universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Conic and Quadratic Optimization tools for Optimization under Uncertainty“

verfasst von / submitted by

Markus Gerald Alexander Gabl, Bakk. phil. BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 794 370 136
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Statistik und Operations Research
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Immanuel Bomze

Acknowledgements

Doing research is a privilege for which I feel immense gratitude. Many people have contributed sometimes substantial effort so I would be able to pursue this profession in the past four years and I want to take this opportunity to express my gratitude towards some of them.

First and foremost I want to thank my supervisor Professor Mag. Dr. Immanuel Bomze. At the Workshop on Optimization, Game Theory, and Data Analysis 2018 held in celebration of Prof. Bomze's 60-th birthday many speakers recalled how Prof. Bomze changed their careers by challenging them to do things they did not believe themselves to be capable of. I certainly can attest to that experience. Prof. Bomze took the risk of hiring me, a person with then little mathematical background, for a research position in applied mathematics. The person I was then would have been in disbelief of all the things I learned, came to understand and of the skills I developed under Prof. Bomze's supervision. I will be forever grateful for that, for the many times he was so much more optimistic about my research then I was myself, for the countless hours we spent discussing ideas and teaching me how to transform valuable ideas and research into publications of scientific merit. I will always think of him as the prime example of a superior who leads by motivation.

I also want to thank Professor Dr. Mirjam Dür and Professor Daniel Kuhn PhD, MSc for hosting me for three months at the University of Augsburg and the EPFL in Lausanne respectively. The time I spent there was both insightful and enjoyable. I also want to thank their research groups and all the people there who are too numerous to include a list of and who made me feel so welcome and made my time abroad so much easier for me.

Further, I would like to thank all the people who were involved in the VGSCO program. First of all, thanks to all the faculty members who collaborated to make this excellent program a reality. Thanks as well to all the guest lecturers, your courses made me regret not to have studied mathematics much earlier in my life. Especially, I want to thank Professor Kurt Anstreicher PhD, BA whose lectures of convex reformulations on nonconvex QCQPs are at the heart of much of my research. Also I want to thank my fellow PhD candidates who participated in the program and with whom I spent so many enjoyable evenings at Luxor. Among those, I want to especially thank my colleague Michal Kahr for allowing me to use his implementation of the Infectio-Immunization algorithm. Thanks also goes to the Austrian Science Fund (FWF, project W 1260-N35) who provided the funding for this program. On that note, special thanks are in order to Professor Dr. Radu Ioan Boţ for enabling additional funds for the prolongation of my position at the University during these difficult times of the covid crisis.

I also want to thank my co-authors Professor Francesca Maggioni PhD, MSc and Professor Mag. Dr. Georg Pflug for the wonderful collaboration and thoughtful discussions that went into creating our joint paper on the two stage stochastic standard quadratic problem.

Finally, I want to thank my family and friends for supporting me. Especially I want to thank my parents who enabled me to go to university in the first place and who worked so hard for the renovation of my apartment. Also, special thanks goes to my good friend Ghobat Shukoor BSc whose idea to start a company for private lectures allowed me to finance my masters degree.

Statement of Originality

As a researcher I do not wish to take credit for other peoples contributions, and therefore, I want to ease the assessment of what is my original contribution in this text. There are three main reasons that necessitate such a clarification.

The present thesis is in the form of a self contained monograph rather than a concatenation of individual papers. For that reason I included chapters that have introductory character and draw heavily from the works of other people.

Also, as it unfortunately happens, the first part of my research on the application of conic reformulation of QCQPs to robust optimization was independently but in parallel developed by the research group of Grani A. Hanasusanto et al. They used almost identical approaches and were able to publish their work before we, my supervisor and I, did. In an effort to minimize the damage we were able to publish a collection of results that were related to but independent of the work of Hanususanto et al in [14], but for the purpose of this thesis I want to present our results in the way that is closer to what we originally planned to do. We will shortly discuss the overlap of our work with theirs in this section. However, regarding these overlaps, I want to clarify that I do not think my research was on par with theirs. They were able to use the approach to much greater effect, especially with regards to applications and utilization of the modeling power of the approach.

Finally, Chapter 6 is a condensed version of a paper that is a collaboration between Prof. Francesca Maggioni, Prof. Georg Pflug, Prof. Immanuel Bomze and myself. I therefore want to clarify what parts of this collaboration were my personal contribution. Also, I want to make clear that I condensed the paper such that the remaining parts are focused on my work. This editing decision was made with the conciseness of the text in mind and not to diminish the contribution of my coauthors. Also, most of my work is coauthored by my supervisor Prof. Bomze, and we worked under the typical arrangement where the main body of work is carried out by the student while the supervisor gives feedback and enacts quality control. In our case this often came down to improvements of the editing of the papers, streamlining of the presentation of some proofs and overall improving the presentation of the papers as well as decisions on what to publish and lots of patient corrections of typos.

I will use the remainder of this section to give a detailed explanation of what comprises my original contribution in this thesis.

Chapter 1 and 2 are introductory and draw from literature referenced therein. Proofs discussed there are mostly simplified versions of their original counterparts in literature that are presented for the readers convenience, the exception being Theorem 2 which is a slight generalization of the original theorem in [11]. Chapter 3 is an original contribution that is heavily inspired by the results in [3, 27, 86]. Chapter 4 is mostly original, however, there is some overlap with [53] which we explain in greater detail in the remark in Section 4.4. Nonetheless, the main results, i.e. Theorem 17, Theorem 20, Theorem 26 are entirely independent of their work.

Chapter 5 is introductory, and the sources are referenced there. Chapter 6 is a shortened version of the joint paper with Prof. Francesca Maggioni, Prof. Georg Pflug, Prof. Immanuel Bomze and myself. The initial idea for for the two-stage stochastic standard quadratic problem came from Prof. Bomze. Prof. Maggioni and Prof. Pflug contributed the parts that are concerned with stochastic optimization theory, which for the purpose of this text are condensed to parts of the introduction and the discussion in Section 6.2.3. The upper and lower bounds based on quadratic optimization theory constitute my contribution. The numerical experiments were conducted by myself as far as implementation is concerned, except for the first of the two random instance generators, which was implemented by Prof. Maggioni and Prof. Pflug. The design of the experiments is the result of joint discussion and effort of each of the coauthors.

Chapter 7 exhibits a lot of overlapping with [84, 46]. We explain the general reformulation strategy that was used in these papers, and that we intended to use if it wasn't for the parallel developments. We then apply the general reformulation strategy to various instances of robust optimization. The applications in Section 7.4 were independently developed by us but eventually published here [84] by Hanasusanto et al. The applications in Section 7.4.2 was not a parallel development, but developed and published in [84] first. What we contributed here are slight generalizations that were not considered in the original publication by Hanasusanto et al. The rest of the chapter is our original contribution although it is of course closely related to the works of Hanasusanto et al. since it deals with the application of quadratic optimization theory to robust optimization.

To the best of our knowledge, the entirety of Chapter 8 is original up to the introduction where we explain some preliminaries about endogenously robust optimization from [56, 62].

Abstract

The present thesis is an amalgamation of four papers: [14], which was recently published in Mathematical Methods of Operations Research, [13] which is currently under major revision for SIAM-Journal on Optimization, [15] which was recently submitted to European Journal of Operations Research, and [45] that is currently under construction and soon to be submitted. These papers are connected by the overarching theme of using conic and quadratic optimization in the context of robust and stochastic optimization. However, for the purpose of this book these papers were re-edited to form a self-contained monograph that we organized in three parts.

The first part is concerned with the quadratically constrained quadratic optimization problems (QCQPs) and their convex reformulations. We give an introduction to this topic in the first two sections of that part of the book. QCQPs form a large and versatile class of optimization problems which are in general nonconvex. However one can obtain exact convexifications by lifting the space of variables and restricting them to an appropriately structured convex set. The cones of set-completely positive matrices (for example the cone of positive semidefinite matrices, or the completely positive matrix cone) are a central ingredient for the characterizations of these lifted convex sets. We present two new results on this topic.

The first one is a generalization of a known result that is based on Pataki's rank theorem. It allows us to reformulate homogeneous QCQPs that fulfill a certain geometric condition. Based on that, we can derive a generalized version of the well known S-lemma and infer some geometric properties of the so called joint numerical range, which is another central object in the study of QCQPs.

The second result is about reduced conic reformulations of QCQPs, where we proof the exactness of a reformulation that involves a lifting of the

variables into a smaller space compared to traditional approaches. We contextualize this result in the theory of matrix completion and show that these to fields have mutual implications. In fact, our exactness result implies a new result in set-completely positive matrix completion.

In the second part of the book, we apply quadratic optimization theory to stochastic and robust optimization problems. We start out by considering a two-stage stochastic version of the well known standard quadratic optimization problem. We consider lower bounds based on a scenario approach where the probability measures is discretized and dissected in order to obtain a chain of lower bounds. For each level of the dissection one has to solve a number of quadratic optimization problems for which we provide lower bounds based on the reduced conic reformulation mentioned above and upper bounds based on the infection-immunization-dynamics algorithm for standard quadratic optimization problems. We investigate the efficacy of these methods in an array of numerical experiments.

We also discuss a general strategy of using convex reformulations of QCQPs in order to obtain finite reformulations of the semi-infinite constraints present in robust optimization for the case where the left-hand side depends quadratically on the index. We apply this general strategy to various instances of robust optimization such as adjustable robust optimization and convex-quadratic robust optimization. We also investigate some theoretical aspects of the problem associated with dual attainability of the QCQP reformulations when applying the general strategy. Following a recent development in literature, we also show that adjustable robust optimization theory also can be used to obtain an adjustable robust reformulation of a certain class of QCQPs. This reformulation is in turn amendable to the general strategy so that we can obtain a lower bound for the original problem. We compare this lower bound to alternative approaches in a set of numerical experiments.

The final part of the book is concerned with a new concept of robust optimization based on the recently introduced setting of endogenously uncertain optimization. In this setting the decision maker may influence parts of the uncertainty that she is facing, for example by investing in market research. This influence is typically modeled as a discrete decision on the uncertainty regime under which the decision maker wishes to operate. However, this situation gives rise to a new phenomenon: Each of the different uncertainty regimes has its own robustly optimal solutions. But these solutions may not only vary with respect to their worst case performance but, due to the different structures of the uncertainty in each individual regime, may also exhibit different qualities with respect to best case performance, regret and predictability. A decision maker who wishes to have a performance guarantee, i.e. wants to select one of the robustly optimal solutions, may still want to take these characteristics into account as additional decision criteria. We conceptualize such a desire as uncertainty preference. In the final chapter of the book we introduce the respective mathematical model and discuss solution methods and modeling capabilities. We also apply this new framework to the endogenously uncertain shortest path problem and provide some numerical experiments that show that we can solve this problem using mixed integer conic optimization techniques and that the obtained solutions have a meaningful impact on the decision making process.

Abstrakt

Die vorliegende Arbeit ist eine Zusammenführung von vier Artikeln: [14], welcher vor Kurzem in Mathematical Methods of Operations Research erschienen ist, [13], welcher sich derzeit im Prozess der "Major-Revision" für SIAM-Journal on Optimization befindet, [15] welcher vor Kurzem bei European Journal of Operations Research eingereicht wurde und [45], an welchem zum Zeitpunkt der Einreichung noch gearbeitet wurde um alsbald eingereicht zu werden. Zum Zwecke der Erstellung dieses Buches wurden die besagten Texte neu editiert und zu einer selbsttragenden Monographie zusammengeführt, die in drei Teilen organisiert ist.

Der erste Teil beschäftigt sich mit quadratisch beschränkten, quadratischen Optimierungsproblemen (QCQPs) und deren konvexen Reformulierungen. Die ersten beiden Kapitel dienen der Einführung in diese Materie. Bei QCQPs handelt es sich um eine große und ausdrucksstarke Klasse von, im allgemeinen nicht-konvexen, Optimierungsproblemen. Es ist jedoch oft möglich sie zu konvexifizieren indem man die zulässigen Menge der Entscheidungsvariablen in einem höher-dimensionalen Raum einbettet, wobei eine geeignete konvexe Teilmenge in diesen Räumen beschrieben werden muss. Die Kegel der mengenspezifischen komplett-positiven Matrizen (set-completely postive matrices), wie etwa der Kegel der positive semidefiniten Matrizen oder der Kegel der komplett-positiven Matrizen, sind zentrale Zutaten in diesem Unterfangen. Wir präsentieren zwei neue Resultate auf diesem Forschungsfeld.

Bei dem Ersten handelt es sich um eine Verallgemeinerung eines bekannten Resultats, das auf Pataki's Rang-Theorem beruht. Es ermöglicht die Reformulierung von homogenen QCQPs die eine bestimmte geometrische Voraussetzung erfüllen. Darauf aufbauend können wir eine Verallgemeinerung des weithin bekannten S-Lemmas herleiten, und erhalten darüber hinaus Einsichten in geometrische Eigenschaften des so genannten gemeinsamen numerischen Wertebereichs, der ein weiteres zentrales Objekt in der Analyse von QCQPs darstellt.

Das Zweite Resultat behandelt reduzierte konische Reformulierungen von QCQPs, bei denen, wie bei der traditionellen Herangehensweise, die zulässigen Menge der Entscheidungsvariablen in einem höher-dimensionalen Raum eingebettet wird, dessen Dimension aber niedriger ist als bei bereits bekannten Reformulierungen. Wir diskutieren dieses Resultat im Kontext von Matrixvervollständigungsproblemen und zeigen auf, dass die beiden Gebiete eng zusammenhängen. Tatsächlich erlaubt unser Reformulierungsresultat die Herleitung eines neuartigen Resultats zu Vervollständigung von mengenspezifisch komplett positiven Matrizen.

Im zweiten Teil des Buches wenden wir die oben beschriebene Theorie auf Probleme der Stochastischen und Robusten Optimierung an. Als erstes diskutieren wir eine zwei-stufige, stochastische Version des Standard-Quadratischen Optimierungsproblems (StQP). Wir beschäftigen uns mit unteren Schranken basierend auf einem Szenario-Ansatz, bei dem das zugrundeliegende Wahrscheinlichkeitsmaß diskretisiert und unterteilt wird um eine Hierarchie an unteren Schranken zu erstellen. Für jeden der unterschiedlichen Grade der Unterteilung lösen wir die sich ergebenden Subprobleme mit Hilfe der unteren und oberen Schranken die wir aus der reduzierten konischen Reformulierung des entsprechenden QCQPs gewinnen, sowie einer oberen Schranke auf Basis des Infection-Immunization-Dynamics Algorithmus für StQPs. Wir untersuchen die Effektivität dieser Methoden in einer Reihe von numerischen Experimenten.

Weiters diskutieren wir eine allgemeine Strategie für die Verwendung von konvexen Reformulierung von QCQPs zum Zwecke der Reformulierung von semi-infiniten Nebenbedingungen mit quadratischer Abhängigkeit vom Index, die in der Robusten Optimierung auftreten. Wir wenden diese Strategie auf verschiedene Probleme der Robusten Optimierung an, wie etwa der Anpassungsfähigen Robust Optimierung (adjustable robust optimization) und der Konvex-Quadratischen Robusten Optimierung. Wir diskutieren darüber hinaus theoretische Aspekte des, für die Anwendung der allgemeinen Strategie kritischen, Problems der vollen, starken Dualität. Inspiriert durch jüngste Entwicklungen in der Literatur untersuchen wir die Anwendung von Anpassungsfähiger Robust Optimierung für die Reformulierung einer bestimmten Klasse von QCQPs. Diese Reformulierung des QCQPs als anpassungsfähiges, robustes Optimierungsproblem ist ihrerseits geeignet für die allgemeine Reformulierungsstrategie, was uns erlaubt untere Schranken für das ursprüngliche QCQP herzuleiten. Wir vergleichen diese Schranken mit anderen Schranken aus der Literatur im Rahmen von numerischen Experimenten.

Im Finalen Teil des Buches führen wir ein neues Konzept für die Robuste Optimierung ein, basierend auf der relativ neuen Idee der Optimierung mit endogener Unsicherheit. Dabei wird angenommen, dass die Entscheidungsträger*in Einfluss auf die Art der Unsicherheit nehmen kann, etwa durch Investition in Markt-Forschung oder genauere Maschinen. Dieser Einfluss wird typischerweise über diskrete Entscheidungsvariablen modelliert, die steuern, welche Art von Unsicherheit sich die Entscheidungsträger*in aussetzen will. Diese Konstellation bedingt aber ein neues Phänomen: Jedes der verschiedenen Unsicherheitsregime hat eine andere robuste, optimale Lösung. Doch diese unterschiedlichen Lösungen unterscheiden sich nicht nur hinsichtlich ihrer Performancegarantie sondern auch hinsichtlich anderer Qualitäten wie etwa der Vorhersehbarkeit der Performance, der möglichen Reue im Nachhinein und der potentiell bestmöglichen Performance. Eine Entscheidungsträger*in die vornehmlich an einer Performancegarantie interessiert ist, möchte diese zusätzlichen Qualitäten vielleicht trotzdem berücksichtigen. Wir beschreiben eine solche Haltung als Unsicherheitspräferenz. Im letzten Kapitel des Buches führen wir die entsprechenden mathematischen Modelle ein und diskutieren mögliche Lösungsmethoden. Wir wenden dieses Konzept auf das Problem des Kürzesten Pfads mit endogener Unsicherheit an und zeigen, dass man das resultierende Problem mit Standardmethoden der diskreten, konischen Optimierung lösen kann und dass die Ergebnisse dazu geeignet sind bei der Entscheidungsfindung zu assistieren.

Contents

List of Figures

Part I

Convex Reformulations of Quadratically Constrained Quadratic Optimization Problems

Chapter 1

Quadratically constrained quadratic optimization

A quadratically constrained quadratic optimization problem (QCQP) consists of minimizing a quadratic function subject to quadratic constraints, formally given by

$$\inf_{\mathbf{x}\in\mathcal{K}} \left\{ \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega \colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x} + \mathbf{a}_i^\mathsf{T}\mathbf{x} \leq b_i,\ i \in [1\!:\!m] \right\} \tag{1.1}$$

where Q, A_i are real, symmetric matrices of order n and $\mathbf{q}, \mathbf{a}_i$ are real vectors, b_i are real numbers. $\mathcal{K} \subseteq \mathbb{R}^n$ is a cone which one could choose to be any polyhedral- or quadratically representable cone without leaving the realm of quadratically constrained quadratic optimization and many of the results in Chapter 2 will work under much weaker assumptions on $\mathcal{K}$. However, for our purpose we mostly will have $\mathcal{K} \in \left\{\mathbb{R}^n, \mathbb{R}^n_+, \mathcal{SOC}^n\right\}$. In general, QCQPs are NP-hard as they contain many NP-hard problems as special cases (see e.g. [63]). However, there are numerous solution approaches, which we will not discuss here since we are mainly interested in one particular approach, i.e. convex reformulation. The introductory part of this book regarding this approach will be quite exhaustive so that in this chapter we merely discuss the expressiveness of QCQPs and some notable examples.

1.1 Expressiveness of QCQPs

The class of QCQPs is an extremely versatile class of optimization problems. Many of the most important instances of convex and nonconvex optimization can be cast as QCQPs, these include

Linear optimization: The generic linear optimization problem

$$\inf_{\mathbf{x}\in\mathbb{R}^n_+} \left\{ \mathbf{q}^\mathsf{T}\mathbf{x} \colon \mathbf{a}_i^\mathsf{T}\mathbf{x} \leq b_i,\ i \in [1\!:\!m] \right\}$$

can be obtained from (1.1) by setting all matrices to the zero matrix and $\mathcal{K} = \mathbb{R}^n_+$.

Second Order Cone optimization: A generic second order cone constraint given by

$$\|\mathsf{A}\mathbf{x} + \mathbf{b}\| \leq \mathbf{a}^\mathsf{T}\mathbf{x} + d$$

can be reformulated using the fact that $\|\mathsf{A}\mathbf{x} + \mathbf{b}\| \geq 0$ so that $\|\mathsf{A}\mathbf{x} + \mathbf{b}\| \leq \mathbf{a}^\mathsf{T}\mathbf{x} + d \Leftrightarrow \|\mathsf{A}\mathbf{x} + \mathbf{b}\|^2 \leq (\mathbf{a}^\mathsf{T}\mathbf{x} + d)^2$ and $\|\mathsf{A}\mathbf{x} + \mathbf{b}\|^2 = \mathbf{x}^\mathsf{T}\mathsf{A}^\mathsf{T}\mathsf{A}\mathbf{x} + 2\mathbf{b}\mathsf{A}\mathbf{x} + \mathbf{b}^\mathsf{T}\mathbf{b}$. Note, that the generic second order cone constraint includes the class of rotated second order cone constraints. In fact the class second order cone optimization problems is exactly the class of convex QCQPs, i.e. QCQPs where all the matrices are positive semidefinite (see [1]).

Mixed Integer Quadratic Optimization The mixed integer quadratic optimization problem

$$\inf_{\mathbf{x}} \left\{ \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega \colon \mathbf{x} \in \mathbb{R}^{n_1}_+ \times \mathbb{Z}^{n_2},\ \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x} + \mathbf{a}_i^\mathsf{T}\mathbf{x} \leq b_i,\ i \in [1\colon m] \right\}$$

can be reformulated as a QCQP in case we have a bound on each integer variable. W.l.o.g assume that for an for $x \in \mathbb{Z}$ we have $x \geq 0$ and $x \leq z_{max}$ is an upper bound. Then the integer constraint is equivalent to $x = \sum_{i=1}^{z_{max}} iy_i,\ \mathbf{y} \in \{0,1\}^{z_{max}}$ and $\mathbf{e}^\mathsf{T}\mathbf{y} = 1$. Further, $y_i \in \{0,1\},\ i \in [1 : z_{max}]$ is equivalent to $y_i^2 - y_i = 0,\ i \in [1\colon z_{max}]$ so that $x \in \mathbb{Z}$ can be replaced by two linear and z_{max} quadratic equalities.

Bilinear Optimization A bilinear optimization problem is given by

$$\min_{\mathbf{x}\in\mathcal{K}_1, \mathbf{y}\in\mathcal{K}_2} \left\{ \mathbf{x}^\mathsf{T}\mathsf{A}_0\mathbf{y} + \mathbf{b}_0^\mathsf{T}\mathbf{x} + \mathbf{c}_0^\mathsf{T}\mathbf{y} + d_0 \colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{y} + \mathbf{b}_i^\mathsf{T}\mathbf{x} + \mathbf{c}_i^\mathsf{T}\mathbf{y} \leq d_i, i \in [1\colon m] \right\}$$

where $\mathsf{A}_i \in \mathbb{R}^{n_1 \times n_2},\ \mathbf{b}_i \in \mathbb{R}^{n_1},\ \mathbf{c}_i \in \mathbb{R}^{n_1},\ i \in [0\colon m]$, and $\mathcal{K}_i \subseteq \mathbb{R}^{n_i},\ i \in [1\colon 2]$. This can be seen as a special case of a QCQP where the symmetric matrices are of the form $\mathsf{Q}_i = \mathsf{M}(\mathsf{O}, 0.5\mathsf{A}_i, \mathsf{O})$. Also, any QCQP can be restated as a bilinear optimization problem by introducing an auxiliary variable $\mathbf{y}$ and decomposing the quadratic terms $\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}$ into a bilinear term $\mathbf{x}^\mathsf{T}\mathbf{y}$ and an additional constraint $\mathbf{y} = \mathsf{Q}\mathbf{x}$.

Semidefinite and Set-Completely Positive Optimization A set-completely positive optimization problem

$$\min_{\mathsf{X}\in\mathcal{S}^n} \{\mathsf{C} \bullet \mathsf{X} \colon \mathsf{X} \in \mathcal{CPP}(\mathcal{K}),\ \mathsf{A}_i \bullet \mathsf{X} = b_i,\ i \in [1\colon m]\}$$

is linear up to the conic constraint (we discuss the set-completely positive cones in Appendix A.2). By Caratheodory's Theorem we have $\mathsf{X} \in \mathcal{CPP}(\mathcal{K})$ if and only if $\mathsf{X} = \sum_{i=1}^{k} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}},\ \mathbf{x}_i \in \mathcal{K},\ i \in [1\!:\!k]$ where $k = \dim \mathcal{S}^n$ (see Carathéodory's Theorem in [73, Theorem 17.1]). The matrix equality can be read entrywise as $(\mathsf{X})_{lj} = \sum_{i=1}^{k} (\mathbf{x}_i)_l (\mathbf{x}_i)_j$ which is a quadratic constraint. If $\mathcal{K} = \mathbb{R}^n$ we recover the class of semidefinite optimization problems (see Appendix A.2).

1.2 Notable problem classes

Some examples of QCQPs have received special interest in the literature due to their usefulness or their particular importance to some other subject. We discuss two of these examples here, as they are also important later in the text.

1.2.1 The Standard Quadratic Optimization Problem

The standard quadratic optimization problem (StQP) is defined as

$$\min_{\mathbf{x}, \in \mathbb{R}^n_+} \left\{ \mathbf{x}^{\mathsf{T}} \mathsf{Q} \mathbf{x} \colon \mathbf{e}^{\mathsf{T}} \mathbf{x} = 1 \right\} \tag{1.2}$$

so that the feasible set is the standard simplex (a.k.a the probability simplex). Note, that the homogeneous case is actually generic since $\mathbf{x}^{\mathsf{T}} \mathsf{Q} \mathbf{x} + \mathbf{q}^{\mathsf{T}} \mathbf{x} = \mathbf{x}^{\mathsf{T}} \left(\mathsf{Q} + 0.5 \left(\mathbf{q} \mathbf{e}^{\mathsf{T}} + \mathbf{e} \mathbf{q}^{\mathsf{T}} \right) \right) \mathbf{x}$, exploiting the fact that $\mathbf{e}^{\mathsf{T}} \mathbf{x} = 1$. The StQP appears in various different areas of mathematics. In [60] it was shown that the size of the maximum clique of a graph is given by $(1 - f^*)^{-1}$ where f^* optimal value of an StQP where the symmetric matrix is specified as the adjacency matrix of the graph. Connections between StQPs, game theory and dynamical systems are established in [16] where these insight are also used to derive an algorithm for finding maximum cliques in a graph. An application to machine learning is provided in [68] wher so called e dominant sets of weighted graphs are introduced, which also can be identified by solving an StQP that is parametrized by the graph. This approach is useful for image segmentation. Portfolio selection problems can also be cast as StQPs (see [17]). Further, StQPs play a central role in copositive optimization as discussed in [11], where a copositive reformulation of the StQP is derived as well.

In [22], a powerful algorithm based on immunization infection dynamics of evolutionary games was introduced that can be used to locally solve StQPs. The algorithm has local convergence guarantees and can be warm started at arbitrary points in the standard simplex. In Chapter 6 we will make use of this algorithm, however we do not explain the details of it here, since that would require us to introduce evolutionary game theory and dynamical systems, which is way beyond the scope of this text.

1.2.2 The Trust-Region Subproblem

The trust-region subproblem (TRS) is defined as

$$\min_{\mathbf{x},\in\mathbb{R}^n} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega \colon \|\mathbf{x}\| \leq 1\right\}, \tag{1.3}$$

that is, the minimization of a (possibly nonconvex) quadratic function over the unit ball. This problem arises in the context of the so called trust region methods for nonconvex unconstrained optimization. These methods consecutively minimize quadratic models of the objective function over a region where these models are trusted to be faithful. The minimizer of an iteration becomes the focal point for the construction of the quadratic model for the next iteration. The TRS is a very well studied problem ([32] and references therein) and is known to be solvable in polynomial time (see [43]). Also, several generalizations have been considered in the literature such as TRS with additional linear constraints (see [27] and references therein) and additional convex quadratic constraints (e.g. the Two-Trust-Region Subproblem, TTRS, see [30]).

In [72] it was shown that (1.3) can be reformulated as a semidefinite optimization problem. This is also possible for some of its generalizations such as the ones considered in [29, 86]. Interestingly, an exact convex reformulation of the TTRS is not known despite the exploitation of several families of cuts (see [87]) and despite the fact that it can be solved in polynomial time (see [10]).

Chapter 2

Core concepts for convex reformulations of QCQPs

General QCQPs are hard, and form a large and quite versatile class of optimization problems. Neither the objective function nor the feasible set of a QCQP need to be convex, the latter may even be disconnected. One way to (approximately) solve QCQPs is to look for convex relaxations which in the best case yield exact reformulations. In this chapter we will review two prominent strains of literature that deal with the problem of obtaining convex reformulations.

2.1 The convex hull of the lifted feasible set

An important strategy for achieving such reformulations is to lift the space of variables, thereby linearizing the quadratic terms and convexifying the lifted feasible set, such that its extreme points correspond to rank-one matrices, which will decompose into vectors feasible to the original problem:

Theorem 1. *Let* $\mathcal{F} \coloneqq \left\{ \mathbf{x} \in \mathcal{K} \colon \mathbf{x}^\mathsf{T} \mathsf{A}_i \mathbf{x} + \mathbf{a}_i^\mathsf{T} \mathbf{x} \leq b_i ,\ i \in [1\!:\!m] \right\} \subseteq \mathbb{R}^n$ *be a feasible set of a QCQP and denote by*

$$\mathcal{G}(\mathcal{F}) \coloneqq \operatorname{clconv} \left\{ (\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T}) : \mathbf{x} \in \mathcal{F} \right\} ,$$

where $\operatorname{clconv}(\mathcal{A})$ *stands for the closure of the convex hull of a set* $\mathcal{A}$*.*
Then for any $\mathsf{Q} \in \mathcal{S}^n$ *and* $\mathbf{q} \in \mathbb{R}^n$ *we have*

$$\inf_{\mathbf{x} \in \mathcal{F}} \left(\mathbf{x}^\mathsf{T} \mathsf{Q} \mathbf{x} + \mathbf{q}^\mathsf{T} \mathbf{x} + \omega \right) = \inf_{(\mathbf{x}, \mathsf{X}) \in \mathcal{G}(\mathcal{F})} \left(\operatorname{trace}(\mathsf{Q}\mathsf{X}) + \mathbf{q}^\mathsf{T} \mathbf{x} + \omega \right) .$$

Proof. See, e.g. [27, 41]. For the readers convenience we repeat the argument here. We refer to the QCQP as (P) and to the reformulation as (R). Let $\mathbf{x}$ be feasible for (P), then $\left(\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T}\right)$ is feasible for with identical objective function value given by $\mathsf{Q} \bullet \mathbf{x}\mathbf{x}^\mathsf{T} + \mathbf{c}^\mathsf{T}\mathbf{x} + \omega = \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} + \omega$. Thus val(R)$\leq$ val(P). For the converse, let $(\mathbf{x}, \mathsf{X})$ be optimal for (R). Then $(\mathbf{x}, \mathsf{X}) = \sum_{i=1}^{k} \lambda_i \left(\mathbf{x}_i, \mathbf{x}_i\mathbf{x}_i^\mathsf{T}\right)$ with $\mathbf{x}_i \in \mathcal{F}$, $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0,\ i \in [1\!:\!k]$ since $\mathcal{G}(\mathcal{F})$ is closed and convex. We have $\mathbf{x}_i^\mathsf{T}\mathsf{Q}\mathbf{x}_i + \mathbf{c}^\mathsf{T}\mathbf{x}_i \geq$ val(P), $i \in [1:k]$ and thus val(R)$=$ $\mathsf{Q} \bullet \mathsf{X} + \mathbf{c}^\mathsf{T}\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i^\mathsf{T}\mathsf{Q}\mathbf{x}_i + \mathbf{c}^\mathsf{T}\mathbf{x}_i \geq \sum_{i=1}^{k} \lambda_i$val(P)$=$ val(P). $\square$

We want to highlight the fact that in the above theorem no convexity assumptions are imposed on $\mathcal{F}$. Convexity enters in the form of th central object in the theorem that is $\mathcal{G}(\mathcal{F})$. Its characterization is the major challenge when employing the reformulation strategy depicted in Theorem 1 and a general workable description of $\mathcal{G}(\mathcal{F})$ is not known. There are, however, characterizations for specific instances of $\mathcal{F}$. In practice such characterizations are regularly given by a conic intersection involving some matrix cone $\mathcal{C}$ and appropriate linear constraints parametrized by matrices $\bar{\mathsf{A}}_i$ and real numbers $\bar{b}_i, i \in [1\!:\!\bar{m}]$. Thus, typically we have reformulations of the form:

$$\inf_{(\mathbf{x},\mathsf{X}) \in \mathcal{G}(\mathcal{F})} \left(\mathsf{Q} \bullet \mathsf{X} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega\right) = \inf_{\mathsf{Y} \in \mathcal{C}} \{\mathsf{M} \bullet \mathsf{Y} \colon \bar{\mathsf{A}}_i \bullet \mathsf{Y} \leq \bar{b}_i\,,\ i \in [1\!:\!\bar{m}]\}\,,$$

where M, $\bar{\mathsf{A}}_i$ and $\bar{b}_i$ depend on the problem data. Important examples for such reformulations can be found in [3, 25, 28, 29, 41, 86], some of which we will discuss in more detail later in the text.

2.1.1 A first step: The Shor relaxation

As stressed in the previous section, the key problem when employing Theorem 1 is to characterize the set $\mathcal{G}(\mathcal{F})$, which of course depends heavily on the choice of $\mathcal{F}$. A natural starting point for its construction is can be inferred from the so called Shor relaxation whose construction we will present shortly. The discussion in this and many following sections will feature the so called set-completely positive matrix cones defined as $\mathcal{CPP}^n(\mathcal{K}) \coloneqq \operatorname{conv}\left\{\mathbf{x}\mathbf{x}^\mathsf{T} \colon \mathbf{x} \in \mathcal{K}\right\}$, for a cone $\mathcal{K} \subseteq \mathbb{R}^n$. We provide a detailed introduction of these cones and their dual cones $\mathcal{COP}(\mathcal{K})$, i.e. the set-copositive matrix cone in Appendix A.2. Readers who are not familiar with these objects might want to read that section first before continuing here.

In the following derivation of the Shor relaxation we will be using the fact that by Lemma 54 we have ext$\mathcal{CPP}(\mathcal{K}) = \left\{\mathbf{x}\mathbf{x}^\mathsf{T} \colon \mathbf{x} \in \mathcal{K}\right\}$ and the shorthand notation $\mathsf{Y}(\mathbf{x}, \mathsf{X})$

explained in Appendix B. The construction proceeds as follows:

$$
\begin{aligned}
&\min_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+\mathbf{q}^\mathsf{T}\mathbf{x}+\omega\colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+\mathbf{a}_i^\mathsf{T}\mathbf{x}\le\alpha_i,\ i\in[1\!:\!m]\}\\
&=\min_{(x_0,\mathbf{x})\in\mathbb{R}_+\times\mathcal{K}}\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+x_0\mathbf{q}^\mathsf{T}\mathbf{x}+\omega x_0^2\colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+x_0\mathbf{a}_i^\mathsf{T}\mathbf{x}\le\alpha_i,\ i\in[1\!:\!m],\ x_0^2=1\}\\
&=\min_{(x_0,\mathbf{x})\in\mathbb{R}_+\times\mathcal{K}}\begin{pmatrix} w & \frac{1}{2}\mathbf{q}^\mathsf{T}\\ \frac{1}{2}\mathbf{q} & \mathsf{Q}\end{pmatrix}\bullet\begin{pmatrix} x_0^2 & x_0\mathbf{x}^\mathsf{T}\\ x_0\mathbf{x} & \mathbf{x}\mathbf{x}^\mathsf{T}\end{pmatrix}\\
&\qquad\text{s.t.}:\begin{pmatrix} 0 & \frac{1}{2}\mathbf{q}^\mathsf{T}\\ \frac{1}{2}\mathbf{q} & \mathsf{A}_i\end{pmatrix}\bullet\begin{pmatrix} x_0^2 & x_0\mathbf{x}^\mathsf{T}\\ x_0\mathbf{x} & \mathbf{x}\mathbf{x}^\mathsf{T}\end{pmatrix}\le\alpha_i\\
&\qquad\qquad\begin{pmatrix} 1 & \mathbf{o}^\mathsf{T}\\ \mathbf{o} & \mathsf{O}\end{pmatrix}\bullet\begin{pmatrix} x_0^2 & x_0\mathbf{x}^\mathsf{T}\\ x_0\mathbf{x} & \mathbf{x}\mathbf{x}^\mathsf{T}\end{pmatrix}=1\quad\left(\text{set } \mathbf{y}=\begin{pmatrix} x_0\\ \mathbf{x}\end{pmatrix}\right)\\
&=\min_{\mathbf{y}\in\mathbb{R}_+\times\mathcal{K}}\left\{\bar{\mathsf{Q}}\bullet\mathbf{y}\mathbf{y}^\mathsf{T}\colon\ \bar{\mathsf{A}}_i\bullet\mathbf{y}\mathbf{y}^\mathsf{T}\le\alpha_i,\ i\in[1\!:\!m],\ y_0^2=1\right\}\\
&=\min_{\mathsf{Y}}\left\{\bar{\mathsf{Q}}\bullet\mathsf{Y}\colon \mathsf{Y}\in\mathrm{ext}\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K}),\ \bar{\mathsf{A}}_i\bullet\mathsf{Y}\le\alpha_i,\ i\in[1\!:\!m],\ \mathsf{Y}_{0,0}=1\right\}\\
&\ge\min_{\mathsf{Y}}\left\{\bar{\mathsf{Q}}\bullet\mathsf{Y}\colon \mathsf{Y}\in\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K}),\ \bar{\mathsf{A}}_i\bullet\mathsf{Y}\le\alpha_i,\ i\in[1\!:\!m],\ \mathsf{Y}_{0,0}=1\right\}\\
&=\min_{\mathsf{X},\mathbf{x}}\{\mathsf{Q}\bullet\mathsf{X}+\mathbf{q}^\mathsf{T}\mathbf{x}+\omega\colon \mathsf{A}_i\bullet\mathsf{X}+\mathbf{a}_i^\mathsf{T}\mathbf{x}\le\alpha_i,\ i\in[1\!:\!m],\mathsf{Y}(\mathbf{x},\mathsf{X})\in\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K})\}
\end{aligned}
$$

The latter problem is the desired Shor relaxation. The core idea was introduced in [76] albeit for the special case of $\mathcal{K}=\mathbb{R}^n$, in which case $\mathcal{CPP}\left(\mathbb{R}_+\times\mathbb{R}^n\right)=\mathcal{S}_+^{n+1}$ (see Proposition 53 in Appendix A). Also, in case the problem has no linear terms in $\mathbf{x}$, it is not necessary to introduce the auxiliary variable x_0 and the final problem will only involve the simpler matrix cone $\mathcal{CPP}\left(\mathcal{K}\right)$. Note, that the only inequality occurs in the forth line when we replace $\mathrm{ext}\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K})$ by its convex hull, i.e. $\mathcal{CPP}\left(\mathbb{R}_+\times\mathcal{K}\right)$. The linearizion itself does not incur a cost as long as we are optimizing over the nonconvex set $\mathrm{ext}\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K})$. But the convexification might introduce extreme points and rays that are not rank-1 which causes the relaxation gap as the feasible set of the Shor relaxation falls short of describing $\mathcal{G}\left(\mathcal{F}\right)$. We will illustrate this point by considering some examples:

Example 1. *Consider the following QCQP and its Shor relaxation to its right:*

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^2}\ & q_{11}x_1^2+2q_{12}x_1x_2+q_{22}x_2^2\\
\text{s.t.}:\ & 2x_1^2+x_2^2\le 12,\\
& 3x_1^2+3x_2^2+2x_1x_2\le 6\\
& 3x_1^2+3x_2^2-2x_1x_2\ge 6,\\
& 3x_1^2+3x_2^2\ge 6,
\end{aligned}
\qquad\qquad
\begin{aligned}
\min_{\mathsf{X}\in\mathcal{S}_+^2}\ & q_{11}\mathsf{X}_{11}+2q_{12}\mathsf{X}_{21}+q_{22}\mathsf{X}_{22}\\
\text{s.t.}:\ & 2\mathsf{X}_{11}+\mathsf{X}_{22}\le 12,\\
& 3\mathsf{X}_{11}+3\mathsf{X}_{22}+2\mathsf{X}_{12}\le 6,\\
& 3\mathsf{X}_{11}+3\mathsf{X}_{22}-2\mathsf{X}_{12}\ge 6,\\
& 3\mathsf{X}_{11}+3\mathsf{X}_{22}\ge 6,
\end{aligned}
$$

Figure 2.1 depicts the feasible set of the Shor relaxation which is the intersection of the positive semidefinite cone and three halfspaces. The rank-1 matrices are at

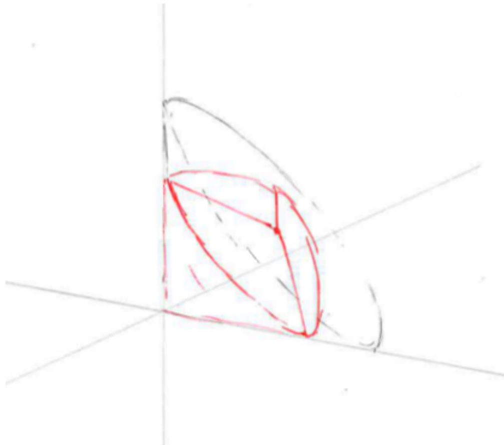


Figure 2.1

Shor relaxation with one non-dyadic extreme point

the boundary of $\mathcal{S}_+^n$ *and for* $n = 2$ *this boundary is entirely comprised out of rank-1 matrices so that it is actually* $\mathrm{ext}\mathcal{CPP}(\mathbb{R}^2)$*. However, since we consider the convex hull of the latter, i.e.* $\mathcal{S}_+^2$*, we produced an extreme point in the interior of* $\mathcal{S}_+^2$*, which thus is not rank-1. For certain choices of the objective function coefficients there will therefore be a gap between the two optimization problems (actually it is those coefficients where the negative of the associated vector is a normal vector to the feasible set at the non-dyadic extreme point). We also see that by tilting one of the hyperplanes we can push the problematic extreme point arbitrarily deep into the psd-cone and away from the intersections of* $\mathrm{ext}\mathcal{CPP}(\mathbb{R}^2)$ *and the three halfspaces, which illustrates that the Shor relaxation can be an arbitrarily bad approximation. On the bright side, we also see that even if we are far from describing* $\mathcal{G}(\mathcal{F})$*, the Shor relaxation can be exact for some choices of the objective function coefficients.*

Example 2. *Now, consider another pair of QCQP and Shor relaxation:*

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^2_+} q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2 \\ \text{s.t.}: 3x_1^2 + 1x_2^2 = 6 \\ x_1^2 + 3x_2^2 = 6, \end{aligned} \qquad \begin{aligned} \min_{\mathsf{X}\in\mathcal{CPP}(\mathbb{R}^2_+)} q_{11}\mathsf{X}_{11} + 2q_{12}\mathsf{X}_{21} + q_{22}\mathsf{X}_{22} \\ \text{s.t.}: 3\mathsf{X}_{11} + \mathsf{X}_{22} = 6, \\ \mathsf{X}_{11} + 3\mathsf{X}_{22} = 6, \end{aligned}$$

The quadratic equality system as exactly one nonnegative solution which is $\left(\sqrt{3/2}, \sqrt{3/2}\right)$*. But when we look at the depiction of the feasible set of the Shor relaxation in Figure 2.2 we see that that we have multiple feasible points, two of which are extreme but only one of them is rank-1. Again, it is the convexification that gave rise to this spurious extreme point, however, this time it is not the constellation* *of the linearized equalities but the geometry of the convex hull of* $\mathrm{ext}\mathcal{CPP}(\mathbb{R}^2_+)$*, i.e.* $\mathcal{CPP}\left(\mathbb{R}^2_+\right) = \mathcal{S}^2_+ \cap \mathcal{N}_2$ *that birthed the problem.*

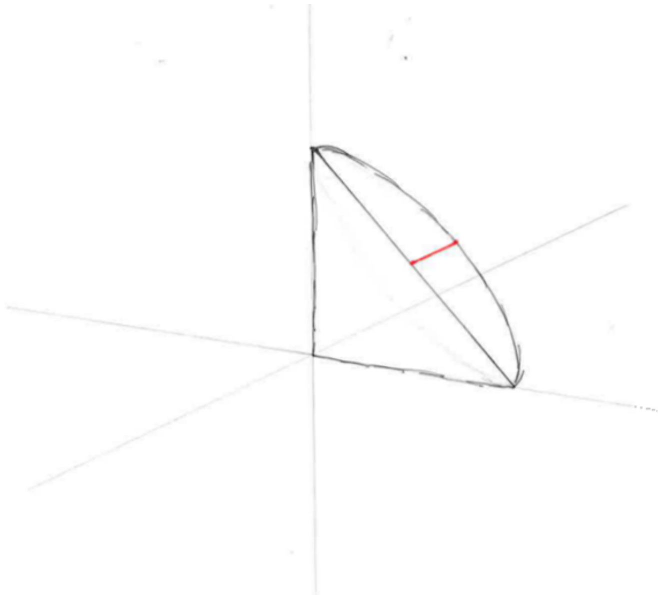


Figure 2.2

Shor relaxation with only one dyadic extreme point.

The above discussion makes it apparent that the Shor relaxation is not necessarily tight, since its feasible set $\mathcal{F}_{Shor}$ can have extreme points that are not rank one matrices. However, if we find that the optimal solution of the Shor relaxation is a rank one matrix, then we have solved the original QCQP. If we can show that a rank one solution exists for any choice of coefficients $(\mathsf{Q}, \mathbf{q}, \omega)$, we have indeed shown that $\mathcal{F}_{Shor} = \mathcal{G}(\mathcal{F})$ (see e.g. [86] for a detailed argument for this claim). But in general $\mathcal{F}_{Shor} \supseteq \mathcal{G}(\mathcal{F})$ so that one has to find a strengthening of the Shor relaxation and/or operate under more strict assumptions if the relaxation gap is to be closed (e.g. see [25, 26, 41]). In the following sections we will discuss some of the strategies that have achieved precisely that.

2.1.2 Convex reformulation of the generalized StQP

One of the earliest results in the copositive optimization literature was the exact copositive reformulation of the StQP (see [11]). It was shown that for

$$\mathcal{F} = \left\{\mathbf{x} \in \mathbb{R}^n_+ : \mathbf{e}^\mathsf{T}\mathbf{x} = 1\right\} \text{ we have } \mathcal{G}(\mathcal{F}) = \left\{\mathsf{X} \in \mathcal{CPP}\left(\mathbb{R}^n_+\right) : \mathsf{E} \bullet \mathsf{X} = 1\right\}.$$

We will slightly generalize this result and consider $\mathcal{F} = \left\{\mathbf{x} \in \mathcal{K} \colon \mathbf{a}^\mathsf{T}\mathbf{x} = 1\right\}$ where $\mathcal{K} \subseteq \mathbb{R}^n$ is a cone whose convex hull is pointed and $\mathbf{a} \in \operatorname{int}(\mathcal{K}^*)$. This will make for a nice illustration of the techniques involved in such a result and contextualize the discussion in later chapters (more specifically Chapter 4 and Chapter 6).

Under the aforementioned requirements on $\mathcal{K}$ and $\mathbf{a}$ we define the generalized StQP as

$$\min_{\mathbf{x}\in\mathcal{K}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} \colon \mathbf{a}^\mathsf{T}\mathbf{x} = 1\right\} \tag{GStQP}$$

and we denote its feasible set as $\mathcal{F}_{GStQP} \coloneqq \left\{\mathbf{x} \in \mathcal{K} \colon \mathbf{a}^\mathsf{T}\mathbf{x} = 1\right\}$.

Theorem 2. *$\mathcal{G}(\mathcal{F}_{GStQP}) = \left\{\mathsf{X} \in \mathcal{CPP}(\mathcal{K}) \colon \mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{X} = 1\right\}$ and we have*

$$\min_{\mathbf{x}\in\mathcal{K}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} \colon \mathbf{a}^\mathsf{T}\mathbf{x} = 1\right\} = \min_{\mathsf{X}\in\mathcal{CPP}(\mathcal{K})} \left\{\mathsf{Q} \bullet \mathsf{X} \colon \mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{X} = 1\right\}.$$

Proof. We will denote the feasible set of the linear conic optimization problem as $\mathcal{F}_{Shor}$. Observe that $\mathbf{a} \in \operatorname{int}(\mathcal{K}^*)$ implies that $\mathbf{v}^\mathsf{T}\mathbf{a}\mathbf{a}^\mathsf{T}\mathbf{v} = (\mathbf{v}^\mathsf{T}\mathbf{a}) > 0 \ \forall \mathbf{v} \in \mathcal{K}$ so that $\mathbf{a}\mathbf{a}^\mathsf{T} \in \operatorname{int}\mathcal{COP}(\mathcal{K})$ by Proposition 53 (see Appendix A). Also $\mathcal{COP}(\mathcal{K}) = \mathcal{CPP}(\mathcal{K})^*$ by Proposition 52 so that $\mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{V} > 0 \ \forall \mathsf{V} \in \mathcal{CPP}(\mathcal{K})$. We argue that therefore the extreme points of this set are precisely the intersections of the extreme rays of $\mathcal{CPP}(\mathcal{K})$ and the hyperplane $\mathcal{H} \coloneqq \left\{\mathsf{X} \in \mathcal{S}^n \colon \mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{X} = 1\right\}$, so that these extreme points are actually rank-1 matrices. Indeed if $\mathsf{Y} \in \operatorname{ext}\mathcal{F}_{shor}$ but $\{\lambda\mathsf{Y} :- \lambda \geq 0\}$ was not an extreme ray of $\mathcal{CPP}(\mathcal{K})$ then by [73, Corollary 18.5.2] $\mathsf{Y} = \sum_{i=1}^k \lambda_i\mathsf{Y}_i, \ \sum_{i=1}^k \lambda_i = 1, \ \lambda_i \geq 0, \ i \in [1\colon k]$ for some generators of extreme rays $\mathsf{Y}_i \in \operatorname{ext}\mathcal{CPP}(\mathcal{K}), \ i \in [1\colon k]$. All these extreme rays have to meet the hyperplane $\mathcal{H}$ since $\mathbf{a}\mathbf{a}^\mathsf{T} \in \operatorname{int}\mathcal{COP}(\mathcal{K})$ so that there are $\mu_i > 0$ with $\mathbf{a}\mathbf{a}^\mathsf{T} \bullet (\mu_i\mathsf{Y}_i) = 1$ so acutally $\mu_i = 1/\left(\mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{Y}_i\right)$. But this implies $\mathsf{Y} = \sum_{i=1}^k \lambda_i\mathsf{Y}_i = \sum_{i=1}^k \frac{\lambda_i}{\mu_i}\mu_i\mathsf{Y}_i$ where $\mu_i\mathsf{Y}_i \in \mathcal{F}_{Shor}$ and $\sum_{i=1}^k \frac{\lambda_i}{\mu_i} = \sum_{i=1}^k \lambda_i\left(\mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{Y}_i\right) = \mathbf{a}\mathbf{a}^\mathsf{T} \bullet \sum_{i=1}^k \lambda_i\mathsf{Y}_i = \mathbf{a}\mathbf{a}^\mathsf{T} \bullet \mathsf{Y} = 1$ so that Y cannot be an extreme point of $\mathcal{F}_{Shor}$. Conversely, all extreme rays of $\mathcal{CPP}(\mathcal{K})$ eventually meet $\mathcal{H}$ at a single point which is of the form $\mathbf{y}\mathbf{y}^\mathsf{T}$ for some $\mathbf{y} \in \mathcal{K}$ by Lemma 54 and by the same lemma have to be extreme in $\mathcal{F}_{Shor}$ since rank-1 matrices cannot be obtained as a convex combination of other rank-1 matrices. This shows that $\mathcal{F}_{Shor} = \mathcal{G}(\mathcal{F})$ and the exactness of the relaxation follows from Theorem 1 ☐

2.1.3 Kim et. al's geometric approach to Burer's convex Reformulation

One of the most celebrated examples of an application of Theorem 1 is Burer's completely positive reformulation of a quite large class of QCQPs:

Theorem 3. *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed, convex cone and assume, that $\mathcal{L} := \{\mathbf{x} \in \mathcal{K} \colon \mathsf{A}\mathbf{x} = \mathbf{b}\}$ is bounded and that*

$$\mathbf{x}^\mathsf{T}\mathsf{Q}_i\mathbf{x} + \mathbf{q}_i^\mathsf{T} + \omega_i \geq 0 \text{ for all } \mathbf{x} \in \mathcal{L} \text{ and } i \in [1\colon l],$$

which is referred to as the "key assumption". Then, any feasible QCQP of the form

$$\min_{x \in \mathcal{K}} \left\{ \mathbf{x}\mathsf{Q}\mathbf{x}^\mathsf{T} + \mathbf{q}^\mathsf{T}\mathbf{x} \colon \mathsf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x}^\mathsf{T}\mathsf{Q}_i\mathbf{x} + \mathbf{q}_i^\mathsf{T} + \omega_i = 0,\ i \in [1\colon l] \right\}$$

is equivalent to

$$\begin{aligned}
\min_{\mathbf{x},\mathsf{X}} \ & \mathsf{Q} \bullet \mathsf{X} + \mathbf{q}^\mathsf{T}\mathbf{x} \\
\text{s.t. :} \ & \mathsf{A}\mathbf{x} = \mathbf{b}, \\
\operatorname{Diag}\left(\mathsf{A}\mathsf{X}\mathsf{A}^\mathsf{T}\right) & = \mathbf{b} \circ \mathbf{b}, \\
\mathsf{Q}_i \bullet \mathsf{X} + \mathbf{q}_i^\mathsf{T} + \omega_i & = 0,\ i \in [1\colon l], \\
\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathsf{X} \end{pmatrix} & \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}).
\end{aligned}$$

Proof. See [26] for the original proof for $\mathcal{K} = \mathbb{R}^n_+$, see [41] for the proof under the assumption that $\mathcal{K}$ is a norm-cone. Considering the results in [54], which we will discuss shortly, no such assumption is needed for the theorem to hold. $\square$

The original proof of the theorem is quite algebraic and seems somewhat removed from the simple, geometric motivation of Theorem 1. Thankfully, Kim et. al recently proposed a geometrical perspective on the proof strategy (see [54]). The concepts they introduce are quite versatile and allow proofs for generalizations of Theorem 3 as well exactness proofs for relaxations of polynomial optimization problems. In Chapter 4 we will use their ideas in order to investigate a new type of convex reformulation. In order to make the results there more transparent we will give a condensed review of the main results in [54] here. The theorems presented in the remainder of this section are therefore simplified (and thus less powerful) versions of results in [54], as are the proofs. This will allow us to keep the discussion concise and on an introductory level while engaging with the subject deep enough so we will still be able to fashion the alternative proof of Theorem 3 and understand the discussion in Chapter 4. What follows is, thus, a didactic presentation of results of which we do not claim authorship!

We start out be investigating a more general question. Let $\mathbb{V}$ be a vector space of dimension n. For a (possibly nonconvex) cone $\mathbb{K} \subseteq \mathbb{V}$, and vectors $\mathsf{Q}, \mathsf{H}_0 \in \mathbb{V}$ and a convex set $\mathbb{J} \subseteq \operatorname{conv}(\mathbb{K})$, we want to know when we have the equality:

$$\min_{\mathsf{X} \in \mathbb{V}} \{\langle \mathsf{Q}, \mathsf{X} \rangle \colon \mathsf{X} \in \mathbb{K} \cap \mathbb{J},\ \langle \mathsf{H}_0, \mathsf{X} \rangle = 1\} \overset{?}{=} \min_{\mathsf{X} \in \mathbb{V}} \{\langle \mathsf{Q}, \mathsf{X} \rangle \colon \mathsf{X} \in \mathbb{J},\ \langle \mathsf{H}_0, \mathsf{X} \rangle = 1\}.$$

Defining $\mathbb{H} \coloneqq \{\mathsf{X}\colon \langle \mathsf{H}_0, \mathsf{X}\rangle = 1\}$, we can equivalently ask for conditions for the equality

$$\operatorname{conv}(\mathbb{H} \cap \mathbb{K} \cap \mathbb{J}) = \mathbb{H} \cap \mathbb{J}.$$

The following theorem gives an answer based on convex geometry.

Theorem 4. *For $\mathbb{H}, \mathbb{K}, \mathbb{J}$ as above, assume that $\mathbb{H} \cap \mathbb{J} \neq \emptyset$ is bounded and that $\mathbb{J}$ is a face of $\operatorname{conv}(\mathbb{K})$. Then $\operatorname{conv}(\mathbb{H} \cap \mathbb{K} \cap \mathbb{J}) = \mathbb{H} \cap \mathbb{J}$.*

Proof. For the "$\subseteq$"-inclusion, since $\mathbb{H} \cap \mathbb{K} \cap \mathbb{J} \subseteq \mathbb{H} \cap \mathbb{J}$ and the latter set is convex, there is nothing left to show. For the converse, let $\mathsf{X} \in \mathbb{H} \cap \mathbb{J}$. Then $\mathsf{X} \in \operatorname{conv}(\mathbb{K})$ since $\mathbb{J} \subseteq \operatorname{conv}(\mathbb{K})$, so that $\mathsf{X} = \sum_{i=1}^{n} \mathsf{X}_i$ with $\mathsf{X}_i \in \mathbb{K}$ but also $\mathsf{X}_i \in \mathbb{J}$ since $\mathbb{J}$ is a face of $\operatorname{conv}(\mathbb{K})$ so that $\mathsf{X}_i \in \mathbb{K} \cap \mathbb{J}$. Now, $\langle \mathsf{H}_0, \mathsf{X}_i\rangle > 0$ since $\mathbb{H} \cap \mathbb{J} \neq \emptyset$ is bounded. Define $\lambda_i = \langle \mathsf{H}_0, \mathsf{X}_i\rangle$. We have $\langle \mathsf{H}_0, \mathsf{X}\rangle = \sum_i \langle \mathsf{H}_0, \mathsf{X}_i\rangle = \sum_i \lambda_i = 1$ and $\lambda_i^{-1}\mathsf{X}_i \eqqcolon \bar{\mathsf{X}}_i \in \mathbb{K} \cap \mathbb{J}$ and thus $\mathsf{X} = \sum_i \lambda_i \bar{\mathsf{X}}_i \in \operatorname{conv}(\mathbb{H} \cap \mathbb{K} \cap \mathbb{J})$. $\square$

This theorem motivates the search for a condition that lets us identify faces of convex cones, which are provided in the following theorem.

Theorem 5. *Assume that $\mathbb{J} = \{\mathsf{X} \in \operatorname{conv}(\mathbb{K})\colon \langle \mathsf{A}_i, \mathsf{X}\rangle = 0,\ i \in [1\!:\!m]\}$ and define $\mathbb{J}_p \coloneqq \{\mathsf{X} \in \operatorname{conv}(\mathbb{K})\colon \langle \mathsf{A}_i, \mathsf{X}\rangle = 0,\ i \in [1\!:\!p]\}$ so that $\mathbb{J}_m = \mathbb{J}$ and $\mathbb{J}_0 = \operatorname{conv}(\mathbb{K})$. If $\mathsf{A}_p \in \mathbb{J}_{p-1}^*, i \in [1\!:\!m]$ then $\mathbb{J}$ is a face of $\operatorname{conv}(\mathbb{K})$.*

Proof. Since a face of a face a convex set is itself a face of that set, the claim will follow by induction if we can show that

$$\mathsf{A}_p \in \mathbb{J}_{p-1}^* \implies \mathbb{J}_p \text{ is a face of } \mathbb{J}_{p-1}.$$

So let $\mathbb{J}_p \ni \mathsf{X} = \mathsf{X}_1 + \mathsf{X}_2$ with $\mathsf{X}_i \in \mathbb{J}_{p-1},\ i \in [1:S]$. We have $\langle \mathsf{A}_p, \mathsf{X}_i\rangle \geq 0$ since $\mathsf{A}_p \in \mathbb{J}_{p-1}^*$ so that $0 = \langle \mathsf{A}_p, \mathsf{X}\rangle = \langle \mathsf{A}_p, \mathsf{X}_1\rangle + \langle \mathsf{A}_p, \mathsf{X}_2\rangle$ implies that actually $\langle \mathsf{A}_p, \mathsf{X}_i\rangle = 0$ and we indeed have $\mathsf{X}_i \in \mathbb{J}_p,\ i \in \{1, 2\}$. $\square$

To see how this is relevant for the reformulation of QCQPs consider the following simple reformulation steps:

$$\begin{aligned}
&\min_{\mathbf{x}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}_0\mathbf{x} + \mathbf{q}_0^\mathsf{T}\mathbf{x} + \omega\colon \mathsf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in \mathcal{K},\ \mathbf{x}^\mathsf{T}\mathsf{Q}_i\mathbf{x} + \mathbf{q}_i^\mathsf{T}\mathbf{x} + \omega = 0,\ i \in [1\!:\!m]\right\}\\
&= \min_{\mathbf{y}} \left\{\bar{\mathsf{Q}}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T}\colon (-\mathbf{b}, \mathsf{A})\mathbf{y} = \mathbf{o},\ y_0 = 1,\ \mathbf{y} \in \mathbb{R}_+ \times \mathcal{K},\ \mathsf{A}_i \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 0,\ i \in [1\!:\!m]\right\}\\
&= \min_{\mathbf{y}} \left\{\bar{\mathsf{Q}}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T}\colon \mathbf{y}^\mathsf{T}(-\mathbf{b}, \mathsf{A})^\mathsf{T}(-\mathbf{b}, \mathsf{A})\mathbf{y} = 0,\ y_0^2 = 1,\ \mathbf{y} \in \mathbb{R}_+ \times \mathcal{K},\ \mathsf{A}_i \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 0,\ i \in [1\!:\!m]\right\}\\
&= \min_{\mathbf{y}} \left\{\bar{\mathsf{Q}}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T}\colon \mathsf{H}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 1,\ \mathsf{A}_i \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 0,\ i \in [0\!:\!m],\ \mathbf{y} \in \mathbb{R}_+ \times \mathcal{K}\right\}\\
&= \min_{\mathbf{y}} \left\{\bar{\mathsf{Q}}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T}\colon \mathsf{H}_0 \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 1,\ \mathsf{A}_i \bullet \mathbf{y}\mathbf{y}^\mathsf{T} = 0,\ i \in [0\!:\!m],\ \mathbf{y}\mathbf{y}^\mathsf{T} \in \operatorname{ext}\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K})\right\}\\
&= \min_{\mathsf{Y}} \left\{\bar{\mathsf{Q}}_0 \bullet \mathsf{Y}\colon \mathsf{H}_0 \bullet \mathsf{Y} = 1,\ \mathsf{A}_i \bullet \mathsf{Y} = 0,\ i \in [0\!:\!m],\ \mathsf{Y} \in \operatorname{ext}\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K})\right\}
\end{aligned}$$

where $\bar{\mathsf{Q}}_0 := \mathsf{M}(\omega_0, \mathbf{q}_0, \mathsf{Q}_0)$, $\mathsf{A}_i := \mathsf{M}(\omega_i, \mathbf{q}_i, \mathsf{Q}_i)$, $i \in [1:m]$, $\mathsf{A}_0 := (-\mathbf{b}, \mathsf{A})^\mathsf{T}(-\mathbf{b}, \mathsf{A})$ and $\mathsf{H}_0 := \mathsf{M}(1, \mathbf{o}, \mathsf{O})$. The final product is of the desired form with $\mathbb{K} = \mathrm{ext}\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K})$, $\mathbb{H} = \{\mathsf{Y} \in \mathcal{S}^{n_1} : \mathsf{H}_0 \bullet \mathsf{Y} = 1\}$ and by Theorem 4 we can show the equivalence

$$\begin{aligned} &\min_{\mathsf{Y}} \left\{ \bar{\mathsf{Q}}_0 \bullet \mathsf{Y} \colon \mathsf{H}_0 \bullet \mathsf{Y} = 1,\ \mathsf{A}_i \bullet \mathsf{Y} = 0,\ i \in [0:m],\ \mathsf{Y} \in \mathrm{ext}\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) \right\} \\ = &\min_{\mathsf{Y}} \left\{ \bar{\mathsf{Q}}_0 \bullet \mathsf{Y} \colon \mathsf{H}_0 \bullet \mathsf{Y} = 1,\ \mathsf{A}_i \bullet \mathsf{Y} = 0,\ i \in [0:m],\ \mathsf{Y} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) \right\} \end{aligned}$$

by proving that $\mathbb{J} = \{\mathsf{Y} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) \colon \mathsf{A}_i \bullet \mathsf{Y} = 0,\ i \in [0:m]\}$ is a face of $\mathrm{conv}\,\mathbb{K} = \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K})$. and that $\mathbb{J} \cap \mathbb{H} \neq \emptyset$ is bounded. This amounts to a proof of Theorem 3, but we will merely give a sketch of the argument to close this section. The boundedness and nonemptyness of $\mathbb{J} \cap \mathbb{H} \neq \emptyset$ follows from the fact that we only consider feasible QCQPs and from the boundedness of $\mathcal{L} := \left\{ \mathbf{x} \in \mathbb{R}^n_+ \colon \mathsf{A}\mathbf{x} = \mathbf{b} \right\}$. The latter can be infered by considering element the recession cone of $\{\mathsf{Y} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) \colon \mathsf{A}_0 \bullet \mathsf{Y} = 0,\ \mathsf{H}_0 \bullet \mathsf{Y} = 1\}$ which can be decomposed into nonzero elements of the recession cone of $\mathcal{L}$ which only contains the origin by assumption. Further $\{\mathsf{Y} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) \colon \mathsf{A}_0 \bullet \mathsf{Y} = 0\}$ is easily shown to be a face of $\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K})$ by using the positive semidefiniteness of A_0. The fact we obtain faces of this face when adding the additional constraints is guaranteed by the key assumption and Theorem 5. The reformulation that we obtain in this manner is slightly different from the one depicted in Theorem 3 but is an elementary exercise to show that these two are equivalent.

Remark. *Readers who are familiar with the original proof of Theorem 3 in [25] may recall the elegant use of the equality case of the Cauchy-Schwartz inequality and wonder why the above sketch of the alternative proof, seemingly without effort, bypasses this aspect of the proof entirely. The reason for this is that we presupposed the characterization of the extreme rays of the set-completely positive cone which we derive in Lemma 54 in Appendix A.2. The proof of this lemma indeed rests on the equality case of the Cauchy-Schwartz inequality so that in this manner it enters the alternative proof strategy as well.*

2.2 QCQP-duality and convex reformulations

2.2.1 The S-Procedure

Another strain of literature that is concerned with a dual perspective on QCQPs is the theory of the *S-procedure*. The central question is when for a given set of matrices

$\{\mathsf{Q}, \mathsf{A}_1, \dots, \mathsf{A}_m\}$ the following two statements are equivalent

i) the following system of inequalities has no solution $\mathbf{x} \in \mathcal{K}$:
$\mathbf{x}^{\mathsf{T}}\mathsf{Q}\mathbf{x} < 0 \quad \text{and} \quad \mathbf{x}^{\mathsf{T}}\mathsf{A}_i\mathbf{x} \leq 0 \; \forall i \in [1 \colon m]$;

ii) the following conic inequality has a solution $\boldsymbol{\lambda} \in \mathbb{R}^m_+$:
$$\mathsf{Q} + \sum_{i=1}^{m} \lambda_i \mathsf{A}_i \in \mathcal{COP}\left(\mathcal{K}\right) .$$

If equivalence between $i)$ and $ii)$ can be established, we say that the S-procedure is exact for this set of matrices. Actually, if $\mathcal{K} = \mathbb{R}^n$ and the equivalence holds for arbitrary Q, then copositivity over $\bar{\mathcal{K}} \coloneqq \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x}^{\mathsf{T}}\mathsf{A}_i\mathbf{x} \leq 0, \; i \in [1 : m]\}$ is characterized by $\mathcal{COP}(\mathcal{K}) = \{\mathsf{Q} \colon \mathsf{Q} + \sum_{i=1}^m \lambda_i \mathsf{A}_i \in \mathcal{S}^n_+ \text{ for some } \boldsymbol{\lambda} \in \mathbb{R}^m_+\}$. Note that '$\supseteq$' always holds since obviously $ii)$ always implies $i)$. The first exactness result, known as the S-Lemma, has been established in [85] for the case $m = 1$ under the condition that $\mathbf{x}_0^{\mathsf{T}}\mathsf{A}_1\mathbf{x}_0 < 0$ holds for some $\mathbf{x}_0 \in \mathbb{R}^n$. The result is obtained by invoking a convexity result derived in [38] concerning the so called *joint numerical range* of quadratic functions defined as $\mathcal{J}(\mathsf{Q}, \mathsf{A}_1, \dots, \mathsf{A}_m, \mathcal{K}) \coloneqq \left\{(\mathbf{x}^{\mathsf{T}}\mathsf{Q}\mathbf{x}, \mathbf{x}^{\mathsf{T}}\mathsf{A}_1\mathbf{x}, \dots, \mathbf{x}^{\mathsf{T}}\mathsf{A}_m\mathbf{x}) \colon \mathbf{x} \in \mathcal{K}\right\}$. It was shown that $\mathcal{J}(\mathsf{Q}, \mathsf{A}_1, \mathbb{R}^n)$ is a convex cone. From this convexity result it is deduced that whenever $\mathcal{J}$ is disjoint from the negative quadrant, i.e., $i)$ holds, these two sets can be separated by a hyperplane. The coefficients of that hyperplane can further be shown to be two nonnegative numbers, where the first is actually positive, by using the assumption on $\mathbf{x}_0$, which eventually implies $ii)$.

Later in the text we will discuss other results on the S-procedure. Also, we will give sufficient conditions for it to be exact for arbitrary m, and discuss the connection of this result to the ones given in [71]. For further references, the interested reader may be referred to [31, 50, 70] and references therein.

2.2.2 Semi-lagrangian duality

Another way of establishing exactness of the Shor relaxation stems from the fact, that the conic dual of the Shor relaxation and the Semi-Lagrangian dual of the underlying QCQP (see [42]) take the same form. To see this consider the following derivation of the Semi-Lagrangian dual of a QCQP:

$$
\begin{aligned}
&\inf_{\mathbf{x}\in\mathcal{K}}\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+\mathbf{q}^\mathsf{T}\mathbf{x}+\omega\colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+\mathbf{a}_i^\mathsf{T}\mathbf{x}\le b_i,\ i\in[1\!:\!m]\}\\
=&\inf_{(x_0,\mathbf{x})\in\mathbb{R}_+\times\mathcal{K}}\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+x_0\mathbf{q}^\mathsf{T}\mathbf{x}+\omega x_0^2\colon \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+x_0\mathbf{a}_i^\mathsf{T}\mathbf{x}\le b_i,\ i\in[1\!:\!m],\ x_0^2=1\}\\
\ge&\sup_{\boldsymbol{\lambda}\in\mathbb{R}_+^{m+1}}\inf_{(x_0,\mathbf{x})\in\mathbb{R}_+\times\mathcal{K}}\left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+x_0\mathbf{q}^\mathsf{T}\mathbf{x}+x_0^2\omega+\lambda_0(x_0^2-1)+\sum_{i=1}^m\lambda_i(\mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+x_0\mathbf{a}_i^\mathsf{T}\mathbf{x}-b_i)\right\}\\
=&\sup_{\boldsymbol{\lambda}\in\mathbb{R}_+^{m+1}}\left\{-\boldsymbol{\lambda}^\mathsf{T}\mathbf{b}-\lambda_0+\inf_{(\mathbf{x},x_0)\in\mathbb{R}_+\times\mathcal{K}}\left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+x_0\mathbf{q}^\mathsf{T}\mathbf{x}+x_0^2\omega+\lambda_0x_0^2+\sum_{i=1}^m\lambda_i(\mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+x_0\mathbf{a}_i^\mathsf{T}\mathbf{x})\right\}\right\}\\
=&\sup_{\boldsymbol{\lambda}\in\mathbb{R}_+^{m+1}}\left\{-\boldsymbol{\lambda}^\mathsf{T}\mathbf{b}-\lambda_0:\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}+x_0\mathbf{q}^\mathsf{T}\mathbf{x}+x_0^2\omega+\lambda_0x_0^2+\sum_{i=1}^m\lambda_i(\mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x}+x_0\mathbf{a}_i^\mathsf{T}\mathbf{x})\ge 0\quad\forall(x_0,\mathbf{x})\in\mathbb{R}_+\times\mathcal{K}\right\}\\
=&\sup_{\boldsymbol{\lambda}\in\mathbb{R}_+^{m+1}}\left\{-\boldsymbol{\lambda}^\mathsf{T}\mathbf{b}-\lambda_0:\mathsf{M}(\omega,\mathbf{q},\mathsf{Q})+\lambda_0\mathsf{M}(1,\mathbf{o},\mathsf{O})+\sum_{i=1}^m\lambda_i\mathsf{M}(0,\mathbf{a}_i,\mathsf{A}_i)\in\mathcal{COP}(\mathbb{R}_+\times\mathcal{K})\right\}.
\end{aligned}
$$

In the second to last equality we used the fact that for any quadratic form $q(\mathbf{x})$ and any cone $\bar{\mathcal{K}}$ we have

$$
\inf_{\mathbf{x}\in\bar{\mathcal{K}}}q(\mathbf{x})=\begin{cases}0 & \text{if } q(\mathbf{x})\ge 0\ \forall\ \mathbf{x}\in\bar{\mathcal{K}}\\ -\infty & \text{otherwise.}\end{cases}
$$

Here, we speak of a Semi-Lagrangian dual, since the constraints $(\mathbf{x},x_0)\in\mathcal{K}\times\mathbb{R}_+$ is not dualized in that we don't introduce respective dual variables. The final equality follows from the definition of $\mathcal{COP}(\cdot)$ and in fact yields the conic dual of the Shor relaxation after a small reformulation,

$$
\begin{aligned}
&\inf_{\mathsf{X},\mathbf{x}}\{\mathsf{Q}\bullet\mathsf{X}+\mathbf{q}^\mathsf{T}\mathbf{x}+\omega\colon \mathsf{A}_i\bullet\mathsf{X}+\mathbf{a}_i^\mathsf{T}\mathbf{x}\le b_i,\ i\in[1\!:\!m],\mathsf{Y}(\mathsf{X},\mathbf{x})\in\mathcal{CPP}(\mathcal{K}\times\mathbb{R}_+)\}\\
=&\inf_{\mathsf{Y}\in\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K})}\{\mathsf{M}(\omega,\mathbf{q},\mathsf{Q})\bullet\mathsf{Y}\colon \mathsf{M}(0,\mathbf{a}_i,\mathsf{A}_i)\bullet\mathsf{Y}\le b_i,\ i\in[1\!:\!m],\ \mathsf{M}(1,\mathbf{o},\mathsf{O})\bullet\mathsf{Y}=1\}\,.
\end{aligned}
$$

Duality results for QCQPs can be obtained from S-lemma type results in the same way the Farkas Lemma (see [73, Corollary 22.3.1]) can be used to prove strong duality for linear optimization problems. For examples, see [31, 50] and references therein as well as Example 3 which will be discussed shortly. If for a fixed feasible set $\mathcal{F}$ the strong Semi-Lagrangian duality holds for an arbitrary choice of the objective function coefficients, the feasible set of the Shor relaxation is in fact $\mathcal{G}(\mathcal{F})$.

We will now proceed to give some examples that illustrate the interplay between relaxation gaps, duality gaps and exact strengthenings of QCQPs and their relaxations.

Example 3. *(No duality gap for the QCQP, hence no relaxation gap for the Shor relaxation) Consider the optimization problem* $\min_{x\in\mathbb{R}_+}\{-x^2\colon x^2\le 1\}$. *We will now* *give an argument for full strong duality based on the joint numerical range. In this simple special case the difficulties of such an approach are avoided while the core idea of the argument is maintained. The following equivalences are easily checked to be true:*

$$\begin{aligned}&\min_{x\in\mathbb{R}_+}\left\{-x^2\colon x^2\leq 1\right\}\\ \Leftrightarrow &\sup_{t\in\mathbb{R}}\left\{t\colon\; -x^2\geq t\;\forall x\in\mathbb{R}_+\colon x^2\leq 1\right\}\\ \Leftrightarrow &\sup_{t\in\mathbb{R}}\left\{t\colon\;\left\{x\in\mathbb{R}_+\colon\; -x^2<t,\; x^2\leq 1\right\}=\emptyset\right\}.\end{aligned}$$

In order to close the duality gap we have to show that the condition

$$\left\{x\in\mathbb{R}_+\colon\; -x^2<t,\; x^2\leq 1\right\}=\emptyset \tag{2.1}$$

is equivalent to

$$\exists\lambda\geq 0\colon\;(-x^2-t)+\lambda(x^2-1)\geq 0\;\forall x\in\mathbb{R}_+. \tag{2.2}$$

Clearly, (2.2) implies (2.1), since in case (2.1) fails any member of the respective set is an $x\in\mathbb{R}_+$ that invalidates (2.2). The reverse implication can be demonstrated by appealing to the geometry of the joint numerical range $\mathcal{J}(-1,1,\mathbb{R}_+)=\left\{(-x^2,x^2)^\mathsf{T}\colon x\in\mathbb{R}_+\right\}$. In fact, in this simple case we have

$$\mathcal{J}(-1,1,\mathbb{R}_+)=\left\{\delta\begin{pmatrix}-1\\1\end{pmatrix}\colon\delta\geq 0\right\},$$

which is a half line and thus convex (which also illustrates the classical result in [38]). Now assume that (2.1) holds, then $\mathcal{J}(-1,1,\mathbb{R}_+)-t\mathbf{e}_1-\mathbf{e}_2$ does not meet $\operatorname{int}(\mathbb{R}^2_-)$. Since both sets are convex, they can be separated by a hyperplane, i.e. it holds that

$$\exists(\alpha,\beta)^\mathsf{T}\in\mathbb{R}^2_+\setminus\{\mathbf{o}\}\colon\;\alpha(-x^2-t)+\beta(x^2-1)\geq 0\;\forall x\in\operatorname{int}(\mathbb{R}_+)$$

where the nonnegativity of the multipliers follows from the fact that we separate from the interior of the negative orthant and the $\operatorname{int}(\cdot)$-operator can be dropped by continuity. For the latter condition to hold, it must be the case that $\alpha>0$, otherwise we had $\beta>0$ and $x=0$ would take the remaining term to be $-\beta<0$. We thus arrive at (2.2) with $\lambda=\beta/\alpha$. Finally, we have that for any fixed $t\in\mathbb{R}$ (2.2) is equivalent to

$$\begin{aligned}&\exists\lambda\geq 0\colon\;(-x^2-ty^2)+\lambda(x^2-y^2)\geq 0\;\forall(x,y)\in\mathbb{R}_+\times\{1\}\\ \Leftrightarrow&\exists\lambda\geq 0\colon\;(-x^2-ty^2)+\lambda(x^2-y^2)\geq 0\;\forall(x,y)\in\mathbb{R}^2_+\end{aligned}$$

$$\Leftrightarrow \exists \lambda \geq 0\colon\ \mathsf{M}(-1,0,-t)+\lambda \mathsf{M}(1,0,-1) \in \mathcal{COP}(\mathbb{R}^2_+),$$

so that

$$\min_{x\in\mathbb{R}_+} \left\{-x^2\colon x^2 \leq 1\right\} = \sup_{t,\lambda} \left\{t\colon \mathsf{M}(-1,0,-t)+\lambda \mathsf{M}(1,0,-1) \in \mathcal{COP}(\mathbb{R}^2_+),\ \lambda \geq 0\right\}.$$

The latter problem is equivalent to $\sup_\lambda \{-\lambda\colon 1 \leq \lambda,\ \lambda \in \mathbb{R}_+\}$ *which is the dual of the Shor relaxation* $\min_{x\in\mathcal{S}^1_+} \{-x\colon x \leq 1\}$*. Thus, it follows that the Shor relaxation is tight.*

Example 4. *(Positive QCQP-duality gap, inexact Shor relaxation with exact tightening, both with zero duality gap) Consider the dual of* $\min\{x^2 - y^2 + x :\ x + y = 1,\ x, y \geq 0\}$ *given by*

$$\begin{aligned} &\sup\ \lambda_1 + \lambda_2 \\ &\text{s.t.} : \begin{pmatrix} 1 & 0 & \frac{1-\lambda_1}{2} \\ 0 & -1 & \frac{\lambda_1}{2} \\ \frac{1-\lambda_1}{2} & \frac{\lambda_1}{2} & \lambda_2 \end{pmatrix} \in \mathcal{COP}(\mathbb{R}^3_+) = \mathcal{S}^3_+ + \mathcal{N}_3. \end{aligned}$$

The primal QCQP has an optimal value of -1 *while the lagrangian dual is infeasible since any matrix in* $\mathcal{COP}(\mathbb{R}^3_+)$ *has nonnegative diagonal entries. The Shor relaxation given by*

$$\begin{aligned} &\min\ X - Y + x \\ &\quad \text{s.t.} : x + y = 1, \\ &\begin{pmatrix} X & Z & x \\ Z & Y & y \\ x & y & 1 \end{pmatrix} \in \mathcal{CPP}(\mathbb{R}^3_+) = \mathcal{S}^3_+ \cap \mathcal{N}_3. \end{aligned}$$

is unbounded, where for the improving ray we have $x = X = Z = 0$*,* $y = 1$ *and* $Y \geq 1$*. Thus the duality gap for the Shor relaxation is actually zero since the QCQP and its Shor relaxation share the same langrangian dual. We can strengthen the Shor relaxation by introducing the additional constraint* $X + 2Z + Y = 1$*, which is the relaxation of the redundant constraint* $(x+y)^2 = 1$*. This yields an exact reformulation by Theorem 3. After applying the simplification steps outlined there, we obtain another exact reformulation, namely* $\min\left\{2X - Y\colon X + 2Z + Y = 1,\ \mathsf{M}(X, 2Z, Y) \in \mathcal{CPP}(\mathbb{R}^2_+)\right\}$ *which has a Slater point at* $X = Y = 0.4,\ Z = 0.1$*, so that it exhibits no duality gap.*

Example 5. *(Exact Shor reformulation with finite duality gap) Let us consider* $\min\{x_1^2 + x_2^2 \colon x_1^2 = 0,\ x_1x_3 - x_2^2 = -1,\ x_2 \geq 0\}$, *whose optimal value of* 1 *is attained at* $x_1 = 0$, $x_2 = 1$. *The Shor relaxation* $\min\left\{\mathsf{X}_{11} + \mathsf{X}_{22} \colon \mathsf{X}_{11} = 0,\ \mathsf{X}_{13} - \mathsf{X}_{22} = -1,\ \mathsf{X} \in \mathcal{CPP}(\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}) = \mathcal{S}_+^3\right\}$ *has the same optimal value, since* $\mathsf{X}_{11} = 0$ *together with* $\mathsf{X} \in \mathcal{S}_+^3$ *forces* $\mathsf{X}_{1i} = 0,\ i = 2, 3$. *Now the Shor relaxation is in fact a known example from [66] for an SDP with finite positive duality gap.*

In summary we have discussed two routes for achieving a convex reformulation of a QCQP, either via closing a relaxation gap or via closing a duality gap, and we have illustrated the following facts:

- If a QCQP enjoys strong duality (e.g. if the joint numerical range is a convex cone), we also have an exact Shor reformulation.
- If the Shor relaxation is exact and we can close its duality gap we also close the duality gap between the underlying QCQP and its dual.
- If the Shor relaxation is not exact but we can find an exact tightening, the dual of that reformulation can be used as an alternative to the Lagrangian dual of the QCQP.

For some background on these facts see for example [19, 31] who discuss the case $\mathcal{K} = \mathbb{R}_+^n$. To the best of our knowledge these facts have so far not been discussed for general convex $\mathcal{K}$, but the generalization is immediate, as shown above.

Chapter 3

Exactness of the Shor relaxation and the generalized S-Lemma

Typically, in literature, one of the paths, closing the relaxation gap or closing the duality gap, is chosen with no regard for the implications of an obtained result for the constitution of the other route. We will now proceed with a demonstration of the former path which is a nice example in two regards. First of all, the derivation is simple and instructive as it is accessible via some geometric intuition. Second, the conditions under which we close the relaxation gap will not only give rise to new conditions for a generalized version of the S-Lemma to hold, but also allows some insight into the joint numerical range of the quadratic forms which fulfil these conditions.

3.1 Harnessing Pataki's rank result

In this section we provide the proof of the exactness of the Shor relaxation under a geometric condition by iterating an application of Pataki's rank theorem. We follow a proof strategy originally considered in [27], where the claim of Theorem 7 was proved for the special case of $m = 2$.

Theorem 6. *Consider a feasible set of a semidefinite optimization problem in block-standard form*

$$\mathcal{T} \coloneqq \{[\mathsf{X}_j]_{j\in[1:p]} \in \bigtimes_{j=1}^{p} \mathcal{S}_+^{n_j} : \sum_{j=1}^{p} \mathsf{A}_{ji} \bullet \mathsf{X}_j = b_i, i \in [1\!:\!k]\}.$$

Let $[\mathsf{X}_1, \dots, \mathsf{X}_p]$ *be an extreme point of* $\mathcal{T}$ *and let* $r_j \coloneqq \operatorname{rank}(\mathsf{X}_j)$*. Then it holds that* $\sum_{j=1}^{p} r_j(r_j + 1) \leq 2k$*.*

Proof. See [67, Theorem 2.2], which is more general than stated here. The required specialization is, however, immediate. □

We consider a further specialisation, namely

Corollary 1. *Let* X *be an extreme point of* $\mathcal{T} := \{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A} \bullet \mathsf{X} \le a, \mathsf{B} \bullet \mathsf{X} \le b\}$. *Then* $\operatorname{rank}(\mathsf{X}) \le 1$.

Proof. After introducing slack variables, $\mathcal{T}$ is a projection of the set

$$\overline{\mathcal{T}} := \{(\mathsf{X}, s, t) \in \mathcal{S}^n_+ \times \mathbb{R}^2_+ \colon \mathsf{A} \bullet \mathsf{X} + s = a\,,\ \mathsf{B} \bullet \mathsf{X} + t = b\}$$

onto $\mathcal{S}^n$. From Theorem 6 it follows that

$$\operatorname{rank}(\mathsf{X})(\operatorname{rank}(\mathsf{X}) + 1) + \operatorname{rank}(s)(\operatorname{rank}(s) + 1) + \operatorname{rank}(t)(\operatorname{rank}(t) + 1) \le 4$$

and thus $\operatorname{rank}(\mathsf{X}) \le 1$. □

The following geometric condition will allow us to leverage Corollary 1 to cases where more than two linear inequalities are present.

Condition 1. For a collection of matrices $\mathsf{A}_i \in \mathcal{S}^n$ and real numbers b_i, $i \in [1\!:\!m]$ we say that Condition 1 holds if for any $\mathsf{X} \in \mathcal{S}^n_+$ with $\mathsf{A}_i \bullet \mathsf{X} \le b_i$ for all $i \in [1\!:\!m]$,

$$\mathsf{A}_k \bullet \mathsf{X} < b_k \quad \forall k \in [1\!:\!m] \setminus \{i, j\} \text{ whenever } \mathsf{A}_i \bullet \mathsf{X} = b_i \text{ and } \mathsf{A}_j \bullet \mathsf{X} = b_j \text{ for } i \neq j\,.$$

The condition requires that for any feasible $\mathsf{X} \in \mathcal{S}^n$ in above sense that at most two constraints can be binding at the same time. Note that, if $\mathcal{F} := \left\{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_i \bullet \mathsf{X} \le b_i, i \in [1\!:\!m]\right\}$ is bounded (as assumed in Theorem 7), one can check Condition 1 by solving $(m^3 - 3m^2 + 2m)/6$ semidefinite optimization problems of the form $\sup_{\mathsf{X} \in \mathcal{F}} \{\mathsf{A}_k \bullet \mathsf{X} - b_k \colon \mathsf{A}_i \bullet \mathsf{X} = b_i, \mathsf{A}_j \bullet \mathsf{X} = b_j\}$. For Condition 1 to hold, all the optimal values must be strictly smaller than 0.
We are now able to prove the following result for the homogeneous problem:

Theorem 7. *Suppose that Condition 1 holds for the matrices* $\mathsf{A}_i \in \mathcal{S}^n$ *and real numbers* $b_i \in \mathbb{R}$, $i \in [1 : m]$. *Further, suppose that the set* $\mathcal{F} := \left\{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_i \bullet \mathsf{X} \le b_i,\ i \in [1\!:\!m]\right\}$ *is bounded. Then*

$$\inf_{\mathbf{x} \in \mathbb{R}^n} \left\{\mathbf{x}^\mathsf{T} \mathsf{Q} \mathbf{x} \colon \mathbf{x}^\mathsf{T} \mathsf{A}_i \mathbf{x} \le b_i,\ i \in [1\!:\!m]\right\} = \inf_{\mathsf{X} \in \mathcal{S}^n_+} \{\mathsf{Q} \bullet \mathsf{X} \colon \mathsf{A}_i \bullet \mathsf{X} \le b_i,\ i \in [1\!:\!m]\}\,.$$

Proof. It is clear that "$\geq$" has to hold as the SDP is a relaxation of the QCQP. Since the former has linear objective and $\mathcal{F}$ is bounded, its optimal value will be attained at extreme points of $\mathcal{F}$. Let X^* be one such point. By Condition 1 at most two inequalities, say i and j, are binding at X^*. If fewer are binding, then one of these indices or both can be chosen arbitrarily. It follows that X^* is also extremal in the set $\mathcal{F}_{i,j} := \left\{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_i \bullet \mathsf{X} \leq b_i,\ \mathsf{A}_j \bullet \mathsf{X} \leq b_j\right\}$. To see this, note that by the strict inequalities in Condition 1 there is a ball $\mathcal{B}_\epsilon(\mathsf{X}^*)$ centered at X^* with radius $\epsilon \geq 0$ such that $\mathsf{A}_k \bullet \mathsf{X} < b_i,\ k \in [1:m] \setminus \{i,j\}$ whenever $\mathsf{X} \in \mathcal{B}_\epsilon(\mathsf{X}^*) \cap \mathcal{F} \subseteq \mathcal{F}_{i,j}$, so that $\mathcal{B}_\epsilon(\mathsf{X}^*) \cap \mathcal{F} = \mathcal{B}_\epsilon(\mathsf{X}^*) \cap \mathcal{F}_{i,j}$. If there was $\mathsf{X}_1, \mathsf{X}_2 \in \mathcal{F}_{i,j}$ such that X were a convex combination of those, then such points had to exist in $\mathcal{F}$ since the line between X_1 and X_2 would penetrate $\mathcal{B}_\epsilon(\mathsf{X}^*) \cap \mathcal{F}$. By Corollary 1 and extremality of X^* in $\mathcal{F}_{i,j}$ we see that $\mathsf{X}^* = \mathbf{x}^*(\mathbf{x}^*)^\mathsf{T}$ so that $\mathbf{x}^*$ is feasible for the QCQP, and $(\mathbf{x}^*)^\mathsf{T}\mathsf{Q}(\mathbf{x}^*) = \mathsf{Q} \bullet \mathbf{x}^*(\mathbf{x}^*)^\mathsf{T} = \mathsf{Q} \bullet \mathsf{X}^*$. $\square$

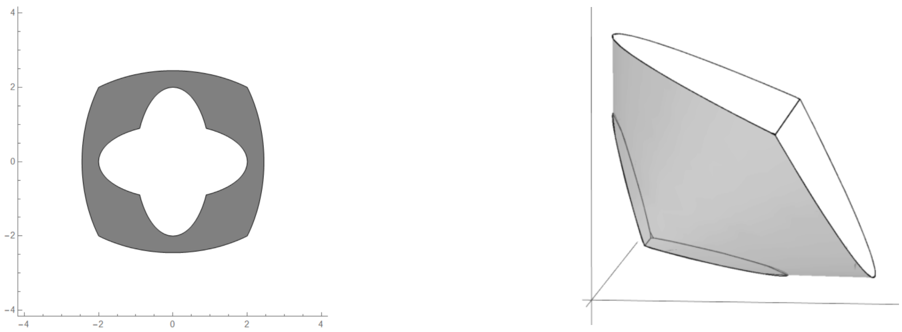
Example 6. *Consider the following quadratic program and its Shor relaxation*

$$
\begin{aligned}
\inf_{\mathbf{x} \in \mathbb{R}^2}\ & q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2 \\
\text{s.t.}:\ & 2x_1^2 + x_2^2 \leq 12, \\
& x_1^2 + 2x_2^2 \leq 12, \\
& 4x_1^2 + x_2^2 \geq 4, \\
& x_1^2 + 4x_2^2 \geq 4,
\end{aligned}
\qquad
\begin{aligned}
\inf_{\mathsf{X} \in \mathcal{S}^2_+}\ & q_{11}x_{11} + 2q_{12}x_{21} + q_{22}x_{22} \\
\text{s.t.}:\ & 2x_{11} + x_{22} \leq 12, \\
& x_{11} + 2x_{22} \leq 12, \\
& 4x_{11} + x_{22} \geq 4, \\
& x_{11} + 4x_{22} \geq 4.
\end{aligned}
$$

The feasible sets of these problems are depicted in Figure 3.1. For the lifted feasible set one can see that no more than two inequalities can be binding at the same time within the semidefinite cone. Also, since all matrices at the boundary of $\mathcal{S}^2_+$ are rank-one matrices, we see that the extreme points of the lifted feasible set have rank one, so that the conclusion of Corollary 1 is quite obvious in this case. By inspection one can clearly see that for $q_{12} = 0,\ q_{11} = q_{22} > 0$ one finds that the optimal set of the QCQP is at the innermost corners, where the third and forth inequality are binding. These are the points $(x_1, x_2) = (\pm 2/\sqrt{5}, \pm 2/\sqrt{5})$. For the Shor relaxation we can also easily see, that the third and fourth inequalities would be binding in the optimum and so would be the semidefiniteness constraint so that $x_{11}x_{22} = x_{12}^2$ for the optimal solution. From $4x_{11} + x_{22} = 4$ and $x_{11} + 4x_{22} = 4$ we get $x_{11} = x_{22} = 4/5$ so that $x_{12} = \pm 4/5$. We have

$$
\begin{pmatrix} \frac{4}{5} & \pm\frac{4}{5} \\ \pm\frac{4}{5} & \frac{4}{5} \end{pmatrix} = \begin{pmatrix} \pm\frac{2}{\sqrt{5}} \\ \pm\frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \pm\frac{2}{\sqrt{5}} \\ \pm\frac{2}{\sqrt{5}} \end{pmatrix}^\mathsf{T},
$$

as claimed in the theorem. Of course, there are more solutions to the Shor relaxation, but these are convex combinations of the two presented here. They can be seen in Figure 3.1 as the line connecting the two lower vertices in the lifted feasible set.



$\mathcal{F}$ $\qquad\qquad$ $\mathcal{G}(\mathcal{F})$

Figure 3.1

The feasible set $\mathcal{F}$ of the QCQP in Example 6 and the feasible set of its Shor relaxation (which by Theorem 7 coincides with $\mathcal{G}(\mathcal{F})$) side by side. The latter figure is based on the projection $\mathsf{M}(x, 2y, z) \mapsto (x, \sqrt{2}y, x)$, and we see the intersection between the cone of positive-semidefinite matrices and four half-spaces.

The optimization problem in Theorem 7 does not involve linear terms, which raises the question whether the result has implications for that case as well. In fact the theorem can be transferred to the case where the quadratic functions are non homogeneous by applying the following simple lemma:

Lemma 8. *The extreme points of* $\mathcal{T} := \{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_1 \bullet \mathsf{X} = b_1,\ \mathsf{A}_i \bullet \mathsf{X} \leq b_i,\ i \in [2\colon m]\}$ *are those extreme points of* $\overline{\mathcal{T}} := \{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_i \bullet \mathsf{X} \leq b_i,\ i \in [1\colon m]\}$ *for which* $\mathsf{A}_1 \bullet \mathsf{X} = b_1$ *holds.*

Proof. Assume $\mathsf{X} \in \operatorname{ext}(\mathcal{T})$, the set of extreme points of $\mathcal{T}$, and assume further that there are $\mathsf{X}_1, \mathsf{X}_2 \in \overline{\mathcal{T}}$ such that $\mathsf{X} = \lambda \mathsf{X}_1 + (1-\lambda)\mathsf{X}_2$ for some $\lambda \in (0,1)$. We have $\mathsf{A}_1 \bullet \mathsf{X}_j \leq \mathbf{b}_1$ for $j = 1, 2$. If $\mathsf{A}_1 \bullet \mathsf{X}_j < \mathbf{b}_1$ for at least one $j \in \{1, 2\}$, then $\mathsf{A}_1 \bullet \mathsf{X} = \lambda \mathsf{A}_1 \bullet \mathsf{X}_1 + (1-\lambda)\mathsf{A}_1 \bullet \mathsf{X}_2 < b_1$, so actually $\mathsf{A}_1 \bullet \mathsf{X}_j = \mathbf{b}_1$ for $j = 1, 2$, implying that $\mathsf{X}_1, \mathsf{X}_2 \in \mathcal{T}$ so that $\mathsf{X} \notin \operatorname{ext}(\mathcal{T})$, a contradiction. Now, assume $\mathsf{X} \in \operatorname{ext}\left(\overline{\mathcal{T}}\right)$ and that $\mathsf{A}_1 \bullet \mathsf{X} = b_1$, so clearly $\mathsf{X} \in \mathcal{T}$. Again, if there was a pair $\mathsf{X}_1, \mathsf{X}_2 \in \mathcal{T}$ such that $\mathsf{X} = \lambda \mathsf{X}_1 + (1-\lambda)\mathsf{X}_2$ for some $\lambda \in (0,1)$ there would be such a pair in $\overline{\mathcal{T}} \supseteq \mathcal{T}$, so that we arrive at an analogous contradiction. $\square$

Theorem 9. *Assume* $\mathsf{A}_i = \mathsf{M}(\omega_i, \mathbf{w}_i, \mathsf{W}_i)$ *and* $b_i = 0$, $i \in [1 : m]$, *together with* $\mathsf{A}_{m+1} := \mathsf{M}(1, \mathbf{o}, \mathsf{O})$ *and* $b_{m+1} = 1$, *fulfill Condition 1 and that the set*

$\mathcal{F}_{Shor} := \left\{\mathsf{M}(1, 2\mathbf{x}, \mathsf{X}) \in \mathcal{S}_+^{n+1} \colon \mathsf{W}_i \bullet \mathsf{X} + \mathbf{w}_i^\mathsf{T}\mathbf{x} + \omega_i \leq 0,\ i \in [1\colon m]\right\}$ *is bounded. Then*

$$\inf_{\mathbf{x}\in\mathbb{R}^n} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} \colon \mathbf{x}^\mathsf{T}\mathsf{W}_i\mathbf{x} + \mathbf{w}_i^\mathsf{T}\mathbf{x} + \omega_i \leq 0,\ i \in [1\colon m]\right\} =$$
$$\inf_{\mathsf{X},\mathbf{x}} \left\{\mathsf{Q} \bullet \mathsf{X} + \mathbf{q}^\mathsf{T}\mathbf{x} \colon \mathsf{W}_i \bullet \mathsf{X} + \mathbf{w}_i^\mathsf{T}\mathbf{x} + \omega_i \leq 0,\ i \in [1\colon m],\ \mathsf{M}(1, 2\mathbf{x}, \mathsf{X}) \in \mathcal{S}_+^{n+1}\right\}.$$

Proof. After reformulating the original problem as

$$\begin{aligned} \inf_{\mathbf{x}\in\mathbb{R}^n, x_0\in\mathbb{R}_+} \ & \mathsf{M}(0, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{M}(x_0^2, 2x_0\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T}) \\ & \mathsf{A}_i \bullet \mathsf{M}(x_0^2, 2x_0\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T}) \leq 0,\ i \in [1\colon m], \\ & \mathsf{A}_{m+1} \bullet \mathsf{M}(x_0^2, 2x_0\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T}) = 1, \end{aligned}$$

we see that its Shor relaxation is given by

$$\begin{aligned} \inf_{x_0,\mathbf{x},\mathsf{X}} \ & \mathsf{M}(0, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{M}(x_0, \mathbf{x}, \mathsf{X}) \\ & \mathsf{A}_i \bullet \mathsf{M}(x_0, \mathbf{x}, \mathsf{X}) \leq 0,\ i \in [1\colon m], \\ & \mathsf{A}_{m+1} \bullet \mathsf{M}(x_0, \mathbf{x}, \mathsf{X}) = 1, \\ & \mathsf{M}(x_0, 2\mathbf{x}, \mathsf{X}) \in \mathcal{S}_+^{n+1}, \end{aligned}$$

where we again make use of the fact that $\mathcal{CPP}(\mathbb{R}_+ \times \mathbb{R}^n) = \mathcal{S}_+^{n+1}$. Note that after resolving the inequality we see that the feasible set of the latter problem is exactly $\mathcal{F}_{shor}$. Proving that all extreme points are rank one is done in a manner analogous to Theorem 7, where the fact that we have one equality instead of an inequality does not interfere due to Lemma 8. In the representation of the reformulation we also combined the constraints $\mathsf{A}_{m+1} \bullet \mathsf{M}(x_0, \mathbf{x}, \mathsf{X}) = 1$ and $\mathsf{M}(x_0, 2\mathbf{x}, \mathsf{X}) \in \mathcal{S}_+^{n+1}$ into $\mathsf{M}(1, 2\mathbf{x}, \mathsf{X}) \in \mathcal{S}_+^{n+1}$ and also resolved the equality. □

Theorem 9 is related to the main result in [86] where the authors proved the following result.

Theorem 10. *Assume $\mathcal{F} \subseteq \mathbb{R}^n$ is the feasible set of a QCQP and assume that the for the set $\bar{\mathcal{F}} := \left\{\mathbf{x} \in \mathcal{F} \colon \mathbf{x}^\mathsf{T}\mathsf{W}_i\mathbf{x} + \mathbf{w}_i^\mathsf{T}\mathbf{x} + \omega_i \leq 0,\ i \in [1\colon m]\right\}$ the additional inequalities introduce non-intersecting hollows in $\mathcal{F}$. Then $\mathcal{G}(\bar{\mathcal{F}}) = \left\{(\mathsf{X}, \mathbf{x}) \in \mathcal{G}(\mathcal{F}) \colon \mathsf{W}_i \bullet \mathsf{X} + \mathbf{w}_i^\mathsf{T}\mathbf{x} + \omega_i \leq 0,\ i \in [1\colon m]\right\}$.*

The non-intersection condition is similar to the condition of Theorem 9. To see this, note that in the Shor relaxation in Theorem 9 the equality associated with A_{m+1} consumes one of two inequalities which can be binding at any time. Thus, the remaining inequalities cannot have intersections within the semidefinite cone, or else Condition 1 fails. Also note, that under the Condition 1 the set $\mathcal{F}_{Shor}$ is bounded if and only if it is bounded by merely one of the m inequalities in its description. From this we see that we can interpret Theorem 9 as a special case of Theorem 10 where $\mathcal{F}$ is given by a single quadratic equality and the non-intersection condition is with respect to the lifted space, i.e. a strictly stronger condition than in Theorem 10. It is an open question whether there is a direct way to deduce Theorem 10 from Theorem 7. However, Theorem 7 cannot be deduced from Theorem 10, and only the former will allow us to derive a generalization of the S-Lemma under geometric condition that is a derivative of Condition 1, as well as allow for some insight in the joint numerical range of the quadratic forms that jointly fulfil said condition.

3.2 Generalized S-lemma

We are now ready to use Theorem 7 in order to establish a generalization of the S-Lemma in the form of a characterization of set-copositivity over a cone given by homogeneous quadratic inequalities. The key to this result will be the fact that a strictly feasible point of a quadratic program guarantees a Slater point in its Shor relaxation, as proved in [81, Theorem 5.1].

Theorem 11. *Let $\mathcal{K} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\mathsf{T}\mathsf{A}_i\mathbf{x} \leq 0,\ i \in [1:m]\}$ with $\mathsf{A}_i \in \mathcal{S}^n$. Assume that there is some $\mathbf{x}_0$ with $\mathbf{x}_0^\mathsf{T}\mathsf{A}_i\mathbf{x}_0 < 0$ for all $i \in [1:m]$. Further, suppose that for all $i \in [1:m]$*

$$\mathsf{X} \in \mathcal{S}^n_+ \setminus \{\mathsf{O}\} \quad \textit{and} \quad \mathsf{A}_i \bullet \mathsf{X} = 0 \implies \mathsf{A}_j \bullet \mathsf{X} < 0 \quad \forall j \in [1:m] \setminus \{i\}\,. \tag{3.1}$$

Then

$$\mathcal{COP}(\mathcal{K}) = \left\{\mathsf{Q} \colon \mathsf{Q} + \sum_{i=1}^m \lambda_i \mathsf{A}_i \in \mathcal{S}^n_+ \textit{ for some } \boldsymbol{\lambda} \in \mathbb{R}^m_+\right\}.$$

Proof. Observe that $\mathsf{Q} \in \mathcal{COP}(\mathcal{K})$ if and only if $q^* := \inf_{\mathbf{x}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} \colon \mathbf{x}^\mathsf{T}\mathsf{I}_n\mathbf{x} \leq 1,\ \mathbf{x} \in \mathcal{K}\right\} \geq 0$, where the latter QCQP has a strictly feasible point $\mathbf{x}_0/(2\|\mathbf{x}_0\|)$. The matrices A_i, $i \in [1:m]$ fulfil (3.1) and thus also fulfil Condition 1 together with the pair I_n and 1. Thus, by Theorem 7 we have $q^* = \inf_{\mathsf{X} \in \mathcal{S}^n_+} \{\mathsf{Q} \bullet \mathsf{X} \colon \mathsf{I}_n \bullet \mathsf{X} \leq 1,\ \mathsf{A}_i \bullet \mathsf{X} \leq 0,\ i \in [1:m]\}$, and the latter set is bounded (due to the constraint $\mathsf{I}_n \bullet \mathsf{X} \leq 1$) and has a Slater point by [81, Theorem 5.1] since the QCQP has a strictly feasible point. Thus by strong duality $q^* = \sup_{\boldsymbol{\lambda} \in \mathbb{R}^m_+} \left\{-\mu \colon \mathsf{Q} + \sum_{i=1}^m \lambda_i \mathsf{A}_i + \mu \mathsf{I}_n \in \mathcal{S}^n_+\right\}$ which is nonnegative if and only if $\mathsf{Q} + \sum_{i=1}^m \lambda_i \mathsf{A}_i \in \mathcal{S}^n_+$ for some $\boldsymbol{\lambda} \in \mathbb{R}^m_+$. $\square$

3.3 Geometry of the joint numerical range

Note that the above theorem is stated in terms of a characterization of copositivity over a cone described by homogeneous quadratic functions. Another way of stating the result would be a theorem of the alternative, which usually is done when discussing the S-procedure, as discussed in the introduction. Such theorems are often derived from results on the joint numerical range of the quadratic forms involved. Specifically, if the joint numerical range can be shown to be a convex cone, an S-Lemma type result can be derived in a straightforward manner (see e.g. [70]). An important example of such a result was proved by Polyak in [71] where he showed that the S-procedure is exact for the case $m = 2$, $n \geq 3$ (under a condition discussed shortly) by invoking a convexity result regarding the joint numerical range of three quadratic functions, also provided in [71], namely:

Theorem 12. *Let $\mathsf{A}_i \in \mathcal{S}^n$, $i = 0, 1, 2$. For $n \geq 3$ the following assertions are equivalent.*

1. *There exists $\boldsymbol{\mu} \in \mathbb{R}^3$ such that*

$$\mu_0 \mathsf{A}_0 + \mu_1 \mathsf{A}_1 + \mu_2 \mathsf{A}_2 \in \mathcal{S}^n_{++}.$$

2. *The joint numerical range*

$$\mathcal{J} \coloneqq \left\{ \left(\mathbf{x}^\mathsf{T} \mathsf{A}_0 \mathbf{x}, \mathbf{x}^\mathsf{T} \mathsf{A}_1 \mathbf{x}, \mathbf{x}^\mathsf{T} \mathsf{A}_2 \mathbf{x} \right) : \mathbf{x} \in \mathbb{R}^n \right\}$$

   is a convex and pointed cone.

The theorem guarantees that whenever 1. holds and $\mathcal{J}$ is disjoint from the interior of the negative orthant, it can be separated from the latter set via a hyperplane. Thus, given the first condition in Theorem 12 and Slater's condition hold, one can prove exactness of the S-procedure, that is, one can prove the following theorem as shown in [71]:

Theorem 13. *Let $n \geq 3$ and $\mathsf{A}_i \in \mathcal{S}^n$, $i = 0, 1, 2$. Assume that there exists $\boldsymbol{\mu} \in \mathbb{R}^3$ such that $\mu_0 \mathsf{A}_0 + \mu_1 \mathsf{A}_1 + \mu_2 \mathsf{A}_2 \in \mathcal{S}^n_{++}$ and that there is an $\mathbf{x}_0 \in \mathbb{R}^n$ such that $\mathbf{x}_0^\mathsf{T} \mathsf{A}_i \mathbf{x}_0 < 0$, $i = 1, 2$, then the following two statements are equivalent*

i) *the following system of inequalities has no solution $\mathbf{x} \in \mathbb{R}^n$:*

$\mathbf{x}^\mathsf{T} \mathsf{A}_0 \mathbf{x} < 0$ *and* $\mathbf{x}^\mathsf{T} \mathsf{A}_i \mathbf{x} \leq 0$, $i = 1, 2$;

ii) *the following conic inequality has a solution $\boldsymbol{\lambda} \in \mathbb{R}^2_+$:*

$\mathsf{A}_0 + \lambda_1 \mathsf{A}_1 + \lambda_2 \mathsf{A}_2 \in \mathcal{S}^n_+$.

Obviously, the condition $\exists \boldsymbol{\mu} \in \mathbb{R}^3 \colon \mu_0 \mathsf{A}_0 + \mu_1 \mathsf{A}_2 + \mu_3 \mathsf{A}_3 \in \mathcal{S}^n_{++}$ is readily fulfilled if for at least one pair of $i \neq j$ there exists $\boldsymbol{\mu} \in \mathbb{R}^2$ with $\mu_1 \mathsf{A}_i + \mu_2 \mathsf{A}_j \in \mathcal{S}^n_{++}$ and we will refer to the later condition as Condition 1'. Thus, if we fix A_1 and A_2 such that Condition 1' is fulfilled, we can characterize copositivity over the cone $\mathcal{K} \coloneqq \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x}^\mathsf{T} \mathsf{A}_1 \mathbf{x} \leq 0, \mathbf{x}^\mathsf{T} \mathsf{A}_2 \mathbf{x} \leq 0\}$ for $n \geq 3$ as

$$\mathcal{COP}(\mathcal{K}) = \left\{ \mathsf{Q} \colon \mathsf{Q} + \lambda_1 \mathsf{A}_1 + \lambda_2 \mathsf{A}_2 \in \mathcal{S}^n_+ \text{ for some } \boldsymbol{\lambda} \in \mathbb{R}^2_+ \right\} .$$

The copositivity condition derived in Theorem 11 is, however, not obtained courtesy of any argument involving the joint numerical range, but rather as a consequence of the exactness of the Shor relaxation and the existence of a Slater point in conjunction with strong conic duality under the assumptions of the theorem. An immediate question is therefore, whether we can learn anything about the joint numerical range of the quadratic forms that fulfil said assumptions. This question is interesting in its own right since the joint numerical range is in general a very ill-behaved object and geometrical results are typically hard to obtain. We will gain at least some insight by proving that Condition 1' is implied by (3.1) and under additional assumptions both coincide.

Lemma 14. *Let* $\mathcal{K}_p \coloneqq \left\{ (\mathsf{A}_1 \bullet \mathsf{X}, \mathsf{A}_2 \bullet \mathsf{X}) \colon \mathsf{X} \in \mathcal{S}^n_+ \right\}$ *and* $\mathcal{K}_d \coloneqq \left\{ \mathbf{y} \in \mathbb{R}^2 \colon y_1 \mathsf{A}_1 + y_2 \mathsf{A}_2 \in \mathcal{S}^n_+ \right\}$*. Then* $\mathcal{K}_d$ *is a closed convex cone,* $\mathcal{K}_p$ *is a convex cone, and* $\mathcal{K}_p^* = \mathcal{K}_d$ *so that* $\mathcal{K}_d^* = \mathcal{K}_p^{**} \supseteq \mathcal{K}_p$*.*

Proof. The claims are immediate from the fact that $\mathcal{S}^n_+$ is a self-dual cone (therefore closed and convex). □

Proposition 15. *For matrices* $\mathsf{A}_1, \mathsf{A}_2 \in \mathcal{S}^n$*, assume that* $\exists \mathbf{x}_0 \in \mathbb{R}^n$ *such that* $\mathbf{x}_0^\mathsf{T} \mathsf{A}_i \mathbf{x}_0 < 0$ *for* $i \in \{1, 2\}$*. Consider the following two statements:*

1. *There exists* $\boldsymbol{\mu} \in \mathbb{R}^2$ *such that*

$$\mu_1 \mathsf{A}_1 + \mu_2 \mathsf{A}_2 \in \mathcal{S}^n_{++}.$$

2. *For any matrix* $\mathsf{X} \in \mathcal{S}^n_+ \setminus \{\mathsf{O}\}$ *and* $i, j \in \{1, 2\}$ *with* $i \neq j$ *we have*

$$\mathsf{A}_i \bullet \mathsf{X} = 0 \implies \mathsf{A}_j \bullet \mathsf{X} < 0,$$

Then, 2. $\implies$ 1. *and if we in addition assume that in the description of the set* $\mathcal{S} \coloneqq \{\mathsf{X} \in \mathcal{S}^n_+ : \mathsf{A}_i \bullet \mathsf{X} \leq 0\, ,\, i = 1, 2\}$ *none of the linear inequalities is redundant, then also* 1. $\implies$ 2.

Proof. For the first claim, assume that property 1. does not hold. We will show that condition 2. then has to fail. Under this hypothesis the linear subspace $\mathcal{M} :=\{\mu_1 \mathsf{A}_1 + \mu_2 \mathsf{A}_2 \colon \boldsymbol{\mu} \in \mathbb{R}^2\}$ is disjoint from $\mathcal{S}^n_{++}$. Thus, by [73, Theorem 11.2] we have a hyperplane containing $\mathcal{M}$, and hence $\mathcal{S}^n_{++}$ lies entirely in one of its associated open halfspaces. Since $\mathcal{M}$ contains the origin, so does said hyperplane; it actually supports $\mathcal{S}^n_+$. Thus, one of its normal vectors, say X, is in $\mathcal{S}^n_+$ by self-duality of the latter matrix-cone. But then for this $\mathsf{X} \in \mathcal{S}^n_+$ we have $\mathsf{A}_1 \bullet \mathsf{X} = \mathsf{A}_2 \bullet \mathsf{X} = 0$, i.e. property 2. fails to hold. For the other direction, assume that property 1. and the additional assumption about $\mathcal{S}$ hold. Consider the cones $\mathcal{K}_p := \{(\mathsf{A}_1 \bullet \mathsf{X}, \mathsf{A}_2 \bullet \mathsf{X}) \colon \mathsf{X} \in \mathcal{S}^n_+\}$ and $\mathcal{K}_d := \{\mathbf{y} \in \mathbb{R}^2 \colon y_1 \mathsf{A}_1 + y_2 \mathsf{A}_2 \in \mathcal{S}^n_+\}$. By Lemma 14 both are convex, $\mathcal{K}_d$ is closed and $\mathcal{K}_p^* = \mathcal{K}_d$. Since the cone of positive definite matrices has nonempty interior, property 1. is sufficient to guarantee that $\mathcal{K}_d$ has nonempty interior. Note that $\mathcal{K}_p$ has to meet the interior of the negative orthant, because of $\mathbf{x}_0^\mathsf{T} \mathsf{A}_i \mathbf{x}_0 < 0,\ i = 1, 2$. Also, $\mathcal{K}_p$ has to meet the interior of second orthant, since otherwise there is no $\mathsf{X} \in \mathcal{S}^n_+$ such that $\mathsf{A}_1 \bullet \mathsf{X} < 0, \mathsf{A}_2 \bullet \mathsf{X} > 0$ or in other words $\{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_1 \bullet \mathsf{X} < 0\} = \{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_1 \bullet \mathsf{X} < 0, \mathsf{A}_2 \bullet \mathsf{X} \leq 0\}$. The closures of the two sets would coincide, and by [73, Theorem 7.6] we had $\{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_1 \bullet \mathsf{X} \leq 0\} = \{\mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{A}_1 \bullet \mathsf{X} \leq 0, \mathsf{A}_2 \bullet \mathsf{X} \leq 0\}$ in contrast to the assumption. By an analogous argument we have nonempty intersection of $\mathcal{K}_p$ and the interior of the fourth orthant. Now assume that condition 2. fails, i.e. there is an $\mathsf{X} \in \mathcal{S}^n_+$ such that $\mathsf{A}_i \bullet \mathsf{X} = 0$ and $\mathsf{A}_j \bullet \mathsf{X} \geq 0$ for $i \neq j$. If in fact $\mathsf{A}_j \bullet \mathsf{X} = 0$ then X is a normal vector to a hyperplane that contains the linear combinations of $\mathsf{A}_1, \mathsf{A}_2$ and which supports $\mathcal{S}^n_+$ by self-duality. This clearly contradicts 1. as such a hyperplane is disjoint from $\mathcal{S}^n_{++}$. Thus assume $\mathsf{A}_j \bullet \mathsf{X} > 0$. Then $(\mathsf{A}_1 \bullet \mathsf{X}, \mathsf{A}_2 \bullet \mathsf{X})$ together with three more points from the intersection of $\mathcal{K}_p$ and the interior of the second, third and fourth orthant respectively, give four vectors whose positive linear combinations span $\mathbb{R}^2$. This span is contained in $\mathcal{K}_p$ so $\mathcal{K}_d = \{0\} = (\mathbb{R}^2)^*$, which also contradicts property 1. □

By this proposition, we could replace (3.1) with Condition 1' in Theorem 11 provided there are no redundant constraints. This gives some insight into the geometry of the joint numerical range $\mathcal{J}(\mathsf{Q}, \mathsf{A}_1, \ldots, \mathsf{A}_m)$ of $m+1$ matrices that satisfy the assumption of the Theorem 11. If $n \geq 3$, every projection of $\mathcal{J}(\mathsf{Q}, \mathsf{A}_1, \ldots, \mathsf{A}_m)$ onto three coordinates is in fact a convex, pointed cone by Theorem 12. We do not know whether these geometric properties of $\mathcal{J}(\mathsf{Q}, \mathsf{A}_1, \ldots, \mathsf{A}_m)$ are enough to ensure linear separability. However, Theorem 11 is not based on a convexity result concerning $\mathcal{J}$, but is a consequence of Pataki's theorem.

Chapter 4

Reducing the size of conic reformulations

Convex reformulations of the style studied in the preceding chapters, though theoretically appealing suffer from the fact that they provide bounds that require solving a set-completely positive optimization problems which have to be approximated by, for example, SDPs (see Appendix A for a list of known approximation schemes). This class of problems is amenable to interior point methods and are thus solvable in polynomial time. But in practice these methods for SDPs are known scale poorly and are not as numerically stable as, say, modern MILP-sovlers. Solving large scale SDPs requires substantial effort and is only possible for problems with exploitable special structure. On typical office-level computers only medium sized general SDPs can be solved. There are LP approximations for $\mathcal{CPP}(\mathbb{R}^n_+)$ but the advantage over SDPs is diminished and might even be outweighed by the size of these approximations. Recent literature (see [4]) introduced cutting plane algorithms for copositive optimization problems, where the subproblem of copositivity detection is solved via MILPs. However, these methods have not yet shown that they can remedy the computational expensiveness of completely positive optimization problems.

This motivates the search for smaller and cheaper convex approximations and reformulations of QCQPs, which may be achieved by exploiting special structure. In this spirit we consider the following QCQP

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_i} \ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathbf{x}^\mathsf{T}\mathsf{B}_i\mathbf{y}_i + \mathbf{y}_i^\mathsf{T}\mathsf{C}_i\mathbf{y}_i + \mathbf{c}_i^\mathsf{T}\mathbf{y}_i \\
\text{s.t.} : \ & \mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i, \quad i \in [1\!:\!S], \\
& \mathbf{x} \in \mathcal{K}_0 \\
& \mathbf{y}_i \in \mathcal{K}_i, \quad i \in [1\!:\!S],
\end{aligned}
\tag{4.1}
$$

where $\mathcal{K}_i \subseteq \mathbb{R}^{n_2}$, $i \in [0:S]$, are closed, convex cones, $\mathsf{A} \in \mathcal{S}^{n_1}$, $\mathbf{a} \in \mathbb{R}^{n_1}$, $\mathsf{B}_i \in \mathbb{R}^{n_1 \times n_2}$, $\mathsf{C}_i \in \mathcal{S}^{n_2}, \mathbf{c}_i \in \mathbb{R}^{n_2}$, $i \in [1:S]$ and $\mathsf{F}_i \in \mathbb{R}^{m_i \times n_1}$, $\mathsf{G}_i \in \mathbb{R}^{m_i \times n_2}$, $\mathbf{r}_i \in \mathbb{R}^{m_i}$, $i \in [1:S]$. The special structure in place here is, firstly, the missing bilinear terms in $\mathbf{y}_i$ and $\mathbf{y}_j$ for $i \neq j$ and secondly the absence of shared constraints between these groups of variables. A structure like this naturally arises in, for example, stochastic optimization, which we will discuss in Chapter 6.

We propose a reduced set-completely positive lower bound

Theorem 16. *The following set-completely positive problem provides a lower bound to (4.1):*

$$\begin{aligned}
\min_{\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i,\mathbf{x},\mathbf{y}_i} \ & \mathsf{A}_i \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathsf{B}_i \bullet \mathsf{Z}_i + \mathsf{C}_i \bullet \mathsf{Y}_i + \mathbf{c}_i^\mathsf{T}\mathbf{y} \\
\text{s.t.}: \ & \mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i, && i \in [1:S], \\
\operatorname{diag}\left(\begin{pmatrix}\mathsf{F}_i & \mathsf{G}_i\end{pmatrix}\begin{pmatrix}\mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i\end{pmatrix}\begin{pmatrix}\mathsf{F}_i^\mathsf{T} \\ \mathsf{G}_i^\mathsf{T}\end{pmatrix}\right) & = \mathbf{r}_i \circ \mathbf{r}_i, && i \in [1:S], \\
\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} & \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i), && i \in [1:S],
\end{aligned} \tag{4.2}$$

Proof. Let $(\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_S)$ be a feasible solution for (4.1). Set $\mathsf{X} = \mathbf{x}\mathbf{x}^\mathsf{T}$, $\mathsf{Y}_i = \mathbf{y}_i\mathbf{y}_i^\mathsf{T}$, $\mathsf{Z}_i = \mathbf{y}_i\mathbf{x}^\mathsf{T}, i \in [1:S]$. By the definition of the Frobenius product we have identical objective function values. Also the first set of constraints is fulfilled by feasibility. For the linear matrix constraints note that for all $i \in [1:S]$,

$$\operatorname{diag}\left(\begin{pmatrix}\mathsf{F}_i & \mathsf{G}_i\end{pmatrix}\begin{pmatrix}\mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i\end{pmatrix}\begin{pmatrix}\mathsf{F}_i^\mathsf{T} \\ \mathsf{G}_i^\mathsf{T}\end{pmatrix}\right) = \mathbf{r}_i \circ \mathbf{r}_i,$$
$$\Updownarrow$$
$$\left(\mathbf{f}_k^i\right)^\mathsf{T} \mathsf{X} \left(\mathbf{f}_k^i\right) + 2\left(\mathbf{f}_k^i\right)^\mathsf{T} \mathsf{Z}_i \left(\mathbf{g}_k^i\right) + \left(\mathbf{g}_k^i\right)^\mathsf{T} \mathsf{Y}_i \left(\mathbf{g}_k^i\right) = \left(r_k^i\right)^2, \ k \in [1:m_i],$$

which given the above choice for the matrix variables is just the squared version of the linear constraints which we already argued to hold. Finally

$$\mathcal{CPP}\left(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i\right) \ni \begin{pmatrix}1 \\ \mathbf{x} \\ \mathbf{y}_i\end{pmatrix}\begin{pmatrix}1 \\ \mathbf{x} \\ \mathbf{y}_i\end{pmatrix}^\mathsf{T} = \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix}, \quad i \in [1:S],$$

which concludes the proof. □

Note that, in case the feasible set is bounded, (4.1) has an exact set-completely positive reformulation described in [26]. The key difference to the bound proposed here is the size of the conic constraints. The traditional results would require to introduce lifted variables that correspond to the bilinear terms between

$\mathbf{y}_i$ and $\mathbf{y}_j$ for $i \neq j$ and to confine all variables appropriately to the matrix cone $\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \cdots \times \mathcal{K}_S)$. In (4.2) we instead use S matrix cones of smaller size and omit the spurious interaction terms entirely. Given that S is sufficiently large, the number of variables can be greatly reduced and the individual, small, conic constraints may have cheaper approximations than their larger counterpart. The cost of these advantages is that exactness results are hard to come by.

The rest of this chapter is organized as follows: In Section 4.1 we will give sufficient conditions for the exactness of (4.2). We will proof in Section 4.2 that the latter problem can itself be lower bounded by an SDP, which is obtained by means of sdp-completion results. In Section 4.3 we will introduce the cone of completable connected components and show that it can be used to obtain an exact reformulation of (4.1) that also has the quality that spurious terms are omitted. In Section 4.4 we show how the aforementioned cone is connected to matrix completion problems and we will use this connection in order to characterize the gap between the exact reformulation and our set-completely positive bound. Finally, in Section 4.5 we show how our results imply a novel theorem for set-completely positive matrix completion problems.

4.1 Exactness in special cases

We start the discussion by introducing some interesting special cases where we can proof that the reduced set-completely positive reformulation is in fact exact. Our result deals with the specialc case $m_i = 1,\ i \in [1\colon S]$. Though restrictive, this case is relevant for the discussion in Chapter 6. We summarize our findings in the following theorem.

Theorem 17. *Consider the problem*

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S}\ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathbf{x}^\mathsf{T}\mathsf{B}_i\mathbf{y}_i + \mathbf{y}_i^\mathsf{T}\mathsf{C}_i\mathbf{y}_i + \mathbf{c}_i^\mathsf{T}\mathbf{y}_i \\
\text{s.t.}:\ & \mathbf{f}_i^\mathsf{T}\mathbf{x} + \mathbf{g}_i^\mathsf{T}\mathbf{y}_i = 1, \quad i \in [1\colon S], \\
& \mathbf{x} \in \mathcal{K}_0, \\
& \mathbf{y}_i \in \mathcal{K}_i, \quad i \in [1\colon S].
\end{aligned}
\tag{4.3}
$$

The following set-completely positive optimization problem gives a lower bound and if $\mathbf{f}_i = \alpha_i\mathbf{f},\ \mathbf{c}_i = \beta_i\mathbf{g}_i,\ \mathsf{B}_i = \mathbf{b}_i\mathbf{g}_i^\mathsf{T},\ \mathbf{f}, \mathbf{b}_i \in \mathbb{R}^{n_1},\ \alpha_i, \beta_i \in \mathbb{R},\ i \in [1\colon S]$, *and we assume the sets* $\mathcal{F}_i = \left\{(\mathbf{x}^\mathsf{T}, \mathbf{y}_i^\mathsf{T})^\mathsf{T} \in \mathcal{K}_0 \times \mathcal{K}_i \colon \mathbf{f}_i^\mathsf{T}\mathbf{x} + \mathbf{g}_i^\mathsf{T}\mathbf{y}_i = 1\right\}, i \in [1\colon S]$ *are bounded then the* *bound is actually tight:*

$$\begin{aligned}
\min_{\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i,\mathbf{y}_i} \ & \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathsf{B}_i \bullet \mathsf{Z}_i + \mathsf{C}_i \bullet \mathsf{Y}_i + \mathbf{c}_i^\mathsf{T}\mathbf{y}_i \\
\text{s.t.}: \quad & \mathbf{f}_i^\mathsf{T}\mathbf{x} + \mathbf{g}_i^\mathsf{T}\mathbf{y}_i = 1, \quad i \in [1\!:\!S] \\
\mathbf{f}_i\mathbf{f}_i^\mathsf{T} \bullet \mathsf{X} + 2\ \mathbf{f}_i\mathbf{g}_i^\mathsf{T} \bullet \mathsf{Z}_i + \mathbf{g}_i\mathbf{g}_i^\mathsf{T} \bullet \mathsf{Y}_i &= 1, \quad i \in [1\!:\!S] \\
\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} &\in \ \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i) \quad i \in [1\!:\!S]
\end{aligned} \tag{4.4}$$

Proof. It is clear that the conic optimization problem gives a lower bound since we merely consider Theorem 16 under the special case $m_i = 1,\ i \in [1\!:\!S]$.

For the converse under the additional assumptions we can always transform the objective function

$$\begin{aligned}
&\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathbf{x}^\mathsf{T}\mathbf{b}_i\mathbf{g}_i^\mathsf{T}\mathbf{y}_i + \mathbf{y}_i^\mathsf{T}\mathsf{C}\mathbf{y}_i + \beta_i\mathbf{g}_i^\mathsf{T}\mathbf{y}_i \\
&= \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathbf{x}^\mathsf{T}\mathbf{b}_i\left(1 - \mathbf{f}_i^\mathsf{T}\mathbf{x}\right) + \mathbf{y}_i^\mathsf{T}\mathsf{C}\mathbf{y}_i + \beta_i(1 - \mathbf{f}_i^\mathsf{T}\mathbf{x}) \\
&= \mathbf{x}^\mathsf{T}\left(\mathsf{A} - 0.5\left(\sum_{i=1}^{S} \mathbf{b}_i\mathbf{f}_i^\mathsf{T} + \mathbf{f}_i\mathbf{b}_i^\mathsf{T}\right)\right)\mathbf{x} + \left(\mathbf{a} + \sum_{i=1}^{S}(\mathbf{b}_i - \beta_i\mathbf{f}_i)\right)^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathbf{y}_i^\mathsf{T}\mathsf{C}\mathbf{y}_i + \beta_i,
\end{aligned}$$

so we only need to consider the case where $\mathsf{B}_i = \mathsf{O},\ \mathbf{c}_i = \mathbf{o},\ i \in [1\!:\!S]$.

From Theorem 3 and the fact that $\mathcal{F}_i,\ i \in [1\!:\!S]$ are bounded we have the following characterization

$$\mathcal{G}(\mathcal{F}_i) = \left\{ \begin{array}{r} \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i) \quad : \qquad\qquad \\ \mathbf{f}_i^\mathsf{T}\mathbf{x} + \mathbf{g}_i^\mathsf{T}\mathbf{y}_i \quad = 1, \\ \mathbf{f}_i\mathbf{f}_i^\mathsf{T} \bullet \mathsf{X} + 2\ \mathbf{f}_i\mathbf{g}_i^\mathsf{T} \bullet \mathsf{Z}_i + \mathbf{g}\mathbf{g}_i^\mathsf{T} \bullet \mathsf{Y}_i \quad = 1, \end{array} \right\}.$$

so that for any element of that set we have a representation

$$\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_I^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} = \sum_{i=1}^{k} \lambda_l \begin{pmatrix} 1 \\ \mathbf{x}_l^i \\ \mathbf{y}_l^i \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x}_l^i \\ \mathbf{y}_l^i \end{pmatrix}^\mathsf{T} : \mathbf{f}_i^\mathsf{T}\mathbf{x}_l^i + \mathbf{g}_i^\mathsf{T}\mathbf{y}_l^i = 1,\ \mathbf{x}_l^i \in \mathcal{K}_0,\ \mathbf{y}_l^i \in \mathcal{K}_i, \tag{4.5}$$

for some $\lambda_l \geq 0,\ l \in [1\!:\!k],\ \sum_{i=1}^{k} \lambda_l = 1$.

Define

$$\phi_i(\mathbf{x}, \mathsf{X}) = \min_{\mathbf{y}_i,\mathsf{Y}_i,\mathsf{Z}_i} \left\{ \mathsf{C} \bullet \mathsf{Y}_i : \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{G}(\mathcal{F}_i) \right\} \tag{4.6}$$

and

$$\mathcal{X}_i = \left\{ (\mathbf{x}, \mathsf{X}) \colon \exists (\mathbf{y}_i, \mathsf{Y}_i, \mathsf{Z}_i) \colon \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{G}\left(\mathcal{F}_i\right) \right\}.$$

The sets $\mathcal{X}_i$ are the projections of $\mathcal{G}(\mathcal{F}_i)$ onto the the $(\mathbf{x}, \mathsf{X})$ coordinates. Since the extreme points of a projection of a convex set correspond to extreme points of the projected set, it follows that the extreme points of $\mathcal{X}_i$ are of the form $(\mathbf{x}, \mathbf{x}\mathbf{x}^\mathsf{T})$.

We will now show that for $i^* \in \arg\min_k \{1/\alpha_k : k \in [1\colon S]\}$ we have $\mathcal{X}_i \subseteq \mathcal{X}_j$ for all $j \in [1\colon S]$. Assume $(\mathbf{x}, \mathsf{X}) \in \mathcal{X}_{i^*}$ and choose any $j \neq i^*$. Since we assume that $\mathcal{F}_i, i \in [1\colon S]$ are bounded we can assume w.l.o.g. that $\alpha_k > 0$ with $(\mathbf{f}, \mathbf{g}_k) \in \operatorname{int}(\mathcal{K}_0^* \times \mathcal{K}_k^*)$ for $k \in [1\colon S]$. By the definition of $\mathcal{X}_{i^*}$ there exist $\mathsf{Y}_{i^*},\ \mathbf{y}_{i^*},\ \mathsf{Z}_{i^*}$ that complete $(\mathbf{x}, \mathsf{X})$ to an element of $\mathcal{G}\left(\mathcal{F}_{i^*}\right)$ so that we have a representation as in (4.5) with

$$\mathbf{f}_{i^*}^\mathsf{T}\mathbf{x}_l + \mathbf{g}_i^\mathsf{T}\mathbf{y}_l^{i^*} = 1,\ \mathbf{x}_l \in \mathcal{K}_0,\ \mathbf{y}_l^{i^*} \in \mathcal{K}_{i^*},\ l \in [1\colon k]$$

suppressing dependence of $\mathbf{x}_l$ on i^*. We then have from $\alpha_{i^*}^{-1}\mathbf{g}_{i^*}^\mathsf{T}\mathbf{y}_l^{i^*} \geq 0$ that

$$\mathbf{f}^\mathsf{T}\mathbf{x}_l \leq \mathbf{f}^\mathsf{T}\mathbf{x}_l + \alpha_{i^*}^{-1}\mathbf{g}_{i^*}^\mathsf{T}\mathbf{y}_l^{i^*} = \alpha_{i^*}^{-1} \leq \alpha_j^{-1}.$$

Now, for any $l \in [1\colon k]$, since $\alpha_j^{-1} - \mathbf{f}^\mathsf{T}\mathbf{x}_l \geq 0$, $\mathbf{g}_j \in \operatorname{int}\mathcal{K}_j^*$ and $\alpha_j > 0$ there exists an $\mathbf{y}_l^j \in \mathcal{K}_j$ so that $\alpha_j^{-1}\mathbf{g}_j^\mathsf{T}\mathbf{y}_l^j = \alpha_j^{-1} - \mathbf{f}^\mathsf{T}\mathbf{x}_l$ and we have $\mathbf{f}^\mathsf{T}\mathbf{x}_l + \alpha_j^{-1}\mathbf{g}_j^\mathsf{T}\mathbf{y}_l^j = \alpha_j^{-1}$ which implies $\mathbf{f}_j^\mathsf{T}\mathbf{x}_l + \mathbf{g}_j^\mathsf{T}\mathbf{y}_l^j = 1$. Thus, we can replace $\mathbf{y}_l^{i^*}$ by $\mathbf{y}_l^j$ in the vectors of the convex-combination representation (4.5) to obtain $(\mathbf{y}_j, \mathsf{Y}_j, \mathsf{Z}_j)$ that certify $(\mathbf{x}, \mathsf{X}) \in \mathcal{X}_j$. We can therefore define $\mathcal{X} \coloneqq \bigcap_{i=1}^S \mathcal{X}_i = \mathcal{X}_{i^*}$ for any $i^* \in \arg\min_k \{1/\alpha_k : k \in [1\colon S]\}$ and rewrite (4.3) as

$$\min_{(\mathbf{x}, \mathsf{X}) \in \mathcal{X}} \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^S \phi_i(\mathbf{x}, \mathsf{X}).$$

The feasible set in the description of $\phi(\mathbf{x}, \mathsf{X})$ in (4.6) is given by (we suppress the indices of the variables for the moment to avoid double indices in the following paragraphs)

$$\mathcal{Y}_i(\mathbf{x}, \mathsf{X}) \coloneqq \left\{ \mathsf{Y} \colon \exists (\mathsf{Z}, \mathbf{y}) \colon \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}^\mathsf{T} \\ \mathbf{y} & \mathsf{Z} & \mathsf{Y} \end{pmatrix} \in \mathcal{G}\left(\mathcal{F}_i\right) \right\},$$

We will now show that

$$\mathcal{Y}_i(\mathbf{x}, \mathsf{X}) \subseteq \left\{ \mathsf{Y} \colon \mathbf{g}_i\mathbf{g}_i^\mathsf{T} \bullet \mathsf{Y} = 1 - 2\mathbf{f}_i^\mathsf{T}\mathbf{x} + \mathbf{f}_i\mathbf{f}_i^\mathsf{T} \bullet \mathsf{X},\ \mathsf{Y} \in \mathcal{CPP}\left(\mathcal{K}_i\right) \right\} =: \bar{\mathcal{Y}}_i(\mathbf{x}, \mathsf{X}).$$

Assume $\mathsf{Y} \in \mathcal{Y}_i(\mathbf{x}, \mathsf{X})$, then $\mathsf{Y} \in \mathcal{CPP}(\mathcal{K}_i)$. Also, by the definition of $\mathcal{Y}_i(\mathbf{x}, \mathsf{X})$ there exist a $\mathsf{Z}, \mathbf{y}$ that completes $(\mathsf{Y}, \mathsf{X}, \mathbf{x})$ to an element of $\mathcal{G}(\mathcal{F}_i)$ so that we again have a representation as in (4.5) with

$$\mathbf{f}_i^{\mathsf{T}} \mathbf{x}_l + \mathbf{g}_i^{\mathsf{T}} \mathbf{y}_l = 1,\ \mathbf{x}_l \in \mathcal{K}_0,\ \mathbf{y}_l \in \mathcal{K}_i,\ l \in [1\!:\!k].$$

This allows us to perform the following reformulation:

$$\begin{aligned}
\mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \mathsf{X} + 2\ \mathbf{f}_i\mathbf{g}_i^{\mathsf{T}} \bullet \mathsf{Z} + \mathbf{g}_i\mathbf{g}_i^{\mathsf{T}} \bullet \mathsf{Y} &= 1, \\
\mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \sum_{i=1}^{k} \lambda_l \mathbf{x}_l\mathbf{x}_l^{\mathsf{T}} + 2\mathbf{f}_i\mathbf{g}_i^{\mathsf{T}} \bullet \sum_{i=1}^{k} \lambda_l \mathbf{y}_l\mathbf{x}_l^{\mathsf{T}} + \mathbf{g}_i\mathbf{g}_i^{\mathsf{T}} \bullet \sum_{i=1}^{k} \lambda_l \mathbf{y}_l\mathbf{y}_l^{\mathsf{T}} &= 1, \\
\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)^2 + 2\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)\left(\mathbf{g}_i^{\mathsf{T}}\mathbf{y}_l\right) + \sum_{i=1}^{k} \lambda_l \left(\mathbf{g}_i^{\mathsf{T}}\mathbf{y}_l\right)^2 &= 1, \\
\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)^2 + 2\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)\left(1 - \mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right) + \sum_{i=1}^{k} \lambda_l \left(\mathbf{g}_i^{\mathsf{T}}\mathbf{y}_l\right)^2 &= 1, \\
\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)^2 + 2\left(\sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right) - \sum_{i=1}^{k} \lambda_l \left(\mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l\right)^2\right) + \sum_{i=1}^{k} \lambda_l \left(\mathbf{g}_i^{\mathsf{T}}\mathbf{y}_l\right)^2 &= 1, \\
2\mathbf{f}_i^{\mathsf{T}}\mathbf{x} - \mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \mathsf{X} + \mathbf{g}_i\mathbf{g}_i^{\mathsf{T}} \bullet \mathsf{Y} &= 1,
\end{aligned}$$

and thus we have $\mathsf{Y} \in \bar{\mathcal{Y}}_i(\mathbf{x}, \mathsf{X})$. Also we get $1 - 2\mathbf{f}_i^{\mathsf{T}}\mathbf{x} + \mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \mathsf{X} = \sum_{i=1}^{k} \lambda_l (1 - \mathbf{f}_i^{\mathsf{T}}\mathbf{x}_l)^2 \geq 0$.

From our discussion in Theorem 2 we see that $\bar{\mathcal{Y}}_i(\mathbf{x}, \mathsf{X})$ are compact, conic intersections that have rank one extreme points. We thus have

$$\min_{\mathsf{Y} \in \bar{\mathcal{Y}}_i(\mathbf{x}, \mathsf{X})} \mathsf{C} \bullet \mathsf{Y} = \mathsf{C} \bullet \bar{\mathbf{y}}\bar{\mathbf{y}}^{\mathsf{T}} \left(1 - 2\mathbf{f}_i^{\mathsf{T}}\mathbf{x} + \mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \mathsf{X}\right) \text{ where } \bar{\mathbf{y}}\bar{\mathbf{y}}^{\mathsf{T}} \in \arg \min_{\mathsf{Y} \in \bar{\mathcal{Y}}_i(\mathbf{o}, \mathsf{O})} \mathsf{C} \bullet \mathsf{Y},$$

so that $\mathbf{g}_i^{\mathsf{T}}\bar{\mathbf{y}} = 1$. Now consider

$$\min_{(\mathbf{x}, \mathsf{X}) \in \mathcal{X}} \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^{\mathsf{T}}\mathbf{x} + \sum_{i=1}^{S} \min_{\mathsf{Y}_i \in \bar{\mathcal{Y}}_i(\mathbf{x}, \mathsf{X})} \mathsf{C}_i \bullet \mathsf{Y}_i. \tag{4.7}$$

We have

$$\begin{aligned}
\text{val}(4.3) &= \min_{(\mathbf{x}, \mathsf{X}) \in \mathcal{X}} \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^{\mathsf{T}}\mathbf{x} + \sum_{i=1}^{S} \phi_i(\mathbf{x}, \mathsf{X}) \\
\geq \text{val}(4.7) &= \min_{(\mathbf{x}, \mathsf{X}) \in \mathcal{X}} \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^{\mathsf{T}}\mathbf{x} + \sum_{i=1}^{S} \mathsf{C}_i \bullet \bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^{\mathsf{T}} \left(1 - 2\mathbf{f}_i^{\mathsf{T}}\mathbf{x} + \mathbf{f}_i\mathbf{f}_i^{\mathsf{T}} \bullet \mathsf{X}\right) \\
&= \mathsf{A} \bullet \bar{\mathbf{x}}\bar{\mathbf{x}}^{\mathsf{T}} + \mathbf{a}^{\mathsf{T}}\bar{\mathbf{x}} + \sum_{i=1}^{S} \mathsf{C}_i \bullet \bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^{\mathsf{T}} \left(1 - \mathbf{f}_i^{\mathsf{T}}\bar{\mathbf{x}}\right)^2,
\end{aligned}$$

since the objective is linear and the extreme points of $\mathcal{X}$ are rank one so that the minimium is attained at some $(\bar{\mathbf{x}}, \bar{\mathbf{x}}\bar{\mathbf{x}}^{\mathsf{T}}) \in \mathcal{X}$. We have $\bar{\mathbf{x}} \in \mathcal{K}_0,\ \left(1 - \mathbf{f}_i^{\mathsf{T}}\bar{\mathbf{x}}\right)\bar{\mathbf{y}}_i \in \mathcal{K}_i,\ i \in$

$[1\!:\!S]$ and $\mathbf{f}_i^\mathsf{T}\bar{\mathbf{x}}+\left(1-\mathbf{f}_i^\mathsf{T}\bar{\mathbf{x}}\right)\mathbf{g}_i^\mathsf{T}\bar{\mathbf{y}} = \mathbf{f}_i^\mathsf{T}\bar{\mathbf{x}}+1-\mathbf{f}_i^\mathsf{T}\bar{\mathbf{x}} = 1$ so that we have a feasible solution for (4.3) that gives the same objective function value so that in total we have

$$\mathrm{val}(4.3) \geq \mathrm{val}(4.4) \geq \mathrm{val}(4.7) \geq \mathrm{val}(4.3),$$

implying that $\mathrm{val}(4.3) = \mathrm{val}(4.4)$.

What remains to be shown is that, in case we implemented changes in the objective function mentioned at the beginning, we can undo these changes in the objective function of the reformulation so that we truly arrive at the reformulation stated in the theorem. So consider a feasible solution to (4.4). We can rearrange the transformed objective function

$$\begin{aligned}
&\left(\mathsf{A}-0.5\left(\sum_{i=1}^{S}\mathbf{b}_i\mathbf{f}_i^\mathsf{T}+\mathbf{f}_i\mathbf{b}_i^\mathsf{T}\right)\right)\bullet\mathsf{X}+\left(\mathbf{a}+\sum_{i=1}^{S}(\mathbf{b}_i-\beta_i\mathbf{f}_i)\right)^\mathsf{T}\mathbf{x}+\sum_{i=1}^{S}+\mathsf{C}\bullet\mathsf{Y}_i+\beta_i\\
=&\mathsf{A}\bullet\mathsf{X}+\mathbf{a}^\mathsf{T}\mathbf{x}+\sum_{i=1}^{S}-0.5\left(\mathbf{b}_i\mathbf{f}_i^\mathsf{T}+\mathbf{f}_i\mathbf{b}_i^\mathsf{T}\right)\bullet\mathsf{X}+(\mathbf{b}_i-\beta_i\mathbf{f}_i)^\mathsf{T}\mathbf{x}+\mathsf{C}_i\bullet\mathsf{Y}_i+\beta_i,
\end{aligned}$$

and consider the S terms in the sum individually. We have $(\mathbf{y}_i,\mathsf{Y}_i)\in\mathcal{Y}_i(\mathbf{x},\mathsf{X}),\ i\in[1\!:\!S]$ so that we again can again consider the above decomposition where $\mathbf{f}_i^\mathsf{T}\mathbf{x}_l+\mathbf{g}_i^\mathsf{T}\mathbf{y}_l^i = 1,\ l\in[1\!:\!S]$ that allows us to rearrange

$$\begin{aligned}
&-0.5\left(\mathbf{b}_i\mathbf{f}_i^\mathsf{T}+\mathbf{f}_i\mathbf{b}_i^\mathsf{T}\right)\bullet\mathsf{X}+(\mathbf{b}_i-\beta_i\mathbf{f}_i)^\mathsf{T}\mathbf{x}+\mathsf{C}_i\bullet\mathsf{Y}_i+\beta_i\\
=&\sum_{l=1}^{k}\lambda_l\left(-0.5\left(\mathbf{b}_i\mathbf{f}_i^\mathsf{T}+\mathbf{f}_i\mathbf{b}_i^\mathsf{T}\right)\bullet\mathbf{x}_l\mathbf{x}_l^\mathsf{T}+(\mathbf{b}_i-\beta_i\mathbf{f}_i)^\mathsf{T}\mathbf{x}_l+\mathsf{C}_i\bullet\mathbf{y}_l^i\left(\mathbf{y}_l^i\right)^\mathsf{T}+\beta_i\right)\\
=&\sum_{l=1}^{k}\lambda_l\left(-\mathbf{x}_l^\mathsf{T}\mathbf{b}_i\left(1-\mathbf{g}_i^\mathsf{T}\mathbf{y}_l^i\right)+\mathbf{b}_i^\mathsf{T}\mathbf{x}_l-\beta_i\left(1-\mathbf{g}_i^\mathsf{T}\mathbf{y}_l^i\right)+\mathsf{C}_i\bullet\mathbf{y}_l^i\left(\mathbf{y}_l^i\right)^\mathsf{T}+\beta_i\right)\\
=&\sum_{l=1}^{k}\lambda_l\left(\mathbf{b}_i\mathbf{g}_i^\mathsf{T}\bullet\mathbf{y}_l^i\mathbf{x}_l+\mathsf{C}_i\bullet\mathbf{y}_l^i\left(\mathbf{y}_l^i\right)^\mathsf{T}+\beta_i\mathbf{g}_i^\mathsf{T}\mathbf{y}_l^i\right)\\
=&\mathsf{B}_i\bullet\mathsf{Z}_i+\mathsf{C}_i\bullet\mathsf{Y}_i+\mathbf{c}_i\mathbf{y}_i,
\end{aligned}$$

so that we recover the original form. □

4.2 A performance guarantee based on SDP-completion

A primitive way to construct a lower bound for (4.1) is the SDP-relaxation given by

$$
\begin{aligned}
\min_{\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i,\mathbf{x},\mathbf{y}_i} \ & \mathsf{A}_i \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathsf{B}_i \bullet \mathsf{Z}_i + \mathsf{C}_i \bullet \mathsf{Y}_{i,i} + \mathbf{c}_i^\mathsf{T}\mathbf{y} \\
\text{s.t.}: \ & \mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i, \qquad i \in [1:S], \\
\operatorname{diag}\left(\begin{pmatrix} \mathsf{F}_i & \mathsf{G}_i \end{pmatrix} \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \begin{pmatrix} \mathsf{F}_i^\mathsf{T} \\ \mathsf{G}_i^\mathsf{T} \end{pmatrix} \right) & = \mathbf{r}_i \circ \mathbf{r}_i, \qquad i \in [1:S], \\
\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_1^\mathsf{T} & \dots & \mathbf{y}_S^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_1^\mathsf{T} & \dots & \mathsf{Z}_S^\mathsf{T} \\ \mathbf{y}_1 & \mathsf{Z}_1 & \mathsf{Y}_{1,1} & \dots & \mathsf{Y}_{1,S} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_S & \mathsf{Z}_S & \mathsf{Y}_{S,1} & \dots & \mathsf{Y}_{S,S} \end{pmatrix} & =: \mathsf{M} \in \mathcal{S}_+^{n_1+Sn_2+1}.
\end{aligned}
\tag{4.8}
$$

The fact that it yields a lower bound follows trivially from a similar argument as in Theorem 16. It is, however, not immediate that it is also a lower bound to (4.2) which we will prove in the sequel. First, we have to establish a technical lemma concerned with the so called specification graph of a partially determined matrix. This type of graphs have a deep connection to matrix completion problems which we explain in greater detail in Appendix A.3. All the terms and theorems that are necessary to understand the following discussion are introduced there. The lemma is more detailed than needed for the discussion in this section, however, the additional details will be needed in later sections so we might as well establish them here.

Lemma 18. *Let $S > 1$ and consider a partial matrix of where the specified entries exhibit an arrow-head structure, i.e.*

$$
\begin{pmatrix}
\mathsf{X} & \mathsf{Z}_1^\mathsf{T} & \mathsf{Z}_2^\mathsf{T} & \dots & \mathsf{Z}_{S-1}^\mathsf{T} & \mathsf{Z}_S^\mathsf{T} \\
\mathsf{Z}_1 & \mathsf{Y}_{1,1} & * & \dots & * & * \\
\mathsf{Z}_2 & * & \mathsf{Y}_{2,2} & \dots & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\mathsf{Z}_{S-1} & * & * & \dots & \mathsf{Y}_{S-1,S-1} & * \\
\mathsf{Z}_S & * & * & \dots & * & \mathsf{Y}_{S,S}
\end{pmatrix},
$$

where $\mathsf{X} \in \mathcal{S}^{n_1}$, $\mathsf{Y}_{i,i} \in \mathcal{S}^{n_2}$, $\mathsf{Z}_i \in \mathbb{R}^{n_2 \times n_1}$, $i \in [1:S]$, and let G_{spec} be its specification graph. Then G_{spec} is chordal. If $n_1 \in \{0,1\}$ then G_{spec} is also a block-clique graph, which is not the case otherwise.

Proof. We start out by showing that G_{spec} is chordal in general. We group the nodes of the specification graph into $S+1$ groups where the first group $g_0 = \{1, \dots n_1\}$ are the nodes that correspond to the first n_1 rows of the matrix and whose internal edges are specified by the north west entries X. The second group $g_1 = \{n_1+1, \dots, n_1+n_2\}$

corresponds to the rows $n_1 + 1$ to $n_1 + n_2$ whose internal edges are specified by the blocks $\mathsf{Y}_{2,2}$ and whose external edges, connecting to neighbors outside of g_1, are specified by Z_1. The construction of the remaining groups proceeds accordingly. We will now show that any cycle of length greater 3 must have a chord. Note that all the groups are cliques since the blocks $\mathsf{Y}_{i,i}$ are fully specified. Thus, a cycle of length greater than 3 must have a chord if it is entirely contained in one of the groups. We therefore only need to consider cycles that are not entirely contained in one group. Also, any member of g_0 is a neighbor to any other node in the graph since the blocks $\mathsf{Z}_i,\ i \in [1\colon S]$ are fully specified. Thus, if a vertex v of g_0 is visited by a cycle, then the edge to any other node in the cycle that is not the predecessor of v gives a chord. A cycle that visits more than one group needs to visit g_0 since the other groups are not connected to one another and thus has a chord.

If $n_1 = 1$ then g_0 is a singleton. A block cannot contain just vertices from multiple $g_i,\ i \in [1 : S]$ since these groups are pairwise disconnected. A connection can only be established by adding g_0 but then the single node in g_0 is a cut vertex, i.e. the subgraph can become disconnected by deleting a single node and its adjacent edges. Hence, a block of G_{spec} must be a subgraph formed from the union of g_0 and one $g_i, i \in [1\colon S]$ and the respective edges. A subgraph formed from all the nodes of such a union, say T, cannot be contained in any other block since the construction of such a block would require to add nodes from a third group. Thus, T is a block, but it is also a clique since the q_i is a clique and the node in g_0 is adjacent to all the members of g_i.

If $n_1 = 0$ then G_{spec} consists of S subgraphs that are cliques and pairwise disconnected, hence they are blocks.

Otherwise, the entire graph is its only block since it cannot become disconnected by deleting a single node and its adjacent edges, but this block is not a clique since $g_i,\ i \in [1\colon S]$ have no inter-group edges. $\square$

Theorem 19. *We have* $\mathrm{val}(4.2) \geq \mathrm{val}(4.8)$ *and the statement also holds if we replace the cones* $\mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i),\ i \in [1\colon S]$ *by any other subset of* $\mathcal{S}^{n_1+n_2+1}_+,\ i \in [1\colon S]$ *as long as the resulting feasible set is non-empty.*

Proof. Clearly, the two problems have the same objective function, so we only need to compare the feasible sets. Let $(\mathsf{X}, \mathsf{Y}_1, \ldots, \mathsf{Y}_S, \mathsf{Z}_i, \ldots, \mathsf{Z}_S, \mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_S)$ be such that

$$\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{S}^{n_1+n_2+1},\ i \in [1\colon S] \tag{4.9}$$

and the linear constraints in (4.2) are fulfilled. Then certainly the linear constraints of (4.8) are fulfilled and what remains to show, is that, after setting $\mathsf{Y}_{i,i} = \mathsf{Y}_i,\ i \in [1\!:\!S]$, we can find $\mathsf{Y}_{i,j},\ i \neq j$ such that the matrix M becomes positive semidefinite. By Theorem 55 it suffices to show that the specification graph of the partial matrix where $\mathsf{Y}_{i,j},\ i \neq j$ are not specified is a chordal graph and that all fully specified principal submatrices are positive semidefinite (see Appendix A.3 for the statement of the theorem and an explanation of all relevant terms). So consider the partial matrix

$$\begin{pmatrix} 1 & \mathbf{x}^{\mathsf{T}} & \mathbf{y}_1^{\mathsf{T}} & \mathbf{y}_2^{\mathsf{T}} & \dots & \mathbf{y}_{S-1}^{\mathsf{T}} & \mathbf{y}_S^{\mathsf{T}} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} & \mathsf{Z}_2^{\mathsf{T}} & \dots & \mathsf{Z}_{S-1}^{\mathsf{T}} & \mathsf{Z}_S^{\mathsf{T}} \\ \mathbf{y}_1 & \mathsf{Z}_1 & \mathsf{Y}_{1,1} & * & \dots & * & * \\ \mathbf{y}_2 & \mathsf{Z}_2 & * & \mathsf{Y}_{2,2} & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{y}_{S-1} & \mathsf{Z}_{S-1} & * & * & \dots & \mathsf{Y}_{S-1,S-1} & * \\ \mathbf{y}_S & \mathsf{Z}_S & * & * & \dots & * & \mathsf{Y}_{S,S} \end{pmatrix}.$$

Since in all but the first two columns (we use this word now referring to literal columns in the above representation) have unspecified blocks one can only obtain fully specified principal submatrices if one deletes all but one of the partially specified columns and all but the respective rows (again in the literal sense). The so obtained blocks are precisely the blocks in (4.9) and are thus positive semidefinite. The chordality of the specification graph follows from Lemma 18. This completes the proof. □

The theorem states that our reduced set-completely positive reformulation is at least as strong as the SDP-relaxation and thus gives a theoretical performance guarantee. It also applies to relaxations of (4.2) such as the $\mathcal{DNN}$-relaxation since $\mathcal{DNN}^n \subseteq \mathcal{S}_+^n$. Another implication is, that (4.2) is exact whenever (4.8) is.

Remark. *We could have arrived at Theorem 19 by using the results in [44] who describe a chordality-detection procedure for SDPs with chordal sparsity pattern. However, this procedure seemed more complicated than proving chordality of arrow-head matrices directly here. It is nonetheless important to note that the above result is not the first of its kind, but can be obtained directly from known results in literature. Still, to the best of our knowledge, the context in which we use this technique is original.*

4.3 A geometrical approach: the cone of completable connected components

As discussed above Section 2.1.3, one way to prove the exactness of set-completely positive reformulations is to show that the feasible set is of the form $\mathbb{J} \cap \mathbb{H}$ where $\mathbb{J}$

is a convex, conic subset of the convex hull an appropriately chosen cone $\mathbb{K}$ and $\mathbb{H}$ is a hyperplane. A natural question is, whether we can execute a similar strategy for proving the exactness of a reduced conic reformulation by replacing the cone of extreme rays of $\mathcal{CPP}(\mathcal{K})$ with another appropriately structured object as our choice for $\mathbb{K}$. In this section we will show that this line of thought leads us at least to an exact, reduced conic reformulation that, like (4.2), does not involve spurious terms but also exhibits a particular structure that, while being fundamentally different from (4.2) might help to get new perspectives on the gap between those two problems.

Two facilitate our endeavor we need to introduce some new objects. Firstly, we define

$$\mathcal{S}_n^{S,k} := \left\{ \left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} \\ \mathsf{Z}_1 & \mathsf{Y}_1 \end{pmatrix}, \dots, \begin{pmatrix} \mathsf{X} & \mathsf{Z}_S^{\mathsf{T}} \\ \mathsf{Z}_S & \mathsf{Y}_S \end{pmatrix} \right] : \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^{\mathsf{T}} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{S}_n,\ i \in [1\!:\!S],\ \mathsf{X} \in \mathcal{S}_k \right\},$$

i.e. the set of (row-)vectors of S symmetric matrices of order n connected by a component of order $k \leq n$. In order to distinguish elements of $\mathcal{S}_n^{S,k}$ from normal matrices we use san-serif letters braced by rectangular braces, for example $[\mathsf{A}]$. Note, that $\mathcal{S}_n^{S,k}$ is isomorphic to the space of arrowhead matrices by the isomorphism

$$\Gamma \colon \mathcal{S}_n^{S,k} \to \mathcal{S}_{k+Sn},\ [\mathsf{A}] \mapsto \Gamma\left([\mathsf{A}]\right) := \begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} & \dots & \mathsf{Z}_S \\ \mathsf{Z}_1 & \mathsf{Y}_1 & \dots & \mathsf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Z}_S & \mathsf{O} & \dots & \mathsf{Y}_S \end{pmatrix},$$

where for the inverse we have

$$\Gamma^{-1} \begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} & \dots & \mathsf{Z}_S \\ \mathsf{Z}_1 & \mathsf{Y}_1 & \dots & \mathsf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Z}_S & \mathsf{O} & \dots & \mathsf{Y}_S \end{pmatrix} = \left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} \\ \mathsf{Z}_1 & \mathsf{Y}_1 \end{pmatrix}, \dots, \begin{pmatrix} \mathsf{X} & \mathsf{Z}_S^{\mathsf{T}} \\ \mathsf{Z}_S & \mathsf{Y}_S \end{pmatrix} \right] \in \mathcal{S}_n^{S,k}.$$

Thus, $\mathcal{S}_n^{S,k}$ is a vector space with a natural inner product $[\mathsf{A}] \odot [\mathsf{B}] := \Gamma\left([\mathsf{A}]\right) \bullet \Gamma\left([\mathsf{A}]\right)$, sum $[\mathsf{A}] \oplus [\mathsf{B}] := \Gamma\left([\mathsf{A}]\right) + \Gamma\left([\mathsf{A}]\right)$ and scalar product $\lambda\,[\mathsf{A}] := \Gamma^{-1}\left(\lambda \Gamma\left([\mathsf{A}]\right)\right)$. For notational convenience we will expand the meaning of the inverse Γ^{-1} so that it is applicable to non-arrowhead matrices as well, where the nonzero off-diagonal blocks are treated as though they were blocks of zeros as in the definition above. Also, we define a second, analogous isomorphism $\Gamma_*(\cdot)$ that maps into the space of partial matrices where the blocks of zeros in the definition of $\Gamma(\cdot)$ are not specified (for the definition of partial matrices see Appendix A.3). We also will use the shorthand notation

$$\left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} \\ \mathsf{Z}_1 & \mathsf{Y}_1 \end{pmatrix}, \dots, \begin{pmatrix} \mathsf{X} & \mathsf{Z}_S^{\mathsf{T}} \\ \mathsf{Z}_S & \mathsf{Y}_S \end{pmatrix} \right] = \left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^{\mathsf{T}} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \right]_{i \in [1:S]}.$$

The central object we are interested in is the following subset of $\mathcal{S}_n^{S,k}$:

$$\mathcal{CMP}_n^{S,k}(\times_{i=0}^S \mathcal{K}_i) := \operatorname{conv}\left\{ \left[\begin{pmatrix} \mathbf{x} \\ \mathbf{y}_i \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y}_i \end{pmatrix}^{\mathsf{T}} \right]_{i\in[1:S]} : \begin{pmatrix} \mathbf{x} \\ \mathbf{y}_i \end{pmatrix} \in \mathcal{K}_0 \times \mathcal{K}_i,\ i \in [1\colon S] \right\}.$$

We call $\mathcal{CMP}_n^{S,k}$ the cone of *completable connected components* and we will justify that name in the succeeding section. Note, that the extreme rays of this cone are given by all its members where the components are rank-1.

In the following theorem we show that by choosing $\mathbb{K} = \operatorname{ext}\mathcal{CMP}_n^{S,k}\left((\mathbb{R}_+ \times \mathcal{K}_0) \times_{i=1}^S \mathcal{K}_i\right)$ and $\mathbb{J}$ and $\mathbb{H}$ appropriately we can use Theorem 4 in order to obtain an exact conic reformulation of (4.1) that strongly resembles (4.2).

Theorem 20. *Considering (4.1), assume* $\mathcal{L}_i := \left\{\left(\mathbf{x}^{\mathsf{T}}, \mathbf{y}_i^{\mathsf{T}}\right) \in \mathcal{K}_0 \times \mathcal{K}_i \colon \mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i\right\}$ *are nonempty bounded sets. Then (4.1) is equivalent to*

$$\begin{aligned} \min_{[\mathsf{X}]\in\mathcal{S}_{n_1+n_2+1}^{S,n_1+1}} & [\mathsf{C}] \odot [\mathsf{X}] \\ \text{s.t.} :& [\mathsf{H}_0] \odot [\mathsf{X}] = 1, \\ & [\mathsf{F}_i] \odot [\mathsf{X}] = 0,\ i \in [1\colon S], \\ & [\mathsf{X}] \in \mathcal{CMP}_{n_1+n_2+1}^{S,n_1+1}\left((\mathbb{R}_+ \times \mathcal{K}_0) \times_{i=1}^S \mathcal{K}_i\right), \end{aligned} \tag{4.10}$$

where $[\mathsf{C}], [\mathsf{F}_i] \in \mathcal{S}_{n_1+n_2+1}^{S,n_1+1}$ *are defined as*

$$[\mathsf{C}] := \Gamma^{-1}\begin{pmatrix} 1 & \frac{1}{2}\mathbf{a}^{\mathsf{T}} & \frac{1}{2}\mathbf{c}_1^{\mathsf{T}} & \dots & \frac{1}{2}\mathbf{c}_S^{\mathsf{T}} \\ \frac{1}{2}\mathbf{a} & \mathsf{A} & \frac{1}{2}\mathsf{B}_1 & \dots & \frac{1}{2}\mathsf{B}_S \\ \frac{1}{2}\mathbf{c}_1^{\mathsf{T}} & \frac{1}{2}\mathsf{B}_1^{\mathsf{T}} & \mathsf{C}_1 & \dots & \mathsf{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\mathbf{c}_S^{\mathsf{T}} & \frac{1}{2}\mathsf{B}_S^{\mathsf{T}} & \mathsf{O} & \dots & \mathsf{C}_i \end{pmatrix}, \quad [\mathsf{H}_0] = \Gamma^{-1}\left(\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}\right),$$

$$[\mathsf{F}_i] := \Gamma^{-1}\left((-\mathbf{r}_i, \mathsf{F}_i, \mathsf{O}, \dots, \mathsf{G}_i, \dots, \mathsf{O})^{\mathsf{T}}(-\mathbf{r}_i, \mathsf{F}_i, \mathsf{O}, \dots, \mathsf{G}_i, \dots, \mathsf{O})\right).$$

Proof. Consider the following equivalences

$$\mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i,\ \mathbf{y} \in \mathcal{K}_i,\ i \in [1\colon S],\ \mathbf{x} \in \mathcal{K}_0,$$

$$\Updownarrow$$

$$\left\| (-\mathbf{r}_i, \mathsf{F}_i, \mathsf{O}, \dots, \mathsf{G}_i, \dots, \mathsf{O}) \begin{pmatrix} x_0 \\ \mathbf{x} \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_S \end{pmatrix} \right\|^2 = \mathbf{o},\ \mathbf{y} \in \mathcal{K}_i,\ i \in [1\colon S],\ \mathbf{x} \in \mathcal{K}_0,\ x_0 = 1,$$

$$\Updownarrow$$

$$[\mathsf{F}_i]\odot\Gamma^{-1}\left(\begin{pmatrix}x_0\\ \mathbf{x}\\ \mathbf{y}_1\\ \vdots\\ \mathbf{y}_S\end{pmatrix}\begin{pmatrix}x_0\\ \mathbf{x}\\ \mathbf{y}_1\\ \vdots\\ \mathbf{y}_S\end{pmatrix}^{\mathsf{T}}\right)=\mathbf{o},\ \mathbf{y}\in\mathcal{K}_i,\ i\in[1\colon S],\ \mathbf{x}\in\mathcal{K}_0,\ x_0=1,$$

$$\Updownarrow$$

$$[\mathsf{F}_i]\odot[\mathsf{X}]=0,\ [\mathsf{H}_0]\odot[\mathsf{X}]=1,\ [\mathsf{X}]\in\mathrm{ext}\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right).$$

Invoking Theorem 4, we specify

$$\begin{aligned}&\mathbb{K}=\mathrm{ext}\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right),\\ &\mathbb{H}=\left\{[\mathsf{X}]\in\mathcal{S}^{S,n_1+1}_{n_1+n_2+1}\colon[\mathsf{H}_0]\odot[\mathsf{X}]=1\right\},\end{aligned}$$

and we need to show is that

$$\mathbb{J}=\left\{[\mathsf{X}]\in\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right):[\mathsf{F}_i]\odot[\mathsf{X}]=0,\ i\in[1\colon S]\right\}$$

is a face of $\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right)$. By Theorem 5, this will follow if we can show that $[\mathsf{F}_i]\odot[\mathsf{X}]\geq 0,\ \forall[\mathsf{X}]\in\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right),\ i\in[1:S]$. Indeed for any of the $[\mathsf{F}_i]$ and any $[\mathsf{X}]\in\mathrm{ext}\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right)$ we have

$$\begin{aligned}[\mathsf{F}_i]\odot[\mathsf{X}]&=\left((-\mathbf{r}_i,\mathsf{F}_i,\dots,\mathsf{G}_i,\dots)^{\mathsf{T}}(-\mathbf{r}_i,\mathsf{F}_i,\dots,\mathsf{G}_i,\dots)\right)\bullet\begin{pmatrix}\mathbf{x}\mathbf{x}^{\mathsf{T}}&\mathbf{x}\mathbf{y}_1^{\mathsf{T}}&\dots&\mathbf{x}\mathbf{y}_S^{\mathsf{T}}\\ \mathbf{y}_1\mathbf{x}^{\mathsf{T}}&\mathbf{y}_1\mathbf{y}_1^{\mathsf{T}}&\dots&\mathsf{O}\\ \vdots&\vdots&\ddots&\vdots\\ \mathbf{y}_S\mathbf{x}^{\mathsf{T}}&\mathsf{O}&\dots&\mathbf{y}_1\mathbf{y}_1^{\mathsf{T}}\end{pmatrix}\\ &=\left(-\mathbf{r}_i,\mathsf{F}_i,\mathsf{G}_i\right)^{\mathsf{T}}\left(-\mathbf{r}_i,\mathsf{F}_i,\mathsf{G}_i\right)\bullet\begin{pmatrix}\mathbf{x}\mathbf{x}^{\mathsf{T}}&\mathbf{x}\mathbf{y}_i^{\mathsf{T}}\\ \mathbf{y}_i\mathbf{x}^{\mathsf{T}}&\mathbf{y}_i\mathbf{y}_i^{\mathsf{T}}\end{pmatrix}\\ &=\left\|\left(-\mathbf{r}_i,\mathsf{F}_i,\mathsf{G}_i\right)\begin{pmatrix}\mathbf{x}\\ \mathbf{y}_i\end{pmatrix}\right\|^2\geq 0.\end{aligned}$$

To complete the proof we need to show that the feasible set is bounded. To this end we consider its recession cone which by [73, Corollary 8.3.3.] is given by

$$0^+\mathcal{F}\coloneqq\left\{[\mathsf{X}]\in\mathcal{CMP}^{S,n_1+1}_{n_1+n_2+1}\left((\mathbb{R}_+\times\mathcal{K}_0)\times_{i=1}^{S}\mathcal{K}_i\right):[\mathsf{H}_0]\odot[\mathsf{X}]=0,\ [\mathsf{F}_i]\odot[\mathsf{X}]=0,\ i\in[1\colon S]\right\}.$$

Take an arbitrary $[\mathsf{X}]\in 0^+\mathcal{F}$, then

$$[\mathsf{F}_i]\odot[\mathsf{X}]=\sum_{l=1}^{k}\lambda_l\left\|\left(-\mathbf{r}_i,\mathsf{F}_i,\mathsf{G}_i\right)\begin{pmatrix}x_l^0\\ \mathbf{x}_l\\ \mathbf{y}_l^i\end{pmatrix}\right\|^2=0,\ i\in[1\colon S]\text{ and}$$

$$[\mathsf{H}_0] \odot [\mathsf{X}] = \sum_{l=1}^{k} \left(x_l^0\right)^2 = 0, \text{ implying that } x_l^0 = 0, \; l \in [1\!:\!k].$$

Thus, for any $i \in [1:S]$ and $l \in [1:k]$ we have $\mathsf{F}_i \mathbf{x}_l + \mathsf{G}_i \mathbf{y}_l^i = \mathbf{o}$ and $(0, \mathbf{x}_l, \mathbf{y}_l^i) \in \mathbb{R}_+ \times \mathcal{K}_0 \times \mathcal{K}_i$ so that we have a element of the recession cone of $\mathcal{L}_i$, which only contains the origin by the boundedness assumption, so that $[\mathsf{X}] = [\mathsf{O}]$. □

The resemblance between (4.2) and (4.10) stems from the fact that they are linear, conic optimization problems, both of which do not involve any of the spurious terms arising from the lifting of the interaction terms $\mathbf{y}_i \mathbf{y}_j^{\mathsf{T}}$, $i \neq j$. In the following section we will investigate the gap between these two problems on the basis of matrix completion problems.

4.4 Connection to matrix completion

We start our discussion by making a basic observation about set-completely positive matrix completion.

Lemma 21. *Let $\mathcal{K}_0 \subseteq \mathbb{R}^{n_1}$, $\mathcal{K}_i \subseteq \mathbb{R}^{n_2}$, $i \in [1\!:\!S]$. Then the matrix partial matrix*

$$\mathsf{M}_* := \begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^{\mathsf{T}} & \dots & \mathsf{Z}_S \\ \mathsf{Z}_1 & \mathsf{Y}_1 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Z}_S & * & \dots & \mathsf{Y}_S \end{pmatrix}, \quad \mathsf{X} \in \mathcal{S}^{n_1}, \; \mathsf{Y}_i \in \mathcal{S}^{n_2}, \; \mathsf{Z}_i \in \mathbb{R}^{n_2 \times n_1}, \; i \in [1\!:\!S].$$

has a completion $\mathsf{M} \in \mathcal{CPP}(\times_{i=0}^{S} \mathcal{K}_i)$ if and only if there are decompositions of the fully specified submatrices

$$\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^{\mathsf{T}} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} = \begin{pmatrix} \bar{\mathsf{X}}\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{X}}\bar{\mathsf{Y}}_i^{\mathsf{T}} \\ \bar{\mathsf{Y}}_i\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{Y}}_i\bar{\mathsf{Y}}_i^{\mathsf{T}} \end{pmatrix}, \; \textit{with} \; \begin{pmatrix} \bar{\mathsf{X}} \\ \bar{\mathsf{Y}}_i \end{pmatrix} \in (\mathcal{K}_0 \times \mathcal{K}_i)^r, \; i \in [1\!:\!S], \; r \in \mathbb{N}$$

Proof. Given said decompositions we can create a matrix

$$\begin{pmatrix} \bar{\mathsf{X}} \\ \bar{\mathsf{Y}}_1 \\ \vdots \\ \bar{\mathsf{Y}}_S \end{pmatrix} \text{ for which } \begin{pmatrix} \bar{\mathsf{X}} \\ \bar{\mathsf{Y}}_1 \\ \vdots \\ \bar{\mathsf{Y}}_S \end{pmatrix} \begin{pmatrix} \bar{\mathsf{X}} \\ \bar{\mathsf{Y}}_1 \\ \vdots \\ \bar{\mathsf{Y}}_S \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} \bar{\mathsf{X}}\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{X}}\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{X}}\bar{\mathsf{Y}}_S^{\mathsf{T}} \\ \bar{\mathsf{Y}}_1\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{Y}}_1\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{Y}}_1\bar{\mathsf{Y}}_S^{\mathsf{T}} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathsf{Y}}_S\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{Y}}_S\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{Y}}_S\bar{\mathsf{Y}}_S^{\mathsf{T}} \end{pmatrix} \in \mathcal{CPP}(\times_{i=0}^{S} \mathcal{K}_i),$$

is the desired completion of M_*. Conversely, if M_* has a completion $\mathsf{M} \in \mathcal{CPP}(\times_{i=0}^{S} \mathcal{K}_i)$ then by definition of the latter cone we have

$$\mathsf{M} = \begin{pmatrix} \bar{\mathsf{X}}\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{X}}\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{X}}\bar{\mathsf{Y}}_S^{\mathsf{T}} \\ \bar{\mathsf{Y}}_1\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{Y}}_1\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{Y}}_1\bar{\mathsf{Y}}_S^{\mathsf{T}} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathsf{Y}}_S\bar{\mathsf{X}}^{\mathsf{T}} & \bar{\mathsf{Y}}_S\bar{\mathsf{Y}}_1^{\mathsf{T}} & \dots & \bar{\mathsf{Y}}_S\bar{\mathsf{Y}}_S^{\mathsf{T}} \end{pmatrix},$$

so that

$$\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} = \begin{pmatrix} \bar{\mathsf{X}}\bar{\mathsf{X}}^\mathsf{T} & \bar{\mathsf{X}}\bar{\mathsf{Y}}_i^\mathsf{T} \\ \bar{\mathsf{Y}}_i\bar{\mathsf{X}}^\mathsf{T} & \bar{\mathsf{Y}}_i\bar{\mathsf{Y}}_i^\mathsf{T} \end{pmatrix}, \text{ with } \begin{pmatrix} \bar{\mathsf{X}} \\ \bar{\mathsf{Y}}_i \end{pmatrix} (\mathcal{K}_0 \times \mathcal{K}_i)^r, \; i \in [1\!:\!S], \; r \in \mathbb{N}.$$

□

The lemma is easily derived, but it highlights the key difficulty for the construction of a completion of the arrow-head arrangement of a set of matrix blocks connected by a common submatrix X. If all of the blocks have representations as a convex-conic combinations (i.e. nonnegative linear combinations) where the parts of the representations that form the connecting X-component are identical for all blocks, obtaining the completion is simply a matter of concatenating the individual factors of the decompositions. However, there is no guarantee that decompositions that are coordinated in this manner do exist. In fact, for the case where $\mathcal{K}_i$ are all nonnegative orthants, Theorem 56 states that there have to be cases where the arrowhead matrix is not completable since its specification graph is in general not a block-clique graph (see Lemma 18). It stands to reason that this phenomenon occurs for more general choices of $\mathcal{K}_i, \; i \in [0\!:\!S]$, although we cannot prove this claim yet.

We can now justify the name of the cone of connected completable components, since from Lemma 21 it is immediate, that

Proposition 22. *We have*

$$\left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix}\right]_{i\in[1:S]} \in \mathcal{CMP}_n^{S,k}(\times_{i=0}^S \mathcal{K}_i)$$

if and only if the matrix

$$\mathsf{M}_* \coloneqq \begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^\mathsf{T} & \dots & \mathsf{Z}_S \\ \mathsf{Z}_1 & \mathsf{Y}_1 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Z}_S & * & \dots & \mathsf{Y}_S \end{pmatrix}$$

has a completion $\mathsf{M} \in \mathcal{CPP}(\times_{i=0}^S \mathcal{K}_i)$.

It is therefore not surprising that we were able to obtain an exact reformulation of (4.1) using the cone of completable connected components. In fact a slight modification of (4.11) given by

$$
\begin{aligned}
\min_{\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i,\mathbf{x},\mathbf{y}_i} \ & \mathsf{A}_i \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{S} \mathsf{B}_i \bullet \mathsf{Z}_i + \mathsf{C} \bullet \mathsf{Y}_i + \mathbf{c}_i^\mathsf{T}\mathbf{y} \\
\text{s.t. :} \ & \mathsf{F}_i\mathbf{x} + \mathsf{G}_i\mathbf{y}_i = \mathbf{r}_i, \quad i \in [1\colon S], \\
\operatorname{diag}\left(\begin{pmatrix} \mathsf{F}_i & \mathsf{G}_i \end{pmatrix} \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \begin{pmatrix} \mathsf{F}_i^\mathsf{T} \\ \mathsf{G}_i^\mathsf{T} \end{pmatrix} \right) & = \mathbf{r}_i \circ \mathbf{r}_i, \quad i \in [1\colon S], \\
\Gamma\left(\left[\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_i^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathbf{y}_i & \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \right]_{i\in[1:S]} \right) & \text{ has a completion in } \mathcal{CPP}(\mathbb{R}_+ \times_{i=0}^{S} \mathcal{K}_i)
\end{aligned}
\tag{4.11}
$$

closes the relaxation gap since the feasible set of (4.11) can easily be projected into the feasible set of the exact set-completely positive reformulation based on Theorem 3.

Remark. *This observation is not new and was recently addressed [53] who used $\mathcal{CPP}$-completion theory in order to reduce the size of completely positive reformulations with block-clique structure. Their approach allows for more general structure in the problem data, since they use the properties of the specification graphs (block-clique structure) of the matrix-valued problem data as qualifiers for describing the family of problems in the scope of their approach. By comparison the structure of (4.1) is more specific. However, they only obtain exactness results for QCQPs where the connecting component of the groups of variables (in our case $\mathbf{x}$) is one-dimensional, an assumption we do not need in Theorem 17.*

The above observation also allows us to further characterize the gap between (4.2) and (4.11). To this end let us define the following generalization of the set-completely positive matrix cone:

$$
\begin{aligned}
\mathcal{CPP}_n^{S,k}\left(\times_{i=0}^{S}\mathcal{K}_i\right) \coloneqq & \left\{ \left[\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \right]_{i\in[1:S]} : \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{CPP}^n, (\mathcal{K}_0 \times \mathcal{K}_i),\ i \in [1\colon S],\ \mathsf{X} \in \mathcal{S}^k \right\} \\
& \subseteq \mathcal{S}_n^{S,k}.
\end{aligned}
$$

It is clear that if we replace $\mathcal{CMP}_n^{S,k}$ with $\mathcal{CPP}_n^{S,k}$ (modulo the respective specifications) in (4.11) we retrieve (4.2). Thus, the gap between these to problems can be characterized by the difference of the two sets.

Proposition 23. *Let $\mathcal{K}_0 \subseteq \mathbb{R}^{n_1},\ \mathcal{K}_i \subseteq \mathbb{R}^{n_2},\ i \in [1\colon S],\ S > 1$. Then*

$$
\mathcal{CMP}_{n_1+n_2}^{S,n_1}\left(\times_{i=0}^{S}\mathcal{K}_i\right) \subseteq \mathcal{CPP}_{n_1+n_2}^{S,n_1}\left(\times_{i=0}^{S}\mathcal{K}_i\right). \tag{4.12}
$$

If $n_1 = 0$ then the sets are identical. If $\mathcal{K}_i,\ i \in [0\colon S]$ are positive orthants then if $n_1 \in \{0,1\}$ the sets are identical otherwise the containment is strict.

Proof. The inclusion follows immediately from the definition of both objects as does the equality in case $n_1 = 0$. The strictness in case all cones are the nonnegative orthant and we don't have $n_1 \notin \{0, 1\}$ follows from the 'only if' direction in Theorem 56 and the fact that the specification graph of arrowhead matrices are then not block-clique graphs by Lemma 18. In case $n_1 \in \{0, 1\}$ these graphs are block-clique also by Lemma 18 so that the statement follows from the 'if'-direction in Theorem 56. $\square$

Example 7. *In order to illustrate the proposition we can consider rearrangement of a matrix that was demonstrated in [8, Example 2.23] to be not completable. It is given by*

$$\mathsf{A} := \begin{pmatrix} 6 & 3 & 0 & 3 \\ 3 & 6 & 3 & 0 \\ 0 & 3 & 2 & * \\ 3 & 0 & * & 2 \end{pmatrix},$$

where $n_1 = 2$ and $n_2 = 1$. The determinant is given $\det(\mathsf{A}) = -27\,(a_{4,3} + 1)^2$ *so that it can't be completed to an element of $\mathcal{CPP}_{n_1+2n_2}$. However, the fully specified submatrices given by*

$$\mathsf{A}_1 := \begin{pmatrix} 6 & 3 & 0 \\ 3 & 6 & 3 \\ 0 & 3 & 2 \end{pmatrix}, \quad \mathsf{A}_2 := \begin{pmatrix} 6 & 3 & 3 \\ 3 & 6 & 0 \\ 3 & 0 & 2 \end{pmatrix},$$

are easily shown to be positive semidefinite and therefore completely positive. Thus, $[\mathsf{A}_1, \mathsf{A}_2] \in \mathcal{CPP}_3^{2,2}\left(\mathbb{R}^3_+\right)$ *but* $[\mathsf{A}_1, \mathsf{A}_2] \notin \mathcal{CMP}_3^{2,2}\left(\mathbb{R}^3_+\right)$

The above discrepancy between the two sets as well as the construction of the results in Theorem 20 might open an new line of attack to exactness results for (4.2). Whenever we can show that the two sets have a common face and that the linear constraints in (4.2) confine the feasible set to that face, we have that the reformulation is actually tight. However, as discussed in [36], the identification of faces of $\mathcal{CPP}\left(\mathbb{R}^n_+\right)$ is a difficult task in and of itself, so we cannot hope to easily obtain results on the faces of $\mathcal{CPP}_n^{S,k}\left(\times_{i=0}^S \mathcal{K}_i\right)$ and $\mathcal{CMP}_n^{S,k}\left(\times_{i=0}^S \mathcal{K}_i\right)$. We leave such a discussion to future research and close the chapter with a discussion on the consequences that Theorem 17 has for matrix completion.

4.5 Conditional set-completely positive matrix completion

We discussed in the previous section how the exactness of reduced conic relaxations can be proved by matrix completion results. However, the implications run both ways:

if we can close the relaxation gap of reduced conic reformulations of a certain type, we may obtain new results on matrix completions. Let $\mathcal{K}_0 \subseteq \mathbb{R}^{n_1}$ and $\mathcal{K}_i \subseteq \mathbb{R}^{n_2}$, $i \in [1\colon S]$ be cones, $f_i\colon \mathcal{S}^{n_1+n_2} \to \mathbb{R}^{m_i}, i \in [1\colon S]$ be linear functions and $\mathbf{r}_i \in \mathbb{R}^{m_i}, i \in [1\colon S]$. For an $\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2}$, let us fix a subdivision

$$\bar{\mathsf{X}} \coloneqq \begin{pmatrix} \mathsf{X} & \mathsf{Z}_1^\mathsf{T} & \dots & \mathsf{Z}_S^\mathsf{T} \\ \mathsf{Z}_1 & \mathsf{Y}_{1,1} & \dots & \mathsf{Y}_{1,S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Z}_S & \mathsf{Y}_{S,1} & \dots & \mathsf{Y}_{S,S} \end{pmatrix},$$

where $\mathsf{X} \in \mathcal{S}^{n_1},\ \mathsf{Z}_i \in \mathbb{R}^{n_2\times n_1},\ i \in [1\colon S],\ \mathsf{Y}_{i,j} \in \mathbb{R}^{n_2\times n_2},\ (i,j) \in [1\colon S]^2$ and define

$$\begin{aligned} \mathcal{F} &\coloneqq \left\{\bar{\mathsf{X}} \in \mathcal{CPP}\left(\times_{i=0}^S \mathcal{K}_i\right) : f_i(\mathsf{X}, \mathsf{Y}_i, \mathsf{Z}_i) = \mathbf{r}_i,\ i \in [1\colon S]\right\}, \\ \mathcal{R}_{ecc} &\coloneqq \left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \mathsf{Y}_{ij} = \mathsf{O}\ \forall (i,j) \in [1\colon S]\colon i \neq j\right\}^\perp, \\ \mathcal{R}(\mathcal{F}) &\coloneqq \left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_{i,i} \end{pmatrix} \in \mathcal{CPP}(\mathcal{K}_0 \times \mathcal{K}_i),\ f_i(\mathsf{X}, \mathsf{Y}_{i,i}, \mathsf{Z}_i) = \mathbf{r}_i,\ i \in [1\colon S]\right\}. \end{aligned}$$

We need the following technical lemma:

Lemma 24. *If $\mathcal{F}$ and the sets*

$$\mathcal{R}_i \coloneqq \left\{\begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_{i,i} \end{pmatrix} \in \mathcal{CPP}(\mathcal{K}_0 \times \mathcal{K}_i),\ f_i(\mathsf{X}, \mathsf{Y}_{i,i}, \mathsf{Z}_i) = \mathbf{r}_i\right\},\ i \in [1\colon S],$$

are compact, then $\mathcal{R}(\mathcal{F})$ and $\mathcal{F} + \mathcal{R}_{ecc}$ are both closed.

Proof. By [73, Corollary 9.1.1.] the Minkowski sum of a compact, convex set and a closed, convex cone are closed since they do not have a common direction of recession. Thus, since $\mathcal{F}$ is compact, convex and $\mathcal{R}_{ecc}$ is a subspace and therefore a closed, convex cone, the set $\mathcal{F} + \mathcal{R}_{ecc}$ is closed. Note that the sets

$$\begin{aligned} &\left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_{i,i} \end{pmatrix} \in \mathcal{R}_i\right\} \\ =&\left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_{i,i} \end{pmatrix} \in \mathcal{R}_i,\ \text{all other entries are zero}\right\} \\ &+ \left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_{i,i} \end{pmatrix} = \mathsf{O}\right\},\ i \in [1\colon S], \end{aligned}$$

and are thus closed by the same argument. We have

$$\mathcal{R}(\mathcal{F}) \coloneqq \bigcap_{i=1}^S \left\{\bar{\mathsf{X}} \in \mathcal{S}^{n_1+Sn_2} \colon \begin{pmatrix} \mathsf{X} & \mathsf{Z}_i^\mathsf{T} \\ \mathsf{Z}_i & \mathsf{Y}_i \end{pmatrix} \in \mathcal{R}_i\right\},$$

so that $\mathcal{R}(\mathcal{F})$ are closed as well. $\square$

The following theorem establishes the connection between the matrix completion and reduced conic reformulations.

Theorem 25. *Assume the conditions in Lemma 24. Then the following statements are equivalent*

1. $\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{R}(\mathcal{F})\right\}=\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{F}\right\}$ *if* $\mathsf{C}\in\mathcal{R}_{ecc}^{\perp}$ *and* $-\infty$ *otherwise.*
2. $\mathcal{R}(\mathcal{F})=\mathcal{F}+\mathcal{R}_{ecc}$
3. *If*
$$\mathsf{M}_i:=\begin{pmatrix}\mathsf{X} & \mathsf{Z}_i^{\mathsf{T}}\\ \mathsf{Z}_i & \mathsf{Y}_i\end{pmatrix}\in\mathcal{CPP}(\mathcal{K}_0\times\mathcal{K}_i),\ f_i(\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i)=\mathbf{r}_i,\ i\in[1\colon S],$$
then $\Gamma_*\left([\mathsf{M}_i]_{i\in[1:S]}\right)$ *has a completion* $\mathsf{M}\in\mathcal{CPP}(\times_{i=0}^{S}\mathcal{K}_i)$.

Proof. We always have we have $\mathcal{R}(\mathcal{F})\supseteq\mathcal{F}+\mathcal{R}_{ecc}$ by the definition of set-completely positive matrix cones. Also

$$\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{F}+\mathcal{R}_{ecc}\right\}=\begin{cases}\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{F}\right\} & \text{if } \mathsf{C}\in\mathcal{R}_{ecc}^{\perp}\\ -\infty & \text{otherwise.}\end{cases}\tag{4.13}$$

1. $\Rightarrow$ 2. Assume 1., then by (4.13) clearly $\mathcal{R}(\mathcal{F})$ and $\mathcal{F}+\mathcal{R}_{ecc}$ have the same supporting hyperplanes which implies that the sets are identical as closed, convex sets are entirely described by their supporting hyperplanes.

2. $\Rightarrow$ 3. Assume 2., then

$$\bar{\mathsf{X}}\in\mathcal{S}^{n_1+Sn_2}\colon \begin{pmatrix}\mathsf{X} & \mathsf{Z}_i^{\mathsf{T}}\\ \mathsf{Z}_i & \mathsf{Y}_i\end{pmatrix}\in\mathcal{CPP}(\mathcal{K}_0\times\mathcal{K}_i),\ f_i(\mathsf{X},\mathsf{Y}_i,\mathsf{Z}_i)=\mathbf{r}_i,\ i\in[1\colon S],$$

implies that $\bar{\mathsf{X}}\in\mathcal{R}(\mathcal{F})$ and since $\mathcal{F}=\mathcal{R}(\mathcal{F})+\mathcal{R}_{ecc}$ the implications follows.

3. $\Rightarrow$ 1. We always have $\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{R}(\mathcal{F})\right\}\leq\min\left\{\mathsf{C}\bullet\bar{\mathsf{X}}\colon \bar{\mathsf{X}}\in\mathcal{F}\right\}$ since $\mathcal{R}(\mathcal{F})\supseteq\mathcal{F}+\mathcal{R}_{ecc}$. Also, if $\mathsf{C}\notin\mathcal{R}_{ecc}$ then the coefficients of $\mathsf{Y}_{ij},\ (i,j)\in[1\colon S]$ in the objective are nonzero and they are not restricted so that the objective is unbounded on $\mathcal{R}(\mathcal{F})$. Now, assume 3. and $\mathsf{C}\in\mathcal{R}_{ecc}^{\perp}$. Then for $\bar{\mathsf{X}}\in\mathcal{R}(\mathcal{F})$ we can replace the $\mathsf{Y}_{i,j},\ i\neq j$ coordinates to obtain an $\bar{\mathsf{X}}\in\mathcal{F}$ which yields the same objective since the coefficients of $\mathsf{Y}_{i,j},\ i\neq j$ are zero in the objective.

□

The above theorem lets us derive sufficient conditions for an arrowhead matrix to be completable to a set-completely positive matrix. Notice that Theorem 17 closes the relaxation gap between (4.1) and (4.2) for the case $n_2 = m_i = 1,\ i \in [1\!:\!S]$, since in that case we can always rewrite the terms in the objective function as needed. Thus, Theorem 25 implies a novel matrix completion result under the conditions of Theorem 17. The details are given in Theorem 26.

Theorem 26. *Assume that for matrices*

$$\mathsf{M}_i := \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & y_i \\ \mathbf{x} & \mathsf{X} & \mathbf{z}_i \\ y_i & \mathbf{z}_i^\mathsf{T} & Y_i \end{pmatrix} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathbb{R}_+) \quad i \in [1\!:\!S],$$

we have $\mathbf{f}_i^\mathsf{T}\mathbf{x} + g_i^\mathsf{T} y_i = 1$ *and* $\mathbf{f}_i\mathbf{f}_i^\mathsf{T} \bullet \mathsf{X} + 2\ g_i\mathbf{f}^\mathsf{T}\mathbf{z}_i + g^2 Y_i = 1$ *with* $\mathbf{f}_i = \beta_i\mathbf{f},\ \mathbf{f} \in \mathbb{R}^{n_1},\ \beta_i \in \mathbb{R},\ \in [1\!:\!S]$ *and that the sets* $\mathcal{F}_i := \left\{(\mathbf{x}^\mathsf{T}, y_i)^\mathsf{T} \in \mathcal{K}_0 \times \mathcal{K}_i \colon \mathbf{f}_i^\mathsf{T}\mathbf{x} + g_i y_i = 1\right\}, i \in [1\!:\!S]$ *are bounded sets. Then* $\Gamma_*\left([\mathsf{M}_i]_{i\in[1:S]}\right)$ *has a completion* $\mathsf{M} \in \mathcal{CPP}\left(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathbb{R}_+^S\right)$.

Proof. Consider the optimization problems

$$\begin{aligned} \min_{\mathbf{x}, y_1,\dots,y_S}\ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^S \mathbf{x}^\mathsf{T}\mathbf{b}_i y_i + C_i y_i^2 + c_i y_i \\ \text{s.t. : } & \mathbf{f}_i^\mathsf{T}\mathbf{x} + g_i y_i = 1, \qquad i \in [1\!:\!S], \\ & \mathbf{x} \in \mathcal{K}_0, \\ & \mathbf{y}_i \in \mathbb{R}_+, \quad i \in [1\!:\!S], \end{aligned} \tag{4.14}$$

and

$$\begin{aligned} \min_{\mathsf{X}, Y_i, \mathbf{z}_i. y_i}\ & \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{i=1}^S \mathbf{b}_i^\mathsf{T}\mathbf{z}_i + C_i Y_i + c_i y_i \\ \text{s.t. : } & \mathbf{f}_i^\mathsf{T}\mathbf{x} + g_i y_i = 1, \qquad i \in [1\!:\!S], \\ & \mathbf{f}_i\mathbf{f}_i^\mathsf{T} \bullet \mathsf{X} + 2\ g_i\mathbf{f}_i^\mathsf{T}\mathbf{z}_i + g_i^2 Y_i = 1, \qquad i \in [1\!:\!S], \\ & \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & y_i \\ \mathbf{x} & \mathsf{X} & \mathbf{z}_i \\ y_i & \mathbf{z}_i^\mathsf{T} & Y_i \end{pmatrix} \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}_0 \times \mathbb{R}_+), \quad i \in [1\!:\!S], \end{aligned} \tag{4.15}$$

and

$$
\begin{aligned}
\min_{\mathsf{X},Y_i,\mathbf{z}_i.y_i} \; & \mathsf{A}\bullet\mathsf{X}+\mathbf{a}^\mathsf{T}\mathbf{x}+\sum_{i=1}^{S}\mathbf{b}_i^\mathsf{T}\mathbf{z}_i+C_iY_i+c_iy_i\\
\text{s.t. :} \quad & \mathbf{f}_i^\mathsf{T}\mathbf{x}+g_iy_i=1, \quad i\in[1\!:\!S],\\
& \mathbf{f}_i\mathbf{f}_i^\mathsf{T}\bullet\mathsf{X}+2\,g_i\mathbf{f}_i^\mathsf{T}\mathbf{z}_i+g_i^2Y_{i,i}=1, \quad i\in[1\!:\!S],\\
& \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & y_1 & \dots & y_S\\ \mathbf{x} & \mathsf{X} & \mathbf{z}_1 & \dots & \mathbf{z}_S\\ y_1 & \mathbf{z}_1^\mathsf{T} & Y_{1,1} & \dots & Y_{1,S}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ y_S & \mathbf{z}_S^\mathsf{T} & Y_{S,1} & \dots & Y_{S,S}\end{pmatrix}\in\mathcal{CPP}\left(\mathbb{R}_+\times\mathcal{K}_0\times\mathbb{R}_+^S\right).
\end{aligned}
\tag{4.16}
$$

Since $\mathcal{F}_i$ are bounded, Theorem 3 establishes val (4.16) = val (4.14) and the statement would remain true if linear terms in $Y_{i,j},\ i\neq j$ would be introduced. Also, if we denote the feasible set of (4.14) by $\mathcal{F}_0$, then the feasible set of (4.16) is exactly $\mathcal{G}(\mathcal{F}_0)$ and therefore compact.

Further, Theorem 17 establishes val (4.15) = val (4.14) since the boundedness of $\mathcal{F}_i$ implies $g_i>0$ and we can always write $\mathbf{b}_i=g_i\left(\mathbf{b}_i/g_i\right)$ and $c_i=g_i\left(c_i/g_i\right)$ so that the conditions on $\mathbf{b}_i,c_i,\ i\in[1:S]$ in Theorem 17 are always fulfilled. However, if linear terms in $Y_{i,j}, i\neq j$ were introduced to the objective, the problem would become unbounded since there is no restriction on these variables in (4.15). Thus, statement 1. in Theorem 25 holds between problems (4.15) and (4.16).

Therefore, our claim will follow if we can show that the conditions of Lemma 24 hold where in our case $\mathcal{F}$ is the feasible set of (4.16) and

$$
\mathcal{R}_i=\left\{\begin{pmatrix}1 & \mathbf{x}^\mathsf{T} & y_i\\ \mathbf{x} & \mathsf{X} & \mathbf{z}_i\\ y_i & \mathbf{z}_i^\mathsf{T} & Y_i\end{pmatrix}\in\mathcal{CPP}(\mathbb{R}_+\times\mathcal{K}_0\times\mathbb{R}_+)\colon \begin{aligned}\mathbf{f}_i^\mathsf{T}\mathbf{x}+g_iy_i &=1\\ \mathbf{f}_i\mathbf{f}_i^\mathsf{T}\bullet\mathsf{X}+2\,g_i\mathbf{f}_i^\mathsf{T}\mathbf{z}_i+g_i^2Y_i &=1,\end{aligned}\right\}, i\in[1\!:\!S].
$$

The former is just $\mathcal{G}(\mathcal{F}_0)$ where $\mathcal{F}_0$ is just the feasible set of (4.1) which under our assumptions in bounded since $\mathcal{F}_i$ are bounded so that $\mathcal{G}(\mathcal{F})$ is compact. Further, $\mathcal{R}_i=\mathcal{G}(\mathcal{F}_0)$ also by Theorem 3 so that they are compact as well. $\square$

The completion theorem is somehow limited but to the best of our knowledge it is the only one of its kind. We hope that based on our geometrical approach we can reveal more such theorems in the future.

Part II

Applications to Decision Making under Uncertainty

Chapter 5

Introduction to decision making under uncertainty

The presence of uncertainty in decision problems has been a driving factor of research in optimization. Poor understanding and recognition of uncertainty in decision making problems may lead to bad performance of a decision made. Sources for uncertainty are numerous. Relevant data may not be available at the time of decision making and have to be estimated. Even if available, relevant data may be riddled with measurement errors, and decisions made may not be implementable exactly. The resulting challenges for optimization have been met under two major paradigms, namely *robust optimization* and *stochastic optimization*.

The latter is concerned with optimizing moments (mostly expected values) of functions that depend on random variables under a known probability law (see e.g. [51, 74]). This approach is appealing as it takes the maximum of what can be known about an uncertain process, i.e. its probability distribution, into account. Unfortunately, this comes at the cost of having to deal with mathematical optimization models that more often than not are intractable, so that their solutions have to be approximated with computationally expensive methods. Also, the probability distribution in question may be unknown so that a new source for uncertainty arises, which itself poses a major challenge. This challenge has in some sense been met by the framework of distributionally robust optimization (see [82] and references therein), but for the sake of brevity we will not go into further detail on in this topic here.

The alternative paradigm of dealing with uncertainty is that of robust optimization. Here, uncertain data is assumed to reside within an appropriately structured uncertainty set and one seeks to find the solution with the best worst-case performance among those solutions that are feasible for all possible realizations of the uncertain data (see e.g. [6, 47]). This approach requires more modest assumptions, as no probability law has to be assumed. In addition, the robust counterpart of an uncertain problem can often be reformulated as a convex optimization problem that can be solved in polynomial time. The major drawback of this approach is that the solutions obtained are often too conservative, courtesy of the pessimistic perspective inherent in robust optimization. Many attempts have been made to alleviate this shortcoming which lead to the development of concepts such as adjustable robustness (see [7]), max-regret robustness (see [55]) among others.

In this chapter we are going to review some basic concepts from robust and stochastic optimization. It is by no means a complete and exhaustive introduction but merely serves the purpose of clarifying basic terms and ideas in as much as they are needed for the understanding of the rest of this text.

5.1 Important concepts from robust optimization

In theory, there are many types of optimization problems that can be solved exactly and efficiently once the structure of the problem, including the relevant data, is known. However in practice the latter is often not the case and one is confronted with an *uncertain optimization problem*:

$$\inf_{\mathbf{x}\in\mathbb{R}^n} \{f_0(\mathbf{x},\mathbf{u}) : f_i(\mathbf{x},\mathbf{u}) \geq 0,\ i \in [1:m]\} \text{ where } \mathbf{u} \in \mathcal{U}. \tag{5.1}$$

The parameters of the functions $f_i,\ i \in [0:m]$ are uncertain and governed by the *uncertainty parameter* $\mathbf{u}$ that lives in an *uncertainty set* $\mathcal{U}$. This set encompasses all possible realizations of $\mathbf{u}$.

Robust optimization aims at making a decision under uncertainty by looking for the decision with the best worst-case performance among those decisions that will be feasible for all realizations of the uncertain data (see [6, 47] and references therein). Formally the so called *robust counterpart* of an uncertain optimization problem is

$$\inf_{\mathbf{x}\in\mathbb{R}^n} \left\{ \sup_{\mathbf{u}\in\mathcal{U}} \{f_0(\mathbf{x},\mathbf{u})\} : f_i(\mathbf{x},\mathbf{u}) \geq 0,\ i \in [1:m],\ \forall \mathbf{u} \in \mathcal{U} \right\}. \tag{5.2}$$

In the rest of the text we will be mainly concerned with cases where $f_i,\ i \in [0:m]$ are linear or quadratic functions in $\mathbf{x}$. For the sake of this introduction we stick to the linear case and consider the quadratic case in as much as $\mathbf{x} \in \{0,1\}^n$ can be considered a quadratic constraint. The case where f_i are more general quadratic functions will be studied in greater detail in Chapter 7.

5.1.1 Uncertain linear and mixed integer linear optimization problems

The *uncertain (mixed integer) linear optimization problem* is given by

$$\inf_{\mathbf{x}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \colon\; \mathbf{x} \in \mathcal{X}_0,\; \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{x} \leq b_i(\mathbf{u}),\, i \in [1\colon m] \right\} \quad \text{where } \mathbf{u} \in \mathcal{U}\,, \tag{5.3}$$

where the problem data are affine functions of $\mathbf{u}$, i.e. $\mathbf{c}(\mathbf{u}) \coloneqq \mathbf{c}_0 + \mathsf{C}\mathbf{u}$, $\mathbf{a}_i(\mathbf{u}) \coloneqq \mathbf{a}_i^0 + \mathsf{A}_i\mathbf{u}$ and $b_i(\mathbf{u}) \coloneqq b_i^0 + \mathbf{b}_i^\mathsf{T}\mathbf{u}$, $i \in [1\colon m]$ and $\mathcal{X}_0 \in \left\{ \mathbb{R}^n_+,\; \mathbb{R}^{n_1}_+ \times \{0,1\}^{n_2} \right\}$. Formulation (5.3) describes a family of problems parametrized by $\mathbf{u}$, the *uncertainty parameter*, which is not under the control of the decision maker and only known to lie in an appropriately structured *uncertainty set* $\mathcal{U} \subseteq \mathbb{R}^q$. Throughout most of the text, unless specified otherwise, we will model uncertainty sets as compact, convex conic intersections, i.e. $\mathcal{U} \coloneqq \left\{ \mathbf{u} \colon (1, \mathbf{u})^\mathsf{T} \in \mathcal{K} \right\}$ where $\mathcal{K}$ is an appropriate, closed, convex, pointed cone with nonempty interior. The problem, where the uncertainty is removed, i.e. $\inf_{\mathbf{x}} \left\{ \mathbf{c}_0^\mathsf{T}\mathbf{x} \colon\; \mathbf{x} \in \mathcal{X}_0,\; (\mathbf{a}_i^0)^\mathsf{T}\mathbf{x} \leq b_i^0,\, i \in [1\colon m] \right\}$, is referred to as the *nominal problem*. We will now briefly discuss some important concepts relating to (5.3).

5.1.1.1 Row-wise and column-wise uncertainty

Two important special cases of (5.3) are given by *row wise* uncertain problems and *column wise* uncertain problems. The former is given by

$$\inf_{\mathbf{x}} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x} \colon\; \mathbf{x} \in \mathcal{X}_0,\; \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}\mathbf{x} \leq b_i(\mathbf{u}_i),\, i \in [1\colon m] \right\} \quad \text{where } \mathbf{u}_i \in \mathcal{U}_i,\, i \in [0\colon m]\,,$$

hence, every row of the problem depends on an uncertainty parameter that is in no way correlated with the uncertainty parameters of any other row. Likewise, for the latter case we have

$$\inf_{\mathbf{x}} \left\{ \sum_{i=1}^{n} c_i(\mathbf{u}_i)x_i \colon\; \mathbf{x} \in \mathcal{X}_0,\; \sum_{i=1}^{n} \mathbf{a}_i(\mathbf{u}_i)x_i \leq \mathbf{b}(\mathbf{u}_{n+1}) \right\} \quad \text{where } \mathbf{u}_i \in \mathcal{U}_i,\, i \in [1\colon n+1]\,,$$

so that every column of the uncertain data is governed by an independent uncertainty parameter. For further information about the relevance of these two types of uncertainty models see [78].

5.1.2 Robust counterparts

Since (5.3) describes a, possibly infinite, family of optimization problems, we need stronger criteria in order to make a decision under uncertainty. The most popular choice here is the worst-case criterion, which leads to the so called *robust counterpart*.

The goal is to find the solution with the best worst-case performance among those solutions that are feasible for all outcomes of the uncertainty. It is thus a pessimistic approach, which finds solutions with performance guarantees that are met in any case and may be exceeded in some cases. It is well known that the robust counterpart of an uncertain linear optimization problem can always be equivalently described as the robust counterpart of a row wise uncertain problem where $\mathcal{U}_i$ are the projections of $\mathcal{U}$ onto the coordinates of the uncertainty parameter present in the i-th constraint. The analogue holds for the objective function. This yields

$$\inf_{\mathbf{x}} \left\{ \sup_{\mathbf{u}_0 \in \mathcal{U}_0} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\} : \ \mathbf{x} \in \mathcal{X}_0, \ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} \leq b_i(\mathbf{u}_i) \ \forall \mathbf{u}_i \in \mathcal{U}_i, \ \ i \in [1\!:\!m] \right\}. \tag{RC}$$

Note that

$$\mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} \leq b_i(\mathbf{u}_i) \, \forall \mathbf{u}_i \in \mathcal{U}_i \ \Leftrightarrow \ \sup_{\mathbf{u}_i \in \mathcal{U}_i} \left\{ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} - b_i(\mathbf{u}_i) \right\} \leq 0 \,,$$

and the objective inside the supremum is linear in $\mathbf{x}$ so that the supremum expression is actually convex in $\mathbf{x}$, rendering (RC) a (mixed integer) convex problem. The standard procedure of deriving a finite reformulation of (RC), the so called *deterministic counterpart*, is by invoking convex duality, in our case,

$$\sup_{\mathbf{u}} \left\{ \mathbf{d}^\mathsf{T} \mathbf{u} \colon \ \left(1, \mathbf{u}^\mathsf{T}\right)^\mathsf{T} \in \mathcal{K} \right\} = \inf_{\rho} \left\{ \rho \colon \ \left(\rho, -\mathbf{d}^\mathsf{T}\right)^\mathsf{T} \in \mathcal{K}^* \right\}.$$

Note that our assumption on $\mathcal{K}$ guarantees the existence of a primal Slater point and thus full strong duality. It is therefore possible to replace the inner suprema:

$$\sup_{\mathbf{u} \in \mathcal{U}} \left\{ \mathbf{a}(\mathbf{u})^\mathsf{T} \mathbf{x} - b(\mathbf{u}) \right\} \leq 0 \ \Leftrightarrow \ \mathbf{a}_0^\mathsf{T} \mathbf{x} - b^0 + \rho \leq 0, \ \left(\rho, \mathbf{b}^\mathsf{T} - \mathbf{x}^\mathsf{T} \mathsf{A}\right)^\mathsf{T} \in \mathcal{K}^*.$$

The supremum in the objective can be treated analogously so that we arrive at the deterministic counterpart given by

$$\begin{aligned}
\min_{\mathbf{x}, \rho_i} \mathbf{c}_0^\mathsf{T} \mathbf{x} + \rho_0 & \\
\text{s.t.} : \rho_i \leq b_0^i - (\mathbf{a}_0^i)^\mathsf{T} \mathbf{x}, & \quad i \in [1\!:\!m], \\
\left(\rho_i, \mathbf{b}_i - \mathsf{A}_i^\mathsf{T} \mathbf{x}\right)^\mathsf{T} \in \mathcal{K}_i^*, & \quad i \in [1\!:\!m], \\
\left(\rho_0, -\mathsf{C}^\mathsf{T} \mathbf{x}\right) \in \mathcal{K}_0^*, & \\
\mathbf{x} \in \mathcal{X}_0. &
\end{aligned} \tag{5.4}$$

5.1.3 Optimistic counterparts

Taking an optimistic approach one arrives at the *optimistic counterpart*, where the decision maker can choose the value of the uncertainty parameter. The respective optimization problem is given by

$$\inf_{\mathbf{x},\mathbf{u}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \colon\; \mathbf{x} \in \mathcal{X}_0,\; \mathbf{u} \in \mathcal{U},\; \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{x} \le b_i(\mathbf{u}),\, i \in [1\!:\!m] \right\}. \tag{5.5}$$

Though seemingly of little practical relevance, (5.5) enjoys some useful theoretical aspects stemming from robust duality theory, which we discuss in more detail in Section 8.2.1.1. Note that (5.5) can generally not be written as the optimistic counterpart of a row wise or column wise uncertain problem. Also, (RC) is tractable if the nominal problem and the uncertainty sets are tractable (see [6]). The same does generally not hold for (5.5) so that it may be intractable even if the uncertainty sets and the nominal problem are well behaved.

5.1.4 Regret robustness

We can think of situations where the ex-ante choice a decision maker would have made in hindsight heavily depends on the outcome of the uncertainty process. In that situation, decision makers run the risk of regretting their decision after the fact. This line of thought leads us the concept of *regret robustness* where the aim is to make a decision so that the maximum regret possible across all realizations of the uncertain process is minimized. The min-max-regret approach has garnered much attention as an alternative concept to the min-max approach in robust optimization and is well studied in the robust optimization literature (see [55]). The *regret-robust counterpart* of an uncertain linear optimization problem is given by

$$\min_{\mathbf{x}\in\mathcal{X}} \sup_{\mathbf{u}\in\mathcal{U}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}\in\mathbb{R}^n} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{y} \le b_i(\mathbf{u}),\, \in [1\!:\!m] \right\} \right\},$$

where $\mathcal{X}$ is a suitably chosen feasible set. The innermost optimization problem represents the decision made in hindsight given a realization of $\mathbf{u}$ and we seek a decision on $\mathbf{x}$ so that the maximum excess performance of the hindsight decision is minimized. The supremum and the inner infimum can actually combined into an inner supremum which then becomes a bilinear optimization problem. These problems form an intractable class of problems in and of itself and the difficulties only increase if one is to minimize such a supremum. We will not go into further detail on how to tackle these problems here. Some techniques will be discussed when we apply the concept of regret in Chapter 8.

5.1.5 Endogenous uncertainty

A recent line of literature introduced the notion of endogenous uncertainty where the decision maker has control over an additional set of (often binary) variables, which govern the structure of the uncertainty set. The endogenous uncertain optimization problem is given by

$$\inf_{\mathbf{x}} \{f_0(\mathbf{x}, \mathbf{u}) : \ \mathbf{x} \in \mathcal{X}_0, \ f_i(\mathbf{x}, \mathbf{u}) \leq 0, \, i \in [1:m]\} \quad \text{where } \mathbf{u} \in \mathcal{U}^s \text{, and } s \in [1:S].$$

The uncertainty sets are modeled as set-valued function $\mathcal{U}^s$ that depend on an additional parameter s that is controlled by the decision maker. This approach allows the modelling of situations where the uncertainty can be influenced by the decision maker by making an investment in e.g. market research, more precise machines etc. Such an investment, in a robust optimization framework, is only acceptable if the expenses do not exceed the gains in worst case pay-off under the new uncertainty set. The arising decision problem was recently addressed in [56, 62], where the authors studied reformulations of the robust counterpart of endogenously uncertain linear optimization problems that are amenable to mixed integer linear programming (MILP) and conic mixed integer linear programming (C-MILP) techniques. Also in [56] it is proved that the class of robust optimization problems under endogenous uncertainty is NP-hard. We will go into more detail on these subjects in Section 8.1.2.

5.1.6 Adjustable robust counterparts

Another way to remedy the problem of overconservatism of robust solutions is known as adjustable robust optimization (ARO). In this setting, the decision variables are grouped into two categories. The *first stage decision variables* represent decisions to be made "here and now", i.e. at a point in time when there is uncertainty in some of the relevant problem data. The *second stage decision variables* represent decisions that can be delayed until uncertainty is removed. Of course, one could in the first stage make a decision on all of the variables and require the solution to be feasible for any realization of the uncertainty parameter. But then potential flexibility is not harnessed since this is equivalent to the classical robust approach and most likely the solution will be unnecessarily conservative. The idea is that, rather than requiring that the decision on all variables is feasible in any case, we merely require our first stage decision to allow for a second stage adjustment that renders the overall solution to be feasible in any case. Said differently, we look for the best among those first stage decisions on $\mathbf{x} \in \mathbb{R}^{n_1}$ for which there exists a function $\mathbf{y} \colon \mathcal{U} \mapsto \mathbb{R}^{n_2}$ that, given any realization of the uncertainty parameter $\mathbf{u}$, maps to a vector for the second stage decision that in total gives a feasible solution to the optimization problem. ARO was introduced in [7] and has received much attention. For a detailed survey see [88]. Generically ARO can be written as

$$\inf_{\mathbf{x}\in\mathbb{R}^{n_1},\mathbf{y}(\mathbf{u})} \left\{ \sup_{\mathbf{u}\in\mathcal{U}} \{f_0(\mathbf{x},\mathbf{y}(\mathbf{u}),\mathbf{u})\} : f_i(\mathbf{x},\mathbf{y}(\mathbf{u}),\mathbf{u}) \geq 0 \ \forall \mathbf{u}\in\mathcal{U},\ i\in[1:m] \right\}. \tag{5.6}$$

The first stage decision $\mathbf{x}\in\mathbb{R}^n$ is again vector valued, but the second stage variable $\mathbf{y}(\mathbf{u})$ is allowed to adapt to the uncertainty and is thus a function of $\mathbf{u}$. Since the space of all functions is intractable, so is (5.6), and thus it is much harder to solve in practice than (5.2). However, there are many powerful approaches to (approximately) solve it (see [88] and references therein).

5.1.6.1 Affine decision rules for adjustable robust mixed integer linear problems

We have seen in Section 5.1.2 that the robust counterpart of a mixed integer linear optimization problem can quite easily be reformulated as a convex optimization problem with finite number of constraints. The same is not true for the adjustable robust counterpart due to the intractability of the search space for the second stage variables. In order to recover tractability one can approximate the space of all functions by the tractable subspace of affine functions. Given an ajdustable robust counterpart of an uncertain mixed integer linear optimization problem:

$$\inf_{\mathbf{x},\mathbf{y}(\mathbf{u})} \left\{ \sup_{\mathbf{u}_0\in\mathcal{U}_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}(\mathbf{x};\mathbf{y}(\mathbf{u}))\right\} :\ \mathbf{x}\in\mathcal{X}_0,\ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}(\mathbf{x};\mathbf{y}(\mathbf{u})) \leq b_i(\mathbf{u}_i)\ \forall \mathbf{u}_i\in\mathcal{U}_i,\ \ i\in[1:m] \right\}. \tag{5.7}$$

one can specify $\mathbf{y}(\mathbf{u}) := \mathsf{Y}\mathbf{u}+\mathbf{y}_0$ to be an affine functions, the coefficients of which become the new decision variables for the adjustable robust counterpart under an affine decision rule. Note that for the left hand side of the constraints we have $\mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}(\mathbf{x};\mathbf{y}(\mathbf{u})) = (\mathsf{A}_i\mathbf{u}_i+\mathbf{a}_i)^\mathsf{T}(\mathbf{x};\mathsf{Y}\mathbf{u}+\mathbf{y}_0)$, which is quadratic in the uncertainty $\mathbf{u}$, likewise we have such terms in the objective function. Thus, the reformulation steps outlined in Section 5.1.2 can only be applied if we require that the coefficients of the second stage variables be independent of $\mathbf{u}$, in other words, we need to assume *fixed recourse*. Under this assumption the adjustable robust counterpart under affine decision rules exhibits the same structer as a robust counterpart and is thus amenable to reformulation based on convex duality. In Section 7.4 we will discuss quadratic decision rules and linear decision rules in case the assumtpion of fixed recourse is dropped.

5.2 Important concepts from Stochastic Optimization

Similar to robust optimization, in stochastic optimization one seeks to make a decision where the relevant data is uncertain. The key difference is that in stochastic optimization the random process generating the data is assumed to be known, hence we assume knowledge on the distribution of the unknown problem data. The aim is then that to find solutions optimizing the expected value of the objective function under the known distribution of the random data. The mathematical formulation of a stochastic optimization problem ist thus given by

$$\min_{\mathbf{x}\in\mathcal{X}} \{\mathbb{E}_{\boldsymbol{\xi}}\left(f(\mathbf{x},\boldsymbol{\xi})\right)\}. \tag{5.8}$$

Here $\mathbb{E}_{\boldsymbol{\xi}}$ indicates the expected value with respect to the random vector $\boldsymbol{\xi}$, which is defined by some probability space $(\Xi, \mathcal{A}, \mathcal{P})$ with support Ξ, probability distribution $\mathcal{P}$ and σ-field $\mathcal{A}$. Again $\mathcal{X} \subseteq \mathbb{R}^n$ is the feasible set of the decision vector $\mathbf{x}$.

5.2.1 Two-Stage Stochastic Optimization

In this text we will be mainly concerned with the-two stage version of this framework. Analogously to the adjustable robust setting one seeks a decision on the first stage variables and on a second stage decision rule that adapts to the realization of the random event. Thus, we are considering

$$\min_{\mathbf{x}\in\mathcal{X}} \left\{f_1(\mathbf{x}) + \mathbb{E}_{\boldsymbol{\xi}}\left(\min_{\mathbf{y}} \{f_2(\mathbf{x},\mathbf{y},\boldsymbol{\xi})\colon \mathbf{y}\in\mathcal{Y}(\mathbf{x},\boldsymbol{\xi})\}\right)\right\}. \tag{5.9}$$

Here we make a choice on the first stage variables $\mathbf{x}$ and second stage policies $\mathbf{y}$ so that we optimize the sum of a deterministic first stage outcome and the expected value of the optimal second stage choice. Problem (5.9) is difficult to solve in general and we will not go into great detail on solution methods here. We will instead focus on the most important concepts that are relevant to the rest of the text.

The function

$$r(\mathbf{x}) := \mathbb{E}_{\boldsymbol{\xi}}\left(\min_{\mathbf{y}} \{f_2(\mathbf{x},\mathbf{y},\boldsymbol{\xi})\colon \mathbf{y}\in\mathcal{Y}(\mathbf{x},\boldsymbol{\xi})\}\right), \tag{5.10}$$

is called the *recourse function*. The solution of $\mathbf{x}^*$ to (5.9) is called the *here and now solution*. A simple lower bound on (5.9) is given by the so called *wait and see approach*

$$\mathbb{E}_{\boldsymbol{\xi}}\left(\min_{\mathbf{x},\mathbf{y}} \{f_1(\mathbf{x}) + f_2(\mathbf{x},\mathbf{y})\colon \mathbf{x}\in\mathcal{X},\ \mathbf{y}\in\mathcal{Y}(\mathbf{x},\boldsymbol{\xi})\}\right), \tag{5.11}$$

which is the expected value of the uncertain problem if we had perfect information, i.e. what the decision maker would be willing to pay in order to obtain perfect information of the future. The difference between (5.9) and (5.11) is called the *value of perfect information*. Another simple approach is the so called *expected value problem* where the uncertain data is replaced by its expected value:

$$\min_{\mathbf{x},\mathbf{y}} \left\{ f_1(\mathbf{x}) + f_2(\mathbf{x}, \mathbf{y}, \mathbb{E}(\boldsymbol{\xi})) \right\}, \tag{5.12}$$

which gives an upper bound on (5.9). The respective optimal first stage solutions $\bar{\mathbf{x}}$ is called the *expected value solution* and the difference of $f_1(\bar{\mathbf{x}}) + r(\bar{\mathbf{x}})$ and (5.9) is called the *value of the stochastic solution*. Finally, bounds can be obtained by discretizing the probability measure of $\boldsymbol{\xi}$, in order to obtain a discrete approximation of the true measure with possible outcomes $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_S\}$ and associated probabilities $\{p_1, \dots, p_S\}$ so that we can construct the so called *scenario problem* given by

$$\min_{\mathbf{x},\mathbf{y}_i} \left\{ f_1(\mathbf{x}) + \sum_{i=1}^{S} p_i f_2(\mathbf{x}, \mathbf{y}_i, \boldsymbol{\xi}_i) \colon \mathbf{x} \in \mathcal{X},\ \mathbf{y}_i \in \mathcal{Y}(\mathbf{x}, \boldsymbol{\xi}_i),\ i \in [1 \colon S] \right\}. \tag{5.13}$$

The scenario problem (5.13) is an approximation of the original problem, the properties of which depend on the problem structure and the way the scenarios are constructed. A specific construction is described in Section 6.2.3. We omit further details here for the sake of brevity.

Chapter 6

Two-Stage Stochastic Standard Quadratic Optimization

6.1 Introduction

Consider the following non-convex *deterministic standard quadratic optimization problem*, which is referred to as DStQP:

$$\text{(DStQP)} \qquad z^*_{\text{det}} := \min \left\{ q(\mathbf{z}) := \mathbf{z}^\mathsf{T} \mathsf{Q} \mathbf{z} : \mathbf{z} \in \Delta^n \right\},$$

where Q is an $(n \times n)$ symmetric matrix, and $\Delta^n = \left\{ \mathbf{z} \in \mathbb{R}^n_+ : \mathbf{e}^\mathsf{T} \mathbf{z} = 1 \right\}$ with $\mathbf{e} = [1, \ldots, 1]^\mathsf{T} \in \mathbb{R}^n$.[1] Note that a linear term $2\mathbf{q}^\mathsf{T}\mathbf{z}$ in the objective $q(\mathbf{z})$ with $\mathbf{q} \in \mathbb{R}^n$ can be incorporated easily in homogenous form by replacing Q with the symmetric matrix $\mathsf{Q} + \mathbf{e}\mathbf{q}^\mathsf{T} + \mathbf{q}\mathbf{e}^\mathsf{T}$.

While the constraints in (DStQP) are out of dispute as probabilities or portfolio shares always are non-negative and sum up to one, the only variable problem data are to be found a the matrix $\tilde{\mathsf{Q}}$. Now suppose a (possibly) small $n_1 \times n_1$ principal submatrix A of $\tilde{\mathsf{Q}}$ is known (more or less) exactly whereas the rest of the problem data are subject to uncertainty with known probability distribution:

$$\tilde{\mathsf{Q}} = \begin{bmatrix} \mathsf{A} & \tilde{\mathsf{B}}^\mathsf{T} \\ \tilde{\mathsf{B}} & \tilde{\mathsf{C}} \end{bmatrix}.$$

Here $\boldsymbol{\xi} := [\tilde{\mathsf{B}}, \tilde{\mathsf{C}}]$ are the uncertain data. We denote with $\mathsf{Q} = \mathbb{E}(\tilde{\mathsf{Q}})$.

Decomposing $\mathbf{z}^\mathsf{T} = [\mathbf{x}^\mathsf{T}, \mathbf{y}^\mathsf{T}(\boldsymbol{\xi})]$ with $(\mathbf{x}, \mathbf{y}(\boldsymbol{\xi})) \in \mathbb{R}^{n_1}_+ \times \mathbb{R}^{n_2}_+$ we arrive at the following problem reformulation under data uncertainty:

$$q(\mathbf{x}, \mathbf{y}(\boldsymbol{\xi})) = \mathbf{x}^\mathsf{T} \mathsf{A} \mathbf{x} + 2\mathbf{x}^\mathsf{T} \tilde{\mathsf{B}}^\mathsf{T} \mathbf{y}(\boldsymbol{\xi}) + \mathbf{y}(\boldsymbol{\xi})^\mathsf{T} \tilde{\mathsf{C}} \mathbf{y}(\boldsymbol{\xi}) .$$

[1] Notice that we use the symbol $\mathbf{e}$ for the vector of ones in varying dimensions. The exact dimension should be clear from the context.

Taking the expectation with respect to the probability distribution of $\boldsymbol{\xi}$, we have that the recourse function is r

$$r(\mathbf{x}) \quad = \quad \mathbb{E}_{\boldsymbol{\xi}} \left[\begin{array}{l} \min_{\mathbf{y}(\boldsymbol{\xi})} 2\mathbf{x}^{\mathsf{T}}\tilde{\mathsf{B}}^{\mathsf{T}}\mathbf{y}(\boldsymbol{\xi}) + \mathbf{y}(\boldsymbol{\xi})^{\mathsf{T}}\tilde{\mathsf{C}}\mathbf{y}(\boldsymbol{\xi}) \\ \\ \text{s.t. } \mathbf{e}^{\mathsf{T}}\mathbf{y}(\boldsymbol{\xi}) = 1 - \mathbf{e}^{\mathsf{T}}\mathbf{x}, \\ \\ \quad \mathbf{y}(\boldsymbol{\xi}) \geq \mathbf{o} \end{array} \right]$$

and the two-stage Stochastic Quadratic Program (SQP) can be formulated as follows:

$$z^*_{\text{stoch}} := \min_{\mathbf{x}\in T^{n_1}} \left\{ s(\mathbf{x}) := \mathbf{x}^{\mathsf{T}}\mathsf{A}\mathbf{x} + r(\mathbf{x}) \right\}.$$

with $T^{n_1} = \operatorname{conv}\{\mathbf{o}, \mathbf{e}^i : i \in [1:n_1]\} = \operatorname{conv}(\Delta^{n_1} \cup \{\mathbf{o}\})$.

6.2 Bounds on the scenario problem

In stochastic optimization it is customary to approximate the stochastic optimization problem by discretizing the underlying random process $\boldsymbol{\xi} \in \Xi$ and subsequently solve the so called scenario problem, i.e. the stochastic optimization problem where the true distribution is replaced by its discretization. There is substantial literature on how to to construct these discretizations, see e.g. [74]. In this section we will discuss upper and lower bounds on the scenario problem based on stochastic and quadratic optimization theory. But first let us derive the actual form of the scenario problem in case of (SQP).

Suppose now that the distribution of $\boldsymbol{\xi}$ has finite support. This means that $\boldsymbol{\xi}$ has a finite number of possible realizations, called scenarios $\xi_s = (\tilde{\mathsf{B}}_s, \tilde{\mathsf{C}}_s)$ with respective positive probabilities p_s, $s \in [1\!:\!S]$, i.e. $\Xi = \{\xi_1, \dots, \xi_S\}$.

For a given $\mathbf{x}$ the recourse function $r(\mathbf{x})$ is equal to the optimal value of the following program

$$\begin{array}{lll} \min\limits_{\mathbf{y}_1,\dots,\mathbf{y}_S} & \sum\limits_{s=1}^{S} p_s(2\mathbf{x}^{\mathsf{T}}\tilde{\mathsf{B}}_s^{\mathsf{T}}\mathbf{y}_s + \mathbf{y}_s^{\mathsf{T}}\tilde{\mathsf{C}}_s\mathbf{y}_s) & \\ & \mathbf{e}^{\mathsf{T}}\mathbf{x} + \mathbf{e}^{\mathsf{T}}\mathbf{y}_s = 1, & s \in [1\!:\!S], \\ & \mathbf{y}_s \geq \mathbf{o}, & s \in [1\!:\!S]. \end{array}$$

The whole two-stage stochastic program can be approximated by the following large scale deterministic equivalent quadratic problem

$$z^*_{\text{stoch}} := \min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S} \quad \mathbf{x}^{\mathsf{T}}\mathsf{A}\mathbf{x} + \sum_{s=1}^{S} p_s(2\mathbf{x}^{\mathsf{T}}\tilde{\mathsf{B}}_s^{\mathsf{T}}\mathbf{y}_s + \mathbf{y}_s^{\mathsf{T}}\tilde{\mathsf{C}}_s\mathbf{y}_s) \tag{6.1}$$

$$\begin{aligned}
\mathbf{e}^\mathsf{T}\mathbf{x} + \mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, &\quad s \in [1\!:\!S],\\
\mathbf{y}_s \geq \mathbf{o}, &\quad s \in [1\!:\!S],\\
\mathbf{x} \geq \mathbf{o}. &
\end{aligned}$$

Since (6.1) describes a class of non-convex QCQPs, which contains the StQP as a special case, it is NP-hard. We will now proceed in presenting some bounds, which can be computed with reasonable effort.

6.2.1 Generating upper bounds based on StQP algorithms

For the StQP, [22] introduced a powerful algorithm, called the Infection-Immunization-Dynamics (IID) Algorithm, whose individual steps are computationally highly efficient and which is guaranteed to converge to a locally optimal solution. Our aim in this section is to harness this powerful method for computing upper bounds for (6.1).

In order to bring the problem (6.1) into the form of a standard quadratic program we use the following approach. We add a redundant constraint

$$S\mathbf{e}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S} \mathbf{e}^\mathsf{T}\mathbf{y}_s = S,$$

which is just the sum of all constraints. After dividing by S and a change of variables $\mathbf{y}_s \to S\mathbf{y}_s$ the constraint together with the non-negativity constraints give a simplex constraint. The problem then reads as

$$\begin{aligned}
&\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S} \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \sum_{s=1}^{S} p_s \left(2S\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{B}}_s\mathbf{x} + S^2\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{C}}_s\mathbf{y}_s\right)\\
&\text{s.t.}: \mathbf{e}^\mathsf{T}\mathbf{x} + S\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \qquad s \in [1\!:\!S],\\
&\qquad \mathbf{e}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S}\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1,\\
&\qquad \mathbf{x} \geq \mathbf{o},\\
&\qquad \mathbf{y}_s \geq \mathbf{o}, \qquad s \in [1\!:\!S].
\end{aligned} \tag{6.2}$$

We now quadratically penalize the all constraints but the simplicial one and we get the following auxiliary problem (AUX):

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S}\ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \sum_{s=1}^{S} p_s \left(2S\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{B}}_s\mathbf{x} + S^2\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{C}}_s\mathbf{y}_s\right) + \lambda\sum_{s=1}^{S}\left(\mathbf{e}^\mathsf{T}\mathbf{x} + S\mathbf{e}^\mathsf{T}\mathbf{y}_s - 1\right)^2 \\
\text{s.t. : } & \mathbf{e}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S}\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \\
& \mathbf{x} \geq \mathbf{o}, \\
& \mathbf{y}_s \geq \mathbf{o}, \qquad s\in[1\!:\!S].
\end{aligned} \tag{AUX}$$

Define $\mathbf{z} = \left(\mathbf{x}^\mathsf{T}, \mathbf{y}_1^\mathsf{T}, \dots, \mathbf{y}_s^\mathsf{T}\right)^\mathsf{T} \in \mathbb{R}^{n_1+Sn_2}$ and an appropriate $S\times(n_1+Sn_2)$ matrix M, so that $\mathsf{M}\mathbf{z} = \mathbf{e}$ is equivalent to $\mathbf{e}^\mathsf{T}\mathbf{x} + S\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1,\ s\in[1\!:\!S]$. We can then write (AUX) as

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S}\ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \sum_{s=1}^{S} p_s \left(2S\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{B}}_s\mathbf{x} + S^2\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{C}}_s\mathbf{y}_s\right) + \lambda\|\mathsf{M}\mathbf{z}-\mathbf{e}\|^2 \\
\text{s.t. : } & \mathbf{e}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S}\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \\
& \mathbf{x} \geq \mathbf{o}, \\
& \mathbf{y}_s \geq \mathbf{o}, \qquad s\in[1\!:\!S],
\end{aligned} \tag{6.3}$$

and calculate $\|\mathsf{M}\mathbf{z}-\mathbf{e}\|^2 = \mathbf{z}^\mathsf{T}\mathsf{M}^\mathsf{T}\mathsf{M}\mathbf{z} - \mathbf{e}^\mathsf{T}\mathsf{M}\mathbf{z} - \mathbf{z}^\mathsf{T}\mathsf{M}\mathbf{e} + \mathbf{e}^\mathsf{T}\mathbf{e}$. Since $\mathbf{e}^\mathsf{T}\mathbf{z} = 1$ for every $\mathbf{z}$ that is feasible for (AUX) we can multiply $\mathbf{e}^\mathsf{T}\mathsf{M}\mathbf{z}$ and $\mathbf{z}^\mathsf{T}\mathsf{M}\mathbf{e}$ by $\mathbf{e}^\mathsf{T}\mathbf{z}$ and $\mathbf{e}^\mathsf{T}\mathbf{e} = S$ by $\mathbf{z}^\mathsf{T}\mathbf{e}\mathbf{e}^\mathsf{T}\mathbf{z}$ so that we can write (AUX) as

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S}\ & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \sum_{s=1}^{S} p_s \left(2S\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{B}}_s\mathbf{x} + S^2\mathbf{y}_s^\mathsf{T}\tilde{\mathsf{C}}_s\mathbf{y}_s\right) + \lambda\mathbf{z}^\mathsf{T}\bar{\mathsf{M}}\mathbf{z} \\
\text{s.t. : } & \mathbf{e}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S}\mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \\
& \mathbf{x} \geq \mathbf{o}, \\
& \mathbf{y}_s \geq \mathbf{o}, \qquad s\in[1\!:\!S],
\end{aligned}$$

where $\bar{\mathsf{M}} := \mathsf{M}^\mathsf{T}\mathsf{M} - \mathbf{e}\mathbf{e}^\mathsf{T}\mathsf{M} - \mathsf{M}\mathbf{e}\mathbf{e} + S\mathbf{e}\mathbf{e}^\mathsf{T}$. The problem is now cast as a (deterministic) standard quadratic problem, which we can attack using the IID algorithm by [22]. We also see that (AUX) approximates (6.2) better and better as $\lambda\to\infty$. Due to the nature of quadratic penalization, the approximation will never be exact. This is due to the fact that the convex quadratic penalization term flattens out near the feasible region, which it enforces. It is therefore most likely optimal to be slightly infeasible, as the penalization incurs a vanishing cost for such violations. We will address this matter further in Section 6.3.

6.2.2 A scalable copositive lower bound

The following theorem is an application of Theorem 17 for the scenario problem.

Theorem 27. *Consider the problem*

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{y}_1,\dots,\mathbf{y}_S} \; & \mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S} \mathbf{x}^\mathsf{T}\mathsf{B}_s\mathbf{y}_s + \mathbf{y}_s^\mathsf{T}\mathsf{C}\mathbf{y}_s + \mathbf{c}_s^\mathsf{T}\mathbf{y}_s \\
\text{s.t. :} \; & \mathbf{e}^\mathsf{T}\mathbf{x} + \mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \quad s \in [1\!:\!S], \\
& \mathbf{x}, \mathbf{y}_1, \dots, \mathbf{y}_S \geq \mathbf{o}.
\end{aligned} \tag{6.4}$$

The following conic optimization problem gives a lower bound and if $\mathbf{c}_s = \alpha_s\mathbf{e},\ \mathsf{B}_s = \mathbf{b}_s\mathbf{e}^\mathsf{T},\ \alpha_s \in \mathbb{R},\ \mathbf{b}_s \in \mathbb{R}^{n_1},\ i \in [1\!:\!S]$ *the bound is actually tight:*

$$\begin{aligned}
\min_{\mathsf{X},\mathsf{Y}_s,\mathsf{Z}_s.\mathbf{y}_s} \; & \mathsf{A} \bullet \mathsf{X} + \mathbf{a}^\mathsf{T}\mathbf{x} + \sum_{s=1}^{S} \mathsf{B}_s \bullet \mathsf{Z}_s + \mathsf{C}_s \bullet \mathsf{Y}_s + \mathbf{c}_s^\mathsf{T}\mathbf{y}_s \\
\text{s.t. :} \; & \mathbf{e}^\mathsf{T}\mathbf{x} + \mathbf{e}^\mathsf{T}\mathbf{y}_s = 1, \quad s \in [1\!:\!S], \\
& \mathsf{E} \bullet \mathsf{X} + 2\,\mathsf{E} \bullet \mathsf{Z}_s + \mathsf{E} \bullet \mathsf{Y}_s = 1, \quad s \in [1\!:\!S], \\
& \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} & \mathbf{y}_s^\mathsf{T} \\ \mathbf{x} & \mathsf{X} & \mathsf{Z}_s^\mathsf{T} \\ \mathbf{y}_s & \mathsf{Z}_s & \mathsf{Y}_s \end{pmatrix} \in \mathcal{CPP}(\mathbb{R}^{n_1+n_2+1}), \quad s \in [1\!:\!S].
\end{aligned} \tag{6.5}$$

Proof. This is a special case of Theorem 17, where all the sets $\mathcal{F}_s$ are identically given by $\left\{\left(\mathbf{x}^\mathsf{T}, \mathbf{y}^\mathsf{T}\right)^\mathsf{T} \in \mathbb{R}^n_+ : \mathbf{e}^\mathsf{T}\mathbf{x} + \mathbf{e}^\mathsf{T}\mathbf{y} = 1\right\}$, which are bounded since they describe the standard simplex. □

Note that (6.4) can be reformulated using Theorem 3 where the size of the completely positive matrix block would be $n_1 + Sn_2 + 1$ so that the number of variables would increase quadratically with S. On the other hand, in Theorem 27 we have S completely positive matrix blocks of order $n_1 + n_2 + 1$ so that the number of variables is linear in S. We can show in numerical examples that this yields substantial advantages in computation time and memory consumption.

6.2.3 Lower bounds based on dissecting the probability measure

In order to have a good description of the uncertain parameters, the number of scenarios S to be considered in z^*_{stoch} is typically very large, making the problem computationally intractable. Nevertheless, chains of lower bounds can be constructed by solving sets of group sub-problems less complex than the original one, and recalculating the probabilities of each scenario in the group accordingly (see [58]). We briefly recall now the construction.

Suppose again that the distribution of $\boldsymbol{\xi}$ has a finite support, i.e. $\boldsymbol{\xi}$ has a finite number of possible scenarios $\xi_s = (\tilde{\mathsf{B}}_s, \tilde{\mathsf{C}}_s)$ with respective positive probabilities p_s, $s = 1, \dots, S$, i.e. $\Xi = \{\xi_1, \dots, \xi_S\}$.

Let

$$z_s^* := \min \left\{ q_s(\mathbf{z}) := \mathbf{z}^\mathsf{T} \tilde{\mathsf{Q}}_s \mathbf{z} : \mathbf{z} \in \Delta^n \right\},$$

be the deterministic optimization problem under scenario $s = 1, \dots, S$. The value

$$z^{1^*} := \sum_{s=1}^{S} p_s z_s^*$$

is the *wait and see* (WS) solution and it gives a lower bound to z^*_{stoch}, i.e. $z^{1^*} \leq z^*_{\text{stoch}}$.

Better lower bounds can be obtained by dissecting the probability measure in convex combinations of probabilities defined in two or more scenarios. We consider the following refinement chain:

$$\begin{array}{c} \Xi = \{\xi_1, \dots, \xi_S\} \\ \vdots \\ (\Xi_1^{(j)}, \Xi_2^{(j)}, \dots, \Xi_{m_j}^{(j)}) \\ \vdots \\ (\Xi_1^{(2)}, \Xi_2^{(2)}, \dots, \Xi_{m_2}^{(2)}) \\ (\{\xi_1\}, \{\xi_2\}, \dots, \{\xi_S\}), \end{array}$$

where each row is a collection of subsets of the probability space Ξ with the property that their union covers the whole space $\Xi = \cup_i \Xi_i^{(j)}$ for all j and that each set $\Xi_i^{(j)}$is the union of sets from the next more refined collection

$$\Xi_i^{(j)} = \cup_{\Xi_k^{(j-1)} \subseteq \Xi_i^{(j)}} \Xi_k^{(j-1)}.$$

Denoting with

$$z^{j^*} = \sum_{i=1}^{m_j} \pi_i^{(j)} z^*(\Xi_i^{(j)}),$$

where $\pi_i^{(j)} = \sum_{\xi_s \in \Xi_i^{(j)}} p_s$, we get to a chain of lower bounds expressed as follows

$$z^{1^*} \leq z^{2^*} \leq \cdots \leq z^{j^*} \leq \cdots \leq z^*_{\text{stoch}}. \tag{6.6}$$

Notice that the higher the index j, the fewer problems have to be solved but with an increasing number of scenarios. We refer to [58] for examples of constructions of refinement chains. Notice that for all these problems z^{j^*} we can use the bounds defined in the previous section.

6.3 Numerical Experiments

We conduct a series of experiments that investigate the performance of the upper and lower bounds proposed in the previous sections. All of the experiments were implemented in Matlab, using the YALMIP environment for communication with solvers and Mosek solver for solving semidefinite programs (SDPs). For solving standard quadratic problems we used an implementation of the IID-algorithm made available by Michael Kahr (as used in [21]). As a benchmark for our upper bounds we employed the *Fmincon* function available in Matlab using an interior point solver. All experiments were conducted on an AMD Ryzen 7 1700X Eight-Core 3.4 GHz CPU, and 16 GB RAM. The experiments are aimed at answering the following questions:

- How do the scalable bounds proposed in Section 6.2.2 perform in terms of computation time and quality of the solution when compared to the bounds based on Theorem 3?
- How does the IID algorithm for computing the upper bound proposed Section 6.2.1 perform in terms of computation time and quality of the solution when compared to *Fmincon*?
- How good are these methods when initialized at a feasible solution obtained from the copositive relaxation?
- When using an approximation based on dissections of the probability measure (see Section 6.2.3), what is the trade off between solution quality an size of the subsets of the probability space?
- How consistent are the results for the approximations described in Section 6.2.3 for a fixed size of subsets, when the grouping of the scenarios is performed randomly?

  In our experiments we generated data from two different data-generating processes P_1 and P_2:

  - For the first one, say P_1, the elements of the matrix $\tilde{\mathsf{Q}}$ represent the mutual distances of $n = n_1 + n_2$ points sampled under an uniform distribution from the unit square $[0, 1] \times [0, 1]$; the n_1 points are fixed and generates the principal submatrix A of dimension $[n_1 \times n_1]$, while for the other n_2 points is only known that they lie in a square with side length 2ε. The actual location of the n_2 unknown points follows a uniform distribution in these squares and generates the matrices $\tilde{\mathsf{B}}_s$ and $\tilde{\mathsf{C}}_s$, $s[1:S]$ of dimension $[n_1 \times n_2]$ and $[n_2 \times n_2]$ respectively. An example is given in Figure 6.1 where the asterisks represent the $n_1 = 3$ fixed points and the squares the uncertainty regions of the other $n_2 = 5$ points.

– For the second one, say P_2, we choose $(\mathsf{A})_{ij} \sim \mathcal{U}_{\{0,1\}}$, $(\mathsf{B})_{ij} \sim \mathcal{U}_{\{0,1\}}$, $(\mathsf{C})_{ij} \sim \mathcal{U}_{[0,0.1]}$, where $\mathcal{U}_{\mathcal{M}}$ is the uniform distribution with support $\mathcal{M}$.

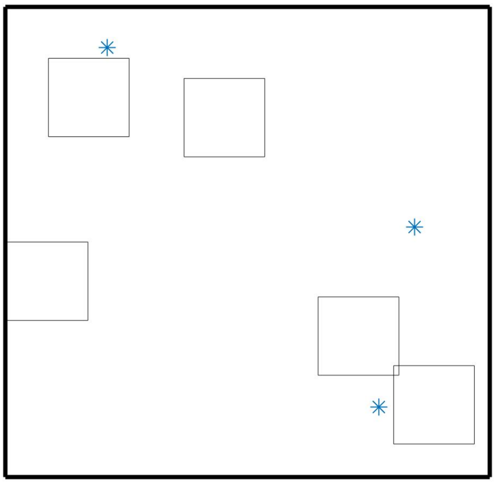


Figure 6.1: Example of $n_1 = 3$ fixed points (*) and the uncertainty region (squares with side length 0.2) of the other $n_2 = 5$ points.

6.3.1 Comparing the copositive bounds

In the following we assessed the value and shortcomings of the compact copositive bound proposed in Section 6.2.2 over the traditional copositive bound from [25]. To this end we picked different configurations for (n_1, n_2, S) and for all of them, we generated 20 instances from P_1 and P_2 respectively. Then we solved the instances using the $\mathcal{DNN}$ relaxations (where the cone of completely positive matrices $\mathcal{CPP}_n$ is replaces by the outer approximation given by the doubly nonegative cone $\mathcal{DNN}_n = \mathcal{S}_+^n \cap \mathcal{N}_n$) of the two copositive reformulations. Note that from each of these relaxation we do not only get a lower bound, i.e. the optimal value of the relaxation, but also an upper bound obtained by considering the feasible values for the considered decision variables.

Table 6.1 summarizes the results obtained under P_1, while Table 6.2 summarizes the results under P_2. In each of them the column **Time** gives the computation time in seconds. In the column **Gap** we give the average percentage gaps between the upper bound and the lower bound produced by the relaxations, and under **max Gap** we give the biggest such gaps that was encountered across all 20 instances. Column DNN1 refers to the results for the traditional copositive bound proposed in [25] and column DNN2 the results for our newly proposed bound given in Section 6.2.2. Finally, column **Vs**. reports the average percentage gap between the two copositive relaxations (avgG) and the biggest such gap across the 20 instances (maxG) respectively. Notice that two objective function values have been considered to be numerically identical if their percentage gap is smaller than 10^{-3} percent, which seems reasonable given the accuracy one can expect from semidefinite optimization software.

(n_1, n_2, S)	**Time**		**Gap**		**max Gap**		**Vs.**	
	DNN1	DNN2	DNN1	DNN2	DNN1	DNN2	avgG	maxG
(5,5,10)	1.438	0.154	0%	0%	0%	0%	0%	0%
(5,10,10)	34.246	0.230	0%	0%	0%	0%	0%	0%
(5,20,10)	1249.161	0.684	0%	0%	0%	0%	0%	0%
(5,40,10)	-	6.228	-	0%	-	0%	-	-
(10,5,10)	2.757	0.865	0%	0%	0%	0%	0%	0%
(20,5,10)	5.010	1.119	0%	0%	0%	0%	0%	0%
(40,5,10)	15.180	3.872	0%	0%	0%	0%	0%	0%
(5,5,20)	35.160	1.186	0%	0%	0%	0%	0%	0%
(5,5,40)	1539.243	1.788	0%	0%	0%	0%	0%	0%
(5,5,60)	-	2.945	-	0%	-	0%	-	-

Table 6.1: Comparison of copositive bounds under P_1.

Table 6.1 shows that both the bounds are tight. It is a known phenomenon, that the bounds obtained from the $\mathcal{DNN}$ relaxations of copositive reformulations is often exact on random instances, and gaps start to appear only after introducing special structures. It is noteworthy, that there is no gap between the two copositive bounds: only the copositive bound from [25] is provably tight, while the bound we proposed is provably tight only in special cases. In conclusion our experiments show that, if there is a gap between the two copositive bounds, then it must be vanishingly small for instances generated from P_1. This is especially encouraging if we take into account the

(n_1, n_2, S)	**Time**		**Gap**		**max Gap**		**Vs.**	
	DNN1	DNN2	DNN1	DNN2	DNN1	DNN2	avgG	maxG
(5,5,10)	1.647	0.191	1.075%	1.373%	13.26%	14.167%	0.019%	0.016%
(5,10,10)	28.05	0.237	0.003%	0.039%	0.056%	0.379%	0.001%	0.001%
(5,20,10)	1132.45	0.754	0.428%	1.281%	5.136%	18.633%	0.001%	0.001%
(10,5,10)	2.942	0.438	1.365%	1.665%	17.201%	18.306%	0.014%	0.011%
(20,5,10)	5.515	0.977	9.432%	9.305%	35.772%	35.824%	0.002%	0.001%
(5,5,20)	37.588	0.546	0.807%	1.394%	13.756%	14.317%	0.029%	0.023%
(5,5,40)	1504.456	0.998	1.09%	1.579%	21.536%	21.625%	0.006%	0.004%

Table 6.2: Comparison of copositive bounds under P_2.

performance with regards to computation time (column **Time**). We see that our new bound scales much nicer with the number of scenarios S, showing a desired property for stochastic optimization problems, since the quality of the solution depends on the number of scenarios that are considered. Even instance types where the traditional lower bound was not solvable due to memory capacity limits (see instances $(5, 40, 10)$ and $(5, 5, 60)$), our scalable lower bound was obtained in few seconds.

In Table 6.2 we see that under P_2 we were able to create instances where the relaxations actually exhibited a gap between the upper and lower bounds they produced (see columns **Gap** and **max Gap**). We found that the scalable copositive bound produced slightly bigger gaps on average. Also the percentage gap between the lower bounds seems to widen but remaining very small. It is questionable whether or not these are reals gaps or whether our threshold for numerical identity (i.e. 10^{-3}) is too large. We have so far not been able to produce instances where the gaps between lower bounds are comfortably outside of what could by numerical inaccuracies.

6.3.2 Warmstarting the upper bound procedures

In Section 6.3.1 we have seen that we can create instances where the upper bounds generated by the copositive relaxation do not close the optimality gap. However, the IID-algorithm allows for a warmstart at the feasible solution generated by the copositive relaxation. The same is true for *Fmincon* available under Matlab environment. In the following experiment we compared the performance of the two methods when warmstarted at a point previously computed by solving the $\mathcal{DNN}$-relaxation of the copositive relaxation. In order to make the initial gaps larger we slightly modified P_2 so that $(\mathsf{A})_{ij} \sim \mathcal{U}_{\{0,1\}}$, $(\mathsf{B})_{ij} \sim \mathcal{U}_{\{0,10\}}$ and $(\mathsf{C})_{ij} \sim \mathcal{U}_{\{0,0.1\}}$.

We summarize the results in Table 6.3. For each configuration of (n_1, n_2, S), indicated in the first column, we generated 20 instances, calculated an upper bound using the $\mathcal{DNN}$ approach and warmstarted both the IID-algorithm and *Fmincon* from that point. The IID-algorithm depends on the parameter λ which is the weight of the quadratic penalization of the non-simplicial constraints. In our implementation we chose two different values $\lambda = 10, 100$ which we indicated in the subscript (IID_λ).

Also, as stated in Section 6.2.1, the IID-algorithm may yield infeasible solutions due to the nature of quadratic penalizations. Thus, we used Mosek in order to calculate the projections of an infeasible solution onto the feasible set, which amounts to solving a second order cone problem. In our experiments the amount of constraint violation was typically very small so that the projections did not affect the optimal values by much, and the additional computation time was small as well. We therefore do not explicitly report the individual components of the optimal value and the computation time, but rather give these values for the overall procedure.

The column **DNNgap** gives the average percentage gap between the upper and the lower bound obtained from the relaxation. The column **NewGap** reports the gap between the $\mathcal{DNN}$-lower bound and the newly computed upper bound from the respective methods. The column **Best** counts the number of times the respective method gave the best upper bond of the three, **IID*** counts the number of times where this was achieved by an IID-algorithm. Finally, the column **Time** gives the average computation time in seconds.

(n_1, n_2, S)	**DNNgap**	**New Gap**			**Best**			**IDD***	**Time**		
		IID_{10}	IID_{100}	Fmincon	IID_{10}	IID_{100}	Fmincon		IID_{10}	IID_{100}	Fmincon
(5,5,10)	72.93%	2.01%	**2.00%**	2.22%	0	2	18	2	4.151	35.883	**1.368**
(5,10,10)	51.94%	0.79%	**0.78%**	2.41%	1	9	10	10	7.086	64.146	**4.134**
(5,20,10)	47.09%	**0.34%**	**0.34%**	3.06%	1	13	6	14	16.255	152.000	**12.632**
(10,5,10)	73.44%	1.82%	**1.8%**	2.48%	0	11	9	11	3.776	30.992	**2.356**
(20,5,10)	94.8%	0.64%	**0.62%**	3.29%	0	13	7	13	3.235	25.976	**3.091**
(5,5,20)	67.62%	1.16%	**1.14%**	1.28%	1	7	12	8	29.600	293.960	**5.061**
(5,5,40)	61.72%	**0.18%**	**0.18%**	**0.18%**	2	8	10	10	245.700	2434.700	**11.847**

Table 6.3: Comparison between IID_{10}, IID_{100} and Fmincon when warmstarted.

We see that all three methods reduced the optimality gap substantially where the IDD_{100} consistently performed best in terms of average percentage gap. However, IDD does not consistently outperform Fmincon in every instance, as in more than half of the instances (72 times), Fmincon gave a better bound. Also, computation time is much better for Fmincon, with IDD's performance being especially bad when the number of scenarios S is high.

6.3.3 Experiments on dissecting the probability measure

This section presents some computational tests for problems being intractable involving a large number of scenarios. We generate data from P_1 with $n_1 = n_2 = 15$, $\epsilon = 0.1$ and considering $S = 200$ equiprobable scenarios. We constructed lower bounds as described in Section 6.2.3 by partitioning the scenario tree in disjoint subgroups.

We now describe the way we constructed the groups of each layer j of the dissection chain. We fix a certain group size, say n_g, with $1 < n_g < S$. The first layer always corresponds to the wait and see (WS) approach, i.e. we consider each scenario independently so that the number of "groups" is $S_1 := S$. In the second layer, we randomly group the scenarios into groups of size n_g. If the total number of groups S_1 is not evenly divisible by n_g, a remainder group is constructed with cardinality S_1 mod n_g. This results in S_2 groups, with $S_2 = S_1$ div n_g if the division has no remainder, or $S_2 = (S_1 \text{ div } n_g) + 1$ if there is a remainder and an additional remainder group is added. These S_2 groups are again randomly grouped into super-groups of n_g groups of the previous layer. Thus, for the j-th layer the number of super-groups is given by

$$S_j := \begin{cases} S_{j-1} \text{ div } n_g & \text{if } S_{j-1}/n_g \in \mathbb{N} \\ (S_{j-1} \text{ div } n_g) + 1 & \text{else.} \end{cases} \tag{6.7}$$

Figure 6.2 provides a boxplot of the percentage gaps obtained from a chain of lower bounds composed by 5 layers over 20 random dissections as described below. The first bound in the chain (z^{1*}) refers to the wait and see (WS) solution and it has been obtained solving the $S_1 = 200$ scenarios independently. The next one (z^{2*}) by partitioning the scenario tree into $S_2 = 40$ subproblems of $n_g = 5$ scenarios each, solving them independently and taking their mean. These sets were again grouped making $S_3 = 8$ subproblems of 25 scenario each to obtain z^{3*}. Finally z^{4*} has been obtained by grouping 5 groups of 25 scenarios in one subproblem of 125 scenario and the remaining 3 groups of 25 scenarios in one of 75 with a total of $S_4 = 2$. The last level in the chain z^{5*} refers to the original problem z^*_{stoch}. The decision which scenarios, or sets of scenarios were grouped was done at random and we created 20 random dissections in this manner. For each of the dissections we calculated z^{j^*} for each layer $j[1:5]$ using the scalable copositive bounds (which empirically we have found to be tight for P_1) and subsequently calculated the percentage gaps between z^{j^*} and $z^*_{stoch} \equiv z^{5^*}$.

Results show that even in the second layer, where groupings of five scenarios are considered, the values for z^{2^*} substantially outperform z^{1^*} across all 20 dissections

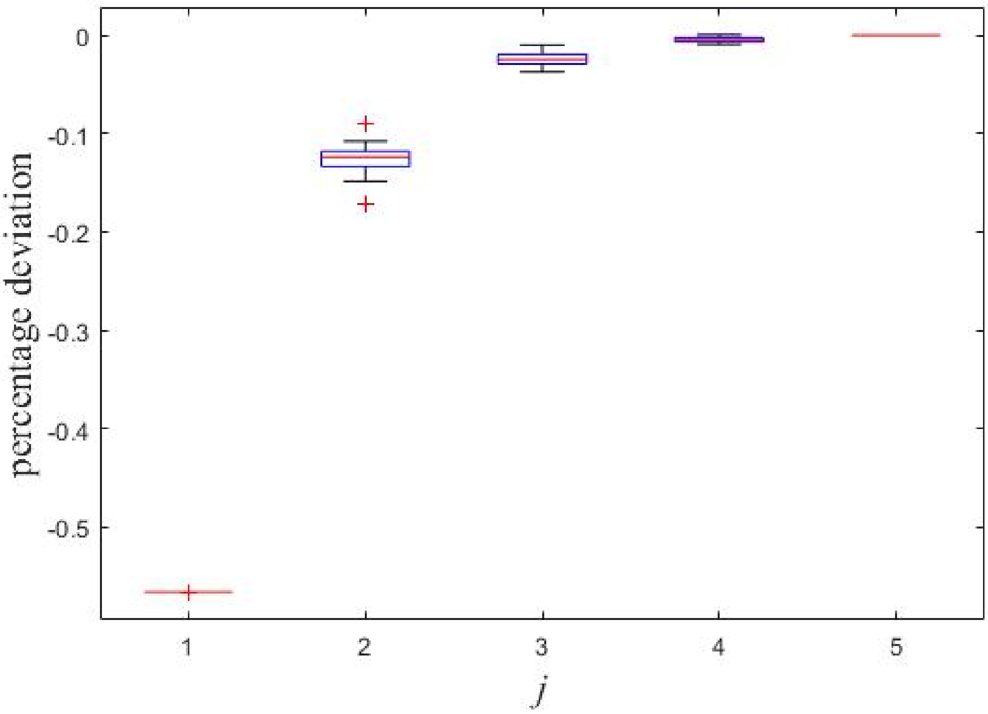


Figure 6.2: Percentage deviation from the optimal objective value z^*_{stoch} for level j of a refinement chain composed by 5 layers.

passing from a percentage gap of -0.5658 to -0.1236. For deeper layers z^{3*} and z^{4*}, the quality of the bounds as well as the consistency between dissections improve, but the marginal improvement diminishes.

We repeated the experiment, this time considering a new tree with again 200 scenarios but this time with $n_g = 2$. Figure 6.3 summarizes the results of this process confirming the fact that increasing the layer j, the quality of the bound improve, while the marginal improvement diminishes. In particular we notice that the seventh layer seems to outperform the eighth and the "true value" z^*_{stoch}. The reason for this phenomenon is the fact that deeper into the layers we have to solve exponentially larger semidefinite optimization problems which are known to have problems with scaling. Specifically, the accuracy of the solution suffers from the scale of the problem so that in fact the final few layers are numerically indistinguishable when taking into account the noise added by the solution method. This shortcoming, at least in our experiment, is partly redeemed by the fact that already for the shallow layers the quality of the bounds is satisfactory since the gap is close to zero.

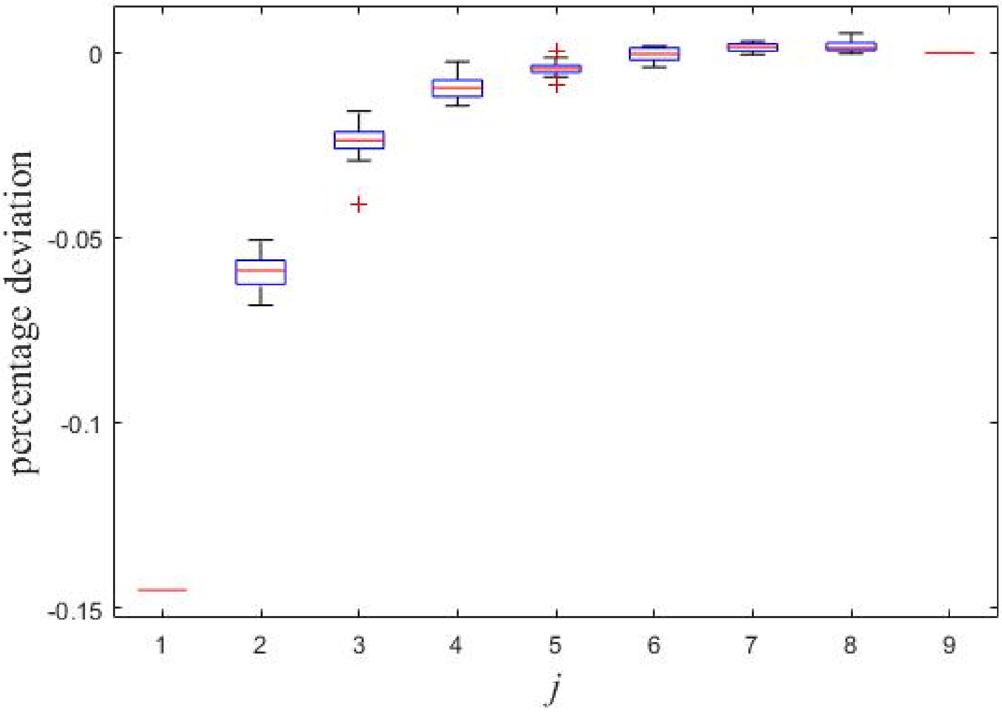


Figure 6.3: Percentage deviation from the optimal objective value z^*_{stoch} for level j of a refinement chain composed by 9 layers.

Chapter 7

Interplay of non-convex quadratically constrained quadratic problems with adjustable robust optimization

7.1 Introduction

In this chapter we explore the connections between robust optimization and quadratically constrained quadratic optimization. In Section 7.2 we discuss a strategy by which convex reformulations of QCQPs have been used in literature for the purpose of reformulating semi-infinite constraints with quadratic index. The strategy is not new, but it has so far not been discussed in a general manner. We think this generic perspective is instructive and helps familiarize the two fields with each other, and also allows for a discussion of the critical elements of that approach. We give small illustrative examples, also to highlight the important role of dual attainability in the reformulation.

Section 7.3 gives a short discussion on the types of uncertainty sets for which the general strategy is feasible and review how these have been used in literature so far.

In Section 7.4 we will apply the general strategy to various instances of robust optimization. We show that we can achieve reformulations of ARO under linear decision rules with uncertain recourse, ARO under quadratic decision rules and robust convex optimization in a single stage and two stage setting. We provide numerical experiments to demonstrate the efficacy of these techniques in Section 7.4.3.

In Section 7.5 we give an alternative, copositive perspective on the general reformulation strategy for semi-infinite constraints. The so obtained reformulation can be used to characterize the gap that occurs if the underlying QCQP-reformulation fails to exhibit dual attainability.

Finally, in Section 7.6, we show that a class of QCQPs can be reformulated as adjustable robust optimization problems which themselves allow for a bound based on the general reformulation strategy for semi-infinite constraints (see Theorem 41). The resulting lower bound is tested against known lower bounds found in literature. Empirical evidence suggests that for some range of the problem parameters our lower bound can be beneficial in terms of computation time and solution quality.

7.2 Convex reformulations of QCQPs and robust optimization

We will now proceed in describing the general strategy by which reformulation results such as the ones presented in Chapter 1 can be leveraged in order to cope with cases where the semi-infinite constraints present in robust optimization depend quadratically on the uncertainty vector $\mathbf{u}$. Assume that

$$f(\mathbf{x},\mathbf{u}) = \mathbf{u}^\mathsf{T}\mathsf{Q}(\mathbf{x})\mathbf{u} + \mathbf{q}(\mathbf{x})^\mathsf{T}\mathbf{u} + \omega(\mathbf{x}), \tag{7.1}$$

where $\mathsf{Q}(\mathbf{x}), \mathbf{q}(\mathbf{x}), \omega(\mathbf{x})$ are appropriate matrix-, vector- and scalar-valued functions of the decision vector $\mathbf{x}$. For ease of notation we refer to them as $\mathsf{Q}, \mathbf{q}$ and ω, suppressing the dependence on $\mathbf{x}$. We thus consider the semi-infinite constraint and its (possibly non-smooth) single-constraint equivalent

$$\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega \geq 0 \ \forall \mathbf{u} \in \mathcal{U} \quad \iff \quad \inf_{\mathbf{u}\in\mathcal{U}} \left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega\right] \geq 0\,. \tag{7.2}$$

Now suppose that for generic $\mathsf{Q}, \mathbf{q}$ and ω the minimization problem in (7.2) has an exact convex, conic reformulation of the following form, using an appropriate matrix cone $\mathcal{C}$ and appropriate matrices $\mathsf{A}_i \in \mathcal{S}^{n+1}$, real numbers $b_i \in \mathbb{R},\ i \in [1\!:\!m]$, so that,

$$\inf_{\mathbf{u}\in\mathcal{U}} \left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega\right] = \inf_{\mathsf{Y}\in\mathcal{C}} \{\mathsf{M}(\omega,\mathbf{q},\mathsf{Q}) \bullet \mathsf{Y} \colon \mathsf{A}_i \bullet \mathsf{Y} \leq b_i\,,\ i \in [1\!:\!m]\}\,.$$

Further, assume that for this reformulation and its dual we can establish zero duality gap and dual attainability (e.g., if the primal satisfies Slater's condition), so that

$$\begin{aligned}&\inf_{\mathsf{Y}\in\mathcal{C}} \{\mathsf{M}(\omega,\mathbf{q},\mathsf{Q}) \bullet \mathsf{Y} \colon \mathsf{A}_i \bullet \mathsf{Y} \leq b_i\,,\ i \in [1\!:\!m]\}\\ &= \sup_{\boldsymbol{\lambda}\in\mathbb{R}^m_+} \left\{-\mathbf{b}^\mathsf{T}\boldsymbol{\lambda} \colon \mathsf{M}(\omega,\mathbf{q},\mathsf{Q}) + \sum_{i=1}^m \lambda_i \mathsf{A}_i \in \mathcal{C}^*\right\}\,.\end{aligned}$$

Since dual attainability guarantees the existence of the dual maximizers, we can enforce the semi-infinite constraint in (7.2) by demanding that

$$\mathbf{b}^{\mathsf{T}}\boldsymbol{\lambda} \leq 0 \quad \text{and} \quad \mathsf{M}(\omega, \mathbf{q}, \mathsf{Q}) + \sum_{i=1}^{m} \lambda_i \mathsf{A}_i \in \mathcal{C}^* \quad \text{for some } \boldsymbol{\lambda} \in \mathbb{R}^m_+ . \tag{7.3}$$

In summary, the strategy is to bypass the need to directly dualize the implicit QCQP in (7.2) by providing a linear conic reformulation, a dual of which can be formed by invoking linear conic duality, which is a very well developed and understood subject. Many results in recent literature have harnessed this general strategy in order to cope with cases where quadratic terms in the uncertainty vector appear (see e.g. [59, 84]). We want to highlight that the critical ingredients for the above strategy are:

- closing the relaxation gap,
- closing the duality gap for the conic reformulation and
- guaranteeing dual attainability.

While the first obstacle seems to be the most challenging, duality also comes with some subtleties attached. In Section 7.5 we will make an effort to understand the reformulation gap, that may occur in case of a failure of dual attainability. For now, we will give some examples that illustrate the above strategy and also demonstrate that the gap can be infinite if dual attainability does not hold.

Example 8. *Consider the robust optimization problem*

$$\begin{aligned} &\min\, x + y \\ &\text{s.t.}: xu_1^2 + yu_2^2 \geq 1 \quad \forall (u_1, u_2) \in \mathbb{R}^2_+ : u_1 + u_2 = 1. \end{aligned} \tag{7.4}$$

If, out of the infinitely many constraints, we enforced merely $0.5^2x + 0.5^2y \geq 1$, we obtained a relaxed problem with an optimal value of 4 attained at $x = y = 2$. This solution is also feasible for the original problem since $\min\left\{2u_1^2 + 2u_2^2 \colon u_1 + u_2 = 1, (u_1, u_2) \in \mathbb{R}^2_+\right\} = 1$, *and thus it is also optimal. We will now show, that the general reformulation strategy yields an equivalent problem with finitely many constraints. It is well known that* $\min\left\{au_1^2 + bu_2^2 + cu_1u_2 \colon u_1 + u_2 = 1,\ (u_1, u_2) \in \mathbb{R}^2_+\right\} = \min\left\{au_{11} + bu_{22} + cu_{21} \colon u_{11} + 2u_{21} + u_{22} = 1,\ \mathsf{U} \in \mathcal{S}^2_+ \cap \mathcal{N}_2\right\}$ *(see [11]). We have* $\mathsf{M}(1/3, 2/6, 1/3) \in \operatorname{int}\left(\mathcal{S}^2_+ \cap \mathcal{N}\right)$ *so by Slater's condition the latter optimization problem is equivalent to* $\max\left\{\lambda \colon \mathsf{M}(a, c, b) - \lambda \mathsf{M}(1, 2, 1) \in \mathcal{S}^2_+ + \mathcal{N}_2\right\}$*, and the maximum is attained. We thus have the equivalences*

$$xu_1^2 + yu_2^2 \geq 1 \quad \forall (u_1, u_2) \in \mathbb{R}^2_+ : u_1 + u_2 = 1,$$

$$\begin{aligned}
&\Leftrightarrow \min\left\{xu_1^2+yu_2^2\colon (u_1,u_2)\in\mathbb{R}^2_+\colon u_1+u_2=1\right\}\geq 1,\\
&\Leftrightarrow \min\left\{xu_{11}+yu_{22}\colon u_{11}+2u_{21}+u_{22}=1,\ \mathsf{U}\in\mathcal{S}^2_+\cap\mathcal{N}_2\right\}\geq 1,\\
&\Leftrightarrow \max\left\{\lambda\colon \mathsf{M}(x,0,y)-\lambda\mathsf{M}(1,2,1)\in\mathcal{S}^2_++\mathcal{N}_2\right\}\geq 1,\\
&\Leftrightarrow \lambda\geq 1,\ \mathsf{M}(x,0,y)-\lambda\mathsf{M}(1,2,1)\in\mathcal{S}^2_++\mathcal{N}_2 \textit{ for some } \lambda\in\mathbb{R}.
\end{aligned}$$

Thus we can replace the constraints in (7.4) in order to obtain

$$\begin{aligned}
&\min\ x+y\\
&\text{s.t.}:\ \lambda\geq 1,\ \mathsf{M}(x,0,y)-\lambda\mathsf{M}(1,2,1)\in\mathcal{S}^2_++\mathcal{N}_2,
\end{aligned}\tag{7.5}$$

and can see by computation that the minimum is attained at $x=y=2$ and $\lambda=1$.

Example 9. *To illustrate the issues of a failure of dual attainability for the conic reformulation we consider the robust optimization problem*

$$\begin{aligned}
&\min\ x+y\\
&\text{s.t.}:\ (x+y)u_1^2\geq u_2u_3\quad \forall(u_1,u_2,u_3)\in[0,1]\times[0,1]\times\{0\},\\
&\qquad\quad x+y\leq 0.
\end{aligned}\tag{7.6}$$

Note that the semi-infinite constraint is equivalent to $x+y\geq 0$, so that the constraints overall imply $x+y=0$ and the minimum is 0. We claim that for any $a\in\mathbb{R}$ we have $q(a):=\min\{au_1^2-u_2u_3\colon (u_1,u_2,u_3)\in[0,1]\times[0,1]\times\{0\}\}=\min\left\{au_{11}-u_{23}\colon u_{11}\leq 1,u_{22}\leq 1,u_{33}=0,\mathsf{U}\in\mathcal{S}^3_+\right\}$. The claim follows easily after considering that $\mathsf{U}\in\mathcal{S}^3_+$ and $u_{33}=0$ imply $u_{23}=0$. Note, that this reformulation is not based on a characterization of $\mathcal{G}(\mathcal{F})$, but is coincidental, and chosen here in order to illustrate a point about dual attainability. Also, we see that $q(a)=a$ if $a<0$ and zero otherwise. The conic reformulation does not have a Slater point, however its dual, given by

$$\begin{aligned}
&\sup\ \lambda_1+\lambda_2\\
&\text{s.t.}:\ \mathsf{M}_D(a,\lambda_1,\lambda_2,\lambda_3):=\begin{pmatrix} a-\lambda_1 & 0 & 0\\ 0 & -\lambda_2 & -1/2\\ 0 & -1/2 & -\lambda_3\end{pmatrix}\in\mathcal{S}^3_+,\quad \lambda_1\leq 0.
\end{aligned}$$

does have a Slater point, where $\lambda_i, i=1,2,3$ are sufficiently small. Thus the conic reformulation and its dual have the same optimal value, but dual attainability is not guaranteed. In fact, some elementary analysis reveals that the dual does not attain its optimum. We thus merely have the following implications

$$(x+y)u_1^2\geq u_2u_3\quad \forall(u_1,u_2,u_3)\in[0,1]\times[0,1]\times\{0\},$$

$$
\begin{aligned}
&\Leftrightarrow \min\left\{(x+y)u_1^2 - u_2u_3 \colon (u_1,u_2,u_3) \in [0,1]\times[0,1]\times\{0\}\right\} \geq 0,\\
&\Leftrightarrow \min\left\{(x+y)u_{11} - u_{23} \colon u_{11} \leq 1, u_{22} \leq 1, u_{33} = 0, \mathsf{U} \in \mathcal{S}^3_+\right\} \geq 0,\\
&\Leftrightarrow \sup\left\{\lambda_1 + \lambda_2 \colon \mathsf{M}_D((x+y),\lambda_1,\lambda_2,\lambda_3) \in \mathcal{S}^3_+, \lambda_1 \leq 0\right\} \geq 0,\\
&\Leftarrow \lambda_1 + \lambda_2 \geq 0,\ \mathsf{M}_D((x+y),\lambda_1,\lambda_2,\lambda_3) \in \mathcal{S}^3_+,\ \lambda_1 \leq 0 \text{ for some } \boldsymbol{\lambda} \in \mathbb{R}^3.
\end{aligned}
$$

The final constraint implies $x + y > 0$ since in case of $x + y = 0$ we had $\lambda_1 = 0$ and λ_2 can neither be zero nor positive. Thus if we replaced the semi-infinite constraint in (7.6) by the latter constraint, the problem would become infeasible. This exemplifies that closing the relaxation gap as well as the duality gap is in general not enough to guarantee an exact reformulation.

7.3 Viable uncertainty sets

The general strategy described in the previous section relies on the existence of a conic reformulation of the implicit QCQP. We discussed such reformulations in the first part of the book. Translating the approach discussed there into the language of robust optimization, we can say that the general strategy is applicable whenever, for an uncertainty set $\mathcal{U}$ we have a conic characterization of $\mathcal{G}(\mathcal{U})$ so that the respective conic reformulation has strong duality and dual attainment. We give a short summary of the types of uncertainty sets for which we have a respective characterization in the theorem below. Of course there are more characterizations than the ones depicted here (see the references is Section 2.1) but we limit the presentation to the ones that either have been used in the robust optimization literature or are considered in the sequel of this text.

Theorem 28. *Consider the following uncertainty sets:*

- $\mathcal{U}_1 \coloneqq \{\mathbf{u} \in \mathbb{R}^m \colon \|\mathbf{u}\| \leq 1\}$.
- $\mathcal{U}_2 \coloneqq \{\mathbf{u} \in \mathbb{R}^m \colon \|\mathbf{u}\| \leq 1, \mathsf{A}\mathbf{u} \leq \mathbf{b}\}$*, where the hyperplanes described by* $\mathsf{A}\mathbf{u} = \mathbf{b}$ *do not intersect inside the unit ball.*
- $\mathcal{U}_3 \coloneqq \left\{\mathbf{u} \in \mathbb{R}^m \colon \|\mathbf{u}\| \leq 1,\ \mathbf{u}^\mathsf{T}\mathsf{W}_i\mathbf{u} + \mathbf{w}_i^\mathsf{T}\mathbf{u} + w_i \leq 0,\ i \in [1\colon k]\right\}$*, where quadratic constraints induce nonintersecting holes into the ellipsoid.*
- $\mathcal{U}_4 \coloneqq \left\{\mathbf{u} \in \mathcal{K} \colon \mathsf{A}\mathbf{u} = \mathbf{b},\ \mathbf{u}^\mathsf{T}\mathsf{Q}_i\mathbf{u} + \mathbf{q}_i^\mathsf{T}\mathbf{u} + \omega_i = 0,\ i \in [1\colon k]\right\}$ *where* $\mathcal{K} \subseteq \mathbb{R}^m$ *is a cone and the qudadratic constraints fulfill the key assumption in Theorem 3.*

Then

- $\mathcal{G}(\mathcal{U}_1) = \{\mathsf{Y}(\mathbf{u}, \mathsf{U}) \in \mathcal{S}^{m+1} \colon \mathrm{trace}(\mathsf{U}) \leq 1\}$,
- $\mathcal{G}(\mathcal{U}_2) = \left\{\mathsf{Y}(\mathbf{u}, \mathsf{U}) \in \mathcal{S}^{m+1} \colon \mathrm{trace}(\mathsf{U}) \leq 1, \begin{array}{l} \|b_i\mathbf{u} - \mathsf{U}\mathbf{a}_i\| \leq b_i - \mathbf{a}_i^\mathsf{T}\mathbf{u},\ i \in [1\!:\!k], \\ b_i\mathbf{a}_j^\mathsf{T}\mathbf{u} + b_j\mathbf{a}_i^\mathsf{T}\mathbf{u} - \mathbf{a}_i^\mathsf{T}\mathsf{U}\mathbf{a}_j \leq b_ib_j,\ (i,j) \in [1\!:\!k]^2 \end{array}\right\}$,
- $\mathcal{G}(\mathcal{U}_3) = \left\{\mathsf{Y}(\mathbf{u}, \mathsf{U}) \in \mathcal{S}^{m+1} \colon \mathrm{trace}(\mathsf{U}) \leq 1,\ \mathsf{W}_i \bullet \mathsf{U} + \mathbf{w}_i^\mathsf{T}\mathbf{u} + w_i \leq 0,\ i \in [1\!:\!k]\right\}$,
- $\mathcal{G}(\mathcal{U}_4) = \left\{\mathsf{Y}(\mathbf{u}, \mathsf{U}) \in \mathcal{CPP}(\mathbb{R}_+ \times \mathcal{K}) : \begin{array}{l} \mathsf{A}\mathbf{u} = \mathbf{b},\ \mathrm{diag}\left(\mathsf{AUA}^\mathsf{T}\right) = \mathbf{b} \circ \mathbf{b}, \\ \mathsf{Q}_i \bullet \mathsf{U} + \mathbf{q}_i^\mathsf{T}\mathbf{u} + \omega_i = 0,\ i \in [1\!:\!k] \end{array}\right\}$.

Proof. The first characterization can be inferred to either from the S-Lemma, the results in [72], from Theorem 9 or from Theorem 10. The second is due to [29]. The third one is from Theorem 10 and the final one is given by Theorem 3. □

From [81, Theorem 5.1] we get that $\mathcal{G}(\mathcal{U})$ has interior whenever $\mathcal{U}$ does so that $\mathcal{G}(\mathcal{U}_i),\ i = 1, 2, 3$ can always be modelled so that they have a Slater point. That is, however, not generally the case with $\mathcal{G}(\mathcal{U}_4)$ and we will discuss this matter in greater detail in Section 7.5.

$\mathcal{U}_1$ is the classical ellipsoidal uncertainty set for which one can find ample discussion in [6]. The characterization of $\mathcal{G}(\mathcal{U}_1)$ is due to the S-Lemma and is therefore widely konwn and has been put to use for the reformulation of semi-infinite constraints with quadratic index for example in [7] for the reformulation of the adjustable robust counterpart of linear optimization problem under a affine decision rule with uncertain recourse. $\mathcal{U}_2$ can be used for modeling uncertainty sets which are intersections of a box uncertainty set, another classical uncertainty set and an ellipsoidal uncertainty set (as long as as the geometric condition is not violated) as suggested in [84]. In [46] the capability of $\mathcal{U}_4$ to encode binary variables was used in order to model uncertainty sets with mixed integer components which they then applied to robust convex quadratically constrained quadratic problems using the respective specification of $\mathcal{G}(\mathcal{U}_4)$. Unfortunately, neither have we found a good application of $\mathcal{U}_3$ ourselves nor in literature. However, we discuss preliminary ideas on this matter in Section 7.4.3.

7.4 Application to robust and adjustable robust optimization

In the following, we look at specific instances of robust optimization where the quadratic dependence on $\mathbf{u}$ arises naturally from the models in questions. In this section we are mainly concerned with the models and their reformulation. Applications of these models and numerical experiments are to be found in the succeeding section.

7.4.1 Adjustable robust counterparts of linear optimization problems

We consider the Adjustable Robust Linear Optimization Problem (ARO):

$$\min_{\mathbf{x}\in\mathcal{X},\mathbf{y}(.)} \left\{\mathbf{c}^\mathsf{T}\mathbf{x}\colon \mathsf{A}(\mathbf{u})\mathbf{x} + \mathsf{B}(\mathbf{u})\mathbf{y}(\mathbf{u}) \geq \mathbf{d}(\mathbf{u}) \quad \forall \mathbf{u}\in\mathcal{U}\right\}, \tag{7.7}$$

where $\mathcal{U} \subset \mathbb{R}^m$ is the uncertainty set of possible values for the uncertainty parameter $\mathbf{u} \in \mathcal{U}$. The first stage decision variable is $\mathbf{x} \in \mathbb{R}^{n_1}$ and the second stage decision is modeled as a function $\mathbf{y}\colon \mathcal{U} \to \mathbb{R}^{n_2}$. Further $\mathsf{A}(\mathbf{u}) : \mathcal{U} \mapsto \mathbb{R}^{k\times n_1}, \mathsf{B}(\mathbf{u}) : \mathcal{U} \mapsto \mathbb{R}^{k\times n_2}$ are functions that are linear in the uncertainty parameter. We model the i-th row of $\mathsf{A}(\mathbf{u})$ as $(\mathsf{A}(\mathbf{u}))_i^\mathsf{T} = \mathsf{A}_i\mathbf{u} + \mathbf{a}_i$ and similarly the i-th row of $\mathsf{B}(\mathbf{u})$ will be given by $(\mathsf{B}(\mathbf{u}))_i^\mathsf{T} = \mathsf{B}_i\mathbf{u} + \mathbf{b}_i$, where $\mathsf{A}_i \in \mathbb{R}^{n_1\times m}$, $\mathbf{a}_i \in \mathbb{R}^{n_1}$, $\mathsf{B}_i \in \mathbb{R}^{n_2\times m}$ and $\mathbf{b}_i \in \mathbb{R}^{n_2}$. The right hand side $\mathbf{d}(\mathbf{u}) : \mathcal{U} \mapsto \mathbb{R}^k$ depends elementwise quadratically on the uncertainty parameter so that for the i-th entry we have $(\mathbf{d}(\mathbf{u}))_i = \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i$ for some $\mathsf{D}_i \in \mathcal{S}^m$, $\mathbf{d}_i \in \mathbb{R}^m$ and $d_i \in \mathbb{R}$. $\mathcal{X} \subset \mathbb{R}^{n_1}$ is a tractable feasible set for the first stage decision, which is not affected by uncertainty. Finally, $\mathbf{c} \in \mathbb{R}^{n_1}$ are certain objective function coefficients. Since any uncertainty in the objective function can be moved into the constraints (by means of introducing an epigraphical variable) this formulation is in fact generic. The second stage decision variable is modelled as a function, so that the search space is the set of all functions which map from the uncertainty set into $\mathbb{R}^{n_2}$, which is an intractable search space.

7.4.1.1 ARO under affine policy with uncertain recourse

The standard procedure to meet the difficulty of optimizing over the space of functions is to restrict that space to the set of linear functions. So the second stage response is subjected to what is called an affine policy, which is a standard approach first outlined in [7]. Thus regarding the i-th constraint of ARO we go from

$$(\mathsf{A}_i\mathbf{u} + \mathbf{a}_i)^\mathsf{T}\mathbf{x} + (\mathsf{B}_i\mathbf{u} + \mathbf{b}_i)^\mathsf{T}\mathbf{y}(\mathbf{u}) + \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u}\in\mathcal{U},$$

to

$$(\mathsf{A}_i\mathbf{u} + \mathbf{a}_i)^\mathsf{T}\mathbf{x} + (\mathsf{B}_i\mathbf{u} + \mathbf{b}_i)^\mathsf{T}(\mathsf{Y}\mathbf{u} + \mathbf{y}_0) + \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u}\in\mathcal{U},$$

by choosing $\mathbf{y}(\mathbf{u}) = \mathsf{Y}\mathbf{u} + \mathbf{y}_0$ where $\mathsf{Y} \in \mathbb{R}^{n_2\times m}$ and $\mathbf{y}_0 \in \mathbb{R}^{n_2}$. The second stage variables are now no longer functions but rather vector valued coefficients of a vector valued function linear in $\mathbf{u}$. Overall we see, that the constraint now is a non-homogeneous quadratic function in $\mathbf{u}$ and therefore we cannot apply reformulation strategies based on linear optimization duality. Note that that is the case even if the term $\mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u}$ were omitted. In fact the assumptions of uncertainty recourse already gives rise to quadratic terms due to the interaction of the affine functions $\mathsf{B}_i(\mathbf{u})$ and $\mathbf{y}(\mathbf{u})$. However, we may apply the reformulation steps outlined in Section 7.2 on order to obtain a conic reformulation of the constraint. We see, thus, that the applicability of linear decision rules is greatly enhanced.

Theorem 29. *Assume $\mathcal{U} \subseteq \mathbb{R}^m$ is a bounded set such that for any $\mathsf{Q} \in \mathcal{S}^m$, $\mathbf{q} \in \mathbb{R}^m$, $\omega \in \mathbb{R}$ we have*

$$\inf_{\mathbf{u}\in\mathcal{U}} \left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega\right] = \inf_{\mathsf{Y}\in\mathcal{C}}\{\mathsf{M}(\omega, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{Y} \colon\, \mathsf{G}_l \bullet \mathsf{Y} \leq g_l\,,\; l \in [1\!:\!m]\},$$

and the dual of the right hand side problem attains its optimum. Then (7.7) with $\mathbf{y}(\mathbf{u}) = \mathsf{Y}\mathbf{u} + \mathbf{y}_0$ can be reformulated as

$$\begin{aligned}
&\min_{\mathbf{x},\mathsf{Y},\mathbf{y}_0,\boldsymbol{\lambda}_i} \; \mathbf{c}^\mathsf{T}\mathbf{x} \\
&\text{s.t. : } \mathbf{g}^\mathsf{T}\boldsymbol{\lambda}_i \leq 0, \qquad i \in [1\!:\!k], \\
&\begin{pmatrix} \mathbf{a}_i^\mathsf{T}\mathbf{x} + \mathbf{b}_i^\mathsf{T}\mathbf{y}_0 + d_i & \frac{1}{2}\left(\mathsf{A}_i^\mathsf{T}\mathbf{x} + \mathsf{Y}^\mathsf{T}\mathbf{b}_i + \mathsf{B}_i^\mathsf{T}\mathbf{y}_0 + \mathbf{d}_i\right)^\mathsf{T} \\ \frac{1}{2}\left(\mathsf{A}_i^\mathsf{T}\mathbf{x} + \mathsf{Y}^\mathsf{T}\mathbf{b}_i + \mathsf{B}_i^\mathsf{T}\mathbf{y}_0 + \mathbf{d}_i\right) & \mathsf{D}_i + \frac{1}{2}\left(\mathsf{B}_i^\mathsf{T}\mathsf{Y} + \mathsf{Y}^\mathsf{T}\mathsf{B}_i\right) \end{pmatrix} + \sum_{l=1}^m \lambda_l^i \mathsf{G}_l \in \mathcal{C}^*, \qquad i \in [1\!:\!k], \\
&\boldsymbol{\lambda}_i \in \mathbb{R}^m_+, \qquad i \in [1\!:\!k], \\
&\mathbf{x} \in \mathcal{X},\; \mathsf{Y} \in \mathbb{R}^{n_2\times m},\; \mathbf{y}_0 \in \mathbb{R}^{n_2}.
\end{aligned}$$

Proof. The theorem follows immediately from the discussion in Section 7.2 and the stucture of the semi-infinite constraint

$$(\mathsf{A}_i\mathbf{u} + \mathbf{a}_i)^\mathsf{T}\mathbf{x} + (\mathsf{B}_i\mathbf{u} + \mathbf{b}_i)^\mathsf{T}(\mathsf{Y}\mathbf{u} + \mathbf{y}_0) + \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}.$$

Note that in order to symmetrize the quadratic term we use $\mathbf{u}^\mathsf{T}\mathsf{B}_i^\mathsf{T}\mathsf{Y}\mathbf{u} = \mathbf{u}\frac{1}{2}\left(\mathsf{B}_i^\mathsf{T}\mathsf{Y} + \mathsf{Y}^\mathsf{T}\mathsf{B}_i\right)\mathbf{u}$. □

7.4.1.2 ARO with quadratic policy

As discussed in Section 5.1.6.1 the appeal of linear policies is the fact that the resulting semi-infinite constraint is easily reformulated using standard convex duality results. More elaborate decision rules such as piecewise affine decision rules already fail to have this quality in the general case and thus have require an approximation scheme themselves [34]. Using the framework described in Section 7.2, we are able to employ quadratic decision rules which deliver two major benefits. First, quadratic decision rules engulf affine decision rules as a spacial case. Hence, the resulting conservative approximations of ARO are at least as good as the ones obtained under affine policy. Second, given we can exactly reformulate the quadratic optimization problem implicit in the semi-infinite constraint, the reformulation of ARO with quadratic decision rules will be exact.

The model we are now interested in is again

$$\min_{\mathbf{x}\in\mathcal{X},\mathbf{y}(.)} \left\{\mathbf{c}^\mathsf{T}\mathbf{x}\colon\ \mathsf{A}(\mathbf{u})\mathbf{x}+\mathsf{B}\mathbf{y}(\mathbf{u}) \geq \mathbf{d}(\mathbf{u}) \quad \forall \mathbf{u}\in\mathcal{U}\right\} \tag{7.8}$$

where the only difference to the preceding sections is that now we have to assume that the matrix B (whose k rows are denoted by $\mathbf{b}_i^\mathsf{T}$) is not affected by uncertainty. This assumption of *fixed recourse* prevents occurrence of third order terms when implementing the quadratic decision rule. The i-th constraint of (7.8) is given by

$$(\mathsf{A}_i\mathbf{u})^\mathsf{T}\mathbf{x}+\mathbf{b}_i^\mathsf{T}\mathbf{y}(\mathbf{u})+\mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u}+\mathbf{d}_i^\mathsf{T}\mathbf{u} \geq 0 \quad \forall \mathbf{u}\in\mathcal{U}.$$

Now the entries of $\mathbf{y}(\mathbf{u})$ will be restricted to be quadratic functions in $\mathbf{u}$:

$$\mathbf{y}(\mathbf{u}) = \begin{pmatrix} \mathbf{u}^\mathsf{T}\mathsf{Y}_1\mathbf{u}+\mathbf{y}_1^\mathsf{T}\mathbf{u}+y_1 \\ \cdots \\ \mathbf{u}^\mathsf{T}\mathsf{Y}_{n_2}\mathbf{u}+\mathbf{y}_{n_2}^\mathsf{T}\mathbf{u}+y_{n_2} \end{pmatrix},$$

so that the robust constraint can be written as

$$\mathbf{u}^\mathsf{T}\left[\sum_{j=1}^{n_2} b_{ij}\mathsf{Y}_j+\mathsf{D}_i\right]\mathbf{u}+\left[\sum_{j=1}^{n_2} b_{ij}\mathbf{y}_j+\mathsf{A}_i^\mathsf{T}\mathbf{x}+\mathbf{d}_i\right]^\mathsf{T}\mathbf{u}+\mathbf{a}_i^\mathsf{T}\mathbf{x}+\sum_{j=1}^{n_2} b_{ij}y_j+d_i \geq 0 \quad \forall \mathbf{u}\in\mathcal{U},$$

which is of the form (7.1) so that it is amenable to the general reformulation strategy.

Theorem 30. *Assume $\mathcal{U}\subseteq\mathbb{R}^m$ is a bounded set such that for any $\mathsf{Q}\in\mathcal{S}^m$, $\mathbf{q}\in\mathbb{R}^m$, $\omega\in\mathbb{R}$ we have*

$$\inf_{\mathbf{u}\in\mathcal{U}}\left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u}+\mathbf{q}^\mathsf{T}\mathbf{u}+\omega\right] = \inf_{\mathsf{Y}\in\mathcal{C}}\left\{\mathsf{M}(\omega,\mathbf{q},\mathsf{Q})\bullet\mathsf{Y}\colon\ \mathsf{G}_l\bullet\mathsf{Y}\leq g_l\,,\ l\in[1\colon m]\right\},$$

and the dual of the right hand side problem attains its optimum. Then (7.7) with $(\mathbf{y}(\mathbf{u}))_j=\mathbf{u}^\mathsf{T}\mathsf{Y}_j\mathbf{u}+\mathbf{y}_j^\mathsf{T}\mathbf{u}+y_j$, $j\in[1\colon n_2]$ can be reformulated as

$$\begin{aligned} &\min_{\mathbf{x},\mathsf{Y}_j,\mathbf{y}_j,y_j,\boldsymbol{\lambda}_i} \mathbf{c}^\mathsf{T}\mathbf{x} \\ &\text{s.t. : } \mathbf{g}^\mathsf{T}\boldsymbol{\lambda}_i \leq 0, && i\in[1\colon k], \\ &\begin{pmatrix} \mathbf{a}_i^\mathsf{T}\mathbf{x}+\sum_{j=1}^{n_2} b_{ij}y_j+d_i & \frac{1}{2}\left(\sum_{j=1}^{n_2} b_{ij}\mathbf{y}_j+\mathsf{A}_i^\mathsf{T}\mathbf{x}+\mathbf{d}_i\right)^\mathsf{T} \\ \frac{1}{2}\left(\sum_{j=1}^{n_2} b_{ij}\mathbf{y}_j+\mathsf{A}_i^\mathsf{T}\mathbf{x}+\mathbf{d}_i\right) & \mathsf{D}_i+\sum_{j=1}^{n_2} b_{ij}\mathsf{Y}_j \end{pmatrix}+\sum_{l=1}^{m}\lambda_l^i\mathsf{G}_l\in\mathcal{C}^*, && i\in[1\colon k], \\ &\boldsymbol{\lambda}_i\in\mathbb{R}^m_+, && i\in[1\colon k], \\ &\mathsf{Y}_j\in\mathcal{S}^m,\ \mathbf{y}_j\in\mathbb{R}^{n_2},\ y_j\in\mathbb{R}, && j\in[1\colon n_2], \\ &\mathbf{x}\in\mathcal{X}. \end{aligned}$$

Proof. The theorem follows immediately from the discussion in Section 7.2 and the stucture of the semi-infinite constraint

$$\mathbf{u}^\mathsf{T}\left[\sum_{j=1}^{n_2} b_{ij}\mathsf{Y}_j + \mathsf{D}_i\right]\mathbf{u} + \left[\sum_{j=1}^{n_2} b_{ij}\mathbf{y}_j + \mathsf{A}_i^\mathsf{T}\mathbf{x} + \mathbf{d}_i\right]^\mathsf{T}\mathbf{u} + \mathbf{a}_i^\mathsf{T}\mathbf{x} + a_i + \sum_{j=1}^{n_2} b_{ij}y_j + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}.$$

□

7.4.2 Convex quadratic robust optimization

We will now discuss cases where the function in the robust constraint will be a concave quadratic function in $\mathbf{x}$. We will review the single stage case and then discuss various ways in which an adjustable second stage can be modeled.

7.4.2.1 Reformulation of the single stage case

The model of interest here is

$$\min_{\mathbf{x}\in\mathcal{X}} \left\{\mathbf{c}^\mathsf{T}\mathbf{x}\colon\ -\left\|\mathsf{A}_i(\mathbf{x})\mathbf{u}\right\|^2 + \left(\mathbf{a}_i(\mathbf{x})\right)^\mathsf{T}\mathbf{u} + a_i(\mathbf{x}) + \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}\,,\ i \in [1\!:\!k]\right\}, \tag{7.9}$$

where $\mathsf{A}_i(\mathbf{x})$ and $\mathbf{a}_i(\mathbf{x})$ are affine matrix and vector pencils respectively, and $a_i(\mathbf{x})$ is a real-valued, affine function of $\mathbf{x}$. This case was recently adressed in [46] in a way similar to the approach presented here. We will slightly generalize their result. It is clear that we can reformulate (7.9) as

$$\begin{aligned} \mathbf{b}^\mathsf{T}\boldsymbol{\lambda} &\leq 0, \\ \mathsf{M}\left(d_i + a_i(\mathbf{x}),\ \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\ -\mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x}) + \mathsf{D}_i\right) + \sum_{l=1}^{m}\lambda_l\mathsf{G}_l &\in \mathcal{C}^*. \end{aligned} \tag{7.10}$$

The entries of the south-east diagonal block of the constraints matrix in (7.10) are now quadratic functions. In case $\mathcal{C}^* = \mathcal{COP}(\mathcal{K})$ for some cone $\mathcal{K} \subseteq \mathbb{R}^m$(which is the case for all the conic reformulations of QCQPs discussed in this text) we can linearize the constraints by employing the following lemma, which is a straightforward generalization of [46, Lemma 4].

Lemma 31. *Assume $\mathcal{C}^* = \mathcal{COP}(\mathcal{K})$ for some cone $\mathcal{K} \subseteq \mathbb{R}^m$. Then, a vector $\mathbf{x} \in \mathbb{R}^n$ fulfills the conic constraint in (7.10) if and only if there exists a matrix $\mathsf{H} \in \mathcal{S}^m$ such that*

$$\mathsf{M}\left(d_i + a_i(\mathbf{x}),\, \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\, -\mathsf{H} + \mathsf{D}_i\right) + \sum_{l=1}^{m}\lambda_l\mathsf{G}_l \in \mathcal{COP}(\mathcal{K}) \quad \text{and} \tag{7.11}$$

$$\mathsf{M}\left(\mathsf{H},\, 2\mathsf{A}_i(\mathbf{x}),\, \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m}. \tag{7.12}$$

Proof. Assume that $\mathbf{x}$ fulfills (7.10), then obviously it fulfills (7.11) and (7.12) by setting $\mathsf{H} = \mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x})$, which implies $\mathsf{H} \succeq \mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x})$ which by Schur complementation implies (7.12). So assume $\mathbf{x}$ and H fulfill (7.11) and (7.12). Define

$$\mathsf{U} \coloneqq \mathsf{M}\left(d_i + a_i(\mathbf{x}),\, \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\, -\mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x}) - \mathsf{D}_i\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l,$$
$$\mathsf{V} \coloneqq \mathsf{M}\left(d_i + a_i(\mathbf{x}),\, \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\, -\mathsf{H} + \mathsf{D}_i\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l.$$

Then

$$\begin{aligned}
&\begin{pmatrix} w \\ \mathbf{u} \end{pmatrix}^\mathsf{T} (\mathsf{U}-\mathsf{V}) \begin{pmatrix} w \\ \mathbf{u} \end{pmatrix} = \\
=&\begin{pmatrix} w \\ \mathbf{u} \end{pmatrix}^\mathsf{T} \left[\begin{pmatrix} d_i + a_i(\mathbf{x}) & \frac{1}{2}\left[\mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i\right]^\mathsf{T} \\ \frac{1}{2}\left[\mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i\right] & -\mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x}) + \mathsf{D}_i \end{pmatrix} - \begin{pmatrix} d_i + a_i(\mathbf{x}) & \frac{1}{2}\left[\mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i\right]^\mathsf{T} \\ \frac{1}{2}\left[\mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i\right] & \mathsf{D}_i - \mathsf{H} \end{pmatrix}\right] \begin{pmatrix} w \\ \mathbf{u} \end{pmatrix} = \\
=&\mathbf{u}^\mathsf{T}\left[\mathsf{H} - \mathsf{A}_i(\mathbf{x})^\mathsf{T}\mathsf{A}_i(\mathbf{x})\right]\mathbf{u} \geq 0 \quad \forall (\mathbf{u}, w).
\end{aligned}$$

where the inequality holds due to (7.12). But since $\mathsf{V} \in \mathcal{COP}(\mathcal{K})$, $\mathsf{U}-\mathsf{V} \in \mathcal{S}^n_+$ implies $\mathsf{U} \in \mathcal{COP}(\mathcal{K})$ for any cone $\mathcal{K}$, we have shown that (7.10) holds as well. □

Using this lemma we can derive the following theorem

Theorem 32. *Assume $\mathcal{U} \subseteq \mathbb{R}^m$ is a bounded set such that for any $\mathsf{Q} \in \mathcal{S}^m$, $\mathbf{q} \in \mathbb{R}^m$, $\omega \in \mathbb{R}$ we have*

$$\inf_{\mathbf{u}\in\mathcal{U}} \left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega\right] = \inf_{\mathsf{Y}\in\mathcal{C}} \left\{\mathsf{M}(\omega, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{Y} \colon \mathsf{G}_l \bullet \mathsf{Y} \leq g_l\,,\, l \in [1\!:\!m]\right\},$$

where $\mathcal{C}^* = \mathcal{COP}(\mathcal{K})$ for some cone $\mathcal{K} \subseteq \mathbb{R}^m$ and the dual of the right hand side problem attains its optimum. Then (7.9) can be reformulated as

$$\begin{aligned}
\min_{\mathbf{x},\boldsymbol{\lambda}_i,\mathsf{H}}\ & \mathbf{c}^\mathsf{T}\mathbf{x} \\
\text{s.t. : } & \mathbf{g}^\mathsf{T}\boldsymbol{\lambda}_i \leq 0, && i \in [1\!:\!k], \\
\mathsf{M}\left(d_i + a_i(\mathbf{x}),\, \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\, -\mathsf{H} + \mathsf{D}_i\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l & \in \mathcal{COP}(\mathcal{K}), && i \in [1\!:\!k], \\
\mathsf{M}\left(\mathsf{H},\, 2\mathsf{A}_i(\mathbf{x}),\, \mathsf{I}_n\right) & \in \mathcal{S}^{n+m}_+, && i \in [1\!:\!k], \\
\boldsymbol{\lambda}_i & \in \mathbb{R}^m_+, && i \in [1\!:\!k], \\
\mathbf{x} & \in \mathcal{X}.
\end{aligned}$$

Proof. The proof is immediate from the above discussion and the discussion in Section 7.2. □

7.4.2.2 Extension to the adjustable robust setting

The strategy above can immediately be transferred to adjustable robust constraints, where the first stage decision enters convex-quadratically. The second stage decision can enter the constraints in several different ways. It may enter linearly but with uncertain recourse. In this case, an affine policy has to be introduced, such that the constraint is overall quadratic in $\mathbf{u}$. It is also possible to introduce a quadratic decision rule, but then recourse has to be fixed as well. A third possibility is given by having the second stage decision enter the constraint convex-quadratically. There are two ways to do this: either with or without interaction with the first stage decision. In any case, the recourse has to be fixed in order to obtain a constraint that is quadratic in $\mathbf{u}$. The following proposition summarizes these cases and states their respective set-copositive reformulations

Proposition 33. *Assume $\mathcal{U} \subseteq \mathbb{R}^m$ is a bounded set such that for any $\mathsf{Q} \in \mathcal{S}^m$, $\mathbf{q} \in \mathbb{R}^m$, $\omega \in \mathbb{R}$ we have*

$$\inf_{\mathbf{u}\in\mathcal{U}} \left[\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} + \mathbf{q}^\mathsf{T}\mathbf{u} + \omega\right] = \inf_{\mathsf{Y}\in\mathcal{C}} \left\{\mathsf{M}(\omega, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{Y} \colon \mathsf{G}_l \bullet \mathsf{Y} \leq g_l \,,\, l \in [1 \colon m]\right\},$$

where $\mathcal{C}^* = \mathcal{COP}\left(\mathcal{K}\right)$ for some cone $\mathcal{K} \subseteq \mathbb{R}^m$ and the dual of the right hand side problem attains its optimum. Then the following list represents equivalent pairs of constraints:

1. *Linear in the second stage with affine decision and uncertain recourse*

$$-\|\mathsf{A}_i(\mathbf{x})\mathbf{u}\|^2 + (\mathsf{B}_i\mathbf{u} + \mathbf{b}_i)^\mathsf{T}(\mathsf{Y}\mathbf{u} + \mathbf{y}_0) + \mathbf{a}_i^\mathsf{T}(\mathbf{x})\mathbf{u} + a_i(\mathbf{x}) + \mathbf{u}^\mathsf{T}\ \mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U},$$

$$\Updownarrow$$

$$\mathsf{M}\left(d_i + a_i(\mathbf{x}) + \mathbf{b}_i^\mathsf{T}\mathbf{y}_0,\, \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i + \mathsf{B}_i^\mathsf{T}\mathbf{y}_0,\, \mathsf{D}_i - \mathsf{H} + \frac{1}{2}[\mathsf{B}_i^\mathsf{T}\mathsf{Y} + \mathsf{Y}^\mathsf{T}\mathsf{B}_i]\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l \in \mathcal{COP}(\mathcal{K}),$$
$$\mathsf{M}\left(\mathsf{H},\, 2\mathsf{A}_i(\mathbf{x}),\, \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m}.$$

2. *Linear in the second stage with quadratic decision rule and fixed recourse*

$$-\|\mathsf{A}_i(\mathbf{x})\mathbf{u}\|^2 + \sum_{j=1}^{n_2} b_{ji}(\mathbf{u}^\mathsf{T}\mathsf{Y}_j\mathbf{u} + \mathbf{u}^\mathsf{T}\mathbf{y}_j + y_j) + \mathbf{a}_i^\mathsf{T}(\mathbf{x})\mathbf{u} + a_i(\mathbf{x}) + \mathbf{u}^\mathsf{T}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^\mathsf{T}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}$$

$$\Updownarrow$$

$$\sum_{j=1}^{n_2} b_{ji}\mathsf{M}(y_j, \mathbf{y}_j, \mathsf{Y}_j) + \mathsf{M}(d_i + a_i(\mathbf{x}), \mathbf{a}_i^T(\mathbf{x}) + \mathbf{d}_i, \mathsf{D}_i - \mathsf{H}) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l \in \mathcal{COP}(\mathcal{K}),$$
$$\mathsf{M}\left(\mathsf{H},\ 2\mathsf{A}_i(\mathbf{x}),\ \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m}.$$

3. *Quadratic in the second stage with interaction effect, linear decision rule and fixed recourse*

$$-\|\mathsf{A}_i(\mathbf{x})\mathbf{u} + \mathsf{B}\mathsf{Y}\mathbf{u}\|^2 + \mathbf{a}_i^{\mathsf{T}}(\mathbf{x})\mathbf{u} + a_i(\mathbf{x}) + \mathbf{u}^{\mathsf{T}}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^{\mathsf{T}}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U},$$

$$\Updownarrow$$

$$\mathsf{M}\left(d_i + a_i(\mathbf{x}), \mathbf{a}_i^{\mathsf{T}}(\mathbf{x}) + \mathbf{d}_i, \mathsf{D}_i - \mathsf{H}\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l \in \mathcal{COP}(\mathcal{K}),$$
$$\mathsf{M}\left(\mathsf{H}, 2[\mathsf{A}_i(\mathbf{x}) + \mathsf{B}_i\mathsf{Y}], \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m}.$$

4. *Quadratic in the second stage without interaction effect, linear decision rule and fixed recourse*

$$-\|\mathsf{A}_i(\mathbf{x})\mathbf{u}\|^2 - \|\mathsf{B}_i(\mathsf{Y}\mathbf{u})\|^2 + \mathbf{a}_i^{\mathsf{T}}(\mathbf{x})\mathbf{u} + a_i(\mathbf{x}) + \mathbf{u}^{\mathsf{T}}\mathsf{D}_i\mathbf{u} + \mathbf{d}_i^{\mathsf{T}}\mathbf{u} + d_i \geq 0 \quad \forall \mathbf{u} \in \mathcal{U},$$

$$\Updownarrow$$

$$\mathsf{M}\left(d_i + a_i(\mathbf{x}),\ \mathbf{a}_i(\mathbf{x}) + \mathbf{d}_i,\ \mathsf{D}_i - \mathsf{H}_1 - \mathsf{H}_2\right) + \sum_{l=1}^{m} \lambda_l \mathsf{G}_l \in \mathcal{COP}(\mathcal{K}),$$
$$\mathsf{M}\left(\mathsf{H}_1,\ 2\mathsf{A}_i(\mathbf{x}),\ \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m},$$
$$\mathsf{M}\left(\mathsf{H}_2, 2\mathsf{B}_i\mathsf{Y}, \mathsf{I}_n\right) \in \mathcal{S}_+^{n+m}.$$

Proof. The proof is obvious from our general reformulation strategy plus straightforward adaptations of Lemma 31. □

In order to avoid bilinear terms in the offdiagonal blocks of the matrices in the reformulation in 3. and 4., it is here necessary to restrict the second-stage decision to a linear rather than allowing for an affine policy.

In the models described above it is important not to confuse the functional form of the second-stage decision with the functional form of its effect on any given constraint. The latter is a design choice, namely whether the second-stage decision enters linearly or convex quadratically into any given constraint, which is determined by the model of the problem. The former however concerns a decision variable we need to optimize, where restricting it to affine or quadratic functions is merely a simplification we undertake to render the problem tractable.

7.4.3 Numerical experiments

In this section we will apply the theory developed in this chapter to illustrate its usefulness in improving existing techniques and for solving robust optimization problems.

7.4.3.1 Quadratic decision rules versus affine policy

We will now present numerical evidence that the quadratic decision rule may outperform the affine policy, and that modifications of the ellipsoidal uncertainty set may lead to an improvement of the optimal value due to a reduced degree of conservativeness. All computations were performed using the YALMIP interface with SDPT3 as solver and were carried out on a system with an Intel Core i7-6700 2x3.40 GHz processor and 32GB RAM. The instances were created by randomly generating $\mathsf{A}_i \in \mathbb{R}^{m\times n_1}, \mathbf{b}_i \in \mathbb{R}^{n_2}, \mathbf{c} \in \mathbb{R}^{n_1}, \mathsf{D} \in \mathbb{R}^{m\times m}, \mathsf{G} \in \mathbb{R}^{n_2\times n_2}, \mathbf{s} \in \mathbb{R}^{n_2}$ to set up the problem

$$\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^{n_1}_+,\mathbf{y}(u)} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\
\mathbf{u}^{\mathsf{T}}\mathsf{A}_i\mathbf{x} + \mathbf{b}_i^{\mathsf{T}}\mathbf{y}(\mathbf{u}) \geq \mathbf{u}^{\mathsf{T}}\mathsf{D}\mathbf{u}, \quad & \forall \mathbf{u} \in \mathcal{U},\ i \in [1\!:\!k], \\
\mathsf{G}\mathbf{y}(\mathbf{u}) \leq \mathbf{s}, \quad & \forall \mathbf{u} \in \mathcal{U}, \\
\mathbf{y}(\mathbf{u}) \geq \mathbf{o}, \quad & \forall \mathbf{u} \in \mathcal{U}.
\end{aligned} \tag{7.13}$$

where by default $\mathcal{U} = \{\mathbf{u} \in \mathcal{K} \colon u_1 = 1\}$ for our particular choices of $\mathcal{K}$. The generation of the data was set up such that the coefficients of the first and second stage variables were positive in each of the first k constraints. Also the objective function coefficients were positive and the the right-hand sides were positive for all realizations of $\mathbf{u} \in \mathcal{U}$. The second set of constraints can be seen as resource constraints (i.e. $\mathbf{s}$ is also positive), where the matrix G ensures, that the second stage variables compete for recourses. We used three types of uncertainty sets with corresponding choices for $\mathcal{K}$ given by $\mathcal{K}_1 \coloneqq \{\mathbf{x} \in \mathbb{R}^m \colon \mathbf{x}^{\mathsf{T}}\mathsf{L}\mathbf{x} \leq 0, \mathbf{e}_1^{\mathsf{T}}\mathbf{x} \geq 0\} = \mathcal{SOC}^m$, $\mathcal{K}_2 \coloneqq \{\mathbf{x} \in \mathbb{R}^m \colon \mathbf{x}^{\mathsf{T}}\mathsf{L}\mathbf{x} \leq 0, \mathbf{x}^{\mathsf{T}}\mathsf{A}\mathbf{x} \leq 0, \mathbf{e}_1^{\mathsf{T}}\mathbf{x} \geq 0\}$, $\mathcal{K}_3 \coloneqq \{\mathbf{x} \in \mathbb{R}^m \colon \mathbf{x}^{\mathsf{T}}\mathsf{L}\mathbf{x} \leq 0\,, \mathsf{C}\mathbf{x} \geq \mathbf{o}\,, \mathbf{e}_1^{\mathsf{T}}\mathbf{x} \geq 0\}$, where $\mathsf{L} = \mathrm{Diag}(-1, 1, \dots, 1)$, $\mathsf{A} = \mathrm{Diag}(0.5, -1, \dots, -1)$ and $\mathsf{C}^{\mathsf{T}} = [\frac{1}{2}\mathbf{e}_1 + \mathbf{e}_2, \frac{1}{2}\mathbf{e}_1 - \mathbf{e}_2]$. These cones are the second order cone; a second order cone from which a smaller copy of it is removed; and a second order cone intersected with two half spaces such that two opposite parts of the cone are removed. The resulting three types of uncertainty sets are: the unit ball, the unit ball with a spherical hole in the center and the unit ball where two opposite caps were removed (i.e., an ellipsoids with two cuts).

Table 7.1 reports average results from 100 randomly generated instances of (7.13) where we investigate the gap between the objective function values under an affine

instance size	ellipsoid	ellipsoid with hole	ellipsoid with two cuts
(n_1, n_2, k, m)	avg. rel. gap (#nz)	avg. rel. gap (#nz)	avg. rel. gap (#nz)
(5,2,2,2)	21.55% (36)	33.22% (45)	24.98% (46)
(15,3,5,3)	48.34% (79)	44.49% (78)	40.33% (69)
(15,8,8,8)	60.01% (100)	61.60% (97)	57.51% (99)
(30,10,10,10)	46.99% (100)	56.28% (97)	48.17% (99)

Table 7.1: Affine vs. quadratic decision rule

policy versus under a quadratic decision rule. The heading "instance size" refers to numbers of the first stage and second stage decisions, as well as the number of constraints and the dimension of the uncertainty set. We report average relative gap, i.e., the average improvement of the quadratic decision rule as a percentage over the optimal value under affine decision rule. The number in the brackets states the number of instances with a positive gap. The table suggests that the possible benefit from employing the quadratic policy can be substantial but not necessarily so. Depending on the parameters we were able to generate instances where the gain was zero or negligible but also instances with noticeable gain. Overall we observed relative gaps from anywhere between 0% up to 100% (i.e. the quadratic policy achieved the 'hard' lower bound of 0, where the requirements of the first set of constraints are met entirely in the second stage). Recall that this data represents a sample of instances which were set up in the specific way outlined above. No inferences about the correlation between the degree of superiority of the quadratic decision rule and the structure of the uncertainty set should be made based on this sample alone. However, providing conditions under which superiority (or even optimality) of the quadratic decision rule is guaranteed is an interesting area of future research.

7.4.3.2 An adjustable robust model for ressource purchase

We will now resort to a simple model concerning ressource purchase in order to illustrate the usefulness of our approach. Suppose you want to produce products $i = [1:l]$ by processing resources $j = [1:n]$. Production should start immediately but only limited amounts of resources are available right now. In a second stage, additional resources will be available for purchase, but prices may fluctuate for exogenous reasons, as does the budget for purchasing. Let z_i denote amount of product i produced and x_j, y_j the amount of resource j purchased in the first and second stage respectively. Further $p_j, \tilde{p}_j$ are the prices of resource j in the first and second stage respectively. The contribution margin of product i is given by $c_i > 0$ and $a_{ij} > 0$ give the amount of resource j consumed by production of one unit of product i. Finally s_j denotes the maximum amount of resource j available in the first stage and $\tilde{b}$ denotes budget for resource purchase over the planning horizon. Since $\tilde{p}_j$ and $\tilde{b}$ are subject to uncertainty, both will be modeled as functions depending on an uncertainty vector $\mathbf{u}$ which lives in an uncertainty set $\mathcal{U} := \mathcal{K} \cap \mathcal{H}$. As the problem at hand is a textbook example of linear programming, we omit stating the deterministic version of the decision problem an turn to the adjustable robust counterpart, which is given by

$$
\begin{aligned}
\max_{\mathbf{x},\mathbf{y}(\cdot)} \quad & \mathbf{c}^\mathsf{T}\mathbf{z} \\
\text{s.t.}: \mathbf{p}^\mathsf{T}\mathbf{x} + \tilde{\mathbf{p}}(\mathbf{u})^\mathsf{T}\mathbf{y}(\mathbf{u}) &\leq \tilde{b}(\mathbf{u}), && \forall \mathbf{u} \in \mathcal{U}, \\
\mathsf{A}\mathbf{z} &\leq \mathbf{x} + \mathbf{y}(\mathbf{u}), && \forall \mathbf{u} \in \mathcal{U}, \\
\mathbf{x} &\leq \mathbf{s}, \\
\mathbf{z} \geq \mathbf{o},\ \mathbf{x} \geq \mathbf{o},\ \mathbf{y}(\mathbf{u}) &\geq \mathbf{o}, && \forall \mathbf{u} \in \mathcal{U}.
\end{aligned}
$$

Specifying $\tilde{\mathbf{p}} := \mathsf{P}\mathbf{u}$ and $\tilde{b}(\mathbf{u}) := \mathbf{u}^\mathsf{T}\mathsf{B}\mathbf{u}$ and employing the reformulations steps as well as the affine policy as before yield and equivalent linear conic optimization problem. We consider three uncertainty sets of the form $\mathcal{U}_i = \{\mathbf{u} \in \mathcal{K}_i : \mathbf{e}_1^\mathsf{T}\mathbf{u} = 1\}$, $i \in \{1, 2, 3\}$, where $\mathcal{K}_1$, and $\mathcal{K}_2$ are defined as in the previous section and $\mathcal{K}_3 := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\mathsf{T}\mathsf{L}\mathbf{x} \leq 0, \mathbf{e}_1^\mathsf{T}\mathbf{x} \geq 0, \mathbf{e}_2^\mathsf{T}\mathbf{x} \geq 0\}$. The third cone is now a second order cone sliced in half along the second coordinate, and the uncertainty set is half a unit ball.

The choice of uncertainty sets here is illustrative in the following sense. The ellipsoidal uncertainty is a standard way of modeling uncertainty but due to additional information about the uncertain outcome we might be able to exclude certain possibilities from consideration. Thus, one might wish to cut away bits and pieces of the uncertainty set, such that the worst case may be tendered. Of course reducing the uncertainty gives rise to less conservative optimal robust strategies. An interesting application of such a modeling choice could be a setting where we have countermeasures available to prevent particular outcomes of the uncertain parameters (e.g. buying options to hedge against fluctuations in resource prices). The cost of such a countermeasure could be weighed against the gain in objective function value when employing an appropriately cut uncertainty set. Working out the details of such a strategy is an interesting direction for future research. However, for now we will merely give a proof of concept, by demonstrating that our toy example behaves as expected. For this purpose we have created three instances each with the certain data given by $\mathbf{c} = (1, 1)^\mathsf{T}, \mathsf{A} = (1, 1; 1, 1)$ and $\mathbf{p} = (5, 5)^\mathsf{T}$. The uncertain data depends on an uncertainty vector of dimension $m = 3$ where the first coordinate was fixed to

	instance 1	**instance 2**	**instance 3**
P	$\begin{pmatrix} 5 & -1 & 1 \\ 5 & -1 & 1 \end{pmatrix}$	$\begin{pmatrix} 5 & 1 & -1 \\ 5 & 1 & -1 \end{pmatrix}$	$\begin{pmatrix} 5 & 1 & 1 \\ 5 & 1 & 1 \end{pmatrix}$
$\mathbf{b}^{\mathsf{T}}$	$(500, 100, -100)$	$(500, -100, 100)$	$(500, 100, 100)$
quadratic	no	no	yes
$\mathcal{K}_1$	24.27	24.27	36.29
$\mathcal{K}_2$	24.27	24.27	38.06
$\mathcal{K}_3$	27.5	24.27	36.29
nominal case	37.5	37.5	37.5

Table 7.2: Comparison of different uncertainty sets

1. The first column of P and the first entry of $\mathbf{b}$ give the nominal cases, where the second and third column/entry give the magnitude and the direction of variations of $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{b}}$ as $\mathbf{u}$ varies within the uncertainty set. For each instance, P and $\mathbf{b}$ are given in the respective column of Table 7.2.

The first instance was set up such that the worst case would materialize at the boundary of the "lower" half of the unit disc that is the uncertainty set. Cutting away the center of that disk thus does not change the optimal value, but cutting away the lower half does. In the second instance, the worst case can be found at the boundary of the upper half of the uncertainty set, so that neither curtailing has an effect. In the third instance, by means of positive definite choice for $\mathsf{B} \in \mathcal{S}^n_+$, the situation was set up in a way, that the uncertainty will mostly work in our favor, so was indeed the worst case was very close the nominal case in the center of the unit disk. Thus cutting away the center consequently resulted in the highest objective function value.

7.5 An alternative perspective via set-copositivity

In the following section we want to discuss the fact that the semi-infinite constraints, which arise in robust and adjustable robust optimization, can be interpreted as set-copositivity constraints. Careful analysis of these constraints will reveal that they can be decomposed into linear and copositivity constraints of simpler structure. This decomposition will allow for some interesting insights in the following sense. We have seen so far, that we can use quadratic optimization tools in conjunction with conic duality in order to reformulate semi-infinite constraints that quadratically depend on an index. However, if we can merely close the duality gap without at the same time guaranteeing for dual attainability, the feasible set described by the reformulation shrinks. The decomposition will reveal more details on the nature of this shrinkage. We will discuss this phenomenon in the context of using results in [25, 41] for the reformulation of semi-infinite constraints. Let us reconsider the semi-infinite constraints we have encountered in (7.2):

$$\mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u}+\mathbf{q}^\mathsf{T}\mathbf{u}+\omega\geq 0 \quad \forall \mathbf{u}\in\mathcal{U}\subseteq\mathbb{R}^n. \tag{7.14}$$

The quadratic inequality can be homogenized in order to obtain a more compact formulation, i.e.,

$$\begin{pmatrix} t \\ \mathbf{u}\end{pmatrix}^\mathsf{T}\begin{pmatrix}\omega & \frac{1}{2}\mathbf{q}^\mathsf{T} \\ \frac{1}{2}\mathbf{q} & \mathsf{Q}\end{pmatrix}\begin{pmatrix} t \\ \mathbf{u}\end{pmatrix}\geq 0 \quad \forall\,(t,\mathbf{u})^\mathsf{T}\in\{1\}\times\mathcal{U}\subseteq\mathbb{R}^{n+1}\,.$$

By the definition of set-copositivity the above constraint is equivalent to

$$\mathsf{M}(\omega,\mathbf{q},\mathsf{Q})\in\mathcal{COP}(\mathrm{cone}[\{1\}\times\mathcal{U}])\subseteq\mathcal{S}^{n+1}. \tag{7.15}$$

We thus see that the copositivity constraint (7.15) can also be enforced by (7.3). Further, from $\mathrm{cone}\,[\{1\}\times(\cup_i\mathcal{U}_i)] = \cup_i\mathrm{cone}\,[\{1\}\times\mathcal{U}_i]$ together with $\mathcal{COP}(\cup_i\mathcal{U}_i)=\bigcap_i\mathcal{COP}(\mathcal{U}_i)$ we see that

$$\mathcal{COP}\left(\mathrm{cone}\,[\{1\}\times(\cup_i\mathcal{U}_i)]\right)=\cap_i\mathcal{COP}\left(\mathrm{cone}\,[\{1\}\times\mathcal{U}_i]\right). \tag{7.16}$$

Thus, as a side note, we can handle unions of sets $\mathcal{U}=\cup_i\mathcal{U}_i$ in (7.2) whenever we can characterize $\mathcal{COP}\left(\mathrm{cone}\,[\{1\}\times\mathcal{U}_i]\right)$ for each individual $\mathcal{U}_i$. However, the complicated structure of the set over which we we require the quadratic form to be non-negative seems to be prohibitive. Nevertheless, this issue can be resolved quite easily. Assume that $\mathcal{U}$ is the compact intersection of a (not necessarily convex) cone $\mathcal{K}$ with a hyperplane:

$$\mathcal{U}\coloneqq\{\mathbf{u}\in\mathcal{K}\colon\mathbf{b}^\mathsf{T}\mathbf{u}=c\}\,.$$

We will prove a result which states that the seemingly complicated copositivity constraint in (7.15) can be replaced by a linear constraint and a copositivity constraint over $\mathcal{K}$. The key idea is that $\mathrm{cone}[\{\mu\}\times(\mathcal{K}\cap\mathcal{H})]$ is actually $\{0\}\times\mathcal{K}$ under a linear transformation. To this end, we need some auxiliary results. First, for any cone $\mathcal{K}\subseteq\mathbb{R}^n$ observe that

$$\mathcal{COP}(\{0\}\times\mathcal{K})=\{\mathsf{M}(\mathsf{R},\mathbf{f},b)\in\mathcal{S}^n\colon\mathsf{R}\in\mathcal{COP}(\mathcal{K})\,,\,\mathbf{f}\in\mathbb{R}^n\,,\,b\in\mathbb{R}\}\,, \tag{7.17}$$

and for any nonsingular $\mathsf{A}\in\mathbb{R}^{n\times n}$

$$\mathcal{COP}(\mathsf{A}\mathcal{K})=\{\mathsf{A}^{-\mathsf{T}}\mathsf{R}\mathsf{A}^{-1}\in\mathcal{S}^n\colon\mathsf{R}\in\mathcal{COP}(\mathcal{K})\}\,. \tag{7.18}$$

Lemma 34. *Let $\mu \in \mathbb{R}$ and $\mathcal{K}$ be a closed cone with pointed convex hull and $\mathcal{H} = \{\mathbf{u} \in \mathbb{R}^n \colon \mathbf{a}^\mathsf{T}\mathbf{u} = b\}$ be a hyperplane such that $\mathcal{H} \cap \operatorname{conv}(\mathcal{K})$ is compact and contains a point $\mathbf{u} \neq \mathbf{o}$. Then, there exists an invertible matrix T of order $n+1$ such that*

$$\operatorname{cone}[\{\mu\} \times (\mathcal{K} \cap \mathcal{H})] = \mathsf{T}(\{0\} \times \mathcal{K})\,.$$

Proof. Without loss of generality we can assume that $\mathbf{a} = \mathbf{e}_1$ and $b = 1$ (so $\mathbf{u} \in \mathcal{H}$ if and only if $u_1 = 1$). Choose T as the following square matrix of order $n+1$:

$$\mathsf{T} \coloneqq \mathsf{I}_{n+1} + \mathbf{v}\mathbf{e}_1^\mathsf{T} = \begin{pmatrix} 1 & 0 & \mathbf{o}^\mathsf{T} \\ \mu & 1 & \mathbf{o}^\mathsf{T} \\ \mathbf{o} & \mathbf{o} & \mathsf{I}_{n-1} \end{pmatrix} \quad \text{with} \quad \mathbf{v} = \begin{pmatrix} \mu \\ \mathbf{o} \end{pmatrix} \in \mathbb{R}^{n+1}\,.$$

Then, for any $\mathbf{u} \in \mathbb{R}^{n+1}$ we have $\mathsf{T}\mathbf{u} = \mathbf{u} + u_1\mathbf{v}$. Further $\mathsf{T}^{-1} = \mathsf{I}_{n+1} - \mathbf{v}\mathbf{e}_1^\mathsf{T}$ exists. We will now show that $\operatorname{cone}(\{\mu\} \times (\mathcal{K} \cap \mathcal{H})) = \mathsf{T}(\{0\} \times \mathcal{K})$. Now $\mathbf{o}$ is contained in both cones, so let $\mathbf{u} \neq \mathbf{o}$ henceforth. First, assume $\mathbf{u} \in \mathsf{T}(\{0\} \times \mathcal{K})$, so that there is an $\mathbf{u}_0 \in \{0\} \times \mathcal{K}$ with $\mathsf{T}\mathbf{u}_0 = \mathbf{u}$. By Lemma 35 we have an $\lambda \geq 0$ such that $\lambda\mathbf{u}_0 \in \{0\} \times (\mathcal{K} \cap \mathcal{H})$ and $\lambda\mathbf{u} = \lambda\mathbf{u}_0 + \mathbf{v} \in \{\mu\} \times (\mathcal{K} \cap \mathcal{H})$ and so $\mathbf{u} \in \operatorname{cone}(\{\mu\} \times (\mathcal{K} \cap \mathcal{H}))$. On the other hand, assume $\mathbf{u} \in \operatorname{cone}(\{\mu\} \times (\mathcal{K} \cap \mathcal{H}))$, then $\lambda\mathbf{u} \in \{\mu\} \times (\mathcal{K} \cap \mathcal{H})$ for some $\lambda \geq 0$ implying $\lambda\mathsf{T}^{-1}\mathbf{u} = \lambda\mathbf{u} - \mathbf{v} \in \{0\} \times (\mathcal{K} \cap \mathcal{H})$ and so $\mathsf{T}^{-1}\mathbf{u} \in \{0\} \times \mathcal{K}$, which gives $\mathbf{u} = \mathsf{T}\mathsf{T}^{-1}\mathbf{u} \in \mathsf{T}(\{0\} \times \mathcal{K})$. $\square$

Lemma 35. *Let $\mathcal{M} \coloneqq \{\mathbf{u} \in \mathbb{R}^n \colon \mathsf{A}\mathbf{u} = \mathbf{b}\}$ be an affine subspace where $\mathsf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m \setminus \{\mathbf{o}\}$. Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed cone with pointed convex hull $\overline{\mathcal{K}} \coloneqq \operatorname{conv}(\mathcal{K})$ such that $\overline{\mathcal{K}} \cap \mathcal{M}$ is compact. Then the following two statements hold.*

a) There are $\boldsymbol{\alpha} \in \operatorname{int}(\mathcal{K}^*)$ and $b_0 > 0$ such that $\mathcal{K} \cap \mathcal{M} = \{\mathbf{u} \in \mathcal{K} \colon \mathsf{A}\mathbf{u} = \mathbf{b},\ \boldsymbol{\alpha}^\mathsf{T}\mathbf{u} = b_0\}$.

b) If $m = 1$, then $\operatorname{cone}(\mathcal{M} \cap \mathcal{K}) = \mathcal{K}$.

Proof. Let $\mathcal{M}_0 \coloneqq \{\mathbf{u} \in \mathbb{R}^n \colon \mathsf{A}\mathbf{u} = \mathbf{o}\} = \ker \mathsf{A}$ denote the recession cone of $\mathcal{M}$. Then, by compactness, we have $\mathcal{M}_0 \cap \overline{\mathcal{K}} = \{\mathbf{o}\}$. We look for a supporting hyperplane of $\overline{\mathcal{K}}$ which meets this cone only at $\mathbf{o}$ (i.e., whose normal vector is interior to $\overline{\mathcal{K}}^*$) and which contains $\mathcal{M}_0$ (i.e., it is also a supporting hyperplane of $\mathcal{M}_0$). Such a hyperplane exists exactly if $\operatorname{int}(\overline{\mathcal{K}}^*) \cap \mathcal{M}_0^* \neq \emptyset$, which we will prove to be the case. Assume that $\operatorname{int}(\overline{\mathcal{K}}^*) \cap \mathcal{M}_0^* = \emptyset$, then by [73, Theorem 11.2] there is a hyperplane $\mathcal{H}$ which contains $\mathcal{M}_0^* = \{\mathsf{A}^\mathsf{T}\boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}^m\}$ while $\operatorname{int}(\overline{\mathcal{K}}^*)$ is contained in one of the open halfspaces associated with $\mathcal{H}$. We have $\mathbf{o} \in \mathcal{M}_0^* \subset \mathcal{H}$ so that $\mathcal{H}$ is a supporting hyperplane of $\overline{\mathcal{K}}^*$. This implies that one of the normal vectors of $\mathcal{H}$, say $\bar{\mathbf{a}}$, is an element of $\overline{\mathcal{K}}$. Note that $\mathcal{M}_0 = (\mathcal{M}_0)^{**} = (\mathcal{M}_0^*)^\perp$ (recall that $\mathcal{M}_0$ is a linear subspace) and hence

$\overline{\mathbf{a}} \in \mathcal{M}_0$. Thus, $\mathcal{M}$ and $\overline{\mathcal{K}}$ have a common direction of recession, which contradicts the compactness of $\overline{\mathcal{K}} \cap \mathcal{M}$. Since $\mathcal{K}^* = \overline{\mathcal{K}}^*$, indeed $\boldsymbol{\alpha} = \mathsf{A}^\mathsf{T}\boldsymbol{\lambda} \in \operatorname{int}(\mathcal{K}^*)$ for some appropriate $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\boldsymbol{\alpha}^\mathsf{T}\mathbf{u} = \boldsymbol{\lambda}^\mathsf{T}\mathsf{A}\mathbf{u} = \boldsymbol{\lambda}^\mathsf{T}\mathbf{b}$ is redundant for $\mathbf{u} \in \mathcal{M}$. Assume w.l.o.g. that $b_1 > 0$. Clearly if $m = 1$ we have $\mathbf{a}_1 \in \operatorname{int}(\mathcal{K}^*)$ so that for all $\mathbf{u} \in \mathcal{K}$ we have $\mathbf{a}_1^\mathsf{T}\mathbf{u} > 0$ and thus there is an $\lambda \geq 0$ such that $\mathbf{a}_1^\mathsf{T}(\lambda\mathbf{u}) = b_1$. This proves $\operatorname{cone}(\mathcal{M} \cap \mathcal{K}) \supseteq \mathcal{K}$ since $\mathbf{o}$ is in both cones. The converse is trivial. □

Remark. *The following procedure can be followed for explicitly determining $\boldsymbol{\alpha}$ and b_0, given $\mathsf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m \setminus \{\mathbf{o}\}$ and $\mathcal{K} \subset \mathbb{R}^n$ as in the existence result in Lemma 35. We first calculate $(\boldsymbol{\lambda}^*, \varepsilon^*) = \arg\max_{(\boldsymbol{\lambda},\varepsilon)\in\mathbb{R}^{m+1}} \left\{\varepsilon \colon (\mathsf{A}^\mathsf{T}\boldsymbol{\lambda})^\mathsf{T}\mathbf{v} \geq \varepsilon \ \forall \mathbf{v} \in \mathcal{K} \text{ with } \mathbf{a}^\mathsf{T}\mathbf{v} = 1\right\}$ where $\mathbf{a}$ is a known element of $\operatorname{int}(\mathcal{K}^*)$ (e.g. $\mathbf{a} = \mathbf{e}$ if $\mathcal{K}$ is the non-negative orthant or $\mathbf{a} = \mathbf{e}_1$ if $\mathcal{K}$ is the second order cone). Note that Lemma 35 guarantees the existence of an feasible solution $(\boldsymbol{\lambda}, \varepsilon)$ such that $\mathsf{A}^\mathsf{T}\boldsymbol{\lambda} \in \operatorname{int}(\mathcal{K}^*) = \left\{\mathbf{y} \colon \mathbf{v}^\mathsf{T}\mathbf{y} > 0 \ \forall \mathbf{v} \in \mathcal{K} \setminus \{\mathbf{o}\}\right\}$ and $\varepsilon > 0$. The maximization problem is equivalent to $\max_{(\boldsymbol{\lambda},\varepsilon,w)\in\mathbb{R}^{m+2}} \left\{\varepsilon \colon w \geq \varepsilon,\ \mathsf{A}^\mathsf{T}\boldsymbol{\lambda} - w\mathbf{a} \in \mathcal{K}^*\right\}$ by a standard duality argument (the dual variable w can be eliminated, which we avoided in order to preserve clarity). Finally we can set $\boldsymbol{\alpha} = \mathsf{A}^\mathsf{T}\boldsymbol{\lambda}^*$ and $b_0 = \boldsymbol{\alpha}^\mathsf{T}\mathbf{x}$ for any $\mathbf{x} \in \mathcal{M}$.*

Now we are ready to prove the main result in this section.

Theorem 36. *Let $\mu \in \mathbb{R}$, $\mathcal{K} \cap \mathcal{H} \subseteq \mathbb{R}^n$ and $\mathcal{K}$ be a closed cone with pointed convex hull and $\mathcal{H} = \{\mathbf{x} \colon \mathbf{a}^\mathsf{T}\mathbf{u} = b\} \subset \mathbb{R}^n$ a hyperplane, such that $\operatorname{conv}(\mathcal{K}) \cap \mathcal{H}$ is compact and nontrivial (neither empty nor a singleton). Then there is a nonsingular matrix A of order $n+1$ such that $\operatorname{cone}(\{\mu\} \times (\mathcal{K} \cap \mathcal{H})) = \mathsf{A}(\{0\} \times \mathcal{K})$ and with*

$$\mathcal{C} := \left\{\mathsf{A}^{-\mathsf{T}}\mathsf{M}(b, \mathbf{f}, \mathsf{R}+)\mathsf{A}^{-1} \in \mathcal{S}^{n+1} \colon \mathsf{R} \in \mathcal{COP}(\mathcal{K})\,,\, \mathbf{f} \in \mathbb{R}^n\,,\, b \in \mathbb{R}\right\}$$

we have

$$\mathcal{COP}(\operatorname{cone}[\{\mu\} \times (\mathcal{K} \cap \mathcal{H})]) = \mathcal{C}\,.$$

Proof. We have

$$\begin{aligned}
\mathcal{COP}(\operatorname{cone}[\{\mu\} \times (\mathcal{K} \cap \mathcal{H})]) &= \mathcal{COP}(\mathsf{A}[\{0\} \times \mathcal{K}]) && \text{by Lemma 34} \\
&= \mathsf{A}^{-\mathsf{T}}\mathcal{COP}(\{0\} \times \mathcal{K})\mathsf{A}^{-1} && \text{by Equation (7.18)} \\
&= \mathcal{C}\,. && \text{by Equation (7.17)}
\end{aligned}$$

□

The above theorem allows us to harness other results in quadratic optimization in order to characterize or at least approximate set-copositive constraints in (7.15). As an example we use the main result in [41], which is a generalization of the celebrated result in [25], in the following theorem.

Theorem 37. *Let* $\mathcal{U}_0 := \{\mathbf{u}_0 \in \mathcal{K}_0 \colon \mathsf{A}_0\mathbf{u}_0 = \mathbf{b}_0\,,\ (\mathbf{u}_0)_j \in \{0,1\},\ j \in \mathcal{I}\} \neq \{\mathbf{o}\}$ *be a closed, bounded and nonempty set, where* $\mathcal{K}_0 \subseteq \mathbb{R}^{n_0}$ *is a closed, convex and pointed cone,* $\mathsf{A}_0 \in \mathbb{R}^{k_0 \times n_0}$*,* $\mathbf{b}_0 \in \mathbb{R}^{k_0}$ *and* $\mathcal{I} \subseteq [1\!:\!n_0]$*. Further, let* $\mathcal{K} := \operatorname{cone}(\mathcal{U}_0)$*. Define*

$$\mathcal{COP}_{prx}(\mathcal{K}) :=$$
$$\left\{ \mathsf{Q} \in \mathcal{S}^n \colon \begin{array}{c} \mathsf{Q} + \mathsf{A}^\mathsf{T}\boldsymbol{\lambda}\boldsymbol{\alpha}^\mathsf{T} + \boldsymbol{\alpha}\boldsymbol{\lambda}^\mathsf{T}\mathsf{A} + \mathsf{A}^\mathsf{T}\operatorname{Diag}(\boldsymbol{\mu})\mathsf{A} \\ -\mathsf{I}^\mathsf{T}\boldsymbol{\nu}\boldsymbol{\alpha}^\mathsf{T} - \boldsymbol{\alpha}\boldsymbol{\nu}^\mathsf{T}\mathsf{I} + 2\mathsf{I}^\mathsf{T}\operatorname{Diag}(\boldsymbol{\nu})\mathsf{I} + \delta\boldsymbol{\alpha}\boldsymbol{\alpha}^\mathsf{T} \in \mathcal{COP}\left(\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z}\right) \\ \textit{and} \quad 2\boldsymbol{\lambda}^\mathsf{T}\mathbf{b} + \boldsymbol{\mu}^\mathsf{T}\mathbf{b}^2 + \delta \leq 0 \quad \textit{for some } (\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\delta) \in \mathbb{R}^{2k+|\mathcal{I}|+1} \end{array} \right\}$$

for some appropriate matrices A*,* I *and* Q*, vectors* $\mathbf{b}$*,* $\mathbf{b}^2$*,* $\boldsymbol{\alpha}$ *and numbers* n_y*,* n_z *and* k *(which can be formed with minimal effort). Then, the following inclusions hold:*

$$\mathcal{COP}(\mathcal{K}) \supseteq \mathcal{COP}_{prx}(\mathcal{K}) \supseteq \operatorname{int}\mathcal{COP}(\mathcal{K})\,.$$

Proof. In order to apply [41, Theorem 2.6] we have to guarantee a version of Burer's key condition holds for $\mathcal{U}_0$, namely

$$\mathsf{A}_0\mathbf{u}_0 = \mathbf{b}_0 \text{ and } \mathbf{u}_0 \in \mathcal{K}_0 \implies 0 \leq (\mathbf{u}_0)_j \leq 1,\ j \in \mathcal{I}.$$

Assume $(\mathbf{u}_0)_j \leq 1$ is not implied for n_y different $j \in \mathcal{I}$ and $0 \leq (\mathbf{u}_0)_j$ is not implied for n_z different $j \in \mathcal{I}$ (i.e. in the worst case $n_y = n_z = |\mathcal{I}|$, in the best case both are zero). This forces us to introduce n_y constraints $(\mathbf{u}_0)_j + y_j = 1$, and n_z constraints $(\mathbf{u}_0)_j - z_j = 0$, where $\mathbf{y} \in \mathbb{R}_+^{n_y}$ and $\mathbf{z} \in \mathbb{R}_+^{n_z}$ are slack variables, which are necessary since all linear constraints have to be equality constraints. Note that all these additional constraints are redundant. Let $n := n_0 + n_y + n_z$ and $k := k_0 + n_y + n_z$. We define $\mathbf{b} := (\mathbf{b}_0; \mathbf{e}; \mathbf{o}) \in \mathbb{R}^k$, where $\mathbf{e} \in \mathbb{R}^{n_y}$ is the all-ones vector and $\mathbf{o} \in \mathbb{R}^{n_z}$. Further, $\mathbf{u} := (\mathbf{u}_0; \mathbf{y}, \mathbf{z}) \in \mathcal{K} \times \mathbb{R}_+^{n_y+n_z}$ and $\mathsf{Q} := (\mathsf{Q}_0, \mathsf{O}; \mathsf{O}, \mathsf{O}) \in \mathcal{S}^n$. Also we abbreviate by $\mathbf{b}^2$ the vector whose i-th entry is given by b_i^2. Now let $\mathsf{A} \in \mathbb{R}^{n \times k}$ be such that $\mathsf{A}\mathbf{u} = \mathbf{b}$ gives line by line the equalities $\mathsf{A}_0\mathbf{u}_0 = \mathbf{b}_0$, $(\mathbf{u}_0)_j + y_j = 1$ and $(\mathbf{u}_0)_j - z_j = 0$. It is easy to check, that for $\mathcal{U} := \{\mathbf{u} \in \mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z} : \mathsf{A}\mathbf{u} = \mathbf{b}\,,\ u_j \in \{0,1\},\ j \in \mathcal{I}\}$ we have $\inf\limits_{\mathbf{u}\in\mathcal{U}} \mathbf{u}^\mathsf{T}\mathsf{Q}\mathbf{u} = \inf\limits_{\mathbf{u}_0\in\mathcal{U}_0} \mathbf{u}_0^\mathsf{T}\mathsf{Q}_0\mathbf{u}_0$, that $\mathcal{U}$ is a compact set, and that $\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z}$ is a closed, convex, pointed cone. Thus, by Lemma 35, we can add a redundant constraint $\boldsymbol{\alpha}^\mathsf{T}\mathbf{u} = \beta$ where $\boldsymbol{\alpha} \in \operatorname{int}((\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z})^*)$ to the definition of $\mathcal{U}$. By redundancy and since $\mathbf{y}^\mathsf{T}\boldsymbol{\alpha} > 0\ \forall \mathbf{y} \in \mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z} \setminus \{0\}$, we see that $\beta > 0$ so that after a scaling, the constraint can be written as $\boldsymbol{\alpha}^\mathsf{T}\mathbf{u} = 1$. Then [41, Theorem 2.6] along with a straightforward adaptation of the *reduction step* outlined in [26, Theorem 2] gives us the equivalence between the following two optimization problems

$$\begin{aligned}
p^* = \inf_{\mathbf{u}_0 \in \mathcal{K}_0} \quad & \mathbf{u}_0^\mathsf{T} \mathsf{Q}_0 \mathbf{u}_0 \\
\text{s.t.}: \quad & \mathsf{A}_0^\mathsf{T} \mathbf{u}_0 = \mathbf{b}_0, \\
& (\mathbf{u}_0)_j \in \{0,1\} \quad j \in \mathcal{I}
\end{aligned}
\qquad
\begin{aligned}
q^* = \inf_{\mathsf{U} \in \mathcal{C}} \quad & \mathsf{Q} \bullet \mathsf{U} \\
\text{s.t.}: \quad & \mathbf{a}_i^\mathsf{T} \mathsf{U} \boldsymbol{\alpha} = b_i, && i \in [1:k], \\
& \mathbf{a}_i^\mathsf{T} \mathsf{U} \mathbf{a}_i = b_i^2, && i \in [1:k], \\
& \mathbf{e}_j^\mathsf{T} \mathsf{U} \mathbf{e}_j = \boldsymbol{\alpha} \mathsf{U} \mathbf{e}_j, && j \in \mathcal{I}, \\
& \boldsymbol{\alpha}^\mathsf{T} \mathsf{U} \boldsymbol{\alpha} = 1
\end{aligned}
\tag{7.19}$$

where $\mathcal{C} := \mathcal{CPP}(\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z})$, so that the latter problem has the following dual, abbreviating $\mathsf{I} = \mathsf{I}_{|\mathcal{I}|}$:

$$\begin{aligned}
d^* = \sup_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\delta} \quad & -(2\boldsymbol{\lambda}^\mathsf{T}\mathbf{b} + \boldsymbol{\mu}^\mathsf{T}\mathbf{b}^2 + \delta) \\
\text{s.t.}: \quad & \mathsf{Q} + \mathsf{A}^\mathsf{T}\boldsymbol{\lambda}\boldsymbol{\alpha}^\mathsf{T} + \boldsymbol{\alpha}\boldsymbol{\lambda}^\mathsf{T}\mathsf{A} + \mathsf{A}^\mathsf{T}\operatorname{Diag}(\boldsymbol{\mu})\mathsf{A} \\
& \quad - \mathsf{I}^\mathsf{T}\boldsymbol{\nu}\boldsymbol{\alpha}^\mathsf{T} - \boldsymbol{\alpha}\boldsymbol{\nu}^\mathsf{T}\mathsf{I} + 2\mathsf{I}^\mathsf{T}\operatorname{Diag}(\boldsymbol{\nu})\mathsf{I} \\
& \quad + \delta\boldsymbol{\alpha}\boldsymbol{\alpha}^\mathsf{T} \in \mathcal{COP}(\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z}),
\end{aligned}$$

where the dual variables $(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\delta) \in \mathbb{R}^{2k+|\mathcal{I}|+1}$ and where $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ were subject to a change of variables from $\boldsymbol{\lambda} \Rightarrow 2\boldsymbol{\lambda}$ and $\boldsymbol{\nu} \Rightarrow 2\boldsymbol{\nu}$ to avoid fractions in the expression. Now, since $\boldsymbol{\alpha} \in \operatorname{int}\,(\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z})^*$ we have $\boldsymbol{\alpha}\boldsymbol{\alpha}^\mathsf{T} \in \operatorname{int}\,\mathcal{COP}(\mathcal{K}_0 \times \mathbb{R}_+^{n_y+n_z})$, so that the dual problem has an obvious Slater point and indeed $d^* = q^*$. We are now ready to prove the inclusions in the theorem. Note that $\mathbf{o} \notin \mathcal{U}$ under the hypothesis of the theorem, hence $p^* > 0$ if and only if $\mathsf{Q} \in \operatorname{int}\,\mathcal{COP}(\mathcal{K})$. In this case $d^* = p^* > 0$ implies that $\mathsf{Q} \in \mathcal{COP}_{prx}(\mathcal{K})$. If $\mathsf{Q} \in \mathcal{COP}_{prx}(\mathcal{K})$ then $p^* = d^* \geq 0$ which implies that $\mathsf{Q} \in \mathcal{COP}(\mathcal{K})$. $\square$

The argument in the proof is analogous to the general strategy outlined in Section 7.2, the difference being that the objective in the minimization problem is homogeneous. Also, we do not assume dual attainability which is why the reverse inclusions cannot be established. The theorem highlights the fact, that in absence of dual attainability the feasible set is shrinking, and locates the points that are lost at the boundary of the set-copositive cone we sought to characterize. The reformulation results in [25, 41] have been used in [59, 84] for the sake of reformulating semi-infinite constraints with quadratic dependency on the uncertainty vector. We want to point out that in these papers the problem of dual attainability was not sufficiently considered. In addition, these papers do not harness the *reduction step*, which by Lemma 35 is always applicable if the uncertainty set $\mathcal{U}$ is compact. The reduction step reduces the dimension of the set-copositive constraint in the copositive reformulation and at least opens the possibility for a primal Slater point to exist, which would never be the case if that reduction step was omitted as was pointed out in [26, Section 3]. To better understand the situation, consider the example given there more carefully. It states, that

$$\begin{aligned}
&\min\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x}\colon \mathbf{e}^\mathsf{T}\mathbf{x} = 1, \mathbf{x} \in \mathbb{R}^n_+\}\\
=&\min\{\mathsf{Q}\bullet\mathsf{X}\colon \mathbf{e}^\mathsf{T}\mathbf{x} = 1,\ \mathbf{e}^\mathsf{T}\mathsf{X}\mathbf{e} = 1,\ \mathsf{Y}(\mathsf{X},\mathbf{x}) \in \mathcal{CPP}(\mathbb{R}^{n+1}_+)\}\\
=&\min\{\mathsf{Q}\bullet\mathsf{X}\colon \mathbf{e}^\mathsf{T}\mathsf{X}\mathbf{e} = 1,\ \mathsf{X} \in \mathcal{CPP}(\mathbb{R}^n_+)\}.
\end{aligned}$$

The first equality is due to the main result in [26] and the second one is the reduction step outlined in Section 3 of that paper. In fact, equality of the first and the last problem was already established at the origins of copositive programming in [11]. Take any interior point $\mathsf{X} \in \mathrm{int}\mathcal{CPP}(\mathbb{R}^n_+)$. Since $\mathbf{e}\mathbf{e}^\mathsf{T} \in \mathrm{int}\mathcal{COP}(\mathbb{R}^n_+)$ one can rescale X such that $\mathbf{e}^\mathsf{T}\mathsf{X}\mathbf{e} = 1$ holds, and we see that the last minimization problem indeed has a primal Slater point. In fact, the same line of reasoning remains valid if we replace $\mathbb{R}^n_+$ by any convex cone $\mathcal{K}_0$ and $\mathbf{e}$ by any $\boldsymbol{\alpha} \in \mathcal{K}_0^*$ and 1 by any positive number b. The only difference is that we then have to appeal to the results in [26] or [41] for the reformulation step (in this thesis we actually presented elementary arguments in Section 2.1.2). In such cases we would have $\mathcal{K} = \mathrm{cone}\{\mathbf{x} \in \mathcal{K}_0\colon \boldsymbol{\alpha}^\mathsf{T}\mathbf{x} = b\} = \mathcal{K}_0$, i.e. $\mathcal{X}_0$ would be a base of $\mathcal{K}_0$. The question is, whether we have any other cases where we end up with a primal Slater point after applying the reduction step? Proposition 38 unfortunately gives a negative answer to this question: it states that a primal Slater point occurs exactly in the case where there is just a single equality constraint. Therefore, Theorem 37 is only an approximate result and does not give rise to an exact copositivity characterization in general. This unfortunate circumstance is at least partly redeemed by the fact that we lose at most some points at the boundary of the set.

Proposition 38. *Consider the conic reformulation in (7.19), i.e.*

$$\begin{aligned}
q^* = \inf_{\mathsf{U}\in\mathcal{C}}\quad & \mathsf{Q}\bullet\mathsf{U} \\
\text{s.t. :}\quad & \mathbf{a}_i^\mathsf{T}\mathsf{U}\boldsymbol{\alpha} = b_i, \qquad i \in [1\colon k]\\
& \mathbf{a}_i^\mathsf{T}\mathsf{U}\mathbf{a}_i = b_i^2, \qquad i \in [1\colon k]
\end{aligned} \tag{7.20}$$

$$\mathbf{e}_j^{\mathsf{T}} \mathsf{U} \mathbf{e}_j = \boldsymbol{\alpha} \mathsf{U} \mathbf{e}_j, \quad j \in \mathcal{I}$$
$$\boldsymbol{\alpha}^{\mathsf{T}} \mathsf{U} \boldsymbol{\alpha} = 1$$

This problem has a strictly feasible point if and only if $\mathbf{a}_i = b_i \boldsymbol{\alpha}$ *for all* $i \in [1\!:\!m]$.

Proof. The "if" part is clear since then all constraints but $\boldsymbol{\alpha}^{\mathsf{T}} \mathsf{U} \boldsymbol{\alpha} = 1$ are redundant, $\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathsf{T}} \in \operatorname{int} \mathcal{CPP}(\mathcal{K}_0)^*$ and $1 > 0$ so that any interior point of $\mathcal{CPP}(\mathcal{K}_0)$ can be scaled to fulfil the one remaining constraint. So assume there is a Slater point $\overline{\mathsf{U}} \in \operatorname{int} \mathcal{CPP}(\mathcal{K}_0) \subseteq \operatorname{int} \mathcal{S}_+^n$. Then $\overline{\mathsf{U}}^{1/2}$ exists and is invertible. Set $\overline{\mathbf{a}}_i = \overline{\mathsf{U}}^{1/2} \mathbf{a}_i$ and $\overline{\boldsymbol{\alpha}} = \overline{\mathsf{U}}^{1/2} \boldsymbol{\alpha}$ then the constraints read as $\overline{\boldsymbol{\alpha}}^{\mathsf{T}} \overline{\boldsymbol{\alpha}} = 1$, $\overline{\mathbf{a}}^{\mathsf{T}} \overline{\mathbf{a}}_i = b_i^2$ and $\overline{\boldsymbol{\alpha}}^{\mathsf{T}} \overline{\mathbf{a}}_i = b_i$. The first and the second one imply $\|\overline{\boldsymbol{\alpha}}\| = 1$ and $\|\overline{\mathbf{a}}_i\| = b_i$ but then by the equality case of Cauchy-Schwarz, the third constraint implies $\overline{\mathbf{a}}_i = \beta_i \overline{\boldsymbol{\alpha}}$ for some β_i; we actually have $b_i = \beta_i \|\overline{\boldsymbol{\alpha}}\|^2 = \beta_i$ and finally, by invertibility of $\mathsf{U}^{1/2}$, $\mathbf{a}_i = b_i \boldsymbol{\alpha}$. □

7.6 Adjustable robust reformulation of disjoint convex-non-convex quadratic optimization

So far, we have discussed how to use reformulations of QCQPs for the purpose of reformulating robust optimization problems. However, [90] outlined a way to use an adjustable robust approach in order to tackle disjoint bilinear optimization problems, which are special cases of QCQPs, thus demonstrating that the usefulness runs both ways. In this section we will show that the techniques we discussed in this text can be used to further their approach.

7.6.1 Deriving the reformulation and its convex-conic lower bound

The problem considered in [90] is given by:

$$\inf_{(\mathbf{x},\mathbf{y}) \in \mathcal{X} \times \mathcal{Y}} \mathbf{x}^{\mathsf{T}} \mathsf{Q} \mathbf{y}. \tag{7.21}$$

The following theorem proved in that article establishes a connection between QCQPs and adjustable robust optimization.

Theorem 39. *Let* $\mathcal{Y} := \left\{ \mathbf{y} \in \mathbb{R}_+^{n_y} \colon \mathsf{A}_y \mathbf{y} \geq \mathbf{b}_y \right\}$, *where* $\mathsf{A}_y \in \mathbb{R}^{m_y \times n_y}$ *and* $\mathbf{b}_y \in \mathbb{R}^{m_y}$. *Then (7.21) has the same optimal value as*

$$\sup_{\tau} \left\{ \tau \colon \textit{there is a nonnegative function } \mathbf{z}(\mathbf{x}) \colon \tau \leq \mathbf{b}_y^{\mathsf{T}} \mathbf{z}(\mathbf{x}),\ \mathsf{A}_y^{\mathsf{T}} \mathbf{z}(\mathbf{x}) \leq \mathsf{Q}^{\mathsf{T}} \mathbf{x} \ \ \forall \mathbf{x} \in \mathcal{X} \right\}.$$

We will expand this approach to the case where in addition a possibly nonconvex quadratic term in $\mathbf{x}$ and a convex quadratic term in $\mathbf{y}$ are present in the objective. We will also slightly deviate from the definition of $\mathcal{Y}$ in that we will assume equality constraints. As a consequence, the function $\mathbf{z}(\mathbf{x})$ will not have a sign restriction, which will be useful in in the proof of Theorem 41.

The following argument is a straightforward but tedious generalization of Theorem 39. For the readers' convenience we provide a detailed derivation.

Theorem 40. *Let $\mathsf{Q}_x \in \mathcal{S}^{n_1}$, $\mathsf{Q}_{xy} \in \mathbb{R}^{n_1 \times n_2}$, $\mathsf{F} \in \mathbb{R}^{k \times n_2}$ and $\mathsf{G} \in \mathbb{R}^{r \times n_2}$. Further, assume $\mathcal{X} \subseteq \mathbb{R}^{n_1}$ is a compact set and $\mathcal{Y} := \{\mathbf{y} \in \mathbb{R}^{n_2}_+ \colon \mathsf{F}\mathbf{y} = \mathbf{d}\} \subseteq \mathbb{R}^{n_2}$ has a Slater point and let $\mathcal{Z}(\mathbf{x}) := \{(\mathbf{z}, \mathbf{w}) \colon \mathsf{F}^\mathsf{T}\mathbf{z} + \mathsf{G}^\mathsf{T}\mathbf{w} \le \mathsf{Q}_{xy}^\mathsf{T}\mathbf{x}\}$. Then*

$$\inf_{\mathbf{x}\in\mathcal{X},\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y} + \|\mathsf{G}\mathbf{y}\|^2 \tag{7.22}$$

$$\begin{aligned} = \sup_{\tau}\{\tau \colon \textit{for some } \mathbf{z}(\mathbf{x}), \mathbf{w}(\mathbf{x}) \colon &\mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{d}^\mathsf{T}\mathbf{z}(\mathbf{x}) - \tfrac{1}{4}\|\mathbf{w}(\mathbf{x})\|^2 \ge \tau, \\ &(\mathbf{z}(\mathbf{x}), \mathbf{w}(\mathbf{x})) \in \mathcal{Z}(\mathbf{x}), \, \forall \mathbf{x} \in \mathcal{X}\}, \end{aligned} \tag{7.23}$$

where $\mathbf{z} \colon \mathbb{R}^{n_1} \to \mathbb{R}^k$ and $\mathbf{w} \colon \mathbb{R}^{n_1} \to \mathbb{R}^r$ are function-valued decision variables.

Proof. Observe that

$$\begin{aligned} &\inf_{(\mathbf{x},\mathbf{y})\in\mathcal{X}\times\mathcal{Y}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y} + \|\mathsf{G}\mathbf{y}\|^2\right\} \\ = &\sup_{\tau} \left\{\tau \colon \mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y} + \|\mathsf{G}\mathbf{y}\|^2 \ge \tau, \, \forall \mathbf{x} \in \mathcal{X}, \, \forall \mathbf{y} \in \mathcal{Y}\right\}. \end{aligned}$$

Note that the semi-infinite constraint in the last problem implies a minimization problem of the constraint function with respect to $\mathbf{y}$ explicitly given by

$$\inf_{\mathbf{y}\in\mathbb{R}^{n_2}_+} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y} + \|\mathsf{G}\mathbf{y}\|^2 \colon \mathsf{F}\mathbf{y} = \mathbf{d}\right\}. \tag{7.24}$$

By assumption there is a $\mathbf{y} \in \operatorname{int}(\mathbb{R}^{n_2}_+)$ such that $\mathsf{F}\mathbf{y} = \mathbf{d}$, thus (7.24) fulfils Slater's condition, so that zero duality gap and dual attainability is guaranteed for the dual given by

$$\sup_{(\mathbf{z},\mathbf{w})\in\mathbb{R}^{2k}} \left\{\mathbf{d}^\mathsf{T}\mathbf{z} - \tfrac{1}{4}\|\mathbf{w}\|^2 \colon \mathsf{F}^\mathsf{T}\mathbf{z} + \mathsf{G}^\mathsf{T}\mathbf{w} \le \mathsf{Q}_{xy}^\mathsf{T}\mathbf{x}\right\}. \tag{7.25}$$

The dual feasible set is given by $\mathcal{Z}(\mathbf{x})$. The following equations hold:

$$\begin{aligned} &\sup_{\tau} \left\{\tau \colon \mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y} + \|\mathsf{G}\mathbf{y}\|^2 \ge \tau, \, \forall \mathbf{x} \in \mathcal{X}, \, \forall \mathbf{y} \in \mathcal{Y}\right\} \\ = &\sup_{\tau} \left\{\tau \colon \forall \mathbf{x} \in \mathcal{X}, \text{ there is } (\mathbf{z}, \mathbf{w}) \in \mathcal{Z}(\mathbf{x}) \colon \mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{d}^\mathsf{T}\mathbf{z} - \tfrac{1}{4}\|\mathbf{w}\|^2 \ge \tau, \right\} \end{aligned}$$

$$= \sup_{\tau} \Big\{ \tau \colon \text{for some functions } \mathbf{z}(\mathbf{x}), \mathbf{w}(\mathbf{x}) \colon \mathbf{x}^\mathsf{T} \mathsf{Q}_x \mathbf{x} + \mathbf{d}^\mathsf{T} \mathbf{z}(\mathbf{x}) - \tfrac{1}{4} \|\mathbf{w}(\mathbf{x})\|^2 \geq \tau,$$
$$(\mathbf{z}(\mathbf{x}), \mathbf{w}(\mathbf{x})) \in \mathcal{Z}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{X} \Big\}.$$

□

We end up with an adjustable robust optimization problem with second stage variables $\mathbf{z}(\mathbf{x})$ and $\mathbf{w}(\mathbf{x})$ and uncertainty set $\mathcal{X}$, i.e. the decision vector $\mathbf{x}$ takes the role of the uncertainty parameter.

The fact that $\mathbf{z}(\mathbf{x})$ and $\mathbf{w}(\mathbf{x})$ are function valued variables is a major complication since the space of all function is an intractable search space. The standard way to deal with this issue is to contract the search space as to encompass only linear functions, i.e. $\mathbf{z}(\mathbf{x}) = \mathsf{Z}\mathbf{x} + \mathbf{z}$ and $\mathbf{w}(\mathbf{x}) = \mathsf{W}\mathbf{x} + \mathbf{w}$. The new decision variables then become the coefficients that identify the affine functions. If after the contraction step the robust constraints become linear in the uncertainty parameter one can employ standard reformulation techniques based on linear conic duality in order to obtain a tractable reformulation of the semi-infinite constraints. However, at least one of the adjustable robust constraints in (7.23) depends quadratically on $\mathbf{x}$ rather than linearly. Thus, approximations based on linear duality cannot be employed here.

However, the general reformulation strategy allows us not only to deal with the terms $\mathbf{x}^\mathsf{T} \mathsf{Q}_x \mathbf{x}$ in the adjustable robust reformulation (7.23), it also allows us for the adjustable variable $\mathbf{z}(\mathbf{x})$ to employ a quadratic decision rule, i.e. $(\mathbf{z}(\mathbf{x}))_j = \mathbf{x}^\mathsf{T} \mathsf{Z}_j \mathbf{x} + \mathbf{x}^\mathsf{T} \mathbf{z}_j + z_j,\ j \in [1\colon k]$, which is a class of decision rules which contains affine decision rules as a special case. Thus it can potentially yield tighter approximations than the ones obtainable by employing affine decision rules. On the other hand, $\mathbf{w}(\mathbf{x})$ is subject to a (convex) quadratic constraint. Consequently, an affine policy has to be employed so that $\mathbf{w}(\mathbf{x}) = \mathsf{W}\mathbf{x} + \mathbf{w}$, lest terms of order larger than 2 emerge. Since employing such decision rules is a restriction of the maximization problem, we will get a lower bound on the original minimization problem.

Theorem 41. *Let the problem data be given as in Theorem 40. Assume that for any* $\mathsf{Q} \in \mathcal{S}^{n_1}$, $\mathbf{q} \in \mathbb{R}^{n_1}$ *and* $\omega \in \mathbb{R}^n$ *it holds that*

$$\inf_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\mathsf{T} \mathsf{Q} \mathbf{x} + \mathbf{q}^\mathsf{T} \mathbf{x} + \omega = \inf_{\mathsf{X} \in \mathcal{C}} \left\{ \mathsf{M}(\omega, \mathbf{q}, \mathsf{Q}) \bullet \mathsf{X} \colon \mathsf{A}_j \bullet \mathsf{X} \leq b_j,\ j \in [1\colon m] \right\},$$

for some appropriate matrix cone $\mathcal{C} \subseteq \mathcal{S}^{n_1+1}$, *matrices* $\mathsf{A}_j \in \mathcal{S}^{n_1+1}$, $j \in [1:m]$ *such that there is no duality gap for the conic reformulation and its dual. Then, the* *following optimization problem gives a nontrivial lower bound for (7.23):*

$$\begin{aligned}
&\sup_{\tau,\mathsf{Z}_i,\mathbf{z}_i,z_i,\boldsymbol{\lambda}_l,\mathsf{W},\mathsf{H}} \tau \\
&\text{s.t. : } \mathsf{M}\left(-\tau,\mathbf{o},\mathsf{Q}_x\right)+\sum_{i=1}^{k} d_i\mathsf{M}\left(z_i,\mathbf{z}_i,\mathsf{Z}_i\right)-\tfrac{1}{4}\mathsf{H}+\sum_{j=1}^{m}(\boldsymbol{\lambda}_0)_j\mathsf{A}_j\in\mathcal{C}^*,\\
&\mathsf{M}(0,(\mathsf{Q}_{xy})_l,\mathsf{O})-\sum_{i=1}^{k} f_{il}\mathsf{M}\left(z_i,\mathbf{z}_i,\mathsf{Z}_i\right)-\sum_{i=1}^{r} g_{il}\mathsf{M}\left(w_i,\mathbf{w}_i,\mathsf{O}\right)+\sum_{j=1}^{m}(\boldsymbol{\lambda}_l)_j\mathsf{A}_i\in\mathcal{C}^*, \qquad l\in[1\!:\!n_2],\\
&\mathsf{M}\left(\mathsf{H},2(\mathbf{w},\mathsf{W}),\mathsf{I}_r\right)\in\mathcal{S}_+^{n_1+r+1},\\
&\mathbf{b}^{\mathsf{T}}\boldsymbol{\lambda}_l\le 0, \qquad l\in[0\!:\!n_2],\\
&\mathsf{W}\in\mathbb{R}^{r\times n_1},\\
&\mathbf{w}\in\mathbb{R}^r\\
&\mathsf{M}(z_i,\mathbf{z}_i,\mathsf{Z}_i)\in\mathcal{S}^{n_1+1}, \quad i\in[1\!:\!k],\\
&\boldsymbol{\lambda}_l\in\mathbb{R}_+^m, \qquad l\in[0\!:\!n_2],
\end{aligned}$$

where $(\mathsf{Q}_{xy})_l = \mathrm{Column}_l(\mathsf{Q}_{xy})$, $\mathbf{w}_i = \mathrm{Row}_i(\mathsf{W})$, $f_{il} = (\mathsf{F})_{li}$, $g_{il} = (\mathsf{G})_{li}$, $w_i = (\mathbf{w})_i$ *and* $d_i = (\mathbf{d})_i$.

Proof. Under the assumptions of the theorem we have that, given any $\mathsf{Q}\in\mathcal{S}^{n_1}$, $\mathbf{q}\in\mathbb{R}^{n_1}$ and $\omega\in\mathbb{R}^n$, we can enforce the semi-infinite constraint

$$\mathbf{x}^{\mathsf{T}}\mathsf{Q}\mathbf{x}+\mathbf{q}^{\mathsf{T}}\mathbf{x}+\omega\ge 0 \;\forall\mathbf{x}\in\mathcal{X} \tag{7.26}$$

by demanding that

$$\mathbf{b}^{\mathsf{T}}\boldsymbol{\lambda}\le 0,\; \mathsf{M}\left(\omega,\mathbf{q},\mathsf{Q}\right)+\sum_{j=1}^{m}\lambda_j\mathsf{A}_j\in\mathcal{C}^*,\; \boldsymbol{\lambda}\in\mathbb{R}_+^m, \tag{7.27}$$

as long as full strong duality holds (the consequences of a failure of full strong duality will be discussed at the end of the proof). Setting $(\mathbf{z}(\mathbf{x}))_i = \mathbf{x}^{\mathsf{T}}\mathsf{Z}_i\mathbf{x}+\mathbf{x}^{\mathsf{T}}\mathbf{z}_i+z_i,\; i\in[1\!:\!k]$ and $\mathbf{w}(\mathbf{x}) = \mathsf{W}\mathbf{x}+\mathbf{w}$ we see that the constraint $\mathbf{x}^{\mathsf{T}}\mathsf{Q}_x\mathbf{x}+\mathbf{d}^{\mathsf{T}}\mathbf{z}(\mathbf{x})-\frac{1}{4}\|\mathsf{W}\mathbf{x}+\mathbf{w}\|^2-\tau\ge 0\;\forall\mathbf{x}\in\mathcal{X}$ is equivalent to $\mathbf{x}^{\mathsf{T}}\mathsf{Q}_x\mathbf{x}+\mathbf{d}^{\mathsf{T}}\mathbf{z}(\mathbf{x})-\frac{1}{4}\left(\mathbf{x}^{\mathsf{T}},1\right)\mathsf{H}\left(\mathbf{x}^{\mathsf{T}},1\right)^{\mathsf{T}}-\tau\ge 0\;\forall\mathbf{x}\in\mathcal{X},\; \mathsf{M}\left(\mathsf{H},2\left(\mathbf{w},\mathsf{W}\right),\mathsf{I}_r\right)\in\mathcal{S}_+^{n_1+r+1}$ by Lemma 31 below. The semi-infinite constraint is of the form (7.26) with $\mathsf{Q} = \mathsf{Q}_x-\frac{1}{4}\left(\mathsf{H}\right)_{1:n_1,1:n_1}+\sum_{i=1}^{k} d_i\mathsf{Z}_i$, $\mathbf{q} = \sum_{i=1}^{k} d_i\mathbf{z}_i-\frac{1}{2}\left(\mathsf{H}\right)_{1:n_1,n_1+1}$ and $\omega = \sum_{i=1}^{k} d_iz_i-\tau-\frac{1}{4}\left(\mathsf{H}\right)_{n_1+1,n_1+1}$. Similarly, the constraints $(\mathsf{Q}_{xy}^{\mathsf{T}}\mathbf{x}-\mathsf{F}^{\mathsf{T}}\mathbf{z}(\mathbf{x})-\mathsf{G}^{\mathsf{T}}\mathbf{w}(\mathbf{x}))_l\ge 0\;\forall\mathbf{x}\in\mathcal{X},\; l\in[1:n_2]$ are of the form (7.26) with $\mathsf{Q} = \sum_{i=1}^{k} f_{il}\mathsf{Z}_i$, $\mathbf{q} = (\mathsf{Q}_{xy})_l^{\mathsf{T}}-\sum_{i=1}^{k} f_{il}\mathbf{z}_i-\sum_{i=1}^{r} g_{il}\mathbf{w}_i^{\mathsf{T}}$ and $\omega = \sum_{i=1}^{k} f_{il}z_i$. Now, since we do not assume full strong duality, but merely zero duality gap for the conic reformulation of (7.26), we need to argue that the conclusion of the theorem still holds in absence of dual attainability. In fact, for any $(\mathsf{Q}, \mathbf{q}, \omega)$ for which dual attainability fails, constraints (7.27) may become infeasible. This can only be the case if there is an $\mathbf{x} \in \mathcal{X}$ such that the inequality (7.26) is actually tight. To see this, consider the case when $p^* = \inf_{\mathbf{x}\in\mathcal{X}} \left\{\mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega\right\} > 0$. Since we assume zero duality gap, there is a dually feasible sequence $\boldsymbol{\lambda}_i$ such that $\lim_{i\to\infty} -\mathbf{b}^\mathsf{T}\boldsymbol{\lambda}_i = p^*$ and since the latter quantity strictly exceeds zero, there has to be an i^* such hat $-\mathbf{b}^\mathsf{T}\boldsymbol{\lambda}_{i^*} \geq 0$ so that $(\mathsf{Q}, \mathbf{q}, \omega)$ is feasible for (7.26) and (7.27) simultaneously. However, if $p^* = 0$ there may be no member of the sequence that takes the dual objective to zero, so that any choice on the decision variables that yields such instances of $(\mathsf{Q}, \mathbf{q}, \omega)$ may be infeasible. Such a contraction of the feasible set can only lower the optimal value of the optimization problem, so that the conclusion of the theorem remains valid, if we can show that the feasible set never shrinks to the empty set. Let us again consider (7.23) under a quadratic policy. We need to show that there is a choice on the coefficients of the quadratic policy and τ so that none of the inequalities is ever binding, i.e. they are fulfilled with strict inequality for all $\mathbf{x} \in \mathcal{X}$. It suffices to show that there is a static policy, where $\mathbf{z}(\mathbf{x})$, $\mathbf{w}(\mathbf{x})$ are constant for all $\mathbf{x}$, which accomplishes this goal. The inequality $\mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x} + \mathbf{d}^\mathsf{T}\mathbf{z} - \frac{1}{4}\|\mathbf{w}\|^2 \geq \tau$ is safe, since we can make τ arbitrarily small and the left-hand side is bounded over $\mathcal{X}$ as it is a compact set. Since $\mathcal{Y} := \{\mathbf{y} \in \mathbb{R}^{n_2}_+ : \mathsf{F}\mathbf{y} = \mathbf{d}\} \subseteq \mathbb{R}^{n_2}$ is a bounded conic intersection we invoke Lemma 35 to show that there is a $\bar{\mathbf{z}} \in \mathbb{R}^k$ so that $\mathsf{F}^\mathsf{T}\bar{\mathbf{z}} \in \mathrm{int}\mathbb{R}^{n_2}_+$. Now set $\mathbf{z} = -\mu\bar{\mathbf{z}}$ to see that $\mathsf{F}^\mathsf{T}\mathbf{z} + \mathsf{G}^\mathsf{T}\mathbf{w} \leq \mathsf{Q}_{xy}^\mathsf{T}\mathbf{x}$ can be fulfilled for all $\mathbf{x} \in \mathcal{X}$ with strict inequality by choosing $\mu > 0$ large enough. This concludes the proof.

$\square$

The argument for the linearization of the constraint $\mathsf{H} = (\mathsf{W}, \mathbf{w})^\mathsf{T}(\mathsf{W}, \mathbf{w})$ in the form $\mathsf{M}\left(\mathsf{I}_r, 2(\mathsf{W}, \mathbf{w}), \mathsf{H}\right) \in \mathcal{S}_+^{n_1+r+1}$ is given in the following lemma, which is a straightforward generalization of [59, Lemma 4]. It is slightly more general than needed for Theorem 41 as it might be useful in other contexts such as the ones discussed in [59].

Lemma 42. *Let $\mathsf{Q}(\mathbf{y})\colon \mathbb{R}^k \mapsto \mathbb{R}^{m\times n}$. A vector $\mathbf{y} \in \mathbb{R}^k$ fulfills*

$$\mathbf{x}^\mathsf{T}(\mathsf{D} - \mathsf{Q}^\mathsf{T}(\mathbf{y})\mathsf{Q}(\mathbf{y}))\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega \geq 0 \;\forall \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n \tag{7.28}$$

if and only if there exists a matrix $\mathsf{H} \in \mathcal{S}^n$ such that

$$\mathbf{x}^\mathsf{T}(\mathsf{D} - \mathsf{H})\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega \geq 0 \;\forall \mathbf{x} \in \mathcal{X} \quad \text{and} \tag{7.29}$$

$$\mathsf{M}\left(\mathsf{H},\, 2\mathsf{Q}(\mathbf{y}),\, \mathsf{I}_m\right) \in \mathcal{S}_+^{n+m}. \tag{7.30}$$

Proof. Assume that $\mathbf{y}$ fulfills (7.28). Then, it fulfils (7.29) and (7.30) by setting $\mathsf{H} = \mathsf{Q}^\mathsf{T}(\mathbf{y})\mathsf{Q}(\mathbf{y})$, which implies $\mathsf{H}-\mathsf{Q}^\mathsf{T}(\mathbf{y})\mathsf{Q}(\mathbf{y}) \in \mathcal{S}^n_+$, which by Schur complementation implies (7.12). So assume $\mathbf{y}$ and H fulfil (7.29) and (7.30), then

$$
\begin{aligned}
&\begin{pmatrix}\mathbf{x}\\1\end{pmatrix}^\mathsf{T}\left[\begin{pmatrix}\mathsf{D}-\mathsf{Q}(\mathbf{y})^\mathsf{T}\mathsf{Q}(\mathbf{y}) & \frac{1}{2}\mathbf{q}\\ \frac{1}{2}\mathbf{q}^\mathsf{T} & \omega\end{pmatrix} - \begin{pmatrix}\mathsf{D}-\mathsf{H} & \frac{1}{2}\mathbf{q}\\ \frac{1}{2}\mathbf{q}^\mathsf{T} & \omega\end{pmatrix}\right]\begin{pmatrix}\mathbf{x}\\1\end{pmatrix}\\
=&\mathbf{x}^\mathsf{T}\left[\mathsf{H}-\mathsf{Q}(\mathbf{y})^\mathsf{T}\mathsf{Q}(\mathbf{y})\right]\mathbf{x} \geq 0 \quad \forall \mathbf{x}\in\mathcal{X}\,,
\end{aligned}
$$

where the inequality holds due to (7.30). But since for any pair of matrices $\mathsf{A},\mathsf{B}\in\mathcal{S}^n$ we have that $\mathbf{x}^\mathsf{T}\mathsf{B}\mathbf{x}\geq 0\ \forall\mathbf{x}\in\mathcal{X}$ and $\mathsf{A}-\mathsf{B}\in\mathcal{S}^n_+$ implies $\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x}\geq 0\ \forall\mathbf{x}\in\mathcal{X}$ for any set $\mathcal{X}$, we have shown that (7.28) holds as well. □

7.6.2 Experimental evidence

In the following we provide results from extensive numerical experiments, where we tried to assess the quality of the lower bound from Theorem 41 for the QCQP (7.22). We will pursue two main types of experiments. First, we will compare the quality of our proposed lower bound with lower bounds obtained from typical relaxations of the (set-)copositive reformulations established in [25, 26, 41]. Second, for the case where only the bilinear term is present in the objective function we will compare our lower bound to the one derived in [90]. The main difference between the two approaches is that we, by means of the strategy outlined in Section 7.2, are able to employ quadratic decision rules while [90] employ affine policies. It should be noted that [84] provided empirical evidence for the advantage of quadratic over affine decision rules. However, we believe that the disjoint bilinear optimization problem is an interesting special case due to its applications in game theory, that is significantly different from the problems considered in [84], so that particular consideration is warranted. All experiments were implemented using the `YALMIP` interface. The semidefinite optimization problems were solved using `SDPT3`, while linear problems were solved using `Gurobi`. For the purpose of generating feasible solutions, and thus, upper bounds, we employed `fmincon` as a global solver. The experiments were run on a system with Intel Core i7-4510U CPU and 8GB RAM.

7.6.2.1 Comparison with lower bounds from copositive programming

For the problem

$$\inf_{(\mathbf{x},\mathbf{y})\in\mathcal{X}\times\mathcal{Y}}\left\{\mathbf{x}^\mathsf{T}\mathsf{Q}_x\mathbf{x}+\mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y}+\|\mathsf{G}\mathbf{y}\|^2\right\}, \tag{7.31}$$

when specifying $\mathcal{X} := \{\mathbf{x} \in \mathcal{K} : \mathsf{B}\mathbf{x} = \mathbf{c}\} \subseteq \mathbb{R}^{n_1}$ for some convex cone $\mathcal{K} \subseteq \mathbb{R}^{n_1}$, $\mathsf{B} \in \mathbb{R}^{k_1 \times n_1}$, $\mathbf{c} \in \mathbb{R}^{k_1}$ and $\mathcal{Y} := \{\mathbf{y} \in \mathbb{R}^{n_2}_+ : \mathsf{F}\mathbf{y} = \mathbf{d}\}$ as in Theorem 40, we obtain a QCQP that is amenable to copositive reformulations as demonstrated in [25]. We will only consider the case where $\mathcal{X}$ and $\mathcal{Y}$ are bounded so that by Lemma 35 we can always ad a redundant linear constraint $\boldsymbol{\alpha}^\mathsf{T}(\mathbf{x}^\mathsf{T}, \mathbf{y}^\mathsf{T})^\mathsf{T} = 1$ with $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^\mathsf{T}, \boldsymbol{\alpha}_2^\mathsf{T})^\mathsf{T} \in \operatorname{int}(\mathcal{K}^* \times \mathbb{R}^{n_2}_+)$ which enables us to apply the simplification step outlined in [26, Section 2.3]. The resulting completely positive reformulation involves a conic constraint that restricts the decision matrix to $\mathcal{CPP}(\mathcal{K} \times \mathbb{R}^{n_2}_+)$. This constraint is generally difficult (unless for example $n_2 = 0$ and $\mathcal{K}$ is the second order cone). For the purpose of our experiments we will resort to approximations which are standard in literature. In case $\mathcal{K} = \mathbb{R}^{n_1}_+$ we will replace $\mathcal{CPP}\left(\mathbb{R}^{n_1+n_2}\right)$ by $\mathcal{DNN}_{n_1+n_2} := \mathcal{S}^{n_1+n_2}_+ \cap \mathcal{N}_{n_1+n_2}$. In case $\mathcal{K}$ is the second order cone we will resort to

$$\mathcal{O}_{uter} := \left\{\mathsf{M}\left(\mathsf{C}_1, 2\mathsf{C}_2, \mathsf{C}_3\right) : \mathsf{C}_1 \in \mathcal{CPP}\left(\mathcal{K}\right), \mathrm{Columns}\left(\mathsf{C}_2\right) \in \mathcal{K}, \mathsf{C}_3 \in \mathcal{DNN}_{n_2}\right\},$$

which is an approximation described in [83], where the dual of $\mathcal{O}_{uter}$ was used to approximate $\mathcal{COP}\left(\mathcal{K} \times \mathbb{R}^n_+\right)$. Note that by the S-Lemma, $\mathcal{CPP}\left(\mathcal{K}\right) = \left\{\mathsf{C} \in \mathcal{S}^n_+ : \mathsf{C} \bullet \mathsf{J}_0 \leq 0\right\}$ where $\mathsf{J}_0 = \mathrm{Diag}(-1, 1, \ldots, 1) \in \mathcal{S}_n$ and is thus a tractable matrix cone.

As for the lower bound in Theorem 41 the main assumption is fulfilled as for the implicit minimization problem $\inf_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} + \omega$ we can again invoke [26]. In consequence the matrix cone $\mathcal{C}^*$ given by $\mathcal{COP}(\mathcal{K})$ which is exactly characterized by $\mathcal{COP}(\mathcal{K}) = \left\{\mathsf{C} : \mathsf{C} + \tau\mathsf{J}_0 \in \mathcal{S}^n_+,\ \tau \geq 0\right\}$ in case $\mathcal{K}$ is the second order cone. For each instance type, characterized by the size of n_1, n_2, m_1, m_2 and the choice for $\mathcal{K} \in \{\mathbb{R}^{n_1}_+, \mathcal{SOC}\}$, which are indicated by the tuples $(\mathcal{K}, n_1, n_2, m_1, m_2)$ in the first column, we randomly generated 100 instances. In the tables we give average values for the relative optimality gap (avg. gap), the maximum relative optimality gap (max. gap), the number of times a model gave the a better solution (# best) and their respective average computation time (avg. time). We see that our model can outperform the copositive bound in cases when dimension of $\mathcal{Y}$ is significantly larger than that of $\mathcal{X}$ and the number of constraints in the former is not too small. In these cases computation time seems to be lower for our model. In some of those cases also the solution quality is slightly better. For both models the average optimality gap was almost vanishing that random generation is not a good way to create challenging instances. However, finding special structure that guarantees hardness is a major task in and of it self, and we do not pursue it in this text.

instance	avg. gap		max. gap		# best		avg. time	
	Cop	Adj	Cop	Adj	Cop	Adj	Cop	Adj
($\mathcal{SOC}$,5,5,1,1)	0	0	0.00	0.00	0	0	0.1345	0.1064
($\mathcal{SOC}$,5,5,3,3)	0	0	0.00	0.00	0	0	0.1608	0.1250
($\mathcal{SOC}$,10,10,2,2)	$\sim$0	$\sim$0	13.05	15.69	14	0	0.2621	0.5178
($\mathcal{SOC}$,10,10,5,5)	0	$\sim$0	0.00	3.29	3	0	0.4145	0.8909
($\mathcal{SOC}$,10,10,9,9)	0	0	0.00	0.00	0	0	0.6023	1.2655
($\mathcal{SOC}$,5,30,2,15)	$\sim$0	$\sim$0	0.15	2.00	4	0	1.4370	2.7077
($\mathcal{SOC}$,5,30,5,25)	0	0	0.00	0.00	0	0	1.8083	3.7375
($\mathcal{SOC}$,20,10,10,5)	$\sim$0	$\sim$0	4.33	7.10	7	0	1.4438	5.2094
($\mathcal{SOC}$,5,60,4,55)	$\sim$0	$\sim$0	25.26	15.88	1	2	6.4050	4.7414
($\mathcal{SOC}$,5,60,2,20)	$\sim$0	$\sim$0	1.84	98.40	14	0	6.3771	6.3557
($\mathbb{R}^n_+$,5,5,1,1)	0	$\sim$0	0.00	2.90	1	0	1.0736	1.8231
($\mathbb{R}^n_+$,5,5,3,3)	0	0	0.00	0.00	0	0	1.1228	1.8930
($\mathbb{R}^n_+$,10,10,2,2)	$\sim$0	$\sim$0	0.00	51.06	2	0	1.1978	2.9358
($\mathbb{R}^n_+$,10,10,5,5)	$\sim$0	$\sim$0	29.18	21.47	1	1	1.3812	3.4920
($\mathbb{R}^n_+$,10,10,9,9)	$\sim$0	$\sim$0	0.04	0.04	0	0	1.4416	3.7445
($\mathbb{R}^n_+$,5,30,2,15)	$\sim$0	$\sim$0	1.68	14.27	4	1	1.8561	8.2976
($\mathbb{R}^n_+$,5,30,5,25)	0	0	0.00	0.00	0	0	0.4588	0.6411
($\mathbb{R}^n_+$,20,10,10,5)	$\sim$0	$\sim$0	73.17	65.21	0	5	0.3362	4.3208
($\mathbb{R}^n_+$,5,60,4,55)	$\sim$0	$\sim$0	3.15	2.38	0	4	6.3774	4.1911
($\mathbb{R}^n_+$,5,60,2,20)	0	$\sim$0	0.00	67.42	8	0	6.0446	5.7693

Table 7.3: Comparison between the copositive and the adjustable robust lower bound.

Remark. *It should be noted that for* $\mathcal{CPP}\left(\mathcal{K}\times\mathbb{R}^n_2\right)$, $\mathcal{K}\in\{\mathbb{R}^{n_1}_+,\mathcal{SOC}^n\}$ *we could have also employed approximation hierarchies such as the ones described in [91], where a sequence of cones is provided that approximate* $\mathcal{CPP}\left(\mathcal{K}\times\mathbb{R}^n_2\right)$ *with increasing precision and where the computational cost of the respective bounds increases accordingly. One could argue that a more fair comparison between the two lower bounds would have been the one where we use those approximations for* $\mathcal{CPP}\left(\mathcal{K}\times\mathbb{R}^n_2\right)$ *where computational effort is similar to our proposed bound. However, these hierarchies are notoriously hard to work with and their computational complexity quickly explodes with precision. It is questionable whether we would have succeeded in fairly balancing computational effort.*

7.6.3 Comparison with linear policy for the bilinear case

As shown in [90] we have

$$\inf_{(\mathbf{x},\mathbf{y})\in\mathcal{X}\times\mathcal{Y}}\left\{\mathbf{x}^\mathsf{T}\mathsf{Q}_{xy}\mathbf{y}\right\}=\sup_{\tau,\mathbf{z}(\mathbf{x})}\left\{\tau\colon\tau\leq\mathbf{d}^\mathsf{T}\mathbf{z}(\mathbf{x}),\mathsf{F}^\mathsf{T}\mathbf{z}(\mathbf{x})\leq\mathsf{Q}_{xy}^\mathsf{T}\mathbf{x}\quad\forall\mathbf{x}\in\mathcal{X}\right\},\qquad(7.32)$$

where $\mathbf{z}(\mathbf{x})$ is a function-valued decision variable and $\mathcal{Y}$ is defined as above. If further $\mathcal{X}$ is defined as in Section 7.6.2.1 and we employ an affine decision rule, i.e. $\mathbf{z}(\mathbf{x})=\mathsf{Z}\mathbf{x}+\mathbf{z}$ we can reformulate the latter maximization problem via linear conic duality in a standard manner (see e.g. [7]) as to obtain a lower bound. In contrast, we, in Theorem 41 employ a quadratic decision rule for $\mathbf{z}(\mathbf{x})$ ($\mathbf{w}(\mathbf{x})$ does not appear if the convex quadratic term is dropped), which potentially tightens the bound. However, moving from a linear programming formulation or second order cone formulation to an SDP comes at a significant computational cost. In our experiment we aimed to quantify the benefit in order to see whether the trade off is justified. Table 7.4 summarizes the results of our experiments. We again present the average relative optimality gap (avg. gap) to an upper bound computed using `fmincon`, the maximum relative optimality gap (max. gap) to that upper bound, the number of times the our model outperformed the linear policy (# improved) and the average CPU-time (avg. time). All averages are taken across 100 replications per instance.

We see, that the linear policy is outperformed regularly, where the biggest advantages are achieved if m_2 is small. This makes sense, since the number of coefficients of the linear and the quadratic policy respectively depends on m_2 which in turn impacts the flexibility of these decision rules. Since the quadratic decision rule is inherently more flexible than the linear one, that impact is more pronounced for the latter. The computational cost for the benefit is however substantial, which is not surprising as

SDPs are known to scale poorly unless special structure is exploited. Such special structure is not present in the generic framework under which we operate here. To explore cases where that would be the cases is an interesting topic for future research.

instance	**avg. gap**		**max. gap**		**# improved**	**avg. time**	
	Lin	Quad	Lin	Quad		Lin	Quad
($\mathcal{SOC}$,5,5,1,1)	0.00	0.00	0.00	0.00	0	0.2165	1.2353
($\mathcal{SOC}$,5,5,3,3)	1.52	1.42	148.03	143.64	7	0.2268	1.2946
($\mathcal{SOC}$,10,10,2,2)	8.92	2.37	187.75	166.79	29	0.2268	1.2946
($\mathcal{SOC}$,10,10,5,5)	6.37	1.46	170.6	79.4	43	0.444	3.6402
($\mathcal{SOC}$,10,10,9,9)	$\sim$0.00	$\sim$0.00	0.15	0.15	0	0.3038	2.8084
($\mathcal{SOC}$,5,30,2,15)	12.41	9.12	792.11	791.30	66	0.3545	6.5923
($\mathcal{SOC}$,5,30,2,25)	14.36	5.59	636.58	448.94	63	0.3873	7.3362
($\mathcal{SOC}$,20,10,10,5)	6.78	1.64	153.94	146.37	52	0.4258	7.2313
($\mathcal{SOC}$,5,60,4,55)	0.02	0.02	2.21	2.21	0	0.3451	11.5162
($\mathcal{SOC}$,5,60,2,20)	6.21	3.29	86.37	85.51	85	0.4041	12.7623
($\mathbb{R}^n_+$,5,5,1,1)	0.00	0.00	0.00	0.00	0	0.4608	3.0671
($\mathbb{R}^n_+$,5,5,3,3)	0.00	0.00	0.00	0.00	0	0.4633	3.1145
($\mathbb{R}^n_+$,10,10,2,2)	0.00	$\sim$0.00	0.00	0.00	1	0.4779	4.7075
($\mathbb{R}^n_+$,10,10,5,5)	0.58	0.10	42.14	9.71	7	0.5141	5.1515
($\mathbb{R}^n_+$,10,10,9,9)	0.00	0.00	0.00	0.00	0	0.5883	5.7755
($\mathbb{R}^n_+$,5,30,2,15)	0.42	0.32	23.32	18.16	4	0.557	11.2174
($\mathbb{R}^n_+$,5,30,2,25)	0.33	0.31	14.21	14.21	2	0.1253	0.4108
($\mathbb{R}^n_+$,20,10,10,5)	4.26	0.86	39.28	27.33	54	0.1638	4.2258
($\mathbb{R}^n_+$,5,60,4,55)	$\sim$0.00	$\sim$0.00	0.27	0.27	0	0.4897	17.7068
($\mathbb{R}^n_+$,5,60,2,20)	5.31	2.77	89.22	86.80	69	0.5513	18.7613

Table 7.4: Comparison between linear and quadratic decision rules.

Part III

Uncertainty Preference: A new Concept for Robust Optimization

Chapter 8

Uncertainty preferences in robust linear optimization with endogenous uncertainty

8.1 Introduction

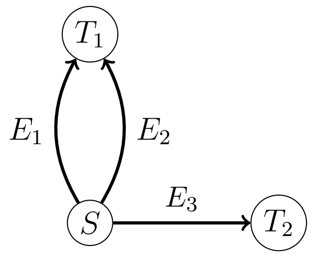
In this chapter, we will expand on the idea of endogenous uncertainty by introducing the notion of uncertainty preference. The framework of robust optimization with endogenous uncertainty allows us to model the trade-off between the cost of uncertainty reduction and worst-case performance. However, and this is the main motivation behind the present chapter, there is another trade-off to be made in such a scenario. Since we influence the structure of the uncertainty by our decision we do not only change the robustly optimal solution, we also affect other qualities such a solution exhibits, and a decision maker may have a certain preference with respect to that qualities. We chose the name *uncertainty preferences* as an overarching title for such preferences regarding the structure of the uncertainty set and the implied quality of robustly optimal solutions. We want clarify this concept in the following section.

8.1.1 The basic idea

In the following we want to illustrate the core idea of uncertainty preference in an easy, yet illustrative example.

Example 10. *Consider the graph depicted in Figure 8.1. The travel costs along the edges are uncertain, and only known to lie in a given interval. We can choose to influence uncertainty either on the edges that lead to T_1 or on edges that lead to T_2. We refer to the uncertainty regime where no influence on the uncertainty is taken as $\mathcal{U}_0$, the one where uncertainty is affected on paths towards T_i is referred to as*

$\mathcal{U}_i$, where $i = 1, 2$. Our aim is to choose a path from the source-vertex S to one of the sink-vertices T_1 or T_2 such that the worst case traveling time is minimized. The following table gives the travel costs on the edges under different uncertainty regimes.



Edge	$\mathcal{U}_0$	$\mathcal{U}_1$	$\mathcal{U}_2$
E_1	[2,5]	[3,4]	[2,5]
E_2	[2,5]	[2,4]	[2,5]
E_3	[2,5]	[2,5]	[1,4]

Figure 8.1

Unless we impose a cost on reducing uncertainty, $\mathcal{U}_0$ is the inferior choice to make. Thus, one would either choose $\mathcal{U}_1$ or $\mathcal{U}_2$ as the uncertainty regime under which to operate. If we choose to operate under $\mathcal{U}_1$, the two routes to T_1 have the same worst case travel cost 4. If we choose to travel along E_1, we face the chance of regretting our decision in case the best case materializes in on E_2. However, if we chose to traverse E_2 the outcome is more uncertain as the range of possible realizations is increased, which may be an undesirable feature of that choice. Under $\mathcal{U}_2$ the robustly optimal choice would be to travel to T_2 via E_3 incurring again a worst case cost of 4. This very choice would on the other hand lead to a cost of 1 in the best case, so that it also outperforms any other choice in the best case. Note that the maximum regret is still high and that the range of possible realizations has the largest attainable width of 3.

Example 10 shows that for different uncertainty regimes, robust solutions to the shortest path problem may exhibit additional features beyond robustness, which may be considered when evaluating such choices. Specifically, under different uncertainty regimes the respective robustly optimal solutions differ with respect to the *predictability* of the outcome, the risk of experiencing *regret* in hindsight, and the *best-case performance*. We can also see that under some uncertainty regimes, a robust solution is not unique, and some of the robustly optimal choices may also exhibit said distinguishing features. We conceptualize the notion of uncertainty preference as the desire to operate under uncertainty regimes where there is a robustly optimal solution that performs well with respect to predictability, regret and/or best-case performance. These attributes can be in conflict with each other, and a decision maker may be oblivious to some while valuing others. The goal of the discussion in this chapter is to provide a model that captures uncertainty preference in a robust optimization problem, and to give examples where these models yield manageable optimization problems which can be handled with standard methods such as mixed integer conic optimization.

Outline

Having now introduced the concept of uncertainty aversion, in Section 8.2 we formulate a mathematical model that captures the respective optimization problem. This leaves us with a bilevel optimization problem. In the top level we choose an uncertainty regime, while in the second level we restrict our decision to the robustly optimal solutions present under this uncertainty regime. The two challenges are thus enforcing optimality for the lower level problem (the robustly optimization problem under the chosen uncertainty regime), and incorporating the preferences addressed above by introducing suitable terms into the objective function. These are addressed in Section 8.2.1 and Section 8.2.2

In case the nominal problem is a linear optimization problem, we characterize the set of robustly optimal solutions based on robust optimization duality theory in Section 8.2.1.1. The resulting reformulation of the bilevel optimization problem is amenable to standard C-MILP solvers. In case the nominal problem is a MILP, we give a finite reformulation of the bilevel optimization problem with potentially exponential number of constraints in Section 8.2.1.2. These reformulations can be tackled using lazy constraints in conjunction with C-MILP solvers.

In Section 8.2.2 we describe the objective function terms that capture preferences for predictability, max-regret and best-case performance. Except for predictability, these formulations are in general not easy to handle. We highlight some special cases where we can achieve reformulations or approximations that can be solved with reasonable effort using C-MILP solvers.

We apply our framework to the shortest path problem under endogenous uncertainty in Section 8.3, which was considered before in [56]. The problem allows for different reformulations depending on the uncertainty set chosen. We consider different special cases of box-uncertainty sets.In numerical experiments in Section 8.4, we demonstrate that these reformulations are handled well by C-MILP solvers, and that the results we obtain are meaningful and can aid a decision process.

Before we continue with these tasks we give a more detailed introduction to endogenously uncertain mixed integer linear optimization problems. This type of robust optimization is central to our idea since the endogeneity of the uncertainty set give rise to multiple robustly optimal solution which can be distinguished based on the characteristics encompassed by the concept of uncertainty preference.

8.1.2 The endogenously uncertain mixed integer linear optimization problem

A central object this chapter is concerned with is the *endogenously uncertain (mixed integer) linear optimization problem*,

$$\inf_{\mathbf{x}} \left\{ \mathbf{c}(\mathbf{u})^{\mathsf{T}}\mathbf{x} \colon\ \mathbf{x} \in \mathcal{X}_0,\ \mathbf{a}_i(\mathbf{u})^{\mathsf{T}}\mathbf{x} \leq b_i(\mathbf{u}),\ i \in [1\!:\!m] \right\} \quad \text{where } \mathbf{u} \in \mathcal{U}^s, \text{ and } s \in [1\!:\!S], \tag{8.1}$$

where the problem data are affine functions of $\mathbf{u}$, i.e. $\mathbf{c}(\mathbf{u}) := \mathbf{c}_0 + \mathsf{C}\mathbf{u}$, $\mathbf{a}_i(\mathbf{u}) := \mathbf{a}_i^0 + \mathsf{A}_i\mathbf{u}$ and $b_i(\mathbf{u}) := b_i^0 + \mathbf{b}_i^{\mathsf{T}}\mathbf{u}$, $i \in [1\!:\!m]$ and $\mathcal{X}_0 \in \left\{\mathbb{R}^n_+,\ \mathbb{R}^{n_1}_+ \times \{0,1\}^{n_2}\right\}$. This problem was introduced in [62], and we will use the remainder of this sections to explain the technical details of their approach.

(8.1) describes a family of problems parametrized by $\mathbf{u}$ and s. The former is the *uncertainty parameter*, which is not under the control of the decision maker and only known to lie in one of the *uncertainty sets* $\mathcal{U}^s \subseteq \mathbb{R}^q$, $s \in [1\!:\!S]$. The latter sets depend on the second parameter s that in fact is controlled by the decision maker so that she can choose the *uncertainty regime* under which she would like to operate. The uncertainty is thus endogenous. Throughout this chapter we will model uncertainty sets as compact, convex conic intersections, i.e. $\mathcal{U}^s := \left\{\mathbf{u} \colon (1,\mathbf{u})^{\mathsf{T}} \in \mathcal{K}^s\right\}$ where $\mathcal{K}^s$ are appropriate, closed, convex, pointed cones. We will further comment on this modeling choice in a remark at the end of this section. Note that these assumptions, together with boundedness of $\mathcal{U}^s$ entail the following implication:

$$(0,\mathbf{x}^{\mathsf{T}})^{\mathsf{T}} \in \mathcal{K}^s \implies \mathbf{x} = \mathbf{o}\,. \tag{8.2}$$

Frequently, for example in case of row- and column-wise uncertainty (see Section 5.1.1), the uncertainty set will be a Cartesian product of sets $\mathcal{U}^s = \mathcal{U}^s_i \times \cdots \times \mathcal{U}^s_k$. The factors of such products are likewise modeled as conic intersections with cones $\mathcal{K}^s_i,\ i \in [1\!:\!k]$.

The choices a decision maker has on $\mathbf{x}$ and s may be entangled (see [62] for examples) so that a decision on $\mathbf{x}$ limits the options on s. Such a situation can always be modeled via constraints of the form

$$\mathsf{H}\mathbf{x} + \mathsf{G}\mathbf{y} \leq \mathsf{F}\mathbf{r}_s + \mathbf{f},$$

where $\mathbf{y}$ is an auxiliary (real or binary) variable $\mathsf{H}, \mathsf{G}, \mathsf{F}, \mathbf{f}$ are appropriate matrices/vectors and the entries of $\mathbf{r}$ are defined by

$$r_i^s = \begin{cases} 1 & \text{if } s = i \\ 0 & \text{else.} \end{cases} \tag{8.3}$$

However, we will omit these constraints from the models we will discuss for the sake of notational simplicity. Also, in some applications it will be practical to push the dependence on s into the functions $\mathbf{c}(\mathbf{u})$, $\mathbf{a}_i(\mathbf{u})$ and $b_i(\mathbf{u})$ while keeping the uncertainty sets constant (see Section 8.3).

We will employ several concepts associated with (8.1) namely: *row/column wise uncertainty*, the *nominal problem* and *robust/optimistic/deterministic counterparts*. For the readers convenience we give a brief review of these topics in Section 5.1.

The mathematical model describing the robust optimization problem under endogenous uncertainty is thus given by $\inf_{s\in[1:S]} \{\phi_s\}$ where

$$\begin{aligned} \phi_s \coloneqq \inf_{\mathbf{x}} \quad & \sup_{\mathbf{u}_0\in\mathcal{U}_0^s} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} \\ \text{s.t.}: \ & \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}\mathbf{x} \le b_i(\mathbf{u}_i), \ \forall \mathbf{u}_i \in \mathcal{U}_i^s, \ i \in [1\!:\!m], \\ & \mathbf{x} \in \mathcal{X}_0. \end{aligned} \tag{RC-s}$$

Hence, ϕ_s is defined by the robust optimization problem under the uncertainty regime indexed by s.

Note that the deterministic counterpart of (RC-s) is given by

$$\begin{aligned} \min_{\mathbf{x},\rho_i} \ & \mathbf{c}_0^\mathsf{T}\mathbf{x} + \rho_0 \\ \text{s.t.}: \ & \rho_i \le b_0^i - (\mathbf{a}_0^i)^\mathsf{T}\mathbf{x}, \ i \in [1\!:\!m], \\ \left(\rho_i, \mathbf{b}_i - \mathsf{A}_i^\mathsf{T}\mathbf{x}\right)^\mathsf{T} & \in (\mathcal{K}_i^s)^* \qquad , \ i \in [1\!:\!m], \\ \left(\rho_0, -\mathsf{C}^\mathsf{T}\mathbf{x}\right) & \in (\mathcal{K}_0^s)^* \\ \mathbf{x} & \in \mathcal{X}_0, \end{aligned} \tag{8.4}$$

Thus, solving $\inf_{s\in[1:S]} \{\phi_s\}$ amounts to solving an C-MILP. For the purpose of solving it, we can model the constraints $\mathbf{y} \in \mathcal{K}_i^s$ as

$$\mathbf{y} = \sum_{j=1}^S \boldsymbol{\delta}_j r_j, \quad \boldsymbol{\delta}_j \in \mathcal{K}_i^j, \ j \in [1\!:\!S],$$

which can always be exactly linearized using McCormick envelopes, as to obtain

$$\mathbf{y} = \sum_{j=1}^S \boldsymbol{\delta}_j, \quad \boldsymbol{\delta}_j \in \mathcal{K}_i^j, \ j \in [1\!:\!S],$$

$$\delta_1^j \leq M_j r_j,\ j \in [1\!:\!S],$$

for some suitable big $M_j \in \mathbb{R}$, $j \in [1\!:\!S]$ and $\mathbf{r} \in \mathbb{R}^S$ modeled as in the introduction. Here we also used the fact that by assumption (8.2) holds, so that it is enough to zero-out to first coordinate of $\boldsymbol{\delta}_j$ in order to have $\boldsymbol{\delta}_j = \mathbf{o}$. The choice of the big-M constants will of course depend on the problem at hand. In the main text, we do not resolve these bilinearities explicitly in the models we present, since that is a well known standard procedure. In [56], linearizations without big Ms are introduced, however in order to keep notation simple and the text focused we will not consider these here. Nonetheless, we do believe, that most of our reformulations can be achieved also under their linearization techniques.

Remark. *We would like to point out, that your modeling choice on $\mathcal{U}^s$, $\mathcal{K}^s$ and $s \in [1:S]$ was made with the desire in mind to have a readable notation that can, with little effort, capture very general structure. We are aware, that these are neither the most natural nor the most practical modeling choices for those objects if we think of more specific instances of robust optimization problems under endogenous uncertainty. Consider for example uncertainty sets $\hat{\mathcal{U}}(\hat{\mathbf{r}}) \coloneqq \{\mathbf{u} \in \mathbb{R}^q_+ \colon u_i \leq 1 + 0.5\hat{r}_i\}$ with $\hat{\mathbf{r}} \in \hat{\mathcal{X}} \coloneqq \left\{\hat{\mathbf{r}} \in \{0,1\}^{\hat{S}} : \mathsf{A}\hat{\mathbf{r}} \leq \mathbf{b}\right\}$ and assume the latter set is bounded. Then one can set $S = |\hat{\mathcal{X}}|$, construct an isomorphism $f \colon \hat{\mathcal{X}} \mapsto [1:S]$ and set $\mathcal{K}^s \coloneqq \operatorname{cone}\left(\{1\} \times \hat{\mathcal{U}}\left(f^{-1}(s)\right)\right)$ in order to obtain $\mathcal{U}^s = \hat{\mathcal{U}}\left(f^{-1}(s)\right)$ so that we have a reformulation that complies to our assumption but is arguably less practical. Firstly, the number of variables has increased to a size perhaps exponential in $\hat{n}_B$ and the size of A. Secondly the implicit constraint $\mathbf{e}^\mathsf{T}\mathbf{r} = 1$ present in our formulation can be much more challenging for MILP solvers than the inequalities in the description of $\hat{\mathcal{X}}$. For practical purposes it is therefore advisable to operate in the original space of variables. We chose our general formulation merely to be able to speak of general structure in simple terms.*

8.2 The model

We study linear and mixed integer linear optimization problems under endogenous uncertainty where the aim is to choose an uncertainty set such there is a robustly optimal solution that optimizes the trade off between worst-case performance, price of implementation of the uncertainty regime and the decision makers uncertainty preference. Thus we want to solve

$$
\begin{aligned}
&\inf_{\mathbf{x},s} \ \gamma_s + \alpha \sup_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\} + \beta \mathcal{M}(\mathbf{x}, s) \\
&\ \text{s.t.} : \mathbf{x} \in \mathcal{R}_{opt}^s, \ s \in [1\!:\!S].
\end{aligned} \tag{UPP}
$$

Here, α, β are weights given to the individual contributers to the objective function that depend on the affinity of the decision maker for either of those. The cost of activating uncertainty regime s is given by γ_s, which is thus a function of s. The term $\sup_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\}$ captures the worst-case performance of a solution $\mathbf{x} \in \mathcal{X}_0$ under uncertainty regime s. $\mathcal{R}_{opt}^s$ is the set of optimal solutions of the robust counterpart of (8.1) under uncertainty regime s (an explicit description of this robust counterpart is given by (RC-s) which can be found in Section 5.1.5), i.e. we define the set of optimal solutions to this problem to be

$$
\mathcal{R}_{opt}^s := \left\{ \mathbf{x} \in \mathbb{R}_+^n : \mathbf{x} \text{ is an optimal solution to (RC-s)} \right\}
$$

We will discuss in Section 8.2.1 how to work with the constraint $\mathbf{x} \in \mathcal{R}_{opt}^s$. Note, that (UPP) is feasible if and only if there is at least one $s \in [1\!:\!S]$ such that (RC-s) is feasible and attains its optimal value.

Another important part of our model is the function $\mathcal{M}(\mathbf{x}, s)$, which gives a measure of the performance of a solution (x, s) with respect to predictability, max-regret and best case performance. This function influences the decision in two ways: firstly, which uncertainty regimes is implemented and secondly, which robustly optimal solution is chosen in case the selected uncertainty regime allows for multiple robustly optimal solutions. We will explore different choices for $\mathcal{M}(\mathbf{x}, s)$ and study their reformulations in Section 8.2.2.

8.2.1 Working with the robustly optimal set $\mathcal{R}_{opt}^s$

The feasible set of the decision vector $\mathbf{x}$ is given by $\mathcal{R}_{opt}^s$, i.e. the set of solutions that are robustly optimal under the uncertainty regime indexed by s. Hence, we have to solve a bilevel optimization problem. This class of problems is known to be challenging, in fact it is strongly NP-hard, as shown in [48]. Even if the optimal set of a lower level problem can be characterized, its parametrization via the top level decision may lead to problem formulations that are not easily handled due to arising bilinear terms. Luckily, the top level decision on s is discrete, so that these bilinearities can be linearized exactly by using McCormick-Envelopes. Thus, the characterization of the robustly optimal set for a fixed choice on s is the main issue which we will address here in the sequel. For cases where $\mathcal{X}_0 = \mathbb{R}^n_+$ we will use robust duality theory to in order to characterize $\mathcal{R}^s_{opt}$. In case the $\mathcal{X}_0 = \mathbb{R}^{n_1}_+ \times \{0,1\}^{n_2}$, this theory does unfortunately not apply, so we provide alternative ways for these cases.

8.2.1.1 Robust duality for characterizing $\mathcal{R}^s_{opt}$

In this section we provide a characterization of $\mathcal{R}^s_{opt}$ for the case when $\mathcal{X}_0 = \mathbb{R}^n_+$. In recent literature many aspects of duality in robust optimization have been studied (see [5, 77]). While the deterministic counterpart of a robust optimization problem, as a convex problem, enjoys rich duality theory, a somewhat surprising observation first made in [5] is that the dual of the robust counterpart of an uncertain linear optimization problem, under mild assumptions, can be identified as the optimistic counterpart of the dual of that optimization problem (where the dualization is with respect to $\mathbf{x}$). Hence the phrase *"primal worst equals dual best"* coined in [5]. In case of (RC-s) the corresponding "dual best" problem is given by

$$\begin{aligned}
&\sup_{\mathbf{y},\mathbf{u}_i} \sum_{i=1}^m b_i(\mathbf{u}_i) y_i \\
&\text{s.t.} : \sum_{i=1}^m \mathbf{a}_i(\mathbf{u}_i) y_i \leq \mathbf{c}(\mathbf{u}_0), \\
&\qquad \mathbf{y} \in \mathbb{R}^m_+, \ \mathbf{u}_i \in \mathcal{U}^s_i, \ i \in [0\!:\!m].
\end{aligned} \tag{8.5}$$

We have the following theorem [5, Theorem 3.1]:

Theorem 43. *The problem (8.5) and the Lagrangian dual of (RC-s) are equivalent.*

Note that (RC-s) is convex in $\mathbf{x}$ as the inner superema are convex functions in $\mathbf{x}$. its deterministic counterpart (8.4) is arrived at merely by finding explicit forms of those functions. Thus the dual of (RC-s) is in fact given by the dual of (8.4) and any condition for (8.4) to exhibit strong duality also implies strong duality between (RC-s) and (8.5).

Corollary 2. *Assume, that strong duality holds between (RC-s) and (8.5) and that both problems attain their optimal value. Then the set of optimal solutions of (RC-s), i.e. $\mathcal{R}^s_{opt}$ is characterized by*

$$\begin{aligned}
\sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x} \right\} &\leq \sum_{i=1}^m b_i(\bar{\mathbf{u}}_i) y_i && \textit{(Zero Duality Gap)} \\
\mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}\mathbf{x} &\leq b_i(\mathbf{u}_i), \ \forall \mathbf{u}_i \in \mathcal{U}^s_i, \ i \in [1\!:\!m], \ \mathbf{x} \in \mathbb{R}^n_+ && \textit{(Primal Feasibility)} \\
\sum_{i=1}^m \mathbf{a}_i(\bar{\mathbf{u}}_i) y_i &\leq \mathbf{c}(\bar{\mathbf{u}}_0), \ \mathbf{y} \in \mathbb{R}^m_+, \ \bar{\mathbf{u}}_i \in \mathcal{U}^s_i, \ i \in [0\!:\!m]. && \textit{(Dual Feasibility)}
\end{aligned}$$

Proof. The proof is immediate from Theorem 43. □

The inequalities in the above corollary may seem difficult, but we have to bear in mind that all of these have convex reformulations. Minding the discussion succeeding Theorem 43 we can see that the assumptions in Corollary 2 can be guaranteed under well known conditions for full strong duality from convex duality theory. We give a summary in the sequel

Condtion 1. *We have strong duality between (RC-s) and (8.5), with attainment of the optimal value in both problems, if*

- $\mathcal{U}_i^s$, *all* $i \in [0\!:\!m]$, *are polyhedral sets and either (RC-s) or (8.5) is feasible with finite optimal value (strong LP duality) or,*
- *there is a feasible point for (RC-s) that fulfills the all inequalities strictly and (RC-s) attains its optimal value (Slater's condition) or,*
- *the feasible set of (RC-s) is bounded and (8.5) attains its optimal value (dual Slater's condition).*

At this point we want to highlight that not all optimistic counterparts are duals of robust counterparts as defined here. In fact, the structural aspect of (8.5) which allows for duality (and, thus, a convex reformulation) is that the uncertainty is column-wise (see [77, Theorem 1]). This fact will in Section 8.2.2.1 allow us to reformulate certain models involving max-regret.

8.2.1.2 Finite reformulations for combinatorial problems

We now consider the case where $\mathcal{X}_0 = \mathbb{R}^{n_1} \times \{0,1\}^{n_2}$ so that (RC-s) is a mixed integer robust optimization problem. In this case, robust duality theory does not apply. However, it is still possible to give a finite reformulation of (UPP) with possibly exponentially many constraints which can be tackled via cutting plane algorithms, especially by using lazy constraints in a branch and bound algorithm. The following theorem is easily checked:

Theorem 44. *Assume* $\mathcal{X}_0 = \mathbb{R}^{n_1} \times \{0,1\}^{n_2}$, *then* $(\mathbf{x}, s)$ *is feasible for (UPP) if and only if the following constraints hold:*

$$
\begin{aligned}
\sup_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\} &\leq \phi_j + (1 - r_j^s) M_j,\ j \in [1\!:\!S], \\
\mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} &\leq b_i(\mathbf{u}_i),\ \forall \mathbf{u}_i \in \mathcal{U}_i^s,\ i \in [1\!:\!m], \\
\mathbf{x} &\in \mathcal{X}_0,\ \ s \in [1\!:\!S],
\end{aligned}
\tag{8.6}
$$

where M_j are upper bounds for $\phi_{max} - \phi_j$, $j \in [1\colon S]$ with

$$\phi_{max} \coloneqq \sup_s \{\phi_s \colon s \in [1\colon S]\},$$

ϕ_j is the optimal value of (RC-s) under uncertainty regime j and $\mathbf{r}_s$ is defined as in (8.3).

Proof. Let $(\mathbf{x}, s)$ be feasible for (UPP), then $s \in [1\colon S]$ and $\mathbf{x} \in \mathcal{R}^s_{opt}$ so that it fulfills the last two constraints of (8.6). Further, by this fact, we have $\sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} = \phi_s \leq \phi_{max} \leq \phi_k + M_k$ for all $k \in [1\colon S] \setminus \{s\}$. Thus $(\mathbf{x}, s)$ is feasible for (8.6). For the converse assume $(\mathbf{x}, s)$ is feasible for (8.6) so that all but the s-th constraint are redundant and by definition and by feasibility we have $\phi_s \leq \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} \leq \phi_s$. So that indeed $\mathbf{x} \in \mathcal{R}^s_{opt}$, and $(\mathbf{x}, s)$ is feasible for (UPP). ☐

Finding appropriate M_j, all $j \in [1\colon S]$ may be a challenge that has to be met using special structure of the problem at hand. As we speak here in quite general terms, we do not have this option here. However, if (RC-s) is such that the uncertainty in the constraints does not depend on s, we can achieve a general reformulation that does not require a big-M. Note, that this covers the case, where there is no uncertainty in the constraints.

Theorem 45. *Assume $\mathcal{X}_0 = \mathbb{R}^{n_1} \times \{0,1\}^{n_2}$ and $\mathcal{U}^s_i = \mathcal{U}_i$ for $i \in [1\colon m]$, i.e. only the uncertainty in the objective of (8.1) is affected by s, then $(\mathbf{x}, s)$ if feasible for (UPP) if and only if the following constraints hold:*

$$\begin{aligned}\sup_{\mathbf{u}_0 \in \mathcal{U}^j_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} \leq \sup_{\mathbf{u}_0 \in \mathcal{U}^j_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{z}_j\right\},\ j \in [1\colon S],\\ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}\mathbf{x} \leq b_i(\mathbf{u}_i),\ \forall \mathbf{u}_i \in \mathcal{U}_i,\ i \in [1\colon m],\\ \mathbf{x} \in \mathcal{X}_0,\ \ s \in [1\colon S],\end{aligned} \tag{8.7}$$

where $\mathbf{z}_j \in \mathcal{R}^j_{opt}$, $j \in [1\colon S]$.

Proof. Assume $(\mathbf{x}, s)$ is feasible for (UPP), then $\mathbf{x} \in \mathcal{R}^s_{opt}$ so that it fulfills the last two constraints of (8.7). Also by this fact we have that $\phi_s = \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} = \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{z}_j\right\} \leq \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{z}_k\right\}$ for all $k \in [1\colon S] \setminus \{s\}$, since all $\mathbf{z}_k$ are feasible for (RC-s). Thus $(\mathbf{x}, s)$ is feasible for (8.7). Conversely, assume $(\mathbf{x}, s)$ are feasible for (8.7). We have $\sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{z}_s\right\} \leq \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} \leq \sup_{\mathbf{u}_0 \in \mathcal{U}^s_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{z}_s\right\}$ where the first inequality holds since $\mathbf{x}$ is feasible for ϕ_s and the second holds by feasibility for (8.7). So that indeed $\mathbf{x} \in \mathcal{R}^s_{opt}$, and $(\mathbf{x}, s)$ is feasible for (UPP). ☐

8.2.2 Modeling uncertainty preference via $\mathcal{M}(\mathbf{x}, s)$

In this section we will discuss different choices for $\mathcal{M}(\mathbf{x}, s)$ and various reformulations of the resulting optimization problems. The goal is to have exact reformulations or at least good relaxations that are amendable to various, general off-the-shelf solution techniques such as mixed integer linear and conic optimization. For some of the models, row-wise uncertainty can be assumed without loss of generality, while in others that is not generally possible. We will point specifically, when this assumption can be made safely and when this is not the case.

8.2.2.1 Regret

However, our aim is not to optimize the maximum regret in an uncertain linear problem. Rather we seek a choice on the uncertainty regime such that the robustly optimal solution performs well with respect to its maximum regret.

Taking into account, that a decision maker wants to make a choice on the uncertainty such that possible regret is minimized, we now consider a model that introduces a cost for running the risk of high regret. For fixed $s \in [1\colon S]$ and $\mathbf{x} \in \mathcal{R}^s_{opt}$ the worst-case regret associated with this solution is given by

$$\rho(\mathbf{x}, s) := \sup_{\mathbf{u}\in\mathcal{U}^s} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}\in\mathbb{R}^n_+} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{y} \le b_i(\mathbf{u}),\ i \in [1\colon m] \right\} \right\},$$

that is, the worst-case difference between the performance of $\mathbf{x}$ and the performance we could have achieved in hindsight. The regret function does not have an equivalent row-wise formulation based on the projection method, which is valid for the robust counterpart.

The model that accounts for the worst-case regret of the outcome of the uncertainty under a chosen uncertainty regime can be formulated as:

$$\begin{aligned}
&\inf_{\mathbf{x},s,\rho} \gamma_s + \alpha \sup_{\mathbf{u}_0\in\mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x} \right\} + \beta(\rho) \\
&\text{s.t.} : \rho \ge \sup_{\mathbf{u}\in\mathcal{U}^s} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}\in\mathbb{R}^n_+} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{y} \le b_i(\mathbf{u}),\ i \in [1\colon m] \right\} \right\}, \\
&\qquad \mathbf{x} \in \mathcal{R}^s_{opt},\ s \in [1\colon S].
\end{aligned} \tag{REG-UPP}$$

Here, $\beta(\cdot)$ is a convex, non-decreasing function which introduces a cost on the regret associated with a chosen solution and ρ is an auxiliary epigraphical variable. Easy viable choices for $\beta(\cdot)$ are the identity and simple linear scalings, but more complicated choices are possible without leaving the realm of convex optimization. The last constraint, despite being convex, presents a major complication in general although in special cases there are manageable reformulations and approximations. In the remainder of the section we discuss problems with appropriate special structure. In Section 8.3.2.1 we also use a reformulation from literature that is specific to the shortest path problem.

Exact reformulation in case of column-wise uncertainty

Note, that the innermost supremum-problem in (REG-UPP) is the optimistic counterpart of the uncertain linear problem (plus an additional robust objective term), which in general can be a nonconvex problem but in special cases is in fact convex, by robust duality theory. Obviously, this is a useful circumstance for reformulating (REG-UPP) , but caution must be taken since not every optimistic counterpart is the dual of a robust optimization problem (see e.g [5, Example 2.1, Remark 3.1]) and therefore convex. In fact, it was proved in [77, Theorem 1] that an optimistic counterpart of an uncertain linear optimization problem with column-wise uncertainty and non-negative decision vector has a convex feasible set, that is, regarding the decision variable $\mathbf{x}$. Further, it was proved that such a problem is the dual of a robust counterpart of an uncertain linear optimization problem with row-wise uncertainty. Thus, in order for the minimization in the definition of $\rho(\mathbf{x}, s)$ to be the "dual best" of a robust linear optimization problem we need the following assumption on the underlying uncertain problem.

Assumption 1. *The underlying uncertain linear optimization problem has*

a) no uncertainty in the objective, i.e. $\mathbf{c}(\mathbf{u}) = \mathbf{c}_0$*,*

b) column-wise uncertainty,

c) the domain of $\mathbf{x}$ *given by* $\mathcal{X}_0 = \mathbb{R}^n_+$*,*

and can thus be written as

$$\inf_{\mathbf{x}} \left\{ \mathbf{c}_0^\mathsf{T}\mathbf{x} \colon \mathbf{x} \in \mathbb{R}^n_+,\ \sum_{i=1}^{n} \mathbf{a}_i(\mathbf{u}_i) x_i \leq \mathbf{b}(\mathbf{u}_{n+1}) \right\} \text{ where } \mathbf{u}_i \in \mathcal{U}_i^s,\ i \in [1 : n+1]. \tag{8.8}$$

where $\mathbf{b}(\mathbf{u}_{n+1}) \coloneqq \mathbf{b}_0 + \mathsf{B}\mathbf{u}_{n+1}$*. For the above formulation we assume that*

d1) either there is an $\mathbf{x} \in \mathbb{R}^n_+$ *such that there exist* $\mathbf{u}_i \in \operatorname{int}(\mathcal{U}_i^s)$ *, all* $i \in [1 : n+1]$ *so that the inequalities are fulfilled, where the* int*-operator can dropped if the set is polyhedral, or* *d2) problem (8.10), i.e. the dual of (8.8) in the robust duality sense, has a Slater point.*

The following lemma provides a useful sufficient condition for Assumption 1 *d1)*:

Lemma 46. *Consider the following condition on (8.8):*

d') There is an $\mathbf{x} \in \mathbb{R}^n_+$ such that there are $\mathbf{u}_i \in \mathcal{U}^s_i$, all $i \in [1:n+1]$ so that the inequalities are fulfilled strictly.

We have that d') implies Assumption 1 d1).

Proof. Take any $\mathbf{u}_i^0 \in \operatorname{int}(\mathcal{U}^s_i)$, all $i \in [1:n+1]$, then for $\bar{\mathbf{u}}_i \coloneqq \mathbf{u}_i^0 - \mathbf{u}_i$ we have $\mathbf{u}_i + \lambda\bar{\mathbf{u}}_i \in \operatorname{int}(\mathcal{U}^s_i)$ for any $\lambda \in (0,1]$ by [73, Theorem 6.1]. Define $\boldsymbol{\varepsilon} \coloneqq \mathbf{b}(\mathbf{u}_{n+1}) - \sum_{i=1}^n \mathbf{a}_i(\mathbf{u}_i)x_i$. Since by assumption $\boldsymbol{\varepsilon} \in \operatorname{int}(\mathbb{R}^m_+)$, we can choose $\lambda \in (0,1]$ such that $\lambda\left(\sum_{i=1}^n \mathsf{A}_i\bar{\mathbf{u}}_i x_i - \mathsf{B}\bar{\mathbf{u}}_{n+1}\right) \leq \boldsymbol{\varepsilon}$. We then have

$$\begin{aligned}
&\sum_{i=1}^n \mathbf{a}_i(\mathbf{u}_i + \lambda\bar{\mathbf{u}})x_i - \mathsf{B}\left(\mathbf{u}_{n+1} + \lambda\bar{\mathbf{u}}_{n+1}\right)\\
=&\sum_{i=1}^n \left(\mathbf{a}_i^0 + \mathsf{A}_i(\mathbf{u}_i + \lambda\bar{\mathbf{u}})\right)x_i - \mathsf{B}\left(\mathbf{u}_{n+1} + \lambda\bar{\mathbf{u}}_{n+1}\right)\\
=&\sum_{i=1}^n \left(\mathbf{a}_i^0 + \mathsf{A}_i\mathbf{u}_i\right)x_i - \mathsf{B}\mathbf{u}_{n+1} + \lambda\left(\sum_{i=1}^n \mathsf{A}_i\bar{\mathbf{u}}x_i - \mathsf{B}\bar{\mathbf{u}}_{n+1}\right)\\
\leq&\sum_{i=1}^n \mathbf{a}_i(\mathbf{u}_i)x_i - \mathsf{B}\mathbf{u}_{n+1} + \boldsymbol{\varepsilon} = \mathbf{b}_0
\end{aligned}$$

and $\mathbf{u}_i + \lambda\bar{\mathbf{u}}_i \in \operatorname{int}(\mathcal{U}^s_i)$, $i \in [1:n+1]$. □

We can generally assume that $\{\mathbf{o}\} \in \operatorname{int}(\mathcal{U}^s)$ since we can always shift the set and adapt the constant terms in $\mathbf{a}(\mathbf{u}), \mathbf{c}(\mathbf{u})$ and $\mathbf{b}(\mathbf{u})$ accordingly. Thus, d') holds if the feasible set of the nominal problem has nonempty interior, which often is established more easily.

Note that the optimistic counterpart of (8.8) given by

$$\begin{aligned}
\inf_{\mathbf{x},\mathbf{u}}\quad & \mathbf{c}_0^{\mathsf{T}}\mathbf{x}\\
& \sum_{i=1}^n \mathbf{a}_i(\mathbf{u}_i)x_i \leq \mathbf{b}(\mathbf{u}_{n+1}),\\
& \mathbf{x} \in \mathbb{R}^n_+,\ \mathbf{u}_i \in \mathcal{U}^s_i,\ i \in [1:n+1],
\end{aligned} \tag{8.9}$$

is precisely of the form (8.5) so that by [5, Theorem 3.1] we can close the duality gap between (8.9) and

$$\begin{aligned}
&\sup_{\mathbf{z}} \inf_{\mathbf{u}_{n+1}\in\mathcal{U}^s_{n+1}} \left\{\mathbf{b}(\mathbf{u}_{n+1})^\mathsf{T}\mathbf{z}\right\}\\
&\quad c_i^0 \geq \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T}\mathbf{z}, \;\; \forall \mathbf{u}_i \in \mathcal{U}_i^s\,, \;\; i\in[1\!:\!n],\\
&\quad -\mathbf{z}\in\mathbb{R}^m_+,
\end{aligned} \tag{8.10}$$

if regularity conditions are met (e.g. Slater's condition or polyhedrality). Mindful of this fact, we see that Assumption 1 suffices to prove the following theorem.

Theorem 47. *Assume that Assumption 1 holds. Then (REG-UPP) is equivalent to the following optimization problem*

$$\begin{aligned}
\inf_{\mathbf{x},\mathbf{z},s,\rho}\; & \gamma_s + \alpha\mathbf{c}_0^\mathsf{T}\mathbf{x} + \beta(\rho)\\
\text{s.t.}:\; & \rho \geq \mathbf{c}_0^\mathsf{T}\mathbf{x} - \mathbf{b}^\mathsf{T}\mathbf{z} + \tau,\\
\left(\tau, \mathbf{z}^\mathsf{T}\mathsf{B}\right)^\mathsf{T} &\in \left(\mathcal{K}^s_{n+1}\right)^*,\\
\left(c_i^0 - (\mathbf{a}_i^0)^\mathsf{T}\mathbf{z},\; -\mathbf{z}^\mathsf{T}\mathsf{A}_i\right)^\mathsf{T} &\in (\mathcal{K}_i^s)^*\,,\; i\in[1\!:\!n],\\
-\mathbf{z}\in\mathbb{R}^n_+,\;\; & \mathbf{x}\in\mathcal{R}^s_{opt},\;\; s\in[1\!:\!S].
\end{aligned} \tag{8.11}$$

Proof. Recall the definition of the max-regret function

$$\rho(\mathbf{x},s) = \sup_{\mathbf{u}\in\mathcal{U}^s}\left\{\mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}\in\mathbb{R}^n_+}\left\{\mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{y}\colon \mathbf{a}_i(\mathbf{u})^\mathsf{T}\mathbf{y}\leq b_i(\mathbf{u}),\; i\in[1\!:\!m]\right\}\right\},$$

which under Assumption 1 reduces to

$$\rho(\mathbf{x},s) = \mathbf{c}_0^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y},\mathbf{u}_i}\left\{\mathbf{c}^\mathsf{T}\mathbf{y}\colon \mathbf{y}\in\mathbb{R}^n_+,\; \mathbf{u}_i\in\mathcal{U}_i^s,\; \sum_{i=1}^n\left(\mathbf{a}_i^0 + \mathsf{A}_i\mathbf{u}_i\right)y_i \leq \mathbf{b}_0 + \mathsf{B}\mathbf{u}_{n+1}\right\}. \tag{8.12}$$

We will first show equivalence between the above infimum and

$$\inf_{\mathbf{w}_0,\mathbf{w}_i,\mathbf{u}_{n+1}}\left\{\mathbf{c}_0^\mathsf{T}\mathbf{w}_0\colon \left(w_i^0,\; \mathbf{w}_i^\mathsf{T}\right)^\mathsf{T}\in(\mathcal{K}_i^s)^*\,,\; \mathbf{u}_{n+1}\in\mathcal{U}^s_{n+1},\; \sum_{i=1}^n \mathbf{a}_i^0 w_i^0 + \mathsf{A}_i\mathbf{w}_i \leq \mathbf{b}_0 + \mathsf{B}\mathbf{u}_{n+1}\right\}, \tag{8.13}$$

minding the fact that by assumption, $\mathcal{U}_i^s = \left\{\mathbf{u}_i\colon \left(1, \mathbf{u}_i^\mathsf{T}\right)^\mathsf{T}\in\mathcal{K}_i^s\right\}$ are compact. If $(\mathbf{y},\mathbf{u}_1,\ldots,\mathbf{u}_{n+1})$ is feasible for (8.12), we set $w_i^0 = y_i$ and $\mathbf{w}_i = y_i\mathbf{u}_i$ for $i\in[1:n]$ to get an equivalent solution for (8.13). Conversely, assuming $(\mathbf{w}_0,\mathbf{w}_1,\ldots,\mathbf{w}_n)$ is feasible for (8.13) we set $y_i = w_i^0$ and $\mathbf{u}_i = \frac{\mathbf{w}_i}{w_i^0}$ if $w_i^0\neq 0$ and $\mathbf{u}_i = \mathbf{o}$ otherwise to obtain an equivalent solution for (8.12) (here we use

the fact that (8.2) holds by assumption). By the same construction, any feasible point in (8.12) with $\mathbf{u}_i \in \operatorname{int}(\mathcal{U}_i^s)$ translates to a Slater point in (8.13). So, in any case, by Assumption 1 d) we have strong duality with its dual given by $\sup_{\mathbf{z}} \left\{ \mathbf{b}^\mathsf{T}\mathbf{z} - \tau \colon -\mathbf{z} \in \mathbb{R}^n_+,\ \left(\tau, \mathbf{z}^\mathsf{T}\mathsf{B}\right)^\mathsf{T} \in \left(\mathcal{K}^s_{n+1}\right)^*,\ \left(c_i^0 - (\mathbf{a}_i^0)^\mathsf{T}\mathbf{z},\ -\mathbf{z}^\mathsf{T}\mathsf{A}_i\right)^\mathsf{T} \in (\mathcal{K}_i^s)^*, i \in [1\colon n] \right\}$.
□

The final supremum problem in the above proof is actually the deterministic counterpart of the "primal worst" (8.10). Thus the above theorem actually rests on robust duality. Rather than referencing this in the proof, we chose to give a simplified, exhaustive proof, firstly for the readers convenience and secondly to make it transparent how Assumption 1 ties in with the final result.

Remark. *A source of faulty intuition is given by the well known fact that for an uncertain (mixed integer) linear optimization problem where the objective is uncertain the robust counterpart can be reformulated, by means of an epigraphical variable t, so that only the constraints are uncertain that is*

$$\begin{aligned}
&\min_{x \in \mathcal{X}_0} \left\{ \sup_{\mathbf{u} \in \mathcal{U}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \right\} : \mathsf{A}(\mathbf{u})\mathbf{x} \le \mathbf{b}(\mathbf{u}) \textit{ for all } \mathbf{u} \in \mathcal{U} \right\} \\
= &\min_{x \in \mathcal{X}_0,\ t \in \mathbb{R}} \left\{ t \colon \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \le t, \mathsf{A}(\mathbf{u})\mathbf{x} \le \mathbf{b}(\mathbf{u}) \textit{ for all } \mathbf{u} \in \mathcal{U} \right\}.
\end{aligned}$$

One might be tempted to assume that the same is possible for the max-regret function, so that Assumption 1 a) can be regarded without loss of generality. However, such an epigraphical reformulation would not resolve the issue that necessitates Assumption 1 a). The reformulation rests on the fact that under Assumption 1 the optimization problem defining the max-regret function is the optimistic counterpart of a column-wise uncertain linear optimization problem. This would no longer be true if Assumption 1 a) was omitted. Column-wise uncertainty is given if for all $i \in [0\colon n]$ the uncertainty parameter $\mathbf{u}_i$ appears always as a factor of the i-th decision variable or always stands alone (as for the right hand side uncertainty). The optimistic counterpart in the max-regret function would, in case of an uncertain objective, exhibit portions of the uncertainty parameters that stand alone (in the term $\mathbf{c}(\mathbf{u})\mathbf{x}$), but also interact with the decision variable $\mathbf{y}$ (in the term $\mathbf{c}(\mathbf{u})\mathbf{y}$) and this cannot be resolved by moving the objective term into the constraints.

LP-relaxation for the max-regret function for combinatorial problems with uncertain objective and box-uncertainty

Consider the uncertain mixed-integer linear problem

$$\inf_{\mathbf{x} \in \{0,1\}^n} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \colon \mathsf{E}\mathbf{x} = \mathbf{d} \right\} \quad \text{where } \mathbf{u} \in \mathcal{U}.$$

The robust counterparts of uncertain MILPs in general do not enjoy strong duality and thus we not have an easy way to resolve the inner supremum problem in (REG-UPP) outside of special cases. In Section 8.3 we will study such cases, where the constraint matrix $\mathsf{E} \in \mathbb{R}^{m\times n}$ is totally unimodular. For discussion in this section we will solely focus on deriving a tractable relaxation of the max-regret function of the uncertain MILP, but already entertain the unimodularity assumption on E. We have

$$\rho(\mathbf{x}, s) = \sup_{\mathbf{u}\in\mathcal{U}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}\in\{0,1\}^n} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathsf{E}\mathbf{y} = \mathbf{d} \right\} \right\} .$$

Assume that $\mathbf{c}(\mathbf{u}) \coloneqq \mathbf{c}_0 + \mathsf{C}\mathbf{u}$ and that $\mathcal{U} \coloneqq [-1,1]^q$. Denoting $\bar{\mathsf{C}} \coloneqq (\mathsf{C}, -\mathsf{C})$, we have

$$\begin{aligned} \rho(\mathbf{x}, s) &= \sup_{\mathbf{u}\in[-1,1]^q, \mathbf{y}\in\mathbb{R}^n_+} \left\{ (\mathbf{c}_0 + \mathsf{C}\mathbf{u})^\mathsf{T}\mathbf{x} - \ (\mathbf{c}_0 + \mathsf{C}\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathsf{E}\mathbf{y} = \mathbf{d} \right\} \\ &= \sup_{\mathbf{u}\in[0,1]^{2q}, \mathbf{y}\in\mathbb{R}^n_+} \left\{ (\mathbf{c}_0 + \bar{\mathsf{C}}\mathbf{u})^\mathsf{T}\mathbf{x} - \ (\mathbf{c}_0 + \bar{\mathsf{C}}\mathbf{u})^\mathsf{T}\mathbf{y} \colon \mathsf{E}\mathbf{y} = \mathbf{d} \right\} . \end{aligned}$$

The first equality holds due to the total unimodularity assumption an the fact that the objective is bilinear. The second one is easily established. We see, that the optimal solutions are contained in $\{0,1\}^{n+2q}$. First, we introduce additional constraints by multiplying the linear constraints with each of the uncertainty parameters, i.e. $\mathsf{E}\mathbf{y}u_i = \mathbf{d}u_i$, $i \in [1\colon q]$. We now lift the space of variables by replacing the bilinear terms by $u_i y_j \to \psi_{ij}$ and construct the McCormick envelope by introducing the constraints

$$\begin{aligned} &\psi_{ij} \leq u_j,\ j \in [1\colon q], \\ &\psi_{ij} \leq y_i,\ i \in [1\colon n], \\ &\psi_{ij} \geq y_j + u_j - 1,\ j \in [1\colon q],\ \ i \in [1\colon n] . \end{aligned}$$

We end up with a linear optimization problem with nonempty feasible set so that it has a representation of the form below and a respective dual

$$\sup_{\mathbf{z}\geq\mathbf{o}} \left\{ \mathbf{f}(\mathbf{x}, s)^\mathsf{T}\mathbf{z} \colon \mathsf{A}\mathbf{z} = \mathbf{b} \right\} = \inf_{\mathbf{w}} \left\{ \mathbf{b}^\mathsf{T}\mathbf{w} \colon \mathsf{A}^\mathsf{T}\mathbf{w} \geq \mathbf{f}(\mathbf{x}, s) \right\} .$$

The final inequality in (REG-UPP) can, thus, be replaced by

$$\rho \geq \mathbf{b}^\mathsf{T}\mathbf{w},\ \mathsf{A}^\mathsf{T}\mathbf{w} \geq \mathbf{f}(\mathbf{x}, s),$$

as to obtain an upper bound for the original problem.

Remark. *The above construction can be seen as an application of the relaxation/reformulation technique outlined in [75]. Since we multiplied the constraints* *only with all the variables from the vector* $\mathbf{u}$ *(whereas in the reference products of powers of all variables are used, where the relaxation improves with the degree and eventually tightens up when the degree equals the number of variables), our construction is actually more conservative than the weakest relaxation described in that reference. Based on their scheme we could have obtained a tight reformulation of the regret-function, but at the cost of introducing an exponential number of constraints, and thus, dual variables. Our construction was motivated by the application to the endogenously uncertain shortest path problem (see Section 8.3 and Section 6.3) and exhibits satisfying performance there.*

8.2.2.2 Predictability

We have constructed uncertainty regimes where the performance of the robust solution is more predictable than in others. This may be a desired property of a solution next to being robust, for example, if, after implementation of the solution, excess optimal value cannot be harnessed due to its unexpected realization (e.g. a delivery that arrives early and cannot be received). The desire to have a predictable outcome can also extend to the constraints. Being able to better estimate the resource consumption of an activity outside the worst case can be advantageous. Thus, a decision maker may prefer to operate under an uncertainty regime where the performance of the robust solution after implementation does not deviate from the robustly optimal value by too much, and where the resource consumption does not deviate too much across scenarios.

Motivated by the desire shrink the gap between best-case and worst-case performance of a robust solution, so to have a predictable outcome, we introduce a the following model:

$$\begin{aligned}
\inf_{\mathbf{x},s,\overline{q}_i,\underline{q}_i} \ & \gamma_s + \alpha \sup_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\} + \sum_{i=0}^{m} \beta_i(\overline{q}_i, \underline{q}_i) \\
\text{s.t.}: \ & \overline{q}_0 \geq \sup_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\}, \\
& \underline{q}_0 \leq \inf_{\mathbf{u}_0 \in \mathcal{U}_0^s} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T} \mathbf{x} \right\}, \\
& \overline{q}_i \geq \sup_{\mathbf{u}_i \in \mathcal{U}_i^s} \left\{ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} - b_i(\mathbf{u}_i) \right\}, \ i \in [1:m], \\
& \underline{q}_i \leq \inf_{\mathbf{u}_i \in \mathcal{U}_i^s} \left\{ \mathbf{a}_i(\mathbf{u}_i)^\mathsf{T} \mathbf{x} - b_i(\mathbf{u}_i) \right\}, \ i \in [1:m], \\
& \mathbf{x} \in \mathcal{R}_{opt}^s, \ s \in [1:S] \,.
\end{aligned} \tag{PDA-UPP}$$

Here, $\overline{q}_i$ and $\underline{q}_i$ are auxiliary epi- and hypographical variables, which we use here in order to have a more readable notation. The right hand side suprema/infima are the worst- and best-case performances of a solution $\mathbf{x}$ under uncertainty regime s in terms of objective function value and in terms of resource consumption. Since the differences between worst and best-case performances are independent across rows, we may assume row wise uncertainty, in a way similar to the robust counterpart. Note, that the suprema are convex functions in $\mathbf{x}$ as they are point wise suprema of linear functions. By the same token, the infima are concave in $\mathbf{x}$. Thus, the respective constraints describe a convex set. The functions $\beta_i(\overline{q}_i, \underline{q}_i)$ are convex function measuring difference between the two arguments. Next to convexity we assume

$$\beta_i(\overline{q}_i, \underline{q}_i) \geq 0, \quad \frac{\partial \beta_i(\overline{q}_i, \underline{q}_i)}{\partial \overline{q}_i} > 0, \quad \frac{\partial \beta(\overline{q}_i, \underline{q}_i)}{\partial \underline{q}_i} < 0 \text{ for all } \overline{q}_i \geq \underline{q}_i. \tag{8.14}$$

Thus, $\beta_i(.,.)$ are convex functions that put a cost on the distance between worst and best-case performance of a feasible solution $\mathbf{x}$. An easy choice for β_i would be $\beta_i := \overline{q}_i - \underline{q}_i$ and in our applications we will exclusively work with this choice. However, more complicated choices are possible since in any case we will end up with a mixed integer convex optimization problem.

Theorem 48. *Problem (PDA-UPP) is equivalent to the following mixed integer conic optimization problem:*

$$\begin{aligned}
\inf_{\mathbf{x},s,\rho_i} \quad & \gamma_s + \alpha\left(\mathbf{c}_0^\mathsf{T}\mathbf{x} + \rho_0\right) + \beta_0(\mathbf{c}_0^\mathsf{T}\mathbf{x} + \rho_0, \mathbf{c}_0^\mathsf{T}\mathbf{x} - \delta_0) + \sum_{i=1}^m \beta_i\left((\mathbf{a}_0^i)^\mathsf{T}\mathbf{x} - b_0^i + \rho_i, (\mathbf{a}_0^i)^\mathsf{T}\mathbf{x} - b_0^i - \delta_i\right) \\
& \text{s.t.}: \left(\rho_0, -\mathbf{x}^\mathsf{T}\mathsf{C}\right)^\mathsf{T} \in (\mathcal{K}_0^s)^*, \\
& \qquad \left(\delta_0, \mathbf{x}^\mathsf{T}\mathsf{C}\right)^\mathsf{T} \in (\mathcal{K}_0^s)^*, \\
& \left(\rho_i, \mathbf{b}_i^\mathsf{T} - \mathbf{x}^\mathsf{T}\mathsf{A}_i\right)^\mathsf{T} \in (\mathcal{K}_i^s)^*, \ i \in [1:m] \\
& \left(\delta_i, \mathbf{x}^\mathsf{T}\mathsf{A}_i - \mathbf{b}_i^\mathsf{T}\right)^\mathsf{T} \in (\mathcal{K}_i^s)^*, \ i \in [1:m] \\
& \qquad \mathbf{x} \in \mathcal{R}_{opt}^s, \ s \in [1:S].
\end{aligned} \tag{8.15}$$

Proof. The proof follows immediately from standard arguments based on convex duality that are well known in robust optimization. For the readers convenience we present the key idea in the notation used here. Frist, by invoking convex duality we have

$$\sup_{\mathbf{u}\in\mathcal{U}} \left\{\mathbf{d}^\mathsf{T}\mathbf{u}\right\} = \sup_{\mathbf{u}} \left\{\mathbf{d}^\mathsf{T}\mathbf{u} \colon \left(1, \mathbf{u}^\mathsf{T}\right)^\mathsf{T} \in \mathcal{K}\right\} = \inf_{\rho} \left\{\rho \colon \left(\rho, -\mathbf{d}^\mathsf{T}\right)^\mathsf{T} \in \mathcal{K}^*\right\}.$$

Here strong duality holds, since we assumed that the uncertainty sets have nonempty interior and thus Slaters condition holds. We thus have the equivalence

$$\sup_{\mathbf{u}\in\mathcal{U}} \left\{\mathbf{d}^\mathsf{T}\mathbf{u} + d_0\right\} \leq 0 \;\Leftrightarrow\; d_0 + \rho \leq 0,\; (\rho, -\mathbf{d})^\mathsf{T} \in \mathcal{K}^*.$$

The reformulation follows by using this equivalence with $\mathbf{d}$ replaced by $\mathbf{x}^\mathsf{T}\mathsf{A}_i - \mathbf{b}_i^\mathsf{T}$ and $\mathbf{x}^\mathsf{T}\mathsf{C}$ respectively. For the infima the analogous procedure holds. Also, we eliminated the epi- and hypographical variables $\overline{q}_i$ and $\underline{q}_i$, all $i \in [0\!:\!m]$. □

From the above theorem we can see that (PDA-UPP) does not add much difficulty over solving (UPP) with $\mathcal{M}(\mathbf{x}, s)$ removed.

8.2.2.3 Best-case performance

In Example 10 we see that there are choices on the uncertainty regime where the resulting robust solution exhibits good performance even in the best case. A decision maker may prefer uncertainty regimes where the conservatism inherent in the robust solution is redeemed by viable best-case performance of that solution. On the other hand, again applying this concept in terms of constraints, a decision maker may value the chance of saving resources in cases other then the worst case.

The model that can account for the best-case performance of the outcome of the uncertainty under a chosen uncertainty regime can be formulated as:

$$\begin{aligned}
&\inf_{\bar{\mathbf{u}},\mathbf{x},s,\underline{q}_i} \gamma_s + \alpha \sup_{\mathbf{u}_0\in\mathcal{U}_0} \left\{\mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x}\right\} + \beta\left(\underline{q}_0, \ldots, \underline{q}_m\right)\\
&\text{s.t.}: \underline{q}_0 \geq \mathbf{c}(\bar{\mathbf{u}})^\mathsf{T}\mathbf{x},\\
&\qquad \underline{q}_i \geq \mathbf{a}_i(\bar{\mathbf{u}})^\mathsf{T}\mathbf{x} - b_i(\bar{\mathbf{u}}), \qquad i \in [1\!:\!m],\\
&\qquad \bar{\mathbf{u}} \in \mathcal{U}^s,\; \mathbf{x} \in \mathcal{R}^s_{opt}, \qquad s \in [1\!:\!S].
\end{aligned} \tag{BCP-UPP}$$

The function $\beta(.)$ is assumed to be a convex, argument wise nondecreasing function and in our applications in later sections is chosen to be the sum of $\underline{q}_i$, all $i \in [0\!:\!m]$. Note that we cannot assume row-wise uncertainty, since not all rows can necessarily attain their minimum simultaneously. We would thus underestimate the optimal $\underline{q}_i$, all $i \in [0 : m]$, if we were to use the projection method, that is valid for the robust counterpart as discussed in the appendix. The problem (BCP-UPP) is a bilinear mixed integer optimization problem and as such very challenging. In case $\mathcal{X}_0 = \{0, 1\}^{n_2}$ one can again use McCormick linearization in order to obtain a problem that is manageable using C-MILP solvers. In case that all coefficients of $\mathbf{x}$ depend on independent portions of the uncertainty vector $\mathbf{u}$ one can pre-compute and replace the uncertain coefficients with their respective minima. In the next section we will discuss an application where both of these strategies are feasible.

Exact reformulation in case of column wise uncertainty

Analogous to the max-regret function, the constraints in (BCP-UPP) can be reformulated exactly in case of column wise uncertainty. For the reformulation we do not need a duality argument so $\mathcal{X}_0$ may be discrete. Also the structure of these "optimistic constraints" is simpler than the optimistic counterpart in the max-regret function, so that we can allow for uncertainty in the objective. Thus, compared to the case of max-regret, merely Assumption 1 b) is needed for the reformulation discussed shortly. Note, that under the assumption of column-wise uncertainty we can write the epigraphical constraints in (BCP-UPP) as

$$
\begin{aligned}
\underline{q}_0 &\geq \sum_{i=1}^{n} c_i(\bar{\mathbf{u}}_i)x_i, \\
\begin{pmatrix} \underline{q}_1 \\ \vdots \\ \underline{q}_m \end{pmatrix} &\geq \sum_{i=1}^{n} \mathbf{a}_i(\bar{\mathbf{u}}_i)x_i - \mathbf{b}(\bar{\mathbf{u}}_{n+1}).
\end{aligned}
$$

For the sake of notational simplicity we introduce

$$
\bar{\mathbf{a}}_i(\mathbf{u}_i) \coloneqq \begin{pmatrix} c_i(\bar{\mathbf{u}}_i) \\ \mathbf{a}_i(\bar{\mathbf{u}}_i) \end{pmatrix}, i \in [1\!:\!n] \quad \text{and} \quad \bar{\mathbf{b}}(\bar{\mathbf{u}}_{n+1}) \coloneqq \begin{pmatrix} 0 \\ \mathbf{b}(\bar{\mathbf{u}}_{n+1}) \end{pmatrix}.
$$

We can then find matrices and vectors $\bar{\mathbf{a}}_i^0,\ \bar{\mathsf{A}}_i,\ i \in [1\!:\!n]$ and of appropriate size so that $\bar{\mathbf{a}}_i(\bar{\mathbf{u}}_i) = \bar{\mathbf{a}}_i^0 + \bar{\mathsf{A}}_i\mathbf{u}_i,\ i \in [1\!:\!n]$. We are now ready to state the following theorem.

Theorem 49. *Assume that Assumption 1 b) holds, so that (BCP-UPP) can be written as*

$$
\begin{aligned}
&\inf_{\bar{\mathbf{u}},\mathbf{x},s,\underline{q}_i} \gamma_s + \alpha \sup_{\mathbf{u}_0 \in \mathcal{U}_0} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x} \right\} + \beta\left(\underline{q}_0, \dots, \underline{q}_m\right) \\
&\text{s.t.} : \underline{\mathbf{q}} \geq \sum_{i=1}^{n} \mathbf{a}_i(\bar{\mathbf{u}}_i)x_i - \mathbf{b}(\bar{\mathbf{u}}_{n+1}), \\
&\qquad \bar{\mathbf{u}} \in \mathcal{U}_i^s, \qquad i \in [1\!:\!n+1], \\
&\qquad \mathbf{x} \in \mathcal{R}_{opt}^s,\ \ s \in [1\!:\!S].
\end{aligned}
$$

Then, this problem is equivalent to

$$
\begin{aligned}
&\inf_{\bar{\mathbf{u}},\mathbf{x},s,\underline{q}_i} \gamma_s + \alpha \sup_{\mathbf{u}_0 \in \mathcal{U}_0} \left\{ \mathbf{c}(\mathbf{u}_0)^\mathsf{T}\mathbf{x} \right\} + \beta\left(\underline{q}_0, \dots, \underline{q}_m\right) \\
&\text{s.t.} : \underline{\mathbf{q}} \geq \sum_{i=1}^{n} \bar{\mathbf{a}}_i^0 x_i + \bar{\mathsf{A}}_i\mathbf{w}_i - \bar{\mathbf{b}}(\bar{\mathbf{u}}_{n+1}) \\
&\begin{pmatrix} x_i \\ \mathbf{w}_i \end{pmatrix} \in \mathcal{K}_i^s, \qquad\qquad i \in [1\!:\!n], \\
&\ \mathbf{u}_{n+1} \in \mathcal{U}^s,\ \ \mathbf{x} \in \mathcal{R}_{opt}^s,\ \ s \in [1\!:\!S].
\end{aligned}
$$

Proof. The proof proceeds as the first part of the proof of Theorem 47. □

8.3 Application to the uncertain shortest path problem with endogenous arc-weight uncertainty

In the following section, we will apply our framework to the shortest path problem, which we believe is a good illustration for our approach for two reasons

- It is a combinatorial problem that, courtesy of total modularity has an exact LP relaxation and thus enjoys strong duality. This fact transfers partly to the robust setting since for some uncertainty models, for example interval uncertainty, the robust objective function remains linear. In these cases we can apply Corollary 2. For more complicated uncertainty models we still can use Theorem 45 in conjunction with lazy constraints. Thus we can test both approaches in a single setting.

- The uncertain shortest path problem appears in pre-desaster network design which has been previously discussed in the context of robust optimization with endogenous uncertainty (see [62]). In this setting, one may be interested in solutions that give a performance guarantee, but not necessarily the best such guarantee if other aspects such predictability, regret and best-case performance can be enhanced. This is precisely the scope of our approach. Since $\mathbf{x} \in \mathcal{R}^s_{opt}$, a solution always comes with a performance guarantee, the robustly optimal value under the uncertainty regime s. But we optimize the uncertainty regime by taking other measures into account. We thus think the uncertain shortest path problem is a natural fit for our framework.

8.3.1 The uncertain shortest path problem and its robust counterpart

Given a directed, weighted Graph $G = (\mathcal{V}, \mathcal{A})$ where $\mathcal{V}$ is a set of vertices and $\mathcal{A}$ a set of arcs. We denote by c_i the arc weight of arc i and vertex 1 is marked as *source* and vertex n is marked as *sink*. The shortest path problem (SPP) is to find a path from the source to the sink such that the sum of the weights of the arcs crossed is minimized. Despite being a combinatorial optimization problem it is well known that the SPP has an exact LP-Formulation

$$\min_{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{A}|}} \left\{ \mathbf{c}^\mathsf{T}\mathbf{x} \colon \mathsf{E}\mathbf{x} = \mathbf{d} \right\},$$

where each (directed) arc corresponds to a variable x_k under some relabeling $\pi(k) = (i,j)$ and

$$e_{i,k} = \begin{cases} -1 & \text{if } \pi(k) = (i,j) \text{ for some } j \\ 1 & \text{if } \pi(k) = (j,i) \text{ for some } j \\ 0 & \text{otherwise} \end{cases}$$

Also $\mathbf{d} = (-1, 0, \ldots, 0, 1)^\mathsf{T} \in \mathbb{R}^{|\mathcal{V}|}$. That integrality can be dropped stems from the fact that E is totally unimodular and $\mathbf{d}$ is integral. The dual of the linearized SPP is given by

$$\max_{\mathbf{y} \in \mathbb{R}^{|\mathcal{V}|}} \left\{ \mathbf{d}^\mathsf{T}\mathbf{y} \colon \mathsf{E}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \right\}.$$

In case of endogenous uncertain arc weights we can formulate the endogenous uncertain SPP as

$$\min_{\mathbf{x} \in \{0,1\}^{|\mathcal{A}|}} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \colon \mathsf{E}\mathbf{x} = \mathbf{d} \right\}, \text{ where } \mathbf{u} \in \mathcal{U}^s, \tag{EU-SPP}$$

and the robust counterpart (for fixed s) is given by

$$\min_{\mathbf{x} \in \{0,1\}^{|\mathcal{A}|}} \left\{ \sup_{\mathbf{u} \in \mathcal{U}^s} \left\{ \mathbf{c}(\mathbf{u})^\mathsf{T}\mathbf{x} \right\} : \mathsf{E}\mathbf{x} = \mathbf{d} \right\}. \tag{R-SPP}$$

In (R-SPP) integrality can not be relaxed in general since the objective is no longer linear, but convex. It was proved in [9] that (R-SPP) under ellipsoidal uncertainty is in fact NP-hard, [89] showed that NP-hardness for finitely generated uncertainty sets. One can formulate it as a conic MILP but then we loose the duality needed for the characterization of $\mathcal{R}^s_{opt}$. In special cases, however, the robust counterpart remains an LP so that we can apply Corollary 2, in other cases we can still work with Theorem 45. In the remainder of this section we discuss various reformulations and approximations of (UPP) for the uncertain SPP under different cases of box uncertainty. All the formulations will involve some bilinear terms in s and $\mathbf{x}$, and since we can model these as bilinear terms between $\mathbf{r}$ and $\mathbf{x}$ and both of these are entry-wise between 0 and 1, the big-M constant for the construction of the McCormick envelope is given by 1. Also we will, in this section, push the dependence on s from the uncertainty set into the description of the uncertain objective coefficients, hence we consider $\mathbf{c}_s(\mathbf{u})$ while the uncertainty set will remain constant.

8.3.2 SPP under structured box uncertainty

Define $\mathbf{c}_s(\mathbf{u}) \coloneqq \mathbf{c}_s^0 + \mathsf{C}_s\mathbf{u}$ with $\mathbf{c}_s^0 \in \mathbb{R}_+^{|\mathcal{A}|}$ and $\mathsf{C}_s \in \mathbb{R}^{|\mathcal{A}|\times q}$, $s \in [1:S]$ and consider the following uncertainty set:

$$\mathcal{U}_{Box} \coloneqq \{\mathbf{u} \in \mathbb{R}^q : \|\mathbf{u}\|_\infty \leq 1\}.$$

There are at least two cases where the (R-SPP) under box uncertainty can be reformulated as an LP.

Interval uncertainty: $q = |\mathcal{A}|$ and C is diagonal:

$$\min_{\mathbf{x}\in\{0,1\}^{|\mathcal{A}|}} \left\{ \sup_{\mathbf{u}\in\mathcal{U}_{Box}} \left\{\mathbf{c}_s(\mathbf{u})^\mathsf{T}\mathbf{x}\right\} : \mathsf{E}\mathbf{x} = \mathbf{d} \right\} = \min_{\mathbf{x}\in\mathbb{R}_+^{|\mathcal{A}|}} \left\{\underline{\mathbf{c}}_s^\mathsf{T}\mathbf{x} : \mathsf{E}\mathbf{x} = \mathbf{d}\right\},$$

where $\underline{\mathbf{c}}$ is the vector of worst case realizations of each individual arc.

Correlated interval uncertainty: $\mathsf{C} \in \mathbb{R}_+^{|\mathcal{A}|\times q}$:

$$\min_{\mathbf{x}\in\{0,1\}^{|\mathcal{A}|}} \left\{ \sup_{\mathbf{u}\in\mathcal{U}_{Box}} \left\{\mathbf{c}_s(\mathbf{u})^\mathsf{T}\mathbf{x}\right\} : \mathsf{E}\mathbf{x} = \mathbf{d} \right\} = \min_{\mathbf{x}\in\mathbb{R}_+^{|\mathcal{A}|}} \left\{(\mathbf{c}_s^0)^\mathsf{T}\mathbf{x} + \|\mathsf{C}_s^\mathsf{T}\mathbf{x}\|_1 : \mathsf{E}\mathbf{x} = \mathbf{d}\right\},$$

where the equality holds since $\|\mathsf{C}_s^\mathsf{T}\mathbf{x}\|_1$ is linear in $\mathbf{x}$ over $\mathbb{R}_+^{|\mathcal{A}|}$, in fact $\|\mathsf{C}_s^\mathsf{T}\mathbf{x}\|_1 = \mathbf{e}^\mathsf{T}\mathsf{C}_s^\mathsf{T}\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}_+^{|\mathcal{A}|}$ and $\mathbf{e} \in \mathbb{R}^q$ is the all ones vector.

Since in both of these cases (R-SPP) can be exactly linearized, we can characterize the respective optimal set using strong linear duality. We can still manage general box-uncertainty by employing Theorem 45.

8.3.2.1 SPP with interval uncertainty

In case of endogenous interval uncertainty, the travel time on an individual arc j is known to lie in an interval $[\overline{c}_s^i, \underline{c}_s^i]$ if uncertainty regime s is active. Thus the worst and best case outcomes do not depend on $\mathbf{x}$ and are easily identified as the outcomes where the travel time is on the upper and lower bound respectively. Let $s \in [1:S]$ and define $\underline{\mathbf{c}}_s/\overline{\mathbf{c}}_s$ to be element-wise the worst/best-case realizations of $\mathbf{c}(\mathbf{u})$ under uncertainty regime s. We have an easily derived mixed integer linear model for each of the uncertainty preferences:

$$\begin{aligned}
&\min_{\mathbf{x},s} \gamma_s + \alpha\underline{\mathbf{c}}_s^\mathsf{T}\mathbf{x} + \gamma\mathcal{M}(\mathbf{x}, s)\\
&\text{s.t.}: \underline{\mathbf{c}}_s^\mathsf{T}\mathbf{x} \leq \mathbf{d}^\mathsf{T}\mathbf{y},\ \mathsf{E}\mathbf{x} = \mathbf{d},\ \mathsf{E}^\mathsf{T}\mathbf{y} \leq \underline{\mathbf{c}}_s,\\
&\qquad \mathbf{x} \in \mathbb{R}_+^{|\mathcal{A}|},\ s \in [1:S].
\end{aligned}$$

Here, we used Corollary 2 in order to characterize $\mathcal{R}^s_{opt}$. We can specify $\mathcal{M}(\mathbf{x}, s)$ in the following ways, as to accommodate the three different types of uncertainty preferences:

$$\begin{aligned}
\text{Predictability: } \mathcal{M}(\mathbf{x}, s) &= \left(\underline{\mathbf{c}}_s - \overline{\mathbf{c}}_s\right)^{\mathsf{T}} \mathbf{x}, \\
\text{Best-case Performance.: } \mathcal{M}(\mathbf{x}, s) &= \overline{\mathbf{c}}_s^{\mathsf{T}} \mathbf{x}, \\
\text{Regret: } \mathcal{M}(\mathbf{x}, s) &= \inf_{\mathbf{z}} \Big\{ \underline{\mathbf{c}}_s^{\mathsf{T}} \mathbf{x} - \mathbf{d}^{\mathsf{T}} \mathbf{z} : \\
&\qquad \mathsf{E}^{\mathsf{T}} \mathbf{z} \leq \overline{\mathbf{c}}_s + \sum_{i=1}^{|\mathcal{A}|} \mathbf{e}_i \left[\underline{\mathbf{c}}_s - \overline{\mathbf{c}}_s\right]_i x_i \Big\}.
\end{aligned}$$

The reformulation of the regret function was introduced in [52], where the concept of regret was dealt with under the name of maximum deviation robustness. The intuition is that, in case of interval uncertainty, the regret is maximized if the ex post solution may only experience the best-case realization on all arcs accept for those who are also crossed by the ex ante solution.

We see that the problem can be solved exactly using MILP-solvers without any further adaptation. Thus, there is at least one setting where our framework can be applied in its entirety, while never leaving the realm of basic and well developed optimization techniques.

8.3.2.2 SPP with correlated interval uncertainty

This type of uncertainty set allows us to model situations where the uncertainty across arcs is positively correlated which is irrelevant for the robust counterpart but impacts the optimistic counterpart es well as the max-regret function, because the optimistic optimizer present in both loses flexibility. In Example 11 we will demonstrate the difference between interval and correlated interval uncertainty.

For the uncertain travel times under uncertainty regime s given by $\mathbf{c}_s(\mathbf{u}) := \mathbf{c}^0_s + \mathsf{C}_s\mathbf{u}$, let C_s be non-negative. Then we have $\|\mathsf{C}_s^{\mathsf{T}}\mathbf{x}\|_1 = \mathbf{e}^{\mathsf{T}}\mathsf{C}_s^{\mathsf{T}}\mathbf{x}$ so that the objective in the robust counterpart has a linear objective and the LP relaxation is still exact. Thus we can again use LP-duality in order to characterize $\mathcal{R}^s_{opt}$ as follows:

$$\begin{aligned}
\min_{\mathbf{x}, s}\ & \gamma_s + \alpha \left((\mathbf{c}^0_s)^{\mathsf{T}} \mathbf{x} + \mathbf{e}^{\mathsf{T}} \mathsf{C}_s^{\mathsf{T}} \mathbf{x} \right) + \gamma \mathcal{M}(\mathbf{x}, s) \\
\text{s.t. : } & (\mathbf{c}^0_s)^{\mathsf{T}} \mathbf{x} + \mathbf{e}^{\mathsf{T}} \mathsf{C}_s^{\mathsf{T}} \mathbf{x} \leq \mathbf{d}^{\mathsf{T}} \mathbf{y}, \\
& \mathsf{E}\mathbf{x} = \mathbf{d}, \\
& \mathsf{E}^{\mathsf{T}} \mathbf{y} \leq \mathbf{c}^0_s + \mathsf{C}_s \mathbf{e}, \\
& \mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}_+,\ s \in [1:S].
\end{aligned}$$

Again, $\mathcal{M}(\mathbf{x}, s)$ can be specified in three different ways according to the respective uncertainty preferences:

$$\begin{aligned}\text{Predictability: } & \mathcal{M}(\mathbf{x}, s) = 2\mathbf{e}^\mathsf{T}\mathsf{C}_s^\mathsf{T}\mathbf{x},\\ \text{Best-case Performance.: } & \mathcal{M}(\mathbf{x}, s) = (\mathbf{c}_s^0)^\mathsf{T}\mathbf{x} - \mathbf{e}\mathsf{C}_s^\mathsf{T}\mathbf{x},\\ \text{Regret: } & \mathcal{M}(\mathbf{x}, s) = \gamma(\rho),\ \rho \geq \mathbf{b}^\mathsf{T}\mathbf{w},\ \mathsf{A}^\mathsf{T} \geq \mathbf{f}(\mathbf{x}, s),\end{aligned}$$

where $\mathbf{b}, \mathsf{A}$, and $\mathbf{f}$ are defined as described in Section 8.2.2.1. We further comment on the usefulness of this bound in section Section 8.3.3 and Section 6.3.

Example 11. *For the best and worst-case calculation, and hence also for predictability, there is no difference between interval uncertainty and correlated interval uncertainty from a modeling standpoint. To see this, let* $\mathsf{C} \in \mathbb{R}_+^{n\times m}$ *and consider*

$$\begin{aligned}&\sup / \inf \left\{\mathbf{x}^\mathsf{T}\mathsf{C}^\mathsf{T}\mathbf{u} \colon \|\mathbf{u}\|_\infty \leq 1\right\} = \pm\mathbf{x}^\mathsf{T}|\mathsf{C}^\mathsf{T}\mathbf{e}|\\ =&\sup / \inf \left\{\mathbf{x}^\mathsf{T}\mathrm{Diag}\left((|\mathsf{C}^\mathsf{T}\mathbf{e}|)_1, \dots, (|\mathsf{C}^\mathsf{T}\mathbf{e}|)_n\right)\mathbf{w} \colon \|\mathbf{w}\|_\infty \leq 1\right\}\end{aligned}$$

However, there is a difference when it comes to calculating the maximum regret. To understand this difference, consider the graph depicted in Figure 8.2 and assume the traveling time on all the edges was known to be between 1 and 2. There are two paths from node 1 to node 4. Under interval uncertainty, the maximum regret, regardless of the path chosen, the maximum regret would be $4 - 2 = 2$*, i.e. the worst case materializes on the path chosen, while the best case materializes for the alternative path. On the other hand, with correlated interval uncertainty we could model the situation with* $\mathsf{C} = 0.5\,\mathbf{e}$ *and* $\mathcal{U} = \{u \colon u \in [-1, 1]\}$ *and* $\mathbf{c}_0 = 1.5\,\mathbf{e}$*. In this case the same outcome would have to occur on all four edges simultaneously and the maximum regret would always be 0.*

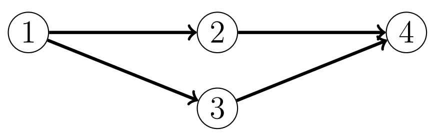


Figure 8.2

8.3.3 SPP with general box uncertainty

If we have box uncertainty where C does not obey the restrictions discussed above, the robust counterpart of the uncertainty does no longer have an exact linearization so that a characterization of $\mathcal{R}^s_{opt}$ based on Corollary 2 is no longer possible, so that Theorem 45 has to be invoked. With $\mathbf{c}_s(\mathbf{u})$ given as above, (UPP) for the SPP with general box uncertainty can be cast as:

$$\begin{aligned} \inf_{\mathbf{x},s}\ \gamma_s + \alpha\left((\mathbf{c}^0_s)^\mathsf{T}\mathbf{x} + \|\mathsf{C}^\mathsf{T}_s\mathbf{x}\|_1\right) + \gamma\mathcal{M}(\mathbf{x},s) & \\ \text{s.t.} : (\mathbf{c}^0_s)^\mathsf{T}\mathbf{x} + \|\mathsf{C}^\mathsf{T}_s\mathbf{x}\|_1 \le (\mathbf{c}^0_s)^\mathsf{T}\mathbf{z}_j + \|\mathsf{C}^\mathsf{T}_s\mathbf{z}_j\|_1,&\ j\in[1\!:\!S], \\ \mathsf{E}\mathbf{x} = \mathbf{d},& \\ \mathbf{x}\in\{0,1\}^{|\mathcal{A}|},\qquad\qquad & s\in[1\!:\!S], \end{aligned} \tag{8.16}$$

where $\mathbf{z}_j \in \mathcal{R}^s_{opt}$ all $s\in[1\!:\!S]$. The respective specifications for $\mathcal{M}(\mathbf{x},s)$ are

$$\begin{aligned} \text{Predictability: } & \mathcal{M}(\mathbf{x},s) = 2\|\mathsf{C}^\mathsf{T}_s\mathbf{x}\|_1, \\ \text{Best case performance.: } & \mathcal{M}(\mathbf{x},s) = \inf_{\mathbf{u}\in\mathcal{U}}\left\{(\mathbf{c}^0_s)^\mathsf{T}\mathbf{x} + \mathbf{u}^\mathsf{T}\mathsf{C}^\mathsf{T}_s\mathbf{x}\right\}, \\ \text{Regret: } & \mathcal{M}(\mathbf{x},s) = \mathbf{b}^\mathsf{T}\mathbf{w},\ \mathsf{A}^\mathsf{T} \ge \mathbf{f}(\mathbf{x},s). \end{aligned}$$

Since $\mathbf{x}$ and s are discrete, we can linearize the multilinear term in the infimum using McCormick envelops iteratively. For the max regret calculation we have to work again with the approximations described in Section 8.2.2.1. However, when we observe the max-regret function for uncertain SPP given by

$$\rho(\mathbf{x},s) = \sup_{\mathbf{u}\in\mathcal{U}}\left\{\mathbf{c}_s(\mathbf{u})^\mathsf{T}\mathbf{x} - \inf_{\mathbf{y}}\left\{\mathbf{c}_s(\mathbf{u})^\mathsf{T}\mathbf{y}\colon \mathbf{y}\in\{0,1\}^{|\mathcal{A}|}_+,\ \mathsf{E}\mathbf{y}=\mathbf{d}\right\}\right\},$$

we see, that we have a mixed integer binary problem where the bilinear terms can be exactly linearized using McCormick. This allows us at least to easily test the quality of our proposed lower upper bound. The same is true for the regret in the correlated intervall uncertainty model. We will report a respective evaluation in Section 6.3.

As stated above, the constraints enforcing $\mathbf{x}\in\mathcal{R}^s_{opt}$ are derived using Theorem 45. The number of these constraints may be exponential (in the sense described in the remark in the appendix) and each requires us to solve an optimization problem in order to determine the right hand side. We circumvent these problems in our implementation by applying these constraints as lazy constraints. Thus, in our implementation described in the next section, the constraints are absent at first, and whenever the branch and bound solver encounters a feasible solution $(\mathbf{x},s)$ it checks whether

$$(\mathbf{c}^0_s)^\mathsf{T}\mathbf{x} + \|\mathsf{C}^\mathsf{T}_s\mathbf{x}\| \le \inf_{\mathbf{y}}\left\{\sup_{\mathbf{u}\in\mathcal{U}}\left\{\mathbf{c}_s(\mathbf{u})^\mathsf{T}\mathbf{y}\right\} : \mathbf{y}\in\{0,1\}^{|\mathcal{A}|}_+,\ \mathsf{E}\mathbf{y}=\mathbf{d}\right\}.$$

If that is not the case, we add the constraint to the model. This process proved to be quite efficient since the robust optimization problem on the right hand side is a MILP that is quite easily tackled by modern solvers.

8.4 Numerical experiments

In this section we will conduct some numerical experiments, where we will apply the concepts introduced in this text to the shortest path problem. We will demonstrate that we can use general MILP solvers in order to solve (UPP) in this setting for interval uncertainty and general box uncertainty. Also we will investigate the trade-off between the different uncertainty preferences and worst-case performance. All instances were solved using Gurobi, where the implementation was done in Python. The experiments were run using an AMD Ryzen 7 1700X Eight-Core 3.4 GHz CPU, and 16 GB RAM.

For our experiments we considered the uncertain SPP under two types of endogenous box uncertainty: interval uncertainty and general box uncertainty. For both types we want to illustrate the trade-off between different contributors of the objective function of (UPP), namely worst-case performance, predictability, maximum regret and best-case performance. We compared the trade-off between two of these at a time so that in total 6 comparisons were made for both types of uncertainty models. The attributes not considered were skipped from the objective function, while the two that were considered had randomly chosen coefficients from the unit box. We probed 14 random coefficients per instance and generated 20 instances per pair of attributes so that in total we solved 280 problems per pair. The size of the instances was $|\mathcal{V}| = 300,\ S = 10$ for the interval uncertainty model and $|\mathcal{V}| = 50,\ S = 10,\ q = 10$ for the general uncertainty model. Arcs between nodes were established randomly, where an arc between nodes i and j was created if a uniformly distributed random number from the interval $[0, 1]$ was smaller than $0.3/|i - j|$. For the general box uncertainty we used the upper bound from Section 8.2.2.1 for the computation of regret. However, for each solution $(\mathbf{x}, s)$ we obtained, we also calculated $\rho(\mathbf{x}, s)$ exactly and compared that value to the value from the upper bound. In 1678 of 1680 cases the bound was in fact tight, indicating that it worked reasonably well for problems of the size considered here. In our experience the quality of the bound diminishes as $|\mathcal{V}|$ and $|\mathcal{A}|$ grow but is acceptable for medium size problems.

We summarize the results in Table 8.1 and Table 8.2. The tables report the average computation time in seconds (T) as well as the range of computation times ([T min, T max]). The column (#TO) gives am measure of how intensive the trade-off between different attributes has been in our experiments. It reports the number of instances for which the value of the compared attributes changed with different objective function coefficients. The columns (Δ_1%) and (Δ_2%) give the percentage difference between the maximum and the minimum value of the first and second attribute within an instance, averaged over all 20 instances. ΔOpt% is the analogous number for optimal values.

Further, Figure 8.3 and Figure 8.4 illustrate what can be thought of as combined, normalized efficiency frontiers. For every problem solved, we normalized the outcome of the attributes on a scale between 0 and 1 using the formula

$$a_{\%} = \frac{a - a_{min}}{a_{max} - a_{min}},$$

where a_{min} and a_{max} are the minimum and the maximum the attribute attained in that instance over all 14 probings, and a is the unnormalized value of that attribute in the current problem. After normalizing one can plot the 14 pairs of outcomes for the two attributes being compared in a scatter plot. What is depicted in the figures is the combination of all of these plots per pair. We only included those instances where there was a trade off between the two attributes.

We summarize the key takeaways from our experiments as follows:

- Even the most demanding model with general box uncertainty was manageable using standard software. Solutions were mostly obtained in a matter of minutes.
- Trade-offs between two attributes were encountered regularly indicating that different uncertainty preferences have a notable impact on the decision.
- In instances where there was a trade-off, outcomes of the attributes could vary substantially across the 14 probings.

preferences	T	$[T_{min}, T_{max}]$	$\#TO$	$\Delta_1\%$	Δ_2 %	$\Delta_{\text{opt}}\%$
predi vs. bestc	131.63	[26.16, 418.12]	20	61.81	16.81	90.57
predi vs. regret	18.48	[4.51, 72.24]	9	16.15	30.38	96.44
bestc vs. regret	115.77	[9.75, 328.43]	9	3.21	20.35	88.29
worstc vs. bestc	91.82	[3.57, 318.92]	20	14.38	13.03	85.88
worstc vs. regret	13.68	[5.63, 43.56]	8	1.64	18.04	94.41
worstc vs predi	5.04	[1.66, 16.14]	5	2.37	9.39	84.16

Table 8.1: Results for instances with general box uncertainty

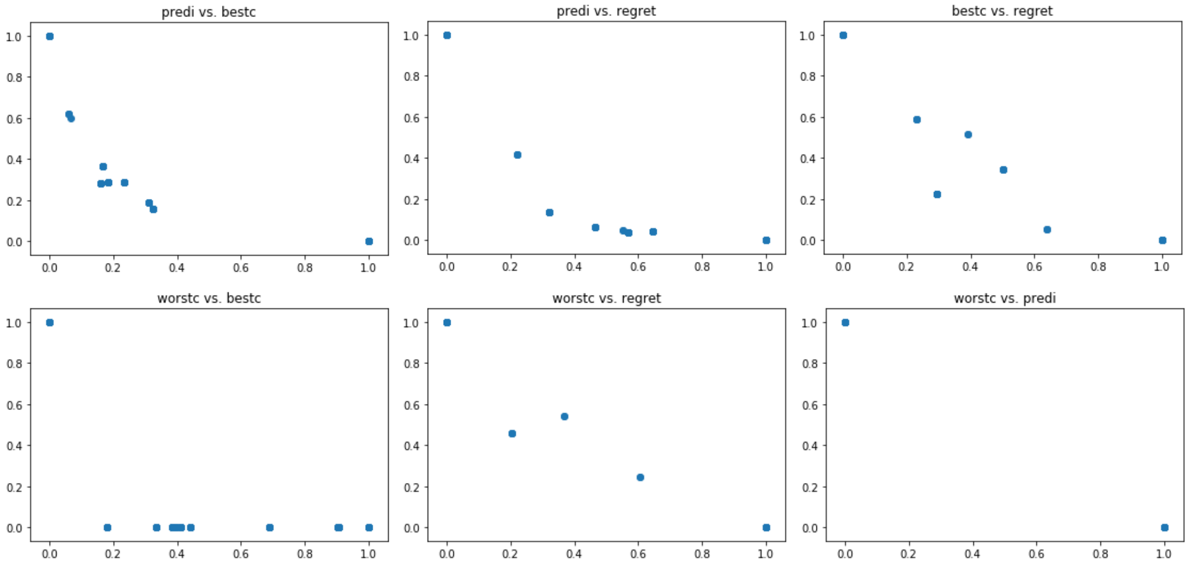


Figure 8.3

Scatter plots for instances with general box uncertainty.

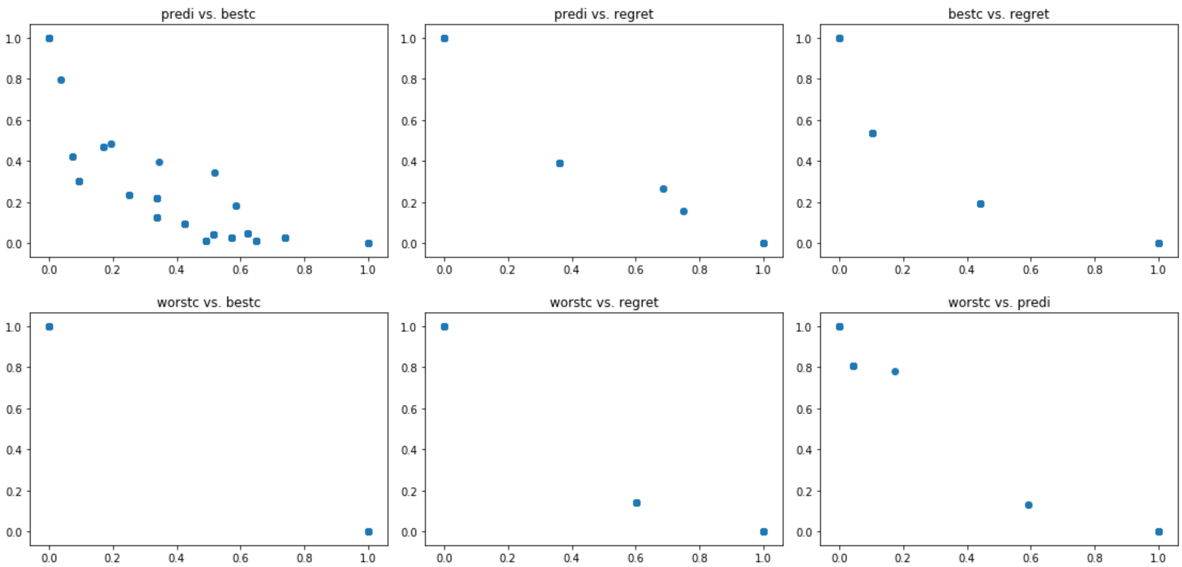


Figure 8.4

Scatter plots for instances with interval uncertainty.

Open questions

We understand that there are many open questions regarding this new model some of which we want to discuss briefly

- In our discussions on different choices for $\mathcal{M}(\mathbf{x}, s)$, we introduced convex func-

preferences	T	$[T_{min}, T_{max}]$	$\#TO$	$\Delta_1\%$	Δ_2 %	$\Delta_{\text{opt}}\%$
predi vs. bestc	14.89	[7.76, 28.7]	19	44.87	42.95	83.86
predi vs. regret	24.8	[7.48, 85.56]	15	21.9	23.91	86.62
bestc vs. regret	22.35	[9.24, 75.66]	9	7.25	19.73	85.16
worstc vs. bestc	14.0	[7.56, 47.97]	9	4.12	11.29	71.31
worstc vs. regret	20.88	[8.44, 36.87]	3	2.46	6.86	75.82
worstc vs predi	15.44	[6.8, 55.16]	12	7.12	15.78	81.15

Table 8.2: Results for instances with interval uncertainty

tions $\beta(.)$ and $\beta_i(.)$ but settled for very simple choices of those functions in our applications. Obviously other choices would be viable from a computational standpoint, but the question is whether there are meaningful choices that can be justified based on some knowledge of the preference of a decision maker.

- Some of our reformulations involve big-M constants for which we cannot hope to find a general and workable recipe. If one was to explore other applications of our framework, finding appropriate constant can pose an interesting challenge.
- In our numerical experiments, our proposed upper bound for the max regret proved to perform quite well. However, as stated above, the quality of the bound deteriorated with the size of the problem. It would be desirable to have stronger bounds and also bounds that are applicable to other problem classes.
- Some of the appeal of robust optimization stems from the fact, that many robust counterparts can be solved using standard optimization solvers. One of our goals was to show that this is at least partly true for our framework. However, in order to have broader range of possible applications and also to have better performance of the solution process, it would be worthwhile to investigate specialized algorithms that are able to solve (UPP) in specified settings.

Appendix A

Copostitive matrices and conic optimization

In this section we will review some basic concepts from convex analysis, conic optimization and set-copositive and set-completely positive matrix cones in so far they are relevant for the understanding of this thesis.

A.1 Convex cones and conic optimization problems

Let $\mathcal{V}$ be either $\mathbb{R}^n$ or $\mathcal{S}^n$, whereby we use the standard inner product in $\mathbb{R}^n$, while in $\mathcal{S}^n$ the inner product is given by the Frobenius product $\mathsf{A} \bullet \mathsf{B} = \mathrm{trace}(\mathsf{AB})$. For an arbitrary cone $\mathcal{K} \subseteq \mathcal{V}$ we denote the dual cone by $\mathcal{K}^* \subseteq \mathcal{V}$ where

$$\mathcal{K}^* \coloneqq \{\mathbf{y} \in \mathcal{V} \colon \langle \mathbf{x}, \mathbf{y} \rangle \geq 0 \ \text{ for all } \mathbf{x} \in \mathcal{K}\} = [\mathrm{cl}\, \mathcal{K}]^* \,.$$

It is well known that $\mathcal{K}^{**} = \mathrm{clconv}(\mathcal{K})$. The closure and the convex-hull operations can be omitted if $\mathcal{K}$ is closed and/or convex, respectively. We also have

$$\mathrm{int}(\mathcal{K}^*) = \{\mathbf{y} \in \mathcal{V} \colon \langle \mathbf{x}, \mathbf{y} \rangle > 0 \ \text{ for all } \mathbf{x} \in \mathcal{K} \setminus \{\mathbf{o}\}\} \,.$$

We say a convex cone $\mathcal{K}$ is *pointed* if it does not contain a line or, equivalently, if $\mathrm{int}(\mathcal{K}^*) \neq \emptyset$. For such a cone, any point $\mathbf{y} \in \mathcal{K}^*$ is a normal vector to a supporting hyperplane $\mathcal{H}_{\mathbf{y}} \coloneqq \{\mathbf{x} \in \mathcal{V} \colon \langle \mathbf{y}, \mathbf{x} \rangle = 0\}$ of $\mathrm{cl}(\mathcal{K})$, since supporting hyperplanes of convex cones are homogeneous. If $\mathbf{y} \in \mathrm{int}(\mathcal{K}^*)$ we have $\mathcal{H}_{\mathbf{y}} \cap \mathrm{cl}(\mathcal{K}) = \{\mathbf{o}\}$ since then, both sets do not have a common direction of recession. It follows that for $\mathbf{y} \in \mathrm{int}(\mathcal{K})$ and $\mathcal{H}_{\mathbf{y},\alpha} \coloneqq \{\mathbf{x} \in \mathcal{V} \colon \langle \mathbf{y}, \mathbf{x} \rangle = \alpha\}$ we have $\mathcal{H}_{\mathbf{y},\alpha} \cap \mathcal{K} = \emptyset$ if $\alpha < 0$ and $\mathcal{H}_{\alpha} \cap \mathcal{K} \neq \emptyset$ otherwise. Moreover, if $\mathrm{int}(\mathcal{K}) \neq \emptyset$ and $\alpha > 0$, then $\mathcal{H}_{\mathbf{y},\alpha} \cap \mathrm{int}(\mathcal{K}) \neq \emptyset$.

Throughout the text, we will make use of conic linear optimization tools. A conic linear optimization problem is given by

$$p^* = \inf_{\mathbf{x}\in\mathcal{K}} \left\{ \langle \mathbf{c}, \mathbf{x} \rangle \colon \langle \mathbf{a}_i, \mathbf{x} \rangle = b_i \, i \in [1\colon m] \right\} \,, \tag{P}$$

which is just a linear optimization problem with an extra constraint that restricts the decision variable $\mathbf{x}$ to lie in a closed, convex cone $\mathcal{K}$. The dual problem is given by

$$d^* = \sup_{\boldsymbol{\lambda}\in\mathbb{R}^m} \left\{ \sum_{i=1}^{m} -\lambda_i b_i \colon \mathbf{c} + \sum_{i=1}^{m} \lambda_i \mathbf{a}_i \in \mathcal{K}^* \right\} . \tag{D}$$

Slater's condition says that a relative interior feasible point of (P) guarantees $p^* = d^*$ and existence of an optimal solution to (D) if it is feasible, and likewise with (D) and (P) interchanged.

A.2 Set-completely positive and set-copositive matrix cones

A large portion of our discussion revolves around *set-copositive* and *set-completely positive* matrix cones, which for arbitrary cones $\mathcal{K} \subseteq \mathbb{R}^n$ are defined in the following way, putting $k = \binom{n+1}{2} = \dim \mathcal{S}^n$:

$$\begin{aligned}
\mathcal{COP}_n(\mathcal{K}) &\coloneqq \left\{ \mathsf{Q} \in \mathcal{S}^n \colon \mathbf{x}^{\mathsf{T}}\mathsf{Q}\mathbf{x} \geq 0 \;\text{ for all } \mathbf{x} \in \mathcal{K} \right\} \,, \\
\mathcal{CPP}_n(\mathcal{K}) &\coloneqq \left\{ \sum_{i=1}^{k} \mathbf{x}_i \mathbf{x}_i^{\mathsf{T}} \colon \mathbf{x}_i \in \mathcal{K}, \; i \in [1\colon k] \right\} \\
&= \operatorname{conv} \left\{ \mathbf{x}\mathbf{x}^{\mathsf{T}} \colon \mathbf{x} \in \mathcal{K} \right\} = \left\{ \mathsf{X}\mathsf{X}^{\mathsf{T}} \colon \mathsf{X} \in \mathbb{R}^{n\times k} \,, \, \mathbf{x}_i \in \mathcal{K}, \, i \in [1\colon k] \right\} \\
&\subseteq \mathcal{S}^n .
\end{aligned}$$

The index n will henceforth be suppressed if the dimension is clear from the context. Whenever we talk about copositive matrices and completely positive matrices we refere to matrices that are set-copostive or set-completely positive over $\mathcal{K} = \mathbb{R}^n_+$ respectively. The following proposition summarizes characterizations of some prominent examples of set-completely positive cones:

Proposition 50. *We have the following characterizations and approximations of set-completely positive cones*

$$\begin{aligned}
\mathcal{CPP}(\mathbb{R}^n) &= \mathcal{S}^n_+, \\
\mathcal{CPP}(\mathcal{SOC}^n) &= \left\{ \mathsf{X} \in \mathcal{S}^n_+ \colon \mathsf{X} \bullet \mathsf{J}_0 \leq 0 \right\} ,
\end{aligned}$$

$$\mathcal{CPP}(\mathbb{R}^n_+) \subseteq \mathcal{S}^n_+ \cap \mathcal{N}^n \ \ (\textit{equality if } n \leq 4)\,,$$

where $\mathcal{SOC}^n := \{(x_0, \mathbf{x}) \subseteq \mathbb{R}^n \colon \|\mathbf{x}\| \leq x_0\}$ *is the second order cone and* $\mathsf{J}_0 :=$ $\mathrm{Diag}(-1, 1, \ldots, 1)$.

Proof. The first equality is a standard characterization of $\mathcal{S}^n_+$. The second one is a consequence of the S-Lemma (see [85]) and the duality results to be discussed shortly. For the final result see [8]. □

Likewise we have for set-copositive cones we have the following characterizations:

Proposition 51. *We have the following characterizations and approximations of set-copositive cones*

$$\begin{aligned}
\mathcal{COP}(\mathbb{R}^n) &= \mathcal{S}^n_+,\\
\mathcal{COP}(\mathcal{S}\mathsf{O}\mathcal{C}) &= \{\mathsf{X} \in \mathcal{S}^n \colon \mathsf{X} - \tau\mathsf{J}_0 \succeq 0, \tau \geq 0\}\,,\\
\mathcal{CPP}(\mathbb{R}^n_+) &\supseteq \mathcal{S}^n_+ + \mathcal{N}^n \ \ (\textit{equality if } n \leq 4)\,,
\end{aligned}$$

where $\mathcal{SOC}^n := \{(x_0, \mathbf{x}) \subseteq \mathbb{R}^n \colon \|\mathbf{x}\| \leq x_0\}$ *is the second order cone and* $\mathsf{J}_0 :=$ $\mathrm{Diag}(-1, 1, \ldots, 1)$.

Proof. The proofs are due to the same references as in Proposition 50. □

The above matrix cones are convex cones that are dual to each other if $\mathcal{K}$ is closed by the following proposition:

Proposition 52. *We have the following duality correspondence*

$$\mathcal{COP}(\mathcal{K})^* = \mathrm{cl}\,[\mathcal{CPP}(\mathcal{K})] = \mathcal{CPP}(\mathrm{cl}\,\mathcal{K}).$$

Proof. See [80, Proposition 1 and Lemma 1]. □

In fact, an easy continuity argument shows that $\mathcal{COP}(\mathcal{K})$ is always closed even if $\mathcal{K}$ is not. Since $\mathcal{S}^n_+ = \mathcal{COP}(\mathbb{R}^n) = \mathcal{CPP}(\mathbb{R}^n)$, this also shows that the cone of positive-semidefinite matrices is self-dual, i.e. $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$.

The concept of copositivity was first introduced in [61] for the case $\mathcal{K} = \mathbb{R}^n_+$, however the standard reference for an exhaustive introduction is given by [8], another very thorough treatment is found in [36]. To determine whether a given matrix is an element of either $\mathcal{COP}(\mathbb{R}^n_+)$ or $\mathcal{CPP}(\mathbb{R}^n_+)$ is NP-hard. This leaves the impression that using these cones in optimization yields a complication rather than a solution of a problem. However, both cones can be approximated by tractable matrix cones, and

symbol	mode	method	ref.
$\mathcal{E}_d^n$	outer	LP	[20, 33]
$\mathcal{Y}_d^n$	outer	LP	[2]
$\mathcal{B}_d^n$	outer	LP	[24, 23]
$\mathcal{B}_d^n(\mathcal{M})$	outer	LP	[12]
$\mathcal{D}_d^n$	inner	LP	[24, 23]
$\mathcal{D}_d^n(\mathcal{M})$	inner	LP	[79]
$\mathcal{I}_d^n$	inner	LP	[64, 65]
$\mathcal{K}_d^n$	inner	SDP	[64, 65]
$\mathcal{Q}_d^n$	inner	SDP	[69]
$\mathcal{L}_d^n(\mu, T)$	outer	SDP	[57]
$\mathcal{L}_d^n(\mu, T; \mathcal{A})$	outer	SDP	[37]

Table A.1: Summary of approximation hierarchies for the cone $\mathcal{COP}^n$, from [12]

recent literature shows that even simple approximations yield good bounds for several kinds of optimization problems. We will not go into further detail but instead refer to [18, 40, 49]. Here, let us just mention the most elementary examples of tractable approximations given by

$$\begin{aligned} &\mathcal{COP}(\mathbb{R}_+^n) \supseteq \mathcal{S}^n + \mathcal{N}_n =: \mathcal{NND}_n, \\ &\mathcal{CPP}(\mathbb{R}_+^n) \subseteq \mathcal{S}_+^n \cap \mathcal{N}_n =: \mathcal{DNN}_n, \end{aligned} \tag{A.1}$$

where the mnemonics $\mathcal{NND}$, $\mathcal{DNN}$ abbreviate *non-negative-decomposable* and *doubly-non-negative* respectively. As mentioned above the inclusions in (A.1) hold with equality in case $n < 5$. More elaborate approximations are listed in Table A.1 along their respective reference.

We also have some useful inequalities

Proposition 53. *For any cones $\mathcal{K}, \mathcal{K}_1, \mathcal{K}_2 \subseteq \mathbb{R}^n$ we have the following equalities*

1. $\mathcal{COP}(\mathbb{R}^n \times \mathbb{R}_+) = \mathcal{CPP}(\mathbb{R}^n \times \mathbb{R}_+) = \mathcal{S}_+^{n+1}$,
2. $\mathcal{CPP}(\mathrm{conv}\mathcal{K}) \supseteq \mathcal{CPP}(\mathcal{K})$ *with equality iff* $\mathcal{K}$ *is convex,*
3. $\mathcal{CPP}(\mathcal{K}_1 \cup \mathcal{K}_2) = \mathcal{CPP}(\mathcal{K}_1) + \mathcal{CPP}(\mathcal{K}_2)$,
4. $\mathrm{int}\mathcal{CPP}(\mathcal{K}) = \left\{\sum_{i=1}^k \mathbf{x}_i\mathbf{x}_i^\mathsf{T} : \mathbf{x}_i \in \mathrm{int}\mathcal{K}, k \in \mathbb{N}\right\}$ *if* $\mathcal{K}$ *is convex and* $\mathrm{int}\mathcal{K} \neq \emptyset$
5. $\mathcal{COP}(\mathcal{K}) = \mathcal{COP}(-\mathcal{K}) = \mathcal{COP}(\mathcal{K} \cup -\mathcal{K})$,
6. $\mathcal{COP}(\mathcal{K}) = \mathrm{cl}\mathcal{COP}(\mathcal{K}) = \mathcal{COP}(\mathrm{cl}\mathcal{K}) = \mathcal{COP}(\mathrm{ri}\mathcal{K})$,
7. $\mathcal{COP}(\mathrm{conv}\mathcal{K}) \subseteq \mathcal{COP}(\mathcal{K})$ *with equality iff* $\mathrm{cl}\mathcal{K}$ *is convex,*

8. $\mathcal{COP}(\mathcal{K}_1 \cup \mathcal{K}_2) = \mathcal{COP}(\mathcal{K}_1) \cap \mathcal{COP}(\mathcal{K}_2)$,

9. $\operatorname{int}\mathcal{COP}(\mathcal{K}) = \left\{\mathsf{Q} \colon \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} > 0 \ \forall \mathbf{x} \in \mathcal{K} \setminus \{0\}\right\}$

Proof. 1. holds since $\mathcal{COP}(\mathbb{R}^n \times \mathbb{R}_+) \supseteq \mathcal{S}_+^{n+1}$ while on the other hand $(\mathbf{x}^\mathsf{T}, x_0)\mathsf{A}(\mathbf{x}; x_0) \geq 0 \ \forall (\mathbf{x}, x_0) \in \mathbb{R}^n \times \mathbb{R}_+$ implies $(\mathbf{x}^\mathsf{T}, x_0)\mathsf{A}(\mathbf{x}; x_0) = (-1)^2(-\mathbf{x}^\mathsf{T}, -x_0)\mathsf{A}(-\mathbf{x}; -x_0) \geq 0 \ \forall (\mathbf{x}, x_0) \in \mathbb{R}^n \times \mathbb{R}_-$. The equality $\mathcal{CPP}(\mathbb{R}^n \times \mathbb{R}_+) = \mathcal{S}_+^{n+1}$ then follows from selfduality of $\mathcal{S}_+^{n+1}$ and the fact that $\mathbb{R}^n \times \mathbb{R}_+$ is a closed, convex cone and hence $\mathcal{COP}(\mathbb{R}^n \times \mathbb{R}_+) = \mathcal{CPP}(\mathbb{R}^n \times \mathbb{R}_+)^*$.

The containment in 2. is trivial as is the equality if $\mathcal{K}$ is convex. Thus, assume that the equality holds but there is an $\mathbf{x} \in \operatorname{conv}\mathcal{K} \setminus \mathcal{K}$, then $\mathbf{x}\mathbf{x}^\mathsf{T} \in \mathcal{CPP}(\operatorname{conv}\mathcal{K})$ but $\mathbf{x}\mathbf{x}^\mathsf{T} \notin \mathcal{CPP}(\mathcal{K})$ since it is a rank one matrix that cannot be generated as a convex combination of other rank ones matrices.

3. is immediate from the definition of $\mathcal{CPP}(\mathcal{K})$. 4. is derived in [35] for the case $\mathcal{K} = \mathbb{R}^n_+$, but the argument can be followed through in case $\mathcal{K}$ is merely convex and has interior. 5. follows from $\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} = (-\mathbf{x})\mathsf{A}(-\mathbf{x})$. 6. follows from the continuity of quadratic functions. 7. follows from 2. and conic duality. 8. is immediate from the definitions of $\mathcal{COP}(\mathcal{K})$.

9. follows from

$$
\begin{aligned}
\operatorname{int}\mathcal{COP}(\mathcal{K}) &= \{\mathsf{Q} \colon \mathsf{Q} \bullet \mathsf{X} > 0 \ \forall \mathsf{X} \in \mathcal{CPP}(\mathcal{K}) \setminus \{\mathsf{O}\}\} \\
&= \{\mathsf{Q} \colon \mathsf{Q} \bullet \mathsf{X} > 0 \ \forall \mathsf{X} \in \mathcal{CPP}(\mathcal{K}) \setminus \{\mathsf{O}\}\} \\
&= \{\mathsf{Q} \colon \mathsf{Q} \bullet \mathsf{X} > 0 \ \forall \mathsf{X} \in \operatorname{ext}\mathcal{CPP}(\mathcal{K}) \setminus \{\mathsf{O}\}\} \\
&= \left\{\mathsf{Q} \colon \mathbf{x}^\mathsf{T}\mathsf{Q}\mathbf{x} > 0 \ \forall \mathbf{x} \in \mathcal{K} \setminus \{0\}\right\}.
\end{aligned}
$$

□

Finally we proof that any dyadic matrix $\mathbf{x}\mathbf{x}^\mathsf{T}$ recruited from $\mathbf{x} \in \mathcal{K}$ is a generator of an extreme ray of $\mathcal{CPP}\left(\mathcal{K}\right)$. The statement is well known for the case $\mathcal{K} \in \left\{\mathbb{R}^n_+, \mathbb{R}^n\right\}$, but a direct proof was hard to find in literature so we might as well proof it for the general case here.

Lemma 54. *The set of extreme rays of* $\mathcal{CPP}(\mathcal{K})$*, denoted by* $\operatorname{ext}\mathcal{CPP}(\mathcal{K})$*, is the union of all the sets* $\mathcal{R}(\mathbf{x}\mathbf{x}^\mathsf{T}) \coloneqq \left\{\lambda\mathbf{x}\mathbf{x}^\mathsf{T} \colon \lambda \geq 0\right\}$ *where* $\mathbf{x} \in \mathcal{K} \subseteq \mathbb{R}^n$.

Proof. Since $\mathcal{CPP}_n(\mathcal{K}) = \operatorname{clconv}\left\{\mathbf{x}\mathbf{x}^\mathsf{T} \colon \mathbf{x} \in \mathcal{K}\right\}$ we see that $\mathsf{X} \in \operatorname{ext}\mathcal{CPP}(\mathcal{K})$ implies that $\mathsf{X} = \mathbf{x}\mathbf{x}^\mathsf{T}$ for some $\mathbf{x} \in \mathcal{K}$.

For the converse, take any $\mathbf{z}\mathbf{z}^\mathsf{T}$ with $\mathbf{z} \in \mathcal{K} \subseteq \mathbb{R}^n$. We have to show that for any $\mathbf{x}\mathbf{x}^\mathsf{T}, \mathbf{y}\mathbf{y}^\mathsf{T} \in \mathcal{CPP}(\mathcal{K})$ such that $\mathbf{z}\mathbf{z}^\mathsf{T} = \mathbf{x}\mathbf{x}^\mathsf{T} + \mathbf{y}\mathbf{y}^\mathsf{T}$ we have $\mathbf{x}\mathbf{x}^\mathsf{T} = \alpha\mathbf{z}\mathbf{z}^\mathsf{T}$ and

$\mathbf{y}\mathbf{y}^\mathsf{T} = \beta\mathbf{z}\mathbf{z}^\mathsf{T}$ for some real numbers $\alpha, \beta \geq 0$.

After a relabelling we can assume that $\mathbf{z} = \begin{pmatrix} \bar{\mathbf{z}} \\ \mathbf{z}_0 \end{pmatrix}$ with $\mathbf{z}_0 = \mathbf{o}$ and $\bar{z}_i \neq 0$. Consequently

$$\begin{pmatrix} \bar{\mathbf{z}}\bar{\mathbf{z}}^\mathsf{T} & \bar{\mathbf{z}}\mathbf{z}_0^\mathsf{T} \\ \mathbf{z}_0\bar{\mathbf{z}}^\mathsf{T} & \mathbf{z}_0\mathbf{z}_0^\mathsf{T} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}}\bar{\mathbf{x}}^\mathsf{T} & \bar{\mathbf{x}}\mathbf{x}_0^\mathsf{T} \\ \mathbf{x}_0\bar{\mathbf{x}}^\mathsf{T} & \mathbf{x}_0\mathbf{x}_0^\mathsf{T} \end{pmatrix} + \begin{pmatrix} \bar{\mathbf{y}}\bar{\mathbf{y}}^\mathsf{T} & \bar{\mathbf{y}}\mathbf{y}_0^\mathsf{T} \\ \mathbf{y}_0\bar{\mathbf{y}}^\mathsf{T} & \mathbf{y}_0\mathbf{y}_0^\mathsf{T} \end{pmatrix}$$

implies that $\mathbf{z}_0\mathbf{z}_0^\mathsf{T} = \mathbf{x}_0\mathbf{x}_0^\mathsf{T} + \mathbf{y}_0\mathbf{y}_0^\mathsf{T}$ and since $\mathcal{S}_+^n$ is a pointed cone we have $\mathbf{x}_0 = \mathbf{y}_0 = 0$. We thus may assume that $\mathbf{z}$ has no zero entries. So assume $z_i \neq 0$ for $i \in [1:n]$. $\mathbf{z}\mathbf{z}^\mathsf{T} = \mathbf{x}\mathbf{x}^\mathsf{T} + \mathbf{y}\mathbf{y}^\mathsf{T} \Leftrightarrow z_i^2 = x_i^2 + y_i^2$, and $z_iz_j = x_ix_j + y_iy_j, i \neq j$ also, x_i or y_i are not zero so w.l.o.g. $y_p \neq 0$ for some $p \in [1\!:\!n]$. This implies $z_i = \pm\sqrt{x_i^2 + y_i^2}$, so that

$$\pm\sqrt{x_i^2 + y_i^2}\sqrt{x_j^2 + y_j^2} = x_ix_j + y_iy_j \Leftrightarrow \left\|\begin{matrix} x_i \\ y_i \end{matrix}\right\| \left\|\begin{matrix} x_j \\ y_j \end{matrix}\right\| = \left|\begin{pmatrix} x_i \\ y_i \end{pmatrix}^\mathsf{T} \begin{pmatrix} x_j \\ y_j \end{pmatrix}\right|$$

which implies

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \lambda_{i,j}\begin{pmatrix} x_j \\ y_j \end{pmatrix} \text{ or } \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \mathbf{o},\ k \in \{i, j\},$$

by the equality case of the Cauchy-Schwarz inequality and the latter is impossible. Note that $\lambda_{p,i} \neq 0, i \in [1:n]$ since otherwise $y_p = 0$ contrary to our assumption. Given $\lambda_{p,i} \neq 0$, $\begin{pmatrix} x_p \\ y_p \end{pmatrix} = \lambda_{p,i}\begin{pmatrix} x_i \\ y_i \end{pmatrix}$ $i \in [1\!:\!n]$ and $y_p \neq 0$ we can define $\mu_i := \frac{\lambda_{p,i}x_i}{y_p}$ and we see that

$$\begin{pmatrix} x_p \\ x_i \end{pmatrix} = \begin{pmatrix} \lambda_{p,i}x_i \\ x_i \end{pmatrix} = \frac{\lambda_{p,i}x_i}{y_p}\begin{pmatrix} y_p \\ y_p/\lambda_{p,i} \end{pmatrix} = \mu_i\begin{pmatrix} y_p \\ y_i \end{pmatrix}. \tag{A.2}$$

From this we see that $\mu_i = \frac{x_p}{y_p} =: \mu$ so that $\mathbf{x} = \mu\mathbf{y}$. And thus

$$\mathbf{z}\mathbf{z}^\mathsf{T} = \mathbf{y}\mathbf{y}^\mathsf{T} + \mu^2\mathbf{y}\mathbf{y}^\mathsf{T} = (1 + \mu^2)\mathbf{y}\mathbf{y}^\mathsf{T},$$

so that we have $\mathbf{y}\mathbf{y}^\mathsf{T} = \frac{1}{1+\mu^2}\mathbf{z}\mathbf{z}^\mathsf{T}$ and $\mathbf{x}\mathbf{x}^\mathsf{T} = \frac{\mu^2}{1+\mu^2}\mathbf{z}\mathbf{z}^\mathsf{T}$ with $\frac{\mu^2}{1+\mu^2} \geq 0$ and $\frac{1}{1+\mu^2} > 0$. ☐

A.3 Matrix completion

A major interest in linear algebra is the question when a matrix whose entries are only partially determined can be completed to a matrix with a specific property. In this chapter we will briefly review the most important results on the existence of positive semidefinite, completely positive and doubly nonnegative matrix completions. These results involve some graph theoretical concepts which we will review briefly. These concepts are only relevant for the theorems at the end of this section and are not needed anywhere else in the text.

A graph $G = (\mathcal{V}, \mathcal{E})$ is given by its set of vertices $\mathcal{V} = \{v_1, \ldots, v_n\}$ and its set of edges $\mathcal{E} \subseteq \{\{v, u\} : v, u \in \mathcal{V}\}$, both of which are finite. A subgraph $T = (\mathcal{V}_T, \mathcal{E}_T)$ of a graph G is a graph such that $\mathcal{V}_T \subseteq \mathcal{V}$ and $\mathcal{E}_T \subseteq \mathcal{E}$. Vertex v_j is *adjacent* to v_j and vice versa if $\{v_i, v_j\} \in \mathcal{E}$. If $e = \{v_i, v_j\} \in \mathcal{E}$ then v_i and v_j are *incident* on e. A graph where all vertices are adjacent to one another is called a *complete* graph. A *path* that connects vertex v_i with v_j is given by a sequence of edges so that $\{v_1.v_{k_1}\}, \ldots, \{v_{k_p}, v_j\}$, v_{k_i} are distinct and $p > 1$ is the *length* of that path. A graph is *connected* if any two vertices have a connecting path. A graph that is not connected is *disconnected*. A *cycle* is path that connects a vertex v to itself. A *chord* of a cycle with length greater then 3 is an edge that connects two vertexes who are incident on two different edges of the cycle. A graph is *chordal* if every cycle with length greater 3 has a chord. A subgraph of G that is complete is called a *clique*. A block B of a graph is a subgraph that is connected, has no disconnected subgraph that is obtained by removing just one vertex and its adjacent edges (i.e. a *cut vertex*) from B, and is not contained in any other subgraph with these two properties. A *block-clique graph* is a graph whose blocks are cliques.

A *partial matrix* of order n is a matrix whose entries in the i-th row and the j-th column are determined if and only if $(i, j) \in \mathcal{I} \subseteq [1\!:\!n]^2$ and are undetermined otherwise. A matrix is said to be *partial positive semidefinite/ completely positive/ doubly nonnegative* if and only if every fully determined principal submatrix is positive semidefinite/ completely positive/ doubly nonnegative. A partial matrix is *positive semidefinite/ completely positive/ doubly nonnegative completable* if we can specify the undetermined entries so that the fully specified matrix is semidefinite/ completely positive/ doubly nonnegative.

The *specification graph* of partial matrix A of order n is a graph $G(\mathsf{A})$ with vertices $\mathcal{V} = \{v_i \colon i \in [1\!:\!n]\}$ and edges $\mathcal{E}$ such that $\{v_i, v_j\} \in \mathcal{E}$ if and only if the entry a_{ij} is specified. A symmetric matrix with $G(\mathsf{A}) = G$ is called a *symmetric matrix realization* of G.

The following three theorems give the key results on matrix completion as far as this text is concerned:

Theorem 55. *All partial positive semidefinite symmetric matrix realizations of a graph G are positive semidefinite completable if and only if G is chordal.*

Proof. See [8, Theorem 1.39]. □

Theorem 56. *Every partial completely positive matrix realization of a graph G is completely positive completable if and only if G is a block-clique graph*

Proof. See [8, Theorem 2.33]. □

Theorem 57. *Every partial doubly nonnegative matrix realization of a graph G is doubly nonnegative completable if and only if G is a block-clique graph*

Proof. See [39]. □

Appendix B

Notation

Matrices

A	matriecs are denoted in San-serif capital letters
$(\mathsf{A})_i/\mathbf{a}_i$	the i-th row or column (as inferred by the context) of a matrix A
I_n	the identity matrix of dimension n *
E_n	the all ones matrix of dimension n *
O_n	the identity matrix of dimension n *
$\mathsf{A} \bullet \mathsf{B}$	The Frobenious product of two matrices given by $\operatorname{trace}(\mathsf{A}^\mathsf{T}\mathsf{B}) = \sum_{i=1}^n \sum_{i=1}^m a_{ij}b_{ij}$
$\operatorname{diag}(\mathsf{A})$	the vector of diagonal entries of the matrix A
$\mathsf{M}(\mathsf{A},\mathsf{B},\mathsf{C})$	A composed matrix $\begin{pmatrix} \mathsf{A} & \frac{1}{2}\mathsf{B}^\mathsf{T} \\ \frac{1}{2}\mathsf{B} & \mathsf{C} \end{pmatrix}$
$\mathsf{Y}(\mathbf{x},\mathsf{X})$	A composed matrix $\begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathsf{X} \end{pmatrix}$

Vectors

$\mathbf{a}$	vectors are denoted by small boldface letters
$(\mathbf{a})_i/a_i$	the i-th entry of the vector $\mathbf{a}$
$\mathbf{e}_i$	the i-th membe of the canonical basis

$\mathbf{e}$	the all ones vector, the dimension will always be clear from the context
$\mathbf{o}_n$	the all zeros vector of dimension n *
$\mathbf{a}^\mathsf{T}$	the transpose of a vector $\mathbf{a}$
$\mathrm{Diag}(\mathbf{a})$	a diagonal matrix whose diagonal entries are the entries of the vector $\mathbf{a}$
$\|\mathbf{a}\|_p$	the p-norm of the vector $\mathbf{a}$ given by $\sqrt[\frac{1}{p}]{\sum_{i=1}^n \|a_i\|^p}$

Sets

$[i:j]$	the set of natural numbers from i to j
$\mathbb{N}^n$	the space of vectors of natural numbers of dimension n
$\mathbb{Z}^n$	the space of vectors of integers of dimension n
$\mathbb{R}^n$	the real vector-space of dimension n
$\mathcal{SOC}^n$	the second order cone of dimension n, i.e. $\left\{\left(x_0, \mathbf{x}^\mathsf{T}\right)^\mathsf{T} \in \mathbb{R}^n \colon \|\mathbf{x}\|_2 \leq x_0\right\}$ *
$\mathbb{R}^{n\times m}$	the space of $n \times m$ matrices
$\mathcal{S}^n$	the vector space of symmetric matrices of order n
$\mathcal{S}^n_+/\mathcal{S}^n_{++}$	the cone of symmetric positive semidefinite/positive definite matrices
$\mathcal{S}^n_-/\mathcal{S}^n_{--}$	the cone of symmetric negative semidefinite/negative definite matrices
$\mathcal{N}^n$	the cone of nonnegative symmetric matrices of order n
$\mathcal{CPP}^n\left(\mathcal{K}\right)$	the cone of set-completely positive matrices over a cone $\mathcal{K} \subseteq \mathbb{R}^n$ *
$\mathcal{COP}^n\left(\mathcal{K}\right)$	the cone of set-copositive matrices over a cone $\mathcal{K} \subseteq \mathbb{R}^n$ *
$\mathcal{DNN}^n$	the cone of double nonnegative matrices of order n, i.e. $\mathcal{S}^n_+ \cap \mathcal{N}^n$ *
$\mathcal{NND}^n$	the cone of nonnegative decomposable matrices of order n, i.e. $\mathcal{S}^n_+ + \mathcal{N}^n$ *
$\mathrm{aff}\left(\mathcal{A}\right)$	the affine hull of a set $\mathcal{A}$
$\mathrm{conv}\left(\mathcal{A}\right)$	the convex hull of a set $\mathcal{A}$
$\mathrm{span}\left(\mathcal{A}\right)$	the span of a set $\mathcal{A}$
$\mathrm{int}\left(\mathcal{A}\right)$	interior of the set $\mathcal{A}$

$\mathrm{cl}\,(\mathcal{A})$	closure of the set $\mathcal{A}$
$\mathrm{clconv}\,(\mathcal{A})$	the closure of the convex hull of a set $\mathcal{A}$
$\mathrm{ri}\,(\mathcal{A})$	relative interior of a set $\mathcal{A}$
$\mathrm{bd}\,(\mathcal{A})$	boundary of a set $\mathcal{A}$
$\mathrm{cone}\,(\mathcal{A})$	the conic hull of a set $\mathcal{A}$, we use the definition $\mathrm{cone}\,(\mathcal{A}) \coloneqq \{\lambda\mathbf{x} \colon \mathbf{x} \in \mathcal{A},\ \lambda \geq 0\}$
$\mathrm{ext}\,(\mathcal{A})$	the set of generators of extreme points and rays of a convex set $\mathcal{A}$

* we suppress the indication of the dimension if it is clear from the context

References

[1] Farid Alizadeh and Donald Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.

[2] E Alper Yıldırım. On the accuracy of uniform polyhedral approximations of the copositive cone. *Optimization methods and software*, 27(1):155–173, 2012.

[3] Kurt M Anstreicher and Samuel Burer. Computable representations for convex hulls of low-dimensional quadratic forms. *Mathematical programming*, 124(1-2):33–43, 2010.

[4] Riley Badenbroek and Etienne de Klerk. An analytic center cutting plane method to determine complete positivity of a matrix. *arXiv preprint arXiv:2006.05319*, 2020.

[5] Amir Beck and Aharon Ben-Tal. Duality in robust optimization: primal worst equals dual best. *Operations Research Letters*, 37(1):1–6, 2009.

[6] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, Princeton and Oxford, 2009.

[7] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[8] Abraham Berman and Naomi Shaked-Monderer. *Completely positive matrices*. World Scientific, 2003.

[9] Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization under ellipsoidal uncertainty sets. *Manuscript, MIT*, 9, 2004.

[10] Daniel Bienstock. A note on polynomial solvability of the cdt problem. *SIAM Journal on Optimization*, 26(1):488–498, 2016.

[11] Immanuel Bomze, Mirjam Dür, Etienne De Klerk, Cornelis Roos, Arie Quist, and Tamás Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 24(2):163–185, 2002.

[12] Immanuel Bomze, Mirjam Dür, and Chung-Piaw Teo. Copositive optimization. *Optima - Newsletter of the M(athmetical)O(ptimization)S(ociety)*, 89(1):2–8, 2012.

[13] Immanuel Bomze and Markus Gabl. Uncertainty preferences in robust linear optimization with endogenous uncertainty. Under Revision at SIAM Journal on Optimization, avaiable at https://arxiv.org/abs/2011.14875.

[14] Immanuel Bomze and Markus Gabl. Interplay of non-convex quadratically constrained problems with adjustable robust optimization. *Mathematical Methods of Operations Research*, pages 1–37, 2020.

[15] Immanuel Bomze, Markus Gabl, Francesca Maggioni, and Georg Pflug. Two-stage stochastic standard quadratic optimization. Submitted to European Journal of Operations Research.

[16] Immanuel Bomze, Marcello Pelillo, and Robert Giacomini. Evolutionary approach to the maximum clique problem: Empirical evidence on a larger scale. In *Developments in global optimization*, pages 95–108. Springer, 1997.

[17] Immanuel M Bomze. On standard quadratic optimization problems. *Journal of Global Optimization*, 13(4):369–387, 1998.

[18] Immanuel M Bomze. Copositive optimization–recent developments and applications. *European Journal of Operational Research*, 216(3):509–520, 2012.

[19] Immanuel M Bomze. Copositive relaxation beats Lagrangian dual bounds in quadratically and linearly constrained quadratic optimization problems. *SIAM Journal on Optimization*, 25(3):1249–1275, 2015.

[20] Immanuel M Bomze and Etienne De Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24(2):163–185, 2002.

[21] Immanuel M. Bomze, Michael Kahr, and Markus Leitner. Trust your data or not – StQP remains StQP: Community detection via robust Standard Quadratic Optimization. *Math. Oper. Res.*, online, 2020.

[22] Samuel Rota Bulò and Immanuel M. Bomze. Infection and immunization: A new class of evolutionary game dynamics. *Games and Economic Behavior*, 71(1):193 – 211, 2011. Special Issue In Honor of John Nash.

[23] Stefan Bundfuss and Mirjam Dür. Algorithmic copositivity detection by simplicial partition. *Linear Algebra and its Applications*, 428(7):1511–1523, 2008.

[24] Stefan Bundfuss and Mirjam Dür. An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization*, 20(1):30–53, 2009.

[25] Samuel Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120(2):479–495, 2009.

[26] Samuel Burer. Copositive programming. In *Handbook on semidefinite, conic and polynomial optimization*, pages 201–218. Springer, 2012.

[27] Samuel Burer and Kurt M Anstreicher. Second-order-cone constraints for extended trust-region subproblems. *SIAM Journal on Optimization*, 23(1):432–451, 2013.

[28] Samuel Burer and Hongbo Dong. Representing quadratically constrained quadratic programs as generalized copositive programs. *Operations Research Letters*, 40(3):203–206, 2012.

[29] Samuel Burer and Boshi Yang. The trust region subproblem with nonintersecting linear constraints. *Mathematical Programming*, 149(1-2):253–264, 2015.

[30] Maria Rosa Celis. A trust region strategy for nonlinear equality constrained optimizati on. In *Proceedings of numerical optimization*, pages 71–82. SIAM, 1985.

[31] NH Chieu, V Jeyakumar, and G Li. Classes of nonnegative quadratic optimization problems with exact copositive relaxation. *arXiv preprint arXiv:1707.09486*, 2017.

[32] Andrew R Conn, Nicholas IM Gould, and Philippe L Toint. *Trust region methods.* SIAM, 2000.

[33] Etienne De Klerk and Dmitrii V Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[34] Erick Delage and Dan A Iancu. Robust multistage decision making. In *The Operations Research Revolution*, pages 20–46. INFORMS, 2015.

[35] Peter Dickinson. An improved characterisation of the interior of the completely positive cone. *The Electronic Journal of Linear Algebra*, 20, 2010.

[36] Peter JC Dickinson. *The copositive cone, the completely positive cone and their generalisations*. Citeseer, 2013.

[37] Peter JC Dickinson and Janez Povh. Moment approximations for set-semidefinite polynomials. *Journal of optimization theory and applications*, 159(1):57–68, 2013.

[38] Lloyd L Dines. On the mapping of quadratic forms. *Bulletin of the American Mathematical Society*, 47(6):494–498, 1941.

[39] John H Drew and Charles R Johnson. The completely positive and doubly nonnegative completion problems. *Linear and Multilinear Algebra*, 44(1):85–92, 1998.

[40] Mirjam Dür. Copositive programming–a survey. In Moritz Diehl, Francois Glineur, Elias Jarlebring, and Wim Michiels, editors, *Recent advances in optimization and its applications in engineering*, pages 3–20. Springer, Berlin Heidelberg, 2010.

[41] Gabriele Eichfelder and Janez Povh. On the set-semidefinite representation of nonconvex quadratic programs over arbitrary feasible sets. *Optimization Letters*, 7(6):1373–1386, 2013.

[42] Alain Faye and Frédéric Roupin. Partial lagrangian relaxation for general quadratic programming. *4OR*, 5(1):75–88, 2007.

[43] Minyue Fu, Zhi-Quan Luo, and Yinyu Ye. Approximation algorithms for quadratic programming. *Journal of combinatorial optimization*, 2(1):29–50, 1998.

[44] Mituhiro Fukuda, Masakazu Kojima, Kazuo Murota, and Kazuhide Nakata. Exploiting sparsity in semidefinite programming via matrix completion I: General framework. *SIAM Journal on Optimization*, 11(3):647–674, 2001.

[45] Markus Gabl. On reduced, conic reformulations of qcqps and set-completely positive matrix completion. Work in progress.

[46] Can Gokalp, Areesh Mittal, and Grani A Hanasusanto. Robust convex quadratically constrained quadratic programming with mixed-integer uncertainty. *arXiv preprint arXiv:1706.01949*, 2017.

[47] Bram L Gorissen, Ihsan Yanikoglu, and Dick den Hertog. A practical guide to robust optimization. *Omega*, 53:124–137, 2015.

[48] Pierre Hansen, Brigitte Jaumard, and Gilles Savard. New branch-and-bound rules for linear bilevel programming. *SIAM Journal on scientific and Statistical Computing*, 13(5):1194–1217, 1992.

[49] Jean-Baptiste Hiriart-Urruty and Alberto Seeger. A variational approach to copositive matrices. *SIAM review*, 52(4):593–629, 2010.

[50] Vaithilingam Jeyakumar and Guoyin Li. Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization. *Mathematical Programming*, 147(1-2):171–206, 2014.

[51] Peter Kall, Stein W Wallace, and Peter Kall. *Stochastic programming.* Springer, 1994.

[52] OE Karaşan, MÇ Pinar, and H Yaman. The robust shortest path problem with interval data. *Computers & Operations Research (to appear)*, 2001.

[53] Sunyoung Kim, Masakazu Kojima, and Kim-Chuan Toh. Doubly nonnegative relaxations are equivalent to completely positive reformulations of quadratic optimization problems with block-clique graph structures. *Journal of Global Optimization*, pages 1–29, 2020.

[54] Sunyoung Kim, Masakazu Kojima, and Kim-Chuan Toh. A geometrical analysis on convex conic reformulations of quadratic and polynomial optimization problems. *SIAM Journal on Optimization*, 30(2):1251–1273, 2020.

[55] Panos Kouvelis and Gang Yu. *Robust discrete optimization and its applications (nonconvex optimization and its applications).* Kluwer Academic Publishers Dordrecht, Germany, 1997.

[56] Nikolaos H Lappas and Chrysanthos E Gounaris. Robust optimization for decision-making under endogenous uncertainty. *Computers & Chemical Engineering*, 111:252–266, 2018.

[57] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.

[58] Francesca Maggioni and Georg Ch. Pflug. Bounds and approximations for multi-stage stochastic programs. *SIAM Journal on Optimization*, 26(1):831–855, 2016.

[59] Areesh Mittal, Can Gokalp, and Grani A Hanasusanto. Robust quadratic programming with mixed-integer uncertainty. *INFORMS Journal on Computing*, 2019.

[60] Theodore S Motzkin and Ernst G Straus. Maxima for graphs and a new proof of a theorem of turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[61] TS Motzkin. Copositive quadratic forms. *National Bureau of Standards Report*, 1952:11–22, 1818.

[62] Omid Nohadani and Kartikey Sharma. Optimization under decision-dependent uncertainty. *SIAM Journal on Optimization*, 28(2):1773–1795, 2018.

[63] Panos M Pardalos and Stephen A Vavasis. Quadratic programming with one negative eigenvalue is np-hard. *Journal of Global optimization*, 1(1):15–22, 1991.

[64] Pablo A Parrilo. Semidefinite programming based tests for matrix copositivity. In *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No. 00CH37187)*, volume 5, pages 4624–4629. IEEE, 2000.

[65] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[66] Gábor Pataki. On positive duality gaps in semidefinite programming. Available as arXiv preprint arXiv:1812.11796 (2019/October/30), 2018.

[67] Gábor Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res*, 23(2):177–203, 1998.

[68] Massimiliano Pavan and Marcello Pelillo. Dominant sets and pairwise clustering. *IEEE transactions on pattern analysis and machine intelligence*, 29(1):167–172, 2006.

[69] Javier Pena, Juan Vera, and Luis F Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105, 2007.

[70] Imre Pólik and Tamás Terlaky. A survey of the S-Lemma. *SIAM review*, 49(3):371–418, 2007.

[71] BT Polyak. Convexity of quadratic transformations and its use in control and optimization. *Journal of Optimization Theory and Applications*, 99(3):553–583, 1998.

[72] Franz Rendl and Henry Wolkowicz. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Mathematical Programming*, 77(1):273–299, 1997.

[73] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.

[74] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

[75] Hanif D Sherali and Warren P Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, 1990.

[76] Naum Z Shor. Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences*, 25(6):1–11, 1987.

[77] AL Soyster and FH Murphy. A unifying framework for duality and modeling in robust linear programs. *Omega*, 41(6):984–997, 2013.

[78] Allen L Soyster and Frederic H Murphy. On the integration of row and column uncertainty in robust linear programming. *Journal of Global Optimization*, 66(2):195–223, 2016.

[79] Julia Sponsel, Stefan Bundfuss, and Mirjam Dür. An improved algorithm to test copositivity. *Journal of Global Optimization*, 52(3):537–551, 2012.

[80] Jos F Sturm and Shuzhong Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28(2):246–267, 2003.

[81] Levent Tunçel. On the slater condition for the sdp relaxations of nonconvex sets. *Operations Research Letters*, 29(4):181–186, 2001.

[82] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations research*, 62(6):1358–1376, 2014.

[83] Guanglin Xu and Samuel Burer. A copositive approach for two-stage adjustable robust optimization with uncertain right-hand sides. *Computational Optimization and Applications*, 70(1):33–59, 2018.

[84] Guanglin Xu and Grani A Hanasusanto. Improved decision rule approximations for multi-stage robust optimization via copositive programming. *Optimization Online*, 2018.

[85] Vladimir A Yakubovich. S-procedure in nonlinear control theory. *Vestnik Leningrad University*, 1:62–77, 1971.

[86] Boshi Yang, Kurt Anstreicher, and Samuel Burer. Quadratic programs with hollows. *Mathematical Programming*, 170(2):541–553, 2016.

[87] Boshi Yang and Samuel Burer. A two-variable analysis of the two-trust-region subproblem. *Preprint, submitted to SIAM J. Optim., http://www. optimization-online. org/DB_HTML/2013/07/4126. html*, 2013.

[88] Ihsan Yanikoglu, B Gorissen, and Dick den Hertog. A survey of adjustable robust optimization. *Available online at ResearchGate*, 2017.

[89] Gang Yu and Jian Yang. On the robust shortest path problem. *Computers & operations research*, 25(6):457–468, 1998.

[90] Jianzhe Zhen, Ahmadreza Marandi, Dick den Hertog, and Lieven Vandenberghe. Disjoint bilinear programming: An adjustable robust optimization perspective. Available at http://www.optimization-online.org/DB_HTML/2018/06/6685.html (2019/October/30).

[91] Luis F Zuluaga, Juan Vera, and Javier Peña. LMI approximations for cones of positive semidefinite forms. *SIAM Journal on Optimization*, 16(4):1076–1091, 2006.