we implement the first condition as $(\phi'(\bar{y}, \bar{v}, \bar{\tau}, \bar{\rho}) - \Gamma\bar{\tau})/D < 1 - \text{tol}_\alpha$ and the second condition as $\phi'(\bar{y}, \bar{v})/D < 1 - \text{tol}_\alpha$, where $\text{tol}_\alpha > 0$ is a given small tolerance.

6.6 Computational experiments

All the experiments run on a Ubuntu server with an Intel(R) Xeon(R) Gold 5218 CPU running at 2.30 GHz, with 96 GB of RAM and 8 cores.

As for the optimizer, we use Gurobi at its latest release (10.0.3) with a time limit (TL) of 900 seconds. On Gurobi, to guarantee better numerical precision, we set integrality tolerance *IntFeasTol* to 10^{-9}, feasibility tolerance *FeasibilityTol* to 10^{-9}, the optimality tolerances *OptimalityTol* to 10^{-9} and *MIPGap* to 0. As for the setting of Benders decomposition methods in Gurobi, we enable the parameter for lazy constraints (*LazyConstraints*=1) for the single-tree approaches to avoid any reductions and transformations that are incompatible with lazy constraints, and disable presolve (*Presolve=0*) in all the Benders approaches since during the presolve phase the model is not complete as many constraints are missing and the optimizer could make reductions without knowing the full problem.

We also set the following values for the tolerances defined in the previous section: $\text{tol}_\alpha = 10^{-5}$, $\text{tol}_\beta = 10^{-6}$, $\epsilon = 10^{-8}$.

Testbed

The testbed is made of large-size instances derived from Cordeau et al. [45]. The original instances are tailored for realistic scenarios where the number of customers is much larger than the number of potential facility locations (i.e., $|J| >> |I|$). In particular, we selected instances with 1000 customers and 100 potential facility locations. Customer demand in these instances was generated by Cordeau et al. by uniformly sampling from the range [1,100] and rounding to the nearest integer. The spatial coordinates for both customers and facilities were randomly selected from the interval [0, 30].

We further adapted the testbed to suit our study on robust and congested settings by injecting uncertainty in the customer demand and generating congestion quadratic cost. Specifically, for each customer j, the deviation $\hat{d}_j$ from the nominal demand d_j was set as a random integer drawn from the interval $[0, 20\%\, d_j]$, meaning that the maximum possible deviation from the nominal value is 20%. Customers affected by uncertainty are those for which $\hat{d}_j > 0$. Protection from the uncertainty in customer demand was introduced with varying levels of Γ in $\{2.5\%, 5\%, 10\%, 15\%, 20\%, 30\%, 40\%, 50\%\}|J|$. We limited Γ at 50% of customers to avoid overly conservative models. Congestion costs were computed by carefully setting the parameters a and b to balance linear and quadratic costs, and obtain optimal solutions requiring the opening of less than half of the potential facilities.

We used a set of 153 instances of the CPSCLP to compare the performance of different methods and formulations. The characteristics of each instance, including its optimal value and the number of open facilities at the optimal solution, are reported in Tables B.1-B.2 of the Appendix B. We observe that we excluded from the original set of 160 instances 7

instances where the Gurobi encountered numerical troubles in at least one of the tested methods leading to inconsistent results, following what was done also in the recent study [29] on perspective cuts and perspective reformulation.

Numerical results

The purpose of this section is to evaluate the performance of our Benders approaches as possible competitive exact algorithms to solve robust and congested PSCLPs. To the best of our knowledge, no previous computational study appeared in the literature for such a problem, despite its potential theoretical and practical relevance. For this reason, in this section we compare the performance of our Benders algorithms against the direct use of Gurobi applied as a black-box MIP solver to the perspective reformulation of the robust counterpart of the problem. We also test Gurobi on the extended formulation of the robust counterpart of the problem to show the effectiveness of the perspective reformulation. We choose to compare our Benders algorithms to Gurobi as it is considered one of the state-of-the-art MIP solvers and is usually used as a benchmark to compare the performance of newly developed exact algorithms. Moreover, Gurobi allows the implementation of callbacks on problems with quadratic constraints, a necessary feature for the implementation of our single-tree approaches.

We report the results given by ten exact algorithms: Gurobi on the extended formulation (Ext-GUROBI), Gurobi on the perspective reformulation (GUROBI), and eight versions of our Benders algorithms directly on the perspective reformulation: one single-tree version for each separation choice without and with the ϵ-technique (ST-BEN$_1$, ST-BEN$_2$, STϵ-BEN$_1$, STϵ-BEN$_2$), one multi-tree version for each separation choice without and with the ϵ-technique (MT-BEN$_1$, MT-BEN$_2$, MTϵ-BEN$_1$, MTϵ-BEN$_2$).

Numerical results are shown in Tables B.3-B.8 of the Appendix B. Using the values reported in the tables, we made Table 6.1 with summary results and a graphical representation of the overall performance of the different exact algorithms in Figure 6.5, illustrating the percentage of problems solved by each algorithm as a function of the computational time. The metrics shown in Table 6.1 are: i) the average computational time (AvgTime[s]) expressed in seconds, ii) the average relative gap at the end of the optimization (AvgGap[%]), iii) the average number of Benders cuts generated (AvgBenCuts), and iv) the percentage of solved instances (SolvedInst[%]). It is important to note that the performance profiles in Figures 6.5 do not represent the cumulative solution time but show the percentage of problems the algorithm can solve within a certain amount of time. The best performance are

graphically represented by the curves in the upper part of the plot. From this representation, we can also infer the percentage of instances solved by a specific algorithm. Note that if a line corresponding to a method does not appear in the plot, it is because the method could not solve any of the instances.

Table 6.1: Summary average results on the instances of the CPSCLP. "TL" stands for time limit.

Method	AvgTime[s]	AvgGap[%]	AvgBenCuts	SolvedInst[%]
Ext-GUROBI	TL	58,65	-	0
GUROBI	414,96	0,01	-	99,35
ST-BEN$_1$	893,16	4,93	156,83	1,96
STϵ-BEN$_1$	TL	1,13	86,8	0
ST-BEN$_2$	700,92	3,8	138,39	26,80
STϵ-BEN$_2$	130,88	0,004	27,73	98,69
MT-BEN$_1$	790,40	23,11	39,16	0
MTϵ-BEN$_1$	871,11	0,83	58,09	0
MT-BEN$_2$	658,18	18,09	40,31	18,30
MTϵ-BEN$_2$	142,58	0,03	12,79	90,20

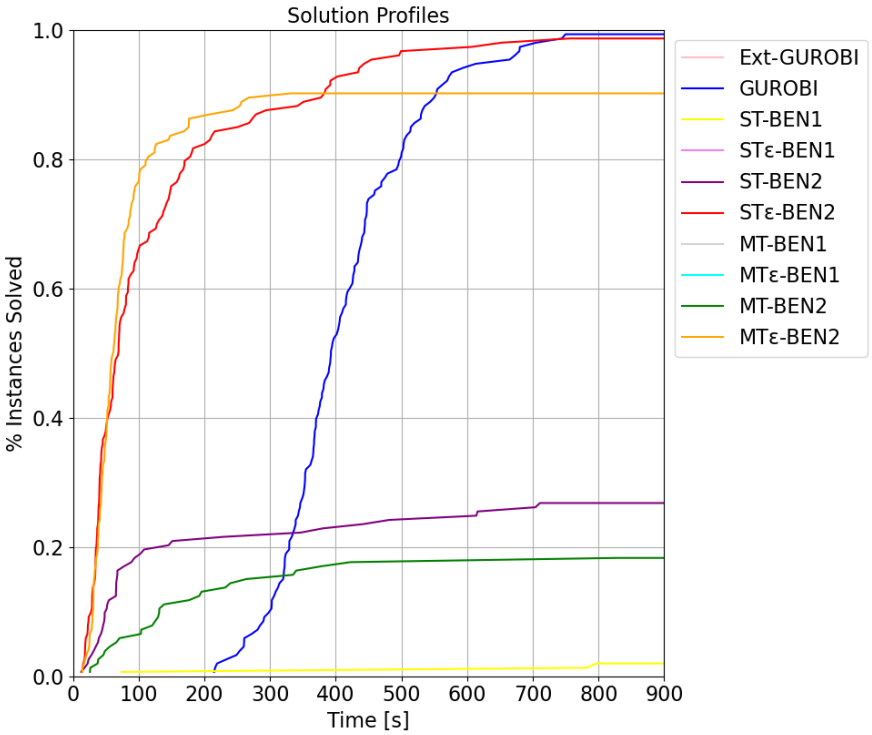

Figure 6.5: Performance profiles.

The profiles show that Ext-GUROBI behaves poorly for this problem as none of the instances can be solved by this method and the relative gaps at the end of the optimization are pretty high, all over 50%. The second worst performing methods for what concerns

the number of solved instances are the Benders approaches for separation choice 1 (with or without the ϵ-technique): STϵ-BEN$_1$ and MTϵ-BEN$_1$ solved 0 instances and the relative gaps at the end of the optimization are around 1% on average; the same methods without the ϵ-technique yield to higher gaps, 5% for ST-BEN$_1$ and 23% for MT-BEN$_1$, with 3 instances solved for ST-BEN$_1$ and 0 instances solved for MT-BEN$_1$. Then, we have the Benders approaches for separation choice 2 without the ϵ-technique: ST-BEN$_2$ solved 41 instances (around 27%) and MT-BEN$_2$ solved 28 instances (around 18%). The Benders methods for separation choice 2 with ϵ-technique, instead, solved most of the instances: STϵ-BEN$_2$ solved 151 instances and MTϵ-BEN$_2$ solved 138 instances (around 90%). Finally, GUROBI solved 152 instances up to 153.

As for resolution times, the algorithms running on the perspective reformulation are extremely faster than the one running on the extended formulation: indeed, Ext-GUROBI cannot solve any instance before the time limit and the gaps are still high at the time limit. For all the other methods, using the perspective reformulation, even if the instances are not solved within the time limit, the gaps are much smaller. The fastest methods are STϵ-BEN$_2$ and MTϵ-BEN$_2$ in a large percentage of the instances considered (more than 98%), and GUROBI in the remaining ones. On solved instances, with STϵ-BEN$_2$ times of GUROBI are reduced on average by almost 70%, with MTϵ-BEN$_2$ times of GUROBI are reduced on average by almost 83%.

For what concerns the direct comparison between the best-performing methods, STϵ-BEN$_2$ and MTϵ-BEN$_2$, we can notice that both approaches are very effective, with MTϵ-BEN$_2$ being slightly superior in terms of solution times, and STϵ-BEN$_2$ being slightly superior in terms of the number of solved instances.

From Figure 6.6 we can further compare the Benders approaches in terms of number of generated cuts. From the comparison of the box plots and Table 6.1 we can assess that the generation of cuts is reduced in both single and multi-tree approaches when using the ϵ-technique in most of the cases. Moreover, the average gap is reduced in all the Benders methods when using the ϵ-technique.

In Figures 6.7-6.8 we report the performance profiles for each fixed value of Γ to assess how Γ affects the performance of each solution algorithm. Whether a line corresponding to a method does not appear in the graph, it is because the method could not solve any of the instances characterized by that specific value of Γ.

From these figures, STϵ-BEN$_2$ and MTϵ-BEN$_2$ consistently emerge as the best-performing methods across all tested values of Γ. Specifically, MTϵ-BEN$_2$ significantly outperforms

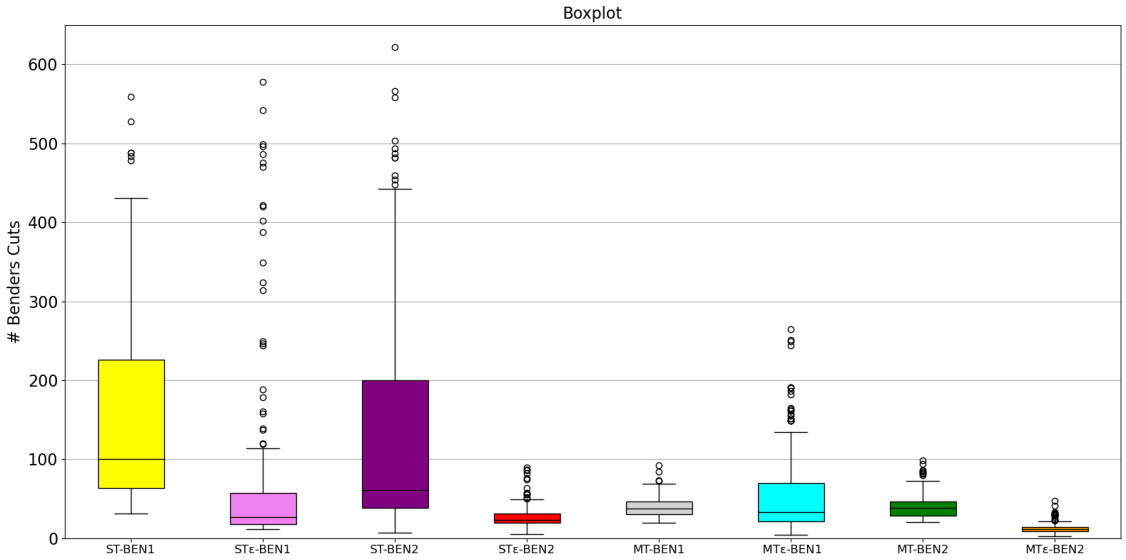

Figure 6.6: Boxplot of the number of Benders cuts generated by Benders decomposition methods.

STϵ-BEN$_2$ for very low values of Γ ($\Gamma = \{2.5\%|J|, 5\%|J|\}$). However, as Γ increases, the performance gap between GUROBI and the Benders methods STϵ-BEN$_2$ and MTϵ-BEN$_2$ narrows. This trend could be attributed to our approach of generating Benders cuts solely at integer points, thereby necessitating a higher number of branching steps in the Benders algorithms. In future work, we intend to explore the impact of generating cuts also at fractional solutions at the root node to understand whether this observed trend persists.

As for Benders methods with separation choice 2 that do not employ the ϵ-trick, they exhibit good performance at low Γ values. Specifically, ST-BEN$_2$ outperforms Gurobi in more than 80% of instances when $\Gamma = 2.5\%|J|$ and in over 20% of instances when $\Gamma = 5\%|J|$. Similarly, MT-BEN$_2$ beats Gurobi in nearly 80% of cases having $\Gamma = 2.5\%|J|$ and in more than 40% when $\Gamma = 5\%|J|$. However, their performance drastically declines as Γ increases, indicating a high limitation in their scalability with respect to increasing levels of protection against uncertainty.

In summary, our findings indicate that: (i) perspective reformulation results in smaller optimality gaps compared to extended formulation; (ii) separation choice 2 outperforms separation choice 1; (iii) employing the ϵ-technique aids in reducing both the number of generated cuts and the optimality gap, thereby accelerating convergence; (iv) The gap between GUROBI's performance and that of the best Benders approaches narrows with increasing values of Γ; (v) the best-performing methods overall are the single-tree and multi-tree Benders approaches with separation choice 2 using the ϵ-technique.

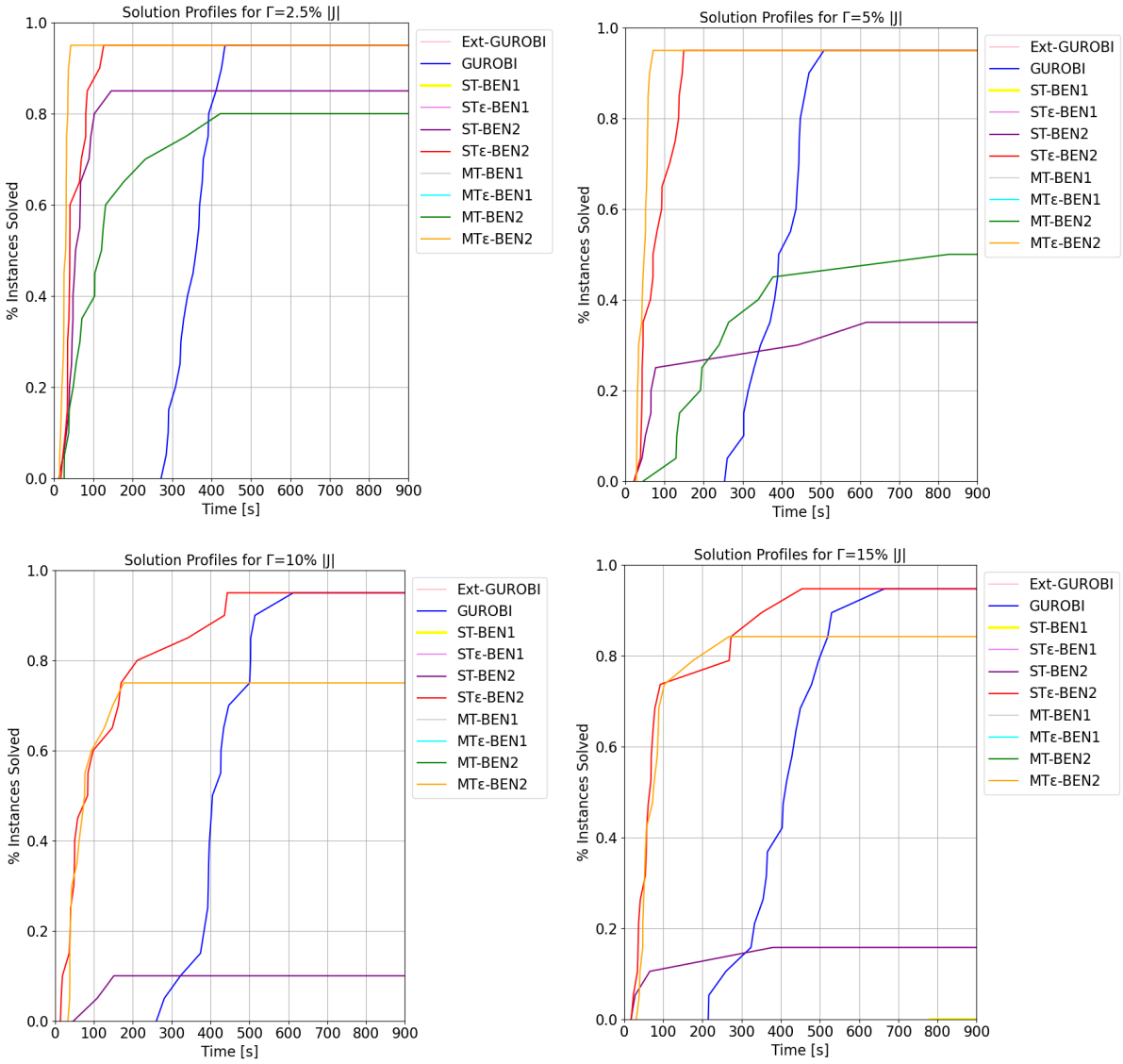

Figure 6.7: Performance profiles for low values of Γ.

6.7 Future work

In our investigation into the problem of congested partial set covering location, we suggested the application of a Benders decomposition method. One compelling advantage of this method is its ability to significantly reduce the number of variables, with a consequent reduction of the computational times when there is no need to generate an abundant number of cuts.

Looking forward, our Benders-based method can also address partial set covering location problems incorporating some additional requirements. Moreover, we propose two possible ways to accelerate the solution of the master problem that we have not tested yet. Finally, we discuss the investigation of the price of robustness.

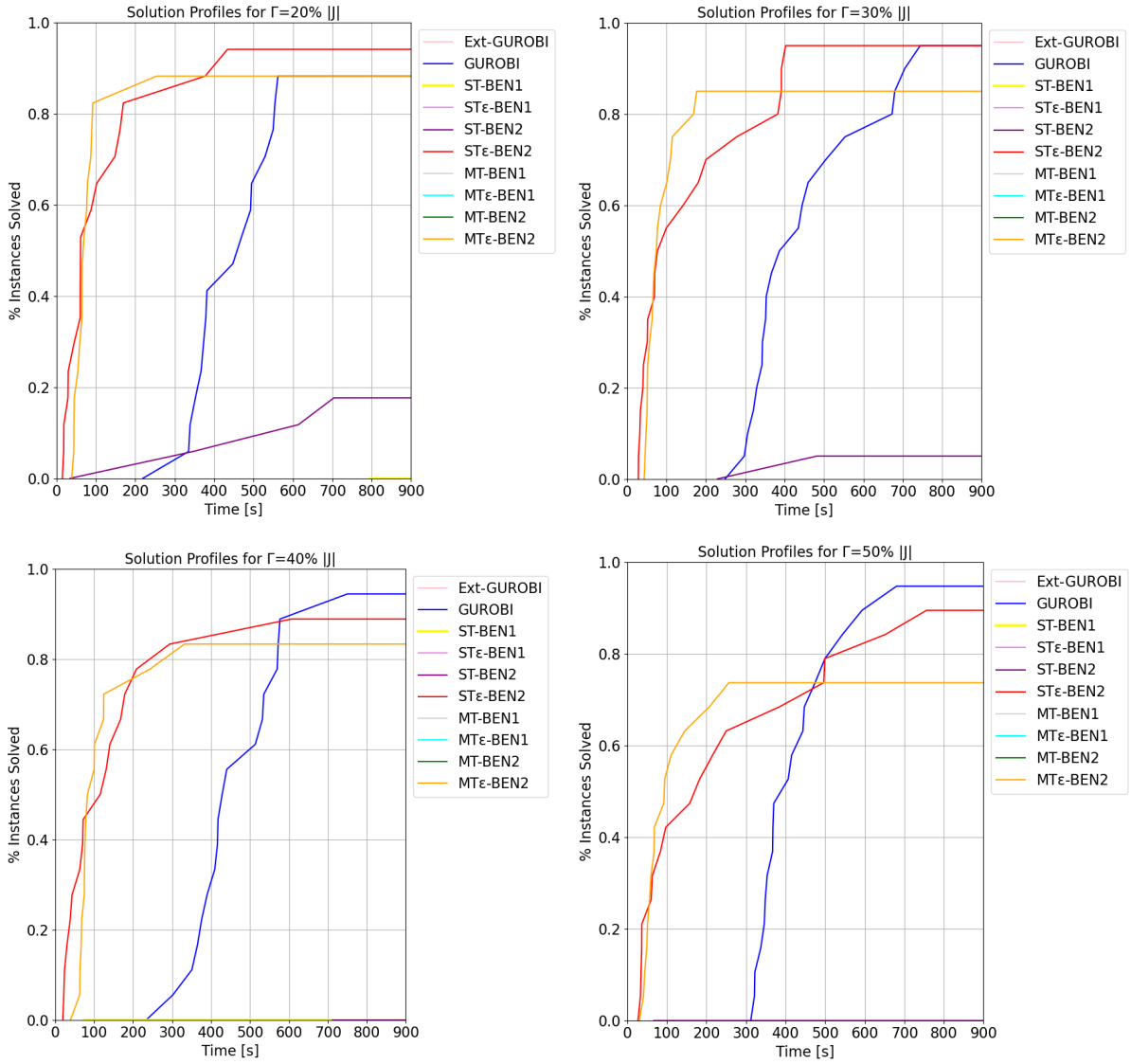

Figure 6.8: Performance profiles for medium to high values of Γ.

Additional constraints

Given the input parameter s_i representing the finite capacity of facility $i \in I$, we can consider the addition of (explicit) capacity constraints $v_i \leq s_i y_i$, imposing an upper limit on the load of each facility i.

Another interesting extension would be to impose a p-median constraint, which is particularly used in contexts where there is a maximum (or a target) number of facilities to open, denoted as p. This constraint can be expressed through $\sum_{i \in I} y_i \leq p$ (or $\sum_{i \in I} y_i = p$).

What we observe is that the Benders-based methods discussed in this chapter are directly applicable to these cases. This is because the variables implicated in capacity and p-median constraints – y and v – belong to the master problem.

Translated into operational terms, the inclusion of such constraints to the problem

would entail a straightforward integration of the same constraints into the master problem, whereas the Benders subproblem and cut remain the same. Therefore, a future study could extend to consider such cases in the testbed, to see if Benders approaches remain competitive compared to state-of-the-art MIP solvers.

Boosting the solution of the master problem

Let us consider the master problem

$$\begin{aligned} \min_{(u,v[,\rho,\tau])\geq 0, y\in\{0,1\}^{|I|}} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} u_i & \\ s.t. \quad & \phi(y, v[,\rho,\tau]) \geq 0 & \\ & v_i^2 - u_i y_i \leq 0 & i \in I \\ & [v_i \geq \Gamma\rho_i] & [i \in I] \end{aligned}$$

where the parts in squared brackets only appear in case we are applying the first separation choice, otherwise they do not appear in the formulation. We can strengthen the master problem by adding a cardinality constraint to the master variables y and, consequently, give a lower bound to the optimal value of the facility opening cost in this way

$$\begin{aligned} \min_{(u,v[,\rho,\tau])\geq 0, y\in\{0,1\}^{|I|}} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} u_i & \\ s.t. \quad & \phi(y, v[,\rho,\tau]) \geq 0 & \\ & v_i^2 - u_i y_i \leq 0 & i \in I \\ & [v_i \geq \Gamma\rho_i] & [i \in I] \\ & \sum_{i\in I} f_i y_i \geq g(I_{min}) & \\ & \sum_{i\in I} y_i \geq I_{min} & \end{aligned}$$

where by I_{min} we denote the minimum number of facilities needed to cover at least D and by $g(I_{min})$ we denote the minimum cost of opening exactly I_{min} facilities. We introduce the 0-1 variables z_j to denote if a customer $j \in J$ is covered or not. To find I_{min} and $g(I_{min})$ we can solve

$$
\begin{aligned}
g(I_{min}) = \min & \sum_{i \in I} f_i y_i \\
s.t. \quad & \sum_{j \in J} d_j z_j - \max_{\{S | S \subseteq J, |S| \leq \Gamma\}} \left\{ \sum_{j \in S} \hat{d}_j z_j \right\} \geq D && (\alpha) \\
& z_j \leq \sum_{i \in I(j)} y_i && j \in J \quad (\beta_j) \\
& y_i \in \{0,1\} && i \in I \\
& z_j \in \{0,1\} && j \in J
\end{aligned}
$$

which can be reformulated as the following MILP

$$
\begin{aligned}
g(I_{min}) = \min_{(\alpha,\beta) \geq 0} \quad & \sum_{i \in I} f_i y_i \\
s.t. \quad & \sum_{j \in J} d_j z_j - \Gamma \alpha - \sum_{j \in J} \beta_j \geq D \\
& z_j \leq \sum_{i \in I(j)} y_i && j \in J \\
& \alpha + \beta_j \geq \hat{d}_j z_j && j \in J \\
& y_i \in \{0,1\} && i \in I \\
& z_j \in \{0,1\} && j \in J
\end{aligned}
\tag{6.46}
$$

since

$$
\max_{\{S | S \subseteq J, |S| \leq \Gamma\}} \left\{ \sum_{j \in S} \hat{d}_j z_j \right\}
$$

is equivalent to

$$
\max \left\{ \sum_{j \in J} \hat{d}_j z_j w_j : \sum_{j \in J} w_j \leq \Gamma, 0 \leq w_j \leq 1 \ \forall j \in J \right\}
$$

which is equivalent to its dual

$$
\min \left\{ \Gamma \alpha + \sum_{j \in J} \beta_j : \alpha + \beta_j \geq \hat{d}_j z_j \ \forall j \in J, \alpha \geq 0, \beta \geq 0 \right\}.
$$

Note that we can relax the integrality of z_j in (6.46). If we denote the optimal solution of (6.46) with $(y^\star, z^\star, \alpha^\star, \beta^\star)$, then I_{min} is given by $\sum_{i \in I} y_i^\star$.

Proposition 6.7.1 *$g(I_{min})$ provides a lower bound to the optimal opening cost of facilities.*

Proof Consider the variable z_j as $z_j = \sum_{i \in I} x_{ij}$ and consider another vector of variables

$v \geq 0$. We can rewrite Problem (6.46) as follows

$$
\begin{aligned}
g(I_{min}) = \min_{(\alpha,\beta)\geq 0} \quad & \sum_{i\in I} f_i y_i + b\sum_{i\in I} v_i + a\sum_{i\in I} v_i^2 & & \\
s.t. \quad & \sum_{j\in J}\sum_{i\in I} d_j x_{ij} - \Gamma\alpha - \sum_{j\in J}\beta_j \geq D & & \\
& \sum_{i\in I} x_{ij} \leq \sum_{i\in I(j)} y_i & & j \in J \\
& \alpha + \beta_j \geq \sum_{i\in I} \hat{d}_j x_{ij} & & j \in J \\
& \sum_{i\in I} x_{ij} \leq 1 & & j \in J \\
& v_i \geq 0 & & i \in I \\
& 0 \leq x_{ij} \leq 1 & & i \in I, j \in J \\
& y_i \in \{0,1\} & & i \in I
\end{aligned}
\tag{6.47}
$$

since the only constraint on v variables is that $v \geq 0$, hence at the optimal solution we have $v = 0$. We observe that Problem (6.47) is a relaxation of the extended formulation (6.13)-(6.18) (and also a relaxation of its perspective reformulation). Hence, its optimal value provides a lower bound to the optimal value of the extended formulation.

□

We observe that

- $y^\star$ is feasible for the robust counterpart of the problem, as it is 0-1 and satisfies the target demand in the worst-case scenario;
- we can use $y^\star$ as a warm start to the master problem;
- we can get an upper bound to the master problem (a cutoff value) by fixing $y^\star$ in the counterpart of the problem and getting the objective value obtained by optimizing with respect to the remaining (continuous) variables.

One more consideration is that, based on the values of the input parameters a and b weighting the components of the congestion cost, the optimal value of the counterpart of the problem is reached at the optimal vector $\tilde{y}$, which could be a sparse vector or a dense vector. From early testing, vector $y^\star$ seems to be a pretty sparse vector. Therefore, another interesting upper bound to the master problem is given by the dense vector $y_i = 1$ for all $i \in I$, namely when we open all the facilities. This feasible solution may be a better

warm start to the problem compared to $y^\star$ when the optimal $\tilde{y}$ is pretty dense, and the corresponding upper bound more strict to the one produced by $y^\star$.

Hence, one strategy could be providing both solutions as warm starts to the problem and choosing as a cutoff value the minimum between the two optimal values given by the two feasible solutions.

We conclude by claiming that usually, the automatic warm start of Gurobi is more optimized than a custom warm start, even if the custom warm start is producing a good initial feasible solution, and for this reason, it is not recommended to change it. We also observe that the presence of cardinality constraints could generate numerical issues when they are combined with the conic constraints of the perspective reformulation. However, we would like in future research to test the use of cardinality constraints and warm start, both together and separated, to assess whether they decrease solution times in Gurobi.

Sensitivity analysis on Γ

In this chapter we studied a robust problem characterized by an unusual feature: the worst realizations occur in the case of increasing demand for what concerns the congestion, and in the opposite case of decreasing demand for the coverage. Therefore, it would be interesting to investigate how the optimal value, solution time, facility loads and number of open facilities vary based on the values of Γ in each of the following scenarios: the deterministic case, the case considering uncertainty only in congestion, the one with uncertainty only in coverage, and the case where uncertainty exists in both congestion and coverage. Exploring these scenarios could provide critical insights into how sensitive the planning is to various types of uncertainty, thus offering a more comprehensive understanding of the problem.

6.8 Conclusions

We presented a novel investigation into solving a robust and convex quadratic variant of the partial set covering location problem using modern Benders decomposition. We addressed a facility location problem affected by congestion, aiming for better resource allocation and balanced solutions. Specifically, we introduced the congested partial set covering location problem, which involves determining a subset of facility locations to open and efficiently allocating customers to these facilities, so as to minimize the combined costs of facility opening and congestion while ensuring target coverage. To enhance the resilience of the solution against demand fluctuations, we proposed to address the case under uncertain customer demand using Γ-robustness. We showed the formulation of the deterministic

problem and its robust counterpart.

However, the size of the robust counterpart grows with the number of customers and facility locations, which can pose challenges in real-world contexts where the customer number is significant. To overcome this issue, we proposed the use of Benders decomposition to effectively reduce the number of variables by projecting out of the master problem all the variables dependent on the number of customers. Although Benders decomposition is a traditional technique, our approach is innovative as it leverages the implementation of callback functions in combination with quadratic constraints, a novel option provided by (few) state-of-the-art solvers. Furthermore, we were able to demonstrate the effectiveness of Benders decomposition, even though the Benders subproblem is non-separable. We illustrated how to incorporate our Benders approach within an MIP solver, addressing explicitly all the ingredients that are instrumental for its success. We discussed and tested both single-tree and multi-tree implementations of Benders decomposition, testing two different separation choices. We also proposed a perturbation technique to generate stronger Benders cuts in case of problem degeneracy.

To validate the effectiveness of our method, we conducted extensive computational experiments on various (adapted) instances from existing literature. Overall, the results showed that the best-performing methods are the single-tree and multi-tree Benders approaches with one particular separation choice and using the perturbation technique.

Other future work will concern the study of the behaviour of Γ in each uncertain term, the use of warm start and/or cardinality constraints in the master problem, and the introduction of capacity and/or p-median constraints.

Chapter 7

Conclusions

In this doctoral thesis, we addressed critical challenges in the design of service and communications networks. With telecommunications playing a crucial role in modern society, the introduction of 5^{th} Generation technology has imposed rigorous network quality standards. Consequently, network providers contend with the complexities of optimal facility location planning and coverage. In response to this need, we investigated models and solution approaches for covering location problems in the telecommunications field.

Our first contribution concerned the design of wireless networks, a typical problem in the telecommunications sector with crucial practical implications. The expansion of new-generation networks, operating in a congested radio spectrum, posed challenges in determining optimal transmitter locations and power emissions. Traditional methods employed by practitioners, relying on simulation and trial-and-error practices, proved limited, necessitating optimization approaches for cost reduction and meeting service quality standards. Several optimization models have been developed for designing wireless network; however, the natural formulation on which most models are based presents severe limitations, preventing the solution of instances of practical interest. We tackled two primary wireless network design problems: the site and power assignment problem for variable-power emissions and the base station deployment problem for fixed-power emissions. For the variable-power case, we enhanced existing formulations by introducing presolve operations, and valid cutting planes such as cliques and variable upper bounds. Additionally, an aggressive reduction scheme was used to improve the formulation and reduce problem size. In the fixed-power case, we intervened with both strengthening and size reduction operations involving the aggregation of service constraints and the addition of valid inequalities. In both problems, we were able to achieve optimal solutions in solution times in line with the times required in the planning phase on large instances based on an authentic case sourced

from an Italian municipality.

The second contribution, with broader applicability to service and communications networks, focused on a variant of the partial set covering location problem. This problem involves choosing the minimum-cost subset of facilities satisfying a target demand. However, existing literature on the problem does not account for congestion or demand uncertainty. To address this gap, we decided to minimize congestion cost associated with overloaded facilities in a robust framework, accounting for uncertainties in customer demand, to enhance better resource allocation and solution resilience against demand fluctuations. By reducing facility congestion, we not only guarantee to meet delivery schedules and uphold rigorous quality standards, but we also respond to the contemporary society's growing demand for high service quality. The size of the resulting problem grows with the number of customers, which could be very large in real-life networks. To handle this, we employed Benders decomposition to eliminate all the variables dependent on the number of customers, and proposed the use of a perturbation technique to generate stronger Benders cuts. Computational experiments on various large-size instances taken from the literature demonstrated the effectiveness of our solution approach, which outperforms the state-of-the-art solver Gurobi.

In conclusion, this thesis contributes to solving complex location problems with covering constraints in the telecommunications sector, providing solutions that bridge the gap between theoretical research and its applicability to the real-world.

Appendices

Appendix A

Additional Figures

The wireless network design problem

Figures A.1-A.10 refers to the base station deployment problem. In particular, each figure depicts the optimal solution found on an instance using the final formulation. They report the optimal assignment of testpoints to each activated base station via color code. The black dots represent the uncovered testpoints, whereas the grey dots represent the original testpoints not included in the instance. The activated base stations have the same color as the testpoints they serve. The representations show that the assignment of the testpoints to the base stations is consistent with their respective positions.

The direction of the signal emitted by the antennas is illustrated with an arrow. The arrow represents a simplification of the antenna pattern, which actually is similar to a main lobe with side and back lobes, as schematized in Figure A.11. Another way to illustrate how the signal emitted by the antenna radiates into space is to plot the received power values in the target area of a signal emitted by a given antenna. Figure A.12 represents precisely this, reproducing the strongest power values in yellow and the less strong ones in dark blue. The multi-lobed formation is also visible in this image.

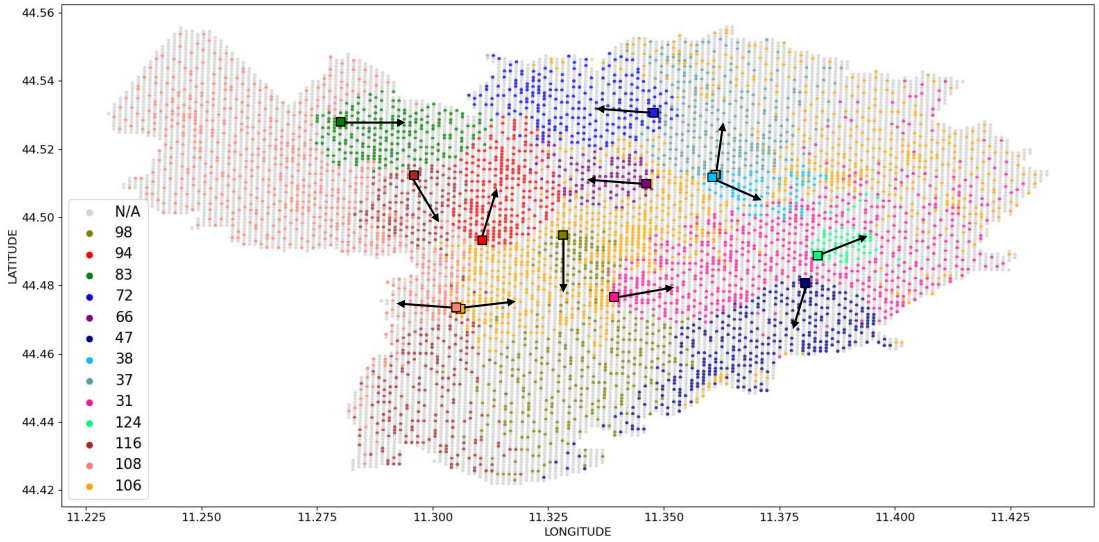

Figure A.1: Optimal assignment of testpoints to the base stations on BOF1.
N/A, not available since the testpoint is not in the instance

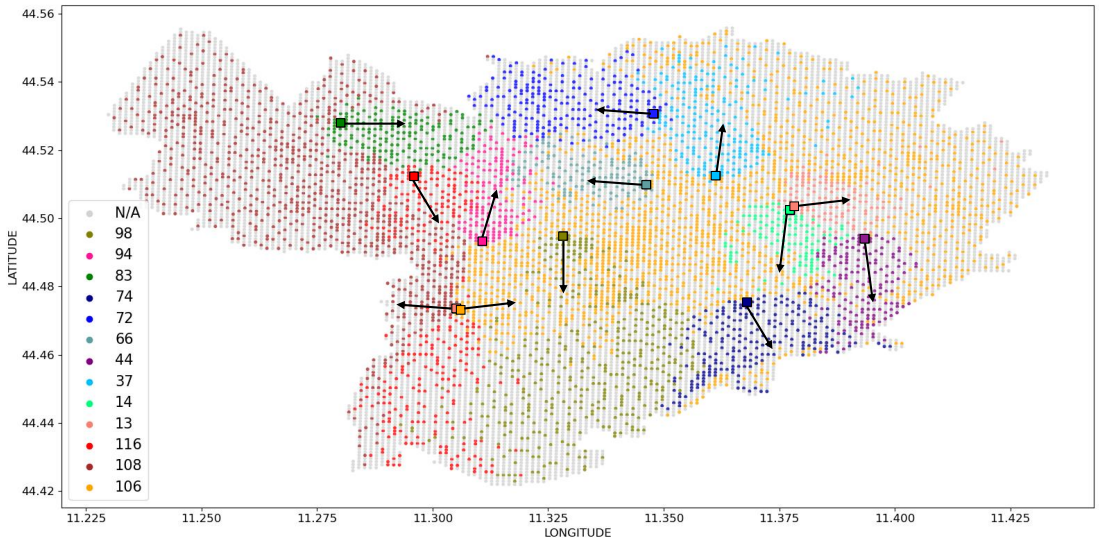

Figure A.2: Optimal assignment of testpoints to the base stations on BOF2.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

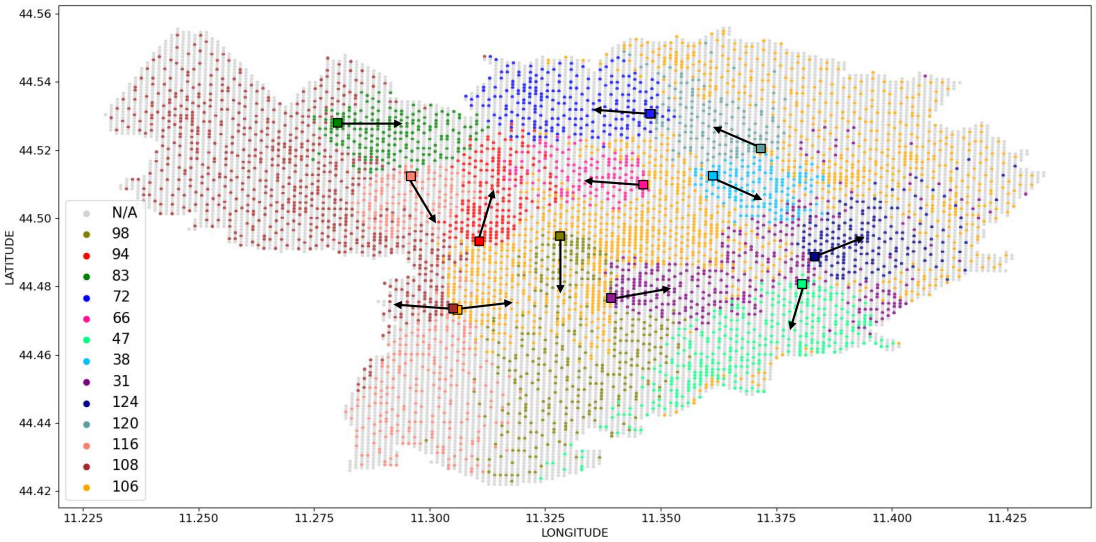

Figure A.3: Optimal assignment of testpoints to the base stations on BOF3.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

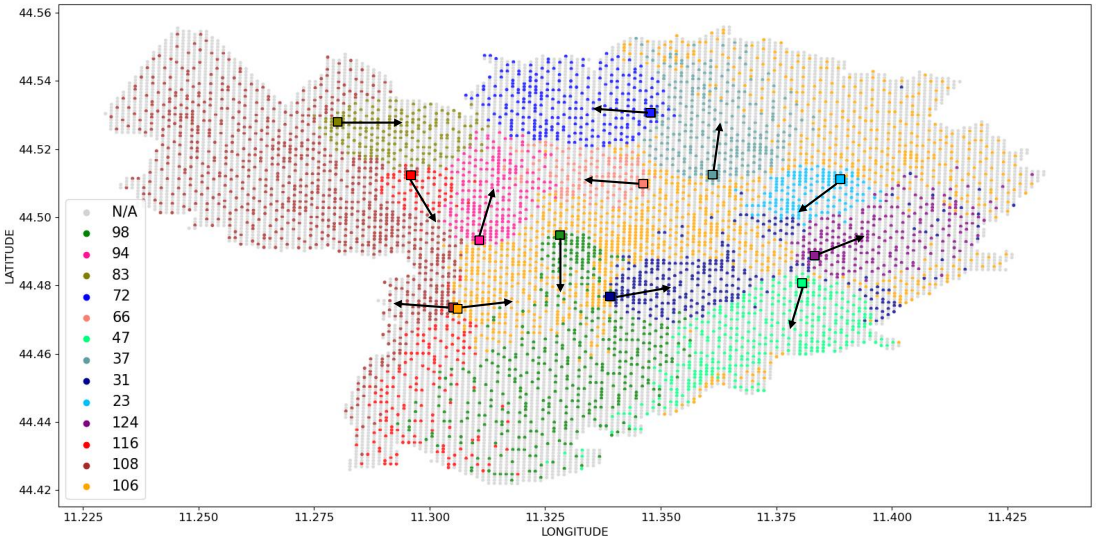

Figure A.4: Optimal assignment of testpoints to the base stations on BOF4.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

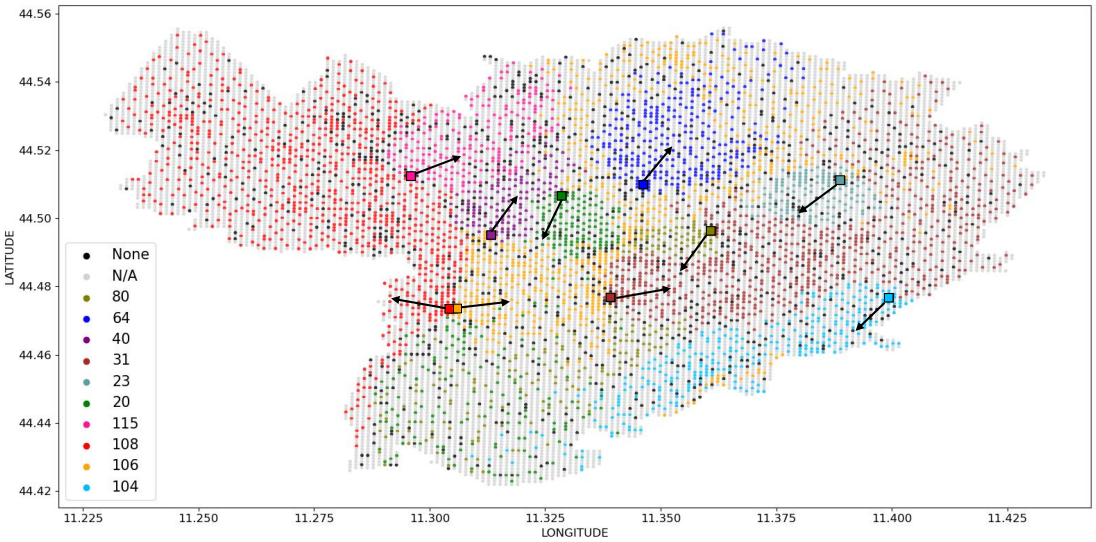

Figure A.5: Optimal assignment of testpoints to the base stations on BOF5.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

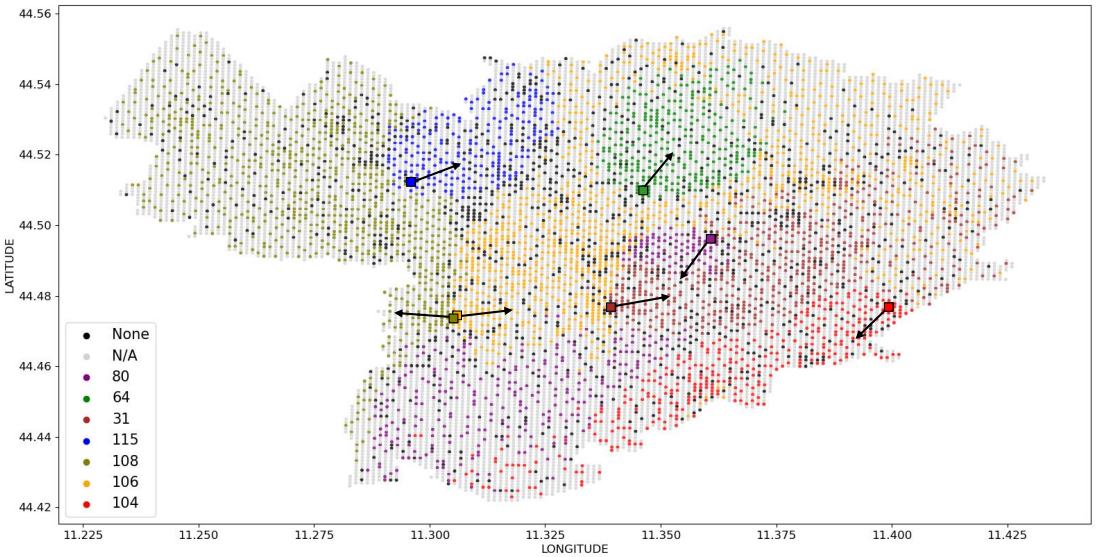

Figure A.6: Optimal assignment of testpoints to the base stations on BOF6.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

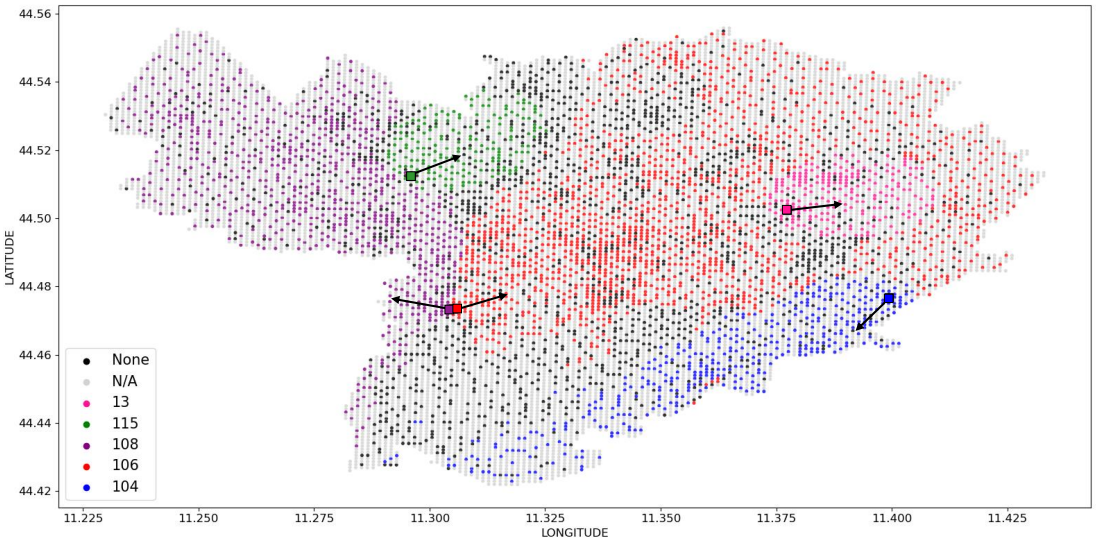

Figure A.7: Optimal assignment of testpoints to the base stations on BOF7.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

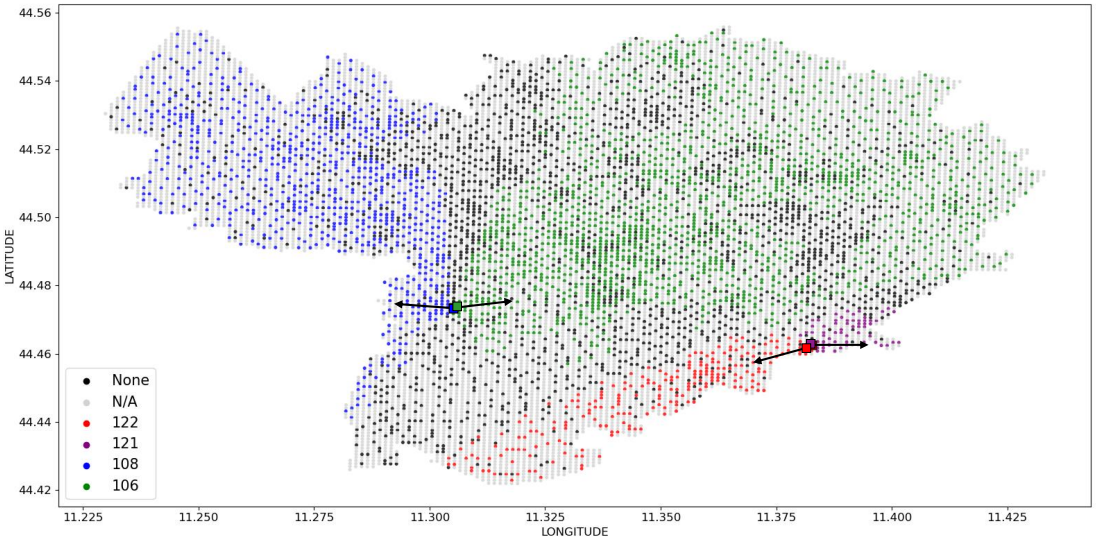

Figure A.8: Optimal assignment of testpoints to the base stations on BOF8.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

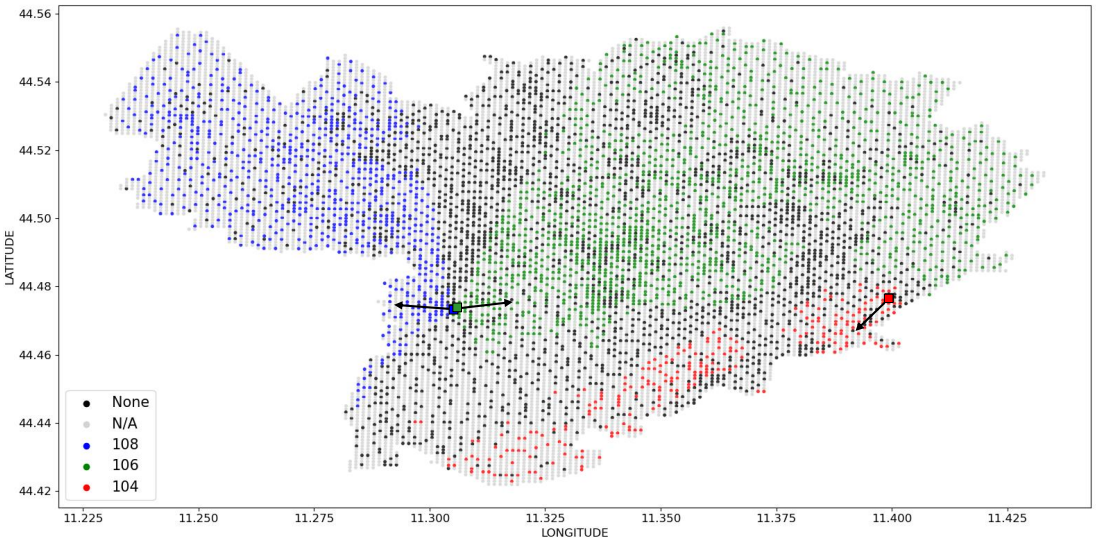

Figure A.9: Optimal assignment of testpoints to the base stations on BOF9.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

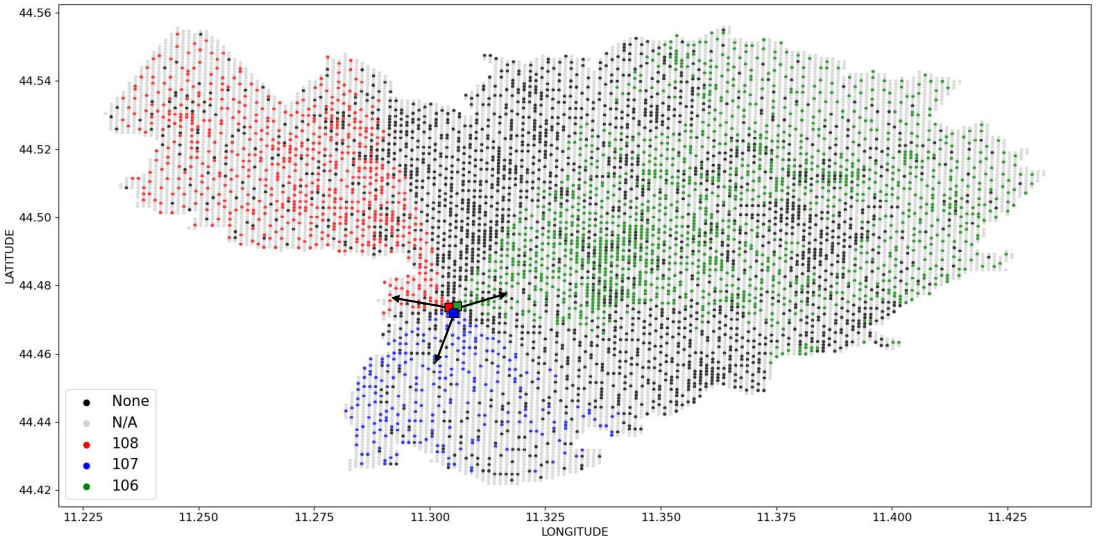

Figure A.10: Optimal assignment of testpoints to the base stations on BOF10.
None, the testpoint is covered by none of the activated antennas; N/A, not available since the testpoint is not in the instance

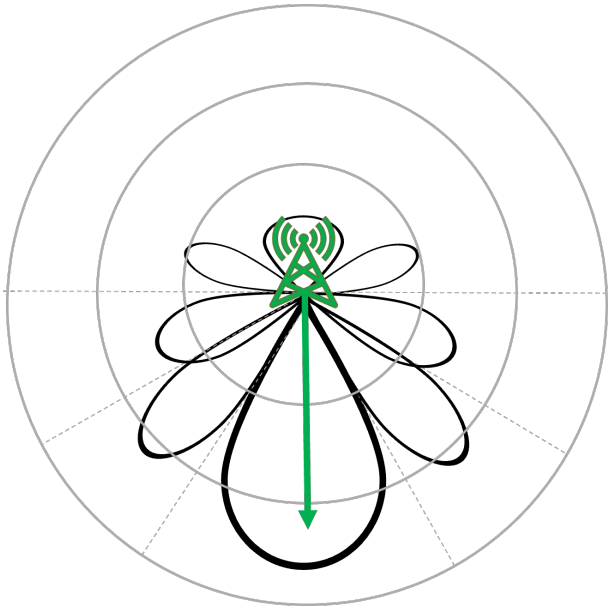

Figure A.11: Scheme of the antenna pattern.

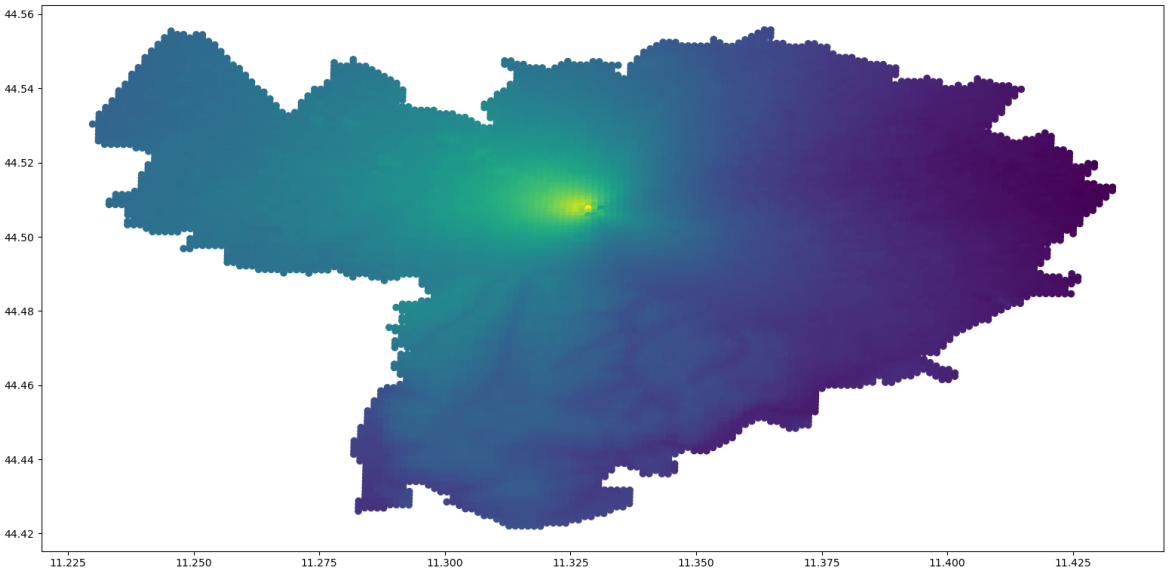

Figure A.12: Antenna pattern through power values.

Appendix B

Additional Tables

The congested partial set covering location problem

Tables B.1-B.2 report the characteristics of the instances of congested partial set covering location. Specifically, for each instance, the following data are reported: a unique identifier ID, a number s linked to the seed of the instance generator, the maximum number of customer deviations Γ we consider for the robust setting, the maximum distance of coverage R, the best upper bound found for instance at hand BestUB, and the number of open facilities OpenFac at the best feasible solution found on the instance.

Tables B.3-B.5 report the results of four exact algorithms: Gurobi on the extended formulation (Ext-GUROBI), Gurobi on the perspective reformulation (GUROBI), and the two best-performing versions of our Benders algorithms using the perspective reformulation: the single-tree version with separation choice 2 using the ϵ-technique (STϵ-BEN$_2$) and the multi-tree version with separation choice 2 using the ϵ-technique (MTϵ-BEN$_2$). The metrics we use are: i) the computational time (Time[s]) expressed in seconds, ii) the relative gap at the end of the optimization (Gap[%]), iii) the number (or the average number for the multi-tree approaches) of branching nodes explored (Nodes) and iv) the number of Benders cuts generated (BenCuts). As for the Gap, the internal relative gap of Gurobi was used where available. For the single-tree method in the absence of feasible solutions and multi-tree method running out of time, the gap was computed as $100\frac{\text{BestUB}-\text{BestLB}}{\text{BestUB}}$, where BestUB is the best upper bound found for instance at hand considering all tested methods (see Tables B.1-B.2), and BestLB is the best lower bound found by the method at hand for the same instance. Tables B.6-B.8 report the results of the single and multi-tree versions of our Benders algorithms using the ϵ-technique with the aim of comparing the two choices of separation. The metrics we use for the comparison are the same as Tables B.3-B.5.

Table B.1: Instances of CPSCLP (part 1).

ID	s	Γ	R	BestUB	OpenFac	ID	s	Γ	R	BestUB	OpenFac
1	1	25	5,5	27239,74	41	39	2	50	6,25	27464,28	41
2	1	25	5,75	27162,26	41	40	2	100	5,5	28240,57	42
3	1	25	6	27080,35	41	41	2	100	5,75	28212,39	42
4	1	25	6,25	27005,71	41	42	2	100	6	28167,52	42
5	1	50	5,5	27952,05	41	43	2	100	6,25	28136,46	42
6	1	50	5,75	27856,05	41	44	2	150	5,5	28467,09	43
7	1	50	6	27769,04	41	45	2	150	5,75	28440,29	42
8	1	50	6,25	27669,14	41	46	2	150	6	28400,78	42
9	1	100	5,5	28450,44	42	47	2	150	6,25	28383,90	42
10	1	100	5,75	28406,33	42	48	2	200	5,5	28673,03	43
11	1	100	6	28368,57	42	49	2	200	6,25	28586,59	43
12	1	100	6,25	28316,45	42	50	2	300	5,5	28976,21	43
13	1	150	5,5	28680,39	42	51	2	300	5,75	28947,27	43
14	1	150	5,75	28631,80	42	52	2	300	6	28899,97	43
15	1	150	6	28597,00	42	53	2	300	6,25	28881,89	43
16	1	150	6,25	28557,92	42	54	2	400	5,5	29170,93	43
17	1	200	5,5	28898,38	42	55	2	400	6	29083,84	43
18	1	200	6	28807,26	42	56	2	500	5,75	29281,64	43
19	1	200	6,25	28763,63	43	57	2	500	6	29221,62	43
20	1	300	5,5	29228,79	42	58	2	500	6,25	29197,22	43
21	1	300	5,75	29168,65	42	59	3	25	5,5	28108,43	35
22	1	300	6	29123,02	42	60	3	25	5,75	28018,51	35
23	1	300	6,25	29072,71	43	61	3	25	6	27951,82	35
24	1	400	5,5	29477,27	43	62	3	25	6,25	27886,82	35
25	1	400	5,75	29401,55	42	63	3	50	5,5	28789,70	36
26	1	400	6	29350,02	43	64	3	50	5,75	28683,83	36
27	1	400	6,25	29295,53	43	65	3	50	6	28608,38	36
28	1	500	5,5	29634,08	43	66	3	50	6,25	28531,91	36
29	1	500	5,75	29559,19	43	67	3	100	5,5	29319,41	36
30	1	500	6	29504,86	43	68	3	100	5,75	29264,81	36
31	1	500	6,25	29450,37	43	69	3	100	6	29245,32	36
32	2	25	5,5	27037,92	41	70	3	100	6,25	29197,75	36
33	2	25	5,75	26973,19	41	71	3	150	5,5	29563,69	37
34	2	25	6	26909,97	41	72	3	150	6	29497,71	36
35	2	25	6,25	26862,46	41	73	3	150	6,25	29466,02	37
36	2	50	5,5	27687,26	42	74	3	200	5,5	29766,99	37
37	2	50	5,75	27614,66	42	75	3	200	5,75	29705,32	37
38	2	50	6	27533,15	42	76	3	200	6	29692,82	37

Table B.2: Instances of CPSCLP (part 2).

ID	s	Γ	R	BestUB	OpenFac	ID	s	Γ	R	BestUB	OpenFac
77	3	200	6,25	29660,42	37	115	4	400	5,75	30941,75	38
78	3	300	5,5	30089,87	37	116	4	400	6	30923,64	38
79	3	300	5,75	30022,47	37	117	4	400	6,25	30905,46	38
80	3	300	6	30010,20	37	118	4	500	5,5	31119,74	39
81	3	300	6,25	29979,06	37	119	4	500	5,75	31094,42	38
82	3	400	5,5	30285,64	37	120	4	500	6	31076,66	38
83	3	400	5,75	30214,90	37	121	4	500	6,25	31057,25	38
84	3	400	6	30202,67	37	122	5	25	5,5	27614,20	40
85	3	400	6,25	30169,95	37	123	5	25	5,75	27515,77	40
86	3	500	5,5	30450,82	37	124	5	25	6	27421,87	40
87	3	500	5,75	30372,28	37	125	5	25	6,25	27346,55	40
88	3	500	6	30356,26	37	126	5	50	5,5	28300,76	41
89	3	500	6,25	30321,74	37	127	5	50	5,75	28207,67	40
90	4	25	5,5	28769,80	36	128	5	50	6	28112,83	40
91	4	25	5,75	28688,22	36	129	5	50	6,25	28034,17	40
92	4	25	6	28626,61	36	130	5	100	5,5	28827,47	41
93	4	25	6,25	28566,13	36	131	5	100	5,75	28801,14	41
94	4	50	5,5	29455,33	37	132	5	100	6	28763,12	41
95	4	50	5,75	29349,76	37	133	5	100	6,25	28716,61	41
96	4	50	6	29265,07	37	134	5	150	5,5	29086,06	41
97	4	50	6,25	29186,63	37	135	5	150	5,75	29060,88	41
98	4	100	5,5	30005,30	38	136	5	150	6	29031,55	41
99	4	100	5,75	29978,57	37	137	5	150	6,25	29000,45	41
100	4	100	6	29954,70	37	138	5	200	5,5	29310,83	41
101	4	100	6,25	29926,45	37	139	5	200	5,75	29284,08	41
102	4	150	5,5	30240,51	38	140	5	200	6	29253,31	41
103	4	150	5,75	30219,40	37	141	5	200	6,25	29227,89	41
104	4	150	6	30204,67	37	142	5	300	5,5	29637,01	42
105	4	150	6,25	30190,79	37	143	5	300	5,75	29608,42	41
106	4	200	5,5	30441,22	38	144	5	300	6	29577,26	41
107	4	200	5,75	30418,71	38	145	5	300	6,25	29544,77	42
108	4	200	6	30403,79	38	146	5	400	5,5	29863,25	42
109	4	200	6,25	30389,21	38	147	5	400	5,75	29832,83	42
110	4	300	5,5	30763,43	38	148	5	400	6	29797,02	42
111	4	300	5,75	30739,45	38	149	5	400	6,25	29763,65	42
112	4	300	6	30723,60	38	150	5	500	5,5	30022,97	42
113	4	300	6,25	30707,81	38	151	5	500	5,75	29992,16	42
114	4	400	5,5	30968,95	38	152	5	500	6	29957,95	42
						153	5	500	6,25	29923,87	42

Table B.3: Performance of the exact procedures on the testbed (part 1). Runs reaching the time limit of 900 seconds are indicated by "TL".

	Ext-GUROBI			**GUROBI**			**STϵ-BEN$_2$**				**MTϵ-BEN$_2$**			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
1	TL	59,20	1673	308,53	0	102	84,52	0	52409	18	25,53	0	2455,18	10
2	TL	60,06	1759	376,45	0	143	39,95	0	6547	23	18,53	0	2247,78	8
3	TL	60,48	1569	392,77	0	97	22,70	0	48320	12	14,35	0	2122,67	5
4	TL	61,09	1559	391,38	0	305	38,02	0	54293	22	23,05	0	1984,63	7
5	TL	58,08	2718	369,75	0	19	92,87	0	7286	23	28,90	0	5167,36	10
6	TL	58,79	2213	444,53	0	38	70,84	0	33649	26	33,13	0	3875,08	11
7	TL	59,94	1577	329,34	0	178	43,02	0	33566	18	29,67	0	3655,60	9
8	TL	60,68	1579	447,25	0	46	113,10	0	35117	19	56,62	0	4500,46	12
9	TL	55,28	8685	392,58	0	36	48,11	0	34184	24	40,95	0	6715,73	10
10	TL	58,32	2983	426,73	0	24	39,78	0	33105	24	42,81	0	4592,82	10
11	TL	59,16	2945	401,47	0	50	36,14	0	30973	20	32,91	0	5077,80	9
12	TL	60,05	1637	612,98	0	25	97,77	0	35565	31	56,37	0	4500,25	15
13	TL	56,32	5801	365,86	0	36	36,87	0	31402	18	52,25	0	5272,92	12
14	TL	58,32	4020	405,90	0	33	35,80	0	31115	20	42,19	0,01	6808,60	4
15	TL	58,55	2135	496,43	0	15	56,55	0	17981	45	31,21	0	6564,18	10
16	TL	60,24	1566	519,91	0	164	69,44	0	34434	24	49,99	0	5358,25	15
17	TL	55,80	7444	366,97	0	18	169,62	0	17233	36	64,77	0	5694,80	14
18	TL	58,46	1732	562,01	0	19	102,04	0	10275	30	38,50	0	4621,58	11
19	TL	59,43	1726	549,85	0	100	60,87	0	12948	27	45,01	0	6232,85	12
20	TL	55,84	5194	297,68	0	295	77,22	0	34450	49	68,22	0	6665,00	16
21	TL	56,94	2055	679,59	0	66	52,49	0	32384	23	48,21	0	6176,33	11
22	TL	58,02	1942	504,38	0	111	51,44	0	31207	35	43,46	0	3442,80	14
23	TL	59,50	1607	705,04	0	25	40,36	0	10603	21	50,96	0	4156,62	12
24	TL	56,43	4204	428,51	0	401	208,53	0	17892	55	82,65	0	4557,50	23
25	TL	57,19	2703	416,35	0	429	178,43	0	17336	57	100,74	0	6828,93	14
26	TL	57,45	1929	350,79	0	282	71,80	0	8701	28	124,56	0	4506,54	12
27	TL	58,58	1678	569,68	0	180	62,96	0	12706	22	39,31	0	5035,36	10
28	TL	56,02	4406	593,17	0	210	384,74	0	40447	57	68,23	0	4545,60	19
29	TL	56,12	2068	415,38	0	161	182,72	0	37938	29	49,30	0	5250,50	13
30	TL	58,11	1661	368,05	0	315	83,76	0	34029	20	41,06	0	6753,55	10
31	TL	59,28	1595	499,51	0	139	36,47	0	4274	23	30,72	0	5593,13	7
32	TL	57,99	3645	369,81	0	344	126,30	0	7410	30	31,85	0	2172,36	10
33	TL	59,56	1483	339,01	0	200	69,43	0	7891	22	35,43	0	1703,08	11
34	TL	58,90	5521	329,45	0	453	116,02	0	7820	31	24,83	0	1389,22	8
35	TL	60,48	3958	353,08	0	348	40,59	0	5612	12	36,95	0	968,78	8
36	TL	57,79	2930	436,19	0	222	149,38	0	22245	32	71,36	0	2351,04	23
37	TL	58,79	1596	469,25	0	344	137,62	0	16935	26	62,03	0	2348,38	15
38	TL	59,39	1594	443,41	0	968	136,17	0	15798	32	51,76	0	1827,73	14
39	TL	59,65	1531	507,80	0	676	127,56	0	12827	25	48,39	0	2111,43	13
40	TL	56,50	4206	446,64	0	258	169,78	0	17333	47	148,35	0	3084,66	31
41	TL	57,88	2087	514,22	0	70	83,66	0	18409	52	176,48	0	3219,17	29
42	TL	58,63	1966	426,08	0	124	146,97	0	16777	40	126,67	0	3629,62	28
43	TL	59,38	1524	433,70	0	95	341,45	0	20984	21	76,69	0	4082,00	14
44	TL	57,21	3210	449,88	0	430	273,04	0	18594	50	176,16	0	4265,24	28
45	TL	57,28	4188	354,73	0	383	268,25	0	16461	51	267,68	0	5025,69	28
46	TL	58,38	2502	323,97	0	265	454,27	0	18607	63	87,04	0	3409,76	24
47	TL	59,85	2086	529,80	0	53	73,41	0	15448	35	47,85	0	3872,69	12
48	TL	56,69	2200	381,80	0	147	434,02	0	22949	76	253,88	0	3754,39	30
49	TL	59,75	1679	492,46	0	349	29,02	0	7464	25	44,10	0	3066,20	14
50	TL	56,01	2533	342,57	0	56	142,44	0	13114	44	176,79	0	3535,42	32
51	TL	57,62	2505	343,76	0	196	382,76	0	18972	86	168,88	0	5048,27	29

Table B.4: Performance of the exact procedures on the testbed (part 2). Runs reaching the time limit of 900 seconds are indicated by "TL".

	Ext-GUROBI			GUROBI			STϵ-BEN$_2$				MTϵ-BEN$_2$			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
52	TL	58,61	2172	306,18	0	249	391,81	0	19718	42	100,90	0	4387,91	21
53	TL	58,94	1589	744,31	0	291	99,73	0	14488	23	84,93	0	4303,31	15
54	TL	55,58	2347	409,46	0	416	606,62	0	3632	42	331,72	0	3190,33	47
55	TL	57,93	1585	572,31	0	79	591,09	0,26	2633	5	243,01	0	4629,81	31
56	TL	56,21	2628	680,26	0	298	TL	0,47	10972	13	207,33	0	3926,31	41
57	TL	57,26	2403	447,42	0	554	250,05	0	14241	87	145,21	0	3706,32	27
58	TL	58,60	1528	543,78	0	13	60,27	0	12305	23	44,04	0	4081,64	13
59	TL	59,30	6007	379,08	0	119	40,24	0	35887	14	29,67	0	4464,75	7
60	TL	60,61	2367	284,72	0	25	33,91	0	12909	15	31,99	0	7112,63	7
61	TL	61,23	4250	322,25	0	66	38,91	0	27138	12	25,13	0	2746,43	6
62	TL	63,20	1651	361,25	0	26	34,41	0	10480	8	20,27	0	4151,33	5
63	TL	58,59	3849	314,39	0	30	63,92	0	18773	31	58,22	0	8848,58	11
64	TL	60,26	1929	253,84	0	1	45,42	0	16264	14	43,11	0	8528,33	8
65	TL	60,93	2273	302,70	0	407	93,94	0	14908	22	54,33	0	8831,78	8
66	TL	61,84	2156	391,95	0	233	80,83	0	16244	20	51,60	0	10989,50	7
67	TL	55,89	10376	394,24	0	85	58,10	0	15290	38	TL	1,39	6478,20	4
68	TL	58,26	6270	374,05	0	41	49,86	0	13492	33	TL	0,24	1285,00	2
69	TL	60,52	1934	404,79	0	46	162,66	0	16191	26	93,57	0	20379,08	12
70	TL	61,43	2072	397,11	0	83	50,44	0	15472	21	69,29	0	14424,33	8
71	TL	56,00	9869	216,39	0	410	78,53	0	17577	29	102,60	0	7855,40	19
72	TL	59,74	4459	333,06	0	50	41,40	0	10014	23	77,20	0	12628,69	12
73	TL	61,32	3656	664,75	0	116	68,50	0	13002	21	56,93	0	14201,60	9
74	TL	56,93	6450	495,12	0	231	148,23	0	57661	45	TL	1,17	11983,00	3
75	TL	59,05	3062	554,48	0	108	61,26	0	52636	16	58,61	0	14578,50	9
76	TL	59,54	2343	339,07	0	193	43,89	0	49224	22	87,15	0	10201,76	16
77	TL	61,48	1647	447,53	0	96	377,27	0	42990	55	78,61	0	17195,36	10
78	TL	56,71	6702	444,47	0	126	278,28	0	21520	89	TL	0,42	10142,33	14
79	TL	57,90	4737	248,94	0	116	69,62	0	13657	35	77,33	0	8646,38	15
80	TL	60,10	2657	434,90	0	15	41,64	0	6473	24	63,77	0	15033,36	10
81	TL	60,15	1972	386,78	0	1	34,06	0	9038	21	57,33	0	13746,73	10
82	TL	57,12	4630	365,26	0	339	293,85	0	6450	29	TL	0,03	9319,00	16
83	TL	58,41	2571	233,69	0	1	130,95	0	56385	23	66,77	0	12757,87	14
84	TL	58,86	1854	302,13	0	1	30,30	0	48489	14	99,85	0	17825,54	12
85	TL	60,45	1492	535,38	0	1	43,33	0	36169	20	74,98	0	13722,09	10
86	TL	56,05	7294	443,79	0	259	215,72	0	21334	82	256,79	0	11377,06	33
87	TL	57,57	4658	337,27	0	24	652,23	0	20990	45	67,81	0	9156,50	11
88	TL	58,72	2639	370,19	0	23	63,28	0	8212	28	TL	0,39	10578,00	6
89	TL	59,76	1902	348,69	0	45	97,08	0	10711	23	315,69	0,07	4150,80	4
90	TL	62,93	532	434,21	0	43	34,49	0	7625	15	14,64	0	8705,00	4
91	TL	-	1	320,02	0	512	17,63	0	5416	13	16,99	0	11139,80	4
92	TL	62,46	1070	367,79	0	551	29,09	0	5279	10	25,30	0	11695,20	4
93	TL	63,26	1509	425,42	0	177	40,72	0	5387	15	31,04	0	6817,57	6
94	TL	59,76	1678	381,32	0	103	22,48	0	4754	17	57,47	0	11655,60	9
95	TL	60,88	2333	260,50	0	18	38,66	0	5468	10	45,21	0	12743,57	6
96	TL	61,73	1477	439,49	0	115	45,28	0	11156	11	55,72	0	13867,11	8
97	TL	62,61	1505	458,36	0	246	42,94	0	6798	14	41,94	0	9426,70	9
98	TL	57,82	6177	260,73	0	283	18,89	0	4966	24	TL	0,03	11236,80	4
99	TL	60,24	3381	500,46	0	47	38,99	0	9855	24	TL	0,31	13626,00	5
100	TL	61,31	1905	395,18	0	39	13,64	0	7790	18	61,81	0	14786,82	10
101	TL	62,33	1466	503,26	0	22	15,70	0	4408	14	75,79	0	24469,30	9
102	TL	58,44	6753	214,83	0	188	57,75	0	8758	26	TL	0,15	14899,89	8

Table B.5: Performance of the exact procedures on the testbed (part 3). Runs reaching the time limit of 900 seconds are indicated by "TL".

	Ext-GUROBI			GUROBI			STϵ-BEN$_2$				MTϵ-BEN$_2$			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
103	TL	59,87	2621	363,29	0	99	61,57	0	10509	23	72,15	0	12724,00	12
104	TL	60,90	1925	415,56	0	28	18,29	0	7792	18	84,67	0	19106,23	12
105	TL	61,92	2189	478,66	0	61	23,65	0	8908	14	88,61	0	19102,46	12
106	TL	58,75	3499	219,25	0	192	18,62	0	4183	19	76,34	0	9639,00	13
107	TL	60,21	1882	335,21	0	64	60,22	0	7885	25	92,26	0	15314,14	13
108	TL	61,30	1502	379,03	0	83	14,72	0	6037	18	89,45	0	20072,50	9
109	TL	62,03	1653	TL	1,16	1	60,37	0	20543	25	64,65	0	13826,40	9
110	TL	57,88	3859	320,76	0	405	29,25	0	5544	16	110,32	0	13377,54	12
111	TL	59,46	3217	351,74	0	127	70,08	0	12416	19	TL	0,004	11293,40	4
112	TL	60,28	1855	366,18	0	204	28,68	0	8212	19	114,90	0	25354,46	12
113	TL	61,66	1534	672,75	0	62	32,35	0	8090	19	74,25	0	12823,82	10
114	TL	58,24	2277	440,30	0	58	19,54	0	4579	19	124,26	0	13152,05	19
115	TL	59,08	1951	417,83	0	183	38,48	0	7837	25	76,18	0	10683,62	12
116	TL	60,60	1453	531,86	0	108	22,01	0	6846	21	62,93	0	10258,92	11
117	TL	61,16	1497	513,56	0	146	115,50	0	23323	25	TL	0,43	12054,33	5
118	TL	57,48	4943	312,10	0	174	36,84	0	5917	29	110,74	0	19407,94	16
119	TL	58,90	2538	321,24	0	73	34,77	0	8577	21	TL	0,34	5060,14	6
120	TL	59,90	1671	367,14	0	54	27,06	0	7415	19	92,17	0	13914,00	11
121	TL	61,13	1512	353,40	0	75	33,27	0	8345	16	TL	0,001	17082,13	14
122	TL	57,51	3068	289,84	0	168	80,79	0	7776	21	35,03	0	2495,00	8
123	TL	57,29	5126	272,01	0	155	80,62	0	8317	28	30,61	0	2050,09	10
124	TL	58,54	4201	291,19	0	57	64,99	0	7622	20	42,72	0	1561,70	9
125	TL	59,63	2818	411,42	0	164	34,63	0	5873	13	31,03	0	2179,44	8
126	TL	56,39	3915	345,55	0	133	41,93	0	6886	17	34,01	0	3814,91	10
127	TL	57,56	2040	421,95	0	167	40,16	0	9004	14	28,52	0	3616,00	9
128	TL	58,61	2707	302,65	0	188	145,59	0	10785	23	30,27	0	2895,89	8
129	TL	59,57	2012	389,37	0	105	70,78	0	9556	14	31,14	0	4039,29	6
130	TL	53,93	8614	281,01	0	191	443,09	0	14115	34	37,88	0	3360,75	11
131	TL	56,64	4641	502,86	0	77	84,52	0	10489	25	37,88	0	5002,91	10
132	TL	58,16	1993	322,66	0	318	211,37	0	11196	25	38,89	0	5101,58	11
133	TL	58,83	1581	383,29	0	144	435,76	0	11346	23	37,52	0	5222,20	9
134	TL	53,71	8310	428,46	0	201	350,88	0	15731	39	54,16	0	5541,36	13
135	TL	55,34	6150	260,03	0	93	54,34	0	9450	19	38,79	0	4049,36	10
136	TL	58,01	3517	438,07	0	361	92,34	0	11821	27	47,65	0	3791,07	13
137	TL	58,65	2613	403,51	0	511	34,07	0	8860	14	40,57	0	5084,00	11
138	TL	54,84	5305	352,51	0	287	161,25	0	12224	29	54,28	0	5593,64	13
139	TL	56,49	2951	372,97	0	386	87,49	0	11994	22	69,62	0	7986,21	13
140	TL	57,79	2306	529,34	0	421	18,17	0	6810	19	43,85	0	7838,27	10
141	TL	59,22	1670	469,44	0	430	29,84	0	7371	21	65,36	0	9565,08	12
142	TL	54,06	6569	329,08	0	557	402,60	0	15702	40	51,43	0	4305,50	13
143	TL	56,12	3915	353,38	0	724	391,73	0	13418	35	52,39	0	5543,93	14
144	TL	57,55	2369	553,27	0	230	200,26	0	13059	26	45,75	0	5122,39	17
145	TL	58,21	1793	459,78	0	551	180,89	0	12057	23	65,79	0	7344,53	16
146	TL	54,52	4205	375,73	0	412	140,38	0	13264	31	78,73	0	6022,95	18
147	TL	56,40	2616	389,91	0	644	70,24	0	11395	28	74,62	0	5953,29	16
148	TL	57,20	1651	749,69	0	223	167,93	0	13985	21	67,90	0	9372,69	15
149	TL	58,68	1587	576,97	0	207	23,92	0	7371	18	63,42	0	12804,29	13
150	TL	54,07	6470	346,15	0	647	756,13	0	23722	74	94,68	0	5418,13	22
151	TL	55,21	4680	406,63	0	683	499,26	0	13988	40	59,99	0	4787,22	17
152	TL	56,30	2028	322,37	0	534	157,85	0	11933	21	56,22	0	6451,77	12
153	TL	58,04	1682	474,56	0	222	496,39	0	11144	25	51,13	0	8109,93	13

Table B.6: Comparison between STϵ-BEN$_1$, MTϵ-BEN$_1$, STϵ-BEN$_2$, MTϵ-BEN$_2$ on the testbed (part 1). Runs reaching the time limit of 900 seconds are indicated by "TL".

	STϵ-BEN$_1$				**MTϵ-BEN$_1$**				**STϵ-BEN$_2$**				**MTϵ-BEN$_2$**			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
1	TL	0,33	25142	470	TL	0,33	634,28	249	84,52	0	52409	18	25,53	0	2455,18	10
2	TL	0,31	24310	422	TL	0,32	710,75	251	39,95	0	6547	23	18,53	0	2247,78	8
3	TL	0,27	27745	499	TL	0,31	536,86	244	22,70	0	48320	12	14,35	0	2122,67	5
4	TL	0,40	40329	402	TL	0,28	557,10	265	38,02	0	54293	22	23,05	0	1984,63	7
5	TL	0,99	21954	64	TL	0,36	2669,42	149	92,87	0	7286	23	28,90	0	5167,36	10
6	TL	0,66	24953	110	TL	0,35	2571,91	165	70,84	0	33649	26	33,13	0	3875,08	11
7	TL	0,83	25108	72	TL	0,50	1866,18	115	43,02	0	33566	18	29,67	0	3655,60	9
8	TL	0,62	25150	119	TL	0,48	1599,34	111	113,10	0	35117	19	56,62	0	4500,46	12
9	TL	0,54	26008	32	TL	0,13	7200,25	55	48,11	0	34184	24	40,95	0	6715,73	10
10	TL	0,40	23938	31	TL	0,04	8291,91	44	39,78	0	33105	24	42,81	0	4592,82	10
11	TL	0,49	25741	37	TL	0,12	6054,18	77	36,14	0	30973	20	32,91	0	5077,80	9
12	TL	0,31	27420	83	TL	0,22	6945,81	75	97,77	0	35565	31	56,37	0	4500,25	15
13	TL	0,67	36797	24	TL	0,28	8664,33	33	36,87	0	31402	18	52,25	0	5272,92	12
14	TL	0,26	29314	30	TL	0,20	7460,66	41	35,80	0	31115	20	42,19	0,01	6808,60	4
15	TL	0,61	29931	22	TL	0,07	7410,04	54	56,55	0	17981	45	31,21	0	6564,18	10
16	TL	0,36	23859	27	TL	0,19	7338,49	53	69,44	0	34434	24	49,99	0	5358,25	15
17	TL	0,20	25456	47	66,89	1,32	3735,57	6	169,62	0	17233	36	64,77	0	5694,80	14
18	TL	0,21	27550	52	TL	0,25	9030,57	49	102,04	0	10275	30	38,50	0	4621,58	11
19	TL	0,23	27365	57	TL	0,27	8810,28	47	60,87	0	12948	27	45,01	0	6232,85	12
20	TL	1,14	25720	21	TL	0,66	9902,57	30	77,22	0	34450	49	68,22	0	6665,00	16
21	TL	0,91	26689	22	TL	0,54	9198,07	29	52,49	0	32384	23	48,21	0	6176,33	11
22	TL	0,57	27626	24	TL	0,49	9571,91	34	51,44	0	31207	35	43,46	0	3442,80	14
23	TL	0,45	32035	24	TL	0,68	9501,76	24	40,36	0	10603	21	50,96	0	4156,62	12
24	TL	1,95	26810	22	TL	1,35	11089,92	12	208,53	0	17892	55	82,65	0	4557,50	23
25	TL	1,92	34775	17	TL	0,84	7928,74	26	178,43	0	17336	57	100,74	0	6828,93	14
26	TL	1,25	28543	22	TL	0,67	11252,30	30	71,80	0	8701	28	124,56	0	4506,54	12
27	TL	0,97	26101	27	TL	0,67	9358,06	33	62,96	0	12706	22	39,31	0	5035,36	10
28	TL	2,19	26582	16	TL	1,25	8868,30	20	384,74	0	40447	57	68,23	0	4545,60	19
29	TL	1,65	30821	19	TL	0,97	10835,43	30	182,72	0	37938	29	49,30	0	5250,50	13
30	TL	1,43	33182	14	TL	1,18	9081,82	21	83,76	0	34029	20	41,06	0	6753,55	10
31	TL	0,81	32596	25	TL	0,91	10423,19	21	36,47	0	4274	23	30,72	0	5593,13	7
32	TL	0,49	39456	324	TL	0,54	1300,64	157	126,30	0	7410	30	31,85	0	2172,36	10
33	TL	0,39	36764	476	TL	0,60	912,60	121	69,43	0	7891	22	35,43	0	1703,08	11
34	TL	0,32	34657	658	TL	0,43	1143,78	190	116,02	0	7820	31	24,83	0	1389,22	8
35	TL	0,38	34150	420	TL	0,43	1148,28	186	40,59	0	5612	12	36,95	0	968,78	8
36	TL	1,80	31756	47	TL	0,73	3705,34	96	149,38	0	22245	32	71,36	0	2351,04	23
37	TL	1,78	31374	34	TL	0,89	2992,58	65	137,62	0	16935	26	62,03	0	2348,38	15
38	TL	1,36	40430	85	TL	0,69	3054,23	102	136,17	0	15798	32	51,76	0	1827,73	14
39	TL	1,33	30313	56	TL	0,85	2843,93	71	127,56	0	12827	25	48,39	0	2111,43	13
40	TL	1,48	27329	14	TL	0,88	9465,65	20	169,78	0	17333	47	148,35	0	3084,66	31
41	TL	0,87	27705	19	TL	0,88	7687,82	21	83,66	0	18409	52	176,48	0	3219,17	29
42	TL	1,14	27611	22	TL	0,80	8652,06	33	146,97	0	16777	40	126,67	0	3629,62	28
43	TL	1,28	21839	17	TL	0,92	9144,59	27	341,45	0	20984	21	76,69	0	4082,00	14
44	TL	1,48	23521	15	TL	0,85	8798,36	28	273,04	0	18594	50	176,16	0	4265,24	28
45	TL	1,42	24721	18	TL	1,04	7956,56	16	268,25	0	16461	51	267,68	0	5025,69	28
46	TL	1,00	24424	20	TL	0,82	9834,19	32	454,27	0	18607	63	87,04	0	3409,76	24
47	TL	1,90	23735	12	TL	0,98	11382,38	16	73,41	0	15448	35	47,85	0	3872,69	12
48	TL	1,60	16969	12	TL	1,36	10544,00	13	434,02	0	22949	76	253,88	0	3754,39	30
49	TL	1,55	28175	13	TL	1,04	9122,76	21	29,02	0	7464	25	44,10	0	3066,20	14
50	TL	1,77	23221	20	TL	1,46	10045,50	18	142,44	0	13114	44	176,79	0	3535,42	32
51	TL	1,83	28937	15	TL	1,65	11296,50	14	382,76	0	18972	86	168,88	0	5048,27	29

Table B.7: Comparison between STϵ-BEN$_1$, MTϵ-BEN$_1$, STϵ-BEN$_2$, MTϵ-BEN$_2$ on the testbed (part 2). Runs reaching the time limit of 900 seconds are indicated by "TL".

	STϵ-BEN$_1$				**MTϵ-BEN$_1$**				**STϵ-BEN$_2$**				**MTϵ-BEN$_2$**			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
52	TL	2,87	27103	11	TL	1,35	9615,43	23	391,81	0	19718	42	100,90	0	4387,91	21
53	TL	1,70	25780	16	TL	1,48	9154,31	16	99,73	0	14488	23	84,93	0	4303,31	15
54	TL	2,26	34552	17	TL	1,81	7321,86	13	606,62	0	3632	42	331,72	0	3190,33	47
55	TL	2,24	37380	12	TL	1,84	11125,45	11	591,09	0,26	2633	5	243,01	0	4629,81	31
56	TL	2,39	25744	11	TL	2,25	10706,73	11	TL	0,47	10972	13	207,33	0	3926,31	41
57	TL	2,05	25726	15	TL	2,34	9367,44	8	250,05	0	14241	87	145,21	0	3706,32	27
58	TL	1,69	28211	13	TL	2,00	9248,18	10	60,27	0	12305	23	44,04	0	4081,64	13
59	TL	0,61	39500	160	TL	0,70	2518,26	82	40,24	0	35887	14	29,67	0	4464,75	7
60	TL	0,16	48174	314	TL	0,42	2544,79	161	33,91	0	12909	15	31,99	0	7112,63	7
61	TL	0,36	47314	388	TL	0,44	2697,70	161	38,91	0	27138	12	25,13	0	2746,43	6
62	TL	0,35	50100	486	TL	0,42	2545,24	149	34,41	0	10480	8	20,27	0	4151,33	5
63	TL	1,26	24309	43	TL	0,44	6024,84	127	63,92	0	18773	31	58,22	0	8848,58	11
64	TL	1,18	27697	46	TL	0,42	4695,28	133	45,42	0	16264	14	43,11	0	8528,33	8
65	TL	1,14	26134	53	TL	0,77	5193,65	62	93,94	0	14908	22	54,33	0	8831,78	8
66	TL	0,76	30377	158	TL	0,47	4277,47	151	80,83	0	16244	20	51,60	0	10989,50	7
67	TL	0,93	26894	31	TL	0,31	11376,29	62	58,10	0	15290	38	TL	1,39	6478,20	4
68	TL	0,15	34126	92	TL	1,33	9796,43	6	49,86	0	13492	33	TL	0,24	1285,00	2
69	TL	0,04	34109	249	TL	0,35	10724,83	53	162,66	0	16191	26	93,57	0	20379,08	12
70	TL	0,59	30067	44	TL	0,39	8820,61	59	50,44	0	15472	21	69,29	0	14424,33	8
71	TL	1,62	47719	21	TL	0,28	8639,59	70	78,53	0	17577	29	102,60	0	7855,40	19
72	TL	0,97	48700	30	TL	2,41	9259,80	4	41,40	0	10014	23	77,20	0	12628,69	12
73	TL	1,39	47710	23	TL	0,27	11694,90	60	68,50	0	13002	21	56,93	0	14201,60	9
74	TL	1,85	31647	17	TL	1,13	16467,83	12	148,23	0	57661	45	TL	1,17	11983,00	3
75	TL	0,95	30166	35	TL	0,88	10471,56	15	61,26	0	52636	16	58,61	0	14578,50	9
76	TL	0,84	43117	35	TL	0,45	10542,40	49	43,89	0	49224	22	87,15	0	10201,76	16
77	TL	1,27	24528	23	78,92	1,27	13890,75	7	377,27	0	42990	55	78,61	0	17195,36	10
78	TL	1,73	31838	16	TL	1,67	12108,89	8	278,28	0	21520	89	TL	0,42	10142,33	14
79	TL	1,69	31830	21	TL	0,84	13021,36	27	69,62	0	13657	35	77,33	0	8646,38	15
80	TL	1,56	33708	20	TL	0,67	11848,43	47	41,64	0	6473	24	63,77	0	15033,36	10
81	TL	1,72	37026	15	TL	1,50	18374,22	8	34,06	0	9038	21	57,33	0	13746,73	10
82	TL	1,30	45277	19	TL	1,22	11329,79	23	293,85	0	6450	29	TL	0,03	9319,00	16
83	TL	1,67	27269	26	TL	1,09	14449,74	27	130,95	0	56385	23	66,77	0	12757,87	14
84	TL	2,21	31289	25	TL	0,81	16081,14	44	30,30	0	48489	14	99,85	0	17825,54	12
85	TL	1,58	50284	19	TL	0,84	14002,65	43	43,33	0	36169	20	74,98	0	13722,09	10
86	TL	2,50	30741	14	TL	1,40	9530,00	29	215,72	0	21334	82	256,79	0	11377,06	33
87	TL	1,68	48351	21	TL	1,23	11097,63	27	652,23	0	20990	45	67,81	0	9156,50	11
88	TL	0,94	41503	28	TL	1,30	11643,12	25	63,28	0	8212	28	TL	0,39	10578,00	6
89	TL	2,01	33741	17	TL	1,26	14320,85	25	97,08	0	10711	23	315,69	0,07	4150,80	4
90	TL	0,27	30855	542	TL	0,63	2413,11	83	34,49	0	7625	15	14,64	0	8705,00	4
91	TL	0,32	38269	496	TL	0,36	1982,10	191	17,63	0	5416	13	16,99	0	11139,80	4
92	TL	0,30	35937	578	TL	0,34	1689,57	191	29,09	0	5279	10	25,30	0	11695,20	4
93	TL	0,25	38629	663	TL	0,35	1983,50	182	40,72	0	5387	15	31,04	0	6817,57	6
94	TL	0,69	24559	114	TL	0,34	3816,97	149	22,48	0	4754	17	57,47	0	11655,60	9
95	TL	0,60	27103	178	TL	0,35	3056,47	163	38,66	0	5468	10	45,21	0	12743,57	6
96	TL	0,58	27252	188	TL	0,35	3864,94	151	45,28	0	11156	11	55,72	0	13867,11	8
97	TL	0,44	36815	349	TL	0,38	3763,53	155	42,94	0	6798	14	41,94	0	9426,70	9
98	TL	0,64	38071	40	TL	0,05	11859,07	66	18,89	0	4966	24	TL	0,03	11236,80	4
99	TL	0,13	36062	76	TL	0,04	10962,34	68	38,99	0	9855	24	TL	0,31	13626,00	5
100	TL	0,85	34662	28	TL	0,69	8039,17	11	13,64	0	7790	18	61,81	0	14786,82	10
101	TL	0,72	35760	31	TL	0,20	9525,53	68	15,70	0	4408	14	75,79	0	24469,30	9
102	TL	0,49	36356	37	TL	0,14	11769,07	56	57,75	0	8758	26	TL	0,15	14899,89	8

Table B.8: Comparison between STϵ-BEN$_1$, MTϵ-BEN$_1$, STϵ-BEN$_2$, MTϵ-BEN$_2$, on the testbed (part 3). Runs reaching the time limit of 900 seconds are indicated by "TL".

	STϵ-BEN$_1$				MTϵ-BEN$_1$				STϵ-BEN$_2$				MTϵ-BEN$_2$			
ID	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**	**Time[s]**	**Gap[%]**	**Nodes**	**BenCuts**
103	TL	0,58	35034	45	TL	0,94	4121,78	8	61,57	0	10509	23	72,15	0	12724,00	12
104	TL	0,42	34677	56	TL	0,63	4248,33	11	18,29	0	7792	18	84,67	0	19106,23	12
105	TL	0,24	34085	94	TL	1,22	7396,57	6	23,65	0	8908	14	88,61	0	19102,46	12
106	TL	0,90	27397	38	TL	0,26	14291,08	25	18,62	0	4183	19	76,34	0	9639,00	13
107	TL	0,90	24502	25	TL	0,23	20584,57	46	60,22	0	7885	25	92,26	0	15314,14	13
108	TL	0,68	28563	38	TL	0,28	15105,79	42	14,72	0	6037	18	89,45	0	20072,50	9
109	TL	0,86	26882	47	TL	1,27	17165,14	6	60,37	0	20543	25	64,65	0	13826,40	9
110	TL	1,88	23885	16	TL	0,39	14259,21	38	29,25	0	5544	16	110,32	0	13377,54	12
111	TL	1,62	27126	18	465,03	0,59	15652,44	17	70,08	0	12416	19	TL	0,004	11293,40	4
112	TL	0,43	30410	57	TL	0,78	13483,00	12	28,68	0	8212	19	114,90	0	25354,46	12
113	TL	1,77	27444	20	TL	0,50	18146,48	29	32,35	0	8090	19	74,25	0	12823,82	10
114	TL	0,63	25712	51	786,73	1,14	14424,40	9	19,54	0	4579	19	124,26	0	13152,05	19
115	TL	1,04	23670	36	TL	0,64	19988,36	33	38,48	0	7837	25	76,18	0	10683,62	12
116	TL	0,84	26099	34	TL	0,79	12102,45	21	22,01	0	6846	21	62,93	0	10258,92	11
117	TL	0,88	28733	33	TL	0,46	20081,28	29	115,50	0	23323	25	TL	0,43	12054,33	5
118	TL	1,48	31814	19	TL	1,20	10087,31	15	36,84	0	5917	29	110,74	0	19407,94	16
119	TL	0,50	36218	69	TL	1,34	8299,20	9	34,77	0	8577	21	TL	0,34	5060,14	6
120	TL	1,39	35405	32	530,56	1,15	12872,67	11	27,06	0	7415	19	92,17	0	13914,00	11
121	TL	1,57	29639	19	TL	0,94	16034,96	24	33,27	0	8345	16	TL	0,001	17082,13	14
122	TL	0,83	31676	120	TL	0,70	2465,43	122	80,79	0	7776	21	35,03	0	2495,00	8
123	TL	0,79	30319	137	TL	0,90	1589,55	67	80,62	0	8317	28	30,61	0	2050,09	10
124	TL	0,51	35800	244	TL	0,60	1920,16	134	64,99	0	7622	20	42,72	0	1561,70	9
125	TL	0,55	38300	247	TL	0,51	1535,47	149	34,63	0	5873	13	31,03	0	2179,44	8
126	TL	1,89	31860	39	TL	0,85	3726,21	73	41,93	0	6886	17	34,01	0	3814,91	10
127	TL	1,75	28423	57	TL	0,85	3853,81	69	40,16	0	9004	14	28,52	0	3616,00	9
128	TL	1,53	29578	58	TL	0,77	3469,18	85	145,59	0	10785	23	30,27	0	2895,89	8
129	TL	1,57	33493	139	TL	0,84	3080,27	88	70,78	0	9556	14	31,14	0	4039,29	6
130	TL	1,06	50041	23	226,74	0,83	8611,45	19	443,09	0	14115	34	37,88	0	3360,75	11
131	TL	1,02	52096	21	TL	0,64	7464,77	39	84,52	0	10489	25	37,88	0	5002,91	10
132	TL	1,10	53730	22	TL	0,61	9904,51	39	211,37	0	11196	25	38,89	0	5101,58	11
133	TL	1,55	54902	24	TL	0,90	8945,49	35	435,76	0	11346	23	37,52	0	5222,20	9
134	TL	0,97	49626	17	117,37	1,81	8277,00	7	350,88	0	15731	39	54,16	0	5541,36	13
135	TL	1,92	51073	15	TL	1,11	8325,31	15	54,34	0	9450	19	38,79	0	4049,36	10
136	TL	0,90	25745	21	TL	0,68	7440,53	45	92,34	0	11821	27	47,65	0	3791,07	13
137	TL	0,92	42426	20	TL	0,67	9790,59	49	34,07	0	8860	14	40,57	0	5084,00	11
138	TL	1,58	22950	14	TL	0,76	8125,08	38	161,25	0	12224	29	54,28	0	5593,64	13
139	TL	0,98	31770	31	TL	0,98	11693,66	29	87,49	0	11994	22	69,62	0	7986,21	13
140	TL	1,42	34371	16	TL	0,84	12821,84	37	18,17	0	6810	19	43,85	0	7838,27	10
141	TL	1,12	25191	25	TL	0,97	10681,56	36	29,84	0	7371	21	65,36	0	9565,08	12
142	TL	1,91	47391	15	TL	1,15	15366,43	30	402,60	0	15702	40	51,43	0	4305,50	13
143	TL	1,93	63662	12	TL	1,88	7780,83	11	391,73	0	13418	35	52,39	0	5543,93	14
144	TL	1,81	53352	15	TL	1,06	13098,12	33	200,26	0	13059	26	45,75	0	5122,39	17
145	TL	1,22	41970	16	TL	1,33	16540,65	22	180,89	0	12057	23	65,79	0	7344,53	16
146	TL	1,55	38465	17	TL	1,51	13188,71	24	140,38	0	13264	31	78,73	0	6022,95	18
147	TL	2,00	35951	16	TL	1,39	11879,11	28	70,24	0	11395	28	74,62	0	5953,29	16
148	TL	2,01	36924	16	508,32	1,65	12333,44	15	167,93	0	13985	21	67,90	0	9372,69	15
149	TL	1,97	34962	16	TL	1,47	20889,16	25	23,92	0	7371	18	63,42	0	12804,29	13
150	TL	2,66	32391	16	TL	1,55	16715,22	23	756,13	0	23722	74	94,68	0	5418,13	22
151	TL	2,21	30141	16	TL	1,54	15266,79	29	499,26	0	13988	40	59,99	0	4787,22	17
152	TL	2,38	27368	22	TL	1,55	16010,70	30	157,85	0	11933	21	56,22	0	6451,77	12
153	TL	1,31	32177	13	TL	1,43	18229,13	31	496,39	0	11144	25	51,13	0	8109,93	13

Bibliography

[1] R. Aboolian, O. Berman, and D. Krass. Profit maximizing distributed service system design with congestion and elastic demand. *Transportation Science*, 46(2):247–261, 2012.

[2] T. Achterberg, R. E. Bixby, Z. Gu, E. Rothberg, and D. Weninger. Presolve reductions in mixed integer programming. *INFORMS Journal on Computing*, 32(2):473–506, 2020.

[3] T. Achterberg, T. Koch, and A. Martin. Branching rules revisited. *Operations Research Letters*, 33(1):42–54, 2005.

[4] B. Adenso-Diaz and F. Rodriguez. A simple search heuristic for the mclp: Application to the location of ambulance bases in a rural region. *Omega*, 25(2):181–187, 1997.

[5] D. Ageyev and A. Al-Ansari. Lte ran and services multi-period planning. In *2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*, pages 272–274. IEEE, 2015.

[6] D. Ageyev and A. Al-Anssari. Optimization model for multi-time period lte network planning. In *2014 First International Scientific-Practical Conference Problems of Infocommunications Science and Technology*, pages 29–30. IEEE, 2014.

[7] R. V. Akhpashev and A. V. Andreev. Cost 231 hata adaptation model for urban conditions in lte networks. In *2016 17th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM)*, pages 64–66. IEEE, 2016.

[8] M. S. Aktürk, A. Atamtürk, and S. Gürel. A strong conic quadratic reformulation for machine-job assignment with controllable processing times. *Operations Research Letters*, 37(3):187–191, 2009.

[9] G. Alexandris, M. Dimopoulou, and I. Giannikos. A three-phase methodology for

developing or evaluating bank networks. *International Transactions in Operational Research*, 15(2):215–237, 2008.

[10] A. Amiri. Solution procedures for the service system design problem. *Computers & operations research*, 24(1):49–60, 1997.

[11] Antonio Ornatelli. Network Control Systems for 5G Multi-Connectivity, 2020. https://iris.uniroma1.it/retrieve/handle/11573/1562255/1878817/Tesi_dottorato_Ornatelli.pdf Last accessed July 18, 2022.

[12] S. Atta. An improved harmony search algorithm using opposition-based learning and local search for solving the maximal covering location problem. *Engineering Optimization*, pages 1–20, 2023.

[13] S. Atta, P. R. Sinha Mahapatra, and A. Mukhopadhyay. Solving maximal covering location problem using genetic algorithm with local refinement. *Soft Computing*, 22:3891–3906, 2018.

[14] P. Avella, A. Calamita, and L. Palagi. Speeding up the solution of the site and power assignment problem in wireless networks. *arXiv preprint arXiv:2210.04022*, 2022.

[15] P. Avella, A. Calamita, and L. Palagi. A compact formulation for the base station deployment problem in wireless networks. *Networks*, 82(1):52–67, 2023.

[16] P. Avella, A. Calamita, and L. Palagi. A computational study of off-the-shelf minlp solvers on a benchmark set of congested capacitated facility location problems. *arXiv preprint arXiv:2303.04216*, 2023.

[17] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Mathematical programming*, 58(1-3):295–324, 1993.

[18] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan. Mixed-integer nonlinear optimization. *Acta Numerica*, 22:1–131, 2013.

[19] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of operations research*, 23(4):769–805, 1998.

[20] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1–13, 1999.

[21] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical programming*, 88:411–424, 2000.

[22] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, 1962.

[23] BEREC. BEREC Guidelines on Very High Capacity Networks, 2020. `https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/9439-berec-guidelines-on-very-high-capacity-networks` (Accessed June 24, 2021).

[24] BEREC. BEREC Guidelines to assist NRAs on the consistent application of Geographical surveys of network deployments, 2020. `https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/9027-berec-guidelines` (Accessed June 24, 2021).

[25] O. Berman, I. Hajizadeh, and D. Krass. The maximum covering problem with travel time uncertainty. *IIE Transactions*, 45(1):81–96, 2013.

[26] O. Berman, D. Krass, and Z. Drezner. The gradual covering decay location problem on a network. *European Journal of Operational Research*, 151(3):474–480, 2003.

[27] D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical programming*, 98(1):49–71, 2003.

[28] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[29] K. Bestuzheva, A. Gleixner, and S. Vigerske. A computational study of perspective cuts. *Mathematical Programming Computation*, pages 1–29, 2023.

[30] N. Bilal, P. Galinier, and F. Guibault. An iterated-tabu-search heuristic for a variant of the partial set covering problem. *Journal of Heuristics*, 20(2):143–164, 2014.

[31] J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[32] B. Boffey, R. Galvao, and L. Espejo. A review of congestion models in the location of facilities with immobile servers. *European Journal of Operational Research*, 178(3):643–662, 2007.

[33] P. Bonami, J. Lee, S. Leyffer, and A. Wächter. More branch-and-bound experiments in convex nonlinear integer programming. *Preprint ANL/MCS-P1949-0911, Argonne National Laboratory, Mathematics and Computer Science Division*, 91, 2011.

[34] O. Bondarenko, D. Ageyev, and O. Mohammed. Optimization model for 5g network planning. In *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, pages 1–4. IEEE, 2019.

[35] A. Calamita, I. Ljubić, and L. Palagi. Benders decomposition for congested partial set covering location with uncertain demand. *arXiv preprint arXiv:2401.12625*, 2024.

[36] M. S. Canbolat and M. von Massow. Planar maximal covering with ellipses. *Computers & Industrial Engineering*, 57(1):201–208, 2009.

[37] A. Capone, L. Chen, S. Gualandi, and D. Yuan. A new computational approach for maximum link activation in wireless networks under the sinr model. *IEEE transactions on wireless communications*, 10(5):1368–1372, 2011.

[38] S. Ceria and J. Soares. Convex programming for disjunctive convex optimization. *Mathematical Programming*, 86:595–614, 1999.

[39] L. Chen, S.-J. Chen, W.-K. Chen, Y.-H. Dai, T. Quan, and J. Chen. Efficient presolving methods for solving maximal covering and partial set covering location problems. *European Journal of Operational Research*, 2023.

[40] L. Chiaraviglio, A. S. Cacciapuoti, G. Di Martino, M. Fiore, M. Montesano, D. Trucchi, and N. B. Melazzi. Planning 5G networks under EMF constraints: State of the art and vision. *IEEE Access*, 6:51021–51037, 2018.

[41] L. Chiaraviglio, C. Di Paolo, and N. B. Melazzi. 5G network planning under service and EMF constraints: Formulation and solutions. *IEEE Transactions on Mobile Computing*, 2021.

[42] T. R. L. Christensen and A. Klose. A fast exact method for the capacitated facility location problem with differentiable convex production costs. *European Journal of Operational Research*, 292(3):855–868, 2021.

[43] R. Church and C. ReVelle. The maximal covering location problem. In *Papers of the regional science association*, volume 32, pages 101–118. Springer-Verlag Berlin/Heidelberg, 1974.

[44] A. A. Coco, A. C. Santos, and T. F. Noronha. Formulation and algorithms for the robust maximal covering location problem. *Electronic Notes in Discrete Mathematics*, 64:145–154, 2018.

[45] J.-F. Cordeau, F. Furini, and I. Ljubić. Benders decomposition for very large scale partial set covering and maximal covering location problems. *European Journal of Operational Research*, 275(3):882–896, 2019.

[46] A. M. Costa. A survey on benders decomposition applied to fixed-charge network design problems. *Computers & operations research*, 32(6):1429–1450, 2005.

[47] R. J. Dakin. A tree-search algorithm for mixed integer programming problems. *The computer journal*, 8(3):250–255, 1965.

[48] F. D'andreagiovanni. Pure 0-1 programming approaches to wireless network design. *4OR*, 10(2):211, 2012.

[49] F. D'Andreagiovanni, C. Mannino, and A. Sassano. GUB covers and power-indexed formulations for wireless network design. *Management Science*, 59(1):142–156, 2013.

[50] G. B. Dantzig. Linear programming under uncertainty. *Management science*, 1(3-4):197–206, 1955.

[51] M. S. Daskin and S. H. Owen. Two new location covering problems: The partial p-center problem and the partial set covering problem. *Geographical Analysis*, 31(3):217–235, 1999.

[52] S. Davari, M. H. F. Zarandi, and A. Hemmati. Maximal covering location problem (mclp) with fuzzy travel times. *Expert Systems with Applications*, 38(12):14535–14541, 2011.

[53] S. Dehghan. A new approach. *3GSM Daily 2005*, 1:44, 2005.

[54] M. Desrochers, P. Marcotte, and M. Stan. The congested facility location problem. *Location Science*, 3(1):9–23, 1995.

[55] M. Dimopoulou and I. Giannikos. Spatial optimization of resources deployment for forest-fire management. *International Transactions in Operational Research*, 8(5):523–534, 2001.

[56] A. Dmytro, A.-A. Ali, and Q. Nameer. Multi-period lte ran and services planning for operator profit maximization. In *The Experience of Designing and Application of CAD Systems in Microelectronics*, pages 25–27. IEEE, 2015.

[57] B. T. Downs and J. D. Camm. An exact algorithm for the maximal covering problem. *Naval Research Logistics (NRL)*, 43(3):435–461, 1996.

[58] Z. Drezner and H. W. Hamacher. *Facility location: applications and theory*. Springer Science & Business Media, 2004.

[59] F. D'andreagiovanni. On improving the capacity of solving large-scale wireless network design problems by genetic algorithms. In *European Conference on the Applications of Evolutionary Computation*, pages 11–20. Springer, 2011.

[60] F. D'andreagiovanni and A. M. Gleixner. Towards an accurate solution of wireless network design problems. In *International Symposium on Combinatorial Optimization*, pages 135–147. Springer, 2016.

[61] F. D'andreagiovanni, H. Lakhlef, and A. Nardin. A matheuristic for joint optimal power and scheduling assignment in dvb-t2 networks. *Algorithms*, 13(1):27, 2020.

[62] F. D'andreagiovanni, C. Mannino, and A. Sassano. Negative cycle separation in wireless network design. In *International Conference on Network Optimization*, pages 51–56. Springer, 2011.

[63] A. Eisenblatter and H.-F. Geerdes. Capacity optimization for umts: Bounds and benchmarks for interference reduction. In *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6. IEEE, 2008.

[64] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on matrix analysis and applications*, 18(4):1035–1064, 1997.

[65] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.

[66] S. Elhedhli. Service system design with immobile servers, stochastic demand, and congestion. *Manufacturing & Service Operations Management*, 8(1):92–97, 2006.

[67] R. Z. Farahani, N. Asgari, N. Heidari, M. Hosseininia, and M. Goh. Covering problems in facility location: A review. *Computers & Industrial Engineering*, 62(1):368–407, 2012.

[68] M. Fischetti, I. Ljubić, and M. Sinnl. Benders decomposition without separability: A computational study for capacitated facility location problems. *European Journal of Operational Research*, 253(3):557–569, 2016.

[69] M. Fischetti, I. Ljubić, and M. Sinnl. Redesigning Benders decomposition for large-scale facility location. *Management Science*, 63(7):2146–2162, 2017.

[70] M. Fischetti, D. Salvagnin, and A. Zanette. A note on the selection of Benders' cuts. *Mathematical Programming*, 124:175–182, 2010.

[71] B. Fortz. Location problems in telecommunications. *Location science*, pages 537–554, 2015.

[72] A. Frangioni and C. Gentile. Perspective cuts for a class of convex 0–1 mixed integer programs. *Mathematical Programming*, 106:225–236, 2006.

[73] A. Frangioni and C. Gentile. Sdp diagonalizations and perspective cuts for a class of nonseparable miqp. *Operations Research Letters*, 35(2):181–185, 2007.

[74] A. Frangioni and C. Gentile. A computational comparison of reformulations of the perspective relaxation: SOCP vs. cutting planes. *Operations Research Letters*, 37(3):206–210, 2009.

[75] A. Frangioni, C. Gentile, and F. Lacalandra. Tighter approximated milp formulations for unit commitment problems. *IEEE Transactions on Power Systems*, 24(1):105–113, 2008.

[76] FUB. Fondazione Ugo Bordoni, 2021. `https://www.fub.it/en/` (Accessed October 4, 2021).

[77] J. D. Gadze, K. A. Agyekum, S. J. Nuagah, and E. Affum. Improved propagation models for lte path loss prediction in urban & suburban ghana. *arXiv preprint arXiv:2001.05227*, 2020.

[78] R. D. Galvão, L. G. A. Espejo, and B. Boffey. A comparison of lagrangean and surrogate relaxations for the maximal covering location problem. *European Journal of Operational Research*, 124(2):377–389, 2000.

[79] R. D. Galvão and C. ReVelle. A lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, 88(1):114–123, 1996.

[80] H. Ganame, L. Yingzhuang, H. Ghazzai, and D. Kamissoko. 5G base station deployment perspectives in millimeter wave frequencies using meta-heuristic algorithms. *Electronics*, 8(11):1318, 2019.

[81] S. García and A. Marín. Covering location problems. *Location science*, pages 93–114, 2015.

[82] A. M. Geoffrion. Generalized benders decomposition. *Journal of optimization theory and applications*, 10:237–260, 1972.

[83] H. Ghazzai, E. Yaacoub, M.-S. Alouini, Z. Dawy, and A. Abu-Dayya. Optimized LTE cell planning with varying spatial and temporal user densities. *IEEE Transactions on vehicular technology*, 65(3):1575–1589, 2015.

[84] O. Günlük and J. Linderoth. Perspective relaxation of mixed integer nonlinear programs with indicator variables. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 1–16. Springer, 2008.

[85] O. Günlük and J. Linderoth. Perspective reformulations of mixed integer nonlinear programs with indicator variables. *Mathematical Programming*, 124(1):183–205, 2010.

[86] O. Günlük and J. Linderoth. Perspective reformulation and applications. In *Mixed Integer Nonlinear Programming*, pages 61–89. Springer, 2012.

[87] O. K. Gupta and A. Ravindran. Branch and bound experiments in convex nonlinear integer programming. *Management science*, 31(12):1533–1546, 1985.

[88] Gurobi. Gurobi Optimizer, 2021. `https://www.gurobi.com/products/gurobi-optimizer/` (Accessed October 08, 2021).

[89] Gurobi. 5 Key Areas Where Mathematical Optimization Can Deliver Benefits for Telecom Companies, 2023. `https://www.gurobi.com/resources/5-key-areas-where-mathematical-optimization-can-deliver-benefits-for\-telecom-companies-in-2021//` (Accessed September 29, 2023).

[90] Gurobi. Telecommunications: Explore the Possibilities, 2023. `https://www.gurobi.com/industry/optimization-for-the-telecommunications-industry/` (Accessed September 29, 2023).

[91] S. L. Hakimi. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, 13(3):462–475, 1965.

[92] W. K. K. Haneveld and M. H. van der Vlerk. Stochastic integer programming: General models and algorithms. *Annals of operations research*, 85:39, 1999.

[93] J. Harkness and C. ReVelle. Facility location with increasing production costs. *European Journal of Operational Research*, 145(1):1–13, 2003.

[94] Infratel Italia. Mappatura 2021 reti fisse a banda ultralarga, 2021. `https://www.infratelitalia.it/archivio-documenti/documenti/mappatura-2021-reti-fisse-a-banda-ultralarga-consultazione-degli-operatori` (Accessed June 24, 2021).

[95] A. Israr, Q. Yang, W. Li, and A. Y. Zomaya. Renewable energy powered sustainable 5g network infrastructure: Opportunities, challenges and perspectives. *Journal of Network and Computer Applications*, 175:102910, 2021.

[96] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on scientific Computing*, 20(1):359–392, 1998.

[97] J. Kennington, E. Olinick, and D. Rajan. *Wireless network design: Optimization models and solution procedures*, volume 158. Springer Science & Business Media, 2010.

[98] J. Krarup, M. Labbé, and G. Rand. Optimal location: minimum versus equilibrium allocation. In *Operational research'87*, pages 718–729. North-Holland, 1987.

[99] A. H. Land and A. G. Doig. An automatic method of solving discrete programming problems. *Econometrica*, 28(3):497–520, 1960.

[100] G. Laporte, S. Nickel, and F. Saldanha-da Gama. *Introduction to location science*. Springer, 2019.

[101] Laurent Lessard. Lecture 13 on Cones and semidefinite constraints of the course in Introduction to Optimization, 2017. `https://laurentlessard.com/teaching/524-intro-to-optimization/` (Accessed October 16, 2023).

[102] B. H. Lee and R. A. Deininger. Optimal locations of monitoring stations in water distribution system. *Journal of Environmental Engineering*, 118(1):4–16, 1992.

[103] J. Lee and S. Leyffer. *Mixed integer nonlinear programming*, volume 154. Springer Science & Business Media, 2011.

[104] T. L. Lei, D. Tong, and R. L. Church. Designing robust coverage systems: a maximal covering model with geographically varying failure probabilities. *Annals of the Association of American Geographers*, 104(5):922–938, 2014.

[105] I. Ljubić, P. Putz, and J.-J. Salazar-González. Exact approaches to the single-source network loading problem. *Networks*, 59(1):89–106, 2012.

[106] L. A. Lorena and M. A. Pereira. A lagrangean/surrogate heuristic for the maximal covering location problem using hillman's edition. *International Journal of Industrial Engineering*, 9:57–67, 2002.

[107] D. Lu. Facility location with economies of scale and congestion. Master's thesis, University of Waterloo, 2010.

[108] T. L. Magnanti and R. T. Wong. Accelerating Benders decomposition: Algorithmic enhancement and model selection criteria. *Operations Research*, 29(3):464–484, 1981.

[109] C. Mannino, S. Mattia, and A. Sassano. Planning wireless networks by shortest path. *Computational Optimization and Applications*, 48(3):533–551, 2011.

[110] C. Mannino, F. Rossi, and S. Smriglio. The network packing problem in terrestrial broadcasting. *Operations Research*, 54(4):611–626, 2006.

[111] V. R. Máximo, J.-F. Cordeau, and M. C. Nascimento. A hybrid adaptive iterated local search heuristic for the maximal covering location problem. *International Transactions in Operational Research*, 2023.

[112] V. R. Máximo and M. C. Nascimento. Intensification, learning and diversification in a hybrid metaheuristic: an efficient unification. *Journal of Heuristics*, 25:539–564, 2019.

[113] V. R. Máximo, M. C. Nascimento, and A. C. Carvalho. Intelligent-guided adaptive search for the maximum covering location problem. *Computers & Operations Research*, 78:129–137, 2017.

[114] N. Megiddo, E. Zemel, and S. L. Hakimi. The maximum coverage location problem. *SIAM Journal on Algebraic Discrete Methods*, 4(2):253–261, 1983.

[115] J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios. Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281, 1995.

[116] A. T. Murray. Maximal coverage location problem: impacts, significance, and evolution. *International Regional Science Review*, 39(1):5–27, 2016.

[117] J. Naoum-Sawaya and S. Elhedhli. A nested benders decomposition approach for telecommunication network planning. *Naval Research Logistics (NRL)*, 57(6):519–539, 2010.

[118] N. Papadakos. Practical enhancements to the Magnanti–Wong method. *Operations Research Letters*, 36(4):444–449, 2008.

[119] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[120] R. Rahmaniani, T. G. Crainic, M. Gendreau, and W. Rei. The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3):801–817, 2017.

[121] M. G. Resende. Computing approximate solutions of the maximum covering problem with grasp. *Journal of Heuristics*, 4(2):161–177, 1998.

[122] M. G. Resende and P. M. Pardalos. *Handbook of optimization in telecommunications*. Springer Science & Business Media, 2008.

[123] C. ReVelle. Facility siting and integer-friendly programming. *European Journal of Operational Research*, 65(2):147–158, 1993.

[124] C. ReVelle, M. Scholssberg, and J. Williams. Solving the maximal covering location problem with heuristic concentration. *Computers & Operations Research*, 35(2):427–435, 2008.

[125] C. S. ReVelle and H. A. Eiselt. Location analysis: A synthesis and survey. *European journal of operational research*, 165(1):1–19, 2005.

[126] C. S. Revelle, H. A. Eiselt, and M. S. Daskin. A bibliography for some fundamental problem categories in discrete location science. *European journal of operational research*, 184(3):817–848, 2008.

[127] M. J. Rosenblatt and H. L. Lee. A robustness approach to facilities design. *International journal of production research*, 25(4):479–486, 1987.

[128] M. Rumney et al. *LTE and the evolution to 4G wireless: Design and measurement challenges*. John Wiley & Sons, 2013.

[129] S. Şelfun. Outer approximation algorithms for the congested p-median problem. Master's thesis, Bilkent Universitesi (Turkey), 2011.

[130] E. L. F. Senne, M. A. Pereira, L. A. N. Lorena, et al. A decomposition heuristic for the maximal covering location problem. *Advances in Operations Research*, 2010, 2010.

[131] S. R. Shariff, N. H. Moin, and M. Omar. Location allocation modeling for healthcare facility planning in malaysia. *Computers & Industrial Engineering*, 62(4):1000–1010, 2012.

[132] L. V. Snyder. Covering problems. *Foundations of location analysis*, pages 109–135, 2011.

[133] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 1973.

[134] R. A. Stubbs and S. Mehrotra. A branch-and-cut method for 0-1 mixed convex programming. *Mathematical programming*, 86:515–532, 1999.

[135] Telco Antennas. 4G LTE Signal Strength Reference Guide, 2021. `https://www.telcoantennas.com.au/blog/guide-to-mobile-networks/4g-lte-signal-strength-reference-guide/` (Accessed May 18, 2021).

[136] C. Toregas, R. Swain, C. ReVelle, and L. Bergman. The location of emergency service facilities. *Operations Research*, 19(6):1363–1373, 1971.

[137] Usat. Understanding LTE Signal Strength, 2021. `https://usatcorp.com/faqs/understanding-lte-signal-strength-values/` (Accessed May 18, 2021).

[138] V. Verter. Uncapacitated and capacitated facility location problems. *Foundations of location analysis*, pages 25–37, 2011.

[139] W. Walker. Using the set-covering problem to assign fire companies to fire houses. *Operations Research*, 22(2):275–277, 1974.

[140] L. A. Wolsey. Valid inequalities for 0–1 knapsacks and mips with generalised upper bound constraints. *Discrete Applied Mathematics*, 29(2-3):251–261, 1990.

[141] Y. Zhang, O. Berman, and V. Verter. Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, 198(3):922–935, 2009.

Acknowledgements

I would like to express my gratitude to those who made this thesis possible through their support and contributions.

I sincerely thank my supervisor Laura Palagi, for your invaluable and constant guidance throughout these three years. Your mentorship, energy and drive have been inspirational to me. A huge thank you for all the opportunities and collaborations that you enabled me to undertake.

I wish to thank Pasquale Avella. Your expertise, in-depth comments, and constructive criticisms made a significant contribution to the first part of this thesis on wireless network design. Your teachings on the general practices of operations research have left a mark on me.

A special thank goes to Ivana Ljubić, who has been a constant source of encouragement, motivation and support. Your constructive inputs and experience have been vital to the development of the second part of this thesis on partial set covering location. I will never forget your genuine enthusiasm for life and research.

I extend acknowledgements to the Fondazione Ugo Bordoni for their funding, which enabled the study on wireless network design. Their competencies have been instrumental in the success of this project.

Sincerely,

Alice

Rome, October 2023