HEC MONTRÉAL
École affiliée à l'Université de Montréal

Linearized Robust Counterparts with Applications in Location and Inventory Management Problems

par
Amir Ardestani-Jaafari

Thèse présentée en vue de l'obtention du grade de Ph. D. en administration
(option Sciences de la décision)

Avril 2016

HEC MONTRÉAL

École affiliée à l'Université de Montréal

Cette thèse intitulée :

Linearized Robust Counterparts with Applications in Location and Inventory Management Problems

Présentée par :

Amir Ardestani-Jaafari

a été évaluée par un jury composé des personnes suivantes :

Michel Denault
HEC Montréal
Président-rapporteur

Erick Delage
HEC Montréal
Directeur de recherche

Michèle Breton
HEC Montréal
Membre du jury

Timothy C. Y. Chan
University of Toronto
Examinateur externe

Jacques Robert
HEC Montréal
Représentant du directeur de HEC Montréal

Résumé

L'optimisation robuste a beaucoup attiré l'attention de la communauté scientifique dans la dernière décennie. Cette approche est maintenant considérée comme une approche standard pour aborder l'incertitude dans les problèmes de prise de décision. Plusieurs études ont démontré la valeur ajoutée de cette approche dans des problèmes industriels pour lesquels il est difficile de représenter l'incertitude à l'aide d'une distribution. Le succès de l'approche est principalement dû à la simplicité du processus de modélisation ainsi qu'à l'efficacité des méthodes de résolution disponibles. Malheureusement, il existe de nombreuses situations, en particulier en ce qui concerne la gestion des stocks et les problèmes de localisation d'installations, pour lesquelles la deuxième propriété est perdue. Cette thèse étudie la traitabilité du problème d'optimisation robuste dans ce type d'applications.

Dans le premier chapitre, nous nous concentrons sur des problèmes d'optimisation robuste dans lesquels la fonction de coût qui doit être « robustifiée » n'est pas linéaire (ou même concave) par rapport aux paramètres incertains. Ces fonctions se présentent, par exemple, dans des problèmes bien connus de marchands de journaux et d'inventaire. Étant donné que ces problèmes sont reconnus pour être insolubles en temps polynomial, nous proposons une nouvelle technique de construction d'une approximation dite conservatrice basée sur la relaxation d'un programme linéaire en nombres entiers quivalent. De plus, nous relions cette technique aux méthodes d'approximation qui sont basées sur l'exploitation de règles de décision affines. Notre nouvelle technique nous permet de proposer deux modèles d'approximation prenant respectivement la forme d'un programme linéaire et d'un problème d'optimisation semi-définie positive. Cette deuxième formulation a le potentiel de fournir des solutions de meilleure qualité au prix de calculs plus laborieux. Finalement, nous identifions des conditions sous lesquelles nos deux modèles d'approximation fournissent des solutions qui sont exactes. En particulier, nous sommes en mesure de proposer les premières reformulations exactes pour la version robuste du problème de vendeur de journaux à plusieurs items pour l'ensemble d'incertitude budgété et du problème de gestion d'inventaire pour une région d'incertitude ayant la forme d'une sphère L_1 ou d'une boite.

Dans le deuxième chapitre, nous étudions la version robuste d'un problème multi période de localisation et de transport à coût fixe avec capacités limitées. Dans ce problème, la localisation et la capacité de chaque installation doivent être déterminés immédiatement, alors que la décision de la quantité à produire et de la distribution peut être retardée jusqu'au moment où les commandes sont reçues (i.e. au début de chaque période). Malheureusement, il est bien connu que ce type de problèmes de décisions à plusieurs périodes sont insolubles en temps polynomial. Pour surmonter cette difficulté, nous proposons un ensemble de modèles d'approximation conservatrice. Chaque modèle exploite à un différent niveau le concept de la flexibilité des décisions futures et atteint un certain compromis entre efficacité de résolution et qualité de la solution obtenue. Un algorithme de génération de lignes est également présenté afin de permettre de résoudre des problèmes de grande taille. Finalement, nous démontrons empiriquement qu'une flexibilité entière n'est souvent pas nécessaire pour obtenir des emplacements et capacités d'installation de très grande qualité ou même optimale.

Dans le troisième chapitre, nous étudions un problème d'optimisation robuste à deux étapes. Compte tenu de la difficulté de résolution de ce type de problème, une approximation conservatrice est proposée. Celle-ci est inspirée par les méthodes de linéarisation employées pour les problèmes d'optimisation bilinéaire. Nous établissons que notre approche peut être considérée équivalente aux approches basées sur l'usage de règles de décision affines et expliquons comment améliorer l'efficacité de ces méthodes. Finalement, nous illustrons comment employer notre méthode d'approximation dans trois applications.

Mots clés : Optimisation robuste, approximation conservatrice traitable, optimisation en deux étapes, inventaire, localisation.

Abstract

Robust optimization has attracted a large amount of attention regarding how to address uncertainty in decision making problems, especially in a vast number of industrial problems where probability distributions are hard to identify. This is mainly due to the simplicity of the modeling process and to the ease of resolution, even for large scale models. Unfortunately, there are many cases, especially with respect to inventory management and facility location problems, wherein the second property is lost. This thesis studies the tractability of robust optimization problem with application to inventory and facility location problems.

In the first chapter, we study robust optimization of problems wherein the cost function that needs to be robustified is not concave (or linear) with respect to the uncertain parameters. Such functions arise, for instance, in the famous newsvendor and inventory problems. Given that these problems are known to be intractable, we propose a new scheme for constructing conservative approximations based on the relaxation of an embedded mixed-integer linear program and relate this scheme to methods that are based on exploiting affine decision rules. Our new scheme gives rise to two tractable models that respectively take the shape of a linear program and a semi-definite program, with the latter having the potential to provide solutions of better quality than the former at the price of heavier computations. We present conditions under which our approximation models are exact. In particular, we are able to propose the first exact reformulations for a robust (and distributionally robust) multi-item newsvendor problem with budgeted uncertainty set and a reformulation for robust multi-period inventory problems that is exact when the uncertainty region reduces to a L_1-norm ball or to a box.

In the second chapter, we study a multi-period robust capacitated fixed-charge location-transportation problem in which, while the location and capacity of each facility need to be determined immediately, the determination of final production and distribution of products can be delayed until actual orders are received in each period. Unfortunately, it is well known that these types of multi-period robust decision problems are computationally intractable. To overcome this difficulty, we propose a set of tractable conservative approximations to the problem that each exploits to a different extent the idea of reducing the flexibility of the delayed decisions. While all of these approximation models outperform previous approximation models that have been proposed for this problem, each of them also has the potential to reach a different level of compromise between efficiency of resolution and quality of the solution. A row generation algorithm is also presented in order to address problem instances of realistic size. We also demonstrate that full flexibility is often unnecessary to reach nearly, or even exact, optimal robust locations and capacities for the facilities.

In the third chapter, we study two-stage robust optimization problem with right-hand side uncertainty. Due to the computational difficulty of this problem, a tractable conservative approximation is developed based on linear programming relaxation. We show some insights about this approximation, including its relation with affinely adjustable robust counterpart, and different methods of improvement of its bound. We further employ our proposed approximation in various applications.

Keywords : Robust optimization, tractable conservative approximation, two-stage optimization, inventory, facility location.

Preface

This thesis includes three articles listed as follows.

1. Ardestani-Jaafari, Amir, and Erick Delage. (2016). Robust optimization of sums of piecewise linear functions with application to inventory problems. Operations Research 64(2): 474-494.

   ** Awarded the Esdras-Minville best student paper award in 2015, HEC Montreal.*

2. Ardestani-Jaafari, Amir, and Erick Delage. (2016). The value of flexibility in robust location-transportation problems. accepted in Transportation Science.

   ** Honorable mention for the best student paper award, the 58th Canadian Operational Research Society (CORS) annual conference, June 2016.*

3. Ardestani-Jaafari, Amir, and Erick Delage. Linearized robust counterparts of two-stage robust optimization problem with applications in operations management, To be submitted.

Table of Contents

List of Tables

List of Figures

To the loving memory of my mother, to my father, and to my dear wife for her unconditional love and support.

Acknowledgements

First, I would like to thank Dr. Erick Delage, my supervisor, for his helpful guidance, support, critical sense and encouragement from beginning to successful completion of my PhD studies. I would like to admit that this PhD would not have been possible without his support.

Second, I would like to thank the rest of my thesis committee: Dr. Timothy C. Y. Chan, Dr. Michel Denault, Dr. Michèle Breton, and also, my proposal committee member, Dr. Vedat Verter. In particular, my sincere thanks goes to Dr. Michèle Breton for her insightful comments and encouragement in all steps of my PhD.

I am grateful to my great friends, Hossein Mirzapour and Marie-Hélène Roy who made my stay at HEC Montréal an enjoyable period of my life. I would also like to thank Yaovi Gassesse (Axel) Siliadin and Mohamed Ali Akari for being warm-hearted colleagues and friends. I also need to thank Dr. Dan Andrei Iancu and Frans de Ruiter for valuable discussions on topics related to this thesis.

Last but not least, I would like to thank my wife, Samaneh, for her love, support, and encouragement.

Introduction

Robust Optimization (RO) is a methodology, combined with computational tools, that allows one to process optimization problems in which some data is uncertain and only known to belong to some uncertainty set (Ben-Tal and Nemirovski 2002). To the best of our knowledge, Soyster (1973) was the first to discuss the use of RO in the context of linear programming models. There were only a few more appearances of this methodology in the two subsequent decades. The main reasons for such impopularity might be attributed to the conception that robust optimization was either too conservative or too computational difficult.

In the 1990s, we have witnessed a rebirth of RO wherein some important works were published in the context of both integer programming (Kouvelis and Yu 1997) and convex programming (Ben-Tal and Nemirovski 1997, 1998, El Ghaoui and Lebret 1997, El Ghaoui et al. 1998). It has been argued in a survey by Bertsimas et al. (2011a) that the work of Ben-Tal and Nemirovski (*e.g.,* Ben-Tal and Nemirovski (1998, 1999, 2000)) and El Ghaoui and Lebret (1997), El Ghaoui et al. (1998) in the late 1990s, coupled with advances in computing technology and the development of fast interior point methods for convex optimization, particularly for semidefinite optimization (*e.g.,* Vandenberghe and Boyd (1996)), has inspired a massive interest in the field of RO. As a consequence, in the last ten years, RO has attracted a large amount of attention regarding how to address uncertainty in decision making problems, especially in a vast number of industrial problems where probability distributions are hard to identify for uncertain elements of the decision problem. Overall, this interest can mainly be attributed to the simplicity of the modeling paradigm and to the ease of resolution of many robust optimization models, even for large scale problems.

Unfortunately, there are still many situations wherein the "ease of resolution" property is lost and we will describe two important cases. The first difficulty arises when the profit (or cost) function that needs to be "robustified" is not convex (or respectively not concave) with respect to the perturbing parameters. Newsvendor problem can be mentioned as an application where this difficulty arises. Unfortunately, robust counterpart (RC) models for these problems are computationally intractable (See Theorem 1.9.4) even when a polyhedral uncertainty set is considered. Furthermore, the classical robust optimization paradigm considers all decisions to be "here-and-now" decisions, *i.e.,* decisions that must be implemented before the uncertain elements will be observed. However this assumption is not always a realistic one to make. To address the uncertainty in problems where some decisions can be delayed until some uncertain elements are observed, the pioneering work of Ben-Tal et al. (2004) proposed an adjustable robust optimization problem. Unfortunately, intractability quickly arises in such adjustable robust optimization problems. In fact, the question of how to address the tractability issues associated to adjustable RO has been an active direction of research since then and is highlighted as an important direction of future research in Bertsimas et al. (2011a).

The popular approach to address the intractability issues associated to the two aforementioned situations is to replace the troublesome non-linear function or adjustable decision variable with an affine function of the perturbing parameters (only the observable ones in the later case). This approach leads to a model known as the affinely adjustable robust counterpart (AARC) of the robust optimization problem and was introduced in Ben-Tal et al. (2004). Although AARC is a tractable approximation and can often address the computational difficulty associated to the two situations, to this day there are still many open questions regarding the efficiency of AARC models. In our opinion, the most important ones that need to be answered are as follows:

- What are the conditions under which an AARC model provides an exact solution?
- How good is the quality of the solutions obtained when the AARC model is an approximation?
- What are efficient procedures that can be used to improve the quality of the solution returned by an AARC model?

While these are questions that require serious consideration, there is only a limited amount of work in the literature that attempts to address them. Consequently, these three questions constitute the main motivation of this thesis with which we aim to pave a way to some answers. Overall, the main ideas of this thesis can be summarized as follows. Let's consider the following robust optimization problem

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \quad h(\mathbf{x},\boldsymbol{\zeta})\,, \tag{1}$$

where $\mathbf{x}$ and $\boldsymbol{\zeta}$ are respectively the decision variable vector in feasible region $\mathcal{X}$ and the perturbation parameter vector defined in uncertainty set $\mathcal{U}$. The function $h(\mathbf{x},\boldsymbol{\zeta})$ can be considered as a profit function that is concave with respect to ζ or a function capturing the sum of future profits in an adjustable model. Our idea relies on proposing a linearization scheme that will correct for the concavity of the objective function involved in the inner problem

$$\min_{\boldsymbol{\zeta}\in\mathcal{U}} \quad h(\mathbf{x},\boldsymbol{\zeta})\,. \tag{2}$$

This linearization scheme will lead to a tractable conservative approximation (*i.e.,* it captures a lower bound on profit) that we will refer as the linearized robust counterpart (LRC) model.

In this thesis, we focus on two important family of robust optimization problems: problems that involve functions that decompose as sums of piecewise linear functions and adjustable two-stage problems with uncertainty in the right-hand side of the recourse problem. For these problems, we show that a basic application of LRC is equivalent to AARC. This results ends up connecting the quality of an AARC model to the tightness of the relaxation obtained from the linearization process applied to the inner problem (2). Consequently, we are able to shed some light on the three questions as follows. An AARC model is exact if the relaxation obtained by linearization is tight. Otherwise, a bound on the quality of optimal solutions obtained using the AARC model should be measured in terms of the largest relaxation gap that can be obtained through linearization. Finally, the quality of solutions obtained from an AARC model can be improved by employing techniques that effectively reduce the relaxation gap.

Next, we explain the main contributions of each chapter as follows.

In the first chapter, we study the robust optimization of a sum of piecewise linear functions. Such functions arise for instance in most inventory management problems, but also in some machine learning and multi-attribute optimization problems. We propose for the first time a LRC model that is based on the linearization and relaxation of a mixed-integer program (MIP) that is embedded in problems of this type. We also relate this new model to models obtained with AARC. In fact, we propose two forms of LRC models: the first one takes the form of a linear program (hence its name LP-RC) that is equivalent to AARC, while the second takes the form of a semi-definite program (SDP-RC) and is guaranteed to provide better solutions at the price of heavier computations. By exploiting the totally unimodularity property of the matrix describing the feasible set of LP-RC, we are able to establish new conditions under which LP-RC (and consequently AARC) provides an exact solution. This is for instance the case for some interesting robust (and distributionally robust) multi-item newsvendor problems.

In the second chapter, we study the application of robust optimization to a multi-period robust capacitated fixed-charge location-transportation problem which naturally decomposes into many stages of decisions: the choice of locations followed by a sequence of decisions about production and transportation of goods. Perhaps surprisingly, we were the first to propose tractable solution methods that account for the dynamics of this robust optimization problem. Our research leads to two interesting improvement for the robust formulation; we first employ the AARC framework to the problem and later improve AARC with a model called extended lifted AARC (ELAARC). The second model exploits for the first time the idea that the approximation is tighter when applying AARC on transportation problems where the constraints are allowed to be violated but at a certain well designed price. Finally, we develop a row-generation algorithm to solve AARC due to computational difficulty of AARC in large instances. This algorithm is developed so that it exploits the structure that emerges in the LRC formulation and allows us to solve problems with 15 facility locations, and 30 possible demand points and 20 periods in a very reasonable amount of time.

In the third chapter, we generalize the results of the first chapter and describe the LRC framework when appied to two-stage robust optimization problems. We show that both problems of previous chapters can be considered as special cases of the canonical form that is studied. By demonstrating the equivalence between LRC and AARC, we offer a valuable connection between AARC and the literature on approximation methods for bilinear optimization problems. This connection allows us to identify new ways of improving AARC that are inspired by the use of linear and conic valid inequalities. We show for instance that adding a special type of valid inequality in LRC can be interpreted as allowing for penalized violation of constraints in the AARC model (a technique that we exploited to obtain the ELAARC model in Chapter two). Finally, we show how one might apply LRC for two-stage robust adaptable problem with uncertainty set that are not polyhedral. We conclude by providing new insights about the newsvendor and facility location problems presented in Chapters 1 and 2 and describe some implications for a multi-product assembly problem.

Chapter 1
Robust Optimization of Sums of Piecewise Linear Functions with Application to Inventory Problems[1]

Amir Ardestani-Jaafari, Erick Delage

Department of Decision Sciences, HEC Montréal, Montréal, Quebec, H3T 2A7, Canada

amir.ardestani-jaafari@hec.ca, erick.delage@hec.ca

Abstract

Robust optimization is a methodology that has gained a lot of attention in the recent years. This is mainly due to the simplicity of the modeling process and ease of resolution even for large scale models. Unfortunately, the second property is usually lost when the cost function that needs to be "robustified" is not concave (or linear) with respect to the perturbing parameters. In this paper, we study robust optimization of sums of piecewise linear functions over polyhedral uncertainty set. Given that these problems are known to be intractable, we propose a new scheme for constructing conservative approximations based on the relaxation of an embedded mixed-integer linear program and relate this scheme to methods that are based on exploiting affine decision rules. Our new scheme gives rise to two tractable models that respectively take the shape of a linear program and a semi-definite program, with the latter having the potential to provide solutions of better quality than the former at the price of heavier computations. We present conditions under which our approximation models are exact. In particular, we are able to propose the first exact reformulations for a robust (and distributionally robust) multi-item newsvendor problem with budgeted uncertainty set and a reformulation for robust multi-period inventory problems that is exact whether the uncertainty region reduces to a L_1-norm ball or to a box. An extensive set of empirical results will illustrate

[1] This article is published online by "Operations Research".

the quality of the approximate solutions obtained using these two models on randomly generated instances of the latter problem.

Keywords

Robust optimization, piecewise linear, linear programming relaxation, semi-definite program, tractable approximations, newsvendor problem, inventory problem.

1.1 Introduction

Since the seminal work of Ben-Tal and Nemirovski (1998), robust optimization is a methodology that has attracted a large amount of attention. Such attention has stemmed in application fields that range from engineering problems like structural design (Ben-Tal and Nemirovski 1997) and circuit design (Boyd et al. 2005), management problems such as portfolio optimization (Goldfarb and Iyengar 2002) and supply chain management (Ben-Tal et al. 2005), to an array of data mining applications such as classification (Xu et al. 2009), regression (El Ghaoui and Lebret 1997) and parameter estimation (Calafiore and El Ghaoui 2001) (see Bertsimas et al. (2011a) for a detailed review of such applications). Two important factors that have contributed to this success are 1) the simplicity of the modeling paradigm, and 2) the tractability of many resulting formulations thus enabling the resolution of problems of scales that can match the practical needs. Unfortunately, the second property is usually lost when the cost function that needs to be "robustified" is not concave (or linear) with respect to the perturbing parameters.

This paper focuses on the following robust optimization problem:

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \;\; \max_{\zeta\in\mathcal{Z}} \;\; \sum_{i=1}^{N} h_i(\boldsymbol{x},\zeta) \;, \tag{1.1}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is a bounded polyhedral set of feasible solution for the decision $\boldsymbol{x}$, $\mathcal{Z} \subseteq \mathbb{R}^m$ is the set containing the possible perturbation ζ and for each i, the cost function $h_i(\boldsymbol{x},\zeta)$ is piecewise linear and convex in both $\boldsymbol{x}$ and ζ (although not necessarily jointly convex). In particular, this means that the cost function can be expressed as follows:

$$h_i(\boldsymbol{x},\zeta) := \max_k \boldsymbol{c}_x^{i,k}(\zeta)^T\boldsymbol{x} + d_x^{i,k}(\zeta) := \max_k \boldsymbol{c}_\zeta^{i,k}(\boldsymbol{x})^T\zeta + d_\zeta^{i,k}(\boldsymbol{x}) \;,$$

for some affine mappings $\boldsymbol{c}_x^{i,k} : \mathbb{R}^m \to \mathbb{R}^n$, $d_x^{i,k} : \mathbb{R}^m \to \mathbb{R}$, $\boldsymbol{c}_\zeta^{i,k} : \mathbb{R}^n \to \mathbb{R}^m$ and $d_\zeta^{i,k} : \mathbb{R}^n \to \mathbb{R}$. Although objective functions that take the form of sums of piecewise linear function abound in practice, exact solutions to the robust version of these problems are often considered impossible to obtain because of the computational difficulties that arise in solving the inner maximization problem (a.k.a. adversarial problem). Beside the two inventory problems that will be discussed later, such structured functions also play an important role in multi-objective optimization and machine learning (see Appendix 1.9.2 for details).

Recently, Gorissen and den Hertog (2013) have made a valuable effort at presenting a comprehensive overview of three families of solution methods that can be employed for this problem: namely, exact methods, tractable conservative approximations[2] , and cutting-plane methods. Unfortunately, while there are a few very special cases for which finding an exact solution is known to be tractable, still very little is known theoretically about the quality of conservative approximations that are available. In this paper, we attempt to reduce this gap by bringing the following contributions:

1. We propose a novel scheme for deriving tractable conservative approximations that connects for the first time the suboptimality of an approximate solution directly to the integrality gap of an associated mixed integer linear program (MILP). This allows us to identify fairly general conditions under which the concept of total unimodularity can be used to establish that the approximate solution obtained by solving a linear program of reasonable size is exactly optimal. The connection to MILP optimization also naturally allows us to propose a tighter conservative approximation model that takes the shape of a semi-definite program. This is perhaps surprising given that it is well known that, while schemes that are based on quadratic decision rules will lead to semi-definite program (SDP) approximation models when the uncertainty set is

[2] In other words, the approximate solution is guaranteed to achieve a worst-case cost that is bounded by the optimal value of the approximate optimization problem.

ellipsoidal, such adjustment functions lead in general to optimization problems that are computationally intractable, for instance when the uncertainty set is polyhedral (see Ben-Tal et al. (2009a) p. 372). Indeed, we show for the first time how to obtain SDP approximation models for such uncertainty sets by employing affine decision rules on a clever reformulation of the objective function.

2. We provide for the first time an exact tractable reformulation for a robust multi-item newsvendor problem with demand uncertainty that is non-rectangular, namely where it takes the shape of a budgeted uncertainty set with an integer budget. A novel tractable reformulation is also presented for the distributionally robust version of this problem in which the distribution information includes a budgeted uncertainty set for the support, the mean vector, and a list of lower bounds on first order partial moments. To the best of our knowledge, this appears to be first exact tractable reformulation for instances of multi-item newsvendor problem where there exists information about how the demand for different items behave jointly, a problem that was left open since the early work of Scarf (1958).

3. We propose a new conservative approximation model for a robust multi-period inventory problem where all orders must be made initially. We prove that this model produces an exact solution when facing a budgeted uncertainty set with a budget equal to one or to the total size of the horizon. Although exact reformulations exist for each of these extreme cases, this is the first model known to be exact for both cases simultaneously. Our empirical study also provides evidence that the suboptimality gap is relatively small with our new model (less than 0.3% gap on average with a maximum observed gap of 5%) when the budget takes on intermediate values. Finally, we present extensive empirical evidence that this model can be used to identify ordering strategies that make better trade-off between performance and robustness in comparison to strategies obtained using existing tractable method in the literature.

The paper is organized as follows. We start in Section 1.2 with a brief review of related work and currently available methods for solving problem (1.1). Section 1.3 presents our notation. In Section 1.4, we introduce our new approximation scheme for the robust optimization problem (1.1) that is based on the fractional relaxation of an associated mixed-integer linear program. In Section 1.5, we present implications of our results for a robust and distributionally robust multi-item newsvendor problem. In Section 1.6, we apply the new models to a robust multi-period inventory problem. Section 1.7 presents experiments on an inventory problem that attempt to evaluate the relative tightness of different approximation schemes and illustrate how one can employ these schemes to explore the trade-offs between expected performance and robustness in choosing an order policy. Finally, we conclude and provide some directions of future research in Section 1.8.

1.2 Background & Prior Work

Our work follows very closely the initiative of Gorissen and den Hertog (2013) who were interested in solving problems of the form (1.1) and where a comprehensive overview of available methods is presented. In Gorissen and den Hertog (2013), the robust optimization of the sums of maxima of linear functions takes the shape of

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad \max_{\zeta\in\mathcal{Z}} \ \ell(\zeta,\boldsymbol{x}) + \sum_{i=1}^{N} \max_k\{\ell_{i,k}(\zeta,\boldsymbol{x})\},$$

where ℓ and $\ell_{i,k}$ are bi-affine functions in the uncertain parameter ζ and the decision variable $\boldsymbol{x} \in \mathbb{R}^n$, and where $\mathcal{Z}$ is the uncertainty set. Given that $\mathcal{Z}$ is convex, the authors first describe an exact solution approach that is based on reducing the worst-case analysis to a search over the vertices of $\mathcal{Z}$ since the objective function is convex in ζ. This leads to the equivalent finite formulation

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} \quad & \max_{v\in\mathcal{V}} \ \ell(\zeta^v,\boldsymbol{x}) + \sum_{i=1}^{N} y_i^v \\ \text{subject to} \quad & y_i^v \geq \ell_{i,k}(\zeta^v,\boldsymbol{x}) \quad \forall i,\, \forall k,\, \forall\, v\in\mathcal{V}, \end{aligned}$$

where $\mathcal{V} = \{1, 2, \cdots, V\}$ with V the number of vertices of $\mathcal{Z}$, and $\{\zeta^v\}_{v=1}^V$ is the finite set of such vertices. The authors do warn their reader that computational complexity of this approach grows exponentially with respect to the number of constraints that define $\mathcal{Z}$.

Gorissen and den Hertog also propose using cutting-plane methods to solve these problems exactly (especially when enumerating the vertices becomes unthinkable). In fact, there is empirical evidence that seems to indicate that such methods are particularly effective in practice for solving two-stage robust optimization problems (see Zeng and Zhao (2013)). In each iteration of a cutting-plane method, there is a need to establish the worst-case ζ for some fixed $\boldsymbol{x}$ in order to produce a cutting-plane: *i.e.,*

$$\max_{\zeta \in \mathcal{Z}} \left\{ \ell(\zeta, \boldsymbol{x}) + \sum_{i=1}^{N} \max_k \{\ell_{i,k}(\zeta, \boldsymbol{x})\} \right\}.$$

While there exists some special cases where an efficient procedure might be identified (see Bienstock and Özbay (2008) for an example), the authors suggest that in general this problem can be solved by solving a MILP similar to

$$\begin{aligned}
\max_{\zeta \in \mathcal{Z}, \boldsymbol{y}, \boldsymbol{z}} \quad & \ell(\zeta, \boldsymbol{x}) + \sum_{i=1}^{N} y_i \\
\text{subject to} \quad & y_i \leq \ell_{i,k}(\zeta, \boldsymbol{x}) + M(1 - z_{i,k}) \quad \forall\, i, \forall\, k \\
& \sum_{k=1}^{K} z_{i,k} = 1 \ \ \forall i \\
& \zeta \in \mathcal{Z}\,, \quad \boldsymbol{z} \in \{0,1\}^{N \times K}\,.
\end{aligned}$$

Unfortunately, although software products that handle such models are well developed, solving this problem is generally NP-hard (see NP-hardness discussion in Section 1.4) thus making this approach prohibitive for large problems. In particular, the experiments we conduct in Section 1.4.4 identified instances of such mixed-integer linear programs reformulations that could not be solved in less than a day of computation already when $N = m = 64$. Finally, polynomial-time solvability of cutting-plane methods is not guaranteed except for the ellipsoid method which is rarely used in practice.

Gorissen and den Hertog finally explain how the theory of affinely adjustable robust counterpart (AARC) proposed by Ben-Tal et al. (2004) can be used to obtain a conservative approximation method. In this case, each convex term of the objective is replaced with an affine function that is adjusted optimally while ensuring that the objective function upper bounds the true objective. The resulting model takes the shape:

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{v},\boldsymbol{w}}{\text{minimize}} \quad & \max_{\zeta \in \mathcal{Z}} \ \ell(\zeta, \boldsymbol{x}) + \sum_{i=1}^{N} (v_i + \boldsymbol{w}_i^T \zeta) \\ \text{subject to} \quad & v_i + \boldsymbol{w}_i^T \zeta \geq \ell_{i,k}(\zeta, \boldsymbol{x}) \, , \ \forall\, i, \, \forall\, k, \, \forall\, \zeta \in \mathcal{Z}, \end{aligned}$$

for which one can easily formulate a finite dimensional linear programming reformulation using duality theory. It is mentioned that this approach can be improved by using a lifting of the uncertainty space (Chen and Zhang 2009) or by involving quadratic decision rules if the uncertainty set is ellipsoidal. Although the AARC approach is often tractable, very little is theoretically known about the suboptimality of the obtained approximate solution (we refer the reader to Iancu et al. (2013) for the most general results to date on this topic).

As mentioned in the introduction, many robust inventory problems can be considered a special case of problem (1.1). In Bertsimas and Thiele (2006), the authors seem to have been the first to propose an approximation method to solve such problems. Their approach relies on finding the worst-case cost of each period individually before summing the results over all periods. Effectively, they replace problem (1.1) with the following:

$$\underset{\boldsymbol{x} \in \mathcal{X}}{\text{minimize}} \ \sum_{i=1}^{N} \max_{\zeta \in \mathcal{Z}} \ h_i(\boldsymbol{x}, \zeta) \, .$$

Interestingly, when the uncertainty set takes the shape of the budgeted uncertainty set (see Bertsimas and Sim (2004)), they show that the optimal robust policy is equivalent to the optimal policy of the nominal problem under a specifically designed demand vector. In spite of having been used in many occasions (*e.g.,* José Alem and Morabito (2012), Wei et al. (2011)), as noted in Gorissen and den Hertog (2013), this conservative approximation does not impose any relation between the worst-case ζ used to evaluate the different periods.

It appears that Ben-Tal et al. (2004) were the first to address a robust inventory problem in which there is a possibility to make adjustments to future orders as information about demand becomes available. They provide conservative approximations of the problem by applying the concept of affine decision rules. In Ben-Tal et al. (2005), similar ideas are applied to a supply chain problem. Interestingly, the empirical experiments presented there seem to indicate that AARC can perform surprisingly well. A similar success was achieved in Ben-Tal et al. (2009b) as reported in their Section 3.2.

1.3 Notation

We briefly review some notation that is used in the remaining sections. First, let $\boldsymbol{e}_i$ be the i-th column of the identity matrix while $\mathbf{1}$ is the vector of all ones, both of their dimensions should be clear from context. Given two matrices of same sizes, $\boldsymbol{A} \bullet \boldsymbol{B}$ refers to the Frobenius inner product which returns $\sum_{i,j} A_{i,j} B_{i,j}$. We use $\boldsymbol{A}_{i,:}$ to refer to the i-th row of $\boldsymbol{A}$ while $\boldsymbol{A}_{:,j}$ would refer to the j-th column of $\boldsymbol{A}$. For the sake of clarity, given a vector $\boldsymbol{b}$ we might use $(\boldsymbol{b})_i$, instead of b_i, to refer to the i-th term of the vector $\boldsymbol{b}$.

1.4 Mixed-integer Linear Programming based Approximation

In this section, we seek to obtain a conservative approximation of problem (1.1) using linearization schemes that are used in the field of mixed-integer linear programming. In particular, it is well-known that the inner maximization problem:

$$\underset{\zeta \in \mathcal{Z}}{\text{maximize}} \quad \sum_{i=1}^{N} \max_k \boldsymbol{c}_{i,k}^T \zeta + d_{i,k} \ , \tag{1.2}$$

where $\mathcal{Z}$ is polyhedral and where we dropped the dependence of $\boldsymbol{c}_{i,k}$ and $d_{i,k}$ on $\boldsymbol{x}$ for clarity, is NP-hard (see Appendix 1.9.3 for a proof). Given that $\mathcal{Z}$ is polyhedral and bounded, we can assume without loss of generality [3] that it is represented as

$$\mathcal{Z} := \{\zeta \in [-1,1]^m \,|\, \boldsymbol{A}\zeta \leq \boldsymbol{b},\ \ \|\zeta\|_1 \leq \Gamma\}\ ,$$

for some $\boldsymbol{A} \in \mathbb{R}^{p\times m}$ and $\boldsymbol{b} \in \mathbb{R}^p_+$, and some $0 \leq \Gamma \leq m$, and with $\boldsymbol{0} \in \mathcal{Z}$ capturing the "nominal" (*i.e.,* most likely) scenario for ζ. Note that this representation reduces to the budgeted uncertainty set when $\boldsymbol{A} := \boldsymbol{0}$ and $\boldsymbol{b} := \boldsymbol{0}$ which is the most natural way of capturing that each ζ_i is a perturbation of similar magnitudes while one does not expect too many terms of ζ being perturbed simultaneously (see Bertsimas and Sim (2004) and its ubiquitous use in robust optimization applications). In the more general case, we expect this representation to be especially relevant in problems where one wishes to emphasize that the uncertainty region is roughly symmetrical around the nominal scenario $\zeta_0 := \boldsymbol{0}$. Hence, we are left with the following adversarial problem:

$$\begin{aligned}
\underset{\zeta\in\mathbb{R}^m}{\text{maximize}} \quad & \sum_{i=1}^{N} \max_k \boldsymbol{c}_{i,k}^T\zeta + d_{i,k} && \text{(1.3a)}\\
\text{subject to} \quad & \boldsymbol{A}\zeta \leq \boldsymbol{b} && \text{(1.3b)}\\
& \|\zeta\|_\infty \leq 1 && \text{(1.3c)}\\
& \|\zeta\|_1 \leq \Gamma\ . && \text{(1.3d)}
\end{aligned}$$

We will initially present two approximation models that will trade-off between computational requirements and quality of the solution. We will then relate these models to the important family of approximation schemes known as AARCs.

[3] Note that to obtain such a reformulation one might need to identify some $\bar{\zeta} \in \mathcal{Z}$ and $\hat{\zeta} \in \mathbb{R}^m$ such that $\mathcal{Z} \subseteq [\bar{\zeta}_1 - \hat{\zeta}_1, \bar{\zeta}_1 + \hat{\zeta}_1] \times \cdots \times [\bar{\zeta}_m - \hat{\zeta}_m, \bar{\zeta}_m + \hat{\zeta}_m]$ and reformulate the problem in terms of $\zeta'_j := (\zeta_j - \bar{\zeta}_j)/\hat{\zeta}_j$ for all $j = 1, \ldots, m$ with $\zeta' \in \mathcal{Z}' \subseteq [-1,1]^m$. This would lead to reformulating the objective function according to $\boldsymbol{c}_\zeta^{i,k\prime}(\boldsymbol{x}) := \text{diag}(\hat{\zeta})\boldsymbol{c}_\zeta^{i,k}(\boldsymbol{x})$ and $d_\zeta^{i,k\prime}(\boldsymbol{x}) := \boldsymbol{c}_\zeta^{i,k}(\boldsymbol{x})^T\bar{\zeta} + d_\zeta^{i,k}(\boldsymbol{x})$.

1.4.1 Linear Programming Approximation Model

Our first step is to convert this convex maximization problem to a mixed-integer quadratic program by replacing the objective function with

$$\max_{\{\boldsymbol{z}\in\{0,1\}^{N\times K}\,|\,\sum_{k=1}^{K} z_{i,k}=1\,,\forall i\}} \quad \sum_{i=1}^{N}\sum_{k=1}^{K} z_{i,k}(\boldsymbol{c}_{i,k}^T\zeta + d_{i,k})\ ,$$

where we introduced additional adversarial binary decision variables $z_{i,k}$. As is often done for mixed-integer quadratic programs, we will circumvent the difficulty of maximizing the terms that are quadratic in $\boldsymbol{z}$ and ζ by linearizing the objective function, yet only after replacing the perturbation variables by the sum of their positive and negative parts (*i.e.,* $\zeta := \zeta^+ - \zeta^-$). Specifically, this linearization is obtained by replacing instances of $z_{i,k}\cdot\zeta^+$ by $\boldsymbol{\Delta}_{i,k}^+$ and $z_{i,k}\cdot\zeta^-$ by $\boldsymbol{\Delta}_{i,k}^-$. In steps, the objective becomes

$$\begin{aligned}\sum_{i=1}^{N}\sum_{k=1}^{K} z_{i,k}(\boldsymbol{c}_{i,k}^T(\zeta^+ - \zeta^-) + d_{i,k}) &= \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}_{i,k}^T\zeta^+ z_{i,k} - \boldsymbol{c}_{i,k}^T\zeta^- z_{i,k} + d_{i,k}z_{i,k}\\ &= \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}_{i,k}^T\boldsymbol{\Delta}_{ik}^+ - \boldsymbol{c}_{i,k}^T\boldsymbol{\Delta}_{ik}^- + d_{i,k}z_{i,k}.\end{aligned}$$

As for the constraints, one can first make explicit the relation between $\boldsymbol{\Delta}^+$ and ζ^+ by imposing $\sum_{k=1}^{K}\boldsymbol{\Delta}_{i,k}^+ = \sum_{k=1}^{K} z_{i,k}\cdot\zeta^+ = \zeta^+$ and similarly the relation between $\boldsymbol{\Delta}^-$ and ζ^-. One can also add to the model what is implied by every linear constraint $\boldsymbol{a}^T\zeta \le b$ on the $\boldsymbol{\Delta}^+$ and $\boldsymbol{\Delta}^-$, in other words that $\boldsymbol{a}^T(\boldsymbol{\Delta}_{i,k}^+ - \boldsymbol{\Delta}_{i,k}^-) = \boldsymbol{a}^T\zeta\cdot z_{i,k} \le b z_{i,k}$. Overall, it is easy to show that the following MILP is equivalent to problem (1.3):[4]

$$\begin{aligned}\underset{\boldsymbol{z},\zeta^+,\zeta^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-}{\text{maximize}} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}_{i,k}^T(\boldsymbol{\Delta}_{i,k}^+ - \boldsymbol{\Delta}_{i,k}^-) + d_{i,k}z_{i,k} && \text{(1.4a)}\\ \text{subject to} \quad & \boldsymbol{A}(\zeta^+ - \zeta^-) \le \boldsymbol{b} && \text{(1.4b)}\end{aligned}$$

[4] Note that, when replacing $\zeta := \zeta^+ - \zeta^-$, we can replace $\|\zeta^+ + \zeta^-\|_1 \le \Gamma$ with $\|\zeta^+ + \zeta^-\|_1 = \Gamma$ because Γ is assumed smaller or equal to m. In particular, if a candidate solution does not use all the budget it is always possible to find an index for which $\zeta_i^+ + \zeta_i^- < 1$ and add the same amount to both positive and negative term without affecting $\zeta^+ - \zeta^-$.

$$\zeta^+ \geq 0 \;\;\&\;\; \zeta^- \geq 0 \;\;\&\;\; \zeta_j^+ + \zeta_j^- \leq 1 \,, \forall\, j \tag{1.4c}$$

$$\mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \tag{1.4d}$$

$$\sum_{k=1}^{K} z_{i,k} = 1 \,, \forall\, i \tag{1.4e}$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^+ = \zeta^+ \;\;\&\;\; \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^- = \zeta^- \;\;, \forall\, i \tag{1.4f}$$

$$\boldsymbol{A}(\boldsymbol{\Delta}_{i,k}^+ - \boldsymbol{\Delta}_{i,k}^-) \leq \boldsymbol{b} z_{i,k} \,, \forall\, i, \forall\, k \tag{1.4g}$$

$$\boldsymbol{\Delta}_{i,k}^+ \geq 0 \;\;\&\;\; \boldsymbol{\Delta}_{i,k}^- \geq 0 \;\;\&\;\; \boldsymbol{\Delta}_{i,k}^+ + \boldsymbol{\Delta}_{i,k}^- \leq z_{i,k} \,, \forall\, i, \forall\, k \tag{1.4h}$$

$$\sum_{j=1}^{m} (\boldsymbol{\Delta}_{i,k}^+)_j + (\boldsymbol{\Delta}_{i,k}^-)_j = \Gamma z_{i,k} \,, \forall\, i, \forall\, k \tag{1.4i}$$

$$z_{i,k} \in \{0, 1\} \,, \forall\, i, \forall\, k \,. \tag{1.4j}$$

Based on the observation that the fractional relaxation of problem (1.4) provides an upper bound for the mixed-integer version, we can already conclude that replacing the adversarial problem in (1.1) with this fractional relaxation will provide us with a conservative approximation for problem (1.1).

Proposition 1.4.1 *The optimization model*

$$\underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\Delta},\nu,\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{w},\boldsymbol{\psi},\boldsymbol{\theta}}{\text{minimize}} \quad \mathbf{1}^T\boldsymbol{\Delta} + \Gamma\nu + \boldsymbol{b}^T\boldsymbol{\rho} + \mathbf{1}^T\boldsymbol{\gamma} \tag{1.5a}$$

$$\text{subject to} \quad \nu \geq \sum_{i=1}^{N} \boldsymbol{\lambda}_i^+ - \boldsymbol{A}^T\boldsymbol{\rho} - \boldsymbol{\Delta} \tag{1.5b}$$

$$\nu \geq \sum_{i=1}^{N} \boldsymbol{\lambda}_i^- + \boldsymbol{A}^T\boldsymbol{\rho} - \boldsymbol{\Delta} \tag{1.5c}$$

$$\gamma_i \geq \boldsymbol{b}^T\boldsymbol{w}_{i,k} + \mathbf{1}^T\boldsymbol{\psi}_{i,k} + \Gamma\theta_{i,k} + d_{i,k}(\boldsymbol{x}) \;\; \forall i, \; \forall k \tag{1.5d}$$

$$\theta_{i,k} \geq -\boldsymbol{\lambda}_i^+ - \boldsymbol{A}^T\boldsymbol{w}_{i,k} - \boldsymbol{\psi}_{i,k} + \boldsymbol{c}_{i,k}(\boldsymbol{x}) \;\; \forall i, \; \forall k \tag{1.5e}$$

$$\theta_{i,k} \geq -\boldsymbol{\lambda}_i^- + \boldsymbol{A}^T\boldsymbol{w}_{i,k} - \boldsymbol{\psi}_{i,k} - \boldsymbol{c}_{i,k}(\boldsymbol{x}) \;\; \forall i, \; \forall k \tag{1.5f}$$

$$\boldsymbol{\rho} \geq 0 \;,\; \boldsymbol{\Delta} \geq 0 \;,\; \boldsymbol{w}_{i,k} \geq 0 \;,\; \boldsymbol{\psi}_{i,k} \geq 0 \;, \forall i, \; \forall k \;, \tag{1.5g}$$

where $\boldsymbol{\rho} \in \mathbb{R}^p$, $\boldsymbol{\Delta} \in \mathbb{R}^m$, $\nu \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^N$, $\boldsymbol{\lambda}_i^+ \in \mathbb{R}^m$, $\boldsymbol{\lambda}_i^- \in \mathbb{R}^m$, $\boldsymbol{w}_{i,k} \in \mathbb{R}^p$, $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^m$, *and* $\theta_{i,k} \in \mathbb{R}$, *is a conservative approximation of problem* (1.1). *Specifically, let* $\hat{\boldsymbol{x}}^*$ *and* $\hat{v}^*$

be respectively the optimal solution and optimal value of this problem, $\hat{v}^*$ is an optimized upper bound for the best achievable worst-case cost, as measured by problem (1.1)*, and $\hat{\boldsymbol{x}}^*$ is guaranteed to achieve a lower worst-case cost than $\hat{v}^*$.*

Proof This is simply obtained by constructing the dual of problem (1.4). Specifically, we identified the dual variables of constraints (1.4b) to (1.4i) respectively as $\boldsymbol{\rho}$, $\boldsymbol{\Delta}$, ν, $\boldsymbol{\gamma}$, $(\boldsymbol{\lambda}_i^+, \boldsymbol{\lambda}_i^-)$, $\boldsymbol{w}_{i,k}$, $\boldsymbol{\psi}_{i,k}$, and $\theta_{i,k}$. Since problem (1.4) is a linear program for which we can show that there is always a feasible solution, one can confirm that duality is strict. After combining the dual problem to the outer minimization in $\boldsymbol{x}$ we obtain problem (1.5). A feasible solution for problem (1.4) can be identified using $\zeta_0 := \mathbf{0}$. We first assign $\zeta_j^+ := \epsilon_j$ and $\zeta_j^+ := -\epsilon_j$ for some $\epsilon \in \mathbb{R}^m$ chosen so that constraints (1.4b)-(1.4d) are satisfied (see Footnote 4). Given any binary assignment for $z_{i,k}$ that satisfies constraint (1.4e), one can complete the solution by setting $(\boldsymbol{\Delta}_{i,k}^+)_j := \zeta_j^+ z_{i,k}$ and $(\boldsymbol{\Delta}_{i,k}^-)_j := \zeta_j^- z_{i,k}$. ∎

At this point, one should wonder how good this approximation scheme is and whether it can be compared to other schemes that have been proposed in the literature. While we will later establish valuable connections to existing approximation methods, we will first shed light on how the quality of our approximation is related to the notion of integrality gap of mixed-integer programs and whether we can bound it.

Definition The integrality gap for a class of mixed-integer programs, where the objective function is maximized and the optimal value is known to be positive, is the supremum of the ratio between the optimal value achieved by a fractional solution and the optimal value achieved by an integer one. Specifically, if we seek $\max_{\zeta \in \mathcal{U} \cap \mathcal{I}} f(\zeta)$ for instances described by $(\mathcal{U}, \mathcal{I}, f) \in \mathbb{F}$, where $\mathcal{U}$ is polyhedron, $\mathcal{I}$ imposes that a set of terms of ζ be integer valued, and $\mathbb{F}$ refers to a certain family of problems that is being studied, then

$$\text{integrality gap} = \sup_{(\mathcal{U}, \mathcal{I}, f) \in \mathbb{F}} f^*_{\mathcal{U}} / f^*_{\mathcal{U} \cap \mathcal{I}} ,$$

where $f^*_{\mathcal{Z}} := \max_{\zeta \in \mathcal{Z}} f(\zeta)$.

Proposition 1.4.2 *Given that for each i and for all $\boldsymbol{x} \in \mathcal{X}$, $h_i(\boldsymbol{x}, \cdot)$ is positive definite on $\mathcal{Z}$, let $\hat{\boldsymbol{x}}$ and $\hat{v}(\hat{\boldsymbol{x}})$ respectively be the optimal solution and optimal value obtained from problem* (1.5)*, then both the true worst-case value for $\hat{\boldsymbol{x}}$ and the value $\hat{v}(\hat{\boldsymbol{x}})$ are less than a factor of γ away from the optimal value of problem* (1.1)*, where γ is the integrality gap for problem* (1.4)*.*

Proof The integrality gap of problem (1.4) with positive optimal value is defined as a bound on the largest achievable ratio between the optimal value obtained by a fractional solution and the one obtained by an integer solution for this type of problem instances. Given that the integrality gap for problem (1.4) is γ, this indicates that the worst-case value achieved by any feasible $\boldsymbol{x}$ is bounded between $\hat{v}(\boldsymbol{x})$ and $\gamma\hat{v}(\boldsymbol{x})$. Hence,

$$\max_{\zeta \in \mathcal{Z}} \ \sum_{i=1}^{N} h_i(\hat{\boldsymbol{x}}, \zeta) \leq \hat{v}(\hat{\boldsymbol{x}}) \leq \hat{v}(\boldsymbol{x}^*) \leq \gamma \max_{\zeta \in \mathcal{Z}} \ \sum_{i=1}^{N} h_i(\boldsymbol{x}^*, \zeta) \ ,$$

where $\boldsymbol{x}^*$ is the optimal solution for problem (1.1), and where we used the fact that $\hat{\boldsymbol{x}}$ is the minimizer for $\hat{v}(\cdot)$ over $\mathcal{X}$, and that γ is the integrality gap for $\max_{\zeta \in \mathcal{Z}} \ \sum_{i=1}^{N} h_i(\boldsymbol{x}^*, \zeta)$. The arguments for linking $\hat{v}(\hat{\boldsymbol{x}})$ to the true optimal value is exactly the same. ∎

Our main attempt toward bounding the integrality gap consists of identifying three sets of conditions on problem (1.4) under which there is no integrality gap for problem (1.4).

Proposition 1.4.3 *For any fixed $\boldsymbol{x} \in \mathbb{R}^n$, given that $\boldsymbol{A} := \boldsymbol{0}$ and $\boldsymbol{b} := \boldsymbol{0}$, problem* (1.4) *and its fractional relaxation have the same optimal value under either of the following sets of conditions:*

1. *The budget Γ is equal to one (L_1-norm ball).*
2. *The budget Γ is equal to m (Box uncertainty set) and there exists $\alpha_{i,k} : \mathbb{R}^n \to \mathbb{R}$ and $\beta_l : \mathbb{R}^n \to \mathbb{R}$ such that for every (i, k) pair $\boldsymbol{c}_{i,k}(\boldsymbol{x}) = \alpha_{i,k}(\boldsymbol{x}) \sum_{l<i}(\beta_l(\boldsymbol{x})\boldsymbol{e}_l)$.*
3. *The budget Γ is integer and there exists $\alpha_{i,k} : \mathbb{R}^n \to \mathbb{R}$ such that for every (i, k) pair $\boldsymbol{c}_{i,k}(\boldsymbol{x}) = \alpha_{i,k}(\boldsymbol{x})\boldsymbol{e}_i$.*

The proof of this proposition is deferred to Appendix 1.9.1 and relies, in the cases of conditions 1-3, on verifying total unimodularity of a matrix that defines an associated polytope to confirm that this polytope has integer vertices. Based on the above result, we can right away conclude about an important property of problem (1.1).

Corollary 1.4.4 *Given that the set of subfunctions* $\{h_i(\boldsymbol{x},\zeta)\}_{i=1}^N$ *satisfies one of the sets of conditions described in Proposition 1.4.3, then problem* (1.5) *is equivalent to problem* (1.1).

Hence, we have in hand a conservative approximation of problem (1.1) that is known to be exact under a fairly general set of conditions. Reading through the three sets of conditions, we might first recognize that Condition 1 reduces the uncertainty set to $\{\zeta | \|\zeta\|_1 \leq 1\}$, a case for which a tractable robust counterpart can also be obtained through vertex enumeration. For the second set, the uncertainty set reduces to $\{\zeta | \|\zeta\|_\infty \leq 1\}$, a set for which the number of vertices is exponential, yet in this case the adversarial problem reduces to a model that was well studied in Bertsimas et al. (2010). One might consider the more important contribution to be related to the third condition which impose that each term of the objective function involve a different term of ζ, and that Γ be integer. Actually, the fact that this results requires the integrality of Γ indicates that it cannot be explained through any of the special cases identified in Gorissen and den Hertog (2013) and Chapter 12 of Ben-Tal et al. (2009a). Furthermore, it extends in a non-trivial way the result of Denton et al. (2010), where $h_i(\boldsymbol{x},\zeta) := \max\{\,0\,,\,\boldsymbol{c}_\zeta^i(\boldsymbol{x})\zeta_i + d_\zeta^i(\boldsymbol{x})\}$ in order to capture delays that are caused by uncertain duration of surgeries in operating rooms.

Remark Note that while one might be able to design a tractable oracle for providing the value and sub-gradient in $\boldsymbol{x}$ of the objective function for problems that satisfy these conditions and thus rely on a cutting-plane method to achieve optimality, the proposed linear programming reformulation has better worst-case convergence rate and can easily be modified to handle binary decision variables.

Remark Note that many different MILP formulations could have been used to replace problem (1.4). Namely, the MILP proposed in Gorissen and den Hertog (2013) takes the form:

$$\begin{aligned}
\underset{z,\zeta,\boldsymbol{y}}{\text{maximize}} \quad & \sum_{i=1}^{N} y_i \\
\text{subject to} \quad & y_i \leq \boldsymbol{c}_{i,k}^T \zeta + d_{i,k} + M(1 - z_{i,k}) \ , \forall\, i, \forall\, k \\
& \boldsymbol{A}\zeta \leq \boldsymbol{b} \\
& -1 \leq \zeta \leq 1 \ \ \& \ \ \sum_{j=1}^{m} |\zeta_j| \leq \Gamma \\
& \sum_{k=1}^{K} z_{i,k} = 1 \ \ \& \ \ z_{i,k} \in \{0, 1\} \ , \forall\, i, \forall\, k \ .
\end{aligned}$$

Our particular choice of formulation is our own best attempt at strategically tightening the integrality gap of the resulting model without paying too much of a price in terms of model size. A side product of our analysis will be to present a perhaps surprising connection to a family of affine approximations used in robust optimization problem where decisions are adjustable.

Remark In recent years, total unimodularity has been somewhat of a fruitful tool for identifying simpler reformulation of risk aware decision problems. In van der Vlerk (2004), the authors show how a two-stage stochastic linear program with binary recourse variables can in some cases be reformulated as a two-stage problem with continuous recourse yet under a different probability measure. In Candia-Véjar et al. (2011), it is a maximum regret minimization problem involving binary variables (*e.g.,* assignment problem) that is reformulated as a simple mixed-integer linear program. In the context of robust optimization, Düzgün and Thiele (2010) introduced an extension to the budgeted uncertainty set that allows parameters to take on values in different sets of intervals and show that the convex hull of possible realizations has a tractable representation. In Mak et al. (2015), the authors exploit some "hidden convexity" to identify a tractable reformulation for a distributionally robust appointment scheduling problem with marginal moment information. Unlike our work which employs the budgeted uncertainty set, the model that is analysed does not capture any correlation between parameters. However, the authors do identify a clever representation of the

$\sum_i h_i(\boldsymbol{x}, \zeta)$ that allows them to handle additional terms that are non-linear function of some ζ_i.

1.4.2 Semi-Definite Programming Approximation Model

Although Section 1.5 and Section 1.6 will present important applications for which one of the three conditions laid out in Proposition 1.4.3 is satisfied, there are still many instances of robust optimization model for which problem (1.5) is inexact. For those instances, there might be a need to dedicate additional computing resources in order to get a better approximation. Drawing from the techniques used to solve or bound the value of mixed-integer quadratic programs, we explore the use of semi-definite programming formulations that might help tighten the integrality gap.

Following the ideas presented in Lovász and Schrijver (1991), we first introduce additional quadratic constraints that are redundant for the mixed-integer program:

$$z_{i,k}^2 = z_{i,k} \;\; \& \;\; z_{i,k} z_{i,k'} = 0 \;, \; \forall\, i, \, \forall\, k \neq k', \, \forall\, i$$
$$0 \leq (\zeta_j^+)^2 \leq \zeta_j^+ \;\; \& \;\; 0 \leq (\zeta_j^-)^2 \leq \zeta_j^- \;, \; \forall\, j \;.$$

Our next step is to introduce a set of N matrices $\boldsymbol{\Lambda}_i \in \mathbb{R}^{K \times K}$, with $i = 1, 2, ..., N$, and two matrices $\boldsymbol{\Lambda}^+, \boldsymbol{\Lambda}^- \in \mathbb{R}^{m \times m}$ as new decision variables that will help characterize the quadratic interactions in the model through $\boldsymbol{\Lambda}_i = \boldsymbol{z}_{i,:} \boldsymbol{z}_{i,:}^T$, $\boldsymbol{\Lambda}^+ = \zeta^+ \zeta^{+T}$, and $\boldsymbol{\Lambda}^- = \zeta^- \zeta^{-T}$. Indeed, we would need that the following constraints be satisfied:

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \mathrm{mat}(\boldsymbol{\Delta}_{i,:}^+) \\ \mathrm{mat}(\boldsymbol{\Delta}_{i,:}^+)^T & \boldsymbol{\Lambda}^+ \end{bmatrix} = \begin{bmatrix} \boldsymbol{z}_{i,:}^T \\ \zeta^+ \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_{i,:} & \zeta^{+T} \end{bmatrix} \;, \; \forall\, i \tag{1.6a}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \mathrm{mat}(\boldsymbol{\Delta}_{i,:}^-) \\ \mathrm{mat}(\boldsymbol{\Delta}_{i,:}^-)^T & \boldsymbol{\Lambda}^- \end{bmatrix} = \begin{bmatrix} \boldsymbol{z}_{i,:}^T \\ \zeta^- \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_{i,:} & \zeta^{-T} \end{bmatrix} \;, \; \forall\, i \tag{1.6b}$$

$$\boldsymbol{\Lambda}_i = \mathrm{diag}(\boldsymbol{z}_{i,:}^T) \;, \; \forall\, i \tag{1.6c}$$

$$\Lambda_{j,j}^+ \leq \zeta_j^+ \;\; \& \;\; \Lambda_{j,j}^- \leq \zeta_j^- \;, \; \forall\, j \tag{1.6d}$$

$$\boldsymbol{\Lambda}^+ \geq 0 \;\; \& \;\; \boldsymbol{\Lambda}^- \geq 0 \;, \tag{1.6e}$$

where $\text{mat}(\boldsymbol{\Delta}^+_{i,:})$ refers to the K by m matrix composed of the terms of $\boldsymbol{\Delta}^+_{i,:}$ organized such that $(\text{mat}(\boldsymbol{\Delta}^+_{i,:}))_{k,j} = (\boldsymbol{\Delta}^+_{i,k})_j$. On the other hand, $\text{diag}(\cdot)$ is an operator that creates a diagonal matrix from a vector: *e.g.*, $(\text{diag}(\boldsymbol{z}^T_{i,:}))_{k,k} = z_{ik}$ while $(\text{diag}(\boldsymbol{z}^T_{i,:}))_{k,k'} = 0$ for all $k \neq k'$.

Unfortunately, equality constraints (1.6a) and (1.6b) are not acceptable in a convex optimization model hence we relax them using a matrix inequality:

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \text{mat}(\boldsymbol{\Delta}^+_{i,:}) \\ \text{mat}(\boldsymbol{\Delta}^+_{i,:})^T & \boldsymbol{\Lambda}^+ \end{bmatrix} \succeq \begin{bmatrix} \boldsymbol{z}^T_{i,:} \\ \zeta^+ \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_{i,:} & \zeta^{+T} \end{bmatrix} , \forall i$$
$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \text{mat}(\boldsymbol{\Delta}^-_{i,:}) \\ \text{mat}(\boldsymbol{\Delta}^-_{i,:})^T & \boldsymbol{\Lambda}^- \end{bmatrix} \succeq \begin{bmatrix} \boldsymbol{z}^T_{i,:} \\ \zeta^- \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_{i,:} & \zeta^{-T} \end{bmatrix} , \forall i$$

which can easily be reformulated as linear matrix inequalities and leads to the following mixed-integer semi-definite program:

$$\underset{\boldsymbol{z},\zeta^+,\zeta^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-,\boldsymbol{\Lambda}_i,\boldsymbol{\Lambda}^+,\boldsymbol{\Lambda}^-}{\text{maximize}} \quad \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}^T_{i,k}(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) + d_{i,k} z_{i,k} \tag{1.7a}$$
$$\text{subject to} \quad \boldsymbol{A}(\zeta^+ - \zeta^-) \leq \boldsymbol{b} \tag{1.7b}$$
$$\zeta^+ \geq 0 \;\;\&\;\; \zeta^- \geq 0 \;\;\&\;\; \zeta^+_j + \zeta^-_j \leq 1 \, , \forall j \tag{1.7c}$$
$$\mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \tag{1.7d}$$
$$\sum_{k=1}^{K} z_{i,k} = 1 \, , \forall i \tag{1.7e}$$
$$\sum_{k=1}^{K} \boldsymbol{\Delta}^+_{i,k} = \zeta^+ \;\;\&\;\; \sum_{k=1}^{K} \boldsymbol{\Delta}^-_{i,k} = \zeta^- \;\; , \forall i \tag{1.7f}$$
$$\boldsymbol{A}(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) \leq \boldsymbol{b} z_{i,k} \, , \forall i, \forall k \tag{1.7g}$$
$$\boldsymbol{\Delta}^+_{i,k} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}^-_{i,k} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}^+_{i,k} + \boldsymbol{\Delta}^-_{i,k} \leq z_{i,k} \, , \forall i, \forall k \tag{1.7h}$$
$$\sum_{j=1}^{m} (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j = \Gamma z_{i,k} \, , \forall i, \forall k \tag{1.7i}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \text{mat}(\boldsymbol{\Delta}^+_{i,:}) & \boldsymbol{z}^T_{i,:} \\ \text{mat}(\boldsymbol{\Delta}^+_{i,:})^T & \boldsymbol{\Lambda}^+ & \zeta^+ \\ \boldsymbol{z}_{i,:} & \zeta^{+T} & 1 \end{bmatrix} \succeq 0 \ , \forall i \tag{1.7j}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \text{mat}(\boldsymbol{\Delta}^-_{i,:}) & \boldsymbol{z}^T_{i,:} \\ \text{mat}(\boldsymbol{\Delta}^-_{i,:})^T & \boldsymbol{\Lambda}^- & \zeta^- \\ \boldsymbol{z}_{i,:} & \zeta^{-T} & 1 \end{bmatrix} \succeq 0 \ , \forall i \tag{1.7k}$$

$$\boldsymbol{\Lambda}_i = \text{diag}(\boldsymbol{z}^T_{i,:}) \ , \forall i \tag{1.7l}$$

$$\Lambda^+_{j,j} \leq \zeta^+_j \ \ \& \ \ \Lambda^-_{j,j} \leq \zeta^-_j \ , \forall j \tag{1.7m}$$

$$\boldsymbol{\Lambda}^+ \geq 0 \ \ \& \ \ \boldsymbol{\Lambda}^- \geq 0 \tag{1.7n}$$

$$z_{i,k} \in \{0, 1\} \ , \forall i, \forall k \ . \tag{1.7o}$$

Since this mixed-integer semi-definite program is equivalent to problem (1.3) yet contains additional constraints compared to problem (1.4), we can expect that its fractional relaxation will lead to a tighter conservative approximation for problem (1.1).

Actually, there are a number of different ways one might choose to tighten the relaxation of problem (1.4) through the addition of linear cuts or lifting in the space of positive semi-definite cones. Problem (1.7) is one such example that leads to a somewhat concise semi-definite program. We refer the interested reader to Lasserre (2002) and Ghaddar et al. (2011) for a hierarchy of polynomial size semi-definite programming relaxation of mixed-integer quadratic programs for which the integrality gap is known to converge to 1. Based on Proposition 1.4.2, it is therefore theoretically possible to find a semi-definite programming model of polynomial size that will generate a solution within a constant factor of the optimal one. Unfortunately, this might often be of little practical relevance. First, this would require us to assess the integrality gap for each of the models in this hierarchy, which can be hard if not impossible to do. Second, the model that is found to achieve a given factor of optimality might be of a size that cannot be solved in a reasonable amount of time. To help resolve this issue, we actually show that one can potentially confirm after solving a conservative approximation model of smaller size that the approximate solution obtained is indeed optimal for problem (1.3). As shown in the following proposition, this is done by verifying whether there is an optimal assignment for the associated relaxed adversarial problem that lies in the convex hull of integer solutions. We refer the reader to Appendix 1.9.4 for a complete proof.

Proposition 1.4.5 *Given a robust optimization problem of the form*

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{\zeta\in\mathcal{U}\cap\mathcal{I}} h(\boldsymbol{x},\zeta) \ ,$$

where $\mathcal{X} \subset \mathbb{R}^n$ *and* $\mathcal{U} \subset \mathbb{R}^m$ *are both bounded convex sets,* $\mathcal{I} = \{\zeta \in \mathbb{R}^m \,|\, \zeta_i \text{ is integer } \forall i \leq q\}$ *for some* $q \leq m$*, hence imposing that a set of terms of* ζ *be integer valued, and* $h(\boldsymbol{x},\zeta)$ *is real valued, convex in* $\boldsymbol{x}$ *and linear in* ζ*. Let* $\hat{\boldsymbol{x}}$ *be the solution of the conservative approximation*

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{\zeta\in\mathcal{U}} h(\boldsymbol{x},\zeta) \ .$$

If there exists a $\hat{\zeta} \in \arg\max_{\zeta\in\mathcal{U}} \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\zeta)$ *that is a member of the convex hull of* $\mathcal{U}\cap\mathcal{I}$*, denoted as* $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$*, then* $\hat{\boldsymbol{x}}$ *is optimal according to the original robust optimization problem.*

When $\mathcal{X}$ does not impose integer constraints, this proposition allows us to imagine a solution scheme in which one generates progressively, based on the current pair $(\hat{\boldsymbol{x}},\hat{\zeta})$, new tightening constraints based on cutting-planes that separates the current $\hat{\zeta}$ from $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$, adds them to problem (1.4), and re-solves the associated conservative approximation.[5] This process has reached optimality whenever it is impossible to separate $\hat{\zeta}$ from $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$. Note that if $\mathcal{U}$ only has integer vertices, although one might not be aware of it, then optimality of $\boldsymbol{x}$ is confirmed instantly when failing to separate the first proposal for $\hat{\zeta}$.

Corollary 1.4.6 *If* $\mathcal{P}(\mathcal{U}\cap\mathcal{I}) = \mathcal{U}$*, then the optimality of* $\hat{x} \in \arg\min_{\boldsymbol{x}\in\mathcal{X}} \ \max_{\zeta\in\mathcal{U}} h(\boldsymbol{x},\zeta)$ *is necessarily confirmed when verifying that* $\hat{\zeta} \in \arg\max_{\zeta\in\mathcal{U}} \ \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\zeta)$ *is a member of the convex hull of* $\mathcal{U}\cap\mathcal{I}$*.*

---

[5] New variables, such as positive semi-definite matrices, can also be added to the adversarial model, in a lifting and projection fashion, as long as these variables are known to be bounded.

Remark For completeness, we briefly outline an algorithm that can be used to determine whether $\hat{\zeta}$ is in the convex hull of $\mathcal{U} \cap \mathcal{I}$. First, let us recall an equivalent definition for convex hull

$$\mathcal{P}(\mathcal{U} \cap \mathcal{I}) = \left\{ \zeta \in \mathbb{R}^m \;\middle|\; \boldsymbol{c}^T \zeta \leq \sup_{\zeta' \in \mathcal{U} \cap \mathcal{I}} \boldsymbol{c}^T \zeta' \;,\; \forall \boldsymbol{c} \in \mathcal{B}(1) \right\} \;,$$

where $\mathcal{B}(1) = \{\boldsymbol{c} \in \mathbb{R}^m \mid \|\boldsymbol{c}\|_2 \leq 1\}$. Based on this definition, verifying membership of $\hat{\zeta}$ to the convex hull reduces to validating whether $\min_{\boldsymbol{c} \in \mathcal{B}(1)} \sup_{\zeta' \in \mathcal{U} \cap \mathcal{I}} \boldsymbol{c}^T(\zeta' - \hat{\zeta})$ is greater or equal to zero or not (if not the argument that minimizes this expression can be used to generate a cutting-plane). Finding the minimum of such an expression can be done using a cutting-plane algorithm as long as one has an efficient algorithm to solve $\sup_{\zeta' \in \mathcal{U} \cap \mathcal{I}} \boldsymbol{c}^T(\zeta' - \hat{\zeta})$ when $\boldsymbol{c}$ is fixed. In practice, one might use CPLEX to do so.

1.4.3 Relation to Affinely Adjustable Robust Counterparts

We now provide an explicit connection between our approach and affinely adjustable robust counterparts methods. In fact,we demonstrate below that any model that is obtained by exploiting affine decision rules can also be motivated using our mixed-integer linear programming based approximation scheme. This is interesting since it indicates that our new scheme is somewhat more flexible and, perhaps more importantly, implies the possibility of generalizing the techniques discussed in Section 1.4.2 so that they can be used to improve the quality of solutions obtained from any AARC approximation of robust multi-stage problems.

Proposition 1.4.7 *Given an adversarial problem of the type*

$$\underset{\zeta \in \mathcal{A}}{\text{maximize}} \quad \sum_{i=1}^{N} \max_k \boldsymbol{c}_{i,k}^T \zeta + d_{i,k} \;,$$

where $\mathcal{A} = \{\zeta | \boldsymbol{A}\zeta \leq \boldsymbol{b}\}$ is a bounded polyhedron, the optimal value of its affinely adjustable robust counterpart

$$\begin{aligned}
\underset{\boldsymbol{\lambda},\boldsymbol{\gamma}}{\text{minimize}} \quad & \max_{\zeta \in \mathcal{A}} \sum_{i=1}^{N} \boldsymbol{\lambda}_i^T \zeta + \gamma_i && \text{(1.8a)} \\
\text{subject to} \quad & \boldsymbol{\lambda}_i^T \zeta + \gamma_i \geq \boldsymbol{c}_{i,k}^T \zeta + d_{i,k} \;\;, \forall\, i,\, \forall\, k,\, \forall\, \zeta \in \mathcal{A} && \text{(1.8b)}
\end{aligned}$$

is equal to the optimal value of the fractional relaxation of the mixed-integer linear programming problem

$$\begin{aligned}
\underset{\boldsymbol{z},\zeta,\boldsymbol{\Delta}}{\text{maximize}} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}_{i,k}^T \boldsymbol{\Delta}_{i,k} + d_{i,k} z_{i,k} && \text{(1.9a)}\\
\text{subject to} \quad & \boldsymbol{A}\zeta \leq \boldsymbol{b} && \text{(1.9b)}\\
& \sum_{k=1}^{K} z_{i,k} = 1\,,\ \forall\, i && \text{(1.9c)}\\
& \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = \zeta\ ,\ \forall\, i && \text{(1.9d)}\\
& \boldsymbol{A}\boldsymbol{\Delta}_{i,k} \leq \boldsymbol{b} z_{i,k}\ ,\ \forall\, i,k && \text{(1.9e)}\\
& z_{i,k} \in \{0,1\}\ ,\ \forall\, i,k\ . && \text{(1.9f)}
\end{aligned}$$

Proof The optimal value of the fractional relaxation of problem (1.9), can be presented in the form

$$\begin{aligned}
\max_{\zeta\in\mathcal{A}} \qquad \max_{\boldsymbol{z},\boldsymbol{\Delta}} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}_{i,k}^T \boldsymbol{\Delta}_{i,k} + d_{i,k} z_{i,k} && \text{(1.10a)}\\
\text{subject to} \quad & \sum_{k=1}^{K} z_{i,k} = 1\,,\ \forall\, i && \text{(1.10b)}\\
& \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = \zeta\ ,\ \forall\, i && \text{(1.10c)}\\
& \boldsymbol{A}\boldsymbol{\Delta}_{i,k} \leq \boldsymbol{b} z_{i,k}\ ,\ \forall\, i,k && \text{(1.10d)}\\
& z_{i,k} \geq 0 \ \ \forall i,k. && \text{(1.10e)}
\end{aligned}$$

For any fixed $\zeta \in \mathcal{A}$, since the inner problem is linear and has a feasible solution, strict duality applies and its optimal value is equal to the optimal value of its dual problem. Hence, this inner problem can be reformulated as

$$\begin{aligned}
\underset{\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \quad & \sum_{i=1}^{N} \gamma_i + \boldsymbol{\lambda}_i^T \zeta && \text{(1.11a)}\\
\text{subject to} \quad & \gamma_i - \boldsymbol{b}^T \boldsymbol{\psi}_{i,k} \geq d_{i,k}\,,\ \forall\, i,k && \text{(1.11b)}\\
& \boldsymbol{A}^T \boldsymbol{\psi}_{i,k} + \boldsymbol{\lambda}_i = \boldsymbol{c}_{i,k}\,,\ \forall\, i,k && \text{(1.11c)}
\end{aligned}$$

$$\boldsymbol{\psi}_{i,k} \geq 0\,,\ \forall\, i,k\ , \tag{1.11d}$$

where $\gamma_i \in \mathbb{R}$, $\boldsymbol{\lambda}_i \in \mathbb{R}^m$, and $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^p$ are respectively the dual variables associated to constraints (1.10b), (1.10c), and (1.10d). Since $\mathcal{A}$ is bounded and convex and the objective function (1.11a) is bilinear in ζ and the set of variable $\{\boldsymbol{\gamma}, \boldsymbol{\lambda}, \boldsymbol{\psi}\}$, Sion's minimax theorem guarantees that the maximin formulation is equal to the minimax one, hence the optimal value of problem (1.10) is equal to the optimal value of

$$\underset{\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \quad \max_{\zeta \in \mathcal{A}} \sum_{i=1}^{N} \gamma_i + \boldsymbol{\lambda}_i^T \zeta \tag{1.12a}$$

$$\text{subject to} \quad \gamma_i - \boldsymbol{b}^T \boldsymbol{\psi}_{i,k} \geq d_{i,k}\,,\ \forall\, i,k \tag{1.12b}$$

$$\boldsymbol{A}^T \boldsymbol{\psi}_{i,k} + \boldsymbol{\lambda}_i = \boldsymbol{c}_{i,k}\,,\ \forall\, i,k\ . \tag{1.12c}$$

It can be shown that constraints (1.12b) and (1.12c) are equivalent to robust constraint (1.8b). Specifically, for all index pair (i,k) the right-hand side of the robust constraint of (1.8b) can be formulated as

$$\underset{\zeta}{\text{maximize}} \quad \boldsymbol{c}_{i,k}^T \zeta - \boldsymbol{\lambda}_i^T \zeta \tag{1.13a}$$

$$\text{subject to} \quad \boldsymbol{A}\zeta \leq \boldsymbol{b}\ . \tag{1.13b}$$

Since strict duality applies once again, problem (1.13) gives the same optimal value as:

$$\begin{aligned} \underset{\boldsymbol{\psi}_{i,k}}{\text{minimize}} \quad & b^T \boldsymbol{\psi}_{i,k} \\ \text{subject to} \quad & \boldsymbol{A}^T \boldsymbol{\psi}_{i,k} = \boldsymbol{c}_{i,k} - \boldsymbol{\lambda}_i\,,\ \forall\, i,k \\ & \boldsymbol{\psi}_{i,k} \geq 0\,,\ \forall\, i,k\ , \end{aligned}$$

where each $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^p$ contains the dual variables associated to constraint (1.13b). It is therefore clear that robust constraint (1.8b) can be reformulated as

$$\begin{aligned} \gamma_i - d_{i,k} \geq \boldsymbol{b}^T \boldsymbol{\psi}_{i,k}\,,\ \forall\, i,k \\ \boldsymbol{A}^T \boldsymbol{\psi}_{i,k} = \boldsymbol{c}_{i,k} - \boldsymbol{\lambda}_i\,,\ \forall\, i,k. \end{aligned}$$

We conclude that formulation (1.9) gives the same optimal value as problem (1.8). ∎

This result is interesting as it establishes that any robust counterpart that is obtained using an affine decision rule scheme can be thought of in terms of replacing the adversarial problem with the fractional relaxation of an equivalent MILP. In particular, one can easily verify that using affine decision rules under the particular lifting $\zeta := \zeta^+ - \zeta^-$, with $\zeta^+ \geq 0$ and $\zeta^- \geq 0$, is equivalent to problem (1.5).

Corollary 1.4.8 *The fractional relaxation of problem* (1.4) *is equivalent to the affinely adjusted approximation of problem* (1.3) *applied to the lifted set of perturbation* $\zeta = \zeta^+ - \zeta^-$, *where* ζ^+ *and* ζ^- *are the positive and negative parts of* ζ. *Specifically, it achieves the same optimal value as the problem:*

$$\begin{array}{lll} \underset{\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\gamma}}{\text{minimize}} & \displaystyle\max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} & \displaystyle\sum_{i=1}^{N} \boldsymbol{\lambda}_i^{+T}\zeta^+ + \boldsymbol{\lambda}_i^{-T}\zeta^- + \gamma_i \\ \text{subject to} & \multicolumn{2}{l}{\boldsymbol{\lambda}_i^{+T}\zeta^+ + \boldsymbol{\lambda}_i^{-T}\zeta^- + \gamma_i \geq \boldsymbol{c}_{i,k}^T(\zeta^+ - \zeta^-) + d_{i,k}\ ,\ \forall\, i,\ k,\ \forall\,(\zeta^+,\zeta^-) \in \mathcal{Z}'\ ,} \end{array}$$

where

$$\mathcal{Z}' = \left\{ (\zeta^+,\zeta^-) \;\middle|\; \begin{array}{l} \zeta^+ \geq 0\ ,\ \zeta^- \geq 0 \\ \zeta_j^+ + \zeta_j^- \leq 1\ \forall\, j \\ \sum_j \zeta_j^+ + \zeta_j^- = \Gamma \\ \boldsymbol{A}(\zeta^+ - \zeta^-) \leq \boldsymbol{b} \end{array} \right\} .$$

While there is a lot of empirical evidence supporting the strength of affinely adjusted approximation schemes (see for instance Ben-Tal et al. (2005)), very little is actually known theoretically about the quality of the approximations that are obtained with these methods, either in terms of optimal value or optimal solution. To the best of our knowledge, the authors of Iancu et al. (2013) are the ones that have identified to this date the most general class of problems for which the approximation was exact. We refer the readers in particular to Theorem 3 in their article which could potentially provide an alternative method for deriving our Proposition 1.4.3 given the connection established in Corollary 1.4.8. Note however that while Iancu et al. do identify problem instances where conditions (P1) and

(P2) of Theorem 3 in their paper are satisfied under box uncertainty or the simplex (given the implicit sublattice structure of these uncertainty sets), they left open the question of identifying such instances for a general budgeted uncertainty set with integer budget.

Overall, our new interpretation of AARC methods is particularly interesting as it states that asking whether AARC methods are exact is equivalent to asking whether an associated MILP (problem (1.4) for example) has an integrality gap or not. Given the extensive efforts that have been dedicated in the last few decades both to developing approximation methods for MILP that are based on fractional relaxation schemes and to measuring the quality of these bounds, we have good hope to find innovative ways of improving the performance of these AARC models.

We conclude this section with a result that establishes the connection between the semi-definite programming based conservative approximation presented in Section 1.4.2 and the theory of AARCs. The proof of this final connection can be found in Appendix 1.9.5.

Proposition 1.4.9 *The optimal value of the fractional relaxation of problem* (1.7) *is equal to the optimal value of the affinely adjustable robust counterpart of*

$$\begin{aligned}
\underset{\substack{\{\boldsymbol{v}_i\}_{i=1}^N,\boldsymbol{w}^+,\boldsymbol{w}^-,\\ \{\boldsymbol{Q}_i^+,\boldsymbol{V}_i^+,\boldsymbol{S}_i^+\}_{i=1}^N\\ \{\boldsymbol{q}_i^+,\boldsymbol{p}_i^+,r_i^+\}_{i=1}^N\\ \{\boldsymbol{Q}_i^-,\boldsymbol{V}_i^-,\boldsymbol{S}_i^-\}_{i=1}^N\\ \{\boldsymbol{q}_i^-,\boldsymbol{p}_i^-,r_i^-\}_{i=1}^N}}{\text{minimize}} \quad & \max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} \; (\boldsymbol{w}^+)^T\zeta^+ + (\boldsymbol{w}^-)^T\zeta^- + \textstyle\sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + 2\boldsymbol{p}_i^{-T}\zeta^- && (1.14\text{a})\\
& \quad + \max_k\{\boldsymbol{c}_{i,k}^T(\zeta^+-\zeta^-) + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + 2(\boldsymbol{V}_i^-)_{k,:}\zeta^- + (\boldsymbol{v}_i)_k\}\\
\text{subject to} \quad & (\boldsymbol{v}_i)_k = (\boldsymbol{Q}_i^+ + \boldsymbol{Q}_i^-)_{k,k} + 2(\boldsymbol{q}_i^+ + \boldsymbol{q}_i^-)_k + r_i^+ + r_i^- \;, \forall i,\, k && (1.14\text{b})\\
& \begin{bmatrix} \boldsymbol{Q}_i^+ & \boldsymbol{V}_i^+ & \boldsymbol{q}_i^+\\ \boldsymbol{V}_i^{+T} & \boldsymbol{S}_i^+ & \boldsymbol{p}_i^+\\ \boldsymbol{q}_i^{+T} & \boldsymbol{p}_i^{+T} & r_i^+ \end{bmatrix} \succeq 0, \; \begin{bmatrix} \boldsymbol{Q}_i^- & \boldsymbol{V}_i^- & \boldsymbol{q}_i^-\\ \boldsymbol{V}_i^{-T} & \boldsymbol{S}_i^- & \boldsymbol{p}_i^-\\ \boldsymbol{q}_i^{-T} & \boldsymbol{p}_i^{-T} & r_i^- \end{bmatrix} \succeq 0 \;, \forall i && (1.14\text{c})\\
& \sum_{i=1}^N \boldsymbol{S}_i^+ \leq \operatorname{diag}(\boldsymbol{w}^+) \;, \; \sum_{i=1}^N \boldsymbol{S}_i^- \leq \operatorname{diag}(\boldsymbol{w}^-) \;, \forall i && (1.14\text{d})\\
& \boldsymbol{w}^+ \geq 0 \quad \boldsymbol{w}^- \geq 0 \;, && (1.14\text{e})
\end{aligned}$$

where $\boldsymbol{w}^+ \in \mathbb{R}^m$, $\boldsymbol{w}^- \in \mathbb{R}^m$, *while for each* i, $\boldsymbol{v}_i \in \mathbb{R}^K$, $\boldsymbol{Q}_i^+ \in \mathbb{R}^{K\times K}$, $\boldsymbol{Q}_i^- \in \mathbb{R}^{K\times K}$, $\boldsymbol{V}_i^+ \in \mathbb{R}^{K\times m}$, $\boldsymbol{V}_i^- \in \mathbb{R}^{K\times m}$, $\boldsymbol{q}_i^+ \in \mathbb{R}^K$, $\boldsymbol{q}_i^- \in \mathbb{R}^K$, $\boldsymbol{S}_i^+ \in \mathbb{R}^{m\times m}$, $\boldsymbol{S}_i^- \in \mathbb{R}^{m\times m}$, $\boldsymbol{p}_i^+ \in \mathbb{R}^m$, $\boldsymbol{p}_i^- \in \mathbb{R}^m$, $r_i^+ \in \mathbb{R}$, $r_i^- \in \mathbb{R}$, *and finally where*

$$\mathcal{Z}' = \left\{ (\zeta^+, \zeta^-) \left| \begin{array}{l} \zeta^+ \geq 0\ ,\ \zeta^- \geq 0 \\ \zeta_j^+ + \zeta_j^- \leq 1\ \forall j \\ \sum_j \zeta_j^+ + \zeta_j^- = \Gamma \\ \boldsymbol{A}(\zeta^+ - \zeta^-) \leq \boldsymbol{b} \end{array} \right. \right\} .$$

One might actually notice that in problem (1.14), when all the variables that are minimized over are set to zero, which is a feasible assignment, the problem reduces to problem (1.3) with the lifted set of perturbation $\zeta = \zeta^+ - \zeta^-$. Intuitively, the SDP model is able to obtain a tighter bound by adding some affine perturbations that have a positive net effect on the evaluation of the objective function yet might allow to reach a lower amount when affine decision rules are introduced. In particular, in the case where $\zeta^- := \boldsymbol{0}$, which we assume for simplicity of exposure, and $0 \leq \zeta^+ \leq 1$, when constraints (1.14b) to (1.14e) are satisfied, then the objective function can be shown to satisfy

$$\begin{aligned}
&(\boldsymbol{w}^+)^T\zeta^+ + \sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + \max_k\{\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + (\boldsymbol{Q}_i^+)_{kk} + 2(\boldsymbol{q}_i^+)_k + r_i^+\} \\
&\geq\ (\boldsymbol{w}^+)^T(\zeta^+)^2 + \sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + \max_k\{\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + (\boldsymbol{Q}_i^+)_{kk} + 2(\boldsymbol{q}_i^+)_k + r_i^+\} \\
&=\ (\boldsymbol{w}^+)^T(\zeta^+)^2 + \sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \sum_{k=1}^K z_k\{\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + (\boldsymbol{Q}_i^+)_{kk} + 2(\boldsymbol{q}_i^+)_k + r_i^+\} \\
&\geq\ \sum_{i=1}^N \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \zeta^{+T}\boldsymbol{S}_i\zeta^+ + 2\boldsymbol{p}_i^{+T}\zeta^+ + \sum_{k=1}^K z_k(\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + (\boldsymbol{Q}_i^+)_{kk} + 2(\boldsymbol{q}_i^+)_k + r_i^+) \\
&\geq\ \sum_{i=1}^N \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \sum_{k=1}^K z_k(\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k}) \quad = \quad \sum_{i=1}^N \max_k\{\boldsymbol{c}_{i,k}^T\zeta^+ + d_{i,k}\}\ ,
\end{aligned}$$

where we used the fact that $\zeta^+ \in [0,1]^m$ to get the first inequality, and constraint (1.14d) to get the second. We finally used the fact that $(\boldsymbol{Q}_i^+)_{kk}z_k = \sum_{k'}(\boldsymbol{Q}_i^+)_{kk}z_k z_{k'}$ over the feasible region that is considered, and used constraint (1.14c) which implies that

$$\zeta^{+T}\boldsymbol{S}_i\zeta^+ + 2\boldsymbol{p}_i^{+T}\zeta^+ + 2\boldsymbol{z}^T(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + \boldsymbol{z}^T\boldsymbol{Q}_i^+\boldsymbol{z}^T + 2\boldsymbol{z}^T\boldsymbol{q}_i^+ + r_i^+ \geq 0\ .$$

A similar argument can be made when we involve the $\zeta^- > 0$. Hence, problem (1.14) necessarily provides a tight upper bound to problem (1.3).

Overall, the connection established in Proposition 1.4.9 indicates that the scheme we adopt in this paper allows to identify tractable conservative approximations that provide tighter bounds than the well known applications of affine decision rules. Indeed, while schemes that are based on quadratic decision rules can in some cases lead to SDP approximation models, such adjustment functions lead in general to optimization problems that are computationally intractable when the uncertainty set is polyhedral (see Ben-Tal et al. (2009a) p. 372). While one might be able to further approximate those models by applying Sum-Of-Squares techniques as were proposed in Bertsimas et al. (2011b), such an approach leads to SDP models of much larger size than the model presented here.

1.4.4 Empirical Evaluation of Integrality Gap

We briefly present a set of empirical experiments that illustrates the trade-off that need to be made between computational effort and quality of the upper bound obtained for problem (1.3) with $\boldsymbol{A} := \boldsymbol{0}$ and $\boldsymbol{b} := \boldsymbol{0}$ using three different fractional relaxation schemes. The first bound is obtained by applying affine decision rules directly on ζ; this method will be referred as AARC. We also compare the two improved bounds based on linear program (1.4), and semi-definite program (1.7). These methods compete on a set of 100 randomly generated instances of problem (1.3) which we solved exactly using CPLEX. Each problem instance is generated by sampling each parameters of the objective function uniformly between -1 and 1, and then ensuring that the optimal value is positive by adding a constant term that makes $\sum_i \max_k \boldsymbol{c}_{i,k}^T \boldsymbol{0} + d_{i,k} = 0$; furthermore, a random integer budget of Γ is generated uniformly between 1 to N. Based on the results presented in Table 1–1, we see that the quality of each bound degrades as N increases yet an approach based on semi-definite programming will achieve significant improvement in tightness. On the other hand, there is a heavier computational price to pay for the semi-definite programming model. It is also observed that LP (1.4) provides better results than AARC. Note that all linear programming models

Table 1–1: Empirical evaluation of integrality gap and resolution time for a set of randomly generated convex maximization problems of form (1.3). All problems have size $N = m = n$ and $K = 2$. Note that in the case of $N = 64$, it took longer than a full day to solve the MILP with CPLEX. We therefore choose to report the minimum, maximum and average relative improvement (over 100 instances) of the bounds obtained by each model compared to the bound obtained with AARC. Note that an integrality gap of one is optimal.

Size		AARC	LP (1.4)	SDP (1.7)
$N = 8$	CPU time	0.062 sec	0.064 sec	1.564 sec
	Gap = 1 instances	14%	29%	59%
	Largest gap	1.70	1.58	1.09
	Average gap	1.26	1.14	1.01
$N = 16$	CPU time	0.17 sec	0.18 sec	27.89 sec
	Gap = 1 instances	3%	6%	11%
	Largest gap	2.32	2.12	1.14
	Average gap	1.82	1.49	1.05
$N = 32$	CPU time	10 sec	10 sec	28.9 min
	Gap = 1 instances	0%	0%	0%
	Largest gap	2.94	2.80	1.22
	Average gap	2.61	1.96	1.10
$N = 64$	CPU time	34 sec	70 sec	19 h
	Min improvement	-	0%	43%
	Max improvement	-	54%	72%
	Avg. improvement	-	26%	68%

were solved using CPLEX 12.4 while the SDP was solved using DSDP 5.8 (Benson et al. 2000).

1.5 Robust Multi-Item Newsvendor Problem

We now pay closer attention to the multi-item newsvendor problem as described in a general form through the following model:

$$\underset{\boldsymbol{x}\in\mathcal{X},0\leq\boldsymbol{y}\leq\min(\boldsymbol{x},\boldsymbol{w})}{\text{maximize}} \quad \sum_{i=1}^{m} v_i y_i - c_i x_i + g_i(x_i - y_i) - b_i(w_i - y_i) \ ,$$

where x_i represents the amount of item i ordered, w_i the demand for this item while $\mathcal{X}$ captures the set of feasible orders and y_i is the (second-stage) amount sold once the demand is known. We also denote the following terms: $c_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}^n$ are respectively the per unit ordering cost and retail prices, $b_i(\cdot)$ is a piecewise linear convex increasing stock-out cost, and $g_i(\cdot)$ is a piecewise linear decreasing concave salvage prices. We assume that for each item $\partial g_i(z)/\partial z < c_i < v_i$ whenever the derivative exists, *i.e.,* that the unit ordering cost is always larger than the marginal salvage price and always lower than the retail price. One can therefore not make profits out of salvaging his products. Since $y_i^* = \min(x_i, w_i)$, the two-stage model is equivalent to:

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sum_{i=1}^{m} c_i x_i - v_i \min(x_i, w_i) - g_i(x_i - \min(x_i, w_i)) + b_i(w_i - \min(x_i, w_i)) \ ,$$

which is presented in terms of minimizing negative profits to be coherent with problem (1.1).

Further manipulations of the model will lead to a form that makes the connection with problem (1.1) explicit:

$$\begin{aligned}
&(c_i x_i - v_i \min(x_i, w_i)) - g_i(x_i - \min(x_i, w_i)) + b_i(w_i - \min(x_i, w_i)) \\
&= (c_i - v_i)w_i + c_i(x_i - w_i)^+ - (c_i - v_i)(w_i - x_i)^+ - g_i((x_i - w_i)^+) + b_i((w_i - x_i)^+) \\
&= (c_i - v_i)w_i + (c_i(x_i - w_i) - g_i(x_i - w_i))^+ + ((v_i - c_i)(w_i - x_i) + b_i(w_i - x_i))^+ \\
&= \max\left(c_i x_i - v_i w_i - g_i(x_i - w_i)\ ,\ (c_i - v_i)x_i + b_i(w_i - x_i)\right) \\
&= \max\left(\max_{k\in\{1,2,\ldots,K^g\}} c_i x_i - v_i w_i - \alpha_{i,k}^g(x_i - w_i) - \beta_{i,k}^g\ ,\ \max_{k\in\{1,2,\ldots,K^b\}} (c_i - v_i)x_i + \alpha_{i,k}^b(w_i - x_i) + \beta_{i,k}^b\right) \\
&= \max_k \alpha_{i,k}^x x_i + \alpha_{i,k}^w w_i + \beta_{i,k} \ ,
\end{aligned}$$

where we exploited the piecewise linear concave and convex structures of

$$g_i(y) = \min_{k\in\{1,2,\ldots,K^g\}} \alpha_{i,k}^g y + \beta_{i,k}^g$$

and

$$b_i(y) = \max_{k\in\{1,2,\ldots,K^b\}} \alpha_{i,k}^b y + \beta_{i,k}^b$$

respectively, and later combined the indexes of the two layers of maximum operators so that

$$\alpha_{i,k}^x = \begin{cases} c_i - \alpha_{i,k}^g & \text{if } k \leq K^g \\ c_i - v_i - \alpha_{i,k-K^g}^b & \text{if } k > K^g \end{cases}$$

$$\alpha_{i,k}^w = \begin{cases} \alpha_{i,k}^g - v_i & \text{if } k \leq K^g \\ \alpha_{i,k-K^g}^b & \text{if } k > K^g \end{cases}$$

$$\beta_{i,k} = \begin{cases} -\beta^g_{i,k} & \text{if } k \le K^g \\ \beta^b_{i,k-K^g} & \text{if } k > K^g \end{cases} .$$

When considering robustness in the multi-item newsvendor problem, we introduce a budgeted uncertainty set for the demand vector. Specifically, we assume that the nominal demand vector takes the form $\bar{w}$, that each term is known to lie in the interval $w_i \in [\bar{w}_i - \hat{w}_i,\ \bar{w}_i + \hat{w}_i]$ and that we do not expect the total perturbation to exceed a budget of Γ. Hence, the robust model takes the form:

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{\zeta\in\mathcal{Z}(\Gamma)} \sum_{i=1}^{m} \max_k \alpha^x_{i,k} x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} \ , \tag{1.15}$$

where $\mathcal{Z}(\Gamma) := \{\zeta \in \mathbb{R}^m | \|\zeta\|_\infty \le 1,\ \|\zeta\|_1 \le \Gamma\}$. One might easily recognize in this form that each term of the objective function depends on a different component of ζ. It is therefore clear based on Corollary 1.4.4 that using the robust counterpart presented in problem (1.5) will provide an exact solution. The proof of the following corollary, presented in Appendix 1.9.6, serves to justify how to obtain a more compact robust counterpart.

Corollary 1.5.1 *Given that* Γ *is a strictly positive integer, then the robust multi-item newsvendor problem* (1.15) *is equivalent to the following linear program:*

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathcal{X},\nu,\boldsymbol{\gamma},\boldsymbol{\psi}}{\text{minimize}} \quad & \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma} \\ \text{subject to} \quad & \gamma_i \ge \psi_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} \ \ \forall i,\ \forall k \\ & \psi_{i,k} + \nu \ge \alpha^w_{i,k}\hat{w}_i \ \ \forall i,\ \forall k \\ & \psi_{i,k} + \nu \ge -\alpha^w_{i,k}\hat{w}_i \ \ \forall i,\ \forall k \\ & \psi_{i,k} \ge 0 \ , \forall i,\ \forall k \ , \end{aligned}$$

where $\nu \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^m$, *and* $\psi_{i,k} \in \mathbb{R}$.

Given that the distribution-free version of the newsvendor problem has received so much attention over the last fifty years (see Scarf (1958), Gallego and Moon (1993), Moon and Silver (2000), Wang et al. (2015), Hanasusanto et al. (2014), and Wiesemann et al. (2014)

for some examples), we provide below an exact reformulation for a model that seeks the distributionally robust newspaper quantities when the only information that is available about the distribution includes that the support is $\mathcal{Z}(\Gamma)$, the mean vector is μ, and a list of lower bounds on first order partial moments. To the best of our knowledge, this appears to be the first tractable exact reformulation for such a problem when there exists information about how the demand for different items behave jointly.

Proposition 1.5.2 *The distributionally robust optimization model*

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{F\in\mathcal{D}} \; \mathbb{E}_F\left[\sum_{i=1}^{m} \max_k \alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k}\right] , \tag{1.16}$$

where

$$\mathcal{D} = \left\{ F \in \mathcal{M} \;\middle|\; \begin{array}{l} \mathbb{P}_F(\zeta \in \mathcal{Z}(\Gamma)) = 1 \\ \mathbb{E}_F[\zeta] = \boldsymbol{\mu} \\ \mathbb{E}_F[(\zeta - \boldsymbol{\mu})^+] \geq \boldsymbol{r}^+ \\ \mathbb{E}_F[(\boldsymbol{\mu} - \zeta)^+] \geq \boldsymbol{r}^- \end{array} \right\} ,$$

is equivalent to the following linear program

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathcal{X},t,\boldsymbol{q},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\nu,\boldsymbol{\gamma},\boldsymbol{\psi}^+,\boldsymbol{\psi}^-}{\text{minimize}} \quad & t + \boldsymbol{\mu}^T\boldsymbol{q} - (\boldsymbol{r}^+)^T\boldsymbol{\lambda}^+ - (\boldsymbol{r}^-)^T\boldsymbol{\lambda}^- && \text{(1.17a)}\\
\text{subject to} \quad & t \geq \Gamma\nu + \boldsymbol{1}^T\boldsymbol{\gamma} && \text{(1.17b)}\\
& \gamma_i \geq \psi^+_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} - \lambda^+_i\mu_i \;\; \forall i,\ \forall k && \text{(1.17c)}\\
& \gamma_i \geq \psi^-_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} + \lambda^-_i\mu_i \;\; \forall i,\ \forall k && \text{(1.17d)}\\
& \psi^+_{i,k} + \nu \geq \alpha^w_{i,k}\hat{w}_i - q_i + \lambda^+_i \;\; \forall i,\ \forall k && \text{(1.17e)}\\
& \psi^+_{i,k} + \nu \geq -\alpha^w_{i,k}\hat{w}_i + q_i - \lambda^+_i \;\; \forall i,\ \forall k && \text{(1.17f)}\\
& \psi^-_{i,k} + \nu \geq \alpha^w_{i,k}\hat{w}_i - q_i - \lambda^-_i \;\; \forall i,\ \forall k && \text{(1.17g)}\\
& \psi^-_{i,k} + \nu \geq -\alpha^w_{i,k}\hat{w}_i + q_i + \lambda^+_i \;\; \forall i,\ \forall k && \text{(1.17h)}\\
& \psi^+_{i,k} \geq 0 \;\; , \quad \psi^-_{i,k} \geq 0 \;\; \forall i,\ \forall k && \text{(1.17i)}\\
& \boldsymbol{\lambda}^+ \geq 0 \;\; , \;\; \boldsymbol{\lambda}^- \geq 0\,. && \text{(1.17j)}
\end{aligned}$$

Proof Applying duality theory for semi-infinite linear programs to the inner problem of the distributionally robust problem (1.16), we obtain the following reformulation (see Delage and Ye (2010) for derivation details):

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathcal{X},t,\boldsymbol{q},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-}{\text{minimize}} \quad & t + \boldsymbol{\mu}^T\boldsymbol{q} - (\boldsymbol{r}^+)^T\boldsymbol{\lambda}^+ - (\boldsymbol{r}^-)^T\boldsymbol{\lambda}^- \\
\text{subject to} \quad & t \geq \sum_{i=1}^{m} \max_k \alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} - q_i\zeta_i \\
& \quad + \lambda^+_i \max(0, \zeta_i - \mu_i) + \lambda^-_i \max(0, \mu_i - \zeta_i) \ \ \forall \zeta \in \mathcal{Z}(\Gamma) \\
& \boldsymbol{\lambda}^+ \geq 0 \ \ , \ \ \boldsymbol{\lambda}^- \geq 0.
\end{aligned}$$

One might realize that the right-hand side equation of the infinite set of constraint indexed by ζ is the sum of piecewise linear convex functions in ζ with $2K$ pieces; for each i, each affine piece gives the highest value over k between either of the two following functions

$$\alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} - q_i\zeta_i + \lambda^+_i(\zeta_i - \mu_i)$$

or

$$\alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} - q_i\zeta_i + \lambda^-_i(\mu_i - \zeta_i) \ .$$

Applying similar steps as provided in the proof of Corollary 1.5.1, we obtain a conservative approximation of the right-hand side equation

$$\begin{aligned}
t \ \geq \quad \min_{\nu,\boldsymbol{\gamma},\boldsymbol{\psi}^+,\boldsymbol{\psi}^-} \quad & \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma} \\
\text{subject to} \quad & \gamma_i \geq \psi^+_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} - \lambda^+_i\mu_i \ \ \forall i, \ \forall k \\
& \gamma_i \geq \psi^-_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} + \lambda^-_i\mu_i \ \ \forall i, \ \forall k \\
& \psi^+_{i,k} + \nu \geq \alpha^w_{i,k}\hat{w}_i - q_i + \lambda^+_i \ \ \forall i, \ \forall k \\
& \psi^+_{i,k} + \nu \geq -\alpha^w_{i,k}\hat{w}_i + q_i - \lambda^+_i \ \ \forall i, \ \forall k \\
& \psi^-_{i,k} + \nu \geq \alpha^w_{i,k}\hat{w}_i - q_i - \lambda^-_i \ \ \forall i, \ \forall k \\
& \psi^-_{i,k} + \nu \geq -\alpha^w_{i,k}\hat{w}_i + q_i + \lambda^+_i \ \ \forall i, \ \forall k \\
& \psi^+_{i,k} \geq 0 \ \ , \ \ \psi^-_{i,k} \geq 0 \ \ \forall i, \ \forall k \ .
\end{aligned}$$

This constraint can then easily be re-inserted in the main problem to obtain the model presented in (1.17). Furthermore, since for each i, each affine pieces only depend on ζ_i, we conclude that this approximation is exact based on Condition 3 of Corollary 1.4.4 being satisfied. $\blacksquare$

We refer the reader to Appendix 1.9.7 for an additional exact reformulation of a distributionally robust multi-item newsvendor problem in which one instead imposes lower bounds on the probability that the realization occurs in each of a set of nested budgeted uncertainty regions: $\mathcal{Z}(1)$, $\mathcal{Z}(2)$, $\mathcal{Z}(3)$, etc.

1.6 Robust Multi-Period Inventory Problem

In robust multi-period inventory problem (RMIP), the inventory manager's objective is to minimize the long term cost of inventory over a horizon of T periods. This long term cost might be composed for each period t of an ordering cost of c_t per unit, a fixed cost of K_t if an order is delivered at time t, a shortage cost of p_t per units of unsatisfied demand, and a holding cost h_t per unit held in storage. In each period, the ordered stocks are first used to satisfy the back-orders and then the current demand if possible. Any extra inventory is held until the next period after paying the associated holding cost. Unfortunately, since future demand is usually not fully determined at the time of making orders, one might require that orders are made such that the worst-case long term cost is as low as possible. This gives rise to the following robust optimization model

$$\begin{aligned}
&\underset{\boldsymbol{u},\boldsymbol{v}}{\text{minimize}} && \max_{\zeta\in\mathcal{Z}} \sum_{t=1}^{T} c_t u_t + K_t v_t + \max(h_t x_{t+1}(\boldsymbol{u},\zeta), -p_t x_{t+1}(\boldsymbol{u},\zeta)) && (1.18\text{a})\\
&\text{subject to} && 0 \le u_t \le M v_t, v_t \in \{0,1\} \;\; \forall t, && (1.18\text{b})
\end{aligned}$$

where $\boldsymbol{v} \in \{0,1\}^T$ and $\boldsymbol{u} \in \mathbb{R}^T$ represent respectively for each t the decision of making an order or not that will be delivered at time t, and the amount to be delivered, and where

$$x_{t+1}(\boldsymbol{u},\zeta) = x_1 + \sum_{j=1}^{t}(u_j - (\bar{w}_j + \hat{w}_j\zeta_j)) \;\; \forall t.$$

Problem (1.18) can be considered as a special case of problem (1.1) where $\boldsymbol{c}_{\zeta}^{t,k}(\boldsymbol{u},\boldsymbol{v}) = \alpha_{t,k}\sum_{j\le t} e_j\hat{w}_j$ and $d_{\zeta}^{t,k}(\boldsymbol{u},\boldsymbol{v}) = c_t u_t + K_t v_t - \alpha_{t,k}(x_1 + \sum_{j\le t}(u_j - \bar{w}_j))$ where $\alpha_{t,1} = -h_t$ and $\alpha_{t,2} = p_t$. We can therefore easily obtain a conservative approximation based on Proposition 1.4.1:

$$
\begin{aligned}
\text{(LP-RC)} \quad & \underset{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{\gamma},\boldsymbol{\Delta},\nu\\ \boldsymbol{\theta},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\psi}}}{\text{minimize}} \quad && \sum_{t=1}^{T}(c_t u_t + K_t v_t + \boldsymbol{\gamma}_t + \Delta_t) + \Gamma\nu && \text{(1.19a)}\\
& \text{subject to} && 0 \le u_t \le M v_t, \;\; v_t \in \{0,1\} \;\; \forall t && \text{(1.19b)}\\
& && \nu + \boldsymbol{\Delta} \ge \sum_{t=1}^{T} \boldsymbol{\lambda}_t^+ && \text{(1.19c)}\\
& && \nu + \boldsymbol{\Delta} \ge \sum_{t=1}^{T} \boldsymbol{\lambda}_t^- && \text{(1.19d)}\\
& \gamma_t \ge \mathbf{1}^T\boldsymbol{\psi}_{t,k} + \Gamma\theta_{t,k} - \alpha_{t,k}(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j)) \;\; \forall t, \; \forall k && && \text{(1.19e)}\\
& && (\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \ge -(\boldsymbol{\lambda}_t^+)_j + \alpha_{t,k}\hat{w}_j \;\; \forall t, \forall j \le t, \; \forall k && \text{(1.19f)}\\
& && (\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \ge -(\boldsymbol{\lambda}_t^-)_j - \alpha_{t,k}\hat{w}_j \;\; \forall t, \forall j \le t, \; \forall k && \text{(1.19g)}\\
& && (\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \ge -(\boldsymbol{\lambda}_t^+)_j \;\; \forall t, \forall j > t, \; \forall k = 1,2 && \text{(1.19h)}\\
& && (\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \ge -(\boldsymbol{\lambda}_t^-)_j \;\; \forall t, \forall j > t, \; \forall k = 1,2 && \text{(1.19i)}\\
& && \boldsymbol{\psi}_{t,k} \ge 0, \forall t, \forall j \le t, \; \forall k = 1,2 \;, && \text{(1.19j)}
\end{aligned}
$$

where $\boldsymbol{\gamma} \in \mathbb{R}^T$, $\boldsymbol{\Delta} \in \mathbb{R}^T$, $\nu \in \mathbb{R}$, $\boldsymbol{\theta} \in \mathbb{R}^{K\times T}$, $\boldsymbol{\lambda}_t^+ \in \mathbb{R}^T$, $\boldsymbol{\lambda}_t^- \in \mathbb{R}^T$, and $\boldsymbol{\psi}_{t,k} \in \mathbb{R}^T$. Note that the expression $c_t u_t + K_t v_t$ would initially appear in the fourth constraint based on model (1.5) but was carried to the objective function given that it is independent of k. Interestingly, based on Corollary 1.4.4, we have conditions under which this approximation scheme returns an optimal robust solution.

Corollary 1.6.1 *The conservative approximation model* (1.19) *is equivalent to problem* (1.18) *when* $\Gamma = 1$ *or* $\Gamma = T$*.*

Following the spirit of Theorem 3.2 in Bertsimas and Thiele (2006), one can also relate the solution of this approximation model to a solution that would be obtained for a specific sequence of deterministic orders.

Proposition 1.6.2 *(Optimal robust policy) Let* $\boldsymbol{\psi}^*$ *and* $\boldsymbol{\theta}^*$ *be optimal assignments in an optimal solution of problem* (1.19). *The optimal robust policy of the problem* (1.19) *is equivalent to the optimal policy of the deterministic version of problem* (1.18) *with demand set to* $w'_t = \bar{w}_t + \Upsilon_t - \Upsilon_{t-1}$ *where* $\Upsilon_0 = 0$ *and* $\Upsilon_t := (B_{t,2} - B_{t,1})/(h_t + p_t)$ *for* $B_{t,k} = \mathbf{1}^T \boldsymbol{\psi}^*_{t,k} + \Gamma\theta^*_{t,k}$.

Proof Proposition 1.6.2. Given an optimal solution tuple $(\boldsymbol{u}^*.\boldsymbol{v}^*, \boldsymbol{\gamma}^*, \boldsymbol{\Delta}^*, \nu^*, \boldsymbol{\theta}^*, \boldsymbol{\lambda}^{+*}, \boldsymbol{\lambda}^{-*}, \boldsymbol{\psi}^*)$ for problem (1.19), it is clear that $\boldsymbol{u}^*$, $\boldsymbol{v}^*$, and $\boldsymbol{\gamma}^*$ would also be the optimal solution of problem (1.19) if the remaining variable were fixed to $\boldsymbol{\Delta}^*$, ν^*, $\boldsymbol{\theta}^*$, $\boldsymbol{\lambda}^{+*}$, $\boldsymbol{\lambda}^{-*}$, and $\boldsymbol{\psi}^*$. Problem (1.19) is therefore equivalent to

$$\underset{u}{\text{minimize}} \quad \sum_{t=1}^{T}(c_t u_t + K_t 1_{\{u_t>0\}} + \max(h_t\bar{x}_{t+1} + B_{t,1}, -p_t\bar{x}_{t+1} + B_{t,2}) + \Delta^*_t) + \Gamma\nu^*, \quad (1.20)$$

where $\bar{x}_{t+1} = x_1 + \sum_{j\leq t}(u_j - \bar{w}_j)$, $B_{t,k} = \mathbf{1}^T\boldsymbol{\psi}^*_{t,k} + \Gamma\theta^*_{t,k}$, and where we use $1_{u_t>0}$ as the indicator function that returns one if u_t is strictly positive and zero otherwise. Let us define variable x'_t according to the linear equation $x'_{t+1} = \bar{x}_{t+1} + \frac{B_{t,1}-B_{t,2}}{h_t+p_t}$. This way we have that

$$\max(h_t\bar{x}_{t+1} + B_{t,1}, -p_t\bar{x}_{t+1} + B_{t2}) = \max(h_t x'_{t+1}, -p_t x'_{t+1}) + \frac{h_t B_{t,2} + p_t B_{t,1}}{h_t + p_t} \ ,$$

therefore, the problem (1.20) can be shown equivalent to

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \sum_{t=1}^{T}(c_t u_t + K_t 1_{u_t>0} + \max(h_t x'_{t+1}, -p_t x'_{t+1}) + \frac{h_t B_{t,2} + p_t B_{t,1}}{h_t + p_t} + \Delta^*_t) + \Gamma\nu^* \ .$$

Based on the equation $x'_{t+1} = x'_t + u_t - (\bar{w}_t + \Upsilon_t - \Upsilon_{t-1})$ where $\Upsilon_t := (B_{t,2} - B_{t,1})/(h_t + p_t)$, we can conclude that the optimal robust policy of problem (1.20) is equivalent to the optimal policy of nominal problem with demand $w'_t = \bar{w}_t + \Upsilon_t - \Upsilon_{t-1}$. ∎

Remark The optimal cost of the problem (1.19) is equal to the optimal cost for the nominal problem with the modified demand, w'_t, added to $\sum_{t=1}^{T}(\frac{h_t B_{2,t} + p_t B_{t,1}}{h_t+p_t} + \Delta^*_t) + \Gamma\nu$.

Remark This robust inventory problem was first addressed in Bertsimas and Thiele (2006), where the authors proposed a conservative approximation that relies on reversing the order of the maximization over ζ and summation over t. The resulting model, which we will later refer to as BT-RC can be reformulated as

$$
\begin{aligned}
\text{(BT-RC)} \quad \underset{\boldsymbol{u},\boldsymbol{v},\boldsymbol{y},\boldsymbol{q},\boldsymbol{r}}{\text{minimize}} \quad & \sum_{t=1}^{T}(c_t u_t + K_t v_t + y_t) \\
\text{subject to} \quad & y_t \geq h_t(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j) + q_t\Gamma + \sum_{j=1}^{t} r_{j,t}) \;\; \forall t \\
& y_t \geq p_t(-x_1 - \sum_{j=1}^{t}(u_j - \bar{w}_j) + q_t\Gamma + \sum_{j=1}^{t} r_{j,t}) \;\; \forall t \\
& q_t + r_{j,t} \geq \hat{w}_j \;\; \forall t, \; \forall j \leq t \\
& q_t \geq 0, \; r_{j,t} \geq 0 \;\; \forall t \; \forall j \leq t \\
& 0 \leq u_t \leq M v_t, \;\; v_t \in \{0,1\} \;\; \forall t \;,
\end{aligned}
$$

where $\boldsymbol{y} \in \mathbb{R}^T$, $\boldsymbol{q} \in \mathbb{R}^T$, and $\boldsymbol{r} \in \mathbb{R}^{T\times T}$. Note that to reduce the conservativeness of their approach, the authors use a different budget for each time period which we choose to omit doing for the sake of comparing similar models.

This model can actually be interpreted as an AARC of problem (1.18) where the affine decision rule is constrained to be constant over ζ. As was already recognized in Gorissen and den Hertog (2013), this indicates that their approximation can already be tightened by using affine decision rules, and, based on the results of Section 1.4.3, tightened even further by using problem (1.19) since the later is equivalent to applying AARC on a lifting involving ζ^+ and ζ^-. For completeness, we present AARC applied directly on ζ for this inventory problem:

$$
\begin{aligned}
\text{(AARC)} \quad \underset{\boldsymbol{u},\boldsymbol{v},\boldsymbol{\gamma},\boldsymbol{\Delta},\nu,\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \quad & \sum_{t=1}^{T}(c_t u_t + K_t v_t + \boldsymbol{\gamma}_t + \Delta_t) + \Gamma\nu \\
\text{subject to} \quad & 0 \leq u_t \leq M v_t, \;\; v_t \in \{0,1\} \;\; \forall t \\
& \nu + \Delta_j \geq \left| \sum_{t=1}^{T} \lambda_{j,t} \right|, \;\; \forall j
\end{aligned}
$$

$$\gamma_t \geq \sum_{j=1}^{T} \psi_{j,t,k} + \Gamma\theta_{t,k} - \alpha_{t,k}(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j)) \ \ \forall t, \ \forall k = 1, 2$$
$$\psi_{j,t,k} + \theta_{t,k} \geq |\lambda_{j,t} - \alpha_{t,k}\hat{w}_j| \ \ \forall t, \forall j \leq t, \ \forall k = 1, 2$$
$$\psi_{j,t,k} + \theta_{t,k} \geq |\lambda_{j,t}| \ \ \forall t, \forall j > t, \ \forall k = 1, 2$$
$$v \geq 0, \ \psi_{j,t,k} \geq 0, \ \ \theta_{t,k} \geq 0 \ \ \forall t, \forall j \leq t, \ \forall k = 1, 2,$$

where $\boldsymbol{\gamma} \in \mathbb{R}^T$, $\boldsymbol{\Delta} \in \mathbb{R}^T$, $\nu \in \mathbb{R}$, $\boldsymbol{\theta} \in \mathbb{R}^{T\times 2}$, $\boldsymbol{\lambda} \in \mathbb{R}^{T\times T}$, $\boldsymbol{\psi} \in \mathbb{R}^{T\times T\times 2}$, and $\alpha_{t,1} = -h_t \ \ \forall t$ and $\alpha_{t,2} = p_t \ \forall t$.

Remark One might notice that in this section we focused on an inventory problem where all ordering decisions must be made at time zero and there is no room for adjustment as time unfolds. While this might appear a bit limiting, our reasons for doing so are two folds. First, we believe this static version of the robust inventory problem is interesting in its own right based on the fact that in some contexts delivery contracts give no freedom to make adjustments to the orders as time evolves; even if there is some freedom, then the formulation studied in this section still gives a meaningful initial ordering plan that can later be improved on by solving an updated version of the model. Secondly, beside the special cases described in Bertsimas et al. (2010), very little is actually known about how to get exact solutions to the static or dynamic version of this robust model. Our hope is that by focusing on the static version of the problem we might understand what are the tools that can provide better near-optimal solutions.

1.7 Numerical Experiments

In this section, we present numerical experiments for the robust multi-period inventory problem discussed in Section 1.6. We initially present the performance of four different approximation methods for the instance that was studied in Bertsimas and Thiele (2006). This will illustrate how the worst-case bound can be gradually improved by using more computationally demanding models. In order from most tractable to most precise, we have the following list of formulation: the BT-RC model, the AARC model, the LP-RC model, a conservative approximate robust counterpart based on the SDP bound presented in (1.7) and referred as SDP-RC, and the exact robust model solved using a cutting-plane method[6] . In order to study to which extent these conclusions can be generalized, we later extend the numerical analysis to a set of randomly generated instances of the multi-period inventory problem where every parameters (*e.g.,* ordering cost, holding cost, amount of uncertainty, etc.) are non-stationary. In doing so, we also explore what is the "price of robustness" (as coined in Bertsimas and Sim (2004)) in this class of problems.

1.7.1 Instance Studied by Bertsimas and Thiele

The instance studied in Section 5.2 of Bertsimas and Thiele (2006) is an inventory problem with $T = 20$, $c_t = 1$, $K_t = 0$, $h_t = 4$, $p_t = 6$, $\bar{w}_t = 100$ and $\hat{w}_t = 40$. Note that this problem is stationary in the sense that the above parameters do not depend on t. Under this context, Table 1–2 presents the optimal worst-case bound obtained with each method, the true worst-case cost achieved by their respective approximate solution, and their respective suboptimality gap. As expected, when $\Gamma = 0$ all four methods give the same optimal bound and solution, which is the optimal solution of the nominal problem. The performance start to differ as Γ is increased. We can first confirm that, for any value of Γ, suboptimality gap is always reduced as we move away from the BT-RC model and use more sophisticated versions of our mixed-integer linear programming based approach. We can also confirm that, when Γ equals 1 or 20, the LP-RC and SDP-RC approach are exact, which was guaranteed by Theorem 1.6.1. This is also the case for AARC for this instance but will not be the case in general (see Table 1–3 for some evidence). In the case of $\Gamma = 10$, we see that using the SDP formulation allows us to reduce the suboptimality to a negligible amount, while it is slightly insufficient for $\Gamma = 15$. Although the suboptimality gap of all approximations methods are somewhat small in this example and the improvements obtained are rather limited, these results already illustrate the key differences between the

---

[6] The exact robust model is solved using an analytic center cutting-plane method up to a precision of 10^{-6} where cuts are generated by using CPLEX to solve the inner mixed-integer linear program.

four different approximation schemes. We expect these differences to be magnified in a richer experimental context.

Finally, it is worth noting that, although some of the solutions obtained are suboptimal, the worst-case cost achieved by a suboptimal solution often indicates exactly the approximate bound returned by the approximation model. For instance, one can observe in Table 2 that for the solutions provided by BT-RC, the worst-case bound approximations that BT-RC provides for $\Gamma \in \{0, 1, 10, 15, 20\}$ are exact. Yet, BT-RC does not return truly optimal solutions except when $\Gamma = 0$. One should therefore treat with care the fact that worst-case cost is equal to the worst-case bound for a solution that is returned by a conservative approximation scheme as there might actually exist other solutions that achieve better worst-case cost but for which the worst-case bound is very inaccurate.

Remark It is worth mentioning that Bertsimas and Thiele (2006) considered an uncertainty set that imposes multiple budgets on ζ, *i.e.,* where $\sum_{k=1}^{t} |\zeta_k|$ is limited to a budget Γ_t, $\forall\, t = 1, \ldots, T$. They proposed a method for calibrating the values for $[\Gamma_1, \ldots, \Gamma_T]$ in a way that makes the BT-RC model achieve its full potential. In particular, for the instance that is considered here, they established that these should be set as $[\Gamma_1, \Gamma_2, \ldots, \Gamma_{20}] =$ [0.51, 0.72, 0.88, 1.02, 1.14, 1.25, 1.35, 1.44, 1.53, 1.61, 1.69, 1.77, 1.79, ..., 1.79]. When we replicate the same experiments as described above using such an uncertainty set, we reach similar conclusions as before, namely that the quality of the bound is improved by employing the AARC model, and further improved using LRC. Specifically, the worst-case bound of BT-RC, AARC, and LP-RC are respectively of 7626, 7394, and 4244, while the worst-case cost of approximate solutions are respectively of 4582, 6434, and 4244 (with a loss in performance when going from BT-RC to AARC which is corrected for by the LP-RC solution). In comparison, the true optimal worst-case cost is of 3938.

1.7.2 Robust Performances on Randomly Generated Instances

In this second set of experiments, we consider a family of randomly generated instances of the robust multi-period inventory problem for a horizon of $T = 10$ and $T = 100$. Specifically,

Table 1–2: Comparison of the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution methods to the inventory problem presented in Bertsimas and Thiele (2006).

	Method	Budget of uncertainty				
		0	1	10	15	20
Worst-case bound	BT-RC	2000	5848	31840	39560	42480
	AARC	2000	5800	31457	39306	41818
	LP-RC	2000	5800	31360	38976	41818
	SDP-RC	2000	5800	31360	38940	41818
Worst-case cost	BT-RC	2000	5848	31840	39560	42480
	AARC	2000	5800	31457	39306	41818
	LP-RC	2000	5800	31360	38976	41818
	SDP-RC	2000	5800	31360	38940	41818
	Exact	2000	5800	31360	38933	41818
Sub-optim. gap	BT-RC	0	0.83%	1.53%	1.61%	1.58%
	AARC	0	0.00%	0.31%	1.10%	0
	LP-RC	0	0	0.00%	0.11%	0
	SDP-RC	0	0	0.00%	0.02%	0

each problem instance is created by randomly generating for each period the values for c_t, h_t, p_t based on a uniform distribution between 0 and 10, and the values for $\bar{w}_t$ and $\hat{w}_t$ from a uniform distribution over the interval $[0, 100]$ and $[0, \bar{w}_t]$ respectively. Note that the choice of using a uniform distribution can be motivated by the fact that it is known to be the distribution that maximizes entropy among all distributions supported on a given interval (see page 124 of Stone (2015) for additional details). Finally, fixed ordering cost K_t are again considered to be null.

Similarly as was done in the previous analysis, we compare in Table 1–3 the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution models, yet this time the table presents the average of each indicator over a set of 1000 randomly generated problem instances. A quick glance at the table should convince the reader that there are obvious gains in terms of worst-case cost for employing either the LP-RC or SDP-RC method instead of BT-RC or AARC. Note for instance that when $\Gamma = 6$, in these experiments the average worst-case cost was 2.2% larger with the AARC method compared to the LP-RC and the SDP-RC methods. More importantly, two valuable insights can be

Table 1–3: Comparison of the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution methods averaged over a set of 1000 randomly generated instances of multi-period inventory problems with 10 periods.

	Method	Budget of uncertainty						
		1	2	3	4	5	6	10
Average worst-case bound	BT-RC	4455	6050	7135	7871	8355	8657	8976
	AARC	3620	4861	5725	6321	6720	6972	7252
	LP-RC	3477	4597	5420	6031	6481	6798	7252
	SDP-RC	3477	4592	5412	6024	6476	6794	7252
Average worst-case cost	BT-RC	4049	5400	6363	7071	7581	7936	8438
	AARC	3619	4832	5677	6272	6681	6946	7252
	LP-RC	3477	4597	5420	6031	6481	6798	7252
	SDP-RC	3477	4591	5411	6023	6475	6794	7252
	True	3477	4585	5407	6020	6474	6794	7252
Average sub-optimality gap	BT-RC	16.6%	18.0%	18.0%	17.8%	17.5%	17.2%	16.8%
	AARC	4.4%	5.7%	5.2%	4.3%	3.3%	2.3%	0
	LP-RC	0	0.3%	0.3%	0.2%	0.1%	0.1%	0
	SDP-RC	0	0.1%	0.1%	0.1%	0.0%	0.0%	0

obtained from this table. First, we can estimate from this table that suboptimality of the approximate solution is reduced by a factor of about 4 when replacing the BT-RC method with AARC, by an additional factor of about 10 when replacing the AARC method with LP-RC, and by a final factor of at least 2 when replacing LP-RC with SDP-RC. Secondly, for general inventory problems it is uncommon for the worst-case bound obtained by any of these methods to be exactly equal to the true worst-case cost for the retrieved approximate solution. This can be observed by comparing the average worst-case bound to the average actual worst-case cost. The only exception is for LP-RC and SDP-RC when $\Gamma = 1$ and $\Gamma = 10$ as our theory previously established. Perhaps surprisingly, in the case of LP-RC, it also appears that for other integer Γ's it is quite rare that the worst-case bound is inexact at the proposed approximate solution, even when this solution is suboptimal.

Table 1–4 presents further statistics regarding the suboptimality of each method. Specifically, for each value of Γ studied, the table indicates for what proportion of random instances was each method able to recover a solution that was below either 0.0001%, 1%, and 10% of optimality. One can easily see that optimality is significantly improved by using the LP-RC

Table 1–4: Proportion of random instances (in 1000 trials) where methods achieved a set of target levels with respect to suboptimality for different values of Γ. The worst suboptimality gap attained by each method is also reported.

		Method			
Γ	Interval	BT-RC	AARC	LP-RC	SDP-RC
1	≤0.0001	0.0%	14.1%	100%	100%
	≤1	0.0%	26.1%	100%	100%
	≤10	19.9%	88.7%	100%	100%
	Maximum gap	56.7%	24.9%	0%	0%
3	≤0.0001	0.0%	0.4%	52.6%	56.6%
	≤1	0.2%	6.2%	90.4%	98.1%
	≤10	15.7%	91.1%	100.0%	100.0%
	Maximum gap	54.6%	23.0%	4.6%	2.0%
5	≤0.0001	0.0%	0.2%	57.3%	63.5%
	≤1	0.0%	9.3%	96.6%	99.70%
	≤10	16.5%	98.6%	100.0%	100.0%
	Maximum gap	52.2%	14.9%	2.6%	1.3%

or SDP-RC methods. In terms of worst suboptimality gap obtained over these instances, one can remark an improvement of a factor of 4, 6, and 2 for migrating from the BT-RC approximation to the AARC, and to the LP-RC and SDP-RC respectively when the budget is set to 5.

In practice, it is often the case that robust optimization, and especially with the budgeted uncertainty set, is used in a context where uncertain parameters are considered to be drawn from a distribution. *e.g.,* the distribution of historical values. For this reason, we next attempt to measure the inherent trade-off that can be observed between expected cost and value at risk when using different levels of budgets for uncertainty. Namely, given a problem instance, taking the shape of a set of parameters $\{c_t, h_t, p_t, \bar{w}_t, \hat{w}_t\}_{t=1}^T$, we assume that each interval $[\bar{w}_t - \hat{w}_t, \bar{w}_t + \hat{w}_t]$ describes the support of a uniform distribution for the random demand at time period t and that demand is independent between each period. We then compare the performance of the different robust solutions in terms of expected value and the 90th percentile of the total cost achieved when different values of Γ are used.

Figure 1–1 presents the average expected cost and average 90th percentile[7] achieved over 1000 problem instances by the five types of approximately robust solutions given different level of conservativeness expressed through Γ. We also present these same results in Figure 1–2 where the averaged (expected cost, value at risk) pair is plotted for each level of the budget, thus allowing us to identify the general structure of the Pareto frontier identified using each approximate models. These figures clearly present how the solution obtained from the nominal problem, when $\Gamma = 0$, can be improved both with respect to expected cost and to value at risk by considering a robust alternative. Although this behavior might seem surprising, it is in agreement with remark 3.2 of Delage et al. (2014) that claims that the mean value problem, where one replaces every random variables with its expected value, actually provides an optimistic solution (*i.e.,* solution based on best-case distribution) to stochastic programs when the objective function is convex with respect to uncertain parameters.

[7] Specifically, each statistic is estimated using 100 samples of random demand vector and averaged over 1000 problem instances.

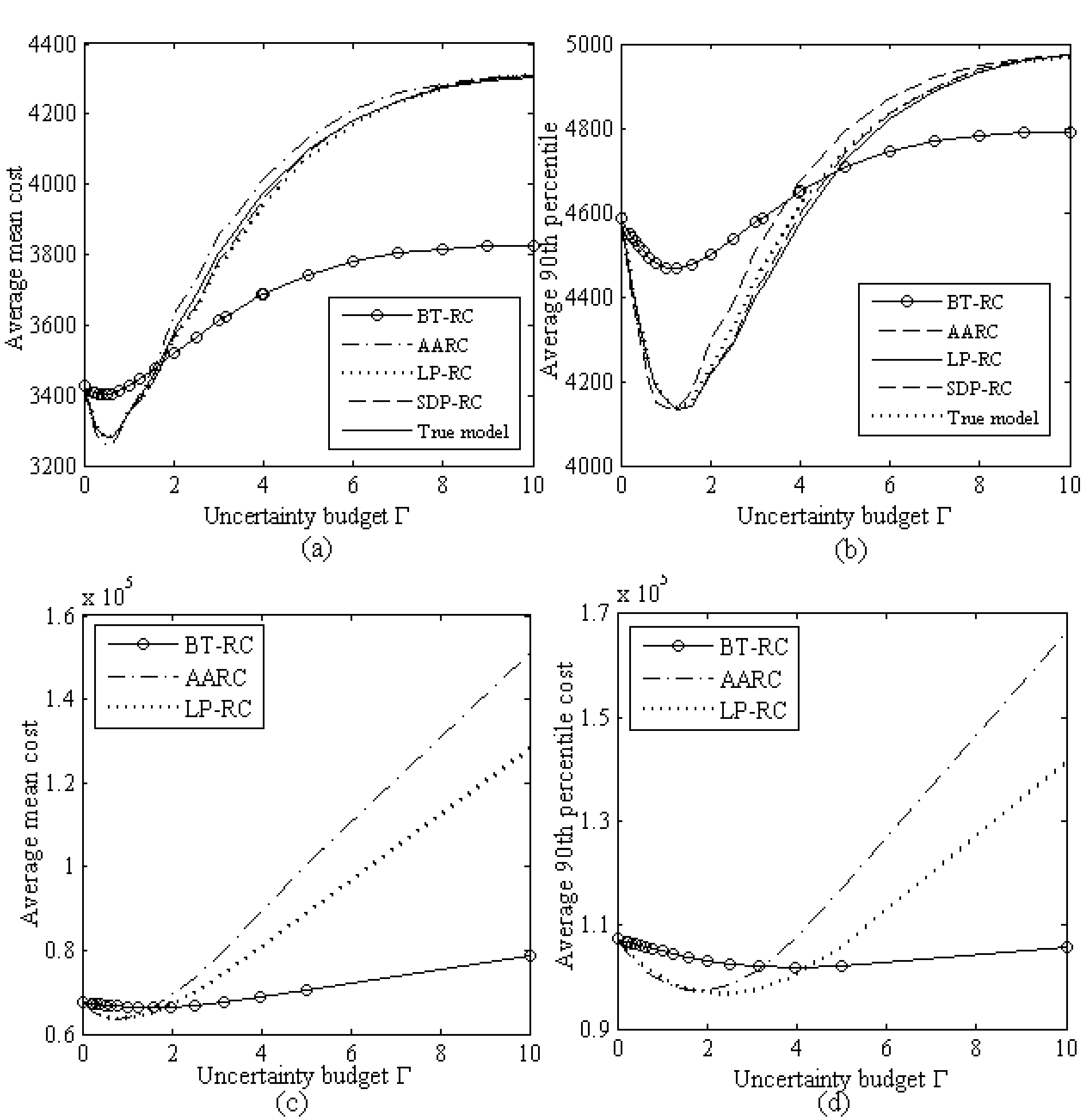

Figure 1–1: Average expected cost and average 90th percentile cost (over 1000 problem instances) achieved by robust solutions given different level of conservativeness. While Figure (a) presents average expected cost and Figure (b) presents average 90th percentile for a horizon of 10 periods, Figures (c) and (d) present the same statistics for a horizon of 100 periods.

For this family of problems, it also appears that there is a threshold above which the budget Γ leads to solutions that cannot be statistically motivated (*i.e.,* dominated in terms of both expected value and 90th percentile). This threshold appears to be respectively 1.5 and 3 in the case of problems with 10 periods and 100 periods. Based on Figures 1–1 (a) and (b), it also appears that although there is a lot to gain from using a more sophisticated model than BT-RC, the statistical performances of models that obtain the robust solution with greater precision than AARC are highly comparable for a short horizon. The difference between AARC and LP-RC is a bit more noticeable when the horizon is larger as portrayed in Figure 1–1 (c) and (d) where we see that the LP-RC dominates AARC for nearly all values of Γ. Note that in our experiments with $T = 100$, we omitted to include the performance of SDP-RC since it was too computationally demanding and since the performance seemed highly comparable to LP-RC.

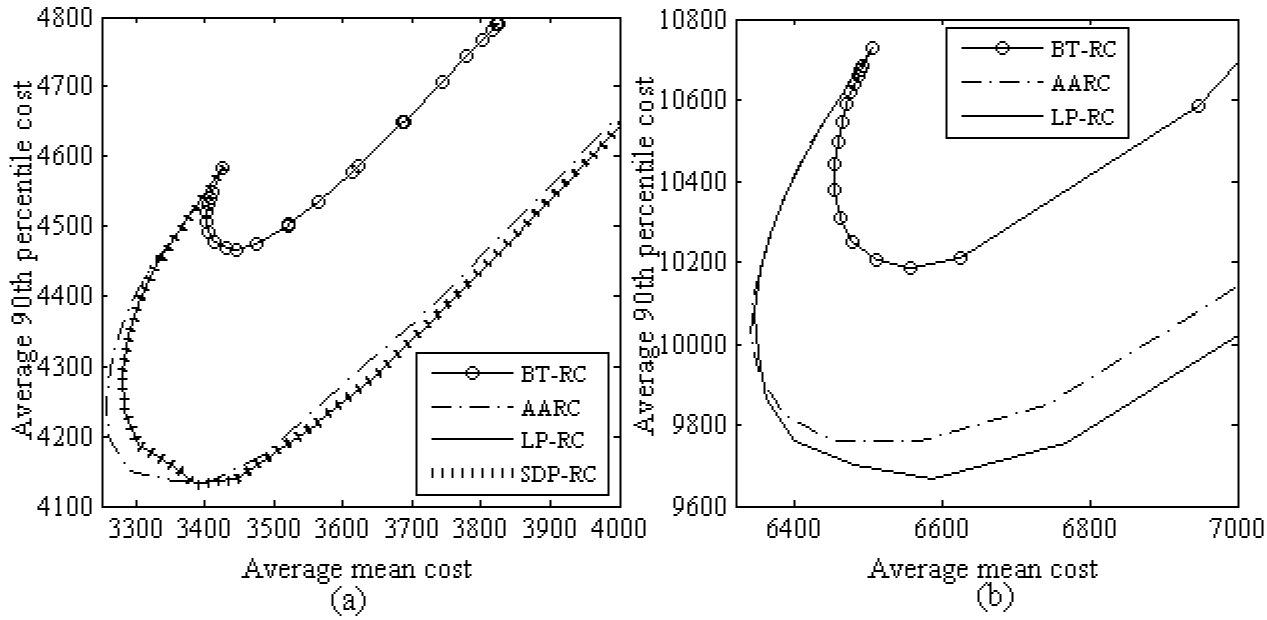

Figure 1–2: Average expected cost versus average 90th percentile achieved by the different robust methods when adjusting the level of conservativeness Γ. Figure (a) presents the achieved risk-return trade-offs for a 10 days horizon while (b) presents it for a 100 days horizon. Note that the two figures focus their attention on a region of the plane where the robust optimization methods are able to achieve their respective best risk-return trade-off. All methods (and especially AARC, LP-RC, and SDP-RC) propose robust solutions that are dominated by the deterministic solution when Γ becomes large.

1.8 Conclusion

In this article, we proposed a new scheme that can be used to generate conservative approximations of robust optimization problems involving the sum of piecewise linear functions and a polyhedral uncertainty set. This scheme exploits the fractional relaxation of a MILP known to be equivalent to the adversarial problem and can be used to identify two specific approximation models that respectively take the shape of a linear program and a SDP. While the linear programming model is shown to be equivalent to an application of AARC on a lifting of the parameter space, the SDP model clearly departs from previously known approximation techniques. Our approximation scheme also allows us to exploit the concept of total unimodularity to establish new conditions under which our LP-RC (and implicitly the AARC approach) model provides exact solutions. In particular, we identified the first exact reformulations for a robust (and distributionally robust) multi-item newsvendor problem with budgeted uncertainty set and a reformulation for robust multi-period inventory problems that is exact whether the uncertainty region reduces to a L_1-norm ball or to a box. An extensive set of empirical results finally illustrates the quality of the solutions obtained from different approximation schemes on randomly generated instances of the latter field of application.

Although very relevant to the discussion of this paper, we leave open the question of how to extend our approximation scheme to uncertainty set that are not polyhedral. In this regard, it appears that one might be able to reach interesting conclusions in contexts where the uncertainty set takes the shape of $\mathcal{Z} := \{\zeta \,|\, g_j(\zeta) \leq b_j \,\forall\, j = 1, 2, \ldots, J\}$ using a set of convex positive homogeneous $g_j(\cdot)$ functions, *i.e.,* functions for which $g_j(\alpha\zeta) = \alpha g_j(\zeta)$ for all $\alpha \geq 0$. This is for example the case when using an ellipsoidal set where $g_j(\zeta) := \|\zeta\|_2$. Under these conditions, the MILP representation of the adversarial problem studied in Section 1.4 will reduce to problem (1.4) with (1.4b) and (1.4g) replaced with

$$g_j(\zeta^+ - \zeta^-) \leq b_j \,,\, \forall\, j$$
$$g_j(\mathbf{\Delta}^+_{i,k} - \mathbf{\Delta}^-_{i,k}) \leq b_j z_{i,k} \,,\, \forall\, i,\, j,\, k \,,$$

This is because we can now exploit that for all i, j, and k, we have that

$$g_j(\mathbf{\Delta}^+_{i,k} - \mathbf{\Delta}^-_{i,k}) = g_j(z_{i,k}(\zeta^+ - \zeta^-)) = z_{i,k} g_j(\zeta^+ - \zeta^-) \leq b_j z_{i,k} \,.$$

Hence, the second constraint is necessarily redundant in the mixed integer program but will lead to a tighter conservative approximation after applying fractional relaxation.

1.9 Appendix

1.9.1 Proof of Proposition 1.4.3

We present this proof in four steps. We first introduce in Section 1.9.1.1 the notion of integral polytope and total unimodularity as these concepts are fundamental components of our proof. We then go through each of the three sets of conditions and present arguments for our conclusions in Sections 1.9.1.2, 1.9.1.3, and 1.9.1.4.

1.9.1.1 Integral Property of Polytope

Let the matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and vector $\boldsymbol{b} \in \mathbb{R}^m$ define the convex polytope $\{\boldsymbol{z} \in \mathbb{R}^n | \boldsymbol{z} \geq 0, \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}\}$.

Definition A matrix $\boldsymbol{A}$ is called totally unimodular if the determinant of every submatrix of $\boldsymbol{A}$ is equal to $+1$ or -1.

Lemma 1.9.1 *(See Truemper (1978) as cited in Grady and Polimeni (2010), page 313) Let the matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and vector $\boldsymbol{b} \in \mathbb{R}^m$ define the convex polytope $\{\boldsymbol{z} \in \mathbb{R}^n | \boldsymbol{z} \geq 0, \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}\}$, all the vertices of this convex polytope are integered valued if $\boldsymbol{A}$ is totally unimodular and $\boldsymbol{b}$ is integer valued.*

Hoffman and Kruskal (1956) proposed the following sufficient conditions for a matrix to be totally unimodular.

Lemma 1.9.2 *Let $\boldsymbol{A}$ be an $m \times n$ matrix containing only elements in the set $\{-1, 0, 1\}$, then $\boldsymbol{A}$ is a totally unimodular matrix if both of the following conditions are satisfied:*

1. *Each column of $\boldsymbol{A}$ contains at most two nonzero elements.*
2. *The rows of $\boldsymbol{A}$ can be partitioned into two sets $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ such that two nonzero entries in a column are in the same set of rows if they have different signs and in different sets of rows if they have the same sign.*

In what follows we will use the following result which we present right away as an example of application of Lemma 1.9.2.

Lemma 1.9.3 *The polytope defined by*

$$\zeta^+ \geq 0 \ \ \& \ \ \zeta^- \geq 0 \ \ \& \ \ \zeta_j^+ + \zeta_j^- \leq 1 \ \ , \forall j \tag{1.21a}$$

$$\mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \tag{1.21b}$$

$$\sum_{k=1}^{K} z_{i,k} = 1 \, , \forall i \tag{1.21c}$$

$$\sum_{k=1}^{K} \Delta_{i,k}^+ = \zeta_i^+ \ \ \& \ \ \sum_{k=1}^{K} \Delta_{i,k}^- = \zeta_i^- \ \ , \forall i \tag{1.21d}$$

$$\Delta_{i,k}^+ \geq 0 \ \ \& \ \ \Delta_{i,k}^- \geq 0 \ \ \& \ \ \Delta_{i,k}^+ + \Delta_{i,k}^- \leq z_{i,k} \, , \forall i, k \, , \tag{1.21e}$$

where $\zeta^+ \in \mathbb{R}^m$*,* $\zeta^- \in \mathbb{R}^m$*,* $\boldsymbol{z} \in \mathbb{R}^{m\times K}$*,* $\boldsymbol{\Delta}^+ \in \mathbb{R}^{m\times K}$ *and* $\boldsymbol{\Delta}^- \in \mathbb{R}^{m\times K}$*, has integer vertices.*

Proof First, let us realize that constraints (1.21c)-(1.21e) make constraint (1.21a) redundant. Hence, constraints (1.21a) to (1.21e) are reduced to:

$$\mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \tag{1.22a}$$

$$\sum_{k=1}^{K} z_{i,k} = 1 \, , \forall i \tag{1.22b}$$

$$\sum_{k=1}^{K} \Delta_{i,k}^+ - \zeta_i^+ = 0 \ \ \& \ \ \sum_{k=1}^{K} \Delta_{i,k}^- - \zeta_i^- = 0 \ \ , \forall i \tag{1.22c}$$

$$\Delta_{i,k}^+ + \Delta_{i,k}^- + s_{i,k} - z_{i,k} = 0 \ \ \forall i, \ k, \tag{1.22d}$$

$$\Delta_{i,k}^+ \geq 0 \ \ \& \ \ \Delta_{i,k}^- \geq 0 \ \ \& \ \ s_{i,k} \geq 0 \, , z_{ik} \geq 0 \ \ \forall i, \ k \ , \tag{1.22e}$$

where constraint (1.22d) is presented in standard form with some additional decision variables $s_{i,k} \in \mathbb{R}$ for all (i, k) pair.

In constraints (1.22a) to (1.22d), matrix of coefficients $\boldsymbol{A}$, contains only elements in the set {-1,0,1}. The columns of matrix $\boldsymbol{A}$ contains two nonzero elements, except for the columns associated with $s_{i,k}$ which only have one. In the columns associated to Δ_{ik}^+ and Δ_{ik}^-, these coefficients are equal to 1 because of constraints (1.22c) and (1.22d) while in columns associated with z_{ik}, each column hold a +1 and -1 coefficient based on constraints (1.22b) and (1.22d). Moreover, in columns associated with ζ_i^+ and ζ_i^-, each column hold a +1 and

-1 coefficient based on constraints (1.22a) and (1.22c). Condition 1 of Lemma 1.9.2 is easily satisfied. As of satisfying Condition 2, it is satisfied if matrix $\boldsymbol{A}$ is partitioned so that $\boldsymbol{A}_1$ contains the coefficients associated with constraints (1.22a) and (1.22c), and $\boldsymbol{A}_2$ contains entries of the other rows. $\blacksquare$

1.9.1.2 Condition 1

We start with the case of $\Gamma = 1$. First, one can show that constraints (1.4b) to (1.4i) reduce to the following:

$$
\begin{aligned}
&\zeta^+ \geq 0 \ \ \& \ \ \zeta^- \geq 0 \\
&\mathbf{1}^T(\zeta^+ + \zeta^-) = 1 \\
&\sum_{k=1}^{K} z_{i,k} = 1 \, , \forall \, i \\
&\sum_{k=1}^{K} \boldsymbol{\Delta}^+_{i,k} - \zeta^+ = 0 \ \ \& \ \ \sum_{k=1}^{K} \boldsymbol{\Delta}^-_{i,k} - \zeta^- = 0 \ \ , \forall \, i \\
&\boldsymbol{\Delta}^+_{i,k} \geq 0 \ \ \& \ \ \boldsymbol{\Delta}^-_{i,k} \geq 0 \, , \forall \, i, k \\
&z_{ik} - \sum_{j=1}^{m} (\boldsymbol{\Delta}^+_{i,k})_j - (\boldsymbol{\Delta}^-_{i,k})_j \geq 0 \, , \forall \, i, k \; .
\end{aligned}
$$

Together, these constraints further imply that $z_{i,k} = \sum_{j=1}^{m} (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j$ since if it was any larger, say by a positive amount $\boldsymbol{\Delta}_{i,k} > 0$ then the sum of $z_{i,k}$ would lead to

$$
\begin{aligned}
\sum_{k=1}^{K} z_{i,k} &= \sum_{k=1}^{K} \left(\sum_{j} (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j \right) + \boldsymbol{\Delta}_{i,k} \\
&= \sum_{j=1}^{m} (\zeta^+_j + \zeta^-_j) + \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = 1 + \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} > 1 \, ,
\end{aligned}
$$

which contradicts the fact that they should sum to one.

This allows us to say that there is always an optimal solution of the fractional relaxation of problem (1.4) that lies at one of the vertices of the polytope described by

$$
\sum_{j=1}^{m} \zeta^+_j + \zeta^-_j = 1
$$

$$
\begin{aligned}
&\sum_{k=1}^{K} \boldsymbol{\Delta}^{+}_{1,k} - \zeta^{+} = 0 \\
&\sum_{k=1}^{K} \boldsymbol{\Delta}^{+}_{i,k} - \sum_{k=1}^{K} \boldsymbol{\Delta}^{+}_{i-1,k} = 0 \;,\; \forall\, 2 \le i \le N \\
&\sum_{k=1}^{K} \boldsymbol{\Delta}^{-}_{1,k} - \zeta^{-} = 0 \\
&\sum_{k=1}^{K} \boldsymbol{\Delta}^{-}_{i,k} - \sum_{k=1}^{K} \boldsymbol{\Delta}^{-}_{i-1,k} = 0 \;,\; \forall\, 2 \le i \le N \\
&z_{i,k} - \sum_{j=1}^{m} (\boldsymbol{\Delta}^{+}_{i,k})_j + (\boldsymbol{\Delta}^{-}_{i,k})_j = 0, \;\; \forall i,\; k, \\
&\zeta^{+} \ge 0 \;\;\&\;\; \zeta^{-} \ge 0 \\
&\boldsymbol{\Delta}^{+}_{i,k} \ge 0 \;\;\&\;\; \boldsymbol{\Delta}^{-}_{i,k} \ge 0 \;,\; \forall\, i, k \;.
\end{aligned}
$$

One can easily show that the polytope defined by the first five constraints has integral vertices by confirming that, in the representation $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$, where $\boldsymbol{y}$ stands for the vector formed by appending all ζ^{+}, ζ^{-}, $\boldsymbol{\Delta}^{+}_{i,k}$ $\boldsymbol{\Delta}^{+}_{i,k}$ variables, the $\boldsymbol{A}$ matrix is totally unimodular. Indeed, total unimodularity is directly verified on the $\boldsymbol{A}$ matrix capturing the first five constraints. For each of these integer vertices, the projection $z_{i,k} := \sum_{j=1}^{m} (\boldsymbol{\Delta}^{+}_{i,k})_j + (\boldsymbol{\Delta}^{-}_{i,k})_j$ makes $z_{i,k}$ also an integer. This completes the proof that there always exists an optimal solution of the fractional relaxation of problem (1.4) that is integer thus the optimal value of this relaxation is exact.

1.9.1.3 Condition 2

For the second case, we will assume without loss of generality that $m = N$, *i.e.,* that there is one perturbation variable per convex sub-function in the objective function. We then observe that the description of the polytope reduces to the following set of constraints.

$$
\begin{aligned}
&\zeta^{+} \ge 0 \;\;\&\;\; \zeta^{-} \ge 0 \;\;\&\;\; \zeta^{+}_j + \zeta^{-}_j \le 1 \;,\; \forall\, j \\
&\sum_{k=1}^{K} z_{i,k} = 1 \;,\; \forall\, i
\end{aligned}
$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}^{+}_{i,k} = \zeta^{+} \;\;\&\;\; \sum_{k=1}^{K} \boldsymbol{\Delta}^{-}_{i,k} = \zeta^{-} \;\;, \forall\, i$$

$$\boldsymbol{\Delta}^{+}_{i,k} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}^{-}_{i,k} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}^{+}_{i,k} + \boldsymbol{\Delta}^{-}_{i,k} \leq z_{i,k} \;\;, \forall\, i,\, k\ .$$

Since an integer solution is feasible with respect to these constraints, we know that the fractional relaxation must necessarily achieve a higher value than problem (1.3). Yet, we now show that this relaxed problem is also upper bounded by problem (1.3) when $\boldsymbol{c}_{i,k}$ has the given structure.

We get the upper bounding problem by first relaxing the feasible set to the following:

$$\begin{aligned}
&\zeta^{+} \geq 0 \;\;\&\;\; \zeta^{-} \geq 0 \;\;\&\;\; \zeta^{+}_{j} + \zeta^{-}_{j} \leq 1 \;\;, \forall\, j \\
&\sum_{k=1}^{K} z_{i,k} = 1\,,\, \forall\, i \\
&\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = (\zeta^{+} - \zeta^{-}) \;\;, \forall\, i \\
&\|\boldsymbol{\Delta}_{i,k}\|_{\infty} \leq z_{i,k} \;\;, \forall\, i,\, k\ .
\end{aligned}$$

For any fixed ζ^{+} and ζ^{-}, the maximum over $\boldsymbol{\Delta}$ and $\boldsymbol{z}$ is upper bounded by its dual problem. Specifically, the optimal value of the problem

$$\underset{\boldsymbol{\Delta},\boldsymbol{z}}{\text{maximize}} \quad \sum_{i=1}^{N}\sum_{k=1}^{K} \boldsymbol{c}^{T}_{i,k}\boldsymbol{\Delta}_{i,k} + d_{i,k}z_{i,k} \tag{1.23a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} z_{i,k} = 1\,,\, \forall\, i \tag{1.23b}$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = \zeta^{+} - \zeta^{-} \;\;, \forall\, i \tag{1.23c}$$

$$\|\boldsymbol{\Delta}_{i,k}\|_{\infty} \leq z_{i,k} \;\;, \forall\, i,\, k \tag{1.23d}$$

is upper bounded by the optimal value of

$$\begin{aligned}
\underset{\boldsymbol{\gamma},\boldsymbol{\lambda}}{\text{minimize}} \quad & \sum_{i=1}^{N} \gamma_i + \boldsymbol{\lambda}^{T}_{i}(\zeta^{+} - \zeta^{-}) \\
\text{subject to} \quad & \gamma_i \geq d_{i,k} + \|\boldsymbol{c}_{i,k} - \boldsymbol{\lambda}_i\|_1 \;\;, \forall\, i,\, k\ ,
\end{aligned}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^N$ and $\boldsymbol{\lambda} \in \mathbb{R}^{N\times m}$ are the dual variables for constraint (1.23b) and (1.23c) respectively.

Based on the observation that reversing the order of $\max_\zeta$ and $\min_{\boldsymbol{\gamma},\boldsymbol{\lambda}}$ can only lead to a further upper bound, we are left with

$$\begin{aligned} \underset{\boldsymbol{\gamma},\boldsymbol{\lambda}}{\text{minimize}} \quad & \max_{\zeta:\|\zeta\|_\infty\leq 1} \sum_{i=1}^N \gamma_i + \boldsymbol{\lambda}_i^T\zeta && (1.24a)\\ \text{subject to} \quad & \gamma_i + \boldsymbol{\lambda}_i^T\zeta \geq d_{i,k} + \boldsymbol{c}_{i,k}^T\zeta \ , \ \forall \|\zeta\|_\infty \leq 1 \ , \forall i,k \ . && (1.24b)\end{aligned}$$

Following some recent results presented in Bertsimas et al. (2010), this last optimization model can be shown to be equivalent to problem (1.3) when $\boldsymbol{c}_{i,k}$ has the given structure. Specifically, given the structure of $\boldsymbol{c}_{i,k}$, problem (1.3) can be seen as evaluating the worst-case cost of linear dynamic system with bounded independent perturbations under a fixed policy

$$\begin{aligned} \underset{\zeta_1,\ldots,\zeta_N,x_1,\ldots,x_N}{\text{maximize}} \quad & \sum_{t=1}^N \max_k \alpha_{t,k}x_t + d_{t,k}\\ \text{subject to} \quad & x_1 = \beta_1\zeta_1\\ & x_t = x_{t-1} + \beta_t\zeta_t \ , \forall\, 2 \leq t \leq N\\ & -1 \leq \zeta_t \leq 1 \ , \forall\, t \ .\end{aligned}$$

In Theorem 3.1 of Bertsimas et al. (2010), the authors proved that an optimized affine running cost can be used instead of a convex one to measure exactly the total cost incurred by such a policy. Since problem (1.24) optimizes over all upper bounding affine running cost, it will necessarily achieve as optimal value the optimal value of problem (1.3). Hence, the fractional relaxation of problem (1.4) is tight.

1.9.1.4 Condition 3

For the third case, we can first realize that the objective function reduces to

$$\sum_{i=1}^N \sum_{k=1}^K \alpha_{i,k}((\boldsymbol{\Delta}_{i,k}^+)_i - (\boldsymbol{\Delta}_{i,k}^-)_i) + d_{i,k}z_{i,k} \ .$$

Hence, it is invariant to our choice for the variables $(\mathbf{\Delta}^+_{i,k})_j$ and $(\mathbf{\Delta}^+_{i,k})_j$ for all $j \neq i$. We are therefore interested in studying the vertices of the projection of the polytope defined by constraints (1.4b) to (1.4i) over the space spanned by the variables $(\mathbf{\Delta}^+_{i,k})_i$ and $(\mathbf{\Delta}^+_{i,k})_i$, $z_{i,k}$, ζ^+, and ζ^-.

When we remove from constraints (1.4b) to (1.4i), any constraint that involves $(\mathbf{\Delta}^+_{i,k})_j$ with $j \neq i$, we get the following set of constraints.

$$
\begin{aligned}
&\zeta^+ \geq 0 \;\;\&\;\; \zeta^- \geq 0 \;\;\&\;\; \zeta^+_j + \zeta^-_j \leq 1 \;\;, \forall j\\
&\mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma\\
&\sum_{k=1}^{K} z_{i,k} = 1\,, \forall i\\
&\sum_{k=1}^{K}(\mathbf{\Delta}^+_{i,k})_i = \zeta^+_i \;\;\&\;\; \sum_{k=1}^{K}(\mathbf{\Delta}^-_{i,k})_i = \zeta^-_i \;\;, \forall i\\
&(\mathbf{\Delta}^+_{i,k})_i \geq 0 \;\;\&\;\; (\mathbf{\Delta}^-_{i,k})_i \geq 0 \;\;\&\;\; (\mathbf{\Delta}^+_{i,k})_i + (\mathbf{\Delta}^-_{i,k})_i \leq z_{i,k} \;\;, \forall i,\, k\,.
\end{aligned}
$$

By construction, the polytope defined by these constraints must include the projection that we seek to define. An important property of this polytope is that it has integer vertices when Γ is an integer (see Lemma 1.9.3 above). Yet, when Γ is integer, it is also a subset of the projection that we are interested in. This can be confirmed by verifying that for any feasible solution of these constraints $(\zeta^+, \zeta^-, (\mathbf{\Delta}^+_{i,k})_i, (\mathbf{\Delta}^-_{i,k})_i, z_{i,k})$ one can create a solution $(\zeta^+, \zeta^-, \mathbf{\Delta}^+_{i,k}, \mathbf{\Delta}^-_{i,k}, z_{i,k})$ with $(\mathbf{\Delta}^+_{i,k})_j := z_{i,k}(\sum_{k'=1}^{K}(\mathbf{\Delta}^+_{j,k'})_j)$ and $(\mathbf{\Delta}^-_{i,k})_j := z_{i,k}(\sum_{k'=1}^{K}(\mathbf{\Delta}^-_{j,k'})_j)$ that is feasible according to constraints (1.4b) to (1.4i). Specifically, we have for all $i = 1, \ldots, N$ and $j = 1, \ldots, m$:

$$
\begin{aligned}
&\sum_{k=1}^{K}(\mathbf{\Delta}^+_{i,k})_j = \sum_{k=1}^{K} z_{i,k}\left(\sum_{k'=1}^{K}(\mathbf{\Delta}^+_{j,k'})_j\right) = \sum_{k'=1}^{K}(\mathbf{\Delta}^+_{j,k'})_j = \zeta^+_j\\
&(\mathbf{\Delta}^+_{i,k})_j + (\mathbf{\Delta}^-_{i,k})_j = z_{i,k}\left(\sum_{k'=1}^{K}(\mathbf{\Delta}^+_{j,k'})_j + (\mathbf{\Delta}^-_{j,k'})_j\right) \leq z_{i,k}\,.
\end{aligned}
$$

For constraints (1.4i), we verify the two cases. If for some (i', k'), $z_{i',k'} = 0$, then

$$\sum_{j=1}^{m} (\boldsymbol{\Delta}^{+}_{i',k'})_j + (\boldsymbol{\Delta}^{-}_{i',k'})_j = 0 \leq 0 \ .$$

Otherwise $x_{i',k'} = 1$ and, since the sum of $z_{i',k}$ over the k's is equal to one, it must be that $\sum_{k=1}^{K} (\boldsymbol{\Delta}^{+}_{i',k})_{i'} = (\boldsymbol{\Delta}^{+}_{i',k'})_{i'} = \zeta^{+}_{i'}$ and similarly for $(\boldsymbol{\Delta}^{-}_{i',k'})_{i'} = \zeta^{-}_{i'}$. Hence, we have that

$$\sum_{j=1}^{m} (\boldsymbol{\Delta}^{+}_{i',k'})_j + (\boldsymbol{\Delta}^{-}_{i',k'})_j = z_{i',k'} \sum_{j=1}^{m} \zeta^{+}_j + \zeta^{-}_j = \Gamma z_{i',k'} \ .$$

We are left to conclude that since the projected polytope has integer vertices, it must be that an optimal solution of problem (1.4) has integer vertices at least for the variables $(\zeta^{+}, \zeta^{-}, (\boldsymbol{\Delta}^{+}_{i,k})_i, (\boldsymbol{\Delta}^{-}_{i,k})_i, z_{i,k})$. Completing this part of the solution with the suggestion above will give an optimal solution of the problem that is completely integer.

1.9.2 Examples of Sums of Piecewise Linear Functions

We briefly summarize two additional examples of problems where sums of piecewise linear functions play an important role and for which robust optimization has the potential of identifying solutions that are immunized against model misspecification.

Example Multi-attribute utility theory: Consider a multi-objective linear program

$$\min_{\boldsymbol{x} \in \mathcal{X}} \ \left\{ \boldsymbol{c}_i^T \boldsymbol{x} + d_i \right\}_{i=1}^{N} \ ,$$

where $\boldsymbol{x} \in \mathbb{R}^n$, and for each i the affine mapping $\boldsymbol{c}_i^T \boldsymbol{x} + d_i$ computes an attribute that should be minimized: *e.g.,* total cost, delivery time, amount of carbon emitted, etc. Multi-attribute utility theory suggests that if the decision maker's preference relation satisfies additive independence then it can be represented by an additive utility function: i.e that he should solve a model of the type

$$\max_{\boldsymbol{x} \in \mathcal{X}} \ \sum_{i=1}^{N} u_i(\boldsymbol{c}_i^T \boldsymbol{x} + d_i) \ ,$$

where $u_i(\cdot)$ is a decreasing function, typically concave, and possibly piecewise linear. One might additionally consider that the value achieved for each objective is linearly influenced by a set of parameters ζ and wonder what would be a decision that maximizes the worst-case multi-attribute utility. Given that each $u_i(\cdot)$ is piecewise linear concave, this question reduces to solving problem (1.1) where $h_i(\boldsymbol{x}, \zeta) := -u_i(\boldsymbol{c}_i(\zeta)^T\boldsymbol{x} + d_i(\zeta)) = -\min_k \alpha_{i,k}(\boldsymbol{c}_i(\zeta)^T\boldsymbol{x} + d_i(\zeta)) + \beta_{i,k}$, with $\boldsymbol{c}_i$ and d_i defined as affine mappings of ζ.

Example Support vector machine: One of the most popular method for classification is known as the support vector machine, whereas one seeks a hyperplane that can separate as well as possible a set of instances $\{(\boldsymbol{\phi}_i, \kappa_i)\}_{i=1}^N$, where each $\boldsymbol{\phi}_i$ is a vector of features in $\mathbb{R}^n$ and $\kappa_i \in \{-1, 1\}$ is a label. In the soft margin method, this hyperplane is obtained by solving the following model:

$$\min_{\boldsymbol{x}, x_0} \ \|\boldsymbol{x}\|^2 + x_0^2 + \alpha \sum_{i=1}^N \max\{ \, 0 \, , \ 1 - \kappa_i(\boldsymbol{\phi}_i^T\boldsymbol{x} - x_0) \, \} \, ,$$

where $\boldsymbol{x} \in \mathbb{R}^n$ and $x_0 \in \mathbb{R}$. Considering that the data points $\boldsymbol{\phi}_i$ that are obtained could be the result of noisy measurements, it might be more appropriate to search for a robust hyperplane in the sense that it is optimal with respect to

$$\min_{\boldsymbol{x}, x_0} \ \|\boldsymbol{x}\|^2 + x_0^2 + \max_{\zeta \in \mathcal{Z}} \alpha \sum_{i=1}^N \max\{ \, 0 \, , \ 1 - \kappa_i(\boldsymbol{\phi}_i(\zeta)^T\boldsymbol{x} - x_0) \, \} \, ,$$

where the second term of the objective is a robust optimization term of the same form as in problem (1.1).

1.9.3 NP-hardness of Problem (1.2)

Proposition 1.9.4 *Evaluating the optimal value of problem* (1.2) *is NP-hard even when* $\mathcal{Z}$ *is polyhedral.*

Proof This result is obtained by showing that the NP-complete 3-SAT problem can be reduced to verifying whether

$$\max_{\zeta \in \mathcal{Z}} \ \sum_{i=1}^N \max_k \boldsymbol{c}_{i,k}^T\zeta + d_{i,k} \geq \gamma \, ,$$

is true or not.

3-SAT problem: Let W be a collection of disjunctive clauses $W = \{w_1, w_2, ..., w_N\}$ on a finite set of variables $V = \{v_1, v_2, ..., v_m\}$ such that $|w_i| = 3\ \forall i \in \{1, ..., N\}$. Let each clause be of the form $w = v_i \vee v_j \vee \bar{v}_k$, where $\bar{v}$ is the negation of v. Is there a truth assignment for V that satisfies all the clauses in W?

Given an instance of the 3-SAT problem, we can attempt to verify whether the optimal value of the following problem is larger or equal to N

$$\max_{\zeta} \quad \sum_{i=1}^{N} h_i(\zeta) \tag{1.25a}$$

$$\text{subject to} \quad 0 \leq \zeta \leq 1\ , \tag{1.25b}$$

where $\zeta \in \mathbb{R}^n$, and where $h_i(\zeta) := \max\{\zeta_{j_1}; \zeta_{j_2}; 1-\zeta_{j_3}\}$ if the i-th clause is $w_i = v_{j_1} \vee v_{j_2} \vee \bar{v}_{j_3}$. It is straightforward to confirm that $\{\zeta \in \mathbb{R}^n \,|\, 0 \leq \zeta \leq 1\}$ is a polyhedron and that each $h_i(\zeta)$ can be expressed as $h_i(\zeta) := \max_k \boldsymbol{c}_{i,k}^T \zeta + d_{i,k}$. More importantly, we have that the answer to the 3-SAT problem is positive if and only if the optimal value of problem (1.25) achieves an optimal value greater or equal to N. ∎

1.9.4 Proof of Proposition 1.4.5

[Proposition 1.4.5.] *Given a robust optimization problem of the form*

$$\underset{\boldsymbol{x} \in \mathcal{X}}{\text{minimize}} \quad \max_{\zeta \in \mathcal{U} \cap \mathcal{I}} h(\boldsymbol{x}, \zeta)\ ,$$

where $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}^m$ are both bounded convex sets, $\mathcal{I} = \{\zeta \in \mathbb{R}^m \,|\, \zeta_i \text{ is integer } \forall i \leq q\}$ for some $q \leq m$, hence imposing that a set of terms of ζ be integer valued, and $h(\boldsymbol{x}, \zeta)$ is real valued, convex in $\boldsymbol{x}$ and linear in ζ. Let $\hat{\boldsymbol{x}}$ be the solution of the conservative approximation

$$\underset{\boldsymbol{x} \in \mathcal{X}}{\text{minimize}} \quad \max_{\zeta \in \mathcal{U}} h(\boldsymbol{x}, \zeta)\ .$$

If there exists a $\hat{\zeta} \in \arg\max_{\zeta \in \mathcal{U}} \min_{\boldsymbol{x} \in \mathcal{X}} h(\boldsymbol{x}, \zeta)$ that is a member of the convex hull of $\mathcal{U} \cap \mathcal{I}$, denoted as $\mathcal{P}(\mathcal{U} \cap \mathcal{I})$, then $\hat{\boldsymbol{x}}$ is optimal according to the original robust optimization problem.

Proof Given that the stated conditions on $\mathcal{X}, \mathcal{U}$, and $h(\boldsymbol{x}, \zeta)$ are satisfied, by Sion's minimax theorem we have that

$$\min_{\boldsymbol{x}\in\mathcal{X}} \max_{\zeta\in\mathcal{U}} h(\boldsymbol{x},\zeta) \;=\; \max_{\zeta\in\mathcal{U}} \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\zeta) \; .$$

and thus that

$$\max_{\zeta\in\mathcal{U}} h(\hat{\boldsymbol{x}},\zeta) = \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\hat{\zeta}) \; .$$

In fact, Sion's minimax theorem further implies that

$$\min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\hat{\zeta}) \le h(\hat{\boldsymbol{x}},\hat{\zeta}) \le \max_{\zeta\in\mathcal{U}} h(\hat{\boldsymbol{x}},\zeta) = \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\hat{\zeta})$$

so that the saddle point property is satisfied:

$$\min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\hat{\zeta}) = h(\hat{\boldsymbol{x}},\hat{\zeta}) = \max_{\zeta\in\mathcal{U}} h(\hat{\boldsymbol{x}},\zeta) \; .$$

Now considering that $\hat{\zeta}$ is in the convex hull of $\mathcal{U}\cap\mathcal{I}$, which we refer to as $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$, we can show that $\hat{\boldsymbol{x}}$ achieves the optimal value in the original robust optimization problem:

$$\begin{aligned}
\min_{\boldsymbol{x}\in\mathcal{X}} \max_{\zeta\in\mathcal{U}\cap\mathcal{I}} h(\boldsymbol{x},\zeta) \;&=\; \min_{\boldsymbol{x}\in\mathcal{X}} \max_{\zeta\in\mathcal{P}(\mathcal{U}\cap\mathcal{I})} h(\boldsymbol{x},\zeta) \;\ge\; \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\hat{\zeta}) \\
&=\; h(\hat{\boldsymbol{x}},\hat{\zeta}) \;=\; \max_{\zeta\in\mathcal{U}} h(\hat{\boldsymbol{x}},\zeta) \\
&\ge\; \max_{\zeta\in\mathcal{U}\cap\mathcal{I}} h(\hat{\boldsymbol{x}},\zeta) \; ,
\end{aligned}$$

where the first equality comes from the fact that the maximum of a linear function over a set of points is achieved on the convex hull of those points. ∎

1.9.5 Proof of Proposition 1.4.9

[**Proposition 1.4.9.**] *The optimal value of the fractional relaxation of problem* (1.7) *is equal to the optimal value of the affinely adjustable robust counterpart of*

$$\begin{aligned}
\underset{\substack{\{\boldsymbol{v}_i\}_{i=1}^N,\boldsymbol{w}^+,\boldsymbol{w}^-,\\ \{\boldsymbol{Q}_i^+,\boldsymbol{V}_i^+,\boldsymbol{q}_i^+,\boldsymbol{S}_i^+,\boldsymbol{p}_i^+,r_i^+\}_{i=1}^N\\ \{\boldsymbol{Q}_i^-,\boldsymbol{V}_i^-,\boldsymbol{q}_i^-,\boldsymbol{S}_i^-,\boldsymbol{p}_i^-,r_i^-\}_{i=1}^N}}{\text{minimize}} \quad & \max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} \; (\boldsymbol{w}^+)^T\zeta^+ + (\boldsymbol{w}^-)^T\zeta^- + \textstyle\sum_{i=1}^N 2{\boldsymbol{p}_i^+}^T\zeta^+ + 2{\boldsymbol{p}_i^-}^T\zeta^- && (1.26a)\\
& \quad + \max_k\{\boldsymbol{c}_{i,k}^T(\zeta^+ - \zeta^-) + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + 2(\boldsymbol{V}_i^-)_{k,:}\zeta^- + (\boldsymbol{v}_i)_k\} \\
\text{subject to} \quad & (\boldsymbol{v}_i)_k = (\boldsymbol{Q}_i^+ + \boldsymbol{Q}_i^-)_{k,k} + 2(\boldsymbol{q}_i^+ + \boldsymbol{q}_i^-)_k + r_i^+ + r_i^- \; , \forall\, i,\, k && (1.26b)
\end{aligned}$$

$$\begin{bmatrix} \boldsymbol{Q}_i^+ & \boldsymbol{V}_i^+ & \boldsymbol{q}_i^+ \\ {\boldsymbol{V}_i^+}^T & \boldsymbol{S}_i^+ & \boldsymbol{p}_i^+ \\ {\boldsymbol{q}_i^+}^T & {\boldsymbol{p}_i^+}^T & r_i^+ \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \boldsymbol{Q}_i^- & \boldsymbol{V}_i^- & \boldsymbol{q}_i^- \\ {\boldsymbol{V}_i^-}^T & \boldsymbol{S}_i^- & \boldsymbol{p}_i^- \\ {\boldsymbol{q}_i^-}^T & {\boldsymbol{p}_i^-}^T & r_i^- \end{bmatrix} \succeq 0 \;, \forall\, i \tag{1.26c}$$

$$\sum_{i=1}^{N} \boldsymbol{S}_i^+ \leq \mathrm{diag}(\boldsymbol{w}^+) \;, \quad \sum_{i=1}^{N} \boldsymbol{S}_i^- \leq \mathrm{diag}(\boldsymbol{w}^-) \;, \forall\, i \tag{1.26d}$$

$$\boldsymbol{w}^+ \geq 0 \quad \boldsymbol{w}^- \geq 0 \;, \tag{1.26e}$$

where $\boldsymbol{w}^+ \in \mathbb{R}^m$*,* $\boldsymbol{w}^- \in \mathbb{R}^m$*, while for each* i*,* $\boldsymbol{v}_i \in \mathbb{R}^K$*,* $\boldsymbol{Q}_i^+ \in \mathbb{R}^{K\times K}$*,* $\boldsymbol{Q}_i^- \in \mathbb{R}^{K\times K}$*,* $\boldsymbol{V}_i^+ \in \mathbb{R}^{K\times m}$*,* $\boldsymbol{V}_i^- \in \mathbb{R}^{K\times m}$*,* $\boldsymbol{q}_i^+ \in \mathbb{R}^K$*,* $\boldsymbol{q}_i^- \in \mathbb{R}^K$*,* $\boldsymbol{S}_i^+ \in \mathbb{R}^{m\times m}$*,* $\boldsymbol{S}_i^- \in \mathbb{R}^{m\times m}$*,* $\boldsymbol{p}_i^+ \in \mathbb{R}^m$*,* $\boldsymbol{p}_i^- \in \mathbb{R}^m$*,* $r_i^+ \in \mathbb{R}$*,* $r_i^- \in \mathbb{R}$*, and finally where*

$$\mathcal{Z}' = \left\{ (\zeta^+, \zeta^-) \;\middle|\; \begin{array}{l} \zeta^+ \geq 0 \;,\; \zeta^- \geq 0 \\ \zeta_j^+ + \zeta_j^- \leq 1 \;\forall\, j \\ \sum_j \zeta_j^+ + \zeta_j^- = \Gamma \\ \boldsymbol{A}(\zeta^+ - \zeta^-) \leq \boldsymbol{b} \end{array} \right\} .$$

Proof We start by employing affine decision rules to produce a conservative approximation of problem (1.26). This leads to the following model.

$$\underset{\substack{\{\boldsymbol{v}_i, \gamma_i, \boldsymbol{\lambda}_i^+, \boldsymbol{\lambda}_i^-\}_{i=1}^N, \boldsymbol{w}^+, \boldsymbol{w}^-, \\ \{\boldsymbol{Q}_i^+, \boldsymbol{V}_i^+, \boldsymbol{q}_i^+, \boldsymbol{S}_i^+, \boldsymbol{p}_i^+, r_i^+\}_{i=1}^N \\ \{\boldsymbol{Q}_i^-, \boldsymbol{V}_i^-, \boldsymbol{q}_i^-, \boldsymbol{S}_i^-, \boldsymbol{p}_i^-, r_i^-\}_{i=1}^N}}{\text{minimize}} \quad \max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} \; (\boldsymbol{w}^+)^T\zeta^+ + (\boldsymbol{w}^-)^T\zeta^- + \sum_{i=1}^N 2{\boldsymbol{p}_i^+}^T\zeta^+ + 2{\boldsymbol{p}_i^-}^T\zeta^- + \gamma_i + {\boldsymbol{\lambda}_i^+}^T\zeta^+ + {\boldsymbol{\lambda}_i^-}^T\zeta^- \tag{1.27a}$$

subject to

$$\gamma_i + {\boldsymbol{\lambda}_i^+}^T\zeta^+ + {\boldsymbol{\lambda}_i^-}^T\zeta^- \geq \boldsymbol{c}_{i,k}^T(\zeta^+ - \zeta^-) + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + 2(\boldsymbol{V}_i^-)_{k,:}\zeta^- + (\boldsymbol{v}_i)_k \,\forall\, i, k, (\zeta^+,\zeta^-) \in \mathcal{Z}' \tag{1.27b}$$

$$(\boldsymbol{v}_i)_k = (\boldsymbol{Q}_i^+ + \boldsymbol{Q}_i^-)_{k,k} + 2(\boldsymbol{q}_i^+ + \boldsymbol{q}_i^-)_k + r_i^+ + r_i^- \;, \forall\, i,\, k \tag{1.27c}$$

$$\begin{bmatrix} \boldsymbol{Q}_i^+ & \boldsymbol{V}_i^+ & \boldsymbol{q}_i^+ \\ {\boldsymbol{V}_i^+}^T & \boldsymbol{S}_i^+ & \boldsymbol{p}_i^+ \\ {\boldsymbol{q}_i^+}^T & {\boldsymbol{p}_i^+}^T & r_i^+ \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \boldsymbol{Q}_i^- & \boldsymbol{V}_i^- & \boldsymbol{q}_i^- \\ {\boldsymbol{V}_i^-}^T & \boldsymbol{S}_i^- & \boldsymbol{p}_i^- \\ {\boldsymbol{q}_i^-}^T & {\boldsymbol{p}_i^-}^T & r_i^- \end{bmatrix} \succeq 0 \;, \forall\, i \tag{1.27d}$$

$$\sum_{i=1}^{N} \boldsymbol{S}_i^+ \leq \mathrm{diag}(\boldsymbol{w}^+) \;, \quad \sum_{i=1}^{N} \boldsymbol{S}_i^- \leq \mathrm{diag}(\boldsymbol{w}^-) \;, \forall\, i \tag{1.27e}$$

$$\boldsymbol{w}^+ \geq 0 \quad \boldsymbol{w}^- \geq 0 \;. \tag{1.27f}$$

Next, the variables γ_i can be replaced with $\gamma_i := \gamma_i' + r_i^+ + r_i^-$, while $\boldsymbol{v}_i$ can be replaced according to constraint (1.27c), in order to bring the terms r_i^+ and r_i^- in the objective function to obtain

$$
\begin{aligned}
\underset{\substack{\{\gamma_i',\boldsymbol{\lambda}_i^+,\boldsymbol{\lambda}_i^-\}_{i=1}^N,\boldsymbol{w}^+,\boldsymbol{w}^-,\\ \{\boldsymbol{Q}_i^+,\boldsymbol{V}_i^+,\boldsymbol{q}_i^+,\boldsymbol{S}_i^+,\boldsymbol{p}_i^+,r_i^+\}_{i=1}^N\\ \{\boldsymbol{Q}_i^-,\boldsymbol{V}_i^-,\boldsymbol{q}_i^-,\boldsymbol{S}_i^-,\boldsymbol{p}_i^-,r_i^-\}_{i=1}^N}}{\text{minimize}} \quad & \max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} \; (\boldsymbol{w}^+)^T\zeta^+ + (\boldsymbol{w}^-)^T\zeta^- + \textstyle\sum_{i=1}^N r_i^+ + r_i^- \\
& \quad + \textstyle\sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + 2\boldsymbol{p}_i^{-T}\zeta^- + \gamma_i' + \boldsymbol{\lambda}_i^{+T}\zeta^+ + \boldsymbol{\lambda}_i^{-T}\zeta^- \\
\text{subject to} \quad & \gamma_i' + \boldsymbol{\lambda}_i^{+T}\zeta^+ + \boldsymbol{\lambda}_i^{-T}\zeta^- \geq \boldsymbol{c}_{i,k}^T(\zeta^+ - \zeta^-) + d_{i,k} + 2(\boldsymbol{V}_i^+)_{k,:}\zeta^+ + 2(\boldsymbol{V}_i^-)_{k,:}\zeta^- \\
& \quad +(\boldsymbol{Q}_i^+ + \boldsymbol{Q}_i^-)_{k,k} + 2(\boldsymbol{q}_i^+ + \boldsymbol{q}_i^-)_k \; \forall\, i,k,(\zeta^+,\zeta^-)\in\mathcal{Z}' \\
& \begin{bmatrix} \boldsymbol{Q}_i^+ & \boldsymbol{V}_i^+ & \boldsymbol{q}_i^+ \\ \boldsymbol{V}_i^{+T} & \boldsymbol{S}_i^+ & \boldsymbol{p}_i^+ \\ \boldsymbol{q}_i^{+T} & \boldsymbol{p}_i^{+T} & r_i^+ \end{bmatrix} \succeq 0, \; \begin{bmatrix} \boldsymbol{Q}_i^- & \boldsymbol{V}_i^- & \boldsymbol{q}_i^- \\ \boldsymbol{V}_i^{-T} & \boldsymbol{S}_i^- & \boldsymbol{p}_i^- \\ \boldsymbol{q}_i^{-T} & \boldsymbol{p}_i^{-T} & r_i^- \end{bmatrix} \succeq 0 \;, \forall\, i \\
& \sum_{i=1}^N \boldsymbol{S}_i^+ \leq \operatorname{diag}(\boldsymbol{w}^+) \;, \; \sum_{i=1}^N \boldsymbol{S}_i^- \leq \operatorname{diag}(\boldsymbol{w}^-) \;, \forall\, i \\
& \boldsymbol{w}^+ \geq 0 \quad \boldsymbol{w}^- \geq 0 \;.
\end{aligned}
$$

We follow the derivation by applying Sion's minimax theorem to invert the order of the minimization and the maximization. Furthermore, since $\mathcal{Z}'$ is bounded and non-empty, constraint (1.27b) can be replaced using duality theory. This finally leads to the following model

$$
\begin{aligned}
\max_{(\zeta^+,\zeta^-)\in\mathcal{Z}'} \; \underset{\substack{\{\gamma_i',\boldsymbol{\lambda}_i^+,\boldsymbol{\lambda}_i^-,\boldsymbol{\psi}_i\}_{i=1}^N,\boldsymbol{w}^+,\boldsymbol{w}^-,\\ \{\boldsymbol{Q}_i^+,\boldsymbol{V}_i^+,\boldsymbol{q}_i^+,\boldsymbol{S}_i^+,\boldsymbol{p}_i^+,r_i^+\}_{i=1}^N\\ \{\boldsymbol{Q}_i^-,\boldsymbol{V}_i^-,\boldsymbol{q}_i^-,\boldsymbol{S}_i^-,\boldsymbol{p}_i^-,r_i^-\}_{i=1}^N}}{\min} \quad & (\boldsymbol{w}^+)^T\zeta^+ + (\boldsymbol{w}^-)^T\zeta^- + \textstyle\sum_{i=1}^N r_i^+ + r_i^- && (1.28a) \\
& \quad + \textstyle\sum_{i=1}^N 2\boldsymbol{p}_i^{+T}\zeta^+ + 2\boldsymbol{p}_i^{-T}\zeta^- + \gamma_i' + \boldsymbol{\lambda}_i^{+T}\zeta^+ + \boldsymbol{\lambda}_i^{-T}\zeta^- \\
\text{subject to} \quad & d_{i,k} - \gamma_i' + \bar{\boldsymbol{b}}^T\boldsymbol{\psi}_{ik} + (\boldsymbol{Q}_i^+ + \boldsymbol{Q}_i^-)_{k,k} + 2(\boldsymbol{q}_i^+ + \boldsymbol{q}_i^-)_k \geq 0, \forall\, i,k && (1.28b) \\
& \boldsymbol{c}_{i,k} - \boldsymbol{\lambda}_i^+ - \mathbb{A}^{+T}\boldsymbol{\psi}_{ik} + 2(\boldsymbol{V}_i^+)_{k,:}{}^T = 0 \,\forall\, i,k && (1.28c) \\
& -\boldsymbol{c}_{i,k} - \boldsymbol{\lambda}_i^- - \mathbb{A}^{-T}\boldsymbol{\psi}_{ik} + 2(\boldsymbol{V}_i^-)_{k,:}{}^T = 0 \,\forall\, i,k && (1.28d) \\
& \begin{bmatrix} \boldsymbol{Q}_i^+ & \boldsymbol{V}_i^+ & \boldsymbol{q}_i^+ \\ \boldsymbol{V}_i^{+T} & \boldsymbol{S}_i^+ & \boldsymbol{p}_i^+ \\ \boldsymbol{q}_i^{+T} & \boldsymbol{p}_i^{+T} & r_i^+ \end{bmatrix} \succeq 0, \; \begin{bmatrix} \boldsymbol{Q}_i^- & \boldsymbol{V}_i^- & \boldsymbol{q}_i^- \\ \boldsymbol{V}_i^{-T} & \boldsymbol{S}_i^- & \boldsymbol{p}_i^- \\ \boldsymbol{q}_i^{-T} & \boldsymbol{p}_i^{-T} & r_i^- \end{bmatrix} \succeq 0 \;, \forall\, i && (1.28e) \\
& \sum_{i=1}^N \boldsymbol{S}_i^+ \leq \operatorname{diag}(\boldsymbol{w}^+) \;, \forall\, i && (1.28f) \\
& \sum_{i=1}^N \boldsymbol{S}_i^- \leq \operatorname{diag}(\boldsymbol{w}^-) \;, \forall\, i && (1.28g)
\end{aligned}
$$

$$\boldsymbol{w}^+ \geq 0 \quad \boldsymbol{w}^- \geq 0\,,\, \boldsymbol{\psi}_{ik} \geq 0\,,\, \forall\, i, k\,, \tag{1.28h}$$

where $\boldsymbol{\psi}_{ik} \in \mathbb{R}^{(p+2+3m)}$ is the dual variable associated to the constraint $\mathbb{A}^+\zeta^+ + \mathbb{A}^-\zeta^- \leq \bar{b}$ with

$$\mathbb{A}^+ := \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{1}^T \\ -\boldsymbol{1}^T \\ \boldsymbol{I} \\ -\boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \qquad \mathbb{A}^- := \begin{bmatrix} -\boldsymbol{A} \\ \boldsymbol{1}^T \\ -\boldsymbol{1}^T \\ \boldsymbol{I} \\ \boldsymbol{0} \\ -\boldsymbol{I} \end{bmatrix} \qquad \bar{\boldsymbol{b}} := \begin{bmatrix} \boldsymbol{b} \\ \Gamma \\ -\Gamma \\ \boldsymbol{1} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}.$$

By applying duality theory one final time to replace the inner minimization problem we obtain the fractional relaxation of problem (1.7). In particular we have $z_{i,k}$, $\boldsymbol{\Delta}_i^+$, $\boldsymbol{\Delta}_i^-$, $\boldsymbol{\Lambda}^+$, and $\boldsymbol{\Lambda}^-$, as dual variable respectively for constraints (1.28b), (1.28c), (1.28d), (1.28f), and (1.28g). This completes the proof. ∎

1.9.6 Proof of Corollary 1.5.1

[**Corollary 1.5.1.**] *Given that* Γ *is a strictly positive integer, then the robust multi-item newsvendor problem* (1.15) *is equivalent to the following linear program:*

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathcal{X},\nu,\boldsymbol{\gamma},\boldsymbol{\psi}}{\text{minimize}} \quad & \Gamma\nu + \boldsymbol{1}^T\boldsymbol{\gamma} && (1.29a)\\
\text{subject to} \quad & \gamma_i \geq \psi_{i,k} + (\alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k}) \;\; \forall i,\ \forall k && (1.29b)\\
& \psi_{i,k} + \nu \geq \alpha^w_{i,k}\hat{w}_i \;\; \forall i,\ \forall k && (1.29c)\\
& \psi_{i,k} + \nu \geq -\alpha^w_{i,k}\hat{w}_i \;\; \forall i,\ \forall k && (1.29d)\\
& \psi_{i,k} \geq 0\ ,\ \forall i,\ \forall k\ , && (1.29e)
\end{aligned}$$

where $\nu \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^m$, *and* $\psi_{i,k} \in \mathbb{R}$.

Proof In the case of the robust newsvendor problem, the fractional relaxation of the adversarial problem (1.4) takes the form

$$\begin{aligned}
\underset{\boldsymbol{z},\zeta^+,\zeta^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-}{\text{maximize}} \quad & \sum_{i=1}^{m}\sum_{k=1}^{K} \alpha_{i,k}^{w}\hat{w}_i((\boldsymbol{\Delta}_{i,k}^{+})_i - (\boldsymbol{\Delta}_{i,k}^{-})_i) + (\alpha_{i,k}^{x}x_i + \alpha_{i,k}^{w}\bar{w}_i + \beta_{i,k})z_{i,k} \\
\text{subject to} \quad & \zeta^+ \geq 0 \;\;\&\;\; \zeta^- \geq 0 \;\;\&\;\; \zeta_j^+ + \zeta_j^- \leq 1 \;\;, \forall\, j \\
& \mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \\
& \sum_{k=1}^{K} z_{i,k} = 1\,, \forall\, i \\
& \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^{+} = \zeta^+ \;\;\&\;\; \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^{-} = \zeta^- \;\;, \forall\, i \\
& \boldsymbol{\Delta}_{i,k}^{+} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}_{i,k}^{-} \geq 0 \;\;\&\;\; \boldsymbol{\Delta}_{i,k}^{+} + \boldsymbol{\Delta}_{i,k}^{-} \leq z_{i,k} \;\; \forall i,\; \forall k \\
& \sum_{j} (\boldsymbol{\Delta}_{i,k}^{+})_j + (\boldsymbol{\Delta}_{i,k}^{-})_j = \Gamma z_{i,k} \;\; \forall i,\; \forall k\;.
\end{aligned}$$

Since for $i \neq j$, the variables $(\boldsymbol{\Delta}_{i,k}^{+})_j$ and $(\boldsymbol{\Delta}_{i,k}^{-})_j$ do not affect the objective value of the problem, we can use the argument presented in Section 1.9.1.4 to justify the following equivalent reformulation for this optimization problem.

$$\begin{aligned}
\underset{\boldsymbol{z},\zeta^+,\zeta^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-}{\text{maximize}} \quad & \sum_{i=1}^{m}\sum_{k=1}^{K} \alpha_{i,k}^{w}\hat{w}_i(\Delta_{i,k}^{+} - \Delta_{i,k}^{-}) + (\alpha_{i,k}^{x}x_i + \alpha_{i,k}^{w}\bar{w}_i + \beta_{i,k})z_{i,k} \\
\text{subject to} \quad & \zeta^+ \geq 0 \;\;\&\;\; \zeta^- \geq 0 \;\;\&\;\; \zeta_j^+ + \zeta_j^- \leq 1 \;\;, \forall\, j \\
& \mathbf{1}^T(\zeta^+ + \zeta^-) = \Gamma \\
& \sum_{k=1}^{K} z_{i,k} = 1\,, \forall\, i \\
& \sum_{k=1}^{K} \Delta_{i,k}^{+} = \zeta_i^+ \;\;\&\;\; \sum_{k=1}^{K} \Delta_{i,k}^{-} = \zeta_i^- \;\;, \forall\, i \\
& \Delta_{i,k}^{+} \geq 0 \;\;\&\;\; \Delta_{i,k}^{-} \geq 0 \;\;\&\;\; \Delta_{i,k}^{+} + \Delta_{i,k}^{-} \leq z_{i,k} \;\; \forall i,\; \forall k\;,
\end{aligned}$$

where $\Delta_{i,k}^{+} \in \mathbb{R}$ is now short for $(\boldsymbol{\Delta}_{i,k}^{+})_i$ and similarly for $\Delta_{i,k}^{-}$.

After replacing ζ_i^+ and ζ_i^- using the third equality constraint and dropping $\zeta_i^+ + \zeta_i^- \leq 1$ which is redundant since $\sum_{k=1}^{K} \Delta_{i,k}^{+} + \Delta_{i,k}^{-} \leq \sum_{k=1}^{K} z_{i,k} \leq 1$, we get the following reduced form:

$$\begin{aligned}
\underset{\boldsymbol{z},\zeta^+,\zeta^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-}{\text{maximize}} \quad & \sum_{i=1}^{m}\sum_{k=1}^{K} \alpha_{i,k}^{w}\hat{w}_i(\Delta_{i,k}^{+}-\Delta_{i,k}^{-}) + (\alpha_{i,k}^{x}x_i + \alpha_{i,k}^{w}\bar{w}_i + \beta_{i,k})z_{i,k} \\
& \sum_{i=1}^{m}\sum_{k=1}^{K} \Delta_{i,k}^{+} + \Delta_{i,k}^{-} = \Gamma \\
& \sum_{k=1}^{K} z_{i,k} = 1 \, , \, \forall i \\
& \Delta_{i,k}^{+} \geq 0 \;\; \& \;\; \Delta_{i,k}^{-} \geq 0 \;\; \& \;\; \Delta_{i,k}^{+} + \Delta_{i,k}^{-} \leq z_{i,k} \;\; \forall i, \; \forall k \; .
\end{aligned}$$

We are left with deriving the dual of this reduced form for the relaxed adversarial problem and re-introducing it back in the outer problem. Hence, we obtain

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathcal{X},\nu,\boldsymbol{\gamma},\boldsymbol{\psi}}{\text{minimize}} \quad & \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma} \\
\text{subject to} \quad & \gamma_i \geq \psi_{i,k} + (\alpha_{i,k}^{x}x_i + \alpha_{i,k}^{w}\bar{w}_i + \beta_{i,k}) \;\; \forall i, \; \forall k \\
& \psi_{i,k} + \nu \geq \alpha_{i,k}^{w}\hat{w}_i \;\; \forall i, \; \forall k \\
& \psi_{i,k} + \nu \geq -\alpha_{i,k}^{w}\hat{w}_i \;\; \forall i, \; \forall k \\
& \psi_{i,k} \geq 0 \; , \forall i, \; \forall k \; .
\end{aligned}$$

This optimization model can be simplified to problem (1.29). Furthermore, since the robust multi-item newsvendor problem presented in (1.15) satisfies Condition 3, we are guaranteed that this robust counterpart is exact by Corollary 1.4.4. ∎

1.9.7 Yet Another Exact Reformulation for a Distributionally Robust Multi-item Newsvendor Problem

Proposition 1.9.5 *Given a vector* $\boldsymbol{p} \in \mathbb{R}^m$ *such that* $0 \leq p_1 \leq p_2 \leq \cdots \leq p_m = 1$*, the distributionally robust optimization model*

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{F\in\mathcal{D}(\boldsymbol{p})} \; \mathbb{E}_F\left[\sum_{i=1}^{m} \max_{k} \alpha_{i,k}^{x}x_i + \alpha_{i,k}^{w}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k}\right] \; , \qquad (1.30)$$

where

$$\mathcal{D}(\boldsymbol{p}) = \{F \in \mathcal{M} | \mathbb{P}_F(\zeta \in \mathcal{Z}(i)) \geq p_i \;\; \forall i = 1, \ldots, m\} \; ,$$

is equivalent to the following linear program

$$\underset{\boldsymbol{x}\in\mathcal{X},t,\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\gamma},\boldsymbol{\psi}}{\text{minimize}} \quad t-\sum_{i=1}^{m} p_i\lambda_i \tag{1.31a}$$

$$\text{subject to} \quad t-\sum_{i=j}^{m}\lambda_i \geq j\nu_j+\sum_{i=1}^{m}\gamma_{i,j} \;\; \forall j=1,\ldots,m \tag{1.31b}$$

$$\gamma_{i,j}\geq \psi_{i,k,j}+\alpha^x_{i,k}x_i+\alpha^w_{i,k}\bar{w}_i+\beta_{i,k} \;\; \forall i,\ \forall k,\ \forall j=1,\ldots,m \tag{1.31c}$$

$$\psi_{i,k,j}+\nu_j\geq \alpha^w_{i,k}\hat{w}_i \;\; \forall i,\ \forall k,\ \forall j=1,\ldots,m \tag{1.31d}$$

$$\psi_{i,k,j}+\nu_j\geq -\alpha^w_{i,k}\hat{w}_i \;\; \forall i,\ \forall k,\ \forall j=1,\ldots,m \tag{1.31e}$$

$$\psi_{i,k,j}\geq 0\ ,\ \forall i,\ \forall k,\ \forall j=1,\ldots,m \tag{1.31f}$$

$$\boldsymbol{\lambda}\geq 0\ . \tag{1.31g}$$

where $\boldsymbol{\lambda}\in\mathbb{R}^m$, $\nu_j\in\mathbb{R}$, $\gamma_{i,j}\in\mathbb{R}$, *and* $\psi_{i,k,j}\in\mathbb{R}$.

Proof Following similar steps as in the proof of Theorem 1 in Wiesemann et al. (2014), one can apply duality theory for semi-infinite linear programs to the inner problem of the distributionally robust problem (1.30) and obtain the following reformulation:

$$\underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{\lambda}}{\text{minimize}} \quad t-\sum_{i=1}^{m-1} p_i\lambda_i \tag{1.32a}$$

$$\text{subject to} \quad \boldsymbol{\lambda}\geq 0 \tag{1.32b}$$

$$t\geq\left(\sum_{i=1}^{m}\max_k \alpha^x_{i,k}x_i+\alpha^w_{i,k}(\bar{w}_i+\hat{w}_i\zeta_i)+\beta_{i,k}\right)+\sum_{i=1}^{m-1}\lambda_i\mathbf{1}_{[\zeta\in\mathcal{Z}(i)]}\ \forall\zeta\in\mathcal{Z}(m)\ , \tag{1.32c}$$

where $\mathbf{1}_{[\zeta\in\mathcal{Z}(i)]}$ is the support function for the set $\mathcal{Z}(i)$, i.e.

$$\mathbf{1}_{[\zeta\in\mathcal{Z}(i)]}:=\begin{cases}1 & \text{if } \zeta\in\mathcal{Z}(i)\\ 0 & \text{otherwise.}\end{cases}$$

Yet, since the sets in $\{\mathcal{Z}(i)\}_{i=1}^m$ are nested such that $\mathcal{Z}(i)\subset\mathcal{Z}(i+1)$ for all i, one can argue that constraint (1.32c) is equivalent to

$$t\geq\left(\sum_{i=1}^{m}\max_k \alpha^x_{i,k}x_i+\alpha^w_{i,k}(\bar{w}_i+\hat{w}_i\zeta_i)+\beta_{i,k}\right)+\sum_{i=j}^{m-1}\lambda_i\ \ \forall\zeta\in\mathcal{Z}(j)\setminus\mathcal{Z}(j-1),\ \forall j=1,\ldots,m\ , \tag{1.33}$$

where for simplicity of exposure it is assumed that $\mathcal{Z}(1) \setminus \mathcal{Z}(0)$ stands for $\mathcal{Z}(1)$. Now since for each j, it is clear that satisfying the two conditions

$$\begin{aligned} t &\geq \left(\sum_{i=1}^{m} \max_k \alpha_{i,k}^x x_i + \alpha_{i,k}^w (\bar{w}_i + \hat{w}_i \zeta_i) + \beta_{i,k} \right) + \sum_{i=j}^{m-1} \lambda_i \;\; \forall \zeta \in \mathcal{Z}(j) \\ t &\geq \left(\sum_{i=1}^{m} \max_k \alpha_{i,k}^x x_i + \alpha_{i,k}^w (\bar{w}_i + \hat{w}_i \zeta_i) + \beta_{i,k} \right) + \sum_{i=j+1}^{m-1} \lambda_i \;\; \forall \zeta \in \mathcal{Z}(j+1) \setminus \mathcal{Z}(j) \end{aligned}$$

implies that

$$t \geq \left(\sum_{i=1}^{m} \max_k \alpha_{i,k}^x x_i + \alpha_{i,k}^w (\bar{w}_i + \hat{w}_i \zeta_i) + \beta_{i,k} \right) + \sum_{i=j+1}^{m-1} \lambda_i \;\; \forall \zeta \in \mathcal{Z}(j+1)$$

is also satisfied due to the fact that $\lambda_j \geq 0$, we can therefore by induction starting from $j = 1$ to $j = m - 1$ demonstrate that any ζ that is feasible according to constraint (1.32c) is also feasible according to

$$t \;\; \geq \;\; \left(\sum_{i=1}^{m} \max_k \alpha_{i,k}^x x_i + \alpha_{i,k}^w (\bar{w}_i + \hat{w}_i \zeta_i) + \beta_{i,k} \right) + \sum_{i=j}^{m-1} \lambda_i \;\; \forall \zeta \in \mathcal{Z}(j), \; \forall j = 1, \ldots, m \, . \tag{1.34}$$

Furthermore, the reverse is also true given that the list of constraints in (1.34) includes the list in (1.33) as a subset. Hence, constraint (1.34) is equivalent to constraint (1.32c). Finally, after observing that each of the robust constraints indexed by j in (1.34) is an instance of the robust objective presented in Corollary 1.5.1, one can employ the reformulation that is proposed in the corollary to obtain problem (1.31). ∎

Chapter 2
The Value of Flexibility in Robust Location-Transportation Problems[1]

Amir Ardestani-Jaafari, Erick Delage

Department of Decision Sciences, HEC Montréal, Montréal, Quebec, H3T 2A7, Canada

amir.ardestani-jaafari@hec.ca, erick.delage@hec.ca

Abstract

This article studies a capacitated fixed-charge multi-period location-transportation problem in which, while the location and capacity of each facility must be determined immediately, the determination of the final production and distribution of products can be delayed until actual orders are received in each period. In contexts where little is known about future demand, robust optimization, namely using a budgeted uncertainty set, becomes a natural method for identifying meaningful decisions. Unfortunately, it is well known that these types of multi-period robust decision problems are computationally intractable. To overcome this difficulty, we propose a set of tractable conservative approximations for the problem that each exploits to a different extent the idea of reducing the flexibility of the delayed decisions. While all of these approximation models outperform previous approximation models that have been proposed for this problem, each also has the potential to reach a different level of compromise between efficiency of resolution and quality of the solution. A row generation algorithm is also presented in order to address problem instances of realistic size. We also demonstrate that full flexibility is often unnecessary to reach nearly, or even exact, optimal robust locations and capacities for the facilities. Finally, we illustrate our findings

---

[1] This article is accepted for publication by "Transportation Science".

with an extensive numerical study where we evaluate the effect of the amount of uncertainty on the performance and structure of each approximate solution that can be obtained.

Keywords

transportation, facility location, robust optimization, flexibility, conservative approximation, demand uncertainty.

2.1 Introduction

Transportation planning can be decomposed into three different levels (Crainic and Laporte 1997): strategic transportation planning, tactical transportation planning, and operational transportation planning. At the highest level of management, an important decision is determining the geographical locations of factories, suppliers and warehouses. Determination of facility location, such as hub locations, supplier locations, air freight hub locations, railway station locations, etc., can significantly impact the design of the strategic networks. Recognizing this fact, researchers (*e.g.,* Christensen et al. (2013) and Abouee-Mehrizi et al. (2014)) have been developing integrated models in order to have better control on the interactions between facility location decisions and transportation strategies.

The traditional way of describing the location-transportation problem (LTP) has been to assume a deterministic environment. In a deterministic setting, *i.e.,* when there is no uncertainty about problem data, a multi-period capacitated fixed-charge LTP, with L facility locations, N customer locations, and T periods, can take the form of the following mixed-integer linear program (MILP):

$$
\begin{aligned}
\text{(Deterministic)} \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{Y},\boldsymbol{P}}{\text{maximize}} \quad & \sum_{t=1}^{T}\sum_{i=1}^{L}\sum_{j=1}^{N}(\eta - d_{ij})Y_{ij}^t - \boldsymbol{c}^T\boldsymbol{P}^t - ({\boldsymbol{c}_0}^T\boldsymbol{Z} + \boldsymbol{f}^T\boldsymbol{I}) && \text{(2.1a)}\\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq \zeta_j^t\,,\ \forall j \in \{1,2,\ldots,N\}\,,\ \forall t \in \{1,2,\ldots,T\} && \text{(2.1b)}\\
& \sum_j Y_{ij}^t \leq P_i^t\,,\ \forall i \in \{1,2,\ldots,L\}\,,\ \forall t \in \{1,2,\ldots,T\} && \text{(2.1c)}\\
& \boldsymbol{P}^t \leq \boldsymbol{Z}\,,\ \forall t \in \{1,2,\ldots,T\} && \text{(2.1d)}\\
& \boldsymbol{Y}^t \geq 0\,,\ \forall t \in \{1,2,\ldots,T\} && \text{(2.1e)}
\end{aligned}
$$

$$\boldsymbol{Z} \leq M\boldsymbol{I}\,,\ \boldsymbol{I} \in \{0,1\}^L\,, \tag{2.1f}$$

where $\boldsymbol{Z} \in \mathbb{R}^L$, $\boldsymbol{Y} \in \mathbb{R}^{L\times N\times T}$,$\boldsymbol{P} \in \mathbb{R}^{L\times T}$, and with M as a constant chosen large enough. This MILP integrates the optimization of both "strategic" and "operational" decisions. At the strategic level, it includes for each candidate location $i = 1, 2, \ldots, L$, the binary decision I_i denoting whether a facility should be opened or not, and the continuous decision Z_i denoting the production capacity of the facility. Once these are decided upon, operational decisions over a horizon of $t = 1, 2, \ldots, T$ include for each period t, P_i^t denoting how many goods are produced at each i-th facility and Y_{ij}^t denoting how many goods are shipped from facility i to customers at location j. The demand during period t for location $j = 1, 2, \ldots, N$ is characterized by ζ_j^t. The total profit generated by the company is computed on the basis of the following: sales revenue (with $\eta > 0$ for the unit price of goods); construction costs (calculated for a given facility i, with a size Z_i, a fixed cost K_i, and variable costs C_iZ_i); production costs c_i for each facility i; and, finally, transportation costs, with d_{ij} being the unit cost for any shipment from location i to j. Note that each parameter η, d_{ij}, and c_i could alternatively be considered time dependant.[2] [3]

In model (2.1), all parameters are considered to be known exactly at the time of making the strategic decision. In practice, however, some parameters, such as the exact size of each demand ζ_j^t,[4] are unknown at the time the facilities are built. In recent years, studies made in a number of field of applications (Bertsimas et al. (2011a), Gabrel et al. (2014b))

[2] Note that in this chapter we follow the notation of Baron et al. (2011). In particular we denote decision variables with upper case letters and parameters with lower case letters. We also distinguish vectors and matrices from scalar values by using a bold font. The distinction between vectors and matrices should be clear based on the context.

[3] It is worth noting that the location-transportation problem that is studied in this article does not consider the possibility of holding inventory from one period to the other. Although it can be considered as a limitation of the model, as mentioned in Baron et al. (2011) this model is perfectly adequate in contexts that involve a make-to order firm as well as in a high volume just-in-time production environment.

[4] While sources of uncertainty other than demand might affect the performance of facility location decisions and it might be interesting to account for them, in this paper, we focus on demand uncertainty as we

have demonstrated the effectiveness of robust optimization (RO) for handling uncertainty, especially in cases where there is no valid argument to justify the choice of a distribution model. A naïve application of robust optimization to LTP under demand uncertainty might lead to the following robust counterpart (RC):

$$
\begin{aligned}
\text{(RC)} \quad & \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{Y},\boldsymbol{P}}{\text{maximize}} \quad && \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - \boldsymbol{c}^T \boldsymbol{P}^t - ({\boldsymbol{c}_0}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) && \text{(2.2a)} \\
& \text{subject to} && \sum_i Y_{ij}^t \leq \zeta_j^t \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D},\, \forall j \,,\, \forall t && \text{(2.2b)} \\
& && \text{(2.1c)} - \text{(2.1f)}\,,
\end{aligned}
$$

where $\mathcal{D}$ is the uncertainty set for the vector composed of all the demands $(\zeta^1, \zeta^2, \ldots, \zeta^T)$.

Although it can be shown that the RC model can be reformulated as a MILP if $\mathcal{D}$ is polyhedral, the solution it provides will often appear overly conservative, *i.e.,* it might suggest opening only a few facilities (if any at all) with very limited capacity. This is actually due to the fact that the RC model completely disregards how operational decisions, namely, the size of production and deliveries, are delayed and can exploit information that becomes available about the demand. This motivates the use of the following multi-period robust location-transportation problem (MRLTP) model:

$$
\begin{aligned}
\text{(MRLTP)} \quad & \underset{\boldsymbol{I},\boldsymbol{Z}}{\text{maximize}} \quad && \min_{\boldsymbol{\zeta} \in \mathcal{D}} \sum_t h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t) - ({\boldsymbol{c}_0}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) && \text{(2.3a)} \\
& \text{subject to} && \boldsymbol{Z} \leq M \boldsymbol{I} \,,\, \boldsymbol{I} \in \{0,1\}^L \,, && \text{(2.3b)}
\end{aligned}
$$

expect it to have the most impact on the quality of the decision that needs to be made. See for instance Delage et al. (2014), where the authors argue that simply using the expected values of parameters that appear in the objective function already generates solutions that can be considered robust for such multi-period problems.

where $h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t)$ is the profit generated during period t, once the demand is revealed for this period, and is defined as

$$
\begin{aligned}
h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t) = \max_{\boldsymbol{Y}^t, \boldsymbol{P}^t} \quad & \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - \boldsymbol{c}^T \boldsymbol{P}^t && \text{(2.4a)} \\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq \zeta_j^t \,, \forall\, j && \text{(2.4b)} \\
& \sum_j Y_{ij}^t \leq P_i^t \,, \forall\, i && \text{(2.4c)} \\
& \boldsymbol{P}^t \leq \boldsymbol{Z} && \text{(2.4d)} \\
& \boldsymbol{Y}^t \geq 0 \,, && \text{(2.4e)}
\end{aligned}
$$

which, in particular, captures the fact that, since it is assumed that goods cannot be stored (or demand backlogged) from one period to the other, neither at the facility nor at the demand locations, it is always possible to design an optimal transportation and production plan that depends only on the currently realized demand.

Finally, we make the common assumption that the demand vector ζ is known to lie in a budgeted uncertainty set (see Bertsimas and Sim (2004)), *i.e.,* that each ζ_i lies in an interval, and that, at most, Γ of the terms across all locations and time periods can take extreme values.

While it appears that the MRLTP does implement as much flexibility as is needed in this problem, Atamtürk and Zhang (2007) established that evaluating the objective is already computationally intractable when $T = 1$. In this paper, we present a set of six conservative approximation models to the problem that each exploits to a different extent the idea of reducing the flexibility of the delayed decisions. These models will allow us to empirically explore the compromises that need to be made between flexibility/conservatism and "tractability".[5] Overall, we consider this article to make the following contributions:

[5] Note that, in this paper, we will consider a model to be tractable if it can be reformulated as a mixed-integer linear program of finite dimension.

1. We present a set of tractable conservative approximations of the MRLTP that each employ different form of the application of affine adjustments proposed in Ben-Tal et al. (2004) and Chen and Zhang (2009). While we do demonstrate empirically for the first time how significant the improvements can be in terms of the quality of the approximate solutions for the MRLTP, especially as compared to the robust model of Baron et al. (2011), we also establish conditions under which some of the simplest approximation schemes already provide optimal solutions. These theoretical results rely on carefully adapting the arguments presented in Ben-Tal et al. (2004) and Bertsimas and Goyal (2012) to our multi-period setting.
2. Two of our formulations, namely, those that will be referred to ELAARC and HD-ELAARC, also provide valuable insights about how better conservative approximation models can be obtained in robust multi-stage optimization problems. With ELAARC, this is done by creating the affine adjustments only after the recourse problem has been replaced by an equivalent penalized formulation. With HD-ELAARC, this is done by letting the affine adjustments depend on the whole history, even though an optimal recourse policy is known to be independent of the history. These two ideas might serve many other instances of robust multi-stage decision problems.
3. We propose a row generation algorithm that employs a parsimonious choice of valid inequalities in order to accelerate the resolution of one of our most complex approximation models, while being easily adaptable to any of our other formulations. Our implementation of this algorithm allows us to reduce the solution time of larger instances by a factor of 16 to 260, and to solve instances with 20 periods, 15 facility locations, and 30 demand locations in less than 3 hours, while an exact method could not converge after running for more than 48 hours.
4. We perform an extensive numerical study in order to analyse the value of flexibility and the robustness-performance trade-off that can be achieved by each approximation model. Furthermore, we provide some insights about the general structure of the

decisions that are proposed by each approximation model on a large set of problem instances.

The remainder of the paper is organized as follows. In Section 2.2, we review prior work about the robust location-transportation problem under demand uncertainty. In Section 2.3, we present six new tractable approximation models for the MRLTP. In Section 2.4, we establish the relation between the bounds that are obtained using each approximation model and identify conditions under which some of the models return exact solutions. Next, we present in Section 2.5, the details of a decomposition scheme that can be used to accelerate the resolution of larger-sized models. In Section 2.6, we provide numerical results; and finally, the conclusions and possible future research directions are presented in Section 2.7.

2.2 Prior Work

To the best of our knowledge, Atamtürk and Zhang (2007) were the first to study a model related to the two-stage robust location-transportation problem (TRLTP), a special case of MRLTP with a single-period $T = 1$, for an application of network flow and design problem where their objective was to minimize worst-case cost over a budgeted uncertainty set. They compared a two-stage robust optimization model with a stochastic program where the objective of the stochastic program, was to minimize the sum of the first-stage cost and the expected value of the second-stage cost. When distribution was captured by 200 demand scenarios, they showed that, while the solution of the two-stage robust optimization model increased the expected cost by 1.1%, it actually decreased by 29.1% the cost incurred under the worst-case scenario. They identified the TRLTP as a special case of the modelling framework, and after recognizing that their problem was NP-hard, proposed to use a cutting-plane algorithm to reach a global optimum.

Recently Gabrel et al. (2014a) and Zeng and Zhao (2013) proposed two cutting-plane methods to solve a TRLTP exactly under the budgeted uncertainty set with an integer budget. Gabrel et al. (2014a) showed that the adversarial problem in the TRLTP could be reformulated as a MILP. The master problem of the TRLTP could then be tackled using

Kelley's cutting-plane algorithm, given that optimality cuts are provided using a MILP solver. Zeng and Zhao (2013) seem to have improved on the solution time by employing a column-and-constraint generation (C&CG) algorithm instead of Kelley's cutting plane algorithm. Finally, in a similar transportation problem, Lei et al. (2015) proposed a two-level cutting plane method for a two-stage mobile-facility fleet sizing and routing problem wherein the fleet sizing and routing plan are determined in the first stage and the allocation of demands to the mobile facilities are determined in the second stage. Although there is empirical evidence that these exact resolution methods are efficient, the adversarial problem that is solved in each case takes the form of a MILP that is inherently NP-hard. There is therefore always a risk of having to endure unbearable computation times before obtaining solutions to any specific problem instance.

In Baron et al. (2011), the authors can be considered to have proposed the first tractable conservative approximation of the MRLTP model. In their paper, the authors proposed a robust optimization model in which static (*i.e.,* inflexible) production and fractional transportation policies are optimized. Indeed, they replaced the Y_{ij}^t variables with $X_{ij}^t \zeta_j^t$, which reflects the notion that X_{ij}^t is the proportion of demand at location j and time t that is satisfied by the facility at location i. Specifically, their proposed fractional variable-based (FVB) model takes the following form:

$$
\begin{aligned}
\text{(FVB)} \quad \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{P}}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij}) X_{ij}^t \zeta_j^t - \boldsymbol{c}^T \boldsymbol{P}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) && \text{(2.5a)}\\
\text{subject to} \quad & \sum_j X_{ij}^t \zeta_j^t \leq P_i^t \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\, \forall i \,,\, \forall t && \text{(2.5b)}\\
& \boldsymbol{P}^t \leq \boldsymbol{Z} \,,\, \forall t && \text{(2.5c)}\\
& \sum_i X_{ij}^t \leq 1 \,,\, \forall j \,,\, \forall t && \text{(2.5d)}\\
& \boldsymbol{X}^t \geq 0 \,,\, \forall t && \text{(2.5e)}\\
& \boldsymbol{Z} \leq M\boldsymbol{I} \,,\, \boldsymbol{I} \in \{0,1\}^L . && \text{(2.5f)}
\end{aligned}
$$

They next studied the impact of two types of uncertainty sets–box and ellipsoidal–on on the structure of the robust solution and compared it to the nominal one. In particular, they paid special attention to the number of opened facilities, the total capacity of facilities, and the number of deliveries made from each facility to the customer locations under different scenarios. Surprisingly, the following example highlights the fact that the solution of the FVB model might drop opportunities of making profits that are arbitrarily large even with respect to the worst-case scenario. In contrast, the simpler RC model actually does not suggest such a conservative solution for the same instances. On the other hand, some might argue that the FVB model provides a transportation policy that can easily be interpreted.

Example Consider an example of MRLTP with $T = 1$ and two customers, such that $\boldsymbol{\zeta} \in [\bar{\boldsymbol{\zeta}} \pm \hat{\boldsymbol{\zeta}}]$ where $\bar{\boldsymbol{\zeta}} = 10000$ and $\hat{\boldsymbol{\zeta}} = 5000$. The locations of customers is considered as the candidate location of facilities $L = 2$. The open facility will cover demand, if possible, with $\eta = 1$, $c_i = 0.1$, $c_{0i} = 0.1$ and $f_i = 3000$ for all i, and the transportation cost between locations is equal to 1. We assume that the budget is $\Gamma = 2$, which leads to a box uncertainty set. As is shown in Appendix 2.8.1, the optimal value of RC model (2.2) is equal to 1000, but the optimal value of the FVB model is zero in this example. This indicates that, while the RC model suggests opening the two facilities, which leads to a worst-case profit of 1000, the FVB model closes everything down. When scaling every parameter in the objective function by some $\alpha > 0$, FVB will let go of an arbitrarily large opportunity to make a profit. Intuitively, the over-conservatism of the FVB model is due to the fact that any feasible candidate for production must satisfy the largest possible demand, because of (2.5b), while the worst-case profits that end up being measured in (2.5a) actually account for the lowest demand. This necessarily leads the FVB model to imply that a lot of the production will be wasted once one attempts to satisfy even a small amount of demand.

Recently, Bertsimas and de Ruiter (2015) proposed applying affine adjustments on a dual reformulation of the TRLTP and showed improved computation time as compared to applying the same type of adjustment on the original TRLTP. Given that the two types of applications of these adjustments are shown to be equivalent, it is likely that their methods could be used to improve the resolution time of the models proposed in this work if one wishes to avoid using dedicated decomposition schemes. Yet, we are still convinced at the time of writing this article that it is necessary to employ row generation algorithms of the type presented here to obtain solutions to the larger instances of the MRLTP problem in a reasonable amount of time.

2.3 Six Conservative Tractable Approximations

In what follows, we provide six progressive ways of improving the quality of the solution obtained from the RC and FVB models. Each will employ the idea of affine adjustments from Ben-Tal et al. (2004) and a version of the splitting-based uncertainty set extensions from Chen and Zhang (2009) to exploit to a different extent the fact that the operational decisions $\boldsymbol{P}$ and $\boldsymbol{Y}$ can be adjusted to the realization of the demand. The type of flexibility added by our models can be divided into three classes. Similarly to what is done in the FVB model, the first class of approximation models, called "customer-driven", will adjust the size of a delivery to a customer simply based on information about that customer's demand, *i.e.,* that $Y_{ij}^t := \pi_{ij}^t(\zeta_j)$ with $\pi_{ij}^t : \mathbb{R} \to \mathbb{R}$. In opposition, the second class of approximation models, called "market-driven," will be more flexible and attempt to optimize delivery policies that take into account the state of the market as a whole, *i.e.,* that $Y_{ij}^t := \pi_{ij}^t(\zeta^t)$ with $\pi_{ij}^t : \mathbb{R}^N \to \mathbb{R}$. This second class will necessarily lead to models that are more computationally demanding yet have the potential to identify better-performing strategies. We will finally introduce a final class of approximation models, referred as "history-driven," that will attempt to exploit the full history of demand, even though we have not yet identified an improvement there that motivates the added computational burden. Note that, in presenting each of the approximation models, we omit to derive and spell out the finite dimensional MILP reformulation that would be obtained by applying duality theory to each robust constraint and objective function, for the sake of keeping the presentation compact.

2.3.1 Customer-driven Affine Adjustments

Our first approximation model will stem from the realization that, in the recourse problem of the MRLTP, namely, problem (2.4), the inequality constraint (2.4c) will be active at optimum and can therefore be replaced with an equality constraint. This argument motivates replacing P_i^t with $\sum_j X_{ij}^t \zeta_j^t$ for all i and t in the FVB model (2.5). Using this replacement, our model effectively fully adapts variable P_i^t to the revealed demand, which was an important issue with the FVB model. In order to ensure that we obtain a tighter approximation than with the RC model, we also propose replacing the fractional adjustment, $Y_{ij}^t := X_{ij}^t \zeta_j^t$ with $Y_{ij}^t := X_{ij}^t \zeta_j^t + W_{ij}^t$ for all i, j, and t. The motivation for using W_{ij}^t is that there are some cases, as shown in Example 1, wherein RC provides a tighter solution than FVB. Introducing the variable W_{ij}^t, namely, the "static" component of the transportation policy, enables us to guarantee that this revised model always provides a solution that is at most as conservative as the solution of the RC model (see Proposition 2.4.1 for more details). Overall, these modifications lead to our revised fractional variable-based (RFVB1) model:

$$
\begin{aligned}
\text{(RFVB1)} \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{W}}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(X_{ij}^t \zeta_j^t + W_{ij}^t) && (2.6a)\\
& \qquad -({\boldsymbol{c}_0}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\
\text{subject to} \quad & \sum_i X_{ij}^t \zeta_j^t + W_{ij}^t \le \zeta_j^t \,,\ \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\ \forall j \,,\ \forall t && (2.6b)\\
& \sum_j X_{ij}^t \zeta_j^t + W_{ij}^t \le Z_i \,,\ \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\ \forall i \,,\ \forall t && (2.6c)\\
& X_{ij}^t \zeta_j^t + W_{ij}^t \ge 0 \,,\ \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\ \forall i \,,\ \forall j \,,\ \forall t && (2.6d)\\
& \boldsymbol{Z} \le M\boldsymbol{I} \,,\ \boldsymbol{I} \in \{0,1\}^L \,. && (2.6e)
\end{aligned}
$$

We next exploit an extended description of the budgeted uncertainty set proposed in Chen and Zhang (2009) in order to optimize customer-driven transportation policies that have a piecewise-linear structure, (We also refer the reader to Georghiou et al. (2015) for details about techniques involving non-linear decision structures.) Specifically, we employ a lifting of the demand uncertainty space

$$\mathcal{D} = \left\{ \boldsymbol{\zeta} \in \mathbb{R}^{N\times T} \middle| \exists(\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2,\, \boldsymbol{\zeta} = \bar{\boldsymbol{\zeta}} + \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^- \right\}$$

where

$$\mathcal{D}_2 = \left\{ (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathbb{R}^{N\times T} \times \mathbb{R}^{N\times T} \middle| \begin{array}{c} \exists(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathbb{R}^{N\times T} \times \mathbb{R}^{N\times T}, \boldsymbol{\delta}^+ \geq 0, \boldsymbol{\delta}^- \geq 0, \|\boldsymbol{\delta}^+ + \boldsymbol{\delta}^-\|_\infty \leq 1, \\ \|\boldsymbol{\delta}^+ + \boldsymbol{\delta}^-\|_1 \leq \Gamma,\, \zeta_j^{t+} = \hat{\zeta}_j^{t+}\delta_j^{t+},\, \zeta_j^{t-} = \hat{\zeta}_j^{t-}\delta_j^{t-}\, \forall j\, \forall t \end{array} \right\}.$$

As illustrated in Figure 2–1, this lifting allows one to define different affine policies for positive perturbations, than those defined for negative perturbations thus giving rise to the possibility of a non-linear adjustment with better performance. For example, by letting $W_{ij} = \alpha\bar{\zeta}_j^t$, $X_{ij}^{t+} = 0$ and $X_{ij}^{t+} = -\alpha$ for some $0 \leq \alpha \leq 1$, the lifting implements the policy $Y_{ij}^t := \alpha \min(\zeta_j^t; \bar{\zeta}_j^t)$ (see Figure 2–1(c)), which can make better use of the capacity Z_i that is made available.

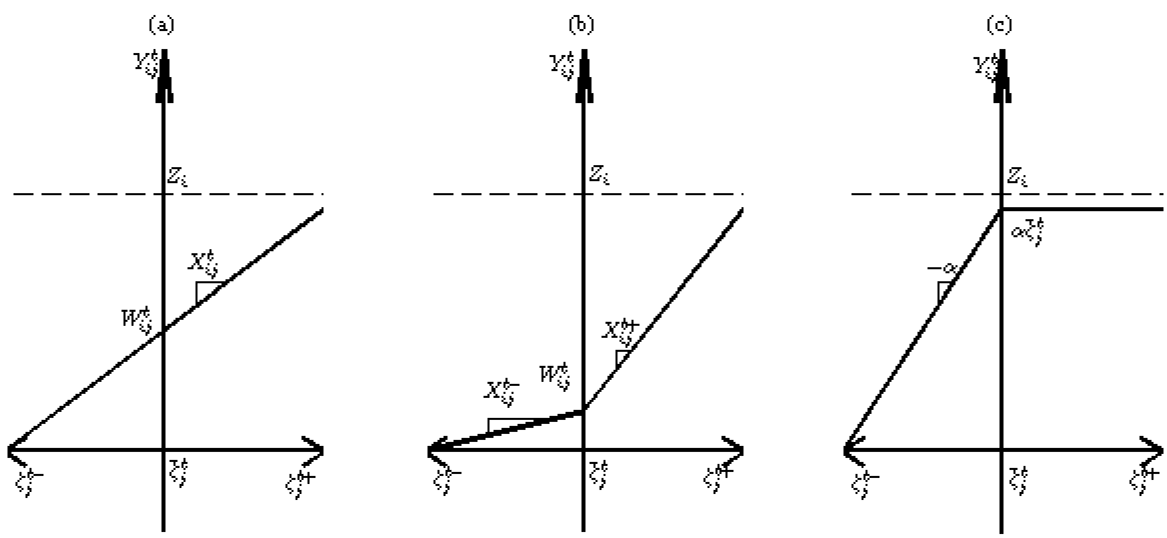

Figure 2–1: Illustrative comparison of an affine adjustment in (a) and an affine adjustment on the lifted space $(\zeta_j^{t+}, \zeta_j^{t-})$ in (b). Finally, (c) presents an example of lifted adjustment that implements $Y_{ij}^t := \alpha \min(\zeta_j^t; \bar{\zeta}_j^t)$ in order to make better use of available capacity.

This manipulation of the model leads to our second revision of the fractional variable-based (RFVB2) model:

$$\text{(RFVB2)} \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^+,\boldsymbol{X}^-,\boldsymbol{W}}{\text{maximize}} \quad \min_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(X_{ij}^{t+}\zeta_j^{t+} + X_{ij}^{t-}\zeta_j^{t-} + W_{ij}^t) - ({\boldsymbol{c}_0}^T\boldsymbol{Z} + \boldsymbol{f}^T\boldsymbol{I}) \quad (2.7a)$$

$$\text{subject to} \quad \sum_i X_{ij}^{t+}\zeta_j^{t+} + X_{ij}^{t-}\zeta_j^{t-} + W_{ij}^t \leq \zeta_j^t \,,\, \forall\,(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2 \,,\, \forall\, j \,,\, \forall\, t \quad (2.7b)$$

$$\sum_j X_{ij}^{t+}\zeta_j^{t+} + X_{ij}^{t-}\zeta_j^{t-} + W_{ij}^t \leq Z_i \,,\, \forall\,(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2 \,,\, \forall\, i \,,\, \forall\, t \quad (2.7c)$$

$$X_{ij}^{t+}\zeta_j^{t+} + X_{ij}^{t-}\zeta_j^{t-} + W_{ij}^t \geq 0 \,,\, \forall\,(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2\,\forall\, i \,,\forall\, j \,,\, \forall t \quad (2.7d)$$

$$\boldsymbol{Z} \leq M\boldsymbol{I} \,,\, \boldsymbol{I} \in \{0,1\}^L \,. \quad (2.7e)$$

2.3.2 Market-driven Affine Adjustments

We now provide three approximation models that will attempt to exploit full market information in making deliveries. The first of these attempts can be considered a direct application of the AARC framework for the MRLTP, as it was initially introduced by Ben-Tal et al. (2004). In such a framework, the adaptive policies for later-stage decisions are considered to be restricted to the set of affine functions of the uncertain parameters. In the context of this problem, this means that each adaptive policy of the MRLTP model (2.3) should take the form $Y_{ij}^t := (\boldsymbol{X}_{ij}^t)^T\boldsymbol{\zeta}^t + W_{ij}^t$ with $\boldsymbol{X}_{ij}^t \in \mathbb{R}^N$ and $W_{ij}^t \in \mathbb{R}$. In other words, this means that the delivery for a customer j can depend on all the orders that are made in this market. Intuitively, this added flexibility might be beneficial, considering that the amount of production is constrained by the capacity of each facility, Z_i; therefore, an increase in demand from a nearby customer might justify reducing the number of goods to transport to a more distant customer in order to improve profitability. We note that, similarly as before, the variable P_i^t of the MRLTP model (2.3) will be replaced by $\sum_j Y_{ij}^t$ in all of our proposed approximations. When restricting our search to affine policies of the $\boldsymbol{\zeta}^t$ vector, the approximation model takes the following form:

$$\text{(AARC)} \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{W}}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)\left((\boldsymbol{X}_{ij}^t)^T\boldsymbol{\zeta}^t + W_{ij}^t\right) - ({\boldsymbol{c}_0}^T\boldsymbol{Z} + \boldsymbol{f}^T\boldsymbol{I}) \quad (2.8a)$$

$$\text{subject to} \quad \sum_i (\boldsymbol{X}_{ij}^t)^T \boldsymbol{\zeta}^t + W_{ij}^t \leq \zeta_j^t \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\, \forall j \,,\, \forall t \tag{2.8b}$$

$$\sum_j (\boldsymbol{X}_{ij}^t)^T \boldsymbol{\zeta}^t + W_{ij}^t \leq Z_i \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\, \forall i \,,\, \forall t \tag{2.8c}$$

$$(\boldsymbol{X}_{ij}^t)^T \boldsymbol{\zeta}^t + W_{ij}^t \geq 0 \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\, \forall i \,, \forall j \,,\, \forall t \tag{2.8d}$$

$$\boldsymbol{Z} \leq M\boldsymbol{I}, \boldsymbol{I} \in \{0,1\}^L \,. \tag{2.8e}$$

Similarly to what was done to obtain the RFVB2 model, AARC can be improved by lifting the uncertainty set. LAARC of MRLTP (2.3) can be obtained by considering policies that are affine in the pair of perturbations $(\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2$, namely, $Y_{ij}^t := (\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t$ with $\boldsymbol{X}_{ij}^{t+} \in \mathbb{R}^N$, $\boldsymbol{X}_{ij}^{t-} \in \mathbb{R}^N$, and $W_{ij}^t \in \mathbb{R}$. This new approximation model takes the following more sophisticated form:

(LAARC)

$$\underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^+,\boldsymbol{X}^-,\boldsymbol{W}}{\text{maximize}} \quad \min_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-) \in \mathcal{D}_2} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) \left((\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \right) - ({\boldsymbol{c}_0}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \tag{2.9a}$$

$$\text{subject to} \quad \sum_i (\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \leq \zeta_j^t \,,\, \forall \, (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2, \, \forall j, \, \forall t \tag{2.9b}$$

$$\sum_j (\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \leq Z_i \,,\, \forall \, (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2, \, \forall i, \, \forall t \tag{2.9c}$$

$$(\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \geq 0 \,, \forall \, (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2, \, \forall i, \forall j, \, \forall t \tag{2.9d}$$

$$\boldsymbol{Z} \leq M\boldsymbol{I}, \; \boldsymbol{I} \in \{0,1\}^L \,. \tag{2.9e}$$

Now, we propose an extension to the LAARC, referred as the ELAARC model, which will benefit from a manipulation of a multi-period robust optimization model. To the best of our knowledge, this is being presented for the first time. The key idea is to reformulate the recourse problem (2.4) in a way that relaxes the constraint that is plagued by uncertainty, without compromising the authenticity of the model. Namely, let us consider the following equivalent reformulation:

$$h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t) = \max_{\boldsymbol{Y}^t, \boldsymbol{P}^t, \boldsymbol{\theta}^t} \quad \sum_i \sum_j (\eta - d_{ij} - c_i) Y_{ij}^t - \sum_j u_j \theta_j^t \tag{2.10a}$$

$$\text{subject to} \quad \sum_i Y_{ij}^t \leq \zeta_j^t + \theta_j^t \,, \forall\, j \tag{2.10b}$$

$$\sum_j Y_{ij}^t \leq Z_i \,, \forall\, i \tag{2.10c}$$

$$\boldsymbol{Y}^t \geq 0 \,, \boldsymbol{\theta}^t \geq 0 \,, \tag{2.10d}$$

where $\boldsymbol{Y}^t \in \mathbb{R}^{L \times N}$, $\boldsymbol{\theta}^t \in \mathbb{R}^N$ and where each u_j is a marginal penalty for violating constraint (2.4b), which is chosen large enough for the optimal value of the optimization problem to remain the same. We refer the reader to Appendix 2.8.2 for a proof that the assignment $u_j = \max_i(\eta - c_i - d_{ij}) \;\; \forall\, j$ meets this criterion.

As for the LAARC model, we adjust the deliveries based on the lifted uncertainty space, $Y_{ij}^t := (\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t$; furthermore, we adjust each new auxiliary variable $\boldsymbol{\theta}_j$ according to $\theta_j^t := S_j^{t+} \zeta_j^{t+} + S_j^{t-} \zeta_j^{t-}$ in order to obtain the ELAARC approximation model

$$\text{(ELAARC)} \quad \underset{\substack{\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{X}^+, \boldsymbol{X}^-, \\ \boldsymbol{W}, \boldsymbol{S}^+, \boldsymbol{S}^-}}{\text{maximize}} \quad \min_{(\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2} \sum_t \sum_i \sum_j \Big((\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \Big)$$
$$- (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) - \sum_t \sum_j u_j (S_j^+ \zeta_j^{t+} + S_j^- \zeta_j^{t-}) \tag{2.11a}$$

$$\text{subject to} \quad \sum_i (\boldsymbol{X}_{ij}^{t+})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \leq \zeta_j^t$$
$$+ S_j^{t+} \zeta_j^{t+} + S_j^{t-} \zeta_j^{t-} \,, \forall\, (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2 \,, \forall\, j \,, \forall\, t \tag{2.11b}$$

$$S_j^{t+} \zeta_j^{t+} + S_j^{t-} \zeta_j^{t-} \geq 0 \,, \forall\, (\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2 \,, \forall\, j \,, \forall\, t \tag{2.11c}$$

$$\text{(2.9c)} - \text{(2.9e)} \,, \tag{2.11d}$$

where $\boldsymbol{S}^+ \in \mathbb{R}^{N \times T}$ and $\boldsymbol{S}^- \in \mathbb{R}^{N \times T}$. Finally, one might realize that, when using this lifted uncertainty space, the worst-case analysis of this optimization model really only depends on negative adversarial perturbations. This will be an interesting feature to exploit when the time comes to implement and solve the model.

Proposition 2.3.1 *The LAARC and ELAARC approximation models can respectively be reduced to the following two optimization problems:*

$$\begin{aligned}
\text{(LAARC2)}\quad \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^{-},\boldsymbol{W}}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}^{-}\in\mathcal{D}_3} \sum_t\sum_i\sum_j(\eta-d_{ij}-c_i)\left((\boldsymbol{X}_{ij}^{t-})^T\zeta^{t-}+W_{ij}^t\right)-({\boldsymbol{c}_0}^T\boldsymbol{Z}+\boldsymbol{f}^T\boldsymbol{I})\\
\text{subject to} \quad & \sum_i(\boldsymbol{X}_{ij}^{t-})^T\zeta^{t-}+W_{ij}^t\le\bar{\zeta}_j^t-\zeta_j^{t-}\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall j\,,\forall t\\
& \sum_j(\boldsymbol{X}_{ij}^{t-})^T\zeta^{t-}+W_{ij}^t\le Z_i\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall i\,,\forall t\\
& (\boldsymbol{X}_{ij}^{t-})^T\zeta^{t-}+W_{ij}^t\ge 0\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall i\,,\forall j\,,\forall t\\
& \boldsymbol{Z}\le M\boldsymbol{I},\ \boldsymbol{I}\in\{0,1\}^L\,,
\end{aligned}$$

and

(ELAARC2)

$$\begin{aligned}
\underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^{-},\boldsymbol{W},\boldsymbol{S}^{-}}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}^{-}\in\mathcal{D}_3} \sum_t\sum_i\sum_j(\eta-d_{ij}-c_i)\left((\boldsymbol{X}_{ij}^{t-})^T\zeta^{-}+W_{ij}\right)\\
& \qquad -({\boldsymbol{c}_0}^T\boldsymbol{Z}+\boldsymbol{f}^T\boldsymbol{I})-\sum_t\sum_j u_j(S_j^{t-}\zeta_j^{t-}) && \text{(2.12a)}\\
\text{subject to} \quad & \sum_i(\boldsymbol{X}_{ij}^{t-})^T\zeta^{-}+W_{ij}\le\bar{\zeta}_j^t-\zeta_j^{t-}+S_j^{t-}\zeta_j^{t-}\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall j\,,\forall t && \text{(2.12b)}\\
& \sum_j(\boldsymbol{X}_{ij}^{t-})^T\zeta^{-}+W_{ij}\le Z_i\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall i\,,\forall t && \text{(2.12c)}\\
& (\boldsymbol{X}_{ij}^{t-})^T\zeta^{-}+W_{ij}\ge 0\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall i\,,\forall j\,,\forall t && \text{(2.12d)}\\
& S_j^{t-}\zeta_j^{t-}\ge 0\,,\ \forall\boldsymbol{\zeta}^{-}\in\mathcal{D}_3\,,\forall j\,,\forall t && \text{(2.12e)}\\
& \boldsymbol{Z}\le M\boldsymbol{I}\,,\ \boldsymbol{I}\in\{0,1\}^L\,, && \text{(2.12f)}
\end{aligned}$$

where

$$\mathcal{D}_3=\left\{\boldsymbol{\zeta}^{-}\in\mathbb{R}^{N\times T}\middle|\exists\boldsymbol{\delta}^{-}\in\mathbb{R}^{N\times T},\,0\le\boldsymbol{\delta}^{-}\le 1,\,\sum_{t=1}^{T}\sum_{j=1}^{N}\delta_j^{t-}\le\Gamma,\,\zeta_j^{t-}=\hat{\zeta}_j^t\delta_j^{t-}\,\forall j\,\forall t\right\}.$$

Proof First, one can easily confirm that both LAARC and ELAARC respectively reduce to LAARC2 and ELAARC2 when the uncertainty set $\mathcal{D}$ is replaced with the following uncertainty set:

$$\mathcal{D}_2':=\mathcal{D}_2\cap\{(\boldsymbol{\zeta}^{+},\boldsymbol{\zeta}^{-})\in\mathbb{R}^{N\times T}\times\mathbb{R}^{N\times T}\,|\,\zeta^{+}=0\}\,.$$

Since $\mathcal{D}_2' \subset \mathcal{D}_2$, it is clear that the optimal values of LAARC2 and ELAARC2 are respectively at least as large as the optimal value of LAARC and ELAARC. Looking more specifically at the LAARC2 model, given any optimal solution $(\boldsymbol{I}^*, \boldsymbol{Z}^*, \boldsymbol{X}^{-*}, \boldsymbol{W}^*)$, it is possible to reconstruct a feasible solution for LAARC simply by considering $\boldsymbol{X}^{+*} = \boldsymbol{X}^{-*}$, which achieves the same objective value as the optimal value identified by LAARC2. Hence, this reconstructed solution is optimal for LAARC. Note that, in confirming feasibility of this reconstructed solution, the difficulty resides in establishing whether the robust demand constraint is satisfied, namely, that for all $j = 1, 2, \ldots, N$ and for all t, one can confirm that

$$\begin{aligned}
&\max_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2} \sum_i ((\boldsymbol{X}_{ij}^{t+*})^T \boldsymbol{\zeta}^{t+} + (\boldsymbol{X}_{ij}^{-*t})^T \boldsymbol{\zeta}^{t-} + W_{ij}^{t*}) - \bar{\zeta}_j^t - \zeta_j^{t+} + \zeta_j^{t-} \\
&\qquad = \max_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2} \sum_i ((\boldsymbol{X}_{ij}^{t-*})^T (\boldsymbol{\zeta}^{t+} + \boldsymbol{\zeta}^{t-}) + W_{ij}^{t*}) - \bar{\zeta}_j^t - \zeta_j^{t+} + \zeta_j^{t-} \\
&\qquad \leq \max_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{D}_2} \sum_i ((\boldsymbol{X}_{ij}^{t-*})^T (\boldsymbol{\zeta}^{t+} + \boldsymbol{\zeta}^{t-}) + W_{ij}^{t*}) - \bar{\zeta}_j^t + (\zeta_j^{t+} + \zeta_j^{t-}) \\
&\qquad = \max_{(0,\zeta^-)\in\mathcal{D}_2} \sum_i ((\boldsymbol{X}_{ij}^{t-*})^T \boldsymbol{\zeta}^{t-} + W_{ij}^{t*}) - \bar{\zeta}_j^t + \zeta_j^{t-} \\
&\qquad = \max_{\boldsymbol{\zeta}^-\in\mathcal{D}_3} \sum_i ((\boldsymbol{X}_{ij}^{t-*})^T \boldsymbol{\zeta}^{t-} + W_{ij}^{t*}) - \bar{\zeta}_j^t + \zeta_j^{t-} \leq 0 \,,
\end{aligned}$$

where we exploited the fact that, for all $(\boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-) \in \mathcal{D}_2$, ζ^+ is non-negative. Finally a similar argument can be made to confirm that the optimal solution of ELAARC2 can be used to obtain an optimal solution to ELAARC, simply by letting $\boldsymbol{X}^{+*} = \boldsymbol{X}^{-*}$ and $\boldsymbol{S}^{+*} = \boldsymbol{S}^{-*}$. ∎

2.3.3 History-driven Affine Adjustments

For completeness, we finally highlight the fact that, in a multi-period setting, one can suppose that an even more flexible transportation strategy can be obtained by employing affine adjustments that depend jointly on all previous realizations of the demand until the implementation of the transportation decision. Mathematically speaking, the injection of such additional flexibility leads to the following structures. For all t and j, in the case of the direct AARC approach, one gets $Y_{ij}^t := \sum_{t'=1}^t (\boldsymbol{X}_{ij}^{tt'})^T \boldsymbol{\zeta}^{t'} + W_{ij}^t$, while the history-driven version of LAARC would employ $Y_{ij}^t := \sum_{t'=1}^t (\boldsymbol{X}_{ij}^{tt'+})^T \boldsymbol{\zeta}^{t'+} + (\boldsymbol{X}_{ij}^{tt'-})^T \boldsymbol{\zeta}^{t'-} + W_{ij}^t$. Finally, the

ELAARC model could additionally employ $\theta_j^t := \sum_{t'=1}^{t} S_j^{tt'}\zeta_j^{t+} + S_j^{tt'}\zeta_j^{t+}$. We present below the history-driven version of ELAARC in its reduced form.

$$
\begin{aligned}
\text{(HD-ELAARC)} \quad \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^-,\boldsymbol{W},\boldsymbol{S}^-}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}^-\in\mathcal{D}_3} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(\sum_{t'=1}^{t} (\boldsymbol{X}_{ij}^{tt'-})^T \boldsymbol{\zeta}^{t'-} + W_{ij}^t) - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\
& \qquad - \sum_t \sum_j u_j (\sum_{t'=1}^{t} S_j^{tt'-}\zeta_j^{t'-}) && \text{(2.13a)} \\
\text{subject to} \quad & \sum_i \sum_{t'=1}^{t} (\boldsymbol{X}_{ij}^{tt'-})^T \boldsymbol{\zeta}^{t'-} + W_{ij}^t \leq \bar{\zeta}_j^t - \zeta_j^{t-} && \text{(2.13b)} \\
& \qquad + \sum_{t'=1}^{t} S_j^{tt'-}\zeta_j^{t'-}\,,\ \forall\, \boldsymbol{\zeta}^- \in \mathcal{D}_3\,,\ \forall\, j\,,\ \forall\, t \\
& \sum_j \sum_{t'=1}^{t} (\boldsymbol{X}_{ij}^{tt'-})^T \boldsymbol{\zeta}^{t'-} + W_{ij}^t \leq Z_i\,,\ \forall\, \boldsymbol{\zeta}^- \in \mathcal{D}_3\,,\ \forall\, i\,,\ \forall\, t && \text{(2.13c)} \\
& \sum_{t'=1}^{t} (\boldsymbol{X}_{ij}^{tt'-})^T \boldsymbol{\zeta}^{t'-} + W_{ij}^t \geq 0\,,\ \forall\, \boldsymbol{\zeta}^- \in \mathcal{D}_3\,,\ \forall\, i\,, \forall\, j\,,\ \forall\, t && \text{(2.13d)} \\
& \sum_{t'=1}^{t} S_j^{tt'-}\zeta_j^{t'-} \geq 0\,,\ \forall\, \boldsymbol{\zeta}^- \in \mathcal{D}_3\,, \forall\, j\,,\ \forall\, t && \text{(2.13e)} \\
& \boldsymbol{Z} \leq M\boldsymbol{I}\,,\ \boldsymbol{I} \in \{0,1\}^L\,, && \text{(2.13f)}
\end{aligned}
$$

where for each i, j, t, and $t' \leq t$, we have that $\boldsymbol{X}_{ij}^{tt'-} \in \mathbb{R}^N$ and $S_j^{tt'-} \in \mathbb{R}$.

While we will show in our numerical experiments that such history-driven models can be used to obtain even tighter bounds than their non-history-driven versions, we note two important drawbacks. First, from a computational perspective, the number of parameters that need to be optimized using this type of adjustment scales the order of $O(LN^2T^2)$. Perhaps as importantly, the decision rules that are obtained with this model will suggest strategies whose structures are incoherent with the most natural structure that would be used by optimal, fully flexible strategies, namely, the fact that the transportation policy for time t only depends on the realized demand for time t. For these two reasons, we will later omit to present a complete numerical analysis of this model.

2.4 Theoretical Analysis of Robust Approximation Models

In this section, we are interested in demonstrating theoretically how better-quality solutions can be obtained by using an approximation model that offers more flexibility for the delayed decisions. In particular, we start by establishing what the respective qualities are of the bounds that obtained from each model regarding the worst-case profit of a candidate solution for facility locations and capacities.

Proposition 2.4.1 *Given some fixed values for the strategic decision vectors* $\boldsymbol{I} \in \{0,1\}^L$ *and* $\boldsymbol{Z} \in \mathbb{R}^L$*, let* $f_{RC}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{MRLTP}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{FVB}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{RFVB1}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{RFVB2}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{AARC}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{LAARC}(\boldsymbol{I},\boldsymbol{Z})$*,* $f_{ELAARC}(\boldsymbol{I},\boldsymbol{Z})$ *and* $f_{HD\text{-}ELAARC}(\boldsymbol{I},\boldsymbol{Z})$ *respectively be the value of the objective functions of approximation models* (2.2)*,* (2.3)*,* (2.5)*,* (2.6)*,* (2.7)*,* (2.8)*,* (2.9)*,* (2.11)*, and* (2.13) *when the rest their respective decision variables are optimized. The following partial ordering is satisfied for any values of* I *and* Z*:*

$$f_{RC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{RFVB1}(\boldsymbol{I},\boldsymbol{Z}) \le f_{RFVB2}(\boldsymbol{I},\boldsymbol{Z}) \le f_{LAARC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{ELAARC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{MRLTP}(\boldsymbol{I},\boldsymbol{Z}),$$

$$f_{FVB}(\boldsymbol{I},\boldsymbol{Z}) \le f_{RFVB1}(\boldsymbol{I},\boldsymbol{Z}) \le f_{AARC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{LAARC}(\boldsymbol{I},\boldsymbol{Z})\,,$$

$$f_{ELAARC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{HD\text{-}ELAARC}(\boldsymbol{I},\boldsymbol{Z}) \le f_{MRLTP}(\boldsymbol{I},\boldsymbol{Z})\,.$$

Proof The function $f_{ELAARC}(\boldsymbol{I},\boldsymbol{Z})$ provides a lower bound on true worst-case profit $f_{\text{MRLTP}}(\boldsymbol{I},\boldsymbol{Z})$ since the adjustable variables that appear in problem (2.10) are limited to affine functions of uncertain parameter. The ELAARC model reduces to the LAARC model when the value of variables S_j^{t+} and S_j^{t-} are forced to take a zero value for all j and t. One can also show that the LAARC model reduces to the AARC model when the constraint $\boldsymbol{X}_{ij}^{t+} = -\boldsymbol{X}_{ij}^{t-}\,\forall\, i,j,t$, is added, thus leading to a lower evaluation of the worst-case multi-period profit. The LAARC model also reduces to the RFVB2 model when adding the constraints that each term of $\boldsymbol{X}_{ij}^{t} \in \mathbb{R}^N$ equals zero except for the j-th term. A similar set of constraints make the AARC model reduce to the RFVB1 model. The RFVB2 model reduces to the RFVB1 model under similar conditions to those that make LAARC reduce to AARC. Lastly, one can show that RFVB1 upper bounds RC since the optimization model becomes equivalent to RC when we force $\boldsymbol{X} = 0$.

Next, assuming that W is fixed to zero, one can show that the evaluation of the worst-case profit obtained from the RFVB1 model is larger than the evaluation from the FVB model since one can replace constraint (2.6c) with

$$\sum_j \zeta_j^t X_{ij}^t \leq P_i^t \,,\, \forall \boldsymbol{\zeta} \in \mathcal{D} \,,\, \forall i \,,\, \forall t \;\;\&\;\; P_i^t \leq Z_i \,,\, \forall i \,,\, \forall t \,,$$

after letting $\boldsymbol{P} \in \mathbb{R}^{L\times T}$ be a set of additional decision variables of the model and since the objective function of the RFVB1 model has the following property:

$$\begin{aligned}
&\min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)\zeta_j^t X_{ij}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\
&\quad = \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)\zeta_j^t X_{ij}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) + \sum_t c^T \boldsymbol{P}^t - c^T \boldsymbol{P}^t \\
&\quad = \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij})\zeta_j^t X_{ij}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) + \sum_t \sum_i c_i(\boldsymbol{P}_i^t - \sum_j \zeta_j^t X_{ij}^t) - \boldsymbol{c}^T \boldsymbol{P}^t \\
&\quad \geq \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t \sum_i \sum_j (\eta - d_{ij})\zeta_j^t X_{ij}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) - \sum_t \boldsymbol{c}^T \boldsymbol{P}^t \,.
\end{aligned}$$

In this derivation, the last inequality comes from the robust constraint $\sum_j \zeta_j^t X_{ij}^t \leq P_i^t \; \forall \boldsymbol{\zeta} \in \mathcal{D}$ for all i and t. Since this last expression is the objective function of the FVB model, it is clear that the optimal value of this problem will be lower than the value of the RFVB1 model. Now, given that, in fact, the RFVB1 optimizes the objective function over all W instead of forcing this decision variable to zero, as assumed earlier, it will necessarily even further increase the difference between the two bounds.

Finally, while it is clear that $f_{\text{ELAARC}}(\boldsymbol{I}, \boldsymbol{Z}) \leq f_{\text{HD-ELAARC}}(\boldsymbol{I}, \boldsymbol{Z})$ since the ELAARC model is equivalent to the HD-ELAARC after we introduce the constraint that $X_{ij}^{tt'} = 0$ for all $t \neq t'$, the case for $f_{\text{HD-ELAARC}}(\boldsymbol{I}, \boldsymbol{Z}) \leq f_{\text{MRLTP}}(\boldsymbol{I}, \boldsymbol{Z})$ needs a little more explanation. To clarify this relation, one needs to remember that, for all $\boldsymbol{\zeta} \in \mathbb{R}^{N\times T}$,

$$\begin{aligned}
\sum_t h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t) = \max_{\{\boldsymbol{Y}^t, \boldsymbol{P}^t\}_{t=1}^T} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - \boldsymbol{c}^T \boldsymbol{P}^t && \text{(2.14a)} \\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq \zeta_j^t \,,\, \forall j \,,\, \forall t && \text{(2.14b)} \\
& \sum_j Y_{ij}^t \leq P_i^t \,,\, \forall i \,,\, \forall t && \text{(2.14c)} \\
& \boldsymbol{P}^t \leq \boldsymbol{Z} \,,\, \forall t && \text{(2.14d)}
\end{aligned}$$

$$\boldsymbol{Y}^t \geq 0\,,\ \forall\, t\,, \tag{2.14e}$$

where all temporal decision variables are optimized jointly in a way that can exploit the full information about ζ, although it is unnecessary to do so, because the problem decomposes. Yet, from this perspective, if we replace each $\boldsymbol{Y}^t$ with a history-driven affine function $Y_{ij}^t := \sum_{t'=1}^{t}(\boldsymbol{X}_{ij}^{tt'})^T\zeta^{t'} + W_{ij}^t$, we necessarily obtain an under evaluation of $\sum_t h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t)$. Note that this argument further indicates that the affine adjustment for each $\boldsymbol{Y}^t$ does not need to be non-anticipative in order to generate a valid lower bound on worst-case profits. ■

The result presented in proposition (2.4.1) can easily be used to establish guarantees with respect to the optimized bound on worst-case profit that are evaluated by each model.

Corollary 2.4.2 *Let* f^*_{RC}, f^*_{MRLTP}, f^*_{FVB}, f^*_{RFVB1}, f^*_{RFVB1}, f^*_{AARC}, f^*_{LAARC}, f^*_{ELAARC} *and* $f^*_{HD\text{-}ELAARC}$ *respectively be the optimal value of* (2.2), (2.3), (2.5), (2.6), (2.7), (2.8), (2.9), (2.11), *and* (2.13). *The following partial ordering is always satisfied:*

$$f^*_{FVB} \leq f^*_{RFVB1} \leq f^*_{RFVB2} \leq f^*_{LAARC} \leq f^*_{ELAARC} \leq f^*_{HD\text{-}ELAARC} \leq f^*_{MRLTP}$$

$$f^*_{RC} \leq f^*_{RFVB1} \leq f^*_{AARC} \leq f^*_{LAARC}\,.$$

Together, these results show that more sophisticated models of this list always provide better conservative approximation of the optimal value of the MRLTP model (See Figure 2–2). In fact, any time one approximation model in this list exactly returns the optimal value of the MRLTP, all models that are higher or equal to it in this ordering are guaranteed to return an exact optimal solution and an exact optimal worst-case bound.

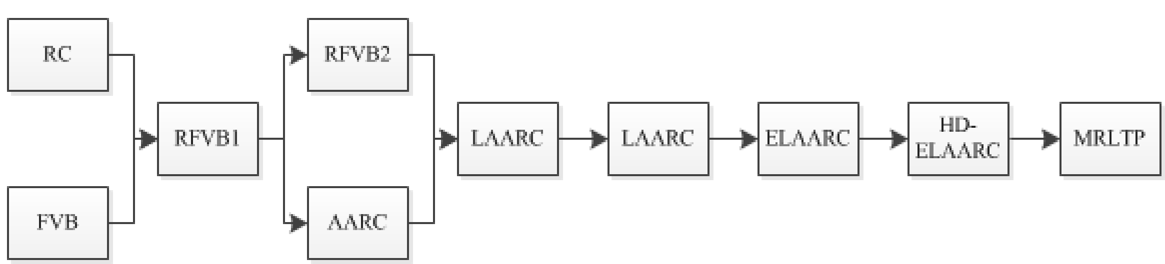

Figure 2–2: Partial ordering of the quality of bounds obtained from the different approximation models. Each arrow connects an approximation model to an approximation model that returns a tighter optimized bound for the optimal worst-case profit of the MRLTP model.

In the following theorem, we present conditions under which some of the proposed approximation models are exact and refer the reader to Appendix 2.8.3 for a detailed proof.

Theorem 2.4.3 *The MRLTP model* (2.3) *is equivalent to:*

- *RFVB1, RFVB2, AARC, LAARC, ELAARC, and HD-ELAARC when* $c_0 = 0$*,*
- *RC, RFVB1, RFVB2, AARC, LAARC, ELAARC, and HD-ELAARC when* $\Gamma = NT$*,*
- *LAARC, ELAARC, and HD-ELAARC when* $\Gamma = 1$*.*

Intuitively, for the cases of $c_0 = 0$ and $\Gamma = NT$, the proof relies on exploiting the fact that the optimization model used to evaluate $f_{\text{MRLTP}}(Z, I)$ can be shown to reduce to a problem in which the uncertainty decomposes over a number of constraints so that an equivalence between static and adjustable decisions identified in Ben-Tal et al. (2004) can be exploited. Otherwise, in the case of $\Gamma = 1$, our proof follows in the spirit of the arguments used to support Theorem 1 of Bertsimas and Goyal (2012), however, they must address differently the fact that none of the delayed decision variables are a mapping of the whole multi-temporal demand vector. We believe this proof contains elements that might pave the way for a possible extension of the result in Bertsimas and Goyal (2012).

Overall, Corollary 2.4.2 and Theorem 2.4.3 imply that LAARC, ELAARC, and HD-ELAARC not only provide tighter bounds than all other proposed approximation models but are also optimal for MRLTP for a number of interesting situations.

2.5 Improving Numerical Efficiency Using A Row Generation Algorithm

In this section, we propose a row generation algorithm as a solution method for ELAARC2 that we expect will be more computationally efficient than feeding the MILP reformulation of the model directly to an off-the-shelf MILP solver. Therefore, we reformulated ELAARC based on the following theorem, the proof of which can be found in Appendix 2.8.4.

Theorem 2.5.1 *The reduced ELAARC model is equivalent to*

$$\underset{\boldsymbol{I},\boldsymbol{Z},\rho}{\text{maximize}} \qquad \rho - (\boldsymbol{c_0}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \tag{2.15a}$$

$$\text{subject to} \quad \rho \leq g(\boldsymbol{Z}) \tag{2.15b}$$

$$\boldsymbol{Z} \leq M\boldsymbol{I}\,,\ \boldsymbol{I} \in \{0,1\}^L\,, \tag{2.15c}$$

where $g(\boldsymbol{Z})$ is defined as

$$\min_{\substack{\boldsymbol{\delta}^-,\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\psi}\\ \boldsymbol{\Theta},\boldsymbol{\Lambda},\boldsymbol{\Psi}}} \quad -({\boldsymbol{c}_0}^T\boldsymbol{Z} + \boldsymbol{f}^T\boldsymbol{I}) + \sum_t\sum_i Z_i\theta_i^t + \sum_t\sum_j \lambda_j^t\bar{\zeta}_j^t - \sum_t\sum_j \Lambda_{jj}^t\hat{\zeta}_j^t \tag{2.16a}$$

$$\text{subject to} \quad \theta_i^t + \lambda_j^t \geq \eta - c_i - d_{ij}\,,\ \forall\, i\,, \forall\, j\,, \forall\, t \tag{2.16b}$$

$$\Theta_{ik}^t + \Lambda_{jk}^t \geq (\eta - c_i - d_{ij})\delta_k^{t-}\,,\ \forall\, i\,,\ \forall\, j\,, \forall k\,, \forall\, t \tag{2.16c}$$

$$\sum_k \Theta_{ik}^t \leq \Gamma\theta_i^t, \Theta_{ik}^t \leq \theta_i^t, \forall i, \forall k, \forall t \tag{2.16d}$$

$$\sum_k \Lambda_{jk}^t \leq \Gamma\lambda_j^t, \Lambda_{jk}^t \leq \lambda_j^t,\ \Lambda_{jk}^t \leq B_j\delta_j^{t-}\,,\ \forall\, j\,,\ \forall\, k\,,\ \forall\, t \tag{2.16e}$$

$$\sum_k \Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-} \leq \Gamma\left(\theta_i^t + \lambda_j^t - \psi_{ij}^t - (\eta - c_i - d_{ij})\right)\,,\ \forall\, i\,, \forall\, j\, \forall\, t \tag{2.16f}$$

$$\Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-} \leq \theta_i^t + \lambda_j^t - \psi_{ij}^t - (\eta - c_i - d_{ij})\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, k\,,\ \forall\, t \tag{2.16g}$$

$$0 \leq \boldsymbol{\delta}^- \leq 1\,,\ \sum_t\sum_j \delta_j^{t-} \leq \Gamma \tag{2.16h}$$

$$\boldsymbol{\lambda} \geq 0\,,\ \boldsymbol{\Lambda} \geq 0\,,\ \boldsymbol{\theta} \geq 0\,,\ \boldsymbol{\Theta} \geq 0\,,\ \boldsymbol{\psi} \geq 0\,,\ \boldsymbol{\Psi} \geq 0\,, \tag{2.16i}$$

with $\boldsymbol{\delta}^- \in \mathbb{R}^{N\times T}$, $\boldsymbol{\theta} \in \mathbb{R}^{L\times T}$, $\boldsymbol{\lambda} \in \mathbb{R}^{N\times T}$, $\boldsymbol{\psi} \in \mathbb{R}^{L\times N\times T}$, $\boldsymbol{\Theta} \in \mathbb{R}^{L\times N\times T}$, $\boldsymbol{\Lambda} \in \mathbb{R}^{N\times N\times T}$, and $\boldsymbol{\Psi} \in \mathbb{R}^{L\times N\times N\times T}$.

Based on Theorem 2.5.1, we propose the use of a row generation algorithm to solve ELAARC2, wherein one goes through the following steps:

Row generation algorithm

Step #1: Set $UB = \infty$ and $LB = -\infty$. Solve the deterministic model with $\boldsymbol{\zeta} = \bar{\boldsymbol{\zeta}}$ to obtain an initial set of facility location $\dot{I}^{(1)}$ and capacities $\dot{Z}^{(1)}$. Let $\kappa = 1$.

Step #2: Solve the following subproblem

$$\text{(SP)} \quad \underset{\boldsymbol{\delta}^-,\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\psi},\boldsymbol{\Theta},\boldsymbol{\Lambda},\boldsymbol{\Psi}}{\text{minimize}} \quad \sum_t \sum_i \dot{Z}_i^{(\kappa)} \theta_i^t + \sum_t \sum_j \lambda_j^t \bar{\zeta}_j^t - \sum_t \sum_j \Lambda_{jj}^t \hat{\zeta}_j^t$$
$$\text{subject to} \quad (2.16\text{b}) - (2.16\text{i}).$$

Set $\dot{\theta}^{(\kappa)}$, $\dot{\lambda}^{(\kappa)}$, $\dot{\Lambda}^{(\kappa)}$, and $(\dot{\delta}^-)^{(\kappa)}$ to their respective values based on the optimal solution of the above SP model. Let ρ^* be the optimal value of the above SP model. Set $LB = \max(LB, \rho^* - (c_0^T \dot{Z}^{(\kappa)} + K^T \dot{I}^{(\kappa)}))$.

Step #3: Let $\kappa := \kappa + 1$ and solve the following master problem:

$$\text{(MP)} \quad \underset{\boldsymbol{I},\boldsymbol{Z},\rho}{\text{maximize}} \quad \rho - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \tag{2.17a}$$
$$\text{subject to} \quad \rho \leq \sum_t \sum_i (\dot{\theta}_i^t)^{(l)} Z_i + \sum_t \sum_j (\dot{\lambda}_j^t)^{(l)} \bar{\zeta}_{jt} - \sum_t \sum_j (\dot{\Lambda}_{jj}^t)^{(l)} \hat{\zeta}_{jt} \quad \forall\, l \in \{1, 2, \ldots, \kappa - 1\} \tag{2.17b}$$
$$\boldsymbol{Z} \leq M \boldsymbol{I},\ \boldsymbol{I} \in \{0,1\}^L. \tag{2.17c}$$

Let $\dot{I}^{(\kappa)}$, $\dot{Z}^{(\kappa)}$, and $\rho^{(\kappa)}$ take on the values of any optimal solution of the master problem (MP). Let $UB = \rho^{(\kappa)} - (\boldsymbol{c}_0^T \dot{Z}^{(\kappa)} + \boldsymbol{f}^T \dot{I}^{(\kappa)})$.

Step #4: If $UB - LB \leq \varepsilon$ then terminate and return $\dot{Z}^{(\kappa)}, \dot{I}^{(\kappa)}$ and $\rho^{(\kappa)}$ as the optimal solution; otherwise, repeat from Step #2. (Note that the termination condition can also be verified at the end of Step #2.)

One can actually improve the convergence speed of the algorithm by exploiting a specific type of valid inequalities for the ELAARC problem. Consider that, in order for a triplet $(\boldsymbol{I}, \boldsymbol{Z}, \rho)$ to be feasible in problem (2.15), for any $\{(\boldsymbol{\zeta}^-)^{(l)}\}_{l \in \Omega} \subset \mathcal{D}_3$, there must exist an assignment for $\boldsymbol{X}^-$, $\boldsymbol{W}$, and $\boldsymbol{S}^-$ such that the following constraint is satisfied:

$$
\begin{aligned}
&\rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) \left((\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l)} + W_{ij}^t \right) - \sum_t \sum_j u_j (S_j^{t-} (\zeta_j^{t-})^{(l)}) \,,\ \forall l \in \Omega \\
&\sum_i (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l)} + W_{ij}^t \leq (\zeta_j^{t-})^{(l)} \,,\ \forall l \in \Omega \,,\ \forall j \,, \forall t \\
&\sum_j (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l)} + W_{ij}^t \leq Z_i \,,\ \forall l \in \Omega \,,\ \forall i \,, \forall t \\
&(\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l)} + W_{ij}^t \geq 0 \,,\ \forall l \in \Omega \,,\ \forall i \,,\ \forall j \,, \forall t \,.
\end{aligned}
$$

This gives rise to the idea of replacing the master problem with

$$
\begin{aligned}
\text{(MP')} \quad & \underset{\boldsymbol{I}, \boldsymbol{Z}, \rho, \boldsymbol{X}^-, \boldsymbol{W}, \boldsymbol{S}^-}{\text{maximize}} && \rho - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\
& \text{subject to} && \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) \left((\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \right) \\
& && \qquad - \sum_t \sum_j u_j S_j^{t-} \zeta_j^{t-} ,\ \forall \boldsymbol{\zeta}^- \in \mathcal{D}_4^\kappa \\
& && \sum_i (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \leq \zeta_j^{t-} \,,\ \forall \boldsymbol{\zeta}^- \in \mathcal{D}_4^\kappa \,,\ \forall j \,, \forall t \\
& && \sum_j (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \leq Z_i \,,\ \forall \boldsymbol{\zeta}^- \in \mathcal{D}_4^\kappa \,,\ \forall i \,, \forall t \\
& && (\boldsymbol{X}_{ij}^{t-})^T \boldsymbol{\zeta}^{t-} + W_{ij}^t \geq 0 \,,\ \forall \boldsymbol{\zeta}^- \in \mathcal{D}_4^\kappa \,,\ \forall i \,,\ \forall j \,, \forall t \\
& && \text{(2.17b)} - \text{(2.17c)} \,,
\end{aligned}
$$

for some well-chosen finite set of feasible demand realizations $\mathcal{D}_4^\kappa$. In particular, our implementation uses $\mathcal{D}_4^\kappa$ as the set that simply contains the most recently identified worst-case demand $\boldsymbol{\zeta}_j^- := \bar{\boldsymbol{\zeta}}_j - \hat{\boldsymbol{\zeta}}_j(\boldsymbol{\delta}_j^-)^{(\kappa)}$.

One can observe in Table 2–1 the effect of including such valid inequalities in the decomposition scheme on a set of four problem instances of different sizes. In particular, it might come as a surprise to realize how much the number of iterations is reduced with this simple improvement.

Table 2–1: Impact of valid inequalities on row generation algorithm.

T	L	N	Γ%	# of iteration			Time (sec)		
				Without VI	With VI	Imp. %	Without VI	With VI	Imp. %
1	10	20	10	34	16	53	6	3	50
			30	46	37	20	11	8	27
			50	30	27	10	7	6	14
			70	27	18	33	5	2	60
			90	23	4	83	5	<1	>80
			100	24	2	92	4	<1	>75
			Avg.	31	17	48	6	<3.5	>42
1	20	40	10	257	163	37	88	46	48
			30	193	177	8	65	52	20
			50	164	135	18	70	49	30
			70	141	93	34	72	57	21
			90	105	20	81	60	22	63
			100	95	2	98	26	<1	>96
			Avg.	159	93	46	64	<37.8	>41
10	10	10	10	162	58	64	25	6	76
			30	174	89	49	28	15	46
			50	181	91	50	29	18	38
			70	159	34	79	25	5	80
			90	159	3	98	26	<1	>96
			100	147	2	99	23	<1	>96
			Avg.	164	46	73	26	<7.6	>71
10	15	15	100	368	63	83	534	88	84
			30	392	109	72	647	143	78
			50	476	121	75	707	173	76
			70	521	99	81	783	134	83
			90	542	15	97	800	20	98
			100	514	2	100	760	2	100
			Avg.	469	68	85	705	93	86

Remark One might alternatively consider the following classical decomposition scheme for robust optimization problems. Start by obtaining the solution of ELAARC for the nominal demand. Then, identify the worst-case realization for the objective and for each constraint. Finally, iterate until convergence, including in each new iteration, the worst-case demand that was generated for each constraint in the previous rounds. Unfortunately, this procedure is somewhat inefficient because of the large difference between the large size of the scenario-based version of ELAARC, which also holds binary variables, and the small size of the linear programming problems that provide the next worst-case demand.

2.6 Numerical Results

In this section, we evaluate the proposed approximation models on a set of randomly generated problem instances. The questions we seek to address are:

- What are the computational requirements of each approximation model and of the proposed row generation algorithm? (Section 2.6.1);
- What is the impact of varying the amount of uncertainty on the quality of the robust strategy and of the optimized bound proposed by each approximation model? (Section 2.6.2);
- What is the potential of each model with respect to trading-off average performance and robustness? (Section 2.6.3);
- Are any interesting insights about the structure of the robust decisions suggested by each approximation model, namely, in terms of number of the open facilities and the total capacity of open facilities, and of statistics about the amount of demand that is covered and the amount of unused capacity under different scenarios? (Section 2.6.4).

Each of these experiments will employ different sets of problem instances generated randomly according to the following procedure. We randomly generate N nodes on a unit square representing the demand points, and randomly choose L nodes of these N nodes as candidate facility locations. The respective unit transportation cost between a facility and a customer location d_{ij} is simply considered equal to the Euclidean distance between the two. For each facility i, we draw a value for each parameter η, c_{0i}, and f_i at random, uniformly and independently from the intervals $[1.5, 2]$, $[0.5, 0.1]$, $[0, 50000]$ respectively, while the production cost parameter is simply set as $c_i = 0.5$. The specific characterization of demand uncertainty is also randomly generated as follows: For each demand location j and period of time t, the nominal demand $\bar{\zeta}_j^t$ is generated uniformly from the interval $[0, 20000]$, and the maximum demand perturbation is set to $\hat{\zeta}_j^t = \varepsilon_j^t \bar{\zeta}_j^t$ where ε_j^t is drawn randomly between 0.15 and 1.

2.6.1 Computational Analysis

In this subsection, we compare the computational time to solve each approximation model when implemented directly using Optimization Programming Language (OPL) within IBM ILOG CPLEX Optimization Studio 12.6.1. For ELAARC, we also evaluated the performance of our novel row generation algorithm. We are especially interested in comparing these computational times to the computational requirements associated with the exact column and constraint (C&CG) algorithm[6] presented in Zeng and Zhao (2013) for varying sizes of problem instances and a budget of uncertainty Γ.

Table 2–2 focuses on single-period problems and presents the computation time for three problem instances of different sizes: the "small" size instance had 10 facilities and 10 demand locations, the medium-sized instance had 10 facility and 20 demand locations, and finally, the large-sized instance had 50 facility and 100 demand locations. For each instance, we measured the impact of varying the budget of uncertainty between different proportions of the total number of locations. A second set of computational experiments involved three multi-period instances of different sizes: the "small" instance had 10 periods, and 10 facility and 10 demand locations, while the largest instance had 20 periods, and 15 facility and 30 demand locations. Again, we attempted to measure the impact of varying the budget of uncertainty, but this time, between different proportions of $T \times N$, which is the size of the uncertain vector D in each problem instance.

Our first observation is that the customer-driven models (*i.e.,* FVB, RFVB1 and RFVB2) benefit from a strong computational efficiency and can actually be solved, even in the case of large problem instances, in a few seconds at most. While the market-driven models are more computationally demanding, we observe a significant reduction in the computational efforts

[6] The column-and-constraint generation algorithm proposed in Zeng and Zhao (2013) was implemented using the two-stage representation of our multi-period problem where the recourse problem takes the form presented in (2.14) and exploits a reduction that relies on the one-sided uncertainty set presented in $\mathcal{D}_3$.

for the LAARC2 and ELAARC2 models as compared to AARC, due to the use of the reduced form identified in Proposition 2.3.1. It also appears that, for medium-sized instances, the ELAARC2 model becomes slightly easier to solve than LAARC2, even though it involves a larger set of decision variables and constraints. Otherwise, although these two market-driven models can be solved in less than an hour for the medium single-period and multi-period instances, it becomes impossible to obtain a solution during our 48 hours time frame for the largest single-period and multi-period instances. One can obviously explain the difficulty of resolving market-driven models by the fact that the number of degrees of freedom for the affine adjustment grow at the rate of $O(LN^2T)$ instead of $O(LNT)$ for customer-driven models. Comparatively, we observe perhaps with surprise that the C&CG algorithm requires much less effort than any of these direct implementations. This seems to indicate that the efficiency of the decomposition scheme used by C&CG compensates for the fact that C&CG requires the solution of a number of outer and inner mixed-integer linear programs. This leaves us with the question of whether our conservative approximation models could also benefit from a well-designed decomposition scheme.

Indeed, looking at the "Row gen." column in both tables, we notice that the time needed to solve the ELAARC2 model can be significantly improved using our proposed row generation algorithm. More precisely, we estimate that this algorithm is responsible for reducing the computation requirements by a factor at least 16 to 260 (see multi-period instance with $\Gamma = 90\%$ where we have $48 \times 3600/663 = 260$) depending on the size of Γ. Practically speaking, we see that this algorithm allows us to identify robust approximate solutions for the largest single-period and multi-period instances in less than three hours (with an average of less than an hour and a half). In comparison, there is also evidence that the C&CG algorithm is unable to converge in less than 48 hours for the single-period instance when Γ equals 30% and 50% of the number of locations, while it is unable to do so for the large multi-period instance when Γ is greater than 30% of the total number of uncertain parameters (except for the trivial case of box uncertainty). One might finally observe that except for AARC the computational time of all models initially increased as the budget was increased, but then later decreased back down to a lower delay. The reason for this trend might be related to the number of extreme points of uncertainty set $\mathcal{D}_3$, which is known to contain the worst-case realizations for at least most of these models.

Regarding the resolution of HD-ELAARC, our experiments indicated that solving this model directly with a MILP solver typically takes about 30 minutes (80$\times$ more difficult than solving ELAARC2) for small-sized multi-period problem (*i.e.,* $T = 10$, $L = 15$, and $N = 15$)). Due to time limitations, we were unable to experiment with larger problem instances.

Conclusions: While both RFVB1 and RFVB2 models can be solved almost as efficiently as the FVB model, market-driven models should only be solved using standard optimization software when the problem instance is of medium size. For larger-sized problems, the use of a row generation algorithm is needed and is highly effective for these models. This allows us to provide nearly exact robust solutions (as shown in the next subsection) for problems where exact solutions are unobtainable. It appears however that much greater algorithmic efforts are needed to provide solutions to HD-ELAARC for problems of such large size.

2.6.2 Optimality Gap Analysis

In this subsection, we attempt to empirically compare the increasing quality of the approximate robust solutions obtained from the different conservative approximation models. Our hope is to quantify, from the perspective of worst-case analysis, what is the actual value in employing a more flexible model. The subsection's development is threefold. We first investigate, in single-period problem instances, the impact of changing the size of the potential demand perturbations ε and of the uncertainty budget Γ on the quality of these solutions. We then perform a similar analysis for the multi-period setting. Finally, we confirm that there exists multi-period problem instances for which the history-driven model HD-ELAARC can indeed be used to obtain a better approximate robust solution than the non-history-driven alternatives.

Table 2–2: Computational time (in seconds) needed for identifying approximate and exact robust solutions for three single-period instances of increasing sizes and varying level of budget (in % of total number of uncertain parameters). The dash "-" denotes situations where the method did not converge in less than 48 hours.

L	N	Γ%	FVB	RFVB1	RFVB2	AARC	LAARC2	ELAARC2	Row gen.	C&CG
10	20	10	<1	<1	<1	4	3	9	3	<1
		30	<1	<1	<1	2	6	10	8	1
		50	<1	<1	<1	11	7	7	6	1
		70	<1	<1	<1	6	13	13	2	1
		90	<1	<1	<1	24	18	28	<1	<1
		100	<1	<1	<1	219	2.6	9	<1	<1
		Avg.	<1	<1	<1	44	8	13	<3.7	<1
20	40	10	<1	<1	<1	521	415	303	46	8
		30	<1	<1	<1	272	264	166	52	11
		50	<1	<1	<1	283	275	191	49	50
		70	<1	<1	<1	581	523	398	57	19
		90	<1	<1	<1	1,747	1,308	1,287	22	3
		100	<1	<1	<1	69,394	2,326	1,011	<1	<1
		Avg.	<1	<1	<1	12,050	852	559	<44	<15
50	100	10	<1	2	6	-	-	-	3,241	8,465
		30	<1	4	11	-	-	-	4,563	-
		50	<1	4	9	-	-	-	8,460	-
		70	<1	5	4	-	-	-	3,781	7,682
		90	<1	4	6	-	-	-	1,382	7
		100	<1	2	2	-	-	-	<1	2
		Avg.	<1	3.5	6.3	-	-	-	<3,572	-

It is worth clarifying that, in what follows, every problem instance was generated using the procedure presented earlier in the introduction of this section, with a single exception concerning the size of the potential demand perturbations ε, which was fixed to specific values in order to monitor the effect of this parameter. Furthermore, in discussing our finding, we will refer to the following values, which are worth defining precisely.

- The "optimized worst-case bound" of a conservative approximation model refers to the best lower bound on worst-case profit that can be achieved according to this model. Mathematically, for some model $\mathcal{M} \neq$ MRLTP, this is measured using $f^*_{\mathcal{M}}$

Table 2–3: Computational time (in seconds) needed for identifying approximate and exact robust solutions for three multi-period instances of increasing sizes and varying level of budgets (in % of total number of uncertain parameters). The dash "-" denotes situations where the method did not converge in less than 48 hours.

T	L	N	Γ%	FVB	RFVB1	RFVB2	AARC	LAARC2	ELAARC2	Row gen.	C&CG
10	10	10	10	<1	<1	<1	25	11	10	6	3
			30	<1	<1	<1	32	25	19	15	1
			50	<1	<1	<1	41	38	21	18	1
			70	<1	<1	<1	115	19	29	5	1
			90	<1	<1	<1	103	23	31	<1	<1
			100	<1	<1	<1	61	32	27	<1	<1
			Avg.	<1	<1	<1	63	25	23	<8.8	<1.5
10	15	15	10	<1	<1	<1	500	342	428	88	1
			30	<1	<1	<1	3,497	1,813	1,916	143	12
			50	<1	<1	<1	4,749	2,770	2,662	173	9
			70	<1	<1	<1	4,815	3,360	3,048	134	36
			90	<1	<1	<1	5,140	3,933	3,681	20	8
			100	<1	<1	<1	6,316	4,431	4,120	2	2
			Avg.	<1	<1	<1	4,170	2,775	2,643	63	11
20	15	30	10	<1	<1	<1	-	-	-	3,781	184
			30	<1	<1	<1	-	-	-	5,646	-
			50	<1	<1	<1	-	-	-	10,567	-
			70	<1	<1	<1	-	-	-	4,445	-
			90	<1	<1	<1	-	-	-	663	-
			100	<1	<1	<1	-	-	-	1	<1
			Avg.	<1	<1	<1	-	-	-	4184	-

- The "achieved worst-case profit" of a strategic decision refers to the actual worst-case profit achieved if this strategic decision is applied. Mathematically, for a strategic decision $(\boldsymbol{I}^*_{\mathcal{M}}, \boldsymbol{Z}^*_{\mathcal{M}})$ obtained using model $\mathcal{M}$, this is measured using $f_{\text{MRLTP}}(\boldsymbol{I}^*_{\mathcal{M}}, \boldsymbol{Z}^*_{\mathcal{M}})$.
- The "optimal worst-case profit" of a problem instance refers to the best worst-case profit that can be achieved for this instance. Mathematically, it is measured using f^*_{MRLTP} and obtained in our experiments by solving the C&CG algorithm (see Footnote 6).
- The "relative optimized bound gap" of a conservative approximation model refers to the relative difference between the optimal worst-case profit for this problem instance

and the optimized worst-case bound of this model. Mathematically, for some model $\mathcal{M} \neq$ MRLTP, it is measured using $(f^*_{\text{MRLTP}} - f^*_{\mathcal{M}})/f^*_{\text{MRLTP}}$.

- The "relative suboptimality" of a strategic decision refers to the relative difference between the optimal worst-case profit for this problem instance and the achieved worst-case profit of this decision. Mathematically, for a strategic decision $(\boldsymbol{I}^*_{\mathcal{M}}, \boldsymbol{Z}^*_{\mathcal{M}})$ obtained using model $\mathcal{M}$, it is measured using $(f^*_{\text{MRLTP}} - f_{\text{MRLTP}}(\boldsymbol{I}^*_{\mathcal{M}}, \boldsymbol{Z}^*_{\mathcal{M}}))/f^*_{\text{MRLTP}}$.

2.6.2.1 Impact of Size of Potential Perturbation on Optimality Gap

We consider 100 randomly generated problem instances with $L = 10$, $N = 10$, and $T = 1$. Table 2–4 presents the average (taken over the set of 100 instances) relative optimized bound gap and the average relative suboptimality gap for the solutions (*i.e.,* identified strategies for $\boldsymbol{I}$ and $\boldsymbol{Z}$) of both customer-driven and market-driven model types, under different budgets of uncertainty Γ when the demand intervals are forced to a relatively small size, *i.e.,* $\varepsilon = 0.15$. Similarly, Tables 2–5 and 2–6 present the same statistics on the same set of instances but with medium-sized $\varepsilon = 0.30$, and large-sized $\varepsilon = 0.45$ demand intervals.

Regarding the quality of the optimized worst-case bound, one might first observe in these tables that, as indicated by Corollary 2.4.2, the optimized bounds always improve when one uses a more flexible approximation model. One might further notice that the most significant improvements appear to occur exactly when passing to models that implement the most significant changes in terms of added flexibility and resulting computational needs, namely, from the FVB model to the RFVB1 model, and later by passing to a market-driven model. When we look at the results for the FVB model and other customer-driven models, we observe that RFVB1 and RFVB2 models reduce by factors of 8 and 13 respectively the quality of the optimized worst-case bound offered by the FVB model. In particular, one might notice that, when $\varepsilon = 0.30$, the company always identifies profitability in servicing its customers under the RFVB1 and RFVB2 models, while the FVB model suggests shutting down all facilities at $\Gamma = 4$. This is serious evidence that the FVB model is overly conservative. Furthermore, it appears that a significant gain is achieved with the introduction of market-driven policies, such that the proposed optimized worst-case bounds are on average always less than 0.59% from being exact. Although the added value of using the LAARC and ELAARC models is not very pronounced (refer to underlined and **bold** entries respectively), the difference becomes more noticeable as the size of demand intervals is increased. Regarding sensitivity to the size of Γ and ε, one might notice that the quality of the optimized worst-case bounds for FVB decreases when the budget of uncertainty increases, unlike the other models. We also estimate the quality of the other model's optimized bound to be less affected by the growth of the size of the demand intervals.

Regarding the quality of the approximate robust solution itself, we can confirm that employing more flexible adjustments clearly improves the chances of identifying good strategic decisions. For instance, in Table 2–6, where there are large demand intervals, for $\Gamma = 5$, the FVB model always suggests that no facilities be built, thereby foregoing all chances of making any profit (*i.e.,* a 100% worst-case profit loss), while ELAARC provides strategic decisions that on average achieve a worst-case profit that is only 0.28% from being the optimal worst-case profit achievable. ELAARC also provides a guaranteed lower bound on worst-case profits that is on average only 0.50% lower than the optimal worst-case profit. It can also be observed that all our proposed methods provide optimal robust solutions for the case with $\Gamma = N$, as predicted by Theorem 2.4.3. Moreover the LAARC and ELAARC models' solutions are also optimal when $\Gamma = 1$.

Table 2–7 provides additional statistics about the relative suboptimality of the different solutions proposed by each approximation model in the 3000 problem instances surveyed in Tables 2–4, 2–5, and 2–6. Specifically, the table indicates, for a number of different percentage gaps, the proportion of instances for which each model was able to identify an approximate robust solution whose relative suboptimality was within that given gap. Each proportion can be interpreted as the likelihood that the solution obtained from a model achieves a worst-case profit that is within some percentage away from being optimal. The table also presents the average and maximum relative suboptimality gap for each model.

In particular, one can observe that the flexibility of ELAARC gives it the best chances of providing a solution that achieves a certain level of relative suboptimality. Yet, one can also note that, in terms of maximum relative suboptimality gap, LAARC was able to perform slightly better. This serves as a reminder that optimizing a tighter lower bound on an objective value does not guarantee that a solution of better quality will be obtained, however; in most cases, one can certainly say that it serves as a great proxy. It is also worth noting that the limited additional flexibility of RFVB1 and RFVB2, compared to FVB, has a significant payoff in terms of relative suboptimality. For instance, the proportion of problem instances where a guaranteed profit is wasted decreases from 75.0% to almost 1% with the RFVB1 and RFVB2 models. Finally, LAARC and ELAARC never forego the potential to make a positive profit in any of these instances.

Table 2–4: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.15

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	31.0	11.9	11.3	4.88	5.47	2.78	0.00	0.00	0	0	0	0
2	52.7	25.5	11.4	4.3	7.32	3.49	0.04	0.03	0.02	0.01	0.02	0.01
3	66.5	36.2	8.55	4.44	7.02	2.39	0.08	0.07	0.04	0.04	**0.03**	0.03
4	76.4	42.7	5.84	3.38	5.61	2.48	0.11	0.09	0.05	0.04	**0.04**	0.04
5	83.2	47.2	3.75	2.34	3.75	2.35	0.11	0.08	0.04	0.03	0.04	0.03
6	87.5	49.9	2.15	1.49	2.15	1.49	0.08	0.06	0.01	0.01	0.01	0.01
7	90.0	54.1	1.08	0.83	1.08	0.83	0.06	0.05	0.01	0.01	0.01	0.01
8	91.2	53.9	0.41	0.35	0.41	0.35	0.03	0.03	0.01	0.01	0.01	0.01
9	91.5	55.5	0.09	0.08	0.09	0.08	0.01	0.01	0.00	0.00	0.00	0.00
10	91.5	55.2	0	0	0	0	0	0	0	0	0	0

Table 2–5: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.30

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	55.0	26.2	22.6	10.2	11.3	6.02	0.00	0.00	0	0	0	0
2	82.4	47.7	24.4	10.5	15.5	7.53	0.12	0.11	0.06	0.05	**0.05**	0.04
3	95.4	74.5	19.2	11.7	15.7	6.14	0.36	0.25	0.12	0.09	**0.08**	0.08
4	99.7	95.4	13.6	8.96	13.1	6.69	0.59	0.38	0.15	0.11	**0.12**	0.11
5	100	100	9.07	6.45	9.07	6.45	0.79	0.46	0.15	0.1	**0.13**	0.09
6	100	100	5.31	4.01	5.31	4.01	0.75	0.45	0.11	0.08	**0.09**	0.06
7	100	100	2.68	2.14	2.68	2.14	0.50	0.33	0.05	0.05	0.05	0.05
8	100	100	1.02	0.9	1.02	0.9	0.2	0.17	0.03	0.03	0.03	0.03
9	100	100	0.22	0.22	0.22	0.22	0.04	0.04	0.00	0.00	0.00	0.00
10	100	100	0	0	0	0	0	0	0	0	0	0

Table 2–6: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when ε=0.45

Γ	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
1	70.7	34.7	34.8	16.5	17.6	9.34	0.00	0.00	0	0	0	0
2	95.5	75.0	39.7	20.4	25.5	14.3	0.20	0.14	0.20	0.14	**0.11**	0.09
3	100.0	98.9	32.6	22.22	26.87	10.74	1.67	0.90	0.40	0.30	**0.26**	0.21
4	100	100	23.9	17.6	23.2	14.0	3.08	1.54	0.57	0.41	**0.44**	0.37
5	100	100	16.2	12.3	16.2	12.3	3.57	1.82	0.56	0.34	**0.50**	0.28
6	100	100	9.67	7.65	9.67	7.65	2.89	1.61	0.46	0.25o	**0.40**	0.22
7	100	100	4.84	4.03	4.84	4.03	1.70	1.15	0.31	0.27	**0.29**	0.25
8	100	100	1.80	1.57	1.80	1.57	0.61	0.52	0.11	0.06	0.11	0.06
9	100	100	0.36	0.36	0.36	0.36	0.11	0.07	0.01	0.01	0.01	0.01
10	100	100	0	0	0	0	0	0	0	0	0	0

Table 2–7: Proportion of the 3000 problem instances analysed in Tables 2–4, 2–5, and 2–6 where the relative suboptimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
$= 0\%$	0.0	32.7	32.7	68.2	83.7	85.7
$\leq 0.1\%$	0.0	35.7	36.4	77.0	90.7	92.6
$\leq 1\%$	1.3	47.5	50.6	92.8	98.8	99.2
$\leq 10\%$	7.4	83.4	88.5	99.9	100.0	100.0
$= 100\%$	75.0	1.1	1.0	0.0	0.0	0.0
Avg. gap	79.81	5.91	4.7	0.26	0.05	0.04
Max gap	100	100	100	34.67	3.32	4.82

2.6.2.2 Optimality Gap Analysis in Multi-period Problems

We consider 100 randomly generated problem instances with $L = 10$, $N = 10$, and either three or five periods. Table 2–8 presents the same statistics as Table 2–4 but for a set of 100 problem instances with three periods $T = 3$, while the demand perturbation size is forced to $\varepsilon = 0.3$. Alternatively, Table 2–9 presents the same statistics for $T = 5$. In these tables, we observe a similar trend as before except for the perhaps unexpected fact that the RFVB2 model seems to provide better-quality solutions and bounds, on average, than AARC. In particular, as underlined in both tables, the average relative optimized bound gap is always very close to 1% or 2% (see the underlined entries), while this same statistic rises to values that are close to 4% or 7% with AARC. This seems to indicate that the flexibility provided by AARC, namely, of adapting to market information, is less useful than the flexibility provided by RFVB2, namely reacting differently to positive and negative perturbations. The same observation can be made when comparing the two models' average relative suboptimality gap. One can additionally confirm that LAARC and ELAARC both provide the best-quality solutions and optimized worst-case bound. Furthermore, ELAARC is able to slightly tighten its optimized bound (as shown in **bold**) and obtain solutions that are slightly less sub-optimal when Γ equals 30% of the total number of uncertain parameters. It finally appears, based on this experiment that, when one uses models other than the FVB, the quality of the approximate robust solutions improve, for any fixed percentage of uncertainty budget, as the number of time periods increases. This appears a little counterintuitive, but one might conjecture from this empirical evidence that, as the horizon becomes longer, it becomes easier to hedge (or perhaps hide from) the risks related to demand perturbations so that approximation models become more effective at identifying good strategies.

Table 2–10 repeats the analysis of Table 2–7 in presenting further statistics regarding the relative suboptimality of the solutions obtained from the different conservative approximation models. All statistics that are presented were assessed on the 1000 problem instances covered in Tables 2–8 and 2–9. Again, we see significant improvement for passing from the FVB model to RFVB1 (with the maximum gap being reduced from 100% to 8.37%), and very good odds (*i.e.,* 99.6%) of achieving less than a 1% relative suboptimality gap with LAARC or ELAARC. Yet, one should realize that the odds of achieving an exact solution with both of these models is significantly reduced in this set of multi-period problem instances, namely, a reduction from above 84% when $T = 1$ to less than 13.2% in this set of multi-period problems.

Table 2–8: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when $T = 3$ and ε=0.3

T	L	N	Γ%	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
				Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
3	10	10	0.1	52.0	6.84	7.67	3.79	0.45	0.20	4.29	1.52	0.06	0.05	0.06	0.04
			0.3	86.5	23.4	9.53	2.89	1.47	0.60	7.24	2.02	0.37	0.21	**0.35**	0.20
			0.5	85.6	23.4	6.94	1.60	2.42	0.86	4.44	1.56	0.71	0.36	**0.70**	0.36
			0.7	84.7	23.8	3.01	1.32	2.15	0.56	1.73	0.97	0.60	0.35	0.60	0.35
			0.9	84.3	24.5	0.29	0.18	0.28	0.18	0.20	0.14	0.10	0.08	0.10	0.08

Table 2–9: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Opt. gap) for the solutions obtained from each approximation model under different values of budget when $T = 5$ and ε=0.3

T	L	N	Γ%	FVB		RFVB1		RFVB2		AARC		LAARC		ELAARC	
				Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap	Bound gap	Opt. gap
5	10	10	0.1	60.4	3.76	5.30	3.00	0.26	0.10	4.22	2.18	0.03	0.02	0.03	0.02
			0.3	73.1	5.11	5.51	2.82	0.68	0.32	4.65	2.12	0.17	0.10	**0.16**	0.10
			0.5	70.4	6.43	4.56	1.65	1.09	0.55	3.33	1.03	0.31	0.15	**0.30**	0.14
			0.7	68.5	8.25	2.61	0.60	1.19	0.50	1.53	0.52	0.33	0.16	0.33	0.16
			0.9	67.5	9.59	0.36	0.15	0.31	0.11	0.18	0.09	0.06	0.04	0.06	0.04

Table 2–10: Proportion of the 1000 problem instances analysed in Tables 2–8 and 2–9 where the relative suboptimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
= 0%	0.0	1.1	2.7	1.0	12.6	13.2
≤ 0.1%	0.2	10.1	23.6	14.6	56.0	56.9
≤ 1%	2.0	43.8	92.2	56.2	99.6	99.6
≤ 10%	67.6	100.0	100.0	100.0	100.0	100.0
= 100%	5.7	0.0	0.0	0.0	0.0	0.0
Avg.	15.97	1.77	0.4	1.17	0.15	0.14
Max gap	100	8.37	3.54	6.18	1.07	1.05

2.6.2.3 Optimized Bound Gap Reduction Using HD-ELAARC

Table 2–11 presents the relative optimized bound gap and the relative suboptimality gap for the solutions of the ELAARC and HD-ELAARC models in a specific multi-period instance where $T = 3$, $L = 10$, and $N = 10$ drawn according to the procedure used previously for different values of the uncertainty budget Γ. The relative gaps that are reported confirm that HD-ELAARC has the potential to identify a tighter optimized worst-case bound for the MRLTP problem, and consequently, provides an approximate robust solution that slightly improves the relative suboptimality gap. Yet, we consider this improvement to be somewhat small for passing from a model whose size grows with $O(LN^2T)$ to $O(LN^2T^2)$.

Conclusions: It appears, based on this analysis, that RFVB2 and ELAARC2 are the two models that have the most to offer compared to other models in their respective class in terms of trading-off speed of resolution and robustness of the facility location strategy that they

Table 2–11: Relative optimized bound gap (Bound gap) and relative suboptimality gap (Opt. gap) for the solutions obtained from ELAARC and HD-ELAARC under different values of budget with ε =0.30.

	ELAARC		HD-ELAARC	
Γ %	Bound gap	Opt. gap	Bound gap	Opt. gap
10	0.32	0.32	0.11	0.11
20	0.60	0.45	0.40	0.29
30	0.77	0.48	0.38	0.21
40	0.74	0.34	0.25	0.09
50	0.93	0.14	0.40	0.07
60	0.83	0.40	0.38	0.17
70	0.44	0.21	0.22	0.13
80	0.42	0.19	0.17	0.07
90	0.17	0.17	0.04	0.04
100	0.00	0.00	0.00	0.00

are able to identify. Additionally, we observed that customer-driven approximation models, in particular the FVB model, are sensitive to the size of potential perturbations, while the performance of market-driven models appear to be a little more stable. It also appears that the performance of solutions from conservative approximation schemes somehow benefit from longer-horizon problems in which there might be more opportunities to hedge or hide from the risk. On the other hand, it appears much more difficult to close the suboptimality gap in larger problems with the type flexibility that is found in customer- and market-driven adjustments. There might still however be some hope to close this gap with a history-driven model like HD-ELAARC; however, one would be left with the challenge of designing an efficient decomposition scheme for this model.

2.6.3 Robustness-performance Trade-off

In this subsection, we study the robustness and performance of the approximate robust solutions obtained using our different approximation models in a pair of experiments. While the first experiment involves a set of 100 medium-sized single-period problem instances, where $L = 10$ and $N = 20$, the second one involves a set of 100 large-sized single-period problem instances, where $L = 50$ and $N = 100$. Each problem instance is generated according to the procedure described in the introduction of this section. Unlike what was done in the numerical studies of previous sections, we do not wish to evaluate the worst-case performance of the solutions obtained but rather estimate what type of balance these solutions can achieve, in terms of the compromise that must be made between potential protection against risk (captured by a percentile) and potential expected profit. More specifically, for each problem instance, we evaluate the statistical performance of each approximate robust solution on a set of 100 demand scenarios. To obtain each of these scenarios, each customer's demand is independently generated from its respective demand interval, using a uniform distribution. In the larger instances, due to the duration of the resolution process, we limit our study to the FVB, RFVB1, and RFVB2 models.

In Figure 2–3, we report the average expected profit and the average 10th percentile profit of each approximation model's solution as the total budget for the uncertainty set is varied. The same results are also presented in Figure 2–4 to highlight what type of compromise can be achieved by adjusting the budget of uncertainty. Considering that a common criticism of robust optimization approaches has been that it provides overly conservative solutions, it might come as a surprise that our results show that a flexible robust optimization approach with an appropriately calibrated uncertainty set (*e.g.,* the LAARC model with $\Gamma = 1$) will provide solutions that outperform the solutions of the deterministic model (2.1), obtained by setting $\Gamma = 0$, in terms of both expected profit and risk exposure, as measured through the 10th percentile. Another interesting observation is that overly conservative solutions might often actually be the result of not injecting enough flexibility in the robust optimization model, as is the case for the FVB and RFVB1 models. The figures clearly show that, whether the instance is small or large, it is always worth employing the slightly more sophisticated RFVB2 model to achieve a significantly better risk and return trade-off. Figure 2–3(a) also demonstrates how performance is improved by employing market-driven models.

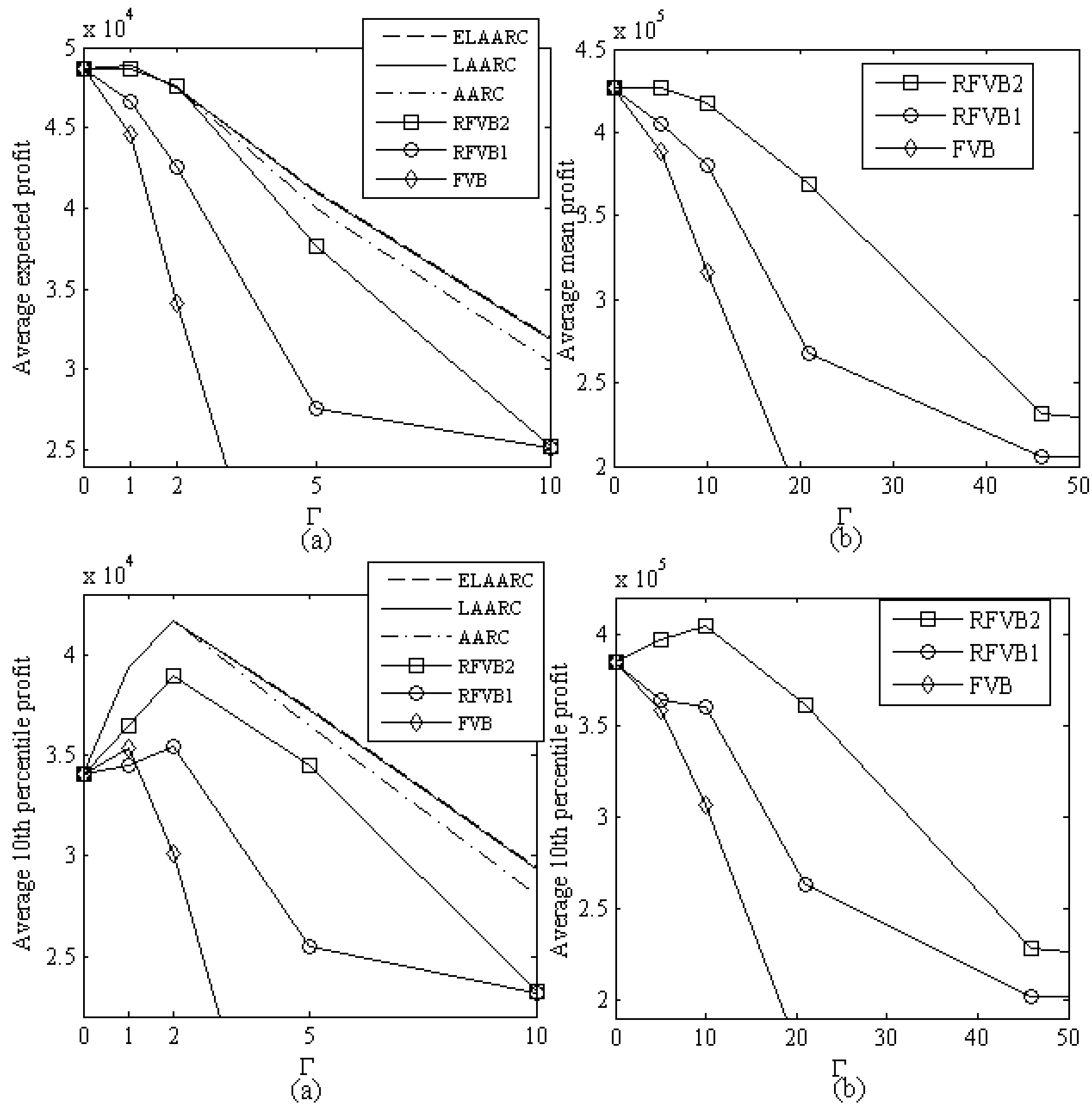

Figure 2–3: Average expected and 10th percentile profit achieved by the different robust methods on 100 problem instances while adjusting the level of conservativeness Γ. Figure (a) and (c) present the average expected and average 10th percentile profit respectively for medium-sized instances with $L = 10$ and $N = 20$, while (b) and (d) present the same statistics for large-sized instances with $L = 50$ and $N = 100$. Note that in (a) and (b), the curves for LAARC and ELAARC were combined since the performances indistinguishable.

Conclusions: Our experiments clearly show that, whether the instance is large or small, it is always worth employing the slightly more sophisticated RFVB2 model to achieve significantly better risk and return trade-off. Figure 2–4 also demonstrates how performance is improved by employing market-driven models. Note, however, that this was not confirmed on large problem instances due to the heavier computational requirements of the resolution methods for these models.

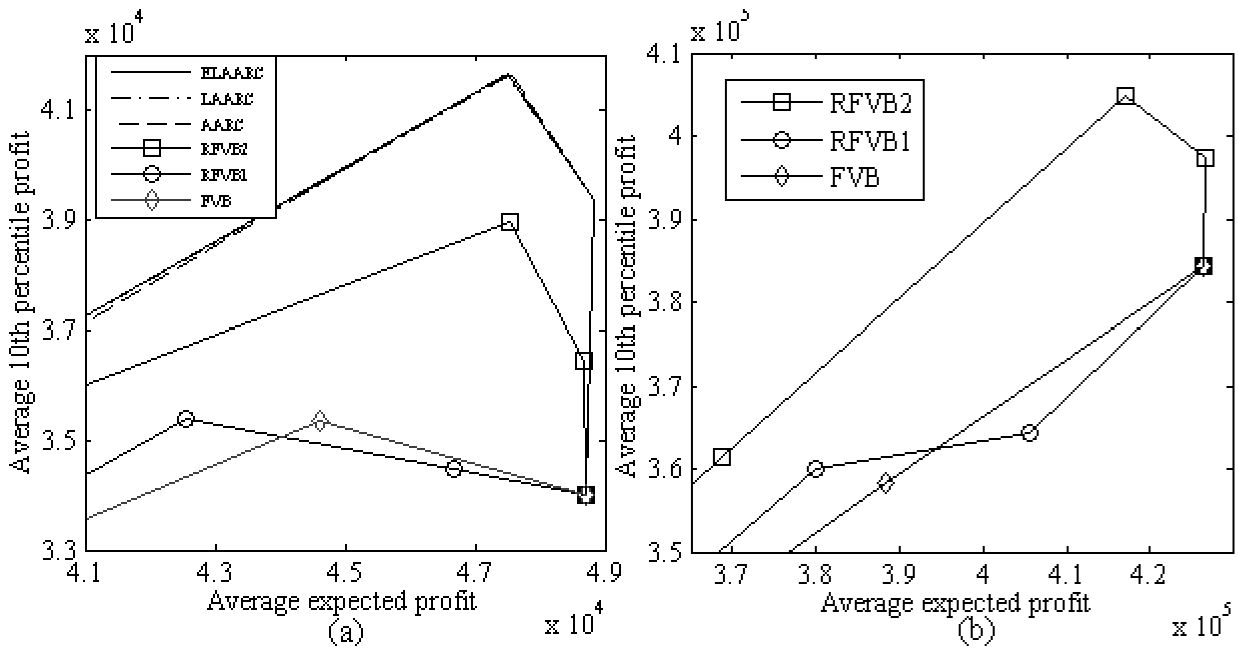

Figure 2–4: Average expected profit versus average 10th percentile profit achieved by the different robust methods on 100 problem instances while adjusting the level of conservativeness Γ. Figure (a) presents the achieved risk-return trade-off for instances of a medium size while (b) presents this for instances of a large size. Note that, in (a), the curves for LAARC and ELAARC were combined since the performances indistinguishable.

2.6.4 Decision Structure

In this subsection, we study the strategies that are obtained from our approximation models. In particular, we look at characteristics such as the number of facilities that are opened and the total production capacities that are installed. To perform this analysis, we replicate the experiments that were done in section 2.6.3 with $L = 10$ and $N = 20$. Statistics of these experiments are reported in Table 2–12. In particular, the table's first set of rows indicates the proportion of problem instances where at least one facility location was proposed for different levels of uncertainty budgets. Once again, the over-conservatism of the FVB model can be observed, as the model refuses to open any facilities in 43% of instances for a relatively small value of $\Gamma = 2$. In contrast, the proportion of problem instances where no facilities are selected is below 15% for all other approximation models. In the other two sets of rows of Table 2–12, we report the number of open facilities and the total capacities of the proposed solution averaged over the instances where at least one facility location was selected. Regarding the strategies proposed by each model, one might notice that more flexible models always propose opening a larger number of facilities. However, the same cannot be said for the total capacity. In fact, it appears that, when $\Gamma = 1$, market-driven models are a bit more cautious with respect to the capacity of its facilities. Increasing the amount of uncertainty has the natural effect of encouraging a smaller number of smaller facilities. It might also be worth emphasizing that, although the FVB model tends to propose the smallest number of facilities, it is misled to promote much larger ones. We believe all these results reaffirm the added value that is obtained by including more flexible policies in the robust optimization model.

We conclude this numerical study with Table 2–13, which describes how much each approximation model is able to cover the realized demand and make efficient use of its capacity as the uncertainty budget Γ is increased. The first observation one can make is that the percentage of covered demand and the percentage of unused capacity displays increased caution, *i.e.,* a decrease of both percentages, as the models account for increased uncertainty through Γ. We also observe that market-driven models have less unused capacity and cover a larger percentage of demand than other models. Among the customer-driven models, the RFVB2 model appears to use a strategy that more closely resembles the strategies of the market-driven models.

Conclusions: In sum, market-driven models propose strategies that open more facilities of smaller size. This strategy seems to allow the decision maker to have more flexibility to choose where goods will be shipped from to meet a given customer's demand. Meanwhile smaller capacities also protect the company from suffering a high rate of unsold products. We also observe that the strategies obtained from market-driven models make better use of the available capacity and cover a larger percentage of the demand than do other models.

Table 2–12: Statistics describing the structure of approximate robust strategies in a set of 100 single-period problem instances with $L = 10$ and $N = 20$.

	Γ	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
# of instances with open facilities	1	75%	88%	93%	95%	95%	95%
	2	57%	86%	90%	94%	94%	94%
	5	10%	85%	85%	91%	91%	91%
Average # of open facility	1	1.56	1.66	1.86	1.86	1.86	1.86
	2	1.21	1.56	1.82	1.81	1.83	1.83
	5	1.20	1.49	1.61	1.62	1.70	1.71
Average total capacity	1	170227	167479	171089	164867	164797	164806
	2	146404	135752	156456	153134	153582	153905
	5	112252	78879	109826	119999	124363	125493

Table 2–13: Proportion of demand that is covered and total capacity that is unused, averaged over the 100 demand scenarios from each problem instance and over a set of 100 problem instances.

Γ	Title	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC
1	Unused capacity (%)	1.18	2.06	1.52	0.74	0.74	0.74
	Covered demand (%)	63.11	72.46	79.25	79.05	79.02	79.03
2	Unused capacity (%)	0.17	0.79	0.74	0.20	0.21	0.22
	Covered demand (%)	41.93	58.15	70.85	73.02	73.21	73.36
5	Unused capacity (%)	0.00	0.00	0.02	0.00	0.00	0.00
	Covered demand (%)	5.80	34.23	47.43	55.55	57.60	58.12

2.7 Conclusion

In this paper, we have studied a multi-period robust location-transportation problem with demand uncertainty that was characterized using the budgeted uncertainty set. In order to overcome the known computational difficulty of solving this model, we presented six new conservative approximation models, each of which implements, to a different extent, the flexibility in the delayed production and transportation decisions. We believe these models, and in particular the RFVB2, ELAARC and HD-ELAARC models, are especially relevant to the transportation literature, given that the only conservative approximation model that had been presented prior to this work was the FVB model, which as demonstrated in Example 2.2 and our empirical results, is overly conservative. While this conservativeness can be easily corrected by adding a small amount of flexibility to the delayed decisions, as is done in the customer-driven RFVB2 model, the solution quality is drastically improved using market-driven models such as the ELAARC. The quality is even further improved using history driven models, *i.e.,* HD-ELAARC, although the number of decision variables in this model quickly becomes prohibitive. As portrayed by Table 2–14, improving solution quality comes at a price in terms of computational requirements. Therefore, we developed a row generation algorithm that enables us to solve market-driven approximations for large instances.

Table 2–14: Summary of the trade-off between flexibility of the adjustments, complexity of the model, and quality of the solution in a multi-period setting. Note that we lack significant evidence about the magnitude of the improvement in quality for HD-ELAARC.

Model	Variables			Total number of variables	Average opt. gap
	P_i^t	Y_{ij}^t	θ_j^t		
HD-ELAARC	$\sum_i Y_{ij}^t$	$\sum_{t'=1}^t \sum_k X_{ijk}^{tt'-}\zeta_k^{t'-} + W_{ij}^t$	$\sum_{t'=1}^t S_j^{tt'-}\zeta_j^{t'-}$	$O(LN^2T^2)$	N/A
ELAARC2	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t-}\zeta_k^{t-} + W_{ij}^t$	$S_j^{t-}\zeta_j^{t-}$	$O(LN^2T)$	$\sim 0.14\%$
LAARC2	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t-}\zeta_k^{t-} + W_{ij}^t$	0	$O(LN^2T)$	$\sim 0.15\%$
AARC	$\sum_i Y_{ij}^t$	$\sum_k X_{ijk}^{t}\zeta_k^{t} + W_{ij}^t$	0	$O(LN^2T)$	$\sim 1.17\%$
RFVB2	$\sum_i Y_{ij}^t$	$X_{ij}^{t+}\zeta_j^{t+} + X_{ij}^{t-}\zeta_j^{t-} + W_{ij}^t$	0	$O(LNT)$	$\sim 0.40\%$
RFVB1	$\sum_i Y_{ij}^t$	$X_{ij}^{t}\zeta_j^{t} + W_{ij}^t$	0	$O(LNT)$	$\sim 1.77\%$
FVB	P_i^t	$X_{ij}^{t}\zeta_j^{t}$	0	$O(LNT)$	$\sim 15.97\%$

A side product of our analysis is to have identified conditions under which full flexibility is not necessary in order to obtain a solution of the best possible quality. This is summarized in Table 2–15. Finally, our numerical study compares the performances of the proposed approximation models in terms of suboptimality of the approximate robust solution, resolution time, achievable risk-return trade-off, and structure of optimal robust decisions.

Although our work focuses on a location-transportation problem, we expect our methods to be applicable to many other multi-stage robust optimization problems with right-hand side uncertainty occurs in the field of transportation, such as network transportation problems,

Table 2–15: Conditions for approximation models to identify optimal robust strategic decisions.

Condition	RC	FVB	RFVB1	RFVB2	AARC	LAARC	ELAARC	HD-ELAARC
$c_0 = 0$	×	×	✓	✓	✓	✓	✓	✓
$\Gamma = 1$	×	×	×	×	×	✓	✓	✓
$\Gamma = N$	✓	×	✓	✓	✓	✓	✓	✓

(*e.g.,* Atamtürk and Zhang (2007)), supply chain network design problems (*e.g.,* Tsiakis et al. (2001)), and hub location-transportation problems (*e.g.,* Oktal and Ozger (2013)).

As a closing remark, one extension of our models that is worth mentioning arises in situations where some facilities may be shut down due to a disruption, such as natural disasters. While we refer the reader to An et al. (2014) and references therein for more details on location-reliability problems, a simple approach consists of considering a set of binary parameters γ_j^t that indicate whether facility j is shut down at time t. One can then replace the maximum production constraint in problem (2.4) with $P_i^t \leq (1-\gamma_i^t)Z_i$, $\forall\, i$, and consider the profit for each period to be a function of $\boldsymbol{I}$, $\boldsymbol{Z}$, $\boldsymbol{\zeta}^t$, and $\boldsymbol{\gamma}^t$. If one assumes that the vector of disruption $\boldsymbol{\gamma}$ lies in a budgeted uncertainty set that is independent from the budgeted uncertainty set used for $\boldsymbol{\zeta}$, then, since $h_t(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}^t, \boldsymbol{\gamma}^t)$ is concave in γ^t for any fixed values of $\boldsymbol{I}$, $\boldsymbol{Z}$, and $\boldsymbol{\zeta}^t$, one can actually relax $\boldsymbol{\gamma}$ to be a vector of fractional value without affecting the model and then employ any version of our different forms of adjustment. For instance, an AARC model would employ the transportation policy $Y_{ij}^t := (\boldsymbol{X}_{ij}^t)^T\boldsymbol{\zeta}^t + (\boldsymbol{O}_{ij}^t)^T\boldsymbol{\gamma}^t + W_{ij}^t$. Alternatively, an ELAARC approach might also employ affine adjustments for penalized excess variables that are used to relax the production constraints. It remains unclear however what might be sufficient conditions for any of these conservative approximations to return exact solutions in this context.

2.8 Appendix

2.8.1 Analytical Solutions to RC and FVB Models in Example 2.2

For the box uncertainty set, the RC model (2.2) takes the following form:

$$\underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{Y},\boldsymbol{P}}{\text{maximize}} \quad \sum_i \sum_j (\eta - d_{ij}) Y_{ij} - \boldsymbol{c}^T \boldsymbol{P} - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \tag{2.19a}$$

$$\text{subject to} \quad \sum_i Y_{ij} \leq \bar{\zeta}_j - \hat{\zeta}_j \, , \forall j \tag{2.19b}$$

$$\sum_j Y_{ij} \leq P_i \, , \forall i \tag{2.19c}$$

$$\boldsymbol{P} \leq \boldsymbol{Z} \, , \ \boldsymbol{Z} \leq M\boldsymbol{I} \tag{2.19d}$$

$$\boldsymbol{Y} \geq 0 \, , \, , \boldsymbol{I} \in \{0, 1\}^L \, . \tag{2.19e}$$

In the optimal solution of RC model (2.19), the value of Y_{ij} is equal to zero, since $\eta - c_i - c_{0i} - d_{ij} < 0$, for all i and j where $i \neq j$, and is equal to $\bar{\zeta}_j - \hat{\zeta}_j = 10000 - 5000 = 5000$ for all i and j when $i = j$. In sequence, the optimal value of variables P_i, Z_i, and I_i are equal to 5000, 5000, and 1 for all i respectively. Therefore, the optimal value of problem (2.19) is equal to 1000. On the other hand, the FVB model (2.5) with box uncertainty set takes the following form

$$\underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{P}}{\text{maximize}} \quad \sum_i \sum_j (\eta - d_{ij})(\bar{\zeta}_j - \hat{\zeta}_j) X_{ij} - \boldsymbol{c}^T \boldsymbol{P} - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \tag{2.20a}$$

$$\text{subject to} \quad \sum_i X_{ij} \leq 1 \, , \forall j \tag{2.20b}$$

$$\sum_j (\bar{\zeta}_j + \hat{\zeta}_j) X_{ij} \leq P_i \, , \forall i \tag{2.20c}$$

$$P_i \leq Z_i \, , \forall i \tag{2.20d}$$

$$X_{ij} \geq 0 \, , \forall i, \forall j \tag{2.20e}$$

$$\boldsymbol{Z} \leq M\boldsymbol{I}, \ \boldsymbol{I} \in \{0, 1\} \, . \tag{2.20f}$$

Similarly to what we concluded above, the optimal solution has $X_{ij} = 0$ for all i and j when $i \neq j$. The optimal value of variable P_i is equal to $(\bar{\zeta}_i + \hat{\zeta}_i) X_{ii}$ for all i, and the optimal value of variable Z_i is equal to that of variable P_i for all i. Therefore, the objective function

(2.20a) can be reformulated as

$$\sum_i \eta(\bar{\zeta}_i - \hat{\zeta}_i)X_{ii} - (c_i + c_{0i})(\bar{\zeta}_i + \hat{\zeta}_i)X_{ii} - f_i 1_{\{X_{ii}>0\}} = \sum_i (2000X_{ii} - 3000 \times 1_{\{X_{ii}>0\}}) \leq 0\,,$$

where the last inequality comes from $\sum_i X_{ij} \leq 1$. Therefore, the optimal value of problem (2.20) is equal to zero in this example.

2.8.2 Selecting Large Enough u for Problem (2.10)

Lemma 2.8.1 *For any* $\boldsymbol{I}$, $\boldsymbol{Z} \geq 0$, $\boldsymbol{\zeta}^t \geq 0$, *the optimal value of problem* (2.10) *is equal to the optimal value of problem* (2.4) *when* $u_j = \max_i(\eta - c_i - d_{ij})\,,\,\forall\, j$.

Proof First, as was argued earlier, in problem (2.4), there is always an optimal solution for which constraint (2.4c) is tight. This implies that the optimal value of problem (2.4) is the same as in

$$
\begin{aligned}
h_t(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}^t) = \max_{\boldsymbol{Y}^t} \quad & \sum_i \sum_j (\eta - d_{ij} - c_i) Y_{ij}^t && (2.21a)\\
\text{subject to} \quad & \sum_i Y_{ij}^t \leq \zeta_j^t\,,\ \forall\, j && (2.21b)\\
& \sum_j Y_{ij}^t \leq \boldsymbol{Z}\ ,\ \forall\, i && (2.21c)\\
& \boldsymbol{Y}^t \geq 0\,. && (2.21d)
\end{aligned}
$$

Now, given that this problem is feasible, strict duality applies, so that its optimal value is equal to the optimal value of the following problem:

$$
\begin{aligned}
h_t(\boldsymbol{I}, \boldsymbol{Z}, \zeta^t) = \min_{\boldsymbol{\lambda}^t, \boldsymbol{\theta}^t} \quad & \sum_i Z_i \theta_i^t + \sum_j \zeta_j^t \lambda_j^t && (2.22a)\\
\text{subject to} \quad & \theta_i^t + \lambda_j^t \geq \eta - d_{ij} - c_i\,,\ \forall\, i\,,\ \forall\, j && (2.22b)\\
& \boldsymbol{\lambda}^t \geq 0,\ \boldsymbol{\theta}^t \geq 0. && (2.22c)
\end{aligned}
$$

where $\boldsymbol{\lambda}^t \in \mathbb{R}^N$ and $\boldsymbol{\theta}^t \in \mathbb{R}^L$ are dual variables for (2.21b) and (2.21c) respectively. It is implied from problem (2.22) that there is an optimal solution for which λ_j^t is smaller or equal to $\max_i(\eta - c_i - d_{ij})$ for all j and t; therefore, one can add to problem (2.22) the constraint that $\lambda_j^t \leq \max_i(\eta - c_i - d_{ij})$ without affecting its optimal value. By applying duality theory a second time, one can easily confirm that he obtains exactly problem (2.10) with $\max_i(\eta - c_i - d_{ij})$ in place of every u_j. Hence, this completes the proof. ∎

2.8.3 Proof of Theorem 2.4.3

2.8.3.1 Proof of Case $c_0 = 0$

First, if all $c_{0i} = 0$, then it is necessarily the case that all Z_i's can be as large as MI_i. Next, we replace variables Y_{ij}^t with $X_{ij}^t \zeta_j^t$ and P_i^t with $\sum_j \zeta_j^t X_{ij}^t$ in the recourse problem (2.4), which makes the recourse problem equivalent to

$$
\begin{aligned}
h_t(\boldsymbol{I}, M\boldsymbol{I}, \zeta^t) := \max_{X^t} \quad & \sum_i \sum_j (\eta - d_{ij} - c_i) X_{ij}^t \zeta_j^t && \text{(2.23a)}\\
\text{subject to} \quad & \sum_i X_{ij}^t \leq 1\,,\ \forall\, j && \text{(2.23b)}\\
& \sum_j X_{ij}^t \zeta_j^t \leq MI_i\,,\ \forall\, i && \text{(2.23c)}\\
& \boldsymbol{X}^t \geq 0\,, && \text{(2.23d)}
\end{aligned}
$$

where $\boldsymbol{X}^t \in \mathbb{R}^{L \times N}$ are the new decision variables for the t-th period. Constraint (2.23c) can be replaced by

$$
X_{ij}^t \leq I_i\,,\ \forall\, i\,,\ \forall j\,, \tag{2.24}
$$

since (2.23c) implies that there can be no shipment when binary variable I_i is equal to 0, and otherwise, the shipment can be as large as M. Therefore, the objective function of the MRLTP can be reformulated as

$$
\begin{aligned}
\min_{\boldsymbol{\zeta} \in \mathcal{D}} \qquad \max_{\boldsymbol{X}} \quad & \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) X_{ij}^t \zeta_j^t && \text{(2.25a)}\\
\text{subject to} \quad & \sum_i X_{ij}^t \leq 1\,,\ \forall\, j\,,\ \forall\, t && \text{(2.25b)}\\
& X^t \leq I_i\,,\ \forall\, i\,,\ \forall j,\ \forall\, t && \text{(2.25c)}\\
& X^t \geq 0\,,\ \forall\, t\,. && \text{(2.25d)}
\end{aligned}
$$

Since both feasible sets for ζ and $\boldsymbol{X}$ are compact, based on Sion's minimax theorem, we can reverse the order of minimization over $\mathcal{D}$ and maximization over $\boldsymbol{X}$, and therefore, problem (2.3) with $c_0 = 0$ can be reduced to

$$\underset{\boldsymbol{I},\boldsymbol{X}}{\text{maximize}} \quad -\boldsymbol{f}^T\boldsymbol{I} + \min_{\boldsymbol{\zeta}\in\mathcal{D}} \sum_t\sum_i\sum_j(\eta - d_{ij} - c_i)X_{ij}^t\zeta_j^t \tag{2.26a}$$

$$\text{subject to} \quad \sum_i X_{ij}^t \leq 1\,,\ \forall\, j\,,\ \forall\, t \tag{2.26b}$$

$$X_{ij}^t \leq I_i\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, t \tag{2.26c}$$

$$\boldsymbol{X}^t \geq 0\,,\ \forall\, t \tag{2.26d}$$

$$\boldsymbol{I} \in \{0,1\}^L\,. \tag{2.26e}$$

where $\boldsymbol{X}^t \in \mathbb{R}^{L\times N}$. Note that problem (2.26) is equivalent to the RFVB1 model when $\boldsymbol{W}$ is fixed to zero in the later one; hence, RFVB1 necessarily achieves an optimal value that is larger since it optimizes over $\boldsymbol{W}$. Given that Proposition 2.4.1 states that RFVB1 optimizes a lower bound on worst-case profit, it is clear that the two models are therefore equivalent. Finally, following Corollary 2.4.2, all tighter approximation models are also equivalent to MRLTP. ∎

2.8.3.2 Proof of Case $\Gamma = NT$

We recall the following theorem from (Ben-Tal et al. 2004).

Theorem 2.8.2 *(Ben-Tal et al. 2004) The adjustable robust counterpart of two-stage robust optimization problem is equivalent to its RC approximation when the uncertainty affecting every one of the constraints is independent of the uncertainty affecting all other constraints (constraint-wise uncertainty).*

For any fixed $\boldsymbol{I}$ and $\boldsymbol{Z}$, the optimal value of the RC model (2.2) can be obtained by solving the following problem:

$$f_{\text{RC}}(\boldsymbol{I},\boldsymbol{Z}) := \max_{\boldsymbol{Y},\boldsymbol{P}} \quad \sum_t\sum_i\sum_j(\eta - d_{ij})Y_{ij}^t - \boldsymbol{c}^T\boldsymbol{P}^t - ({\boldsymbol{c}_0}^T\boldsymbol{Z} + \boldsymbol{f}^T\boldsymbol{I})$$

$$\text{subject to} \quad \sum_i Y_{ij}^t \leq \zeta_j^t\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{D},\ \forall\, j\,,\ \forall\, t$$

$$\sum_j Y_{ij}^t \leq P_i^t \,, \forall i \,, \forall t$$
$$\boldsymbol{P}^t \leq \boldsymbol{Z} \,, \forall t$$
$$\boldsymbol{Y}^t \geq 0 \,, \forall t \,.$$

Noting that, in this problem, when $\mathcal{D}$ is a box uncertainty set, the uncertainty does decompose constraint-wise. Hence, according to Theorem 2.8.2, the optimal value of this problem is equal to the optimal value of the following "wait-and-see" problem:

$$\min_{\boldsymbol{\zeta}\in\mathcal{D}} \; g(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta})$$

where

$$\begin{aligned} g(\boldsymbol{I}, \boldsymbol{Z}, \zeta) := \quad & \max_{\boldsymbol{Y},\boldsymbol{P}} & & \sum_t \sum_i \sum_j (\eta - d_{ij}) Y_{ij}^t - \boldsymbol{c}^T \boldsymbol{P}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\ & \text{subject to} & & \sum_i Y_{ij}^t \leq \zeta_j^t, \forall j \,, \forall t \\ & & & \sum_j Y_{ij}^t \leq P_i^t \,, \forall i \,, \forall t \\ & & & \boldsymbol{P}^t \leq \boldsymbol{Z} \,, \forall t \\ & & & \boldsymbol{Y}^t \geq 0 \,, \forall t \,. \end{aligned}$$

In this problem, all decisions are made once all the information about ζ is obtained. This necessarily leads to an optimal value that is larger than if each $(\boldsymbol{Y}^t, \boldsymbol{P}^t)$ was adjusted only based on the realized ζ^t. We thus conclude that

$$f_{\text{RC}}(\boldsymbol{I}, \boldsymbol{Z}) \leq f_{\text{MRLTP}}(\boldsymbol{I}, \boldsymbol{Z}) \leq \max_{\boldsymbol{\zeta}\in\mathcal{D}} g(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}) = f_{\text{RC}}(\boldsymbol{I}, \boldsymbol{Z}) \,.$$

Furthermore, based on Corollary 2.4.2, RFVB1, RFVB2, AARC, LAARC, ELAARC are optimal in this case and equivalent to the following formulation:

$$\begin{aligned} & \underset{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{Y},\boldsymbol{P}}{\text{maximize}} & & \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) Y_{ij}^t - (\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) \\ & \text{subject to} & & \sum_i Y_{ij}^t \leq \bar{\zeta}_j^t - \hat{\zeta}_j^t, \forall j, \forall t \end{aligned}$$

$$\sum_j Y_{ij}^t \leq Z_i \;,\; \forall\, i,\, \forall t$$
$$\boldsymbol{Y} \geq 0 \,,\, \boldsymbol{I} \in \{0,1\}^L \,. \quad \blacksquare$$

2.8.3.3 Proof of Case $\Gamma = 1$

We start by demonstrating that each $h_t(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}^t)$ is a concave function of $\boldsymbol{\zeta}^t$.

Lemma 2.8.3 *Let $g : \mathbb{R}^m \to \mathbb{R}$ be a function defined as*

$$g(\boldsymbol{x}) := \quad \max_{\boldsymbol{y}\in\mathbb{R}^n} \quad \boldsymbol{c}^T\boldsymbol{y}$$
$$\text{subject to} \quad \boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{x}$$
$$\boldsymbol{y} \in \mathcal{Y}\,,$$

where $\boldsymbol{y} \in \mathbb{R}^n$, for some $\boldsymbol{c} \in \mathbb{R}^n$, some $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, and some compact convex set $\mathcal{Y} \subset \mathbb{R}^n$, and where infeasibility of the optimization problem is interpreted as returning the value $-\infty$. Then, $g(\cdot)$ is a concave function.

Proof Consider two assignments $\boldsymbol{x}_1$ and x_2 for which $g(\boldsymbol{x}_1)$ and $g(\boldsymbol{x}_2)$ are finite valued, we should show that $g(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \geq \theta g(\boldsymbol{x}_1) + (1-\theta)g(\boldsymbol{x}_2)$. To do so, first consider that, since $g(\cdot)$ is finite valued at $\boldsymbol{x}_1$, and $\boldsymbol{x}_2$ and since $\mathcal{Y}$ is compact, there must exist some assignments y_1 and y_2 that respectively achieve the optimum of the optimization problems associated to $g(\boldsymbol{x}_1)$ and $g(\boldsymbol{x}_2)$. Now consider the following:

$$\begin{aligned}
g(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) &= \sup\{c^T\boldsymbol{y} : \boldsymbol{y} \in \mathcal{Y},\, A\boldsymbol{y} \leq \theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2\} \\
&\geq \boldsymbol{c}^T(\theta\boldsymbol{y}_1 + (1-\theta)\boldsymbol{y}_2) \;=\; \theta\boldsymbol{c}^T\boldsymbol{y}_1 + (1-\theta)\boldsymbol{c}^T\boldsymbol{y}_2 \\
&= \theta\sup\{\boldsymbol{c}^T\boldsymbol{y} : \boldsymbol{y} \in \mathcal{Y},\, A\boldsymbol{y} \leq x_1\} + (1-\theta)\sup\{\boldsymbol{c}^T\boldsymbol{y} : \boldsymbol{y} \in \mathcal{Y},\, A\boldsymbol{y} \leq \boldsymbol{x}_2\} \\
&= \theta g(\boldsymbol{x}_1) + (1-\theta)g(\boldsymbol{x}_2)\;,
\end{aligned}$$

where we used the fact that $\boldsymbol{y} := \theta\boldsymbol{y}_1 + (1-\theta)\boldsymbol{y}_2$ is a valid assignment in the first supremum operation since $\mathcal{Y}$ is convex and $\boldsymbol{A}(\theta\boldsymbol{y}_1 + (1-\theta)\boldsymbol{y}_2) = \theta(\boldsymbol{A}\boldsymbol{y}_1) + (1-\theta)\boldsymbol{A}\boldsymbol{y}_2 \leq \theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2$. $\blacksquare$

Since the function $\sum_t h_t(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}^t)$ is jointly concave in $\boldsymbol{\zeta}$ and the budgeted uncertainty set is polyhedral, a worst-case demand necessarily occurs at one of the extreme points of $\mathcal{D}$. There are $2NT + 1$ extreme points in $\mathcal{D}$ when $\Gamma = 1$: *i.e.,* . The nominal demand takes on the role of the first extreme point. In other extreme points, all customers' demand get their nominal value for all periods except for a single customer at a single time period where the demand can be either equal to its largest amount or lowest amount. Let us identify each of these extreme points as $\{(\zeta^t)^{(l,\tau)}\}_{(l,\tau)\in\Omega}$ with $\Omega := \{0, 1, \ldots, 2N\} \times \{1, \ldots, T\}$ and where

$$(\zeta^t)^{(l,\tau)} := \begin{cases} \bar{\zeta}^t & l = 0 \text{ or } \tau \neq t \\ \bar{\zeta}^t + e_l \hat{\zeta}_l^t & l = 1, \ldots, N \\ \bar{\zeta}^t - e_{l-N} \hat{\zeta}_{l-N}^t & l = N+1, \ldots, 2N \end{cases},$$

with e_l as the vector of size N with all elements equal to 0 except for the l-th element, which is equal to 1.

Therefore, for some fixed $\boldsymbol{I}$ and $\boldsymbol{Z}$, and when the budget is equal to one, the optimal value of the MRLTP model is equivalent to

$$f_{\text{MRLTP}}(\boldsymbol{I}, \boldsymbol{Z}) = \max_{\boldsymbol{\zeta} \in \{\zeta^{(l,\tau)}\}_{(l,\tau)\in\Omega}} \sum_t h_t(\boldsymbol{I}, \boldsymbol{Z}, \boldsymbol{\zeta}^t) \,.$$

Following this argument, we have that

$$f_{\text{MRLTP}}(\boldsymbol{I}, \boldsymbol{Z}) = \max_{\boldsymbol{Y},\rho} \quad \rho - \sum_i (c_{0i} Z_i + f_i I_i) \tag{2.27a}$$

$$\text{subject to} \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^t)^{(l,\tau)} \;\; \forall\, (l,\tau) \in \Omega \tag{2.27b}$$

$$\sum_j (Y_{ij}^t)^{(l,\tau)} \leq Z_i \;,\; \forall\, i,\; \forall\, (l,\tau) \in \Omega \tag{2.27c}$$

$$\sum_i (Y_{ij}^t)^{(l,\tau)} \leq (\zeta_j^t)^{(l,\tau)} \;,\; \forall\, j,\; \forall\, (l,\tau) \in \Omega \tag{2.27d}$$

$$(Y_{ij}^t)^{(l,\tau)} \geq 0 \;,\; \forall\, i, \forall\, j,\; \forall\, (l,\tau) \in \Omega \tag{2.27e}$$

$$(Y_{ij}^t)^{(l,\tau)} = (Y_{ij}^t)^{(0,t)} \,,\; \forall\, i,\; \forall\, j,\; \forall\, (l,\tau) \in \Omega \,,\; \forall\, t \neq \tau \,, \tag{2.27f}$$

where $(Y_{ij}^t)^{(l,\tau)} \in \mathbb{R}$ is the recourse decision when scenario (l, τ) occurs, and where the last constraint captures the fact that, in the MRLTP model, the decisions for each $\boldsymbol{Y}^t$ only depend on $\boldsymbol{\zeta}^t$, so that the transportation decision should be the same for all vertices where $\boldsymbol{\zeta}^t = 0$. After replacing the variables $(Y_{ij}^t)^{(l)} := (Y_{ij}^t)^{(l,\tau)} = (Y_{ij}^t)^{(0,t)}$, $\forall\, t \neq \tau$, we alternatively obtain

$$
\begin{aligned}
f_{\text{MRLTP}}(\boldsymbol{I}, \boldsymbol{Z}) = \max_{\boldsymbol{Y},\rho} \quad & \rho - \sum_i (c_{0i} Z_i + f_i I_i) && \text{(2.28a)} \\
\text{subject to} \quad & \rho \leq \sum_{t \neq \tau} \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^t)^{(0)} && \text{(2.28b)} \\
& \qquad + \sum_i \sum_j (\eta - d_{ij} - c_i)(Y_{ij}^\tau)^{(l)} \, , \, \forall\, (l, \tau) \in \Omega \\
& \sum_j (Y_{ij}^t)^{(l)} \leq Z_i \, , \, \forall\, i, \, \forall\, t, \; \forall\, l = 0, \ldots, 2N && \text{(2.28c)} \\
& \sum_i (Y_{ij}^t)^{(l)} \leq (\zeta_j^t)^{(l,t)} \, , \, \forall\, j, \, \forall\, t, \; \forall\, l = 0, \ldots, 2N && \text{(2.28d)} \\
& (Y_{ij}^t)^{(l)} \geq 0 \, , \, \forall\, i, \forall\, j, \, \forall\, t, \; \forall\, l = 0, \ldots, 2N \, . && \text{(2.28e)}
\end{aligned}
$$

Given that, in the LAARC model, the objective function and each robust constraint involve expressions that are linear in ζ, a similar argument as the above can be used to also reformulate this model in terms of vertices of the budgeted uncertainty set. This leads to the following problem:

$$
\begin{aligned}
& f_{\text{LAARC}}(\boldsymbol{I}, \boldsymbol{Z}) = \\
& \max_{\boldsymbol{X}^+, \boldsymbol{X}^-, \boldsymbol{W}, \rho} \quad \rho - \sum_i (c_{0i} Z_i + f_i I_i) \\
& \text{subject to} \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)((\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,\tau)} + W_{ij}^t) \, , \, \forall\, (l, \tau) \in \Omega \\
& \qquad \sum_j ((\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,\tau)} + W_{ij}^t) \leq Z_i \, , \, \forall\, i \, , \, \forall\, t, \, \forall\, (l, \tau) \in \Omega \\
& \qquad \sum_i ((\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,\tau)} + W_{ij}^t) \leq (\zeta_j^t)^{(l,\tau)} \, , \, \forall\, j \, , \, \forall\, t, \, \forall\, (l, \tau) \in \Omega \\
& \qquad (\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,\tau)} + W_{ij}^t \geq 0 \, , \, \forall\, i, \forall\, j, \; \forall\, (l, \tau) \in \Omega \, ,
\end{aligned}
$$

where we characterized the extreme points of $\mathcal{D}_2$ as

$$((\zeta^{t+})^{(l,\tau)}, (\zeta^{t-})^{(l,\tau)}) = \begin{cases} (0,0) & \text{if } t \neq \tau \text{ or } l = 0 \\ (e_l \hat{\zeta}_l^t, 0) & \text{if } t = \tau \text{ and } l = 1, \ldots, N \\ (0, e_{l-N} \hat{\zeta}_l^t) & \text{if } t = \tau \text{ and } l = N+1, \ldots, 2N \end{cases},$$

and with $(\zeta_j^t)^{(l,\tau)} := \bar{\zeta}_j + (\zeta_j^{t+})^{(l,\tau)} - (\zeta_j^{t-})^{(l,\tau)}$.

By exploiting the definition of each $(\zeta^t)^{(l,\tau)}$, one can show that the above equation reduces to

$$\begin{aligned} & f_{\text{LAARC}}(\boldsymbol{I}, \boldsymbol{Z}) = \\ & \max_{\boldsymbol{X}^+, \boldsymbol{X}^-, \boldsymbol{W}, \rho} \quad \rho - \sum_i (c_{0i} Z_i + f_i I_i) \\ & \text{subject to} \\ & \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)((\boldsymbol{X}_{ij}^{\tau+})^T (\zeta^{\tau+})^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{\tau-})^{(l,\tau)}) \\ & \qquad\qquad + (\eta - d_{ij} - c_i) W_{ij}^t \,, \, \forall \, (l, \tau) \in \Omega \\ & \quad \sum_j (\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,t)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,t)} + W_{ij}^t \leq Z_i \,, \, \forall \, i \,, \, \forall \, t \,, \, \forall \, l = 0, \ldots, 2N \\ & \quad \sum_i (\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,t)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,t)} + W_{ij}^t \leq (\zeta_j^t)^{(l,t)} \,, \, \forall \, j \,, \, \forall \, t \,, \, \forall \, l = 0, \ldots, 2N \\ & \quad (\boldsymbol{X}_{ij}^{t+})^T (\zeta^{t+})^{(l,t)} + (\boldsymbol{X}_{ij}^{t-})^T (\zeta^{t-})^{(l,t)} + W_{ij}^t \geq 0 \,, \, \forall \, i, \forall \, j, \, \forall \, l = 0, \ldots, 2N \,. \end{aligned}$$

and further manipulations lead to

$$f_{\text{LAARC}}(\boldsymbol{I}, \boldsymbol{Z}) = \tag{2.30a}$$

$$\max_{\boldsymbol{X}^+, \boldsymbol{X}^-, \boldsymbol{W}, \rho} \quad \rho - \sum_i (c_{0i} Z_i + f_i I_i) \tag{2.30b}$$

$$\text{subject to} \quad \rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t \tag{2.30c}$$

$$\rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t + (\eta - d_{ij} - c_i) X_{ijk}^{\tau+} \hat{\zeta}_k^\tau, \begin{cases} \forall \, k = 1, \ldots, N \\ \forall \, \tau = 1, \ldots, T \end{cases} \tag{2.30d}$$

$$\rho \leq \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i) W_{ij}^t + (\eta - d_{ij} - c_i) X_{ijk}^{\tau-} \hat{\zeta}_k^\tau, \begin{cases} \forall \, k = 1, \ldots, N \\ \forall \, \tau = 1, \ldots, T \end{cases} \tag{2.30e}$$

$$\sum_j W_{ij}^t \leq Z_i\,,\ \forall\, i\,,\ \forall\, t \tag{2.30f}$$

$$\sum_j X_{ijk}^{t+}\hat{\zeta}_k + W_{ij}^t \leq Z_i\,,\ \forall\, i\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N \tag{2.30g}$$

$$\sum_j X_{ijk}^{t-}\hat{\zeta}_k + W_{ij}^t \leq Z_i\,,\ \forall\, i\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N \tag{2.30h}$$

$$\sum_i W_{ij}^t \leq \bar{\zeta}_j\,,\ \forall\, j\,,\ \forall\, t \tag{2.30i}$$

$$\sum_i X_{ijk}^{t+}\hat{\zeta}_k + W_{ij}^t \leq \bar{\zeta}_j + \hat{\zeta}_k 1_{\{j=k\}}\,,\ \forall\, j\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N \tag{2.30j}$$

$$\sum_i X_{ijk}^{t-}\hat{\zeta}_k + W_{ij}^t \leq \bar{\zeta}_j - \hat{\zeta}_k 1_{\{j=k\}}\,,\ \forall\, j\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N \tag{2.30k}$$

$$W_{ij}^t \geq 0\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, t \tag{2.30l}$$

$$X_{ijk}^{t+}\hat{\zeta}_k + W_{ij}^t \geq 0\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N \tag{2.30m}$$

$$X_{ijk}^{t-}\hat{\zeta}_k + W_{ij}^t \geq 0\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, t,\ \forall\, k = 1,\ldots,N\,, \tag{2.30n}$$

where we made use of the fact that

$$(\boldsymbol{X}_{ij}^{t+})^T(\zeta^+)^{(l,\tau)} + (\boldsymbol{X}_{ij}^{t-})^T(\zeta^-)^{(l,\tau)} + W_{ij}^t = \begin{cases} W_{ij} & \text{if } t \neq \tau \text{ or } l = 0 \\ X_{ijl}^{t+}\hat{\zeta}_l + W_{ij} & \text{if } t = \tau \text{ and } l = 1,\ldots,N \\ X_{ij(l-N)}^{t-}\hat{\zeta}_{l-N} + W_{ij} & \text{if } t = \tau \text{ and } l = N+1,\ldots,2N \end{cases}.$$

In problem (2.30), we next reformulate the decision variables W_{ij}, X_{ijl}^+, and $X_{i,j,l-N}^-$ as follows:

$$W_{ij}^t \to \dot{Y}_{ij0}^t,\quad \forall i,\ \forall j,\ \forall t\,,$$

$$X_{ijk}^{t+} \to \frac{\dot{Y}_{ijk}^t - \dot{Y}_{ij0}^t}{\hat{\zeta}_k},\quad \forall i,\ \forall j,\ \forall t,\ \forall k = 1,\ldots,N\,,$$

$$X_{ijk}^{t-} \to \frac{\dot{Y}_{ij(N+k)}^t - \dot{Y}_{ij0}^t}{\hat{\zeta}_k},\quad \forall i,\ \forall j,\ \forall k = 1,\ldots,N\,.$$

Therefore, problem (2.30) can be reformulated as

$$f_{\text{LAARC}}(\boldsymbol{I},\boldsymbol{Z}) = \tag{2.31a}$$

$$\max_{\dot{\boldsymbol{Y}},\rho} \quad \rho - \sum_i (c_{0i}Z_i + f_i I_i) \tag{2.31b}$$

$$\text{subject to} \quad \rho \leq \sum_{t\neq\tau}\sum_i\sum_j (\eta - d_{ij} - c_i)\dot{Y}_{ij0}^t + \sum_i\sum_j (\eta - d_{ij} - c_i)\dot{Y}_{ijl}^\tau\,,\ \forall\,(l,\tau)\in\Omega \tag{2.31c}$$

$$\sum_j \dot{Y}^t_{ijl} \le Z_i \,,\, \forall i \,,\, \forall t \,,\, \forall l = 0, \ldots, 2N \tag{2.31d}$$

$$\sum_i \dot{Y}^t_{ijl} \le \bar{\zeta}_j \,,\, \forall j \,,\, \forall t \tag{2.31e}$$

$$\sum_i \dot{Y}^t_{ijl} \le \bar{\zeta}_j + \hat{\zeta}_j 1_{\{j=l\}} \,,\, \forall j \,,\, \forall t \,,\, \forall l = 1, \ldots, 2N \tag{2.31f}$$

$$\sum_i \dot{Y}^t_{ijl} \le \bar{\zeta}_j - \hat{\zeta}_j 1_{\{j=l-N\}} \,,\, \forall j \,,\, \forall t \,,\, \forall l = N+1, \ldots, 2N \tag{2.31g}$$

$$\dot{Y}^t_{ijl} \ge 0 \,,\, \forall i \,,\, \forall j \,,\, \forall t \,,\, \forall l = 0, \ldots, 2N \,. \tag{2.31h}$$

A careful comparison of problems (2.28) and (2.31) can confirm that these are the same, so they will return the same optimal value and identify the same set of optimal solutions for $\boldsymbol{Z}$ and $\boldsymbol{I}$.

2.8.4 Proof of Theorem 2.5.1

We first derive the robust counterpart of constraint (2.12b) as

$$\exists \boldsymbol{O} \in \mathbb{R}^{N \times T}, \boldsymbol{Q} \in \mathbb{R}^{N \times N \times T},$$

$$\sum_i W^t_{ij} + \Gamma O^t_j + \sum_k Q^t_{jk} \le \bar{\zeta}_{jt} \,,\, \forall j \,,\, \forall t \tag{2.32a}$$

$$O^t_j + Q^t_{jj} \ge \hat{\zeta}_{jt}(1 - S^{t-}_j + \sum_i X^t_{ijj}) \,,\, \forall j \,,\, \forall t \tag{2.32b}$$

$$O^t_j + Q^t_{jk} \ge \hat{\zeta}_{kt} \sum_i X^t_{ijk}, \; \forall j, \forall k \neq j \,,\, \forall t \tag{2.32c}$$

$$\boldsymbol{O} \ge 0 \,,\; \boldsymbol{Q} \ge 0 \,, \tag{2.32d}$$

where $\forall k$ refers to $\forall k = 1, \ldots, N$, as will continue to be the case below. The condition described in (2.32a)-(2.32d) can be considered equivalent to the original constraint, given that strict duality applies, since $\mathcal{D}_3$ is non-empty when $\Gamma \ge 0$.

Similarly, we can derive the robust counterpart of constraint (2.12c) as

$$\exists \boldsymbol{E} \in \mathbb{R}^{L \times T}, \boldsymbol{F} \in \mathbb{R}^{L \times N \times T},$$

$$\sum_j W^t_{ij} + \Gamma E^t_i + \sum_k F^t_{ik} \le Z_i \,,\, \forall i \,, \forall t \tag{2.33a}$$

$$E^t_i + F^t_{ik} \ge \hat{\zeta}^t_k \sum_j X^{t-}_{ijk} \,,\, \forall i \,,\, \forall k \,,\, \forall t \tag{2.33b}$$

$$\boldsymbol{E} \ge 0 \,,\; \boldsymbol{F} \ge 0 \,, \tag{2.33c}$$

and the robust counterpart of constraint (2.12d) as

$$\exists \boldsymbol{G} \in \mathbb{R}^{L\times N\times T}, \boldsymbol{H} \in \mathbb{R}^{L\times N\times N\times T},$$

$$-W_{ij}^t + \Gamma G_{ij}^t + \sum_k H_{ijk}^t \leq 0\,,\ \forall\, i\,,\ \forall\, j\,, \forall\, t \tag{2.34a}$$

$$G_{ij}^t + H_{ijk}^t \geq -\hat{\zeta}_k^t X_{ijk}^{t-}\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, k, \forall t \tag{2.34b}$$

$$\boldsymbol{G} \geq 0\,,\ \boldsymbol{H} \geq 0\,, \tag{2.34c}$$

and finally, the robust counterpart of constraint (2.12e) as

$$S_j^{-t} \geq 0, \forall j\,,\ \forall\, t\,. \tag{2.35a}$$

Therefore the reduced ELAARC can be reformulated as

$$\underset{\substack{\boldsymbol{I},\boldsymbol{Z},\boldsymbol{X}^-,\boldsymbol{W},\boldsymbol{S}^- \\ \boldsymbol{O},\boldsymbol{Q},\boldsymbol{E},\boldsymbol{F},\boldsymbol{G},\boldsymbol{H}}}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}^- \in \mathcal{D}_3} \sum_t \sum_i \sum_j (\eta - d_{ij} - c_i)(\sum_k X_{ijk}^{t-}\zeta_k^{t-} + W_{ij}^t)$$

$$-(\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) - \sum_t \sum_j u_j S_j^{t-} \zeta_j^{t-} \tag{2.36a}$$

$$(2.32\text{a}) - (2.32\text{d})\,,\ (2.33\text{a}) - (2.33\text{c})\,,\ (2.34\text{a}) - (2.34\text{c})\,,\ (2.35\text{a}) \tag{2.36b}$$

$$\boldsymbol{Z} \leq M\boldsymbol{I},\ \boldsymbol{I} \in \{0,1\}^L. \tag{2.36c}$$

Since $\mathcal{D}_3$ is compact and convex, one can apply Sion's minimax theorem to reverse the order of maximization over $\{\boldsymbol{X}^-, \boldsymbol{W}, \boldsymbol{S}^-, \boldsymbol{O}, \boldsymbol{Q}, \boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}\}$ with the minimization over ζ^- and then replace the inner maximization by its dual minimization problem. The dual minimization problem joined with the minimization with respect to ζ^- leads to the following optimization model:

$$\begin{aligned}
\min_{\substack{\boldsymbol{\delta}^-,\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\psi} \\ \boldsymbol{\Theta},\boldsymbol{\Lambda},\boldsymbol{\Psi}}} \quad & -(\boldsymbol{c}_0{}^T \boldsymbol{Z} + \boldsymbol{f}^T \boldsymbol{I}) + \sum_t \sum_i Z_i \theta_{it} + \sum_t \sum_j \lambda_{jt} \bar{\zeta}_{jt} - \sum_t \sum_j \Lambda_{jjt} \hat{\zeta}_{jt} \\
\text{subject to} \quad & \theta_i^t + \lambda_j^t - \psi_{ij}^t = \eta - c_i - d_{ij}\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, t \\
& \Theta_{ik}^t + \Lambda_{jk}^t - \Psi_{ijk}^t = (\eta - c_i - d_{ij})\delta_k^{t-}\,,\ \forall\, i\,,\ \forall\, j\,,\ \forall\, k\,,\ \forall\, t \\
& \sum_k \Theta_{ik}^t \leq \Gamma \theta_i^t, \Theta_{ik}^t \leq \theta_i^t\,,\ \forall\, i\,,\ \forall\, k\,,\ \forall\, t
\end{aligned}$$

$$\sum_k \Lambda_{jk}^t \leq \Gamma\lambda_j^t, \Lambda_{jk}^t \leq \lambda_j^t,\ \Lambda_{jk}^t \leq B_j\delta_j^{t-}\ , \forall\, j\,, \forall\, k\,, \forall\, t$$

$$\sum_k \Psi_{ijk}^t \leq \Gamma\psi_{ij}^t, \Psi_{ijk}^t \leq \psi_{ij}^t\,, \forall\, i\,, \forall\, j\,, \forall\, k\,, \forall\, t$$

$$0 \leq \boldsymbol{\delta}^- \leq 1\,,\ \sum_t \sum_j \delta_j^{t-} = \Gamma$$

$$\boldsymbol{\lambda} \geq 0\,,\ \boldsymbol{\Lambda} \geq 0\,,\ \boldsymbol{\theta} \geq 0\,,\ \boldsymbol{\Theta} \geq 0\,,\ \boldsymbol{\psi} \geq 0\,,\ \boldsymbol{\Psi} \geq 0\,,$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{N\times T}$, $\boldsymbol{\Lambda} \in \mathbb{R}^{N\times N\times T}$, $\boldsymbol{\theta} \in \mathbb{R}^{L\times T}$, $\boldsymbol{\Theta} \in \mathbb{R}^{L\times N\times T}$, $\boldsymbol{\psi} \in \mathbb{R}^{L\times N\times T}$, and $\boldsymbol{\Psi} \in \mathbb{R}^{L\times N\times N\times T}$ are the dual variables associated with constraints (2.32a), (2.32c)-(2.32d), (2.33a), (2.33b), (2.34a), and (2.34b) respectively.

Next, one can further reduce this optimization problem by replacing $\psi_{ij}^t = \theta_i^t + \lambda_j^t - (\eta - c_i - d_{ij})$ and $\Psi_{ijk}^t = \Theta_{ik}^t + \lambda_{jk}^t - (\eta - c_i - d_{ij})\delta_k^{t-}$ everywhere and obtain the model presented in the theorem. It is worth emphasizing that this replacement of variables reduces the rate of growth of the total number of decision variables of the model to $O(LNT)$ instead of $O(LN^2T)$. ∎

Chapter 3
Linearized Robust Counterparts of Two-stage Robust Optimization Problems

Amir Ardestani-Jaafari, Erick Delage

Department of Decision Sciences, HEC Montréal, Montréal, Quebec, H3T 2A7, Canada

amir.ardestani-jaafari@hec.ca, erick.delage@hec.ca

Abstract

In this short article, we discuss an alternative method for deriving conservative approximation models for two-stage robust optimization problems. The method extends in a natural way a linearization scheme that was recently proposed to construct tractable reformulations for robust static problems involving functions that decompose as a sum of piecewise linear convex expressions. Given that this generalized method mainly relies on a linearization scheme employed in bi-linear optimization problems, we will say that it gives rise to the "linearized robust counterpart" model. We identify a close relation between this linearized robust counterpart model and the more popular affinely adjustable robust counterpart model. We also describe a very simple way of modifying both types of models in order to make these approximations less conservative. We finally demonstrate how to employ this new scheme in a set logistic application problems in order to improve the performance and guarantees of robust optimization.

Keywords

Two-stage adjustable robust optimization, linear programming relaxation, affinely adjustable robust counterpart, bilinear optimization.

3.1 Introduction

Classical robust optimization (RO) assumes that all decisions are here-and-now, *i.e.,* they must be made before the realization of uncertainty. However this assumption is not realistic in many real-world problems. In many location-transportation problems, for example, transportation decisions can be delayed until the uncertain demand of customers is revealed. To address the uncertainty in such problems, Ben-Tal et al. (2004) introduced an adjustable robust optimization (ARO) problem that takes the following form on a two-stage setting when the uncertainty can be captured in the right-hand side of the constraint set:

$$\text{(ARO)} \quad \underset{x\in\mathcal{X},y(\zeta)}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) \tag{3.1a}$$

$$\text{subject to} \quad Ax + By(\zeta) \leq \Psi(x)\zeta \,,\; \forall\, \zeta \in \mathcal{U}\,, \tag{3.1b}$$

where $A \in \mathbb{R}^{m\times n_x}$, $B \in \mathbb{R}^{m\times n_y}$, $c \in \mathbb{R}^{n_x}$, $d \in \mathbb{R}^{n_y}$, $\Psi(x) : \mathbb{R}^{n_x} \to \mathbb{R}^{m\times n_\zeta}$ such that $\Psi(x)$ is an affine mapping of variable x, where $\mathcal{U}$ is an uncertainty set for ζ, and where $y : \mathbb{R}^{n_\zeta} \to \mathbb{R}^{n_y}$ is an adjustable decision. Since problem (3.1) is shown to be computationally intractable, Ben-Tal et al. suggested instead solving the affinely adjustable robust counterpart (AARC) of the problem, wherein adjustable decisions are forced to be affine adjustments of the observed uncertain vector ζ, *i.e.*, $y(\zeta) := Y\zeta + y$, for some $Y \in \mathbb{R}^{n_y\times n_\zeta}$ and $y \in \mathbb{R}^{n_y}$; therefore, problem (3.1) is conservatively approximated with

$$\text{(AARC)} \quad \underset{x\in\mathcal{X},y,Y}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} \; c^T x + d^T (Y\zeta + y) \tag{3.2a}$$

$$\text{subject to} \quad Ax + B(Y\zeta + y) \leq \Psi(x)\zeta,\; \forall\, \zeta \in \mathcal{U}. \tag{3.2b}$$

In a seemingly unrelated article, Ardestani-Jaafari and Delage (2016a) recently proposed a scheme for creating conservative approximation models for static robust optimization problems that involve a sum of piecewise linear concave functions. Specifically, they study the problem of

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad g(x) := \min_{\zeta\in\mathcal{U}} \; \sum_{i=1}^{N} \min_{k=1,\ldots,K} c_\zeta^{i,k}(x)^T\zeta + d_\zeta^{i,k}(x)\;, \tag{3.3}$$

where $c_\zeta^{i,k} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_\zeta}$ and $d_\zeta^{i,k} : \mathbb{R}^{n_x} \to \mathbb{R}$ are affine mappings of variable x. Broadly speaking, this scheme consists in linearizing the objective function of a mixed-integer programming representation of the worst-case analysis problem before relaxing the integrality constraints as shown below:

$$g(x) = \begin{array}{rl} \min\limits_{\zeta\in\mathcal{U},\{\lambda^k\}_{k=1}^K} & \sum\limits_{ik} c_\zeta^{i,k}(\boldsymbol{x})^T \zeta \lambda_k^i + d_\zeta^{i,k}(\boldsymbol{x})\lambda_k^i \\ \text{subject to} & \sum\limits_k \lambda_k^i = 1\,,\,\forall\, i = 1,\ldots,N \\ & \lambda^k \in \{0,1\}^N\,,\,\forall\, k \end{array} \geq \begin{array}{rl} \min\limits_{\zeta\in\mathcal{U},\{\lambda^k\}_{k=1}^K,\{\Delta^{ik}\}_{ik}} & \sum\limits_{ik} \boldsymbol{c}_\zeta^{i,k}(\boldsymbol{x})^T \Delta^{ik} + d_\zeta^{i,k}(\boldsymbol{x})\lambda_k^i \\ \text{subject to} & \sum\limits_k \lambda_k^i = 1\,,\,\forall\, i = 1,\ldots,N \\ & \lambda^k \in [0,1]^N\,,\,\forall\, k \\ & \sum\limits_k \Delta^{ik} = \zeta_i\,,\,\forall\, i\,, \end{array}$$

where each $\Delta^{ik} \in \mathbb{R}^{n_\zeta}$ is introduced to capture the relation $\Delta^{ik} = \zeta\lambda_i^k$. Note that the right-hand side model is a linear program that will lead to a compact conservative approximation of problem (3.3) when the dual maximization model associated with this linear program is reintroduced into problem (3.3). The authors show that, for a specific choice of a mixed-integer programming formulation, the resulting "linearized robust counterpart" (LRC) model is equivalent to employing affine adjustments in the following two-stage representation of (3.3):

$$\begin{aligned} & \underset{x\in\mathcal{X},y(\zeta)}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} \sum_{i=1}^N y_i(\zeta) && (3.4a) \\ & \text{subject to} \quad y_i(\zeta) \leq \boldsymbol{c}_\zeta^{i,k}(\boldsymbol{x})^T\zeta + d_\zeta^{i,k}(\boldsymbol{x})\,,\,\forall\, i\,,\,\forall\, k\,,\,\forall\,\zeta\in\mathcal{U}\,. && (3.4b) \end{aligned}$$

They also show that, under some conditions on the structure of the uncertainty set and of the mappings $\boldsymbol{c}_\zeta^{i,k}(\cdot)$ and $d_\zeta^{i,k}(\cdot)$, this conservative approximation is exact. Furthermore, they propose a way of improving this approximation using a semi-definite programming formulation.

In this paper, we extend the scope of the linearization scheme proposed in Ardestani-Jaafari and Delage (2016a) to the set of two-stage ARO problems that take the form described in (3.1). In doing so, we offer the following contributions:

- We introduce a new scheme for constructing conservative approximation models (called linearized robust counterpart models) of two-stage ARO problems with right-hand side uncertainty.

- We establish a new interpretation for the conservative approximation models obtained by employing AARC on a two-stage ARO problem. This interpretation will be based on popular relaxation methods that are used (*e.g.,* in Sherali and Alameddine (1992)) for approximating bilinear optimization problems.
- We provide a methodology to improve LRC, using linear and conic valid inequalities. This will lead to a simple procedure that can be used to improve the approximation obtained using AARC.
- Finally, we discuss the application of the LRC and AARC models in three types of logistics problems wherein it is possible either to demonstrate exactness of these methods or improve the quality of their solution using valid inequalities.

The remainder of the paper is organized as follows. In Section 3.2, we introduce the linearized robust counterpart model associated to an ARO problem with a polyhedral uncertainty set and we define the notion of a relaxation gap that can be used to bound the suboptimality of solutions. Next in in Section 3.3, we establish the equivalence between LRC and AARC models. We follow in Section 3.4 with the description of two methods that can be used to tighten the approximation obtained through LRC or AARC. These methods are heavily inspired from the use of linear and conic valid inequality in the process of linearization of bilinear problems. Section 3.5 briefly describes how one might extend our results to general convex sets. Finally, we present three applications to practical logistics problems in Section 3.6, and we conclude in Section 3.7.

3.2 The linearized robust counterpart model

In order to present the LRC model, we need to make the following three assumptions.

Assumption 1 *Let $\mathcal{U}$ be a bounded and non-empty polyhedral set defined as $\mathcal{U} := \{\zeta | P\zeta \leq q\}$ where $P \in \mathbb{R}^{n_{\mathcal{U}} \times n_\zeta}, q \in \mathbb{R}^{n_{\mathcal{U}}}$.*

Assumption 2 *Let the ARO model possess relatively complete recourse, namely, that*

$$\forall x \in \mathcal{X}\,,\, \exists y(\zeta) :\, Ax + By(\zeta) \leq \Psi(x)\zeta \quad \forall \zeta \in \mathcal{U}\,.$$

Assumption 3 *For all $x \in \mathcal{X}$, there exists a feasible ζ, such that the recourse problem is bounded. In other words, let problem* (3.1) *be bounded.*

The three assumptions described above should not be considered limiting. Considering Assumption 1, it is typically the case that $\mathcal{U}$ includes, at very least, a nominal, or most-likely, scenario, or most likely, scenario, and that all possible scenarios reside in a bounded set. Satisfying Assumption 2 is mostly a matter of formulating $\mathcal{X}$ so that it does not include any solutions for which there might be no feasible second-stage solutions, a situation that is typically associated with an infinite loss. Finally, it is reasonable to assume that problem (3.1) is bounded in realistic practical problems.

Let us now consider the fact that the ARO model can be formulated as

$$\underset{x \in \mathcal{X}}{\text{maximize}} \quad g(x) \tag{3.5a}$$

where $g(x)$ is defined as

$$\begin{aligned} g(x) := \min_{\zeta \in \mathcal{U}} \quad & \max_{y} \quad c^T x + d^T y && (3.6a)\\ & \text{subject to} \quad Ax + By \leq \Psi(x)\zeta \,. && (3.6b) \end{aligned}$$

Based on Assumption 2, one can apply duality theory on the inner maximization problem to show that $g(x)$ is exactly equal to

$$\begin{aligned} g(x) = \quad & \min_{\zeta, \lambda} \quad c^T x + (\Psi(x)\zeta)^T \lambda - (Ax)^T \lambda && (3.7a)\\ & \text{subject to} \quad B^T \lambda = d && (3.7b)\\ & \quad P\zeta \leq q && (3.7c)\\ & \quad \lambda \geq 0 \,, && (3.7d) \end{aligned}$$

where $\lambda \in \mathbb{R}^m$ is the dual variable associated with constraint (3.6b).

Lemma 3.2.1 *Problem* (3.7) *possesses a feasible solution.*

Proof Assumption 3 guarantees that, for all $x \in \mathcal{X}$, there exists a feasible $\bar{\zeta}$ for which the maximization problem in y has a finite optimal value. By the strong duality property, this indicates that, for this same $\bar{\zeta}$, the minimization problem in λ also has a finite optimal value and must therefore have a feasible solution $\bar{\lambda}$. Together, the pair $(\bar{\zeta}, \bar{\lambda})$ constitutes a feasible solution for problem (3.7). ∎

In Sherali and Alameddine (1992), the authors employ a linearization scheme that exploits a set of valid inequalities for a bilinear optimization problem similar to problem (3.7). In the context that we study here, this scheme leads us to consider that

$$
\begin{aligned}
g(x) = \quad \min_{\zeta,\lambda} \quad & c^T x + \operatorname{tr}(\Psi(x)\zeta\lambda^T) - (Ax)^T\lambda && \text{(3.8a)}\\
\text{subject to} \quad & B^T\lambda = d && \text{(3.8b)}\\
& P\zeta \le q && \text{(3.8c)}\\
& \lambda \ge 0 && \text{(3.8d)}\\
& \zeta\lambda^T B = \zeta d^T && \text{(3.8e)}\\
& P\zeta\lambda^T \le q\lambda^T && \text{(3.8f)}\\
& B^T\lambda\lambda^T = d\lambda^T && \text{(3.8g)}\\
& \lambda\lambda^T \ge 0\,, && \text{(3.8h)}
\end{aligned}
$$

where $\operatorname{tr}(\cdot)$ stands for the trace operator, and where, for any two matrices A and B of the same dimension $n \times m$, a constraint $A \le B$ stands for $A_{ij} \le B_{ij}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$, and similarly for the constraint $A = B$. Note that, in this model, constraints (3.8e) to (3.8h) are a set of redundant constraints that were added to problem (3.7). In particular, constraint (3.8e) is implied from

$$\text{(3.7b)} \Rightarrow B^T\lambda\zeta^T = d\zeta^T \Rightarrow \zeta\lambda^T B = \zeta d^T\,.$$

Moreover, constraints (3.8g)-(3.8h) can be similarly derived:

$$\text{(3.7c)\&(3.7d)} \Rightarrow P\zeta\lambda^T \le q\lambda^T\,,$$

$$\text{(3.7b)} \Rightarrow B^T\lambda\lambda^T = d\lambda^T\,,$$

$$(3.7\text{d}) \Rightarrow \lambda\lambda^T \geq 0\,.$$

We next linearize problem (3.8) by introducing the variables $\Delta \in \mathbb{R}^{n_\zeta \times m}$ and $\Lambda \in \mathbb{R}^{m \times m}$, respectively defined as $\Delta := \zeta\lambda^T$ and $\Lambda := \lambda\lambda^T$, such that

$$
\begin{aligned}
g(x) = \quad & \min_{\zeta,\lambda,\Delta,\Lambda} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda && \text{(3.9a)}\\
& \text{subject to} && B^T\lambda = d && \text{(3.9b)}\\
& && P\zeta \leq q && \text{(3.9c)}\\
& && \lambda \geq 0 && \text{(3.9d)}\\
& && \Delta B = \zeta d^T && \text{(3.9e)}\\
& && P\Delta \leq q\lambda^T && \text{(3.9f)}\\
& && B^T\Lambda = d\lambda^T && \text{(3.9g)}\\
& && \Lambda \geq 0 && \text{(3.9h)}\\
& && \Lambda = \lambda\lambda^T && \text{(3.9i)}\\
& && \Delta = \zeta\lambda^T\,. && \text{(3.9j)}
\end{aligned}
$$

A simple relaxation of problem (3.9) will lead to the linearized robust counterpart model for problem (3.1).

Proposition 3.2.2 *The following linearized robust counterpart model is a conservative approximation of problem* (3.1)*:*

$$
\begin{aligned}
\textit{(LRC)} \quad & \underset{x\in\mathcal{X},Y,y,S,s}{\text{maximize}} && c^T x + d^T y - q^T s && \text{(3.10a)}\\
& \text{subject to} && P^T S = Y^T B^T - \Psi(x)^T && \text{(3.10b)}\\
& && Ax + By + S^T q \leq 0 && \text{(3.10c)}\\
& && P^T s = -Y^T d && \text{(3.10d)}\\
& && s \geq 0\,,\ S \geq 0\,, && \text{(3.10e)}
\end{aligned}
$$

where $Y \in \mathbb{R}^{n_y \times n_\zeta}, y \in \mathbb{R}^{n_y}, S \in \mathbb{R}^{n_\mathcal{U} \times m}$*, and* $s \in \mathbb{R}^{n_\mathcal{U}}$*.*

Proof We start that relaxing problem (3.9) by removing constraint (3.9j) to get a lower bound for $g(x)$. Next, we consider that since, when constraints (3.9b) to (3.9f) are satisfied, one can simply let $\hat{\Lambda} := \hat{\lambda}\hat{\lambda}^T$ in order to satisfy constraints (3.9g)-(3.9i), the problem stays the same when disregarding Λ and the three constraints (3.9g)-(3.9i). Hence, we obtain a lower bound for $g(x)$ in the form

$$
\begin{aligned}
g(x) \geq g_{LRC}(x) := \quad & \min_{\zeta,\lambda,\Delta} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && \text{(3.11a)} \\
& \text{subject to} && B^T \lambda = d && \text{(3.11b)} \\
& && P\zeta \leq q && \text{(3.11c)} \\
& && \Delta B = \zeta d^T && \text{(3.11d)} \\
& && P\Delta \leq q\lambda^T && \text{(3.11e)} \\
& && \lambda \geq 0\,. && \text{(3.11f)}
\end{aligned}
$$

Since, based on Lemma 3.2.1, there exists a solution $(\hat{\zeta}, \hat{\lambda})$ that satisfies constraints (3.11b), (3.11c), and (3.11f), one can confirm that the triplet $(\hat{\zeta}, \hat{\lambda}, \hat{\Delta})$, with $\hat{\Delta} := \hat{\zeta}\hat{\lambda}^T$, is a feasible solution to problem (3.11). Hence, strong duality applies for problem (3.11) so that it can be equivalently represented as

$$
\begin{aligned}
g_{LRC}(x) = \quad & \max_{Y,y,S,s} && c^T x + d^T y - q^T s && \text{(3.12a)} \\
& \text{subject to} && P^T S = Y^T B^T - \Psi(x)^T && \text{(3.12b)} \\
& && Ax + By + S^T q \leq 0 && \text{(3.12c)} \\
& && -P^T s = Y^T d && \text{(3.12d)} \\
& && s \geq 0\,,\ S \geq 0\,, && \text{(3.12e)}
\end{aligned}
$$

where the variables $y \in \mathbb{R}^{n_y}$, $s \in \mathbb{R}^{n_\mathcal{U}}$, $Y \in \mathbb{R}^{n_y \times n_\zeta}$, and $S \in \mathbb{R}^{n_\mathcal{U} \times m}$ are the dual variables associated with constraints (3.11b), (3.11c), (3.11d), and (3.11e) respectively. We next combine problem (3.12) with the maximization over variable $x \in \mathcal{X}$, which leads to LRC (3.10). ∎

In order to provide a relation between the optimal solution of ARO and the optimal solution of LRC, we first introduce the notion of a relaxation gap.

Definition Assuming that, for all $x \in \mathcal{X}$, $g_{LRC}(x)$ is positive definite, *i.e.*, $g_{LRC}(x) \geq 0$, the relaxation gap γ of $g_{LRC}(x)$ is defined as the bound on the largest achievable ratio between the value obtained by $g_{LRC}(x)$ and the one obtained by $g(x)$ for any feasible solution x, *i.e.*,

$$g_{LRC}(x) \leq g(x) \leq \gamma g_{LRC}(x)\,, \forall\, x \in \mathcal{X}\,. \tag{3.13}$$

Next, we establish that, if one can identify a bound on the relaxation gap of $g_{LRC}(x)$, then this automatically provides a bound on the suboptimality of the solution returned by this model.

Proposition 3.2.3 *Let $\hat{x}$ and $g_{LRC}(\hat{x})$ respectively be the optimal solution and optimal value of LRC* (3.10)*; then, both the actual worst-case value of $\hat{x}$, i.e., $g(\hat{x})$, and the estimated worst-case value $g_{LRC}(\hat{x})$ are less than a factor of γ away from the optimal value of problem* (3.1)*, where γ is the relaxation gap of $g_{LRC}(x)$.*

Proof Given that x^* is the optimal solution of problem (3.1) and $\hat{x}$ is the optimal solution of problem (3.10), this indicates that

$$g_{LRC}(x^*) \leq g_{LRC}(\hat{x}) \leq g(\hat{x}) \leq g(x^*)\,, \tag{3.14}$$

where the second inequality is implied from the fact that $g_{LRC}(x)$ is a conservative approximation of $g(x)$ for all $x \in \mathcal{X}$. Combining (3.13) and (3.14) results in

$$g_{LRC}(x^*) \leq g_{LRC}(\hat{x}) \leq g(\hat{x}) \leq g(x^*) \leq \gamma g_{LRC}(x^*) \leq \gamma g_{LRC}(\hat{x})\,,$$

and consequently, it results in

$$g_{LRC}(\hat{x}) \leq g(\hat{x}) \leq g(x^*) \leq \gamma g_{LRC}(\hat{x})\,. \quad \blacksquare$$

Furthermore, while it is known that evaluating the worst-case value of a given first-stage solution, *i.e.,* evaluating $g(x)$, is computationally intractable, we can show that it is possible to efficiently evaluate a bound on the worst-case value of the solution of the LRC model.

Lemma 3.2.4 *Given that $\hat{x}$ is the optimal solution of LRC* (3.10) *and ($\hat{\lambda}$, $\hat{\zeta}$, $\hat{\Delta}$) is the optimal solution of problem* (3.11)*, when x is fixed to $\hat{x}$, we have that*

$$0 \leq g(\hat{x}) - g_{LRC}(\hat{x}) \leq \hat{\lambda}^T \Psi(x)\hat{\zeta} - \mathrm{tr}(\Psi(x)\hat{\Delta})$$

Proof Given that $(\hat{\lambda},\, \hat{\zeta})$ is a feasible solution of problem (3.6) when x is fixed to $\hat{x}$, and since $g_{LRC}(\hat{x}) = c^T\hat{x} + \sum_{ij} \Psi(x)_{ij}\hat{\Delta}_{ji} - (A\hat{x})^T\hat{\lambda}$, this indicates that

$$0 \leq g(\hat{x}) - g_{LRC}(\hat{x}) \leq (\Psi(x)\hat{\zeta})^T\hat{\lambda} - \mathrm{tr}(\Psi(x)\hat{\Delta})\,. \qquad \blacksquare$$

3.3 Relation to AARC

In this section, we explain how the LRC model can be considered equivalent to the conservative approximation model obtained with AARC.

Proposition 3.3.1 *LRC* (3.10) *is equivalent to AARC* (3.2)*.*

Proof Proof. As a first step, we reformulate problem (3.11) in terms of an inner and an outer minimization operation:

$$\begin{aligned} g_{LRC}(x) = \min_{\zeta\in\mathcal{U}} \quad & \min_{\lambda,\Delta} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda && (3.15a)\\ & \text{subject to} && B^T\lambda = d && (3.15b)\\ & && \Delta B = \zeta d^T && (3.15c)\\ & && P\Delta \leq q\lambda^T && (3.15d)\\ & && \lambda \geq 0\,. && (3.15e) \end{aligned}$$

We next derive the dual formulation of the inner minimization over λ and Δ as

$$\begin{aligned} \max_{y,Y,S} \quad & c^T x + d^T(y + Y\zeta) && (3.16a)\\ \text{subject to} \quad & Ax + By + S^T q \leq 0 && (3.16b) \end{aligned}$$

$$P^T S = Y^T B^T - \Psi(x)^T \tag{3.16c}$$

$$S \geq 0\,, \tag{3.16d}$$

where $y \in \mathbb{R}^{n_y}, Y \in \mathbb{R}^{n_y \times n_\zeta}$ and $S \in \mathbb{R}^{n_\mathcal{U} \times m}$ are the dual variables associated with constraints (3.15b), (3.15c), and (3.15d) respectively. Based on Sion's minimax theorem, since $\mathcal{U}$ is bounded, the same value for $g_{LRC}(x)$ can be obtained by reversing the order of minimization over ζ and maximization over y, Y, and S. Equivalently, we have that

$$g_{LRC}(x) = \quad \max_{y,Y,S} \quad \min_{\zeta \in \mathcal{U}} c^T x + d^T(Y\zeta + y) \tag{3.17a}$$

$$\text{subject to} \quad Ax + By + S^T q \leq 0 \tag{3.17b}$$

$$P^T S = Y^T B^T - \Psi(x)^T \tag{3.17c}$$

$$S \geq 0\,. \tag{3.17d}$$

We next consider the i^{th} row of constraint (3.17b) and the i^{th} column of constraint (3.17c):

$$A_{i:}x + B_{i:}w + (S_{:i})^T q \leq 0, \tag{3.18a}$$

$$P^T S_{:i} = Y^T (B_{i:})^T - (\Psi(x)_{i:})^T \tag{3.18b}$$

where $A_{i:}$, $B_{i:}$, and $\Psi(x)_{i:}$ denote the i^{th} row of matrices A, B, and $\Psi(x)$ respectively, and $S_{:i}$ denotes the i^{th} column of matrix S. We show that constraints (3.18a) and (3.18b) are equivalent to

$$A_{i:}x + B_{i:}(Y\zeta + y) \leq \Psi(x)_{i:}\zeta,\ \forall\,\zeta \in \mathcal{U}\,. \tag{3.19}$$

We do so by considering that S is not in the objective function so that we can remove S from the set of decision variables and instead replace constraints (3.18a) and (3.18b) with

$$\min_{S_{:i}} \quad A_{i:}x + B_{i:}w + (S_{:i})^T q \quad \leq 0\,, \tag{3.20a}$$

$$\text{subject to} \quad P^T S_{:i} = X^T (B_{i:})^T - (\Psi(x)_{i:})^T\,, \tag{3.20b}$$

where the embedded minimization problem can be replaced by a maximization problem using duality theory. This leads us to considering constraint (3.20) as equivalent to

$$\begin{aligned} \max_{\zeta} \quad & A_{i:}x + B_{i:}w + (X^T(B_{i:})^T - (\Psi(x)_{i:})^T)^T\zeta && \leq 0\,, \\ \text{subject to} \quad & P\zeta \leq q \end{aligned}$$

with ζ as the dual variable of (3.20b). We have thus confirmed that constraints (3.18a) and (3.18b) are equivalent to (3.19), and likewise that constraints (3.17b) and (3.17c) are equivalent to constraint (3.2b). Therefore, the LRC model (3.10) is equivalent to the AARC model (3.2). $\blacksquare$

3.4 Improving LRC and AARC Using Valid Inequalities

In this section, we identify two types of valid inequalities that can be employed to formulate improved versions of LRC that provide tighter conservative approximations. First, we will make use of valid linear inequalities that can be derived from an implicit upper bound on the optimal solution for λ in problem (3.7). This process will lead to a modified LRC model that preserves the computational complexity of LRC. Secondly, we will identify a set of conic valid inequalities that will lead to a semi-definite programming formulation for LRC.

We start with a simple assumption that can be used to generate helpful valid inequalities for problem (3.7) and obtain our Modified LRC (MLRC) model.

Assumption 4 *There exists a bounding vector $u \in \mathbb{R}^m$ such that, for all $x \in \mathcal{X}$ and for all $\zeta \in \mathcal{U}$, there exists an optimal solution $\lambda^* \leq u$ for the problem*

$$\begin{aligned} \underset{\lambda}{\text{minimize}} \quad & (\Psi(x)\zeta - Ax)^T\lambda && (3.21a) \\ \text{subject to} \quad & B^T\lambda = d && (3.21b) \\ & \lambda \geq 0\,. && (3.21c) \end{aligned}$$

Proposition 3.4.1 *Given Assumption 4, the following modified linearized robust counterpart model is a conservative approximation to problem* (3.1)*:*

$$\begin{aligned}
\text{(MLRC)} \quad \underset{x\in\mathcal{X},Y,y,S,s,W,w}{\text{maximize}} \quad & c^T x + d^T y - q^T s - u^T w - u^T W q && \text{(3.22a)}\\
\text{subject to} \quad & Y^T B^T - P^T(S-W) = \Psi(x)^T && \text{(3.22b)}\\
& Ax + By + (S-W)^T q - w \leq 0 && \text{(3.22c)}\\
& -P^T(s + Wu) = Y^T d && \text{(3.22d)}\\
& s \geq 0\,,\ S \geq 0\,,\ w \geq 0\,,\ W \geq 0\,, && \text{(3.22e)}
\end{aligned}$$

where $Y \in \mathbb{R}^{n_y \times n_\zeta}$*,* $y \in \mathbb{R}^{n_y}$*,* $S \in \mathbb{R}^{n_\mathcal{U} \times m}$*,* $s \in \mathbb{R}^{n_\mathcal{U}}$*,* $W \in \mathbb{R}^{n_\mathcal{U} \times m}$*, and* $w \in \mathbb{R}^m$*. Furthermore, the optimal value of problem* (3.22) *is necessarily larger than or equal to the optimal value of LRC* (3.10)*.*

Proof Given that Assumption 4 is satisfied, the following constraints are valid inequalities for problem (3.7) in the sense that they can be added to this problem without affecting its optimal value:

$$\lambda \leq u \tag{3.23a}$$

$$(q - P\zeta)(u - \lambda)^T \geq 0 \tag{3.23b}$$

where constraint (3.23b) can be linearized by replacing $\Delta := \zeta\lambda^T$ as

$$P\Delta \geq q\lambda^T - (q - P\zeta)u^T\,. \tag{3.24}$$

Adding constraints (3.23a) and (3.24) to problem (3.11) leads to MLRC after applying duality theory. ∎

One might consider that Assumption 4 is difficult to exploit in practice since λ is a variable that does not have a clear physical meaning. For this reason, we propose a numerical procedure that can be used, in a specific problem instance, to identify a u that satisfies this useful assumption. We will later show in two practical examples presented in Section 3.6 that a good value for u can also often be obtained analytically.

Proposition 3.4.2 *For any fixed $k = 1, 2, \ldots, m$, one can identify an upper bound u_k that satisfies Assumption 4 by solving the following mixed-integer program*

$$\max_{x\in\mathcal{X},\zeta\in\mathcal{U},y,\lambda,v} \quad \lambda_k \tag{3.25a}$$

$$\text{subject to} \quad Ax + By \leq \Psi(x)\zeta \tag{3.25b}$$

$$A^T\lambda = d \tag{3.25c}$$

$$\lambda_i \leq Mv_i\,,\ \forall\, i \tag{3.25d}$$

$$-a_i^T x - b_i^T y + \psi_i^T\zeta \leq M(1 - v_i)\,,\ \forall\, i \tag{3.25e}$$

$$\lambda \geq 0\,,\ v \in \{0, 1\}^m\,, \tag{3.25f}$$

where M is a large enough constant. Furthermore, this optimization problem reduces to a mixed-integer linear program when $\Psi(x) = \Psi$, i.e., it is independent of x, and the feasible set $\mathcal{X}$ can be represented using linear constraints.

Proof For any fixed $x \in \mathcal{X}$ and $\zeta \in \mathcal{U}$, it is well known that a certain λ^* is optimal in problem (3.21) if and only if it can be paired to some y^* so that together they satisfy the following KKT conditions:

$$Ax + By^* \leq \Psi(x)\zeta \tag{3.26a}$$

$$A^T\lambda^* = d \tag{3.26b}$$

$$\lambda^* \geq 0 \tag{3.26c}$$

$$\lambda_i^*(-a_i^T x - b_i^T y^* + \Psi(x)_i^T\zeta) = 0\,,\ \forall\, i\,. \tag{3.26d}$$

It is therefore clear that, by solving the following optimization problem, one gets a valid candidate for u_k:

$$\underset{x\in\mathcal{X},\zeta\in\mathcal{U},y,\lambda}{\text{maximize}} \quad \lambda_k$$

$$\begin{aligned} \text{subject to} \quad & Ax + By \leq \Psi(x)\zeta \\ & A^T\lambda = d \\ & \lambda \geq 0 \\ & \lambda_i(-a_i^T x - b_i^T y + \Psi(x)_i^T \zeta) = 0\,,\ \forall\, i\,. \end{aligned}$$

One can finally linearize the complementary slackness conditions by introducing binary variables into this problem, *i.e.*, $v_i \in \{0, 1\}$ stating whether λ_i is equal to zero or not. This leads to problem (3.25). ∎

As for the original LRC model, one can uncover an intimate connection between MLRC and conservative approximations that are obtained using AARC. This connection is made explicit in the following proposition.

Proposition 3.4.3 *The MLRC* (3.22) *is equivalent to applying affine adjustments to the following two-stage problem:*

$$\underset{x\in\mathcal{X},y(\zeta),z(\zeta)}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) + u^T z(\zeta) \tag{3.27a}$$

$$\text{subject to} \quad Ax + By(\zeta) \leq \Psi(x)\zeta + z(\zeta),\ \forall\,\zeta\in\mathcal{U} \tag{3.27b}$$

$$z(\zeta) \geq 0\,,\ \forall\,\zeta\in\mathcal{U}\,. \tag{3.27c}$$

Note that the optimal affine adjustments obtained from solving model (3.27) might not be implementable practically, as their only purpose is to identify good first-stage decisions. This means that once x^* is implemented and $\bar{\zeta}$ is observed, it is important to solve the specific recourse problem that is being experienced, *i.e.*, $\text{maximize}_y\, d^T y$, subject to $Ax^* + By \leq \Psi(x^*)\bar{\zeta}$.

Proof Adding constraints (3.23a) and (3.24) to problem (3.11) leads to the following formulation:

$$g_{MLRC}(x) := \quad \min_{\zeta,\lambda,\Delta} \quad c^T x + \text{tr}(\Psi(x)\Delta) - (Ax)^T\lambda \tag{3.28a}$$

$$\text{subject to} \quad B^T \lambda = d \tag{3.28b}$$

$$P\zeta \leq q \tag{3.28c}$$

$$\Delta B = \zeta d^T \tag{3.28d}$$

$$P\Delta \leq q\lambda^T \tag{3.28e}$$

$$0 \leq \lambda \leq u \tag{3.28f}$$

$$P\Delta \geq q\lambda^T - (q - P\zeta)u^T \,. \tag{3.28g}$$

Similarly to what was described in the proof of Proposition 3.3.1, the function $g_{MLRC}(x)$ can be reformulated as

$$g_{MLRC}(x) = \max_{Y,y,W,w,S} \quad \min_{\zeta \in \mathcal{U}} c^T x + d^T(Y\zeta + y) + u^T(w + W^T(q - P\zeta)) \tag{3.29a}$$

$$\text{subject to} \quad P^T S = Y^T B^T - \Psi(x)^T + P^T W \tag{3.29b}$$

$$Ax + Bw + S^T q \leq W^T q + w \tag{3.29c}$$

$$S \geq 0 \tag{3.29d}$$

$$w \geq 0 \,,\, W \geq 0 \,, \tag{3.29e}$$

where $w \in \mathbb{R}^m$ and $W \in \mathbb{R}^{n_{\mathcal{U}} \times m}$ are respectively the dual variables associated with constraints (3.28f) and (3.28g). Again, the constraints (3.29b)-(3.29d) can be replaced with

$$Ax + B(Y\zeta + y) \leq \Psi(x)\zeta + w + W^T(q - P\zeta), \, \forall \, \zeta \in \mathcal{U} \,,$$

and decision variable S removed from the optimization problem. In this way, we obtain

$$g_{MLRC}(x) = \max_{Y,y,W,w} \quad \min_{\zeta \in \mathcal{U}} c^T x + d^T(Y\zeta + y) + u^T(w + W^T(q - P\zeta)) \tag{3.30a}$$

$$\text{subject to} \quad Ax + B(Y\zeta + y) \leq \Psi(x)\zeta + w + W^T(q - P\zeta), \, \forall \, \zeta \in \mathcal{U} \tag{3.30b}$$

$$w \geq 0 \,,\, W \geq 0 \,. \tag{3.30c}$$

We next introduce new variables z and Z as

$$z := w + W^T q \,, \qquad Z := -W^T P \,,$$

where $z \in \mathbb{R}^m$ and $Z \in \mathbb{R}^{m \times n_\zeta}$. Therefore, problem (3.29) can be reformulated as

$$\begin{aligned} g_{MLRC}(x) = \quad & \max_{Y,y,W,Z,z} && \min_{\zeta \in \mathcal{U}} c^T x + d^T(Y\zeta + y) + u^T(z + Z\zeta) && (3.31a) \\ & \text{subject to} && Ax + B(Y\zeta + y) \leq \Psi(x)\zeta + z + Z\zeta, \; \forall \zeta \in \mathcal{U} && (3.31b) \\ & && Z = -W^T P && (3.31c) \\ & && z - W^T q \geq 0 && (3.31d) \\ & && W \geq 0\,. && (3.31e) \end{aligned}$$

We finally show that constraints (3.31c)-(3.31e) are equivalent to the following constraint:

$$z + Z\zeta \geq 0\,, \; \forall \zeta \in \mathcal{U} \tag{3.32}$$

This is done by considering that, for some fixed j, duality can once again be used to reformulate the constraint that

$$\begin{aligned} & \max_{w_{:j}} && z_j - W_{:j}^T q && \geq 0 && (3.33a) \\ & \text{subject to} && Z_{j:}^T = -P^T W_{:j} && && (3.33b) \\ & && W_{:j} \geq 0 && && (3.33c) \end{aligned}$$

as the constraint that

$$\begin{aligned} & \min_{\zeta} && z_j + Z_{j:}\zeta && \geq 0\,. && (3.34a) \\ & \text{subject to} && P\zeta \leq q && && (3.34b) \end{aligned}$$

This completes our proof. ∎

Our second source of improvement for the LRC model comes from considering the following set of quadratic equalities:

$$\begin{bmatrix} \Lambda & \Delta^T \\ \Delta & \Xi \end{bmatrix} = \begin{bmatrix} \lambda \\ \zeta \end{bmatrix} \begin{bmatrix} \lambda^T & \zeta^T \end{bmatrix}, \tag{3.35}$$

where $\Lambda \in \mathbb{R}^{m\times m}$ and $\Xi \in \mathbb{R}^{n_\zeta \times n_\zeta}$ such that, $\Lambda := \lambda\lambda^T$ and $\Xi := \zeta\zeta^T$. It is well known that this system of equations can be relaxed using the following matrix inequality

$$\begin{bmatrix} \Lambda & \Delta^T \\ \Delta & \Xi \end{bmatrix} \succeq \begin{bmatrix} \lambda \\ \zeta \end{bmatrix} \begin{bmatrix} \lambda^T & \zeta^T \end{bmatrix},$$

where $A \succeq B$ indicates that $A - B$ is in the cone of positive semi-definite matrices. This non-linear matrix inequality reduces to a linear matrix inequality after applying Schur's complement

$$\begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix} \succeq 0\,.$$

This constraint can be added to problem (3.11) with additional valid inequalities involving Λ and Ξ to obtain the tighter SDP-LRC model.

Proposition 3.4.4 *Given Assumption 4, the following semi-definite programming linearized robust counterpart is a conservative approximation of problem* (3.1)*:*

$$g_{\text{SDP-LRC}}(x) = \min_{\zeta,\lambda,\Delta,\Lambda,\Xi} \quad c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda \tag{3.36a}$$

$$\text{subject to} \quad B^T\lambda = d \tag{3.36b}$$

$$P\zeta \leq q \tag{3.36c}$$

$$0 \leq \lambda \leq u \tag{3.36d}$$

$$\Delta B = \zeta d^T \tag{3.36e}$$

$$P\Delta \leq q\lambda^T \tag{3.36f}$$

$$P\Delta \geq q\lambda^T - (q - P\zeta)u^T \tag{3.36g}$$

$$\begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix} \succeq 0 \tag{3.36h}$$

$$\Lambda B = \lambda d^T \tag{3.36i}$$

$$P\Xi P^T + qq^T \geq P\zeta q^T + q\zeta^T P^T \quad (3.36j)$$

$$0 \leq \Lambda \leq u\lambda^T\,. \quad (3.36k)$$

Moreover, the optimal value of $\max_{x\in\mathcal{X}} g_{\text{SDP-LRC}}(x)$ *is necessarily larger than or equal to the optimal value of MLRC* (3.22) *and LRC* (3.10)*.*

Proof We simply start by including the new variables Λ and Ξ and constraint (3.36) in problem (3.11). One can then realize that constraints (3.9g) and (3.9h) can now help tighten the feasible region. Finally, a final tightening step can be achieved by exploiting the fact that $P\zeta \leq q$ implies the following:

$$q - P\zeta \geq 0 \Rightarrow (q - P\zeta)(q - P\zeta)^T \geq 0 \Rightarrow P\zeta\zeta^T P^T + qq^T \geq P\zeta q^T + q\zeta^T P^T$$

and that $\lambda \leq u$ implies that $\lambda\lambda^T \leq u\lambda^T$, which together lead to constraint (3.36j) after replacing $\Xi := \zeta\zeta^T$ and $\Lambda := \lambda\lambda^T$. ∎

Note that, although we presented $g_{\text{SDP-LRC}}(x)$ as a minimization problem, a semi-definite programming duality can be employed to obtain a maximization representation of this function that can be integrated with the maximization in x as was done with other LRC models. We however omit the details of this reformulation for aesthetics reasons. Given the connections to AARC that were established regarding the LRC and MLRC models, we suspect that a similar connection could be obtained for the SDP-LRC model. In fact, the authors of Ardestani-Jaafari and Delage (2016a) were able to establish such a connection for a special case of the SDP-LRC model. A quick look at their result suggests that the connection that could be established here is highly technical and would provide rather limited new insights.

3.5 LRC with A General Uncertainty Set

In this section, we extend our LRC model so that it can accomodate general convex uncertainty sets, *i.e.,* $\mathcal{U}$ is not polyhedral. In this regard, we will instead consider uncertainty sets that can be represented as

$$\mathcal{U}_{general} := \{\zeta \in \mathbb{R}^{n_\zeta} \mid f_l(\zeta) \leq q_l,\, \forall\, l = 1, \ldots, L\} \quad (3.37)$$

using a set of convex $f_l(\cdot)$ functions. To establish an extension of LRC, we will need to make use of perspective functions which are defined next.

Definition The perspective of a convex function $f : \mathbb{R}^{n_\zeta} \to \mathbb{R}$ is the function $h : \mathbb{R}^{n_\zeta} \times \mathbb{R} \to \mathbb{R}$ defined as $h(\zeta, t) := tf(\zeta/t)$. It is well known that the function $h(\zeta, t)$ is jointly convex in ζ and t if $f(\zeta)$ is a convex function.

Under the uncertainty set $\mathcal{U}_{general}$, when Assumption 4 is satisfied, the value of $g(x)$ becomes

$$
\begin{aligned}
g(x) = \quad & \min_{\zeta,\lambda} && c^T x + (\Psi(x)\zeta)^T \lambda - (Ax)^T \lambda && \text{(3.38a)} \\
& \text{subject to} && B^T \lambda = d && \text{(3.38b)} \\
& && f_l(\zeta) \leq q_l, \forall\, l && \text{(3.38c)} \\
& && 0 \leq \lambda \leq u && \text{(3.38d)} \\
& && \zeta \lambda^T B = \zeta d^T && \text{(3.38e)} \\
& && \lambda_i f_l(\zeta) \leq \lambda_i q_l\,, \forall\, i\,, \forall\, l && \text{(3.38f)} \\
& && (q_l - f_l(\zeta))(u_i - \lambda_i) \geq 0\,, \forall\, i\,, \forall\, l\,. && \text{(3.38g)}
\end{aligned}
$$

Similarly as before, this model can be linearized as

$$
\begin{aligned}
g_{GLRC}(x) = \quad & \min_{\zeta,\lambda,\Delta} && \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && \text{(3.39a)} \\
& \text{subject to} && B^T \lambda = d && \text{(3.39b)} \\
& && f_l(\zeta) \leq q_l, \forall\, l && \text{(3.39c)} \\
& && 0 \leq \lambda \leq u && \text{(3.39d)} \\
& && \Delta B = \zeta d^T && \text{(3.39e)} \\
& && h_l(\Delta_{:i}, \lambda_i) \leq q_l\,, \forall\, i\,, \forall\, l && \text{(3.39f)} \\
& && h_l(\zeta u_i - \Delta_{:i}, u_i - \lambda_i) \leq q_l(u_i - \lambda_i)\,, && \text{(3.39g)}
\end{aligned}
$$

where constraint (3.39f) is derived from

$$f_l(\zeta) \le q_l \Rightarrow \lambda_i f_l(\zeta) \le q_l \lambda_i \Rightarrow h(\Delta_{:i}, \lambda_i) = \lambda_i f_l\left(\frac{\zeta\lambda_i}{\lambda_i}\right) \le q_l \lambda_i$$

and constraint (3.39g) is derived from

$$\big(q_l - f_l(\zeta)\big)(u_i - \lambda_i) \ge 0 \Rightarrow (u_i - \lambda_i) f_l\left(\frac{\zeta(u_i - \lambda_i)}{(u_i - \lambda_i)}\right) \le q_l(u_i - \lambda_i) \Rightarrow h_l(\zeta u_i - \Delta_{:i}, u_i - \lambda_i) \le q_l(u_i - \lambda_i)$$

for all i and for all l, and where the term $\zeta\lambda_i$ is linearized through $\Delta_{:i}$. One might apply duality theory to problem (3.39) to derive an LRC model for problem (3.1) under general convex uncertainty set. Regarding the relation between this more general LRC model and AARC, our conjecture is that the problem $\text{maximize}_{x \in \mathcal{X}}\, g_{GLRC}(x)$ is exactly equivalent to employing affine adjustments in problem (3.27) with $\mathcal{U}_{general}$.

3.6 Examples

In this section, we provide some examples to show how to apply LRC in practice. The first example is a robust multi-item newsvendor problem that is an instance of problem (3.3) for which it is possible to identify conditions under which the LRC is an exact model, *i.e.,* the relaxation gap γ that is described in Proposition 3.2.3 is equal to one. Moreover, we describe two logistics applications of ARO wherein LRC can be improved by using the linear valid inequalities presented in Proposition 3.4.1.

We note that, in this section, in order to be more concise in our descriptions, we let $\mathcal{I}$ and $\mathcal{J}$ represent the sets $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$ respectively, and consider that $i \in \mathcal{I}$ and $j \in \mathcal{J}$ when the membership for i and j is left unspecified.

Example (Multi-item newsvendor problem) Consider the following robust multi-item newsvendor problem:

$$\max_{x \in \mathcal{X}} \min_{\zeta \in \mathcal{U}} \sum_j r_j \min(x_j, \zeta_j) - c_j x_j + s_j \max_j(x_j - \zeta_j, 0) - p_j \max(\zeta_j - x_j, 0)\,, \qquad (3.40)$$

where r_j, c_j, s_j, and p_j denote price, ordering cost, salvage price, and shortage cost of a unit of the j-th item, $j \in \mathcal{J}$, respectively, and ζ_j denotes the demand for item j for each j.

Problem (3.40) is a special case of ARO, as

$$\begin{aligned}
\max_{x\in\mathcal{X},y(\zeta)} \quad & \min_{\zeta\in\mathcal{U}} \sum_j y_j(\zeta) \\
\text{subject to} \quad & y_j(\zeta) \le r_j\zeta_j - c_jx_j + s_j(x_j-\zeta_j)\,,\ \forall\, j\in\mathcal{J}\,, \forall\,\zeta\in\mathcal{U} \\
& y_j(\zeta) \le (r_j-c_j)x_j + l_j(x_j-\zeta_j)\,,\ \forall\, j\in\mathcal{J}\,, \forall\,\zeta\in\mathcal{U}\,.
\end{aligned}$$

In Ardestani-Jaafari and Delage (2016a), the authors show that $g_{LRC}(x)$ is tight in this problem when $\mathcal{U}$ is defined as the following uncertainty set:

$$\mathcal{U} = \{\zeta|\zeta_j = \bar{\zeta}_j + (\delta_j^+ - \delta_j^-)\hat{\zeta}_j\,,\ \delta_j^- \ge 0, \delta_j^+ \ge 0\,, \delta_j^+ + \delta_j^- \le 1\,,\ \forall\, j\,, \sum_j \delta_j^+ + \delta_j^- = \Gamma\}\,,$$

and when Γ is an integer value. Furthermore, one could explore for other uncertainty set whether it is possible to get a tighter conservative approximation by employing Proposition 3.4.1. Yet, one can show that the dual variables of the recourse problem are implicitly bounded in this multi-item newsvendor problem. Specifically, the dual problem takes the shape of

$$\begin{aligned}
\min_{\lambda^1,\lambda^2} \quad & \sum_j \lambda_j^1(r_j\zeta_j - c_jx_j + s_j(x_j-\zeta_j)) + \sum_j \lambda_j^2((r_j-c_j)x_j + l_j(x_j-\zeta_j)) \\
\text{subject to} \quad & \lambda_j^1 + \lambda_j^2 = 1\,,\ \forall\, j\in\mathcal{J}\,, \\
& \lambda^1 \ge 0\,,\ \lambda^2 \ge 0\,,
\end{aligned}$$

which already implies that $\lambda^1 \le 1$ and $\lambda^2 \le 1$ at optimum. Hence, MLRC model (3.22) with $u = 1$ becomes trivially equivalent to LRC model (3.10) in this case. In order to obtain a tighter conservative approximation, one should instead employ the SDP-LRC approximation model.

Example (Location-transportation problem) The robust location-transportation problem can be formulated as

$$\underset{x,y(\zeta),v}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} \ -\sum_i c_ix_i - k_iv_i + \sum_i\sum_j \eta_{ij}y_{ij}(\zeta) \tag{3.41a}$$

$$\text{subject to} \quad \sum_i y_{ij}(\zeta) \leq \zeta_j \, , \, \forall \, j \in \mathcal{J} \, , \, \forall \, \zeta \in \mathcal{U} \tag{3.41b}$$

$$\sum_j y_{ij}(\zeta) \leq x_i \, , \, \forall \, i \in \mathcal{I} \, , \, \forall \, \zeta \in \mathcal{U} \tag{3.41c}$$

$$y(\zeta) \geq 0 \, , \, \forall \, \zeta \in \mathcal{U} \tag{3.41d}$$

$$0 \leq x_i \leq M v_i \, , \, \forall \, i \in \mathcal{I} \tag{3.41e}$$

$$v_i \in \{0, 1\} \, , \, \forall \, i \in \mathcal{I} \, . \tag{3.41f}$$

In this problem, variable v_i indicates that, if there is an open facility in location i for each $i \in \mathcal{I}$, variable x_i denotes the production capacity of the facility i, and variable y_{ij} denotes how many goods are shipped from facility i to customers at location j, with $j \in \mathcal{J}$. The demand for location j is characterized by ζ_j. Parameter $\eta_{ij} > 0$ denotes the unit revenue of goods shipped from facility i to customer j, while c_i and k_i denote variable and fixed capacity cost for facility i respectively.

Here, the dual formulation of the recourse function takes the form

$$\underset{\lambda^1, \lambda^2, \lambda^3}{\text{minimize}} \quad \sum_j \zeta_j \lambda_j^1 + \sum_i x_i \lambda_i^2 \tag{3.42a}$$

$$\text{subject to} \quad \lambda_j^1 + \lambda_i^2 - \lambda_{ij}^3 = \eta_{ij} \, , \, \forall \, i \in \mathcal{I} \, , \, \forall \, j \in \mathcal{J} \tag{3.42b}$$

$$\lambda^1 \geq 0 \, , \; \lambda^2 \geq 0 \, , \; \lambda^3 \geq 0 \, , \tag{3.42c}$$

where λ^1, λ^2, and $\lambda^3 \in \mathbb{R}^{m \times n}$ are dual variables associated with constraints (3.41b)-(3.41d) respectively. Since the objective function of problem (3.42) is non-decreasing in λ^1 and λ^2, one can conclude that, at optimum, each term of λ^{1*} and λ^{2*} will be such that it will be involved in at least one active constraint among the set of constraints

$$\lambda_j^1 + \lambda_i^2 \geq \eta_{ij} \, , \, \forall \, i \in \mathcal{I} \, , \, \forall \, j \in \mathcal{J} \, .$$

It must therefore be that

$$\lambda_j^{1*} \leq \max_i \eta_{ij} - \lambda_i^{2*} \leq \max_i \eta_{ij} = u_j^1 \, , \, \forall \, j \in \mathcal{J} \, ,$$

and that

$$\lambda_i^2 \leq \max_j \eta_{ij} - \lambda_j^{1*} \leq \max_j \eta_{ij} = u_j^2 \,,\, \forall\, i \in \mathcal{I} \,.$$

Finally, since $\lambda_{ij}^{3*} = \lambda_j^{1*} + \lambda_i^{2*} - \eta_{ij}$, one could conclude that $\lambda_{ij}^3 \leq u_{ij}^3 := \max_i \eta_{ij} + \max_j \eta_{ij} - \eta_{ij}$ but this constraint would be redundant in problem (3.42) after imposing the upper bounds on λ^1 and λ^2, so that one can leave $u^3 = \infty$. Based on Propositions 3.4.1 and 3.4.3, we know that it is possible to obtain a tighter conservative approximation to problem (3.41) by employing affine adjustments in the following augmented model:

$$\begin{aligned}
\underset{x,y(\zeta),z^1(\zeta),z^2(\zeta),z^3(\zeta),v}{\text{maximize}} \quad & \min_{\zeta \in \mathcal{U}} \; -\sum_i c_i x_i - K_i I_i + \sum_i \sum_j \eta_{ij} y_{ij}(\zeta) + {u^1}^T z^1(\zeta) + {u^2}^T z^2(\zeta) \\
\text{subject to} \quad & \sum_i y_{ij}(\zeta) \leq \zeta_j + z_j^1(\zeta) \,,\, \forall\, j \in \mathcal{J} \,,\, \forall\, \zeta \in \mathcal{U} \\
& \sum_j y_{ij}(\zeta) \leq x_i + z_i^2(\zeta) \,,\, \forall\, i \in \mathcal{I} \,,\, \forall\, \zeta \in \mathcal{U} \\
& y(\zeta) \geq 0 \,,\, \forall\, \zeta \in \mathcal{U} \\
& z^1(\zeta) \geq 0 \,,\; z^2(\zeta) \geq 0 \,,\, \forall\, \zeta \in \mathcal{U} \\
& 0 \leq x_i \leq M v_i \,,\, \forall\, i \in \mathcal{I} \\
& v_i \in \{0, 1\} \,,\, \forall\, i \in \mathcal{I} \,.
\end{aligned}$$

where $z^1 : \mathbb{R}^n \to \mathbb{R}^n$ and $z^2 : \mathbb{R}^n \to \mathbb{R}^m$ can be interpreted as violation adjustments for constraints (3.41b) and (3.41c). In Ardestani-Jaafari and Delage (2016b), the authors employed a special case of such a conservative approximation where $z^2 = 0$, and showed that there exist instances of the location-transportation problem for which this conservative approximation is strictly tighter than employing affine adjustments directly on model (3.41).

Example (Multi-product assembly problem (Shapiro et al. 2014) Page. 9) In this problem, a manufacturer produces n products using m different types of parts. It is a two-stage problem wherein the manufacturer pre-orders x_i units for part $i \in \mathcal{I}$ with a cost of c_i per unit in the first stage; and when demand is realized, it must be determined how many products to make, y_j for each product $j \in \mathcal{J}$. The robust multi-product assembly problem can be formulated as follows:

$$\begin{aligned}
\underset{x,y(\zeta)}{\text{maximize}} \quad & \min_{\zeta\in\mathcal{U}} \; -c^T x + (q-l)^T y(\zeta) + s^T(x - Ay(\zeta)) && (3.43\text{a})\\
\text{subject to} \quad & y(\zeta) \le \zeta \,, \forall\, \zeta \in \mathcal{U} && (3.43\text{b})\\
& Ay(\zeta) \le x \,, \forall\, \zeta \in \mathcal{U} && (3.43\text{c})\\
& y(\zeta) \ge 0 \,, \forall\, \zeta \in \mathcal{U} && (3.43\text{d})\\
& 0 \le x \le M \,, && (3.43\text{e})
\end{aligned}$$

where $\zeta \in \mathbb{R}^n$ is the uncertain demand for each product and where parameters q and l denote, respectively, the selling price and the production cost per unit of the products, while s denotes the salvage unit value of unused parts. Finally A_{ij} denotes the number of units of part i that is required to assemble product j.

As was done for the previous example, one can hope to identify a tighter conservative approximation by applying an affine adjustment on the following augmented model:

$$\begin{aligned}
\underset{x,y(\zeta),z^1(\zeta),z^2(\zeta)}{\text{maximize}} \quad & \min_{\zeta\in\mathcal{U}} \; -c^T x + (q-l)^T y(\zeta) + s^T(x - Ay(\zeta)) + {u^1}^T z^1(\zeta) + {u^2}^T z^2(\zeta) && (3.44\text{a})\\
\text{subject to} \quad & y(\zeta) \le \zeta + z^1(\zeta) \,, \forall\, \zeta \in \mathcal{U} && (3.44\text{b})\\
& Ay(\zeta) \le x + z^2(\zeta) \,, \forall\, \zeta \in \mathcal{U} && (3.44\text{c})\\
& y(\zeta) \ge 0 \,, \forall\, \zeta \in \mathcal{U} && (3.44\text{d})\\
& 0 \le x \le M \,, && (3.44\text{e})
\end{aligned}$$

where $z^1 : \mathbb{R}^n \to \mathbb{R}^n$ and $z^2 : \mathbb{R}^n \to \mathbb{R}^m$ can be interpreted as violation adjustments for constraints (3.43b) and (3.43c). Yet, in this case, the u bounds are obtained from the dual problem:

$$\begin{aligned}
\underset{\lambda^1,\lambda^2}{\text{minimize}} \quad & \zeta^T \lambda^1 + x^T \lambda^2 && (3.45\text{a})\\
\text{subject to} \quad & \lambda^1_j + \sum_i A_{ij}\lambda^2_i \ge q_j - l_j + A^T_{:j} s \,, \; \forall\, j \in \mathcal{J} && (3.45\text{b})\\
& \lambda^1 \ge 0 \,, \; \lambda^2 \ge 0 \,, && (3.45\text{c})
\end{aligned}$$

where $\lambda^1 \in \mathbb{R}^n$ and $\lambda^2 \in \mathbb{R}^n$ are the dual variables associated to constraints (3.43b) and (3.43c). Here again, the objective function is non-decreasing in λ^1 and λ^2 so that, at optimum, each term of these two vectors must be involved in at least one active constraint among the following set:

$$\lambda_j^1 + \sum_i A_{ij}\lambda_i^2 \geq q_j - l_j + A_{:j}^T s \, , \, \forall \, j \in \mathcal{J} \; .$$

This indicates to us that

$$\lambda_j^{1*} \leq q_j - l_j + A_{:j}^T s - \sum_i A_{ij}\lambda_i^{2*} \leq q_j - l_j + A_{:j}^T s = u_j^1 \, ,$$

and that

$$\lambda_i^{2*} \leq \max_{j \in \mathcal{J}_i} \frac{1}{A_{ij}}(q_j - l_j + A_{:j}^T s - \lambda_j^1 - \sum_{i' \neq i} A_{ij}\lambda_i^{2*}) \leq \max_{j \in \mathcal{J}_i} \frac{1}{A_{ij}}(q_j - l_j + A_{:j}^T s) = u_i^2 \; ,$$

where the set of indices $\mathcal{J}_i := \{j \,|\, A_{ij} \neq 0\}$.

We conclude this example with a description of the specific context in which exploiting the information about the bound u on λ^* leads to a strictly tighter conservative approximation. In particular, consider a multi-product assembly problem with three products and two different types of parts. The cost of parts A and B are respectively \$25 per unit and \$5 per unit, while the salvage value is \$4 per unit and \$1 per unit. Furthermore, the difference between the selling price and the unit production cost of each product is: \$380/unit, \$800/unit, and \$1200/unit respectively for products #1 to #3. Next, we have that product #1 requires 9 units of both parts, product #2 requires 5 units of part B, and #3 requires 9 units of A and 4 units of B. Finally, for products #1 to #3, the nominal demand is respectively of 9000, 10,000, and 8000 units while the worst-case demand for each is 1000, 2000, and 0 units respectively. In this specific context, one can set the bound vectors as $u^1 := [\; 425 \;\; 805 \;\; 1240 \;]^T$ and $u^2 := [\; 140 \;\; 310 \;]^T$. When the budget of uncertainty is set to $\Gamma = 2$, a direct application of affine adjustments on problem (3.43) will lead to the acquisition of 92,793 parts of type A and 91,000 parts of type B, with a worst-case profit estimated at 2.474 million dollars; meanwhile applying affine adjustments on the equivalent formulation that allows penalized violations, *i.e.,* problem (3.44), results in an order of 81,,000 parts of type A and 91,000 parts of type B to achieve a worst-case profit estimated at 2.722 million dollars (namely a 10% increase in profit). This confirms that the MLRC model can provide a strictly tighter conservative approximation.

3.7 Conclusions

In this paper, we extended the linearization scheme presented in Ardestani-Jaafari and Delage (2016a) so that it can be used to construct tractable conservative approximation models for two-stage adjustable robust optimization problems with right-hand side uncertainty. We showed that, as in Ardestani-Jaafari and Delage (2016a), this scheme provides an alternate interpretation of models obtained through the use of AARC. Yet, by considering the adversarial problem as a bilinear optimization problem that needs to be linearized, it becomes very natural to identify modifications based on linear and conic valid inequalities that will improve LRC and consequently provide tightening procedures for AARC. Based on these results, it is clear that the LRC model can help clarify the quality of solutions obtained from AARC and offers a perspective that might help design better approximation methods for two-stage adjustable robust optimization models. We finally surveyed the types of improvement that the models we suggest might offer in three different applications of logistics problems.

Conclusion

The main contribution of this thesis is to propose a new scheme, named as the linearized robust counterpart (LRC) framework, for constructing tractable conservative approximation models for two-stage robust optimization problem and problems where sums of piecewise linear functions needs to be robustified. In the context of a profit maximization problem, this scheme relies on four steps: 1) exploiting duality to represent the recourse problem as a minimization problem; 2) introducing valid inequalities that involve the problematic bilinearities; 3) linearizing the problem by introducing new decision variables; 4) applying duality to get a compact reformulation in terms of the first-stage decisions. This scheme has many natural applications including multi-period location-transportation problems, multi-product assembly problems, but also applications that involve sums of piecewise linear function such as inventory management problems, newsvendor problems, classification problems, brachytherapy, etc.

While some might consider that this new scheme lacks in originality given that in many occasion we showed that it leads to models that can be obtained by using AARC in a clever way, we argue that this connection instead enriches the LRC framework. Indeed, AARC is a method that has been abundantly used in the past ten years in the context of many important applications, therefore all the results contained in this thesis about conditions for exactness and ways of getting better approximation from LRC can have a direct impact on any of these applications. From a conceptual perspective, we also believe an important contribution to have connected AARC to topics such as relaxation gap of a MILP and bilinear programming which uncovers new possibilities of collaborations between researchers in the fields of robust optimization and integer programming.

Some interesting future directions of research might be worth exploring based on ideas presented in this thesis. The first direction of research could involve applying the LRC framework to multi-stage robust optimization problem, for which the multi-period problem discussed in Chapter 2 might constitute a good starting point. Another interesting extension could be to consider employing the LRC framework in problems where uncertainty affects the left-hand side of the constraints (a.k.a. uncertainty about the technology matrix) which can also be used to capture problems with uncertainty about the objective coefficient. This might actually not be too difficulty given that it simply means that bilinear terms would appear in problem 3.9 which could be themselves linearized. One might also attempt to identify additional valid inequalities that might help improve the LRC model or even establish new conditions that would imply exactness of the LRC model. Last but not least, one should consider whether the LRC framework can offer additional insights about distributionally robust optimization, in the spirit of what was established for the distributionally robust multi-item newsvendor problem in Chapter 1.

Bibliography

Abouee-Mehrizi, H., O. Berman, M. R. Baharnemati. 2014. Designing production-inventory-transportation systems with capacitated cross-docks. *Transportation Science* **48**(1) 121–135.

An, Y., B. Zeng, Y. Zhang, L. Zhao. 2014. Reliable p-median facility location problem: two-stage robust models and algorithms. *Transportation Research Part B: Methodological* **64** 54–72.

Ardestani-Jaafari, A., E. Delage. 2016a. Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research. Advance online publication.* doi:10.1287/opre.2016.1483.

Ardestani-Jaafari, A., E. Delage. 2016b. The value of flexibility in robust location-transportation problem. *Accepted in Transportation Science* .

Atamtürk, A., M. Zhang. 2007. Two-stage robust network flow and design under demand uncertainty. *Operations Research* **55**(4) 662–673.

Baron, O., J. Milner, H. Naseraldin. 2011. Facility location: a robust optimization approach. *Production and Operations Management* **20**(5) 772–785.

Ben-Tal, A., L. El Ghaoui, A. Nemirovski. 2009a. *Robust optimization*. Princeton University Press.

Ben-Tal, A., B. Golany, A. Nemirovski, J. P. Vial. 2005. Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing & Service Operations Management* **7**(3) 248–271.

Ben-Tal, A., B. Golany, S. Shtern. 2009b. Robust multi-echelon multi-period inventory control. *European Journal of Operational Research* **199**(3) 922–935.

Ben-Tal, A., A. Goryashko, E. Guslitzer, A. Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Ben-Tal, A., A. Nemirovski. 1997. Robust truss topology design via semidefinite programming. *SIAM Journal of Optimization* **7**(4) 991–1016.

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23**(4) 769–805.

Ben-Tal, A., A. Nemirovski. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters* **25**(1) 1–13.

Ben-Tal, A., A. Nemirovski. 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* **88**(3) 411–424.

Ben-Tal, A., A. Nemirovski. 2002. Robust optimization–methodology and applications. *Mathematical Programming* **92**(3) 453–480.

Benson, S. J., Y. Ye, X. Zhang. 2000. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM Journal on Optimization* **10**(2) 443–461.

Bertsimas, D., D. B. Brown, C. Caramanis. 2011a. Theory and application of robust optimization. *SIAM Review* **3**(3) 464–501.

Bertsimas, D., F. de Ruiter. 2015. Duality in two-stage adaptive linear optimization: faster computation and stronger bounds. *accpeted in Informs Jounal of Computing* .

Bertsimas, D., V. Goyal. 2012. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming* **134**(2) 491–531.

Bertsimas, D., D. A. Iancu, P. A. Parrilo. 2010. Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research* **35**(2) 363–394.

Bertsimas, D., D. A. Iancu, P. A. Parrilo. 2011b. A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control* **56**(12) 2809–2824.

Bertsimas, D., M. Sim. 2004. The price of robustness. *Operations Research* **52**(1) 35–53.

Bertsimas, D., A. Thiele. 2006. A robust optimization approach to inventory theory. *Operations Research* **54**(1) 150–168.

Bienstock, D., N. Özbay. 2008. Computing robust basestock levels. *Discrete Optimization* **5**(2) 389–414.

Boyd, S., S.-J. Kim, D. Patil, M. Horowitz. 2005. Digital circuit optimization via geometric programming. *Operations Research* **53**(6) 899–932.

Calafiore, G., L. El Ghaoui. 2001. Robust maximum likelihood estimation in the linear model. *Automatica* **37**(4) 573–580.

Candia-Véjar, A., E. Álvarez-Miranda, N. Maculan. 2011. Minmax regret combinatorial optimization problems: an algorithmic perspective. *RAIRO-Operations Research* **45**(2) 101–129.

Chen, X., Y. Zhang. 2009. Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research* **57**(6) 1469–1482.

Christensen, T. R. L., K. A. Andersen, A. Klose. 2013. Solving the single-sink, fixed-charge, multiple-choice transportation problem by dynamic programming. *Transportation Science* **47**(3) 428–438.

Crainic, T. G., G. Laporte. 1997. Planning models for freight transportation. *European Journal of Operational Research* **97**(3) 409 – 438.

Delage, E., S. Arroyo, Y. Ye. 2014. The value of stochastic modeling in two-stage stochastic programs with cost uncertainty. *Operations Research* **62**(6) 1377–1393.

Delage, E., Y. Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* **58**(3) 595–612.

Denton, B. T., A. J. Miller, H. J. Balasubramanian, T. R. Huschka. 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research* **58**(4-Part-1) 802–816.

Düzgün, R., A. Thiele. 2010. Robust optimization with multiple ranges: Theory and application to R&D project selection. Tech. rep., Industrial and Systems Engineering Department, Lehigh University, Bethlehem, Pennslvania.

El Ghaoui, L., H. Lebret. 1997. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications* **18**(4) 1035–1064.

El Ghaoui, L., F. Oustry, H. Lebret. 1998. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization* **9**(1) 33–52.

Gabrel, V., M. Lacroix, C. Murat, N. Remli. 2014a. Robust location transportation problems under uncertain demands. *Discrete Applied Mathematics* **164, Part 1**(0) 100 – 111.

Gabrel, V., C. Murat, A. Thiele. 2014b. Recent advances in robust optimization: An overview. *European Journal of Operational Research* **235**(3) 471 – 483.

Gallego, G., I. Moon. 1993. The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society* (8) 825–834.

Georghiou, A., W. Wiesemann, D. Kuhn. 2015. Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming* **152**(1-2) 301–338.

Ghaddar, B., J. C. Vera, M. F. Anjos. 2011. Second-order cone relaxations for binary quadratic polynomial programs. *SIAM Journal on Optimization* **21**(1) 391–414.

Goldfarb, D., G. Iyengar. 2002. Robust portfolio selection problems. *Mathematics of Operations Research* **28**(1) 1–38.

Gorissen, B. L., D. den Hertog. 2013. Robust counterparts of inequalities containing sums of maxima of linear functions. *European Journal of Operational Research* **227**(1) 30 – 43.

Grady, L. J., J. Polimeni. 2010. *Discrete calculus: Applied analysis on graphs for computational science*. Springer Science & Business Media.

Hanasusanto, G., D. Kuhn, S. W. Wallace, S. Zymler. 2014. Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming* **152**(1) 1–32.

Hiriart-Urruty, J.-B., C. Lemaréchal. 2001. *Fundamentals of Convex Analysis*. Grundlehren Text Editions, Springer-Verlag Berlin Heidelberg.

Hoffman, A. J., J. B. Kruskal. 1956. Integral boundary points of convex polyhedra. *Linear inequalities and related systems* **38** 223–246.

Iancu, D. A., M. Sharma, M. Sviridenko. 2013. Supermodularity and affine policies in dynamic robust optimization. *Operations Research* **61**(4) 941–956.

José Alem, D., R. Morabito. 2012. Production planning in furniture settings via robust optimization. *Computers & Operations Research* **39**(2) 139–150.

Kouvelis, P., G. Yu. 1997. *Robust discrete optimization and its applications*, vol. 14. Springer Science & Business Media.

Lasserre, J. B. 2002. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM Journal on Optimization* **12**(3) 756–769.

Lei, C., W-H. Lin, L. Miao. 2015. A two-stage robust optimization approach for the mobile facility fleet sizing and routing problem under uncertainty. *Computers & Operations Research* **67** 75–89.

Lovász, L., A. Schrijver. 1991. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization* **1**(2) 166–190.

Mak, H. Y., Y. Rong, J. Zhang. 2015. Appointment scheduling with limited distributional information. *Management Science* **61**(2) 316–334.

Moon, I., E. A. Silver. 2000. The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *Journal of the Operational Research Society* **51** 602–608.

Oktal, H., A. Ozger. 2013. Hub location in air cargo transportation: A case study. *Journal of Air Transport Management* **27**(0) 1 – 4.

Scarf, H. 1958. A min-max solution of an inventory problem. H. E. Scarf K. S. Arrow, S. Karlin, ed., *Studies in the mathematical theory of inventory and production*. Stanford University Press, Stanford, CA, 201–209.

Shapiro, A., D. Dentcheva, A. Ruszczynski. 2014. *Lectures on stochastic programming: modeling and theory*, vol. 16. SIAM.

Sherali, H. D., A. Alameddine. 1992. A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization* **2**(4) 379–410.

Soyster, A. L. 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* **21**(5) 1154–1157.

Stone, James V. 2015. *Information Theory: A Tutorial Introduction*. Sebtel Press.

Truemper, K. 1978. Algebraic characterizations of unimodular matrices. *SIAM Journal on Applied Mathematics* **35**(2) 328–332.

Tsiakis, P., N. Shah, C. Pantelides. 2001. Design of multi-echelon supply chain networks under demand uncertainty. *Industrial & Engineering Chemistry Research* **40**(16) 3585–3604.

van der Vlerk, M. H. 2004. Convex approximations for complete integer recourse models. *Mathematical Programming* **99**(2) 297–310.

Vandenberghe, L., S. Boyd. 1996. Semidefinite programming. *SIAM review* **38**(1) 49–95.

Wang, Z., P. W. Glynn, Y. Ye. 2015. Likelihood robust optimization for data-driven problems. *Accepted in Computational Management Science* .

Wei, C., Y. Li, X. Cai. 2011. Robust optimal policies of production and inventory with uncertain returns and demand. *International Journal of Production Economics* **134**(2) 357–367.

Wiesemann, W., D. Kuhn, M. Sim. 2014. Distributionally robust convex optimization. *Operations Research* **62**(6) 1358–1376.

Xu, H., C. Caramanis, S. Mannor. 2009. Robustness and regularization of support vector machines. *Journal of Machine Learning Research* **10** 1485–1510.

Zeng, B., L. Zhao. 2013. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters* **41**(5) 457 – 461.