Positivstellensatz Certificates for Containment of Polyhedra and Spectrahedra

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Informatik und Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von

Kai Kellner

aus Frankfurt am Main

Frankfurt 2015

(D 30)

vom Fachbereich Informatik und Mathematik der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan:
Prof. Dr. Uwe Brinkschulte

Gutachter:
Prof. Dr. Thorsten Theobald
Prof. Dr. Raman Sanyal

Datum der Disputation: 23.06.2015

Deutsche Zusammenfassung

Enthaltenseinsprobleme gehören zu den klassischen Problemen der (konvexen) Geometrie. Im eigentlichen Sinne ist ein Enthaltenseinsproblem die Frage nach der mengentheoretischen Inklusion zweier gegebener Mengen. Diese Aufgabe ist im Allgemeinen sowohl vom theoretischen als auch vom praktischen Standpunkt aus schwer. Im weiteren Sinne gehören auch Radien- [GK92] und Packungsprobleme [BCS99] zu dieser Klasse von Problemen. Für einige Klassen konvexer Mengen wurden Enthaltenseinsprobleme intensiv untersucht. Dazu gehören unter anderem Enthaltenseinsprobleme für Polyeder und Kugeln [FO85] sowie Enthaltenseinsprobleme für Polyeder [GK94], die vor allem gegen Ende des 20. Jahrhunderts wegen ihrer inhärenten Bedeutung in der linearen Optimierung und der Kombinatorik untersucht wurden. Seitdem war der Fortschritt bei der Untersuchung von Enthaltenseinsproblemen dieser Art eher gering. In den letzten Jahren ist großes Interesse an Enthaltenseinsproblemen von Spektraedern, die eine natürliche Verallgemeinerung von Polyedern bilden, aufgekommen [BTN02, HKM13]. Dieses Interesse ist unter anderem in der intrinsischen Bedeutung von Spektraedern und deren Projektionen in der polynomiellen Optimierung [Ble12, GPT10] und der konvexen algebraischen Geometrie [HV07] begründet. Die bisherigen Untersuchungen beschränken sich jedoch auf spezielle Klassen oder Enthaltenseinssituationen, demgegenüber hat eine übergreifende Betrachtung derartiger Probleme noch nicht stattgefunden.

In der vorliegenden Arbeit betrachten wir das Enthaltenseinsproblem für Polyeder, Spektraeder und deren Projektionen aus dem Blickwinkel semialgebraischer Probleme und studieren algebraische Zertifikate für Enthaltensein. Dies führt zu einer neuen und systematischen Herangehensweise an Enthaltenseinsprobleme von (Projektionen von) Polyedern und Spektraedern, und liefert neue und teilweise unerwartete Resultate.

Wir wollen zunächst den Hauptansatz dieser Arbeit skizzieren. Dieser Ansatz ist im Bereich der polynomiellen Optimierung mittlerweile üblich, aber in Bezug auf geometrische Probleme kleinen Grades noch immer wenig verstanden. Wir verstehen lineare Optimierung als eine Anwendung des Lemmas von Farkas, das die (Un-)Lösbarkeit eines Systems linearer Ungleichungen beschreibt. Die affine Version von Farkas' Lemma charakterisiert lineare Polynome, die nichtnegativ auf einem gegebenen Polyeder sind. Lassen wir die Linearitätsbedingung weg, so erhalten wir eine Nichtnegativitätsbedingung auf einer semialgebraischen Menge. Dies führt zu sogenannten Positivstellensätzen (oder genauer Nichtnegativstellensätzen). Ein Positivstellensatz liefert ein *Zertifikat* in Form einer polynomiellen Identität für die Positivität eines Polynoms auf einer semialgebraischen Menge. Wie im linearen Fall sind Positivstellensätze die Grundlage der polynomiellen Optimierung und von Relaxierungsmethoden. Der Übergang von Positivität zu Nichtnegativität ist eine der größten Herausforderungen im Bereich der reellen algebraischen Geometrie und der polynomiellen Optimierung.

Im Bezug auf Enthaltenseinsprobleme ergeben sich daraus mehrere Hauptfragen: Kann das jeweilige Enthaltenseinsproblem als ein polynomielles Nichtnegativitätsproblem (oder Zulässigkeitsproblem) formuliert werden? Falls ja, welchen Zusammenhang zwischen Positivität und Nichtnegativität auf der einen Seite und Enthaltensein auf der anderen Seite gibt es, insbesondere im Hinblick auf deren geometrische Bedeutung? Gibt es einen geeigneten Positivstellensatz für das jeweilige Problem? Den Grad der semialgebraischen Zertifikate betreffend, welcher Grad ist notwendig, welcher hinreichend, um Enthaltensein zu zertifizieren?

Tatsächlich können (nahezu) alle Enthaltenseinsprobleme, die in dieser Arbeit untersucht werden, als ein polynomielles Nichtnegativitätsproblem formuliert werden, was die Anwendung von Positivstellensätzen erlaubt. Im Gegensatz zu diesem allgemeinen Resultat, hängt die Beantwortung der anderen Fragen stark vom jeweiligen Enthaltenseinsproblem, insbesondere im Hinblick auf die zugrunde liegende Geometrie des Problems, ab. Ein wichtiger Punkt ist, ob die Hierarchien, die durch Erhöhen der Grade in den polynomiellen Relaxierungen entstehen, Enthaltensein immer nach endlich vielen Schritten zertifizieren. Wir verstehen *endliche Konvergenz* einer Hierarchie in diesem Sinne (im Gegensatz zu asympotischer Konvergenz).

In den nachfolgenden Abschnitten erläutern wir die Hauptprobleme dieser Arbeit und stellen unseren Beitrag zu deren Verständnis aus dem Blickwinkel der semialgebraischen Geometrie dar. Wir konzentrieren uns auf die folgenden Enthaltenseinsprobleme. Das Enthaltenseinsproblem für Polytope stellt die Frage, ob ein $\mathcal{H}$-Polytop in einem $\mathcal{V}$-Polytop enthalten ist. Die Frage ob ein Spektraeder in einem Spektraeder enthalten ist, bezeichnen wir als das Enthaltenseinsproblem für Spektraeder. Zudem behandeln wir Enthaltenseinsprobleme für Projektionen von Polyedern und Spektraedern. Die Auswahl begründet sich unter anderem auf der wenig untersuchten Geometrie dieser Probleme und deren komplexitätstheoretischer Klassifizierung als schwere Probleme. Für algorithmische Fragestellungen nehmen wir immer an, dass die Eingabedaten rationale Zahlen sind.

Das Enthaltenseinsproblem für Polytope. Ein Polytop kann sowohl als konvexe Hülle endlich vieler Punkte ("$\mathcal{V}$-Polytop") als auch als Durchschnitt endlich vieler Halbräume ("$\mathcal{H}$-Polytop") dargestellt werden. Für Polytope wurden die algorithmische Geometrie und Komplexität von Enthaltenseinsproblemen eingehend untersucht. Ein prominentes Problem in der algorithmischen Polytop-Theorie ist das Folgende.

> Enthaltenseinsproblem für Polytope:
> **Eingabe:** $d \in \mathbb{N}$, ein $\mathcal{H}$-Polytop $P \subseteq \mathbb{R}^d$ und ein $\mathcal{V}$-Polytop $Q \subseteq \mathbb{R}^d$.
> **Aufgabe:** Entscheide, ob $P \subseteq Q$.

Gemäß einem Resultat von Freund und Orlin [FO85] ist das Enthaltenseinsproblem für Polytope co-NP-vollständig (während die Beantwortung der umgekehrten Frage trivial ist). Wenn die Dimension von zumindest einem Polytop fest gewählt ist, dann kann das Enthaltenseinsproblem für Polytope in polynomieller Zeit gelöst werden (Theorem 3.1.3). Trotz seiner grundlegenden Natur ist das Enthaltenseinsproblem für Polytope noch nicht umfassend untersucht worden.

Wir formulieren das Problem als ein bilineares Zulässigkeitsproblem mit zwei disjunkten Familien linearer Nebenbedingungen. Äquivalent dazu können wir das Problem als ein Maximierungsproblem über dem Produkt zweier $\mathcal{H}$-Polytope auffassen (Proposition 4.1.1). Die Formulierung als bilineares Problem erlaubt die Anwendung linearer und semidefiniter Relaxierungen. Während im Fall von *starkem Enthaltensein*, das heißt $P \subseteq Q$ und $\partial P \cap \partial Q = \emptyset$, die Konvergenz der Hierarchien beziehungsweise die Existenz von Zertifikaten aus den Positivstellensätzen von Handelman [Han88] und im anderen Fall von Putinar [Put93] folgen, ist der Fall nicht-starken Enthaltenseins nicht durch die allgemeine Theorie beantwortet.

Handelmans Positivstellensatz (Proposition 2.4.1) liefert Zertifikate für die Positivität von Polynomen auf Polytopen. Die Zertifikate können mithilfe einer Hierarchie linearer Programme bestimmt werden. Neben einem grundlegenden Konvergenzresultat (Theorem 4.2.1) diskutieren wir die weitgehend offene Frage von Gradschranken (Theorem 4.2.5).

Putinars Positivstellensatz (Proposition 2.4.3) liefert Zertifikate für die Positivität von Polynomen auf semialgebraischen Mengen. Die Zertifikate können mithilfe einer Hierarchie semidefiniter Programme bestimmt werden. Als ein Hauptresultat beweisen wir, dass unter schwachen (geometrischen) Voraussetzungen die semidefinite Relaxierung in endlich vielen Schritten ein Zertifikat für Enthaltensein liefert (Theoreme 4.3.1 und 4.3.6). Der Beweis dieser Aussage basiert auf einem hinreichenden Kriterium von Marshall [Mar08].

Das Enthaltenseinsproblem für Spektraeder. Ein Spektraeder $\mathcal{S}$ kann sowohl als Schnitt des Kegels positiv semidefiniter Matrizen mit einem affinen Unterraum als auch als Positivitätsregion eines linearen (Matrix-)Büschels dargestellt werden. Ein lineares Büschel ist dabei eine symmetrische Matrix mit linearen Polynomen als Einträgen. Von letzterer Darstellung lässt sich leicht erkennen, dass jedes Polyeder (in $\mathcal{H}$-Darstellung) ein Spektraeder ist. Dazu schreibt man die definierenden linearen Polynome des $\mathcal{H}$-Polyeders auf die Diagonale. Wir definieren das folgende Problem.

> Enthaltenseinsproblem für Spektraeder:
> **Eingabe:** $d \in \mathbb{N}$, Spektraeder $S_A \subseteq \mathbb{R}^d$ und $S_B \subseteq \mathbb{R}^d$.
> **Aufgabe:** Entscheide, ob $S_A \subseteq S_B$.

Da Polyeder eine Spezialklasse von Spektraedern bilden, umfasst das Enthaltenseinsproblem für Spektraeder insbesondere die Fälle $\mathcal{H}$-in-$\mathcal{H}$, $\mathcal{H}$-in-$\mathcal{S}$ und $\mathcal{S}$-in-$\mathcal{H}$. Während das erste Problem in polynomieller Zeit entscheidbar ist, ist das zweite Problem co-NP-schwer. Somit ist das allgemeine Enthaltenseinsproblem für Spektraeder co-NP-schwer. Der Komplexitätsstatus des dritten Problems ist noch nicht vollständig klassifiziert (vgl. Kapitel 3 und [KTT13]).

Ben-Tal und Nemirovski [BTN02] haben das sogenannte Matrixwürfel-Problem untersucht. Dies entspricht dem Enthaltenseinsproblem für Spektraeder unter der Einschränkung, dass das innere Spektraeder ein Würfel ist. Für das Matrixwürfel-Problem haben Ben-Tal und Nemirovski ein hinreichendes semidefinites Kriterium angegeben. Helton, Klep und McCullough [HKM12, HKM13] haben das Enthaltenseinsproblem für sogenannte freie Spektraeder untersucht. Ein freies Spektraeder lebt in der Vereinigung von Vektorräumen unterschiedlicher Dimension. Aus ihren Untersuchungen konnten Helton et al. ein hinreichendes Kriterium für das Enthaltenseinsproblem für Spektraeder ableiten. Ist das innere Spektraeder ein Würfel (in Standarddarstellung), dann stimmt das hinreichende Kriterium mit jenem von Ben-Tal und Nemirovski überein.

Unser Ausgangspunkt für die Untersuchung des Enthaltenseinsproblem für Spektraeder ist das Enthaltenseinsproblem für $\mathcal{H}$-Polyeder, das als lineares Zulässigkeitsproblem (LFP) formuliert werden kann (Theorem 5.1.1) und daher in polynomieller Zeit lösbar ist. Aus diesem LFP leiten wir ein hinreichendes semidefinites Kriterium zur Entscheidung des Enthaltenseinsproblems für Spektraeder ab. Da das Kriterium im Allgemeinen nicht notwendig für Enthaltensein ist – das gilt bereits in einem geometrisch einfachen, 2-dimensionalem Beispiel –, kommen in natürlicher Weise zwei Hauptfragen auf. Erstens unter welchen zusätzlichen Voraussetzungen das Kriterium notwendig ist und zweitens ob ein besseres Kriterium existiert, in dem Sinne dass es sowohl hinreichend als auch notwendig ist.

Um die zweite Fragestellung anzugehen, formulieren wir das Enthaltenseinsproblem für Spektraeder als ein polynomielles Zulässigkeitsproblem (beziehungsweise als ein quantifiziertes polynomielles Optimierungsproblem). Darauf basierend geben wir eine Hierarchie semidefiniter Zulässigkeitsprobleme an, deren positive Lösung ein hinreichendes Kriterium für Enthaltensein ist. Da für lineare Büschel kein vollwertiges Analogon des Lemmas von Farkas existiert, ergibt sich die Existenz eines Zertifikates im Falle *starken Enthaltenseins*, das heißt $S_A \subseteq \operatorname{int} S_B$, nicht unmittelbar aus der allgemeinen Theorie. Unter einer schwachen Zusatzbedingung an das Büschel des äußeren Spektraeders und der Voraussetzung der Beschränktheit an das innere Spektraeder, liefert der Positivstellensatz von Hol und Scherer die Existenz eines Zertifikates für starkes Enthaltensein (Theorem 5.1.8). Daraus lässt sich die asymptotische Konvergenz der Optimierungsversion ableiten. Endliche Konvergenz oder die Existenz eines Zertifikates im Fall von nicht-starkem Enthaltensein ist ein offenes Problem. Wie im skalaren Fall stellt der Übergang von Positivität zu Nichtnegativität eine der größten Herausforderungen dar.

Aufgrund der Komplexität hoher Relaxierungsstufen und im Hinblick auf die zuvor genannten Probleme, fokussieren wir unsere Untersuchungen auf den *initialen Relaxierungsschritt* der Hierarchie. Wir stellen fest, dass der initiale Relaxierungsschritt mit dem hinreichenden semidefiniten Kriterium übereinstimmt, welches wir zuvor vom $\mathcal{H}$-in-$\mathcal{H}$ Problem abgeleitet haben (Theorem 5.1.11). Die Untersuchung des initialen Relaxierungsschrittes erlaubt uns die Angabe von Zertifikaten für Enthaltensein in einigen wichtigen Fällen und Beispielen.

Im Spezialfall $\mathcal{S}$-in-$\mathcal{H}$ des Enthaltenseinsproblems für Spektraeder genügt unter schwachen – in der semidefiniten Optimierung üblichen – Voraussetzungen die Betrachtung des initialen Relaxierungsschrittes, um das Enthaltenseinsproblem zu entscheiden (Theorem 5.2.3). Die Aussage gilt auch in dem Fall, dass die Koeffizienten des äußeren Büschels simultan diagonalisierbar sind. Der Beweis beruht auf dem geometrischen Verhalten des initialen Relaxierungsschrittes in Kombination mit der Dualitätstheorie semidefiniter Optimierung. Die Aussage ist insofern bemerkenswert, als das Kriterium im Allgemeinen von der Büschel-Darstellung der Spektraeder abhängt.

Im Fall von Spektratopen, das heißt beschränkten Spektraedern, existiert immer ein positiver Skalierungsfaktor, so dass nach der Skalierung eines der beiden Spektratope Enthaltensein in der initialen Relaxierung zertifiziert wird (Theorem 5.3.1).

Das Enthaltenseinsproblem für Spektraeder kann auch vom Standpunkt positiver linearer Abbildungen aus betrachtet werden. Diesen Ansatz haben zuerst Helton et al. [HKM13] – in einem allgemeineren Kontext – gewählt. Eine wohlbekannte Relaxierung der Positivität ist die vollständige Positivität einer linearen Abbildung. Man kann zeigen, dass vollständige Positivität äquivalent zur initialen Relaxierungsstufe ist. Unsere Hierarchie für das Enthaltenseinsproblem für Spektraeder ist daher auch eine Hierarchie zwischen Positivität und vollständiger Positivität. Andererseits erlaubt die Theorie positiver linearer Abbildungen interessante Einblicke in das Enthaltenseinsproblem für Spektraeder. Im Besonderen zeigen wir, dass die initiale Relaxierungsstufe nicht nur hinreichend, sondern auch notwendig für eine bestimmte Familie 2-dimensionaler Spektratope ist (Theorem 5.4.10).

Die Enthaltenseinsprobleme für Projektionen. Die Projektion eines Polyeders ist wieder ein Polyeder, demgegenüber ist die Projektion eines Spektraeders im Allgemeinen kein Spektraeder. Insbesondere ist die Projektion eines Spektraeders nicht notwendigerweise abgeschlossen. Für eine Menge $S \subseteq \mathbb{R}^{d+m}$ bezeichnen wir mit $\pi(S) \subseteq \mathbb{R}^d$ die Projektion von S. Da projizierte Polyeder eine Unterklasse von projizierten Spektraedern bilden, fassen wir die genannten Enthaltenseinsprobleme wie folgt zusammen.

Enthaltenseinsprobleme für Projektionen:
Eingabe: $d, m, n \in \mathbb{N}$, Spektraeder $S_A \subseteq \mathbb{R}^{d+m}$ und $S_B \subseteq \mathbb{R}^{d+n}$.
Aufgabe: Entscheide, ob $\pi(S_A) \subseteq \pi(S_B)$.

Der Komplexitätsstatus von Enthaltenseinsproblemen ändert sich mit dem Übergang vom nicht-projektiven auf den projektiven Fall nicht wesentlich. Während jedoch das Enthaltenseinsproblem zweier $\mathcal{H}$-Polyeder in polynomieller Zeit entscheidbar ist, ist das Enthaltenseinsproblem zweier projizierter $\mathcal{H}$-Polyeder, $\pi\mathcal{H}$-in-$\pi\mathcal{H}$, co-NP-vollständig (Theorem 3.2.4).

Das $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ Enthaltenseinsproblem kann auf einfache Weise als ein bilineares Zulässigkeitsproblem mit zwei disjunkten Familien an Nebenbedingungen geschrieben werden. Interessanterweise sind die linearen Bedingungen des äußeren Polyeders nicht Teil des Zulässigkeitssystems, sondern ausschließlich die Koeffizientenmatrizen der projizierten Variablen. Dafür tauchen neue Variablen auf, deren Anzahl durch die Anzahl der linearen Bedingungen an das äußere Polytop bestimmt ist. Vom Standpunkt der computerorientierten Berechenbarkeit ergeben sich aus dieser Formulierung zwei Probleme. Zum einen sind die neuen Variablen unbeschränkt, zum anderen erfordert das Enthaltenseinszertifikat den Nachweis, dass ein Polynom auf einer polyedrischen Menge identisch null ist. Um diese Probleme zu umgehen, führen wir eine zusätzliche Voraussetzung ein und studieren ihre geometrische Bedeutung (Theorem 6.1.1).

Beim Übergang zum $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Problem ergibt sich das zusätzliche Problem, dass Projektionen von Spektraedern im Allgemeinen nicht abgeschlossen sind. Unter einer zusätzlichen Annahme an das lineare Büschel, können wir dieses Problem jedoch umgehen (Theorem 6.2.1). Wir betrachten die gleichen Probleme wie für den $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ Fall und diskutieren kurz eine alternative Formulierung unter Anwendung des Lemmas von Ramana, einer alternativen Variante des Lemmas von Farkas für Spektraeder (Proposition 6.2.5).

Während wir im allgemeinen Fall, dem $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Problem, auf Probleme in Form des Fehlens einer sauberen Version von Farkas' Lemma und eines auf die Situation passenden Positivstellensatzes stoßen, können wir in den Fällen $\pi\mathcal{H}$-in-$\mathcal{H}$ und $\pi\mathcal{S}$-in-$\mathcal{S}$ verschiedene Resultate aus den nicht-projektiven Fällen übertragen.

Zunächst geben wir ein lineares Kriterium für das $\pi\mathcal{H}$-in-$\mathcal{H}$ Enthaltenseinsproblem an (Theorem 6.3.1) und leiten davon ein hinreichendes semidefinites Kriterium für den $\pi\mathcal{S}$-in-$\mathcal{S}$ Fall ab (Theorem 6.3.3).

Die Formulierung des $\pi\mathcal{S}$-in-$\mathcal{S}$ Enthaltenseinsproblems als ein polynomielles Optimierungsproblem erlaubt die Anwendung des Positivstellensatzes von Hol und Scherer. Der Nachteil dieser Herangehensweise ist das Auftauchen der Projektionsvariablen im quadratischen Modul und daher in der Relaxierung durch Summe von Quadraten. Einer Idee von Gouveia und Netzer folgend, beweisen wir eine verfeinerte Version des Positivstellensatzes von Hol und Scherer für diese Situation, in der die Projektionsvariablen nicht auftreten (Theorem 6.3.5). Wie im nicht-projektiven Fall stimmt der initiale Relaxierungsschritt (basierend auf dem verfeinerten Ansatz) mit dem hinreichenden, geometrischen Kriterium überein. Wir zeigen auch, dass der initiale Relaxierungsschritt im Spezialfall $\pi\mathcal{S}$-in-$\mathcal{H}$ nicht nur hinreichend sondern auch notwendig ist (Theorem 6.3.8).

Abschließend diskutieren wir die mögliche Verallgemeinerung des Konzepts positiver linearer Abbildungen auf den projektiven Fall.

Gliederung der Dissertation. In die für diese Arbeit relevanten Teile der Theorie über Polyeder und Spektraeder sowie über lineare und semidefinite Optimierung führen wir in Kapitel 2 ein.

Kapitel 3 ist der Komplexität von Enthaltenseinsproblemen gewidmet. In Abschnitt 3.1 wiederholen wir bekannte Komplexitätsresultate für den polyedrischen Fall, insbesondere für das Enthaltenseinsproblem für Polytope. Die Abschnitte 3.2 und 3.3 klassifizieren die Komplexität der Enthaltenseinsprobleme für Projektionen von Polyedern und Spektraedern.

Kapitel 4 behandelt das Enthaltenseinsproblem für Polytope. Die Formulierung als bilineares Zulässigkeitsproblem ist in Abschnitt 4.1 zu finden. Die Abschnitte 4.2 und 4.3 behandeln die Anwendung der Positivstellensätze von Handelman beziehungsweise Putinar auf das bilineare Problem.

Das Enthaltenseinsproblem für Spektraeder diskutieren wir in Kapitel 5. Wir untersuchen eine Hierarchie semidefiniter Relaxierungen zur Entscheidung des Enthaltenseinsproblems in Abschnitt 5.1. In den Abschnitten 5.2 und 5.3 behandeln wir Zertifikate für den Fall $\mathcal{S}$-in-$\mathcal{H}$ und einige Beispiele. Die Verbindungen des Enthaltenseinsproblems für Spektraeder und der Theorie positiver linearer Abbildungen sind in Abschnitt 5.4 formuliert. Abschnitt 5.5 diskutiert einen skalarisierten Ansatz basierend auf der Formulierung des Enthaltenseinsproblems für Spektraeder als bilineares Problem.

In Kapitel 6 diskutieren wir die Erweiterung der Konzepte und Resultate aus den vorangegangenen Kapiteln auf die Enthaltenseinsprobleme für Projektionen. Wir beginnen mit dem $\pi\mathcal{H}$-in-$\pi\mathcal{H}$-Problem in Abschnitt 6.1 und behandeln dann das $\pi\mathcal{S}$-in-$\pi\mathcal{S}$-Problem in Abschnitt 6.2. Für den Fall $\pi\mathcal{S}$-in-$\mathcal{S}$ formulieren und beweisen wir eine Erweiterung des Positivstellensatzes von Hol und Scherer in Abschnitt 6.3. In Abschnitt 6.4 stellen wir den Zusammenhang zwischen einer Verallgemeinerung positiver linearer Abbildungen und dem Enthaltenseinsproblem von Projektionen von Spektraedern dar.

Abschließende Bemerkungen und eine kurze Diskussion offener Fragen sind in Kapitel 7 zu finden.

Contents

Deutsche Zusammenfassung **I**

1 Introduction **1**
1.1 Structure of the Thesis . . . 5
1.2 Publication in Advance . . . 6

2 Basic Facts on Polyhedra, Spectrahedra, and Optimization **7**
2.1 Notation . . . 7
2.2 Polyhedra and Linear Programming . . . 8
2.3 Spectrahedra and Semidefinite Programming . . . 11
2.4 Polynomial Optimization and Positivstellensätze . . . 20
2.4.1 Handelman's Positivstellensatz . . . 20
2.4.2 Putinar's Positivstellensatz . . . 21
2.4.3 Hol-Scherer's Positivstellensatz . . . 23

3 Complexity of Containment Problems for Projected Polytopes and Spectrahedra **27**
3.1 Complexity of the POLYTOPE CONTAINMENT Problem . . . 27
3.2 Complexity Concerning Projected Polytopes . . . 29
3.3 Complexity Concerning Projected Spectrahedra . . . 30

4 Positivstellensatz Certificates for the POLYTOPE CONTAINMENT Problem **33**
4.1 A Bilinear Programming Approach to the POLYTOPE CONTAINMENT Problem 33
4.2 Handelman Certificates for the POLYTOPE CONTAINMENT Problem . . . 37
4.2.1 Degree Bounds . . . 39
4.2.2 Examples . . . 40
4.3 Putinar Certificates for the POLYTOPE CONTAINMENT Problem . . . 41
4.3.1 Finite Convergence . . . 45
4.3.2 Examples . . . 47

5 Positivstellensatz Certificates for the SPECTRAHEDRON CONTAINMENT Problem **49**
5.1 Hol-Scherer Certificates for the SPECTRAHEDRON CONTAINMENT Problem . . 49
5.1.1 From the Polyhedral Case to the Spectrahedral Case . . . 50
5.1.2 Hol-Scherer Certificates for SPECTRAHEDRON CONTAINMENT . . . 55
5.1.3 The Initial Step of the Hol-Scherer Hierarchy . . . 57
5.2 Certificates for Containment of $\mathcal{H}$-Polyhedra and Spectrahedra . . . 61
5.2.1 Spectrahedron in $\mathcal{H}$-Polyhedron . . . 62
5.2.2 $\mathcal{H}$-Polyhedron in Spectrahedron – A Note on the Polytope Recognition Problem . . . 64
5.3 Containment Certificates for Some Structured Classes . . . 65
5.3.1 Some Exact Cases . . . 65
5.3.2 Containment of Scaled Spectrahedra . . . 68
5.4 Containment of Spectrahedra and Positive Linear Maps . . . 70
5.4.1 Positive Linear Maps . . . 71
5.4.2 Containment of 2-dimensional Spectrahedra . . . 72

5.5 A Bilinear Programming Approach to Spectrahedron Containment . . . 76

6 The Projection Containment Problems **79**

6.1 A Bilinear Formulation of the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ Containment Problem 79

6.2 A Bilinear Formulation of the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Containment Problem 82

6.2.1 An Alternative Approach to the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Containment Problem . . . 84

6.3 Positivstellensatz Certificates for the $\pi\mathcal{S}$-in-$\mathcal{S}$ Containment Problem 85

6.3.1 From the $\pi\mathcal{H}$-in-$\mathcal{H}$ to the $\pi\mathcal{S}$-in-$\mathcal{S}$ Containment Problem 85

6.3.2 The Naive Way 87

6.3.3 A more Sophisticated Positivstellensatz 87

6.3.4 Examples 90

6.4 Containment of Projected Spectrahedra and Positive Linear Maps 93

7 Final Remarks and Open Questions **97**

Bibliography **99**

1 Introduction

Containment problems belong to the classical problems of (convex) geometry. In the proper sense, a containment problem is the task to decide the set-theoretic inclusion of two given sets, which is hard from both the theoretical and the practical perspective. In a broader sense, this includes, e.g., radii [GK92] or packing problems [BCS99], which are even harder. For some classes of convex sets there has been strong interest in containment problems. This includes containment problems of polyhedra and balls [FO85], and containment of polyhedra [GK94], which have been studied in the late 20th century because of their inherent relevance in linear programming and combinatorics. Since then, there has only been limited progress in understanding containment problems of that type. In recent years, containment problems for spectrahedra, which naturally generalize the class of polyhedra, have seen great interest [BTN02, HKM13]. This interest is particularly driven by the intrinsic relevance of spectrahedra and their projections in polynomial optimization [Ble12, GPT10] and convex algebraic geometry [HV07]. Except for the treatment of special classes or situations, there has been no overall treatment of that kind of problems, though.

In this thesis, we provide a comprehensive treatment of containment problems concerning polyhedra, spectrahedra, and their projections from the viewpoint of low-degree semialgebraic problems and study algebraic certificates for containment. This leads to a new and systematic access to studying containment problems of (projections of) polyhedra and spectrahedra, and provides several new and partially unexpected results.

The main idea – which is meanwhile common in polynomial optimization, but whose understanding of the particular potential on low-degree geometric problems is still a major challenge – can be explained as follows. One point of view towards linear programming is as an application of Farkas' Lemma which characterizes the (non-)solvability of a system of linear inequalities. The affine form of Farkas' Lemma characterizes linear polynomials which are nonnegative on a given polyhedron. By omitting the linearity condition, one gets a polynomial nonnegativity question on a semialgebraic set, leading to so-called Positivstellensätze (or, more precisely, Nichtnegativstellensätze). A Positivstellensatz provides a *certificate* for the positivity of a polynomial function in terms of a polynomial identity. As in the linear case, these Positivstellensätze are the foundation of polynomial optimization and relaxation methods; see, e.g., [BPT13, Las10, Lau09]. The transition from positivity to nonnegativity is still a major challenge in real algebraic geometry and polynomial optimization.

With this in mind, several principal questions arise in the context of containment problems: Can the particular containment problem be formulated as a polynomial nonnegativity (or, feasibility) problem in a sophisticated way? If so, how are positivity and nonnegativity related to the containment question in the sense of their geometric meaning? Is there a sophisticated Positivstellensatz for the particular situation, yielding certificates for containment? Concerning the degree of the semialgebraic certificates, which degree is necessary, which degree is sufficient to decide containment?

Indeed, (almost) all containment problems studied in this thesis can be formulated as polynomial nonnegativity problems allowing the application of semialgebraic relaxations. Other than this general result, the answer to all the other questions (highly) depends on the specific containment problem, particularly with regard to its underlying geometry. An important point is whether the hierarchies coming from increasing the degree in the polynomial relaxations always decide containment in finitely many steps. We understand *finite convergence* in this sense, as opposed to asymptotic convergence.

In the subsequent paragraphs, we outline the main problems tackled in this thesis and our contribution in their understanding from the viewpoint of semialgebraic geometry. We focus on the containment problem of an $\mathcal{H}$-polytope in a $\mathcal{V}$-polytope and of a spectrahedron in a spectrahedron, henceforth referred to as the POLYTOPE CONTAINMENT problem and the SPECTRAHEDRON CONTAINMENT problem, respectively. Moreover, we address containment problems concerning projections of $\mathcal{H}$-polyhedra and spectrahedra, which we subsume under the label PROJECTION CONTAINMENT problems. This selection is justified by the fact that the mentioned containment problems are computationally hard and their geometry is not well understood. For computational questions, we always assume that the input data is given in terms of rational numbers.

The POLYTOPE CONTAINMENT **Problem.** A polytope can be represented as the convex hull of finitely many points ("$\mathcal{V}$-polytope") or as the intersection of finitely many halfspaces ("$\mathcal{H}$-polytope"). For polytopes, the computational geometry and computational complexity of containment problems have been studied in detail. A prominent problem in algorithmic polytope theory is the following.

> POLYTOPE CONTAINMENT:
> **Input:** $d \in \mathbb{N}$, an $\mathcal{H}$-polytope $P \subseteq \mathbb{R}^d$ and a $\mathcal{V}$-polytope $Q \subseteq \mathbb{R}^d$.
> **Task:** Decide whether $P \subseteq Q$.

Due to Freund and Orlin [FO85], the POLYTOPE CONTAINMENT problem is known to be co-NP-complete; note that the converse question $Q \subseteq P$ is trivial to decide. If the dimension of one (or both) polytopes is fixed, then the POLYTOPE CONTAINMENT problem is solvable in polynomial time (Theorem 3.1.3). For the POLYTOPE CONTAINMENT problem, despite of its fundamental nature, there seems to be only limited progress on that problem so far.

We formulate the POLYTOPE CONTAINMENT problem as a disjointly constrained bilinear feasibility problem (or, equivalently, as the maximization problem of a bilinear function on the product of two $\mathcal{H}$-polytopes; Proposition 4.1.1). The bilinear formulation of the problem allows to apply linear relaxations based on Handelman's Positivstellensatz [Han88] respectively semidefinite relaxations based on Putinar's Positivstellensatz [Put93]. While in the case of *strong containment* (i.e., $P \subseteq Q$ and $\partial P \cap \partial Q = \emptyset$) finite convergence comes out of the general theory, in the case of non-strong containment this is a critical issue.

Handelman's Positivstellensatz (Proposition 2.4.1) deals with positivity of polynomials on polytopes and provides a hierarchy of linear programs. Beyond a standard convergence result (Theorem 4.2.1), we provide characterizations on the (widely open) question of degree bounds (Theorem 4.2.5).

Putinar's Positivstellensatz (Proposition 2.4.3) deals with more general, semialgebraic constraint sets and provides a hierarchy of semidefinite programs. As a main result, we show that under mild and explicitly known conditions the Putinar relaxation converges in finitely many steps (Theorems 4.3.1 and 4.3.6), based on the so-called boundary Hessian condition stated by Marshall [Mar08].

While it is a fundamental geometric problem by itself, we sketch an exemplary application scenario in which the POLYTOPE CONTAINMENT problem occurs. Generally, many applications in data analysis or shape analysis of point clouds involve the convex hull of point sets (see, e.g., [BK01]), and a Polytope Containment problem can be used to answer questions about certain (polyhedral) properties on the set.

A textbook type example in diet realization. We consider d different food types and k underlying basic nutrients. Assume that each unit of food j contains A_{ij} units of the ith nutrient a_i. A dietary requirement is described by linear inequalities of the form $\sum_{j=1}^{d} A_{ij}x_j \geq a_i,\ 1 \leq i \leq k$, with positive $A_{i1}, \ldots, A_{id}$ for a minimum of a requirement and negative $A_{i1}, \ldots, A_{id}$ for a maximum of a requirement. Thus the requirements define an $\mathcal{H}$-polytope P. Moreover, assume we are given l fixed combinations of can food, each can containing one unit. Type s consists of an amount of $b_t^{(s)}$ of food t. The convex combinations of the vectors $b^{(1)}, \ldots, b^{(l)}$ correspond to the food which can be combined from the can food, where the convex combination signifies that the resulting food has also the size of one unit. The question is, if every food combination satisfying the dietary constraints can be assembled from the can food in that way. This is a Polytope Containment problem. Similar scenarios occur, e.g., in the mixing of liquids (such as oil).

The Spectrahedron containment Problem. A spectrahedron can be represented as the slice of the cone of positive semidefinite matrices with an affine subspace or as the positivity region of a linear (matrix) pencil, i.e., a symmetric matrix with linear polynomial entries. From this it is evident that every polyhedron (in $\mathcal{H}$-representation) is a spectrahedron, just by writing the linear polynomials defining the halfspaces onto the diagonal. We define the following problem.

Spectrahedron containment:
Input: $d \in \mathbb{N}$, spectrahedra $S_A \subseteq \mathbb{R}^d$ and $S_B \subseteq \mathbb{R}^d$.
Task: Decide whether $S_A \subseteq S_B$.

Since polyhedra are special cases of spectrahedra, the Spectrahedron Containment problem covers, in particular, the $\mathcal{H}$-in-$\mathcal{H}$, $\mathcal{H}$-in-$\mathcal{S}$, and $\mathcal{S}$-in-$\mathcal{H}$ containment problems. While the first problem is known to be solvable in polynomial time, the second problem is co-NP-hard, implying the co-NP-hardness of the Spectrahedron containment problem. The complexity status of the third problem is not completely solved.

Ben-Tal and Nemirovski studied the matrix cube problem [BTN02], which corresponds to the Spectrahedron Containment problem where S_A is a cube. They derived a sufficient semidefinite criterion to decide containment in this situation. In a much more general setting, Helton, Klep, and McCullough [HKM12, HKM13] studied containment problems of matricial positivity domains, also called free spectrahedra, which live in a union of spaces of different dimensions. As a byproduct, they also derived some implications for containment of spectrahedra. In the case of the matrix cube problem, the sufficient criterion of Helton et al. coincides with Ben-Tal-Nemirovski's criterion.

Our point of departure is the containment problem for pairs of $\mathcal{H}$-polyhedra, which can be decided in polynomial time by solving a linear feasibility problem (Theorem 5.1.1). From that we derive a sufficient semidefinite criterion to decide the Spectrahedron Containment problem (Theorem 5.1.3). As the semidefinite feasibility criterion coming from the geometric approach is only sufficient for the Spectrahedron Containment problem and fails to certify containment already in two-dimensional examples (Example 5.1.7), two principal questions arise. First, under which additional assumptions necessity holds, and, second, whether there is a better criterion in the sense that necessity can be achieved.

To tackle the second question, we formulate the Spectrahedron Containment problem in terms of a polynomial feasibility problem (or, equivalently, as a quantified polynomial optimization problem), yielding a hierarchy of sufficient semidefinite criteria to decide containment. Due to the absence of a clean Farkas-type Lemma for linear pencils (i.e., a Farkas Lemma without preconditions) and contrary to the scalar setting, finite convergence in the case of *strong containment* (i.e., $S_A \subseteq \text{int}(S_B)$) is not an immediate consequence of the general theory. However, under a mild technical assumption, called reducedness of a linear pencil, finite convergence in the strong containment case can be achieved by applying Hol-Scherer's Positivstellensatz (Theorem 5.1.8). From this asymptotic convergence can be deduced if the inner set is a spectratope (i.e., a bounded spectrahedron). This being said, the behavior of the hierarchy for non-strong containment or non-containment is uncertain. As in the scalar case, the transition from positive definiteness to positive semidefiniteness is a major challenge.

With this in mind and in order to tackle the mentioned problems, we study the initial relaxation step of the hierarchy in more detail. We start by a fundamental result that stresses the importance of the initial relaxation step of the Hol-Scherer hierarchy. It turns out that the initial step of the hierarchy coincides with the solitary semidefinite criterion coming from the geometric approach (Theorem 5.1.11). Thus studying the initial relaxation step is the same as studying the quality of the solitary criterion. We develop some auxiliary results on the containment criteria, showing that they behave geometrically. From that analysis, we are able to provide (partially explicit) containment certificates for some structured examples and important cases.

Specializing the Spectrahedron Containment problem to the case $\mathcal{S}$-in-$\mathcal{H}$, the containment question can be decided with the initial relaxation step if some mild assumptions hold (Theorem 5.2.3). This remains true if the outer set is given by a linear pencil whose coefficient matrices are simultaneously diagonalizable. The proof of the statement is based on various properties of the containment criteria and combining them with duality theory of semidefinite programming. The exactness for the $\mathcal{S}$-in-$\mathcal{H}$ case is particularly surprising, since a priori the criteria depend on the linear pencil representation of the spectrahedron.

In the case of bounded spectrahedra, certifying containment can always be achieved by an appropriate scaling of one of the spectrahedra involved (Theorem 5.3.1).

Containment of spectrahedra is also linked to positivity of linear maps. This has first been investigated by Helton, Klep, and McCullough [HKM13] in a much more general setting. A well-known relaxation in this setting is complete positivity of a linear map, which can be tested by a semidefinite program. It transpires that complete positivity is equivalent to containment being certified by the initial relaxation step. Thus our hierarchy for the Spectrahedron Containment problem also serves as a hierarchy between positivity and complete positivity. On the other hand, the theory of positive linear maps gives some interesting insights into the Spectrahedron Containment problem. Most notably, we can prove necessity of the initial Hol-Scherer relaxation for a special family of 2-dimensional spectratopes (Theorem 5.4.10).

The Projection Containment **Problems.** It is well-known that the projection of a polyhedron is again a polyhedron. For a spectrahedron this is no longer true. In particular, the projection of a spectrahedron is not necessarily closed. For a set $S \subseteq \mathbb{R}^{d+m}$, denote by $\pi(S)$ the projection of S onto $\mathbb{R}^d$. Since $\pi\mathcal{H}$-polyhedra build a subclass of $\pi\mathcal{S}$-spectrahedra, we subsume the mentioned problems under the following label.

Projection Containment:
Input: $d, m, n \in \mathbb{N}$, spectrahedra $S_A \subseteq \mathbb{R}^{d+m}$ and $S_B \subseteq \mathbb{R}^{d+n}$.
Task: Decide whether $\pi(S_A) \subseteq \pi(S_B)$.

Most results on the complexity classification for containment problems involving polyhedra and spectrahedra can be brought forward to Projection Containment. However, while the $\mathcal{H}$-in-$\mathcal{H}$ containment problem is solvable in polynomial time, deciding containment of two projected $\mathcal{H}$-polytopes (i.e., $\pi\mathcal{H}$-in-$\pi\mathcal{H}$) turns out to be co-NP-complete (Theorem 3.2.4).

The $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment problem can be formulated as a disjointly constrained bilinear feasibility problem in a straightforward way. Interestingly, the projection of the outer polyhedron do not appear in the feasibility system, only the corresponding coefficients in the $\mathcal{H}$-representation and some new variables depending on the size of the representation. From a computational viewpoint, this formulation lacks in two ways. First, the new variables are unbounded and, second, certifying containment requires to test whether a polynomial is identically zero on a polyhedral set. To tackle these problems, we introduce an additional precondition and study its geometric meaning (Theorem 6.1.1).

Continuing with the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem, a polynomial formulation lacks in the fact that projected spectrahedra are not closed in general. However, under an additional assumption, which is common in semidefinite programming, the statement holds (Theorem 6.2.1). We treat the same problems as for $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ and discuss briefly an alternative approach based on the Farkas-type Lemma of Ramana (Proposition 6.2.5).

While for the general case, the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem, we hit on problems like the lack of a clean Farkas Lemma for cones as well as the absence of a sophisticated Positivstellensatz, retreating to the $\pi\mathcal{H}$-in-$\mathcal{H}$ and $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problems allows to bring forward several results from the non-projected case.

To begin with, we establish a linear feasibility criterion for the $\pi\mathcal{H}$-in-$\mathcal{H}$ containment problem (Theorem 6.3.1). From that we deduce a sufficient semidefinite criterion for the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem (Theorem 6.3.3).

The formulation of the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem as a polynomial optimization problem allows to apply Hol-Scherer's Positivstellensatz. The drawback of this approach is the appearance of the projection variables in the quadratic module and thus in the sum of squares hierarchy. Following a result of Gouveia and Netzer [GN11], we establish a more sophisticated Hol-Scherer like Positivstellensatz for this situation, removing the projection variables from the quadratic module (Theorem 6.3.5). As in the non-projected case, the initial step of the sum of squares hierarchy coincides with the sufficient semidefinite criterion coming from the geometric approach. Moreover, as in the non-projected situation, if the outer set is a polyhedron (i.e., the $\pi\mathcal{S}$-in-$\mathcal{H}$ containment problem), then the initial relaxation step is not only sufficient for containment but also necessary (Theorem 6.3.8).

To some extent, the concept of positive linear maps can also be generalized to the projected case (Theorem 6.4.1).

1.1 Structure of the Thesis

We start with the preliminary Chapter 2 introducing notation and relevant background on polyhedra and spectrahedra, as well as linear, semidefinite, and polynomial optimization.

In Chapter 3, we review the complexity classification for containment problems concerning polyhedra and extend it to projections of polyhedra and spectrahedra. The complexity classification of the Polytope Containment problem is reviewed in Section 3.1. Section 3.2 approaches the case involving only polyhedra and their projections. Section 3.3 contains the complexity classification involving also projections of spectrahedra.

Chapter 4 discusses the Polytope Containment problem. The reformulation of the problem as a bilinear feasibility problem is stated in Section 4.1. The Sections 4.2 and 4.3 deal with the application of Handelman's and Putinar's Positivstellensatz, respectively, to the bilinear reformulation.

The Spectrahedron Containment problem is dealt with in Chapter 5. We state and study a hierarchy of semidefinite relaxations coming from a sum of squares approach in Section 5.1. Sections 5.2 and 5.3 show the effectiveness of the approach by providing finite convergence in several important cases, including containment of a spectrahedron in an $\mathcal{H}$-polyhedron. The connections between containment of spectrahedra and positive linear maps are given in Section 5.4. In Section 5.5, we discuss briefly a scalarized sum of squares approach coming from a bilinear formulation of the Spectrahedron Containment problem.

In Chapter 6, we discuss extensions of the concepts and results stated in Chapter 5 to the Projection Containment problems. Starting from the reformulation of the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment problem as a polynomial feasibility problem in Section 6.1, we discuss the adaption to the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem in Section 6.2. For the case $\pi\mathcal{S}$-in-$\mathcal{S}$, an extension of Hol-Scherer's Positivstellensatz is stated and proved in Section 6.3. In Section 6.4, the connections between (a generalization of) positive linear maps and containment of projected spectrahedra are outlined.

Final remarks and a short discussion of open questions can be found in Chapter 7.

1.2 Publication in Advance

Parts of this thesis are published or submitted in advance. Chapter 4 is based on a joint work with Thorsten Theobald that is available online as a preprint and has been accepted for a contributed talk at the MEGA conference series (Trento, 2015) but is yet not published [KT14]. The content of Chapter 5 is based on joint works with Thorsten Theobald and Christian Trabandt [KTT13, KTT15], both published in the SIAM Journal on Optimization.

The complexity classification of containment problems involving polyhedra and spectrahedra as stated in [KTT13] is part of Christian Trabandt's PhD thesis [Tra14]. In Chapter 3, we extend the results to projections of polyhedra and spectrahedra.

[KTT15] is concerned with an optimization approach based on Lasserre's moment method for polynomial matrix inequalities [HL06]. As this is part of [Tra14], we do not state these parts here but have a brief look at some of them from the viewpoint of the thesis at hand in Section 5.5.

2 Basic Facts on Polyhedra, Spectrahedra, and Optimization

In this chapter, we fix terminology and recall the basic concepts of linear, semidefinite, and polynomial optimization. Most proofs of the statements are omitted. We do not give references to all facts stated here as they belong to the basics of the respective theory but give general references at the beginning of the subsections. For an overall survey of linear and semidefinite programming see, e.g., the monograph of Tunçel [Tun10]; regarding semidefinite and polynomial optimization, we refer to the book of Blekherman et. al. [BPT13].

2.1 Notation

The set of positive (resp. nonnegative) integers is denoted by $\mathbb{N}$ (resp. $\mathbb{N}_0$). For two nonnegative integers $m \leq n$, we write $[m,n] = \{m, m+1, \ldots, n\}$ and $[n] = [1,n] = \{1, 2, \ldots, n\}$.

For a set $M \subseteq \mathbb{R}^d$, the *convex hull* of M, denoted by $\operatorname{conv} M$, is the set of all *convex combinations*

$$\sum_{i=1}^{m} \lambda_i x_i \text{ with } \mathbb{1}_m^T \lambda = 1,\ \lambda_i \geq 0\ \forall i \in [m] \text{ for some } x_i \in M,$$

where $\mathbb{1}_m$ denotes the all-ones vector in $\mathbb{R}^m$. The *linear span* or linear hull, $\operatorname{span} M$, of M is the smallest linear subspace containing M. The smallest affine subspace containing M is referred to as the *affine span* $\operatorname{aff} M$ of M. For a subspace $V \subseteq \mathbb{R}^d$, $V^\perp$ denotes the *orthogonal complement* of V with respect to the Euclidean inner product. The interior (resp. closure) of a set M is denoted by $\operatorname{int} M$ (resp. $\operatorname{cl} M$).

By $\mathbb{R}^d_+$ and $\mathbb{R}^d_{++}$ we denote the closed convex cone of componentwise nonnegative and positive vectors, respectively.

The *(geometric) polar* of a subset $M \subseteq \mathbb{R}^d$ is defined as the set

$$M^\circ = \left\{ y \in \mathbb{R}^d \mid y^T x \leq 1\ \forall x \in M \right\} \subseteq \mathbb{R}^d.$$

It is easy to see that M° is a closed convex set containing the origin.

Proposition 2.1.1. *Let $M, N \subseteq \mathbb{R}^d$.*

(1) $M \subseteq N$ *implies* $N^\circ \subseteq M^\circ$.

(2) $M^{\circ\circ} = \operatorname{cl}\operatorname{conv}(M \cup \{0\})$. *In particular, if M is a closed convex set containing the origin, then $M^{\circ\circ} = M$.*

(3) M° *is bounded if and only if* $0 \in \operatorname{int} M$.

The *kernel* or nullspace (resp. *span* or range) of a matrix $M \in \mathbb{R}^{k \times l}$ is defined as the set $\ker M = \{x \in \mathbb{R}^l \mid Mx = 0\}$ (resp. the linear span of its columns, i.e., $\operatorname{span} M = \{y \in \mathbb{R}^k \mid y = Mx,\ x \in \mathbb{R}^l\}$).

For a matrix A, the (i,j)th entry of A is labeled by A_{ij} as usual. For a block matrix B, we label the (i,j)th block by B_{ij} and the (s,t)th entry of B_{ij} by $(B_{ij})_{st}$. A square matrix with 1 in the entry (i,j) and zeros otherwise is denoted by E_{ij}. The $n \times n$ identity matrix is denoted by I_n.

For a square matrix $A \in \mathbb{R}^{k\times k}$, we call $\mathrm{tr}(A) = \sum_{i=1}^{k} A_{ii}$ the *trace* of A. For square matrices $A, B \in \mathbb{R}^{k\times k}$, we define the scalar product

$$\langle A, B\rangle = \mathrm{tr}(B^T A) = \sum_{i,j=1}^{k} A_{ij} B_{ij}.$$

Let $\mathcal{S}^k$ be the space of symmetric $k \times k$-matrices with real entries. For $A \in \mathcal{S}^k$, we write $A \succeq 0$ (resp. $A \succ 0$) if A is positive semidefinite (resp. positive definite). The (non-)closed, convex cone of positive semidefinite (resp. definite) matrices is denoted by $\mathcal{S}^k_+$ (resp. $\mathcal{S}^k_{++}$). For $x = (x_1, \ldots, x_d)$, $\mathcal{S}^k[x]$ denotes the space of symmetric $k \times k$-matrices with entries in the polynomial ring $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_d]$.

For matrices $A \in \mathbb{R}^{k\times k}$ and $B \in \mathbb{R}^{l\times l}$, their *Kronecker product* $A \otimes B$ is the $kl \times kl$-matrix

$$A \otimes B = \begin{bmatrix} A_{11}\, B & \ldots & A_{1k}\, B \\ \vdots & \ddots & \vdots \\ A_{k1}\, B & \ldots & A_{kk}\, B \end{bmatrix}. \tag{2.1.1}$$

It is well-known (see, e.g., [HJ94, Corollary 4.2.13]) that the Kronecker product of two symmetric (positive semidefinite) matrices is again symmetric (positive semidefinite).

Given matrices $M_1 \in \mathbb{R}^{k_1\times l_1}, \ldots, M_d \in \mathbb{R}^{k_d\times l_d}$, we denote by

$$\bigoplus_{i=1}^{d} M_i = \begin{bmatrix} M_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & M_d \end{bmatrix}$$

their *direct sum*, which is a blockdiagonal matrix with $\sum_{i=1}^{d} k_i$ rows and $\sum_{j=1}^{d} l_j$ columns.

2.2 Polyhedra and Linear Programming

We recall relevant notation, existing fundamental results on the theory of polyhedra, and sketch the basic concepts of linear programming as they are relevant to us. As general references, we refer to the books of Schrijver [Sch86] and Ziegler [Zie95].

For $a \in \mathbb{R}^k$ and $A \in \mathbb{R}^{k\times d}$, an *($\mathcal{H}$-)polyhedron* is defined as the intersection of finitely many affine halfspaces,

$$P_A = \left\{ x \in \mathbb{R}^d \mid a - Ax \geq 0 \right\}, \tag{2.2.1}$$

and a bounded polyhedron is called an *$\mathcal{H}$-polytope*. For $B \in \mathbb{R}^{d\times l}$, the set

$$Q_B = \mathrm{conv}(B) = \left\{ x \in \mathbb{R}^d \mid \exists \lambda \in \mathbb{R}^l_+ :\ x = B\lambda,\ \mathbb{1}_l^T \lambda = 1 \right\} \tag{2.2.2}$$

is called a *$\mathcal{V}$-polytope*, where $\mathbb{1}_l$ denotes the all-ones vector in $\mathbb{R}^l$. The subscript in the notion of P_A and Q_B indicates the dependency on the specific representation (a, A) and B, respectively, of the polytopes involved. However, if there is no risk of confusion, we often state P and Q without subscript. Every $\mathcal{H}$-polytope is a $\mathcal{V}$-polytope and vice versa.

A *face* of a polyhedron P is the (nonempty) intersection of P with an affine hyperplane

H such that P is contained in one of the closed halfspaces defined by H. The empty set is an improper face, called the empty face. A face is called *vertex* (resp. *facet*) if its dimension, i.e., the dimension of its affine span, is zero (resp. $\dim(P) - 1$). We denote by $V(P)$ the set of vertices of a polytope P, and by $F(P)$ the set of facets. Clearly, every $\mathcal{H}$-representation of a polytope contains all facet defining inequalities (up to a scalar multiple) and every $\mathcal{V}$-representation contains all the vertices. Inequalities (resp. points) not defining a facet (resp. vertex) are called *redundant constraints*. Note that removing redundant constraints is a polynomial time process; see, e.g., [GK93, Theorem 2.1] for a constructive proof.

By McMullen's Upper bound Theorem [McM70], any d-dimensional polytope, d-polytope for short, with k vertices (resp. facets) has at most

$$\binom{k - \left\lfloor \frac{1}{2}(d+1) \right\rfloor}{k-d} + \binom{k - \left\lfloor \frac{1}{2}(d+2) \right\rfloor}{k-d}$$

facets (resp. vertices). This bound is sharp for neighborly polytopes, such as cyclic polytopes. Thus a $\mathcal{V}$-representation of a given $\mathcal{H}$-polytope (or vice versa) can be exponential in the dimension d and the number of facets k (resp. the number of vertices).

An $\mathcal{H}$-polyhedron $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ contains the origin if and only if $a \geq 0$. Moreover, P contains the origin in its interior $\operatorname{int} P$ if and only if $a > 0$. In this situation, $P = \{x \in \mathbb{R}^d \mid \mathbb{1}_k - Ax \geq 0\}$ after an appropriate scaling of the inequalities.

If $P \subseteq \mathbb{R}^d$ and $Q \subseteq \mathbb{R}^e$ are polytopes, then their product $P \times Q \subseteq \mathbb{R}^{d+e}$ is a polytope of dimension $\dim P + \dim Q$, whose nonempty faces are the products of nonempty faces of P and nonempty faces of Q; see for example [Zie95, Page 10].

The next lemma plays a prominent role in polyhedral and linear programming theory.

Lemma 2.2.1 (Farkas' Lemma I [Zie95, Proposition 1.7])**.** *Let $A \in \mathbb{R}^{k \times d}$ and $z \in \mathbb{R}^k$. Then exactly one of the following two systems has a solution.*

$$\exists x \in \mathbb{R}^d : \; Ax \leq z \tag{2.2.3}$$

$$\exists c \in \mathbb{R}^k : \; c \geq 0, \; c^T A = 0, \; c^T z < 0 \tag{2.2.4}$$

There are many equivalent formulations of Farkas' Lemma (see, e.g., [Zie95, Proposition 1.8 ff.]). For us the affine form of Farkas' Lemma is often more suitable.

Lemma 2.2.2 (Affine form of Farkas' Lemma [Sch86, Corollary 7.1h])**.** *Let the polyhedron $P = \{x \in \mathbb{R}^d \mid l_i(x) \geq 0, \; i \in [m]\}$ with affine functions $l_i : \mathbb{R}^d \to \mathbb{R}$ be nonempty. Then every affine $l : \mathbb{R}^d \to \mathbb{R}$ that is nonnegative on P can be written as $l(x) = c_0 + \sum_{i=1}^m c_i l_i(x)$ with nonnegative coefficients c_i.*

The following lemma characterizes boundedness of polyhedra using the so-called recession cone or characteristic cone.

Lemma 2.2.3 ([Sch86, Section 8.2])**.** *Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ be a nonempty polyhedron. P_A is bounded if and only if its recession cone $R_A = \{x \in \mathbb{R}^d \mid Ax \leq 0\}$ is zero-dimensional, i.e., $R_A = \{0\}$.*

From the above lemma it follows that boundedness of a polyhedron can be formulated as a linear feasibility problem (LFP), which is solvable in polynomial time [Sch86, Theorem 13.4].

The next lemma is a special case of Motzkin's transposition theorem [Sch86, Corollary 7.1k]; see also [Sch86, Section 7.8] and the references therein.

Lemma 2.2.4 (Stiemke's Transposition Theorem)**.** *Let* $A \in \mathbb{R}^{k \times d}$ *and* $a \in \mathbb{R}^k$*. Then*

$$\exists x \in \mathbb{R}^d : \ x > 0, \ Ax = 0 \quad \Longleftrightarrow \quad \left\{ y \in \mathbb{R}^k \mid y^T A \geq 0 \right\} = \left\{ y \in \mathbb{R}^k \mid y^T A = 0 \right\}$$

A famous and fundamental polytope is the *(standard) d-simplex*

$$\Delta^d = \left\{ x \in \mathbb{R}^d \mid \mathbb{1}_d^T x = 1, x \geq 0 \right\}. \tag{2.2.5}$$

Here the prefix d denotes the dimension of the simplex plus one, i.e., $\dim \Delta^d = d - 1$.

Linear Programming. Linear Programming (LP) is the problem of maximizing (resp. minimizing) a linear function over a polyhedron. In this situation the linear function is referred to as the *objective function*. A point of the polyhedron is called *feasible solution*. If the objective function attains its optimum value in a feasible point, then the point is called an *optimal solution*.

Given an $\mathcal{H}$-polyhedron $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ and a vector $c \in \mathbb{R}^d$, we call

$$\begin{aligned} \sup \ & c^T x \\ \text{s.t.} \ & x \in P_A \end{aligned} \tag{2.2.6}$$

the *primal* LP. To the primal problem, we associate the so-called *dual* problem

$$\begin{aligned} \inf \ & a^T y \\ \text{s.t.} \ & y^T A = c \\ & y \geq 0. \end{aligned} \tag{2.2.7}$$

The set of feasible solutions $\{y \geq 0, \ y^T A = c\}$ for (2.2.7) is again a polyhedron (dependent on c).

It is evident from the construction of the problems, that *weak duality* holds, i.e., for every pair of primal and dual feasible solutions $(x, y) \in \mathbb{R}^d \times \mathbb{R}^k$, it holds that $c^T x \leq a^T y$. If both, x and y, are not only feasible but also optimal, then we get a stronger result.

Proposition 2.2.5 (Strong Duality in Linear Programming [Sch86, Corollary 7.1g])**.** *Let* P_A *be a nonempty* $\mathcal{H}$*-polyhedron. Assume that the dual polyhedron is also nonempty. Let* $c \in \mathbb{R}^d$ *be fixed. Then*

$$\max \left\{ c^T x \mid x \in P_A \right\} = \min \left\{ a^T y \mid y \geq 0, \ A^T y = c \right\}.$$

In particular, for every pair of primal and dual optimal solutions (x, y)*, we have* $c^T x = a^T y$*.*

An LP can be solved efficiently from both the theoretical and the practical perspective. In the late 1970s, Khachiyan [Kha80] proved that linear programming (with rational input) is solvable in polynomial time using the ellipsoid method (see also [Sch86, Theorem 13.4]). In practice, however, the ellipsoid method turns out to be impracticable. On the other hand, the famous simplex algorithm (see, e.g., [Sch86, Chapter 11]) is highly efficient in practice while its complexity is unknown. In recent years, interior point algorithms provide (alongside with the simplex algorithm) the most effective solution technique; see, e.g., [GM07]. For a brief history of linear programming we refer to [Bix12]. Latest benchmarks on linear programming can be found on Mittelmann's homepage [Mit].

Projection of Polyhedra. Denote by $\pi : \mathbb{R}^{d+m} \to \mathbb{R}^d$, $(x, y) \mapsto x$ the linear coordinate projection map. First we treat the case of $\mathcal{V}$-presented polytopes.

Lemma 2.2.6. *Let $B = \{b^{(1)}, \ldots, b^{(l)}\} \subseteq \mathbb{R}^{d+m}$ with $b^{(i)} = (x^{(i)}, y^{(i)})$. Denote by $\pi(Q)$ the projection of the $\mathcal{V}$-polytope $Q = \operatorname{conv}(B)$. Then $\pi(Q) = \operatorname{conv}\{x^{(1)}, \ldots, x^{(l)}\}$.*

Proof. Write $X := \operatorname{conv}\{x^{(1)}, \ldots, x^{(l)}\}$ and let $p \in X$. Then there exists a $\lambda \in \mathbb{R}^l_+$ such that $\mathbb{1}_l^T \lambda = 1$ and $p = \sum_{i=1}^l \lambda_i x^{(i)}$. Define $q = \sum_{i=1}^l \lambda_i y^{(i)}$. Then $(p, q) \in Q$ and hence $p \in \pi(Q)$.

For the converse, let $p \in \pi(Q)$. Then there exists $q \in \mathbb{R}^m$ such that $(p, q) \in Q$, i.e., there exists a vector $\lambda \in \mathbb{R}^l_+$ such that $\mathbb{1}_l^T \lambda = 1$ and $(p, q) = \sum_{i=1}^l \lambda_i b^{(i)} = \sum_{i=1}^l \lambda_i (x^{(i)}, y^{(i)}) = (\sum_{i=1}^l \lambda_i x^{(i)}, \sum_{i=1}^l \lambda_i y^{(i)})$. Thus $p \in \operatorname{conv}\{x^{(1)}, \ldots, x^{(l)}\} = X$. □

While for $\mathcal{V}$-polytopes the situation is straightforward, the situation changes for projections of $\mathcal{H}$-polytopes. By Fourier-Motzkin elimination, given an $\mathcal{H}$-polyhedron $P = \{(x, y) \in \mathbb{R}^{d+m} \mid a - Ax - A'y \geq 0\}$, the projection of P onto the x coordinates, $\pi(P)$, is again an $\mathcal{H}$-polyhedron. Unfortunately, a quantifier-free $\mathcal{H}$-description of $\pi(P)$ can be exponential in the input size (d, m, k) respectively (d, n, l); see [Zie95, Sections 1.2 and 1.3] and the references therein.

From Farkas' Lemma one can deduce a description of the polar of a polyhedron and its projection.

Proposition 2.2.7. *Let $\pi(P_A) = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m : \; a - Ax - A'y \geq 0\}$ be the projection of a polyhedron with $a \in \mathbb{R}^k_+$. Define $I = \{i \in [k] \mid a_i > 0\}$ and $\bar{I} = [k] \backslash I = \{j \in [k] \mid a_j = 0\}$. Set $\mathbb{A}^T = [A^T, A'^T]$. Then, after an appropriate rescaling of the rows of A and A',*

$$\begin{aligned} P_A{}^\circ &= \operatorname{conv}\left(\mathbb{A}_I^T \cup \{0\}\right) + \operatorname{cone}\left(\mathbb{A}_{\bar{I}}^T\right) \\ \textit{and} \;\; \pi(P_A)^\circ &= \pi\left(P_A{}^\circ \cap \left(\mathbb{R}^d \times \{0\}\right)\right) \end{aligned}$$

Furthermore, $P_A{}^\circ$ is bounded if and only if $a > 0$.

The decomposition of a polar polyhedron can be seen as the sum of a linear image of the $|I|$-dimensional simplex $\Delta^{|I|} \subset \mathbb{R}^{|I|+1}$ and a linear image of the cone $\mathbb{R}_+^{|[n] \backslash I|}$.

2.3 Spectrahedra and Semidefinite Programming

We state fundamental results on the theory of spectrahedra and sketch the basic concepts of semidefinite programming. As general references we refer to the books of de Klerk [dK02] and Tunçel [Tun10].

For real symmetric matrices $A_0, \ldots, A_d \in \mathcal{S}^k$, an (affine) linear matrix polynomial

$$A(x) = A_0 + \sum_{p=1}^d x_p A_p \in \mathcal{S}^k[x] \tag{2.3.1}$$

is called a *linear (matrix) pencil*. The positivity domain of $A(x)$ is defined as the set of all points in $\mathbb{R}^d$ for which $A(x)$ is positive semidefinite, i.e.,

$$S_A = \left\{x \in \mathbb{R}^d \mid A(x) \succeq 0\right\},$$

where $A(x) \succeq 0$ denotes positive semidefiniteness of the linear pencil evaluated in the point $x \in \mathbb{R}^d$. A linear matrix pencil is *strictly feasible* if it evaluates positive definite at some point, i.e., $A(x) \succ 0$ for some $x \in \mathbb{R}^d$. Note that strict feasibility implies full-dimensionality of S_A, but the converse is not true in general.

The closed, convex and basic closed semialgebraic set S_A is called a *spectrahedron*. Following the notation for bounded polyhedra, we call a bounded spectrahedron a *spectratope*. Sometimes, we are interested in the homogeneous case, i.e.,

$$A(x_0, x) = x_0 A_0 + \sum_{p=1}^{d} x_p A_p \in \mathcal{S}^k[x_0, x].$$

The solution set $\{(x_0, x) \in \mathbb{R}^{d+1} \mid A(x_0, x) \succeq 0\}$ of the homogeneous pencil is a closed convex cone.

Back to the late 1970s, semidefinite programming entered the field of combinatorial optimization through Lovász' semidefinite relaxation on the Shannon capacity of a graph; see [Lov79]. Since then it has been studied as a relaxation technique for many computationally hard combinatorial problems as, e.g., the MAX-CUT problem [GW95]. In the last years, there has been strong interest in understanding the geometry of spectrahedra and their projections (see, e.g., [Bar12, GN11, HN10]), particularly driven by their intrinsic relevance in polynomial optimization [Ble12, GPT10] and convex algebraic geometry [HN12, HV07]. See also Section 2.4.

Every polyhedron $P_A = \{x \in \mathbb{R}^d \mid a + Ax \geq 0\}$ has a natural representation as a spectrahedron called the *normal form* of the polyhedron P_A as a spectrahedron,

$$P_A = \left\{ x \in \mathbb{R}^d \mid A(x) = \bigoplus_{i=1}^{k} a_i(x) = \begin{bmatrix} a_1(x) & & 0 \\ & \ddots & \\ 0 & & a_k(x) \end{bmatrix} \succeq 0 \right\}, \tag{2.3.2}$$

where $a_i(x) = (a + Ax)_i$ for $i \in [k]$. However, the converse is not true, i.e., there exists nondiagonal linear matrix pencils describing polyhedra. Ramana [Ram97b] showed that the *Polyhedrality Recognition Problem* (PRP) for spectrahedra is NP-hard. Extending Ramana's ideas, Bhardwaj, Rostalski, and Sanyal [BRS11] reduced PRP to an $\mathcal{H}$-in-$\mathcal{S}$ containment problem (cf. Section 5.2.2).

Proposition 2.3.1 ([BRS11, Corollary 2.4])**.** *Let $S_A = \{x \in \mathbb{R}^d \mid A(x) \succeq 0\}$ be a spectrahedron with the origin in its interior. S_A is a polyhedron if and only if there exists $M \in \mathrm{GL}_k(\mathbb{R})$ such that*

$$MA(x)M^T = D(x) \oplus Q(x)$$

with a diagonal linear matrix pencil $D(x)$ and $S_A = S_D$.

We call a non-diagonal linear pencil describing a polyhedron an *$\mathcal{S}$-representation* of the polyhedron, or $\mathcal{S}$-polyhedron for short.

A spectrahedron $S_A = \{x \in \mathbb{R}^d \mid A(x) \succeq 0\}$ contains the origin if and only if A_0 is positive semidefinite. Since the class of spectrahedra is closed under translation, this can always be achieved (assuming that S_A is nonempty). Indeed, there exists a point $x' \in \mathbb{R}^d$ such that $A(x') \succeq 0$ if and only if the origin is contained in the set $\{x \in \mathbb{R}^d \mid A(x + x') \succeq 0\}$. In particular, the constant term in the linear pencil $A'(x) = A(x + x')$ is positive semidefinite.

Contrary to the polyhedral situation, positive definiteness of A_0 is only sufficient for the origin being an interior point. However, S_A contains the origin in its interior if and only if there is a linear pencil $A'(x)$ with the same positivity domain such that $A'_0 = I_k$; see [HV07]. We refer to such a pencil as a *monic* linear pencil.

More generally, the interior of S_A does not have to coincide with the positive definiteness region of the pencil (but the interior contains the latter set).

Proposition 2.3.2 ([GR95, Corollary 5])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil and let $N = \bigcap_{i=0}^d \ker(A_i)$ be the intersection of the coefficient nullspaces. Define $A'(x) = V^T A(x) V$, where V is a basis of the orthogonal complement $N^\perp$. If S_A is full-dimensional, then $S_A = S_{A'}$ and* $\operatorname{int} S_A = \operatorname{int} S_{A'} = \{x \in \mathbb{R}^d \mid A'(x) = V^T A(x) V \succ 0\}$.

We call a linear pencil with the property $\operatorname{int} S_A = \{x \in \mathbb{R}^d \mid A(x) \succ 0\}$ a *reduced* linear pencil.

We occasionally assume the matrices $A_1, \ldots, A_d$ to be linearly independent. This assumption is not too restrictive. In order to see this, denote by $\tilde{A}(x) = A(x) - A_0$ the pure-linear part of the linear pencil $A(x)$. Recall the well-known fact that the *lineality space* L_A of a spectrahedron S_A, i.e., the largest linear subspace contained in S_A, is the set $L_A = \{x \in \mathbb{R}^d \mid \tilde{A}(x) = 0\}$; see [GR95, Lemma 3]. Obviously, if the coefficient matrices $A_1, \ldots, A_d$ are linearly independent, then the lineality space is zero-dimensional, i.e., $L_A = \{0\}$. In particular, this is the case whenever the spectrahedron S_A is bounded (and $A_0 \succeq 0$); see Lemma 2.3.6. Conversely, if there are linear dependencies in the coefficient matrices, then we can simply reduce the containment problem to lower dimensions.

Proposition 2.3.3. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils such that S_A is nonempty.*

(1) *$L_A = \{0\}$ if and only if $A_1, \ldots, A_d$ are linearly independent.*

(2) *If $S_A \subseteq S_B$, then $L_A \subseteq L_B$.*

(3) *If $L_A \subseteq L_B$, then $S_A \subseteq S_B$ holds if and only if $S_{A'} \subseteq S_{B'}$ holds, where $S_{A'} = S_A \cap L_A^\perp$ and $S_{B'} = S_B \cap L_A^\perp$.*

To prove the proposition, we need a result concerning the lineality space of a closed convex set.

Lemma 2.3.4 ([Web94, Theorem 2.5.8])**.** *Let S be a nonempty closed convex set in $\mathbb{R}^d$ with lineality space L. Then $S = L + (S \cap L^\perp)$ and the convex set $S \cap L^\perp$ contains no lines.*

Proof of Proposition 2.3.3.

To (1): This follows directly from $L_A = \{x \in \mathbb{R}^d \mid \tilde{A}(x) = 0\}$ and the definition of linear independence.

To (2): If $L_A = \{0\}$, then $L_A \subseteq L_B$ is obviously true. Therefore, assume $L_A \neq \{0\}$. Let $\bar{x} \in S_A \subseteq S_B$ and $0 \neq x \in L_A$. As above, denote by $\tilde{B}(x) = B(x) - B_0 = \sum_{p=1}^d x_p B_p$ the pure-linear part of $B(x)$. Then $A(\bar{x} + tx) \succeq 0$ for all $t \in \mathbb{R}$ and hence $B(\bar{x}) \pm t\tilde{B}(x) = B(\bar{x} \pm tx) \succeq 0$ for all $t \in \mathbb{R}$. Consequently, $\pm\tilde{B}(x) \succeq 0$, i.e., $\tilde{B}(x) = 0$. Thus the linear subspace $\operatorname{span} x$ is contained in L_B. Since $0 \neq x \in L_A$ was arbitrary and L_B is a linear subspace, we have $L_A \subseteq L_B$.

To (3): Assume first $S_A \subseteq S_B$ holds. Then $S_{A'} = S_A \cap L_A^\perp \subseteq S_B \cap L_A^\perp = S_{B'}$. For the converse, note that $S_A = L_A + S_{A'}$. Let $x \in S_A$. Then $x = x_1 + x_2$ with $x_1 \in L_A$ and $x_2 \in S_{A'}$. Since $x_1 \in L_A \subseteq L_B$ and $x_2 \in S_{A'} \subseteq S_{B'} \subseteq S_B$, we have $x \in L_B + S_B = S_B$. □

If $A_0 \in \operatorname{span}\{A_1, \ldots, A_d\}$, then S_A is an affine cone. A cone is called pointed if the lineality space is zero-dimensional.

Proposition 2.3.5. *Let $A(x) \in \mathcal{S}^k[x]$. S_A is an affine cone if and only if $A_0 = \sum_{i\in[d]} \lambda_i A_i$ for some $\lambda \in \mathbb{R}^d$, i.e., $A_0 \in \operatorname{span}\{A_1, \dots, A_d\}$. In this case, S_A equals the shifted recession cone $S_A = R_A - \lambda$. Moreover, S_A is pointed if and only if $A_1, \dots, A_d$ are linearly independent.*

Proof. By Ramana [GR95, Lemma 3], the recession cone R_A of S_A has the form $R_A = \{x \in \mathbb{R}^d \mid \tilde{A}(x) \succeq 0\}$. Let $x \in S_A$. Then $0 \preceq A(1,x) = \sum_{i\in[d]} \lambda_i A_i + \sum_{i\in[d]} x_i A_i = \sum_{i\in[d]} (x_i + \lambda_i) A_i$, implying $x + \lambda \in R_A$. Conversely, let $x \in R_A$. Then $A(1, x - \lambda) = \sum_{i\in[d]} \lambda_i A_i + \sum_{i\in[d]} (x_i - \lambda_i) A_i = \sum_{i\in[d]} x_i A_i \succeq 0$, implying $x - \lambda \in S_A$. Therefore $S_A = R_A - \lambda$. By [GR95, Corollary 6], S_A is (linear) conical if and only if $S_A = R_A$. Hence S_A is affine conical if and only if $S_A = R_A - \lambda$. The statement concerning pointedness follows from Proposition 2.3.3. □

As opposed to the polyhedral case, deciding boundedness of a spectrahedron is somewhat tricky. It can be formulated as a semidefinite feasibility problem (2.3.13), whose complexity is not fully classified. Linear independence of $A_0, A_1, \dots, A_d$ is a necessary condition for boundedness of the spectrahedron S_A.

Lemma 2.3.6 ([HKM13, Proposition 2.6])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil. If $A_0 \succeq 0$ is nonzero and the spectrahedron S_A is bounded, then the coefficient matrices $A_0, \dots, A_d$ are linearly independent.*

Proof. First note that because of $A_0 \succeq 0$, we have $0 \in S_A \neq \emptyset$.

Let $0 \neq (x_0, x) \in \mathbb{R}^{k+1}$ with $x_0 A_0 + \dots + x_d A_d = 0$. Then $x \neq 0$ (since $x = 0$ and $A_0 \neq 0$ implies $x_0 = 0$). If $x_0 = 0$, then $A_0 + \sum_{i=1}^d t x_i A_i = A_0 \succeq 0$ for all $t \in \mathbb{R}$. Thus S_A is unbounded. Otherwise, $x_0 \neq 0$, we have $A_0 + \sum_{i=1}^d (1/x_0) x_i A_i = 0$. Therefore $\sum_{i=1}^d (-t/x_0) x_i A_i = t A_0 \succeq 0$ for all $t > 0$, and hence $A_0 + \sum_{i=1}^d (-t/x_0) x_i A_i \succeq 0$ for all $t > 0$ which again implies unboundedness of S_A. □

The converse of the above statement is not true. Too see this, assume $A_0, A_1, \dots, A_d$ to be linearly independent and positive semidefinite. (Note that this is possible since the cone of positive semidefinite matrices is full-dimensional in the space of symmetric matrices.)

We state analogs of Farkas' Lemma for the spectrahedral situation.

Lemma 2.3.7 (Farkas' Lemma for cones I [GM12, Lemma 4.5.6])**.** *Let $\tilde{A}(x) = \sum_{i=1}^d x_i A_i \in \mathcal{S}^k[x]$ be a pure-linear pencil and let $c \in \mathbb{R}^d$. Then exactly one of the following two systems has a solution.*

$$\exists (Z_j)_j \in \mathcal{S}_+^k : \lim_{j\to\infty} \langle A_i, Z_j \rangle = c_i \ \forall i \in [d] \tag{2.3.3}$$

$$\exists x \in \mathbb{R}^d : \ \tilde{A}(x) \in \mathcal{S}_+^k, \ \langle x, c \rangle < 0 \tag{2.3.4}$$

Lemma 2.3.8 (Farkas' Lemma for cones II [Tun10, Theorem 2.22])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil and denote by $\tilde{A}(x) = \sum_{i=1}^d x_i A_i$ the pure-linear part. Then exactly one of the following two systems has a solution.*

$$\forall \varepsilon > 0 \ \exists A_0' \in \mathcal{S}^k, \ \exists x \in \mathbb{R}^d : \ \|A_0 - A_0'\| < \varepsilon, \ A_0' + \tilde{A}(x) \in \mathcal{S}_+^k \tag{2.3.5}$$

$$\exists Z \in \mathcal{S}^k : \ Z \succeq 0, \ \langle A_i, Z \rangle = 0 \ \forall i \in [d], \ \langle A_0, Z \rangle < 0 \tag{2.3.6}$$

The Farkas type lemmas (and thus the theory of semidefinite programming, see below) lack in the fact that the linear image of the cone of positive semidefinite matrices is not closed in general (cf. Lemma 2.3.9). Additional conditions which leads to more clean formulations are called *constraint qualification*.

Lemma 2.3.9 ([BPT13, Theorem 2.28]). *Let $A(x) \in \mathcal{S}^k[x]$ be a strictly feasible pure-linear pencil. Then the set $\{(\langle A_i, Z\rangle)_{i=1}^d \mid Z \in \mathcal{S}_+^k\}$ is closed and thus either* (2.3.4) *holds or the set $\{Z \in \mathcal{S}_+^k \mid \langle A_i, Z\rangle = c_i,\ i \in [d]\}$ is nonempty.*

Lemma 2.3.10 ([BV04, Example 5.14]). *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil. Assume the condition*

$$\sum_{i=1}^d x_i A_i \succeq 0 \Longrightarrow \sum_{i=1}^d x_i A_i = 0$$

holds for any x. Then either (2.3.6) *has a solution or S_A is nonempty. If $A_1, \ldots, A_d$ are linearly independent, then the above condition can be replaced by the condition*

$$\sum_{i=1}^d x_i A_i \succeq 0 \Longrightarrow x = 0.$$

In view of Propositions 2.3.3 and 2.3.5, the constraint qualification stated in Lemma 2.3.10 says that the recession cone R_A of S_A coincides with the lineality space L_A of S_A or, in the case of linearly independent $A_1, \ldots, A_d$, $R_A = L_A = \{0\}$. We use this constraint qualification in Section 6.2.

In the late 1990s, Ramana stated a new Farkas' type lemma for spectrahedra which closes the gap in the straightforward version of Farkas' Lemma for cones. Unfortunately, Ramana introduces a high amount of additional variables and constraints. It turns out that this is exactly the face reduction method studied by Borwein and Wolkowicz; see [RTW97].

Lemma 2.3.11 (Ramana's Lemma [RTW97, Section 4.1]; see also [Ram97a]). *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil. Then exactly one of the following two systems has a solution.*

$$\exists x \in \mathbb{R}^d :\ A(x) \in \mathcal{S}_+^k \tag{2.3.7}$$

$$\exists W \in \mathcal{W}_t,\ U \in \mathcal{S}_+^k :\ \langle A_i, U + W\rangle = 0\ \forall i \in [d],\ \langle A_0, U + W\rangle < 0, \tag{2.3.8}$$

where $\mathcal{W}_t$ is defined as

$$\begin{aligned}\mathcal{W}_t = \Big\{W_t + W_t^T \in \mathcal{S}^k \mid \exists (U_1, W_1; \ldots; U_t, W_t) \in \mathcal{S}^k \oplus \mathbb{R}^{k\times k} \oplus \cdots \oplus \mathcal{S}^k \oplus \mathbb{R}^{k\times k} :\\ W_0 = 0,\ \Big\langle A_j, U_i + W_{i-1} + W_{i-1}^T \Big\rangle = 0\ \forall j \in [d],\\ \Big\langle A_0, U_i + W_{i-1} + W_{i-1}^T \Big\rangle = 0,\ U_i \succeq W_i W_i^T,\ \forall i \in [t]\Big\},\end{aligned}$$

for some $t \geq 1$.

Since in Lemma 2.3.11 the first three conditions are linear and the last condition can be rephrased as

$$\begin{bmatrix} I & W_i^T \\ W_i & U_i \end{bmatrix} \succeq 0,$$

the set of all $(U_1, W_1; \ldots; U_t, W_t)$ is closed and convex. Moreover, $(\mathcal{W}_t)_{1\leq t\leq d}$ defines an increasing chain of linear subspaces, i.e., $\mathcal{W}_1 \subseteq \cdots \subseteq \mathcal{W}_d \subseteq \mathcal{S}^k$. The proof of these facts is not hard but sophisticated; see [Ram97a]. By a dimension counting argument, the number of iterations t in Lemma 2.3.11 is bounded from above by the minimum of the number of variables d and the size of the linear pencil k, i.e., $t \leq \min\{d, k\}$. As mentioned by Sturm [Stu00], the so-called degree of singularity ds(A(x)) equals the minimum number of steps t needed in Ramana's Lemma. Computing ds(A(x)) is itself a hard problem.

Lemma 2.3.12 (Ramana's Lemma - Dual [RTW97, Section 4.2])**.** *Let $\tilde{A}(x) = \sum_{i=1}^{d} x_i A_i \in \mathcal{S}^k[x]$ be a pure-linear pencil and let $c \in \mathbb{R}^d$. Then exactly one of the following two systems has a solution.*

$$\exists Z \in \mathcal{S}_+^k : \ \langle A_i, Z\rangle = c_i \ \forall i \in [d] \tag{2.3.9}$$

$$\exists Z \in \mathcal{Z}_t, \ x \in \mathbb{R}^d : \ Z + \tilde{A}(x) \succeq 0, \ \langle c, x\rangle < 0, \tag{2.3.10}$$

where $\mathcal{Z}_t$ is defined as

$$\mathcal{Z}_t = \Big\{ Z_t + Z_t^T \in \mathcal{S}^k \mid \exists (x_1, Z_1; \ldots; x_t, Z_t) \in \mathbb{R}^d \oplus \mathbb{R}^{k\times k} \oplus \cdots \oplus \mathbb{R}^d \oplus \mathbb{R}^{k\times k} : \ Z_0 = 0,$$
$$\tilde{A}(x_i) + Z_{i-1} + Z_{i-1}^T \succeq 0, \ \tilde{A}(x_i) \succeq Z_i Z_i^T, \ x_i^T c = 0 \ \forall i \in [t] \Big\},$$

for $t \geq 1$.

Recently, Liu and Pataki consider a new approach based on "elementary reformulations"; see [LP14]. For a sum of squares formulation of semidefinite programming see Klep and Schweighofer [KS13].

From Farkas' Lemma for cones we can deduce a description of the polar of a spectrahedron. This goes back to Goldman and Ramana [GR95]. Define the set $Q_A = \{Z \in \mathcal{S}_+^k \mid \langle A_0, Z\rangle \leq 1\}$. The image of Q_A under the linear map $(\langle A_i, \cdot\rangle)_i : \ \mathcal{S}^k \to \mathbb{R}^d$ is called the *algebraic polar* of S_A, denoted by

$$S_A^* = \Big\{ x \in \mathbb{R}^d \mid \exists Z \in \mathcal{S}_+^k : \ \langle A_0, Z\rangle \leq 1 \text{ and } x_i = -\langle A_i, Z\rangle \ \ \forall i \in [d] \Big\}.$$

It always contains the origin but is not closed in general; see [GR95, Section 3]. It can be shown that the closure of the algebraic polar is equal to the (geometric) polar, ${S_A}^\circ = \operatorname{cl} S_A^*$, provided that $A_0 \in \mathcal{S}_+^k$. We are now able to state the desired description of the polar spectrahedron.

Proposition 2.3.13. *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil with $A_0 \in \mathcal{S}_+^k$. Then*

$$S_A^\circ = \operatorname{cl} S_A^* = \operatorname{cl} \Big\{ v \in \mathbb{R}^d \mid \exists Z \in \mathcal{S}_+^k : \ v_i = -\langle A_i, Z\rangle \ \ \forall i \in [d], \ \langle A_0, Z\rangle \leq 1 \Big\}.$$

Furthermore, if $A(x)$ is strictly feasible, then $S_A^\circ = S_A^*$.

The condition $A_0 \in \mathcal{S}_+^k$ ensures that the origin is contained in the spectrahedron S_A. If this is not the case, one can translate the set by any point of S_A (provided the set is nonempty). Note that this is a semidefinite feasibility problem (2.3.13).

Using Ramana's Lemma for spectrahedra 2.3.11, we can describe the polar of the spectrahedron more explicit in the sense that the closure can be omitted.

Corollary 2.3.14 ([Ram97a, Theorem 2])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil with $A_0 \in \mathcal{S}_+^k$. The polar of S_A has the form*

$${S_A}^\circ = \Big\{ v \in \mathbb{R}^d \mid \exists W \in \mathcal{W}_t, \ Z \in \mathcal{S}_+^k : \ v_i = -\langle A_i, U + W\rangle \ \ \forall i \in [d], \ \langle A_0, U + W\rangle \leq 1 \Big\}$$

for all $t \geq d - 1$.

Note that computing the subspaces $\mathcal{W}_t$ for $t \geq d - 1$ is computationally expensive.

Semidefinite Programming. A (linear) *semidefinite program* (SDP) is an optimization problem of maximizing a linear objective function over a spectrahedron. As in the situation of linear programming a point in the spectrahedron is called *feasible solution* and a feasible point in which the objective attains its optimum value is referred to as an *optimal solution.*

Given a spectrahedron $S_A = \{x \in \mathbb{R}^d \mid A(x) \succeq 0\}$ and a vector $c \in \mathbb{R}^d$, the optimization problem

$$\begin{aligned} \sup \ & c^T x \\ \text{s.t.} \ & x \in S_A \end{aligned} \tag{2.3.11}$$

is called the *primal* SDP. The associated *dual* problem is

$$\begin{aligned} \inf \ & \langle A_0, Y\rangle \\ \text{s.t.} \ & \langle A_i, Y\rangle = c_i, \quad i \in [d] \\ & Y \in \mathcal{S}^k_+. \end{aligned} \tag{2.3.12}$$

The set of dual feasible solutions can easily be rewritten in terms of a linear pencil and thus is a spectrahedron.

As for linear programming, *weak duality* holds, i.e., $c^T x \leq \langle A_0, Y\rangle$ for all pairs (x, Y) of feasible solutions. Strong duality, however, does not hold in this generality. It can be achieved under the assumption of strict feasibility. Recall that for every strictly feasible linear pencil the spectrahedron is full-dimensional but the converse has not to be true.

Proposition 2.3.15 (Strong Duality in Semidefinite Programming [dK02, Theorem 2.2])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil and let $c \in \mathbb{R}^d$ be fixed.*

(1) *Assume $A(x) \in \mathcal{S}^k[x]$ is strictly feasible and the primal SDP* (2.3.11) *has a finite optimal value. Then*

$$\sup \left\{ c^T x \mid x \in S_A \right\} = \min \left\{ \langle A_0, Y\rangle \ \mid Y \succeq 0, \ \langle A_i, Y\rangle = c_i, \ i \in [d] \right\}.$$

In particular, there exists a dual optimal solution.

(2) *Assume the dual problem is strictly feasible, i.e., set of dual feasible solutions contains a positive definite matrix, and the optimal value of the dual problem* (2.3.12) *is finite. Then*

$$\max \left\{ c^T x \mid x \in S_A \right\} = \inf \left\{ \langle A_0, Y\rangle \ \mid Y \succeq 0, \ \langle A_i, Y\rangle = c_i, \ i \in [d] \right\}.$$

In particular, there exists a primal optimal solution.

In theory, SDPs with rational input data are solvable in polynomial time (up to a given $\varepsilon > 0$) by the ellipsoid method. If both the primal and the dual SDP are strictly feasible, the optimal value of the semidefinite programs can be approximated up to a given additive error ε in polynomial time using an interior point method [GM12, Corollary 6.5.3]; see [dK02] for a more detailed treatment.

A *Semidefinite Feasibility Problem* (SDFP) is defined as the following decision problem (see, e.g., [KP97, Ram97a]).

$$\begin{aligned} &\text{Given } d, k \in \mathbb{N} \text{ and rational symmetric } k \times k\text{-matrices } A_0, \ldots, A_d, \\ &\text{decide whether there exist } x \in \mathbb{R}^d \text{ such that } A(x) \succeq 0. \end{aligned} \tag{2.3.13}$$

Equivalently, one can ask whether the spectrahedron S_A is nonempty. Although checking positive semidefiniteness can be done in polynomial time by computing a Cholesky factorization, the complexity classification of the problem SDFP is one of the major open complexity questions related to semidefinite programming (see [dK02, Ram97a]). Using semidefinite programming techniques, a SDFP can be solved efficiently in practice. In our model of computation, the binary Turing machine (cf. Chapter 3), the following is known. (In the BSS-model, SDFP lies in the intersection of NP and co-NP; see [Ram97a, Theorem 25]).

Proposition 2.3.16 ([Ram97a, Theorem 25] and [KP97, Theorem 7])**.** *SDFP is in NP if and only if it is in co-NP. Moreover, unless NP = co-NP, SDFP is neither NP-complete nor co-NP-complete.*

If the number of variables d or the matrix size k is fixed, then SDFP is solvable in polynomial time.

As seen by a standard example in semidefinite programming (see, e.g., [Ali93, GR95]), there exists a spectrahedron whose elements have a coordinate of double-exponential size in the number of variables and hence double-exponential distance to the origin in the number of variables. Therefore we cannot in general expect to attain a certificate for feasibility (or boundedness) of a spectrahedron in the sense of an explicit point $x \in \mathbb{R}^d$ that is polynomial in the input size.

Using different solvers, there may be significant differences in the time need for computation. For example, while for small problems the SDP-solver SDPT3 and SeDuMi have a similar time need, the latter solver is significantly worse in higher dimensions. Mittelmann gives an overview on "The-State-of-the-Art in Conic Optimization Software" in [Mit12]. On his homepage, he states benchmarks for the most common SDP solvers; see [Mit]. As one can see there, at present, the free solver SDPT3 [TTT99] and the commercial solver Mosek [AA00] (free academic licenses available) are the best performing SDP solvers.

Due to Mittelmann's benchmarks and the experiences of the author, Mosek seems to outperforming SDPT3 for most semidefinite programming problems coming from polynomial optimization problems (see Section 2.4). Thus Mosek is the solver of choice for all examples stated in this thesis.

Examples of Spectrahedra. We fix the notation of some well-known spectratopes, which we use in later examples.

An *ellipsoid* with the origin in its interior can be written as the spectratope S_A given by the linear pencil

$$A(x) = A_0 \oplus r + \sum_{p=1}^{d} x_p (E_{p,d+1} + E_{d+1,p}) \in \mathcal{S}^{d+1}[x]$$

with $A_0 \in \mathcal{S}^d_{++}$ and $r > 0$. If the ellipsoid has axis-aligned semiaxes, then it has the form

$$A(x) = I_{d+1} + \sum_{p=1}^{d} \frac{x_p}{a_p} (E_{p,d+1} + E_{d+1,p}) \in \mathcal{S}^{d+1}[x] \qquad (2.3.14)$$

with $(a_1, \dots, a_d) > 0$. We call (2.3.14) the *normal form of the ellipsoid.* Specifically, for the case $a_1 = \dots = a_d =: r$ this gives the *normal form of a ball* with radius r centered at the origin.

The *d-elliptope* is defined by the monic linear pencil

$$A(x) = I_k + \sum_{1 \le i < j \le k} x_{ij} (E_{i,j} + E_{j,i}) \in \mathcal{S}^k[x], \qquad (2.3.15)$$

where $d = k(k-1)/2$ is the dimension of the spectratope. It appears in many applications of semidefinite programming to combinatorial problems as, e.g., the MAX-CUT problem [GW95].

The spectratope (of dual form)

$$\mathbb{T}^k = \left\{ X \in \mathcal{S}^k \mid X \succeq 0,\ \langle I_k, X \rangle = 1 \right\} \tag{2.3.16}$$

is called the *k-spectraplex* (of dimension $k(k+1)/2$) and can be seen as a spectrahedral analog of the simplex (2.2.5).

Projections of Spectrahedra. Given a linear pencil $A(x,y) \in \mathcal{S}^k[x,y]$ with $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_m)$ for some nonnegative integer m, a projection of the spectrahedron S_A is its image under an affine map. By an elementary observation, without loss of generality, we can assume that the affine projection is a coordinate projection.

Proposition 2.3.17 ([GN11, Section 2])**.** *If a set $T \subseteq \mathbb{R}^d$ is the image of a spectrahedron S under an affine map, then there exists a linear pencil $A(x,y) \in \mathcal{S}^k[x,y]$ with $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_m)$ for some nonnegative integer m such that T is a coordinate projection of $S_A \subseteq \mathbb{R}^{d+m}$. Furthermore, if T and S have nonempty interior, then this can be assumed for S_A too.*

Due to the proposition, we always assume that the *projection of spectrahedron S_A* as given by the linear pencil $A(x,y) \in \mathcal{S}^k[x,y]$ with $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_m)$, $m \geq 0$, is the set

$$\pi(S_A) = \left\{ x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m :\ A(x,y) \succeq 0 \right\}, \tag{2.3.17}$$

where $\pi :\ \mathbb{R}^{d+m} \to \mathbb{R}^m$ denotes the coordinate projection. Note that $\pi(S_A) = S_A$ holds for $m = 0$. Hence, if we do not state the converse, all statements for projections of spectrahedra hold for spectrahedra.

While projections of polyhedra are again polyhedral, this is not true for spectrahedra (see, e.g., [BPT13, Section 6.3.1]). Moreover, whereas spectrahedra are basic closed semialgebraic sets (the semialgebraic constraints are given by the nonnegativity condition on the principal minors), projected spectrahedra are generally not. Though they are semialgebraic, they are not closed in general; see Example 6.2.2.

Lemma 2.3.18. *Let $A(x,y) \in \mathcal{S}^k[x,y]$ be a linear pencil.*

(1) $S_A \neq \emptyset \iff \pi(S_A) \neq \emptyset$.

(2) *If S_A is bounded, then $\pi(S_A)$ is bounded.*

Proof. The first part of the statement is obvious.

Let $\pi(S_A)$ be unbounded, i.e., there exists $\bar{x} \in \mathbb{R}^d$ such that $x + t\bar{x} \in \pi(S_A)$ for all $x \in \pi(S_A)$ and $t > 0$. By definition, for all $t > 0$ there exists $y \in \mathbb{R}^m$ such that $A(x+t\bar{x}, y) \succeq 0$, implying the unboundedness of S_A. □

The converse of part (2) in the previous lemma is not true in general.

Example 2.3.19. *Consider the linear pencil*

$$A(x,y) = \begin{bmatrix} 1-x & 0 & 0 \\ 0 & 1+x & 0 \\ 0 & 0 & y \end{bmatrix}.$$

For all $t \in \mathbb{R}_+$ *is* $(x, y) = (0, t) \in S_A$*, i.e.,* S_A *is unbounded, but its projection* $\pi(S_A) = [-1, 1]$ *is bounded.* □

2.4 Polynomial Optimization and Positivstellensätze

Given a set of polynomials $G \subset \mathbb{R}[x]$ in the variables $x = (x_1, \dots, x_d)$, a polynomial optimization problem is to infimize (or supremize) a objective function $f \in \mathbb{R}[x]$ on the semialgebraic set $S_G = \{x \in \mathbb{R}^d \mid g(x) \geq 0 \ \forall g \in G\}$. Most times we only consider finite sets G.

In the subsequent sections we introduce well-known relaxation techniques to polynomial optimization problems based on so-called Positivstellensätze. We refer to the books of Blekherman et. al. [BPT13], Lasserre [Las10] and Marshall [Mar08].

2.4.1 Handelman's Positivstellensatz

For a polytope $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$, define the cone

$$H(P) = \left\{ p \in \mathbb{R}[x] \mid p = \sum_{\alpha \in \mathbb{N}^k} c_\alpha \prod_{j=1}^{k} (a - Ax)_j^{\alpha_j}, \ c_\alpha \in \mathbb{R}_+ \right\} \tag{2.4.1}$$

in the ring $\mathbb{R}[x]$ of real polynomials in $x = (x_1, \dots, x_d)$. Here we assume implicitly that only finitely many scalar multipliers c_α are nonzero.

Proposition 2.4.1 (Handelman's Positivstellensatz [Han88])**.** *Given a nonempty polytope* $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$*, the cone* $H(P)$ *contains every polynomial* $f \in \mathbb{R}[x]$ *positive on* P*.*

Handelman's original statement deals with subfields of the real numbers. This, in particular, implies that whenever $a \in \mathbb{Q}^k$, $A \in \mathbb{Q}^{k \times d}$, and $f \in \mathbb{Q}[x]$ are rational, then there is a (rational) certificate for membership of f in the cone $H(P) \subseteq \mathbb{Q}[x]$ and thus for positivity of f on P.

We denote by $H_t(P)$ the truncated cone

$$H_t(P) = \left\{ p \in \mathbb{R}[x] \mid p = \sum_{\alpha \in \mathbb{N}^k, |\alpha| \leq t} c_\alpha \prod_{j=1}^{k} (a - Ax)_j^{\alpha_j}, \ c_\alpha \in \mathbb{R}_+ \right\},$$

where $|\alpha| := \alpha_1 + \cdots + \alpha_k$. Observe that membership of a polynomial to some fixed $H_t(P)$ can be decided by comparing coefficients, which is a linear programming problem. For the maximization problem $\max\{f(x) \mid x \in P\}$, this leads to a hierarchy of linear programming relaxations

$$\max \{f(x) \mid x \in P\} = \inf \{\nu \mid \nu - f \in H(P)\} \leq \inf \{\nu \mid \nu - f \in H_t(P)\}.$$

Since in the latter program all information is encoded in the coefficients of the polynomials involved, it reduces to a (finite dimensional) linear program. Clearly, it does not make sense to consider a relaxation order t less than the degree of f. Thus we always assume $t \geq \deg(f)$ and call $t = \deg(f)$ the *initial relaxation order*.

There are some advantages and disadvantages of the LP-approach based on Handelman's Theorem (as compared, e.g., with the semidefinite approach in the subsequent section). Since, both in theoretical issues (such as exactness of duality theory) and in practical speed, linear programming has advantages compared to semidefinite programming, this makes the Handelman approach attractive.

On the other hand, if for one of the global maximizers the set of active constraints is empty, i.e., a global maximizer lies in the interior of the polytope, then no Handelman representation exists; see [Las02, Theorem 3.1]. Since in our specific problem (cf. Chapter 4) all maximizers are part of the boundary, this obstacle does not occur. However, because of the large binomial coefficients involved, the LP relaxation is ill conditioned in general. In practice, setting up the problem, i.e., extracting the linear program from the input, takes a lot more time than actually solving the resulting LP.

The Handelman relaxation can also be applied to general semialgebraic sets as in the subsequent section. We do not need this here and refer to Lasserre's book [Las10].

2.4.2 Putinar's Positivstellensatz

Consider a set of polynomials $G = \{g_1, \ldots, g_k\} \subset \mathbb{R}[x]$ in the variables $x = (x_1, \ldots, x_d)$. The quadratic module generated by G is defined as

$$\mathrm{QM}(G) = \left\{ \sigma_0 + \sum_{i=1}^{k} \sigma_i g_i \mid \sigma_i \in \Sigma[x], i \in [k] \right\}, \tag{2.4.2}$$

where $\Sigma[x] \subset \mathbb{R}[x]$ is the set of sum of squares polynomials. A polynomial $p \in \mathbb{R}[x]$ is called *sum of squares* (sos) if it can be written in the form $p = \sum_i h_i(x)^2$ for some $h_i \in \mathbb{R}[x]$. Equivalently, p has the form $[x]_t^T Q [x]_t$, where $[x]_t$ is the vector of all monomials in x up to half the degree of p, $\deg(p) = 2t$, and Q is a positive semidefinite matrix of appropriate size. Checking whether a polynomial is sos is an SDFP (2.3.13). We denote by $\Sigma_t[x]$ the set of sos polynomials of degree at most $2t$.

If a polynomial $f(x) \in \mathbb{R}[x]$ lies in $\mathrm{QM}(G)$, then we say f has a *sum of squares* decomposition in $g_1, \ldots, g_k$. Obviously, every element in $\mathrm{QM}(G)$ is nonnegative on the semialgebraic set $S_G = \{x \in \mathbb{R}^d \mid g(x) \geq 0 \ \forall g \in G\}$. Putinar [Put93] showed that the converse is true under some regularity assumption.

Definition 2.4.2. *A quadratic module* QM *is called* Archimedean *if one of the following equivalent conditions holds.*

(1) *There is a polynomial* $p(x) \in \mathrm{QM}$ *such that the level set* $\{x \in \mathbb{R}^d \mid p(x) \geq 0\}$ *is compact.*

(2) $n - (x_1^2 + \cdots + x_d^2) \in \mathrm{QM}$ *for some integer* $n \geq 1$.

(3) $n \pm x_i \in \mathrm{QM}$ *for* $i \in [d]$ *and some integer* $n \geq 1$.

For a proof that these conditions are indeed equivalent, we refer to Marshall's book [Mar08].

Proposition 2.4.3 (Putinar's Positivstellensatz [Put93] – see also [Mar08, Theorem 5.6.1])**.** *Let* $S_G = \{x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$ *for some polynomials* $g_1, \ldots, g_k \in \mathbb{R}[x]$. *If the quadratic module* $\mathrm{QM}(G)$ *is Archimedean, then* $\mathrm{QM}(G)$ *contains every polynomial* $f \in \mathbb{R}[x]$ *positive on* S_G.

The Archimedean condition in the proposition is not very restrictive. Since every compact semialgebraic set is contained in a scaled unit ball by the definition of boundedness, Archimedeanness can always be achieved by adding the ball defining polynomial to the constraint set. However, in the general situation, the radius is not known a priori. In the case of linear polynomials g_i, the condition is always fulfilled.

Corollary 2.4.4 ([Mar08, Theorem 7.1.3]). *If the polynomials g_i in Proposition 2.4.3 are linear and $P = S_G$ is a polytope, then the quadratic module* $\mathrm{QM}(P)$ *is Archimedean.*

Proof. Let $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ be a nonempty polytope in $\mathcal{H}$-description. Compute the minimum $c > 0$ such that $P \subseteq \{x \in \mathbb{R}^d \mid c \pm x_i \geq 0,\ i \in [d]\}$. By Farkas' Lemma 2.2.2, every linear polynomial $c \pm x_i$ is a nonnegative combination of the rows in $a - Ax$. Thus $c \pm x_i \in \mathrm{QM}(P)$. □

We illustrate this fact by an example.

Example 2.4.5. *The d-dimensional unit cube is given by the* $2d$ *inequalities* $1 \pm x_i$. *Since*

$$2(1 - x_i^2) = (1 - x_i)^2(1 + x_i) + (1 + x_i)^2(1 - x_i) \in \mathrm{QM}(1 \pm x_i)$$

for $i \in [d]$*, the quadratic module is Archimedean.*

Nie and Schweighofer [NS07] stated an exponential upper bound on the degree of a sos-representation for all polynomials f positive on a certain nonempty semialgebraic set S rescaled to fit in the open unit cube. However, in practice, often a small degree suffices; see [NS07, Theorem 6].

Powers recently showed that if the polynomials $f, g_1, \ldots, g_k$ have only rational coefficients and the polynomial $g_{k+1} = N - x^T x$ is a member of the quadratic module for some positive integer N, then there exists a rational certificate, i.e., $\sigma_i \in \mathbb{Q}[x]$, for f to be positive on S in the terms of $g_1, \ldots, g_k, g_{k+1}$; see [Pow11, Theorem 7].

In order to apply Proposition 2.4.3 to polynomial optimization, consider an optimization problem

$$\sup \{f(x) \mid g_i(x) \geq 0,\ i \in [k]\} \tag{2.4.3}$$

with $f, g_1, \ldots, g_k \in \mathbb{R}[x]$. Clearly, this is the same as to infimize a scalar μ such that $\mu - f(x) \geq 0$ on the set S_G. A natural relaxation of the latter reformulation is to replace the nonnegativity condition by an sos-condition. This is a semi-infinite program since deciding membership can be rephrased as a semi-infinite feasibility problem (i.e., an infinite dimensional semidefinite program). In order to get a (finite-dimensional) semidefinite program, we truncate the quadratic module $\mathrm{QM}(G)$ by considering only monomials up to a certain degree $2t$,

$$\mathrm{QM}_t(G) = \left\{ \sigma_0 + \sum_{i=1}^{k} \sigma_i g_i \mid \sigma_0 \in \Sigma_t[x],\ \sigma_i \in \Sigma_{t - \lceil t_G/2 \rceil}[x] \text{ for } i \in [k] \right\} \subseteq \mathbb{R}[x],$$

where t_G denotes the maximum degree of the polynomials $g_1, \ldots, g_k$. The tth sos-relaxation of the polynomial optimization problem (2.4.3) has the form

$$\mu(t) = \inf \{\mu \mid \mu - f(x) \in \mathrm{QM}_t(G)\}. \tag{2.4.4}$$

Clearly, the sequence of *truncated quadratic modules* is increasing with respect to inclusion as t grows. Thus the sequence of optimal values $\mu(t)$ is monotone decreasing and bounded from below by the optimal value of (2.4.3).

The dual problem to (2.4.4) can be formulated in terms of moment matrices, again leading to an SDP relaxation of the polynomial optimization problem (2.4.3). From a computational point of view it is often easier (i.e., faster) to compute the dual side. But extracting a sos-certificate out of the dual optimal solution is not an easy task in general. Since we do not use the dual side here, we refer interested readers to Lasserre's fundamental work [Las01]. See also Lasserre's book on polynomial optimization [Las10].

To end this subsection and to illustrate the approach based on sum of squares techniques, we state a famous example of a globally nonnegative polynomial which is not a sum of squares. In fact, this polynomial stated by Motzkin is the first example of such a polynomial. The example shows the possible advantages of testing nonnegativity with respect to semialgebraic sets.

Example 2.4.6. *The Motzkin polynomial is (globally) nonnegative but not a sum of squares (of polynomials),*

$$f(x,y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1.$$

However, it is a sum of squares of rational functions

$$\begin{aligned} f(x,y) &= (x^2+y^2-3)x^2y^2+1 \\ &= \left[\frac{x^2-y^2}{x^2+y^2}\right]^2 + \left[\frac{xy(x^2+y^2-2)}{x^2+y^2}\right]^2 + \left[\frac{x^2y(x^2+y^2-2)}{x^2+y^2}\right]^2 + \left[\frac{xy^2(x^2+y^2-2)}{x^2+y^2}\right]^2. \end{aligned}$$

Thus the Motzkin polynomial multiplied by $\mathrm{p}^2 = (x^2+y^2)^2$ *is a sum of squares*

$$\mathrm{p}^2 f(x,y) = \left[(x^2-y^2)^2 + (xy(x^2+y^2-2))^2 + (x^2y(x^2+y^2-2))^2 + (xy^2(x^2+y^2-2))^2\right]$$

(implying nonnegativity of f). After homogenization, the Motzkin polynomial is the form

$$f_h(x,y,z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6.$$

Of course, it is again nonnegative but not a sum of squares. On the other hand, since f_h is homogeneous, it is globally nonnegative if and only if it is nonnegative on the unit ball defined by the polynomial $g = 1 - x^2 - y^2 - z^2$. By definition, the quadratic module generated by g is Archimedean and hence a sos-relaxation (based on a homogeneous version of Putinar's Positivstellensatz) yields lower bounds and converges to the optimal value as the relaxation order goes to infinity. For example, using degree four polynomials (the lowest reasonable relaxation degree), YALMIP [LÖ4] computes the lower bound $\mu(4) = -0.0045964$ in 0.41 seconds (for $d = 6$, we get $\mu(6) = -0.00020332$ in 0.73 seconds).

2.4.3 Hol-Scherer's Positivstellensatz

Consider a matrix polynomial $G = G(x) \in \mathcal{S}^k[x]$ in the variables $x = (x_1, \dots, x_d)$, i.e., a matrix whose entries lie in the polynomial ring $\mathbb{R}[x]$. We say G has degree t if the maximum degree of the entries is t, i.e., $t = \max\{\deg(G_{ij}) \mid i,j \in [k]\}$. The quadratic module generated by $G(x)$ is defined as

$$\mathrm{QM}(G) = \left\{ s_0(x) + \langle S(x), G(x)\rangle \mid s_0(x) \in \Sigma[x],\ S(x) \in \Sigma^k[x] \right\} \subseteq \mathbb{R}[x], \tag{2.4.5}$$

where $\Sigma^k[x] \subseteq \mathcal{S}^k$ is the set of sum of squares $k \times k$-matrix polynomials. A matrix polynomial $S = S(x) \in \mathcal{S}^k[x]$ is called *sum of squares* (sos-matrix for short) if it has a decomposition $S = U(x)U(x)^T$ with $U(x) \in \mathbb{R}^{k\times m}[x]$ for some positive integer m. Equivalently, S has the form $(I_k \otimes [x]_t)^T Z (I_k \otimes [x]_t)$, where $[x]_t$ denotes the monomial basis in x up to $t = \max\{\deg(S_{ij}(x))/2 \mid i,j \in [k]\}$ and Z is a positive semidefinite matrix of appropriate size. For $k = 1$, S is a sos polynomial as defined in Section 2.4.2. As in the scalar case, checking whether a matrix polynomial is a sos-matrix is an SDFP (2.3.13). We denote by $\Sigma_t^k[x]$ the set of sos $k \times k$-matrices of degree at most $2t$.

If a polynomial $f(x) \in \mathbb{R}^l[x]$ lies in $\mathrm{QM}(G)$, then we say f has a *sum of squares* decomposition in $G(x)$. Obviously, every element in $\mathrm{QM}(G)$ is nonnegative on the semialgebraic set $S_G := \{x \in \mathbb{R}^d \mid G(x) \succeq 0\}$. As in Section 2.4.2, the quadratic module $\mathrm{QM}(G)$ is called *Archimedean* if $N - x^T x \in \mathrm{QM}(G)$ for some positive integer N. Hol and Scherer [HS04] showed that for polynomials positive on S_G the converse is true under the Archimedeanness condition. Henrion and Lasserre described the dual approach (in terms of moment matrices); see [HL06].

Since spectrahedra are defined by linear pencils, we are also interested in matrix polynomials positive semidefinite on a spectrahedron or, more general, on a semialgebraic set. To that end, for matrices $M = (M_{ij})_{i,j=1}^l \in \mathcal{S}^{kl}$ and $N \in \mathcal{S}^k$, define

$$\langle M, N \rangle_l := \left(\langle M_{ij}, N \rangle \right)_{i,j=1}^l = \sum_{i,j=1}^l E_{ij} \cdot \langle M_{ij}, N \rangle \,, \tag{2.4.6}$$

where E_{ij} denotes the $l \times l$-matrix with one in the (i,j)th entry and zero otherwise. We refer to (2.4.6) as the lth scalar product. It can be seen as a generalization of the Gram matrix representation of a positive semidefinite matrix. Indeed, for positive semidefinite matrices M and N the $l \times l$-matrix $\langle M, N \rangle_l$ is positive semidefinite as well.

For any positive integer l, define the quadratic module

$$\mathrm{QM}^l(G) = \left\{ S_0(x) + \langle S(x), G(x) \rangle_l \;\mid\; S_0(x) \in \Sigma^l[x],\; S(x) \in \Sigma^{kl}[x] \right\} \subseteq \mathcal{S}^l[x]. \tag{2.4.7}$$

Note that $\mathrm{QM}^1(G) = \mathrm{QM}(G)$ as in (2.4.5). If a matrix polynomial $F(x) \in \mathcal{S}^l[x]$ lies in $\mathrm{QM}^l(G)$, then we say F has a *sum of squares* decomposition in $G(x)$. Obviously, every element in $\mathrm{QM}^l(G)$ is positive semidefinite on the semialgebraic set S_G. Hol and Scherer [HS06] showed that for matrix polynomials positive definite on S_G the converse is true under the Archimedeanness condition, analog to Putinar's Positivstellensatz 2.4.3. See [CT14, Proposition 3] for an analog statement to Definition 2.4.2.

Interestingly, the usual quadratic module $\mathrm{QM}^1(G) = \mathrm{QM}(G)$ is Archimedean if and only if $\mathrm{QM}^l(G)$ is for any positive integer l.

Proposition 2.4.7. *The following two statements are equivalent.*

(1) *For some positive integer l, the quadratic module $\mathrm{QM}^l(G)$ is Archimedean.*

(2) *For all positive integers l, the quadratic module $\mathrm{QM}^l(G)$ is Archimedean.*

Furthermore, assume G is a linear pencil. $\mathrm{QM}(G)$ (and thus $\mathrm{QM}^l(G)$ for any l) is Archimedean if and only if the spectrahedron S_G is bounded.

The equivalence of the first two statements was proved by Helton, Klep, and McCullough for monic linear matrix pencils in the language of their matricial relaxation; see [HKM13, Lemma 6.9]. We recapitulate the proof and extend it to quadratic modules generated by arbitrary matrix polynomials.

Proof. The implication (2) $\Longrightarrow$ (1) is obvious. To show the reverse implication, note first that $\mathrm{QM}^l(G)$ is Archimedean if and only if $(N - x^T x) I_l \in \mathrm{QM}^l(G)$ for some positive integer N. Let $m \in \mathbb{N}$ be arbitrary but fixed. We have to show that $(N - x^T x) I_m \in \mathrm{QM}^m(G)$. Denote by E_1 the $m \times m$-matrix with one in the entry $(1,1)$ and zero elsewhere and let Q

be the $l \times m$-matrix with one in the entry $(1,1)$ and zero elsewhere. Clearly, $E_1 = Q^T Q$. Let $(N - x^T x)I_l = S_0 + \langle S, G\rangle_l$ with $S = (S_{ij})_{i,j=1}^l$ be the desired sos-representation. Setting $\tilde{S}_0 := Q^T S_0 Q = (S_0)_{11} E_1 \in \Sigma^m[x]$ and $\tilde{S} = E_1 \otimes S_{11} \in \Sigma^m[x]$, we get

$$(N - x^T x)E_1 = Q^T (N - x^T x) I_l Q = Q^T (S_0 + \langle S, G\rangle_l) Q = \tilde{S}_0 + E_1 \langle S_{11}, G\rangle = \tilde{S}_0 + \left\langle \tilde{S}, G \right\rangle_m .$$

Applying the same to E_i for $i \in [m]$ and using additivity of the quadratic module $\mathrm{QM}^m(G)$ yields $(N - x^T x) I_m \in \mathrm{QM}^m(G)$.

The last statement follows from [KS13, Corollary 4.4.2] (see also [KS11]). □

We state the desired Positivstellensatz of Hol and Scherer. See [KS10] for an alternative proof by Klep and Schweighofer using the concept of pure states.

Proposition 2.4.8 (Hol-Scherer's Positivstellensatz [HS06, Corollary 1])**.** *Let l be a positive integer and let $S_G = \{x \in \mathbb{R}^d \mid G(x) \succeq 0\}$ for a matrix polynomial $G \in \mathcal{S}^k[x]$. If the quadratic module* $\mathrm{QM}^l(G)$ (2.4.7) *is Archimedean, then it contains every matrix polynomial $F \in \mathcal{S}^l[x]$ positive definite on S_G.*

In [HS06, Section 3.4], Hol and Scherer stated an upper bound on the degree of an sos-representation. Among others, the bound depends on the degree of the Archimedeanness certificate, i.e., the degree of an sos polynomial s_0 and an sos-matrix S in a representation $N - x^T x = s_0 + \langle S, G\rangle$ for some positive integer N as well as some Pólya-type quantities.

For $l = 1$ the application of Proposition 2.4.8 to a polynomial optimization problem (2.4.3) is similar to the scalar case in Section 2.4.2. For $l > 1$, there is no natural optimization formulation. But a feasibility problem of the form “Is the matrix polynomial $F \in \mathcal{S}^l[x]$ positive semidefinite on $S_G = \{x \in \mathbb{R}^d \mid G(x) \succeq 0\}$?” can be relaxed to “Is $F \in \mathrm{QM}_t^l[x]$ for some t?”, where the *truncated quadratic module* $\mathrm{QM}_t^l(G)$ is defined as

$$\mathrm{QM}_t^l(G) = \left\{ S_0(x) + \langle S(x), G(x)\rangle_l \;\middle|\; S_0(x) \in \Sigma_t^l[x],\; S(x) \in \Sigma_t^{kl}[x] \right\} \subseteq \mathcal{S}^l[x]. \tag{2.4.8}$$

3 Complexity of Containment Problems for Projected Polytopes and Spectrahedra

In this chapter, we review known complexity classifications and classify the complexity of several containment problems for projections of polytopes and spectrahedra. In particular, the Polytope Containment problem and the Spectrahedron Containment problem are known to be co-NP-complete and co-NP-hard, respectively.

Concerning only polytopes, the computational complexity of containment problems strongly depends on the type of input representations; see [FO85] and [GK93]. Except for the Polytope Containment problem, i.e., $\mathcal{H}$-polytope in $\mathcal{V}$-polytope, all cases are solvable in polynomial time. The situation remains if the inner set is a projected $\mathcal{H}$-polytope ($\pi\mathcal{H}$) or a projected $\mathcal{V}$-polytope ($\pi\mathcal{V}$). In the setting of deciding containment of two projected $\mathcal{H}$-polytopes (i.e., $\pi\mathcal{H}$-in-$\pi\mathcal{H}$), however, the situation changes significantly as the problem turns out to be co-NP-complete (Theorem 3.2.4).

Concerning containment problems involving spectrahedra, the $\mathcal{V}$-in-$\mathcal{S}$ containment problem can be decided in polynomial time and the $\mathcal{S}$-in-$\mathcal{H}$ containment problem can be formulated by the complement of semidefinite feasibility problems (involving also strict inequalities). The remaining cases are co-NP-hard. The transition to projections of spectrahedra does not change the situation significantly. The classification involving polytopes and spectrahedra (but not their projections) is part of [KTT13, Tra14].

Our model of computation is the binary Turing machine: projections of polytopes are presented by certain rational numbers, and the size of the input is defined as the length of the binary encoding of the input data (see, e.g., [GK92]). Consider the linear projection map $\pi : \mathbb{R}^{d+m} \to \mathbb{R}^d$, $(x, y) \mapsto x$. A $\pi\mathcal{V}$-polytope P is given by a tuple $(d; m; k; v^{(1)}, \ldots, v^{(k)})$ with $d, m, k \in \mathbb{N}$, and $v^{(1)}, \ldots, v^{(k)} \in \mathbb{Q}^{d+m}$ such that $P = \pi(\mathrm{conv}\{v^{(1)}, \ldots, v^{(k)}\})$ is the projection of the convex hull onto the first d coordinates. An $\pi\mathcal{H}$-polytope P is given by a tuple $(d; m; k; A; A'; a)$ with $d, m, k \in \mathbb{N}$, matrices $A \in \mathbb{Q}^{k\times d}$ and $A' \in \mathbb{Q}^{k\times m}$, and $a \in \mathbb{Q}^k$ such that $P = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m : \ a + Ax + A'y \geq 0\}$ is bounded. For algorithmic questions, a linear pencil is given by a tuple $(d; m; k; A_0, \ldots, A_d, A'_1, \ldots, A'_m)$ with $d, m, k \in \mathbb{N}$ and $A_0, \ldots, A_d, A'_1, \ldots, A'_m \in \mathbb{Q}^{k\times k}$ rational symmetric matrices such that the projected spectrahedron S is given by $S = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m : \ A(x, y) \succeq 0\}$.

Section 3.1 reviews the complexity classifications of the Polytope Containment problem, and links it with related problems in polytope theory. In the subsequent two sections, we extend the complexity classification to containment problems involving projected polytopes (Section 3.2), and projected spectrahedra (Section 3.3). See Table 3.1 for a summary.

We use the shortcuts (co-)NPH and (co-)NPC for (co-)NP-hard and (co-)NP-complete, respectively. Recall the definition of a semidefinite feasibility problem (SDFP) in (2.3.13).

3.1 Complexity of the Polytope Containment Problem

The computational complexity of containment problems concerning polytopes (or, polyhedra) has been studied extensively by Freund and Orlin [FO85] as well as Gritzmann and Klee [GK93, GK94, GKL95]. The main results are the following.

Proposition 3.1.1 ([FO85, GK94]). *The containment problems $\mathcal{H}$-in-$\mathcal{H}$, $\mathcal{V}$-in-$\mathcal{V}$, and $\mathcal{V}$-in-$\mathcal{H}$ can be decided in polynomial time.*

This positive result is contrasted by the following hardness statement.

Proposition 3.1.2 ([FO85, GK94]). *The* Polytope Containment *problem, i.e., the decision problem whether an $\mathcal{H}$-polytope is contained in a $\mathcal{V}$-polytope, is co-NPC.*

This hardness persists even if the $\mathcal{H}$-polytope is a (standard) cube and the $\mathcal{V}$-polytope is (the affine image of) a cross polytope.

To prove the co-NP-hardness, Freund and Orlin give a reduction of the so-called integer containment problem to the Polytope Containment problem. The integer containment problem asks whether for a given rational matrix A there exists a ± 1-vector x such that the system of linear inequalities $Ax \leq \mathbb{1}$ has a solution. This problem is known to be NP-complete. The membership of the Polytope Containment problem in the class co-NP is then easily seen, as "$\mathcal{H}$ not in $\mathcal{V}$" can be shown by the existence of an extreme point of the $\mathcal{H}$-polytope that is not an element of the $\mathcal{V}$-polytope. Thus co-NP-completeness follows. Note that the co-NP-completeness statement remains valid in the unbounded case, i.e., for $\mathcal{H}$- and $\mathcal{V}$-polyhedra. See [FO85] for a detailed proof.

In general dimension (i.e., if the dimension is not fixed but part of the input) the size of one presentation can be exponential in the size of the other [McM70]. In fixed dimension, $\mathcal{H}$- and $\mathcal{V}$-presentations of a rational polytope can be converted into each other in polynomial time. Consequently, if the dimension of the polytopes is fixed, the Polytope Containment problem can be decided in polynomial time. That result can be strengthened slightly.

Theorem 3.1.3. *Let P and Q be an $\mathcal{H}$-polytope and a $\mathcal{V}$-polytope, respectively. If the dimension of P or the dimension of Q is fixed, then the* Polytope Containment *problem $P \subseteq Q$ can be decided in polynomial time.*

Proof. If the dimension of P is fixed, then first compute the affine hull of P. This can be done in polynomial time. Taking that affine hull as ambient space in fixed dimension, P can be transformed into a $\mathcal{V}$-representation in polynomial time. It remains to decide containment of a $\mathcal{V}$-polytope in a $\mathcal{V}$-polytope, which can be done in polynomial time.

Similarly, if the dimension of Q is fixed, an $\mathcal{H}$-representation of Q can be computed in polynomial time, and the resulting problem of deciding whether an $\mathcal{H}$-polytope is contained in an $\mathcal{H}$-polytope can be decided in polynomial time. □

$\subseteq$	$\mathcal{H}$	$\mathcal{V}$	$\mathcal{S}$	$\pi\mathcal{H}$	$\pi\mathcal{V}$	$\pi\mathcal{S}$
$\mathcal{H}$	P	co-NPC	co-NPH	co-NPC	co-NPC	co-NPH
$\mathcal{V}$	P	P	P	P	P	SDFP
$\mathcal{S}$	SDP	co-NPH	co-NPH	co-NPH	co-NPH	co-NPH
$\pi\mathcal{H}$	P	co-NPC	co-NPH	co-NPC	co-NPC	co-NPH
$\pi\mathcal{V}$	P	P	P	P	P	SDFP
$\pi\mathcal{S}$	SDP	co-NPH	co-NPH	co-NPH	co-NPH	co-NPH

Table 3.1: Computational complexity of containment problems for (projected) polytopes and (projected) spectrahedra.

While in this thesis we concentrate on the Polytope Containment problem, let us briefly mention some related problems. Finding the largest simplex in a $\mathcal{V}$-polytope is an NP-hard problem [GKL95]. However, for that problem Packer has given a polynomial-time approximation [Pac04]. It is well known that the problem of enumerating all facets of a $\mathcal{V}$-polytope (or, equivalently, enumerating all vertices of an $\mathcal{H}$-polytope) can be polynomially reduced to the so-called polytope verification problem, i.e., the decision problem whether a given $\mathcal{H}$-polytope and a given $\mathcal{V}$-polytope coincide; see Avis et. al. [ABS97], Kaibel and Pfetsch [KP03]. Note that enumerating the vertices of an (unbounded) polyhedron is hard [BBE$^+$08]. Moreover, Joswig and Ziegler [JZ04] showed that the polytope verification problem is polynomially equivalent to a geometric polytope completeness problem. Recently, Gouveia et. al. have studied the question which nonnegative matrices are slack matrices [GGK$^+$13], and they establish equivalence of the decision problem to the polyhedral verification problem.

3.2 Complexity Concerning Projected Polytopes

We first extend the complexity classification concerning only polytopes to projected polytopes, starting with the positive results.

Theorem 3.2.1. *The following containment problems can be decided in polynomial time.*

(1) *$\pi\mathcal{V}$-in-$\mathcal{H}$,*

(2) *$\pi\mathcal{V}$-in-$\pi\mathcal{V}$, or $\mathcal{V}$-in-$\pi\mathcal{V}$, or $\pi\mathcal{V}$-in-$\mathcal{V}$,*

(3) *$\pi\mathcal{V}$-in-$\pi\mathcal{H}$, or $\mathcal{V}$-in-$\pi\mathcal{H}$.*

Proof. By Lemma 2.2.6, all stated problems reduce to the case concerning $\mathcal{V}$-polytopes instead of a $\pi\mathcal{V}$-representation. Hence parts (1) and (2) of the theorem follow from Proposition 3.1.1. Part (3) reduces to a linear feasibility problem (LFP) which can be solved in polynomial time, e.g., by the ellipsoid method; see [Sch86, Theorem 13.4]. □

As for the Polytope Containment problem, the situation changes if the inner set is (the projection of) an $\mathcal{H}$-polytope.

Theorem 3.2.2 ($\pi\mathcal{H}$-in-$\pi\mathcal{V}$)**.** *Deciding whether an (projected) $\mathcal{H}$-polytope is contained in a (projected) $\mathcal{V}$-polytope is co-NPC.*

Proof. By Lemma 2.2.6, the problems reduce to $\mathcal{H}$-in-$\mathcal{V}$ and $\pi\mathcal{H}$-in-$\mathcal{V}$. The first one is part of Proposition 3.1.2. Concerning the second statement, it is co-NPH since $\mathcal{H}$-in-$\mathcal{V}$ is co-NPC. Given a certificate for non-containment, i.e., a point p of the $\pi\mathcal{H}$-polytope which is not in the $\mathcal{V}$-polytope, one has to solve two LFP's. Thus the problem is in co-NP. □

In the latter theorems, the statements do not differ from the non-projected cases. The next two theorems show a significant change in the complexity classification when passing from $\mathcal{H}$-polytopes to projected $\mathcal{H}$-polytopes.

Theorem 3.2.3 ($\pi\mathcal{H}$-in-$\mathcal{H}$)**.** *Deciding whether a projected $\mathcal{H}$-polytope is contained in an $\mathcal{H}$-polytope can be done in polynomial time.*

Proof. Let $\pi(P) = \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m : a + Ax + A'y \geq 0\}$ be a projected $\mathcal{H}$-polytope and let $Q = \{x \in \mathbb{R}^d \mid b + Bx \geq 0\}$ be an $\mathcal{H}$-polytope. Embed Q into $\mathbb{R}^{d+m}$ by $Q' = \{(x,y) \in \mathbb{R}^{d+m} \mid b+Bx+0y \geq 0\}$. Then the containment problem $\pi(P) \subseteq Q$ is equivalent to the $\mathcal{H}$-in-$\mathcal{H}$ containment problem $P \subseteq Q'$. The statement then follows from Proposition 3.1.1. □

An algorithmic proof of the above statement is given in Theorem 6.3.1 using Farkas' Lemma. While the containment problem $\pi\mathcal{H}$-in-$\mathcal{H}$ is decidable in polynomial time, the situation changes surprisingly for the reverse containment problem $\mathcal{H}$-in-$\pi\mathcal{H}$.

Theorem 3.2.4 ($\pi\mathcal{H}$-in-$\pi\mathcal{H}$)**.** *Deciding whether an (projected) $\mathcal{H}$-polytope is contained in a projected $\mathcal{H}$-polytope is co-NPC.*

Proof. Consider a $\mathcal{V}$-polytope. It has a representation as the projection of an $\mathcal{H}$-polytope polynomial in the input data. Thus the containment problem $\mathcal{H}$-in-$\pi\mathcal{H}$ is co-NPH since $\mathcal{H}$-in-$\mathcal{V}$ is co-NPC. It is also in the class co-NP since given a certificate for '$\mathcal{H}$ not in $\pi\mathcal{H}$', i.e., a point p, one can test whether $p \in \mathcal{H}$ and $p \notin \pi\mathcal{H}$ by evaluating the linear constraints of $\mathcal{H}$ (all have to be satisfied) and by solving a LFP, which both is in P by [Sch86, Theorem 13.4]. Therefore $\mathcal{H}$-in-$\pi\mathcal{H}$ is co-NPC. Obviously, the proof remains valid when passing to $\pi\mathcal{H}$-in-$\pi\mathcal{H}$. □

3.3 Complexity Concerning Projected Spectrahedra

Theorems 3.3.1 – 3.3.3 give the results concerning (projected) spectrahedra and (projected) $\mathcal{V}$-polytopes. Theorems 3.3.4 – 3.3.7 state the complexity classification concerning (projected) spectrahedra and (projected) $\mathcal{H}$-polytopes.

Theorem 3.3.1 ($\pi\mathcal{V}$-in-$\mathcal{S}$)**.** *Deciding whether a $\mathcal{V}$-polytope or its projection $\pi\mathcal{V}$ is contained in a spectrahedron can be done in polynomial time.*

Proof. Consider a $\mathcal{V}$-polytope $P = \operatorname{conv}\{v^{(1)}, \ldots, v^{(k)}\}$ with $v^{(i)} \in \mathbb{Q}^{d+m}$ for $i \in [k]$. By Lemma 2.2.6, its projection to the first d coordinates equals the convex hull of the projection of the defining points, i.e., $\pi(P) = \operatorname{conv}\{\pi(v^{(1)}), \ldots, \pi(v^{(k)})\}$. Consequently, $\pi(P)$ is contained in a given spectrahedron if and only if the k points $\pi(v^{(1)}), \ldots, \pi(v^{(k)})$ are elements of the spectrahedron. This can be checked by evaluating the linear pencil defining the spectrahedron at the k points which can be decided in polynomial time [GVL96]. □

The positive result is contrasted by the subsequent hardness statement.

Theorem 3.3.2 ($\pi\mathcal{S}$-in-$\pi\mathcal{V}$)**.** *Deciding whether a (projected) spectrahedron is contained in a (projected) $\mathcal{V}$-polytope is co-NPH.*

Proof. By Lemma 2.2.6, the problem reduces to $\pi\mathcal{S}$-in-$\mathcal{V}$. Deciding whether a (projected) spectrahedron is contained in a $\mathcal{V}$-polytope is co-NPH since already the Polytope Containment problem (Proposition 3.1.1) is co-NPC. □

Containment questions for spectrahedra are connected to feasibility questions of semidefinite programs in a natural way. Due to Proposition 2.3.16, the subsequent statements on containment of a (projected) $\mathcal{V}$-polytope in a projected spectrahedron and on containment of a spectrahedron in an $\mathcal{H}$-polytope do not give a complete answer concerning polynomial solvability of these containment questions in the Turing machine model.

Theorem 3.3.3 ($\pi\mathcal{V}$-in-$\pi\mathcal{S}$)**.** *Deciding whether (the projection of) a $\mathcal{V}$-polytope is contained in a projected spectrahedron is an SDFP* (2.3.13)*, whose size is polynomial in the description size of the input data.*

Proof. By Lemma 2.2.6, the problem reduces to $\mathcal{V}$-in-$\pi\mathcal{S}$. The latter is a semidefinite feasibility problem. Indeed, let $V = \{v^{(1)}, \ldots, v^{(k)}\}$ be a set of distinct points in $\mathbb{Q}^d$ and let $B(x, y) \in \mathcal{S}^l[x, y]$ be a linear pencil. Then $\operatorname{conv} V \subseteq \pi(S_B)$ if and only if for all $i \in [k]$ there exists $y_i \in \mathbb{R}^m$ such that $B(v^{(i)}, y_i)$ is positive semidefinite. Thus the $\mathcal{V}$-in-$\pi\mathcal{S}$ containment problem is equivalent to k SDFPs. ☐

In the remaining part of the section, we study the complexity of containment problems involving (projected) spectrahedra and (projected) $\mathcal{H}$-polyhedra.

Theorem 3.3.4 ($\pi\mathcal{S}$-in-$\mathcal{H}$)**.** *The problem of deciding whether a projected spectrahedron is contained in an $\mathcal{H}$-polytope can be formulated by the complement of semidefinite feasibility problems (involving also strict inequalities), whose sizes are polynomial in the description size of the input data.*

Proof. Consider a spectrahedron S_A given by the linear matrix pencil $A(x, y)$ and the coordinate projection of S_A onto the x-variables $\pi(S_A)$. Given an $\mathcal{H}$-polytope $P = \{x \in \mathbb{R}^d \mid b+Bx \geq 0\}$ with rational input $b \in \mathbb{Q}^l$ and $B \in \mathbb{Q}^{l\times d}$, construct for each $i \in [l]$ the semidefinite feasibility problem

$$b_i + \sum_{j=1}^{d} b_{ij}x_j < 0, \quad A(x, y) \succeq 0$$

involving a strict inequality. Then $\pi(S_A) \not\subseteq P$ if one of the l SDFPs is not solvable. ☐

Ben-Tal and Nemirovski showed that the $\mathcal{H}$-in-$\mathcal{S}$ containment problem is co-NPH as the co-NPH problem of maximizing a positive definite quadratic form over the unit cube can be formulated as the containment question whether the unit cube is contained in a spectrahedron (defined by a linear pencil with rank-2 coefficient matrices) [GK89, BTN02]. Theobald, Trabandt, and the author of the thesis at hand give a reduction of the NPC 3-satisfiability problem (3-SAT) to the containment problem of an $\mathcal{H}$-polytope in a ball [KTT13, Tra14]. From that we deduce the subsequent statements for containment of projected spectrahedra.

Theorem 3.3.5 ($\pi\mathcal{H}$- or $\pi\mathcal{S}$-in-$\mathcal{S}$)**.** *Deciding whether an (projected) $\mathcal{H}$-polytope or a (projected) spectrahedron is contained in a spectrahedron is co-NPH. This hardness statement persists if the $\mathcal{H}$-polytope is a standard cube or if the outer spectrahedron is a ball.*

Proof. Since the problem $\mathcal{H}$-in-$\mathcal{S}$ is co-NPH (see [BTN02, Proposition 4.1] and [KTT13, Theorem 3.4]), deciding whether a projected $\mathcal{H}$-polytope or projected spectrahedron is contained in a spectrahedron is co-NPH as well. ☐

While the $\pi\mathcal{S}$-in-$\mathcal{H}$ containment problem is efficiently solvable in practice (due to Theorem 3.3.4), the situation changes if the outer set is given as the projection of an $\mathcal{H}$-polytope.

Theorem 3.3.6 ($\pi\mathcal{S}$-in-$\pi\mathcal{H}$)**.** *Deciding whether a (projected) spectrahedron is contained in the projection of an $\mathcal{H}$-polytope is co-NPH.*

Proof. The statement follows immediately from Theorem 3.2.4 as $\mathcal{H}$ is a subclass of $\mathcal{S}$ (respectively $\pi\mathcal{S}$). ☐

Theorem 3.3.7 ($\pi\mathcal{H}$- or $\pi\mathcal{S}$-in-$\pi\mathcal{S}$)**.** *Deciding whether an (projected) $\mathcal{H}$-polytope or a (projected) spectrahedron is contained in the projection of a spectrahedron is co-NPH.*

Proof. The statement is a consequence of Theorem 3.3.5. ☐

Recall from Section 2.3 that while every $\mathcal{H}$-polytope can be rewritten in terms of a linear pencil representation, there exist linear pencil representations of polytopes which cannot be reduced to a diagonal pencil. We call this a $\mathcal{S}$-representation of a polytope (see Proposition 2.3.1). Since, given a linear pencil, the polyhedral recognition problem (PRP) is NPH, we do not expect to find a polynomial time algorithm for this containment problem.

4 Positivstellensatz Certificates for the Polytope Containment Problem

The Polytope Containment problem can be formulated as a disjointly constrained bilinear feasibility problem or, equivalently, as the maximization problem of a bilinear function on the product of two $\mathcal{H}$-polytopes living in different spaces (of the same dimension). While for linear programs Farkas' Lemma provides a perfect duality theory, the standard Lagrange duality fails for bilinear programs.

One way to tackle this problem is the application of a Positivstellensatz in order to get Farkas' type certificates for positivity of a (bilinear) polynomial on a given set. For the Polytope Containment problem, the general theory automatically implies the existence of such certificates in the situation of strong containment. However, the transition from positivity to nonnegativity, or, in our geometric situation, the transition from strong containment to non-strong containment, is a major challenge.

First we treat Handelman's Positivstellensatz which deals with positivity of polynomials on polytopes and provides hierarchies of linear programs. We show that the approach behaves geometrically. From that we deduce a necessary condition for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the tth relaxation step of the Handelman hierarchy (Theorem 4.2.4). Besides a standard convergence result (Theorem 4.2.1), we discuss the (widely open) question of degree bounds (Theorem 4.2.5) based on Averkov's proof of Handelman's Positivstellensatz.

Subsequently, we consider Putinar's Positivstellensatz. That theorem deals with more general, semialgebraic constraint sets and it provides hierarchies of semidefinite programs. As in the Handelman case, the approach behaves geometrically in our setting and thus allows to state an analog of the necessary condition for the tth relaxation step of the Putinar hierarchy (Theorem 4.3.5). Based on the study of geometric properties of the bilinear reformulation, we extend the convergence result for the strong containment case (Theorem 4.3.1) to the case where the polytopes have (at most) finitely many common boundary points (Theorem 4.3.6).

The chapter is structured as follows. We study geometric properties of a natural bilinear programming formulation in Section 4.1. Section 4.2 deals with linear relaxations of the Polytope containment problem, as based on Handelman's Positivstellensatz. Section 4.3 deals with semidefinite relaxations, as based on Putinar's Positivstellensatz. In Sections 4.2.2 and 4.3.2, we outline the general behavior of the approaches by providing certificates in certain specific structured examples (such as cubes and cross polytopes).

4.1 A Bilinear Programming Approach to the Polytope Containment Problem

For $a \in \mathbb{R}^k$, $A \in \mathbb{R}^{k\times d}$, and $B = [b^{(1)}, \ldots, b^{(l)}] \in \mathbb{R}^{d\times l}$ let

$$P = P_A = \left\{x \in \mathbb{R}^d \mid a - Ax \geq 0\right\} \text{ and } Q = Q_B = \operatorname{conv}(B) = \operatorname{conv}(b^{(1)}, \ldots, b^{(l)})$$

be an $\mathcal{H}$-polytope and a $\mathcal{V}$-polytope, respectively. The subscript in the notion of P and Q indicates the dependency on the specific representation of the polytopes involved. However, if there is no risk of confusion, we often state P and Q without subscript.

Throughout the chapter, we assume boundedness and nonemptyness of P, which both can be tested in polynomial time [Kha80]. W.l.o.g. we assume $0 \in Q$. If this is not the case, one can translate Q and P by the centroid of the vertices of Q. The centroid is clearly contained in Q, and it is an interior point of Q whenever Q has an interior point. Note that by definition, a $\mathcal{V}$-polytope is never empty. Recall that the polar polyhedron of Q is

$$Q^\circ = \left\{ z \in \mathbb{R}^d \mid \mathbb{1}_l - B^T z \geq 0 \right\} ,$$

where $\mathbb{1}_l$ is the all-ones vector in $\mathbb{R}^l$. For a polytope P, the set of vertices and the set of facets are denoted by $V(P)$ and $F(P)$, respectively.

We first collect some geometric properties of the Polytope Containment problem. Our starting point is the following reformulation as a bilinear feasibility problem.

Proposition 4.1.1. *Let the $\mathcal{H}$-polytope $P = P_A$ be nonempty and let the $\mathcal{V}$-polytope $Q = Q_B$ contain the origin.*

(1) *P is contained in Q if and only if*

$$x^T z \leq 1 \quad \text{for all } (x, z) \in P \times Q^\circ.$$

(2) *We have*

$$\sup\{x^T z \mid (x, z) \in P \times Q^\circ\} = 1$$

if and only if $P \subseteq Q$ and $\partial P \cap \partial Q \neq \emptyset$.

Proof. To (1): If $P \subseteq Q$, then for any $x \in P$ we have $x^T z \leq 1$ for all $z \in Q^\circ$. Conversely, if $x^T z \leq 1$ holds for all $z \in Q^\circ$, then for any $x \in P$ we have $x \in Q^{\circ\circ} = Q$.

To (2): Let $P \subseteq Q$ and $\partial P \cap \partial Q$ be nonempty. Then there exists a vertex $v \in V(P)$ and a facet $F \in F(Q)$ such that $v \in F$. F defines a vertex f of the polar Q°. Further, $f^T v = 1$ implies that the supremum is at least one. By part (1) of the statement, the supremum must be exactly one.

Conversely, if the supremum is one, then $x^T z \leq 1$ for all $(x, z) \in P \times Q^\circ$ and since the set $P \times Q^\circ$ is closed, there exists a point $(\bar{x}, \bar{z}) \in P \times Q^\circ$ such that $\bar{x}^T \bar{z} = 1$. Hence $\bar{x}^T z \leq 1$ for all $z \in Q^\circ$ and $\bar{x}^T \bar{z} = 1$, i.e., $\bar{x}$ defines a supporting hyperplane of Q°. Thus $\bar{x}$ is a boundary point of Q. Similarly, $x^T \bar{z} \leq 1$ for all $x \in P$ and $\bar{x}^T \bar{z} = 1$, implying $\bar{x} \in \partial P$. Consequently, $\bar{x} \in \partial Q \cap \partial P$. □

We record the following slightly more general version for the case that the precondition $0 \in \operatorname{int} Q$ is not satisfied.

Corollary 4.1.2. *Let Q_B be an irredundant representation and let the interior of Q_B be nonempty. Denote by $c = \frac{1}{l} \sum_{i=1}^{l} b^{(i)}$ the centroid of Q_B. Then $P_A \subseteq Q_B$ if and only if*

$$x^T z \leq 1 \quad \text{for all } (x, z) \in (P_A - c) \times Q^\circ_{B-c} ,$$

where $P_A - c = \{x \in \mathbb{R}^d \mid a - A(x + c) \geq 0\}$ and $B - c = B - c\mathbb{1}_{d \times l}$.

Proposition 4.1.1 suggests to formulate the Polytope containment problem via a disjointly constrained bilinear program

$$\begin{aligned} \mu^* = \sup \;\; & x^T z \\ \text{s.t.} \;\; & (x,z) \in P \times Q^\circ. \end{aligned} \tag{4.1.1}$$

Since $0 \in Q^\circ$, the supremum is nonnegative whenever P is nonempty. By Proposition 4.1.1, $P \subseteq Q$ if and only if the supremum of (4.1.1) is at most 1. The following characterization of the optimal solutions is based on Konno's work [Kon76] on bilinear programming.

Proposition 4.1.3. *Let P be a nonempty polytope. If the origin is an interior point of Q, then the set of optimal solutions for* (4.1.1) *is a set of proper faces $F \times G$ of $P \times Q^\circ$, and the supremum is finite and attained at a pair of vertices of P and Q°.*

For the convenience of the reader, we recall the short proof.

Proof. The objective function attains its maximum since it is continuous and the feasible region is nonempty, compact by $0 \in \operatorname{int} Q$.

Let $(\bar{x}, \bar{z}) \in P \times Q^\circ$ be an optimal solution. The linear program $\max\{\bar{x}^T z \mid z \in Q^\circ\}$ has a finite optimal value since its feasible region is nonempty and bounded. Hence there exists a vertex $\hat{z}$ of Q° at which the optimal value of the LP is attained. Since $\bar{z} \in Q^\circ$, we get $\bar{x}^T \hat{z} \geq \bar{x}^T \bar{z}$. Analogously, there exists a vertex $\hat{x} \in P$ at which the optimal value of the LP $\max\{x^T \hat{z} \mid x \in P\}$ is attained. Consequently, $\hat{x}^T \hat{z} \geq \bar{x}^T \hat{z} \geq \bar{x}^T \bar{z}$ and, by the optimality of $(\bar{x}, \bar{z})$, the pair of vertices $(\hat{x}, \hat{z})$ is an optimal solution for (4.1.1).

From the above, it is obvious that the set of optimal solutions is contained in the boundary $\partial P \times \partial Q^\circ$. Consider an optimal solution $(\bar{x}, \bar{z}) \in P \times Q^\circ$ to (4.1.1). Let $F_{\bar{x}}$ (resp. $F_{\bar{z}}$) be the minimal face of P (resp. Q°) containing $\bar{x}$ (resp. $\bar{z}$). For all $x \in F_{\bar{x}}$, we have $\bar{x}^T \bar{z} = \sup\{x^T z \mid z \in Q^\circ\}$. The same holds for all $z \in F_{\bar{z}}$. Thus $F_x \times F_z$ is contained in the set of the optimal soltions. For the converse, note that every boundary point of a polytope has an unique minimal face containing it. □

There is a nice geometric interpretation of the latter proposition. Since, in the case $0 \in \operatorname{int} Q$, each vertex of Q° corresponds to a facet of Q and vice versa, an optimal solution $(x,z) \in V(P) \times V(Q^\circ)$ of (4.1.1) yields a pair of a vertex of P and a facet defining normal vector of Q which either certify containment or non-containment. However, since computing the set of vertices $V(Q^\circ)$ is an NP-hard problem, it is not reasonable to reduce the problem to the set of vertices, in general.

The (sign-)oriented Euclidean distance of a point v and a hyperplane H given by $H = \{x \mid 1 - h^T x = 0\}$ is defined as

$$\operatorname{dist}(v, H) = \frac{1 - h^T v}{\|h\|}.$$

If the hyperplane defines a facet of a certain polytope, we call this the distance between the point and the facet. Given two polytopes P and Q, the minimum oriented distance of the vertices of P and the facets of Q is denoted by

$$d(P,Q) := \min\{\operatorname{dist}(v, F) \mid x \in V(P),\; F \in F(Q)\}.$$

Corollary 4.1.4. *Let the origin be an interior point of Q. Denote by $(\bar{x}, \bar{z}) \in V(P) \times V(Q^\circ)$ an optimal solution of* (4.1.1)*. Then the minimum (sign-)oriented distance between the vertices* *of P and the facets of Q is given by the distance of $\bar{x}$ and $F_{\bar{z}} = Q \cap \{x \in \mathbb{R}^d \mid 1 - \bar{z}^T x = 0\}$, i.e., $d(P,Q) = \operatorname{dist}(\bar{x}, F_{\bar{z}}) = \frac{1-\bar{z}^T\bar{x}}{\|\bar{z}\|}$.*

Proof. Let $(v,f) \in V(P) \times V(Q^\circ)$. Since $0 \in \operatorname{int} Q$, f defines a facet $F = Q \cap \{x \in \mathbb{R}^d \mid 1 - f^T x = 0\}$ of Q. The oriented Euclidean distance $\operatorname{dist}(v,F)$ of v and F is $\operatorname{dist}(v,F) = \frac{1-f^T v}{\|f\|}$. For a given F, the distance has a nonnegative value if and only if v is contained in the positive halfspace, i.e., $v \in \{x \mid 1 - f^T v \geq 0\}$.

By Proposition 4.1.1, P is contained in Q if and only if $v^T f \leq 1$ for every $(v,f) \in V(P) \times V(Q^\circ)$. Thus $P \subseteq Q$ if and only if the oriented distance $\operatorname{dist}(v,F)$ is nonnegative for all vertices of P and all facets of Q.

Consider an optimal solution $(\bar{x}, \bar{z}) \in V(P) \times V(Q^\circ)$, which exists by Proposition 4.1.3. Then the distance $\operatorname{dist}(\bar{x}, F_{\bar{z}}) = \frac{1-\bar{z}^T\bar{x}}{\|\bar{z}\|}$ is minimal over the set of vertices $V(P)$ and the set of facets $F(Q)$. □

The optimal value of problem (4.1.1) might be attained by other boundary points than vertices, and, moreover, there might be infinitely many optimal solutions. From a geometric point of view, this only occurs in somewhat degenerate cases.

Lemma 4.1.5. *Let the interior of P be nonempty and let $0 \in \operatorname{int} Q$. The following statements are equivalent.*

(1) *Problem* (4.1.1) *has finitely many optimal solutions.*

(2) *Every optimal solution of Problem* (4.1.1) *is a pair of vertices of P and Q°.*

(3) *Let $\alpha > 0$ be the unique minimal factor so that $\partial P \cap \partial(\alpha Q) \neq \emptyset$. $\partial P \cap \partial(\alpha Q)$ is a finite set and every $v \in \partial P \cap \partial(\alpha Q)$ lies in the relative interior of a facet of αQ.*

(4) *Let $\alpha > 0$ be the unique minimal factor so that $\partial P \cap \partial(\alpha Q) \neq \emptyset$. $\partial P \cap \partial(\alpha Q)$ is a finite set and for every $v \in \partial P \cap \partial(\alpha Q)$ the outer normal cone of v with respect to αQ is 1-dimensional.*

Proof. We consider the equivalence of the first two statements. Clearly, if the set of optimal solutions is a subset of $V(P) \times V(Q)$, then there are only finitely many solutions. For the converse, assume there exists an optimal solution $(\bar{x}, \bar{z}) \in P \times Q^\circ$ which is not a pair of vertices. Since $\bar{z}$ is an optimal solution of the LP $\max\{\bar{x}^T z \mid z \in Q^\circ\}$, every point in the minimal face $F(\bar{z})$ of Q° containing $\bar{z}$ is an optimal solution of the LP with the same optimal value $\mu^* = \bar{x}^T\bar{z}$. Thus the number of optimal solutions is unbounded.

Equivalence of the other statements can be shown in a similar way and is left to the reader. □

Remark 4.1.6. *To some extent, all in this section holds for the unbounded case as well, i.e., given a nonempty $\mathcal{H}$-presented polyhedron P and a $\mathcal{V}$-presented polyhedron $Q = \operatorname{conv}(B) + \operatorname{cone}(C)$ containing the origin, P is contained in Q if and only if*

$$\sup\{x^T z \mid x \in P,\ \mathbb{1}_l - B^T z \geq 0,\ -C^T z \geq 0\} \ \leq\ 1\,.$$

Proposition 4.1.3 indicates to consider vertex tracking algorithms. There is rich literature of cutting plane and branch-and-bound algorithms for bilinear programming following this approach; see, e.g., [GU77, Kon76]. Unfortunately, for bilinear programming problems Lagrange duality fails. So far, no converging algorithm is known based on this approach. In the subsequent sections, we study the bilinear reformulation of the POLYTOPE CONTAINMENT problem from the viewpoint of algebraic certificates and (linear and semidefinite) relaxations.

4.2 Handelman Certificates for the Polytope Containment Problem

We study and discuss certificates for the Polytope Containment Problem coming from Handelman's Positivstellensatz 2.4.1. This approach leads to a hierarchy of LP-relaxations to decide the Polytope Containment question. Besides a standard convergence result (Theorem 4.2.1), we discuss degree bounds on the Handelman representation (Theorem 4.2.5).

To keep notation simple, we denote the (truncated) Handelman cone (2.4.1) generated by the linear constraints $a - Ax$ and $1 - B^T z$ by $H_t(P, Q)$. The Handelman relaxation for the bilinear program (4.1.1) is

$$\nu(t) = \inf\{\nu \mid \nu - x^T z \in H_t(P, Q)\}. \tag{4.2.1}$$

Since the objective function $x^T z$ has degree two, the initial relaxation step is $t = 2$.

Asymptotic convergence of the relaxation in the general case and finite convergence in the strong containment case are direct consequences of Handelman's Positivstellensatz. Here we understand finite convergence in the sense of deciding containment in finitely many steps.

Theorem 4.2.1. *Let P be an $\mathcal{H}$-polytope and Q be a $\mathcal{V}$-polytope with $0 \in \operatorname{int} Q$.*

(1) *If $\nu(t) \leq 1$ for some integer $t \geq 2$, then $P \subseteq Q$.*

(2) *The relaxation* (4.2.1) *converges asymptotically from above to the optimal value μ^* of problem* (4.1.1)*.*

(3) *If P is strongly contained in Q (i.e., $P \subseteq Q$ and $\partial P \cap \partial Q = \emptyset$), then the hierarchy* (4.2.1) *certifies containment in finitely many steps, i.e., there exists an integer $t \geq 2$ such that $\nu(t) \leq 1$.*

Proof. The first statement is clear by construction of the relaxation.

For the third statement, let P be strongly contained in Q. Then the optimal value μ^* of problem (4.1.1) is less than one by part (2) of Proposition 4.1.1 and thus the polynomial $1 - x^T z$ is positive on $P \times Q^\circ$. Handelman's Positivstellensatz 2.4.1 implies the claim.

The second statement follows by blowing up Q such that $P \subseteq Q$ and $\partial P \cap \partial Q = \emptyset$. $\square$

In the following, we deduce a necessary condition for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the tth relaxation step of the Handelman hierarchy (see Theorem 4.2.4 below). Before proving this criterion, we collect some relevant properties of relaxation (4.2.1) in the forthcoming lemmas. These properties also show that the relaxation behaves geometrically in a natural way. Recall that every $\mathcal{H}$-representation of a certain polytope contains the facet defining halfspaces. Similarly, the vertices are part of each $\mathcal{V}$-representation.

Lemma 4.2.2 (Redundant constraints)**.** *Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ and $Q_B = \operatorname{conv}(B)$ be nonempty polytopes with $a \in \mathbb{R}^{k+1}$, $A \in \mathbb{R}^{(k+1)\times d}$, and $B \in \mathbb{R}^{d\times(l+1)}$.*

(1) *If $(a - Ax)_{k+1} \geq 0$ is a redundant inequality in the $\mathcal{H}$-representation of P_A, then it is also redundant in the Handelman representation* (4.2.1)*, i.e., the inclusion of P_A in Q_B is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering $P_{A\setminus A_{k+1}}$ instead.*

(2) *If $b^{(l+1)}$ is a redundant point in the $\mathcal{V}$-representation of Q_B, then it is also redundant in the Handelman representation* (4.2.1)*, i.e., $P_A \subseteq Q_B$ is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering $Q_{B\setminus b^{(l+1)}}$ instead.*

Proof. We only prove statement (1), the proof of part (2) is analogous. Let

$$\nu(t) - x^T z = \sum_{|\alpha,\beta,\gamma|\le t} c_{\alpha,\beta,\gamma}(a - Ax)^\alpha (1 - B^T z)^\beta (a - Ax)_{k+1}^\gamma \in H_t(P_A, Q_B)$$

with nonnegative $c_{\alpha,\beta,\gamma}$ be a Handelman representation of $\nu(t) - x^T z$ for some $t \ge 2$. Since $(a - Ax)_{k+1}$ is redundant in the description of P_A, we can write it as a convex combination of the remaining linear polynomials,

$$(a - Ax)_{k+1} = \lambda^T (a - Ax),\ \lambda^T \mathbb{1}_k = 1,\ \lambda \in \mathbb{R}_+^k.$$

The multinomial theorem implies

$$(a - Ax)_{k+1}^\gamma = \sum_{|\delta|=\gamma} \binom{\gamma}{\delta_1, \dots, \delta_k} \prod_{j=1}^k (\lambda_j (a - Ax)_j)^{\delta_j}.$$

Replacing $(a - Ax)_{k+1}^\gamma$ in the Handelman representation by the above term for any γ yields a Handelman representation of the form

$$\nu(t) - x^T z = \sum_{|\alpha',\beta'|\le t} c_{\alpha',\beta'} \prod_{i=1}^k (a - Ax)_i^{\alpha'_i} (1 - B^T z)^{\beta'} \in H_t(P_{A \setminus A_{k+1}}, Q_B)$$

with $c_{\alpha',\beta'} \ge 0$. □

Removing redundant constraints, which is a polynomial time process, may lead to faster computations; see, e.g., [GK93, Theorem 2.1].

Lemma 4.2.3 (Transitivity)**.**

(1) *Given a $\mathcal{V}$-polytope Q and $\mathcal{H}$-polytopes P and P' such that $P' \subseteq P \subseteq Q$. If there is a Handelman representation of a certain degree $t \ge 2$ certifying containment of P in Q, then it also certifies containment of P' in Q.*

(2) *Given $\mathcal{V}$-polytopes Q and Q', and an $\mathcal{H}$-polytope P such that $P \subseteq Q \subseteq Q'$. If there is a Handelman representation of a certain degree $t \ge 2$ certifying containment of P in Q, then it also certifies containment of P in Q'.*

Proof. Assume $\nu(t) - x^T z \in H_t(P, Q)$ for some $t \ge 2$. By Farkas' Lemma, we can write the linear polynomials defining P as convex combinations of the one defining P'. Using the multinomial theorem as in Lemma 4.2.2 yields a Handelman representation $\nu(t) - x^T z \in H_t(P', Q)$. This proves part (1) of the lemma. The proof of part (2) is analog. □

To end this subsection, we state a necessary characterization for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the tth relaxation step of the Handelman hierarchy. The underlying idea is to consider the $\mathcal{H}$-polytope as a union of points and interpreting each of these points as a (degenerated) polytope in $\mathcal{H}$-representation.

We define the (formal) natural $\mathcal{H}$-representation of a point $\bar{x}$ considered as a d-dimensional cube with edge length 0,

$$C_d(\bar{x}) := \left\{ x \in \mathbb{R}^d \mid \begin{pmatrix} -\bar{x} \\ \bar{x} \end{pmatrix} - \mathbb{I}_d x \ge 0 \right\}, \quad \mathbb{I}_d = \begin{bmatrix} -I_d \\ I_d \end{bmatrix}, \tag{4.2.2}$$

where I_d is the $d \times d$-identity matrix. For $t \geq 2$ and a matrix $B \in \mathbb{R}^{d \times l}$, let $Q = \mathrm{conv}(B)$ and define the set

$$\mathcal{R}_B^t = \left\{ \bar{x} \in \mathbb{R}^d \mid 1 \geq \inf\{\nu \mid \nu - x^T z \in H_t(C_d(\bar{x}), Q)\} \right\} .$$

Clearly, $\mathcal{R}_B^t$ is a subset of Q for each $t \geq 2$.

Theorem 4.2.4. *Let $Q = \mathrm{conv}\, B$ be a fixed $\mathcal{V}$-polytope. If the containment of an $\mathcal{H}$-polytope P in Q is certified by the tth relaxation step of the Handelman hierarchy* (4.2.1)*, then P is a subset of the polytope $\mathcal{R}_B^t$.*

Proof. Assume that the inclusion $P \subseteq Q$ is certified by the tth relaxation step, i.e.,

$$1 \geq \inf\{\nu \mid \nu - x^T z \in H_t(P, Q)\} , \tag{4.2.3}$$

and assume $P \not\subseteq R_B^t$. Then there exists $\bar{x} \in P \backslash R_B^t$. Considering $\bar{x}$ as fixed, we have

$$1 < \alpha := \inf\{\nu \mid \nu - x^T z \in H_t(C_d(\bar{x}), Q)\}.$$

But, by Lemma 4.3.4, this implies a contradiction to (4.2.3).

In order to show that $\mathcal{R}_B^t$ is a polytope, observe that the set $\{(\bar{x}, \nu) \in \mathbb{R}^d \times \mathbb{R} : \nu - x^T z \in H_t(C_d(\bar{x}), Q)\}$ is a polyhedron, and $\mathcal{R}_B^t$ is the projection of this set on the $\bar{x}$-variables. Since $\mathcal{R}_B^t$ is bounded, it is a polytope. □

4.2.1 Degree Bounds

The computational efforts to compute (good) certificates depends on the degree of a Handelman representation for the polynomial $1 - x^T z$. We are not aware of an explicit degree bound for a Handelman representation of this polynomial. However, a quantitative treatment of Averkov's proof of Handelman's Theorem ([Ave13], cf. also [PR01]) allows at least to provide an upper bound related to the Pólya exponent of a suitable polynomial. Here, for a homogeneous polynomial $f : \mathbb{R}^d \to \mathbb{R}$ positive on a simplex $\{x \in \mathbb{R}_+^d | \sum_{i=1}^d x_i = \alpha\}$ the *Pólya exponent* Pólya(f) of f is defined as the minimum N such that $(x_1 + \cdots + x_d)^N f(x)$ has only nonnegative coefficients. The existence of such an N is guaranteed by a classical Theorem of Pólya. Note that due to the homogeneity of f the Pólya exponent is independent of α.

In order to state the connection in an appropriate way, assume that we apply a translation on P and Q° so that they are contained in the positive orthant. Let τ be sufficiently large such that

$$\tau - \sum_{i=1}^{k} (a - Ax)_i - \sum_{i=1}^{l} (1 - B^T z)_i - \sum_{i=1}^{d} x_i - \sum_{i=1}^{d} z_i > 0$$

on $P \times Q^\circ$, and set $h : \mathbb{R}^{d+k} \times \mathbb{R}^{d+l} \to \mathbb{R}$,

$$h(x, z) \ = \ 1 - x^T z + c \sum_{i=1}^{k} (x_{d+i} - (a - Ax)_i)^2 + c \sum_{i=1}^{l} (z_{d+i} - (1 - B^T z)_i)^2$$

for some constant $c > 0$. Let $\bar{h}(x, z, w)$ be the homogenization of h with respect to $\frac{1}{\tau}(w + \sum_{i=1}^{d+k} x_i + \sum_{i=1}^{d+l} z_i)$, which is positive on the simplex $\Delta = \{(x, z, w) \in \mathbb{R}_+^{2d+k+l+1} \mid w + \sum_{i=1}^{d+k} x_i + \sum_{i=1}^{d+l} z_i = 1\}$.

Theorem 4.2.5. *Let P and Q° be in the positive orthant and let P be strongly contained in Q. Then there exists a Handelman representation of the polynomial $1 - x^T z$ whose degree is bounded by $2 +$ Pólya$(\bar{h}(x, z, w))$.*

Proof. This follows from Averkov's proof of Handelman's Theorem in connection with the observation that in our situation, the re-substitutions $x_{d+i} \mapsto (a - Ax)_i$, $z_{d+i} \mapsto (1 - B^T z)_i$, $w \mapsto 1$ do not increase the degree. $\square$

In [PR01, Theorem 1], Powers and Reznick stated a bound on the Pólya exponent in dependence of the minimum objective value (on the underlying ground simplex).

4.2.2 Examples

To illustrate the behavior of the approach, we discuss some (structured) examples.

Example 4.2.6. *Let $P = \{-1 \leq x_i \leq 1,\ i \in [d]\}$ be the d-dimensional unit cube and let $Q^\circ = \{-1 \leq ez_i \leq 1,\ i \in [d]\}$, i.e., Q is the d-dimensional unit cross polytope scaled by $e > 0$. Then*

$$\frac{d}{e} - x^T z = \frac{1}{2e} \sum_{i=1}^{d} (1 - x_i)(1 + ez_i) + \frac{1}{2e} \sum_{i=1}^{d} (1 + x_i)(1 - ez_i) \in H_2(P, Q)$$

is a Handelman representation (2.4.1) of order $t = 2$ certifying the containment $P \subseteq Q$ for $e \geq d$. Indeed, if $e \geq d$, then $1 - x^T z \geq \frac{d}{e} - x^T z \geq 0$, certifying the inclusion $P \subseteq Q$ (with strong containment if $e > d$). If $e < d$, then $1 - x^T z < \frac{d}{e} - x^T z$. This is not a certificate for non-containment since there might be a different Handelman representation. However, in this case, $P \subseteq Q$ if and only if $e \geq d$.

Interestingly, while the hardness result in Proposition 3.1.2 indicates the combinatorial complexity of this problem, the order of the Handelman representation is low ($t = 2$) and the number of summands is only linear in the dimension. $\square$

Example 4.2.7. *Let P be the d-dimensional unit cube in $\mathcal{H}$-representation as in Example 4.2.6, and $Q = \operatorname{conv}(\{-1, 1\}^d)$ be the d-dimensional unit cube in $\mathcal{V}$-representation. Denote by $rP := \{x \in \mathbb{R}^d \mid -r \leq x_i \leq r,\ i \in [d]\}$ the r-scaled unit cube with edge length $2r$. Clearly, $rP \subseteq Q$ if and only if $0 \leq r \leq 1$. This containment problem is combinatorially hard since the number of inequalities is equal to $2d + 2^d$ and thus exponential in the dimension. Consequently, setting up a Handelman representation of degree t considers $\binom{2d+2^d+t}{t}$ possible terms.*

We are interested in the maximal r such that the containment $rP \subseteq Q$ is certified by a certain relaxation degree t. On the other hand, we can ask for the minimal relaxation order t such that $P = 1P \subseteq Q$ is certified. Note that such a t only exists in case of finite convergence.

We show that for $t = 2$, $\frac{1}{d}P \subseteq Q$ is certified and for $t = d + 1$, the maximal inclusion $P \subseteq Q$ is certified. Moreover, in our numerical computations, we get that the tth relaxation step certifies containment of $\frac{t-1}{d}P \subseteq Q$; see Table 4.1. The bottleneck of computation is extracting the LP from the input. Solving the LP is pretty fast.

We write $1 \circ x_i$, where $\circ \in \{+, -\}$, to denote the constraints of P and $1 \circ z_i * z_{j \neq i}$ to denote the constraints of Q where $\circ \in \{+, -\}$ is fixed and $* \in \{+, -\}^{d-1}$ is arbitrary. For $* \in \{+, -\}$, denote by $*^{-1}$ the opposite sign (i.e., if $* = +$, then $*^{-1} = -$, and vice versa).

For $r \leq 1/d$, $rP \subseteq Q$ is certified by the Handelman representation

$$dr - x^T z = \frac{1}{2^d} \sum_{i=1}^{d} \sum_{(\circ, *) \in \{+,-\}^d} (r \circ x_i) \left(1 \circ^{-1} z_i * z_{j \neq i}\right) \in H_2(P, Q).$$

$d\backslash t$	2	3	4	5
2	1/2	1		
3	1/3	2/3	1	
4	1/4	2/4	3/4	1

Table 4.1: Computational test of containment of r-scaled $\mathcal{H}$-unit cube in $\mathcal{V}$-unit cube as described in Example 4.2.7. The entries denote the maximal r such that containment in dimension d is certified by relaxation order t.

The maximal inclusion $P \subseteq Q$ is certified by the Handelman representation of degree $t = d+1$

$$1 - x^T z = \frac{1}{2^d} \sum_{* \in \{+,-\}^d} (1 *_1 x_1) \cdots (1 *_d x_d) \left(1 *_1^{-1} z_1 \cdots *_d^{-1} z_d\right) \in H_{d+1}(P, Q). \tag{4.2.4}$$

Table 4.1 shows the maximal values for r such that containment in dimension d is certified by a given relaxation order t.

While the example problem seems to be easier than the cube-in-crosspolytope problem in Example 4.2.6, the Handelman representation in (4.2.4) *has an exponential number of summands, and we are not aware of a more compact Handelman representation.* □

We end this section with the numerical behavior of the Handelman relaxation (4.2.1) for two non-symmetric examples.

Example 4.2.8. *Consider the $\mathcal{H}$-polytope $P = \{x \in \mathbb{R}^2 \mid \mathbb{1}_4 - Ax \geq 0\}$ and the $\mathcal{V}$-polytopes $Q_1 = \operatorname{conv} B_1$ and $Q_2 = \operatorname{conv} B_2$ defined by*

$$A = \begin{bmatrix} -1 & -1 \\ 0 & -1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -1 & 0 & 2 & 2 & -1 \\ 1 & 3 & 1 & -1 & -1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -1 & -2 & 1 & 2 & 1 \\ 2 & 0 & -2 & 1 & 2 \end{bmatrix}.$$

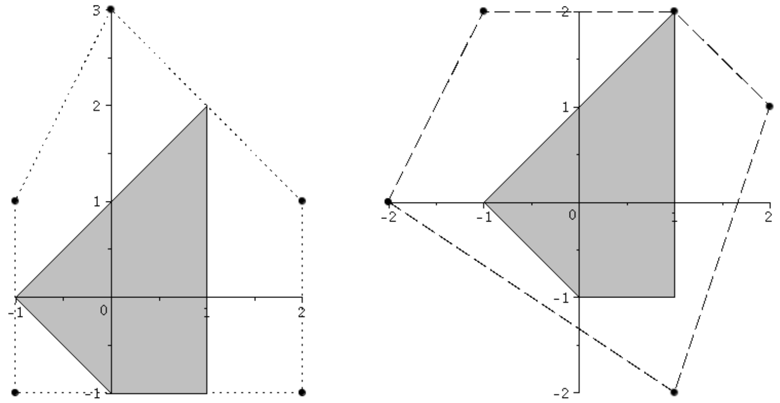
P is contained in both Q_1 and Q_2 but not strongly contained. Q_1 and P share infinitely many boundary points, in fact, the boundary of Q_1 contains a facet of P. Q_2 and P intersect in a single vertex. See Figure 4.2.1.

While for the second problem $P \subseteq Q_2$, we get a numerical certificate even in the initial relaxation step $t = 2$, the first problem $P \subseteq Q_1$ is not certified for $t = 2$. If we scale P as in Example 4.2.7, the maximum scaling factor for which containment is certified is $r = 0.8409$. For $t = 3$, $P \subseteq Q_1$ is certified. □

4.3 Putinar Certificates for the Polytope Containment Problem

In this section, we apply Putinar's Positivstellensatz 2.4.3 to the Polytope Containment problem yielding a hierarchy of semidefinite feasibility problems to decide containment. Our main goal is to show that in generic cases (in a well-defined sense) Putinar's approach yields a certificate for containment after finitely many steps; see Theorems 4.3.1 and 4.3.6.

To keep notation simple, we denote the (truncated) quadratic module (2.4.2) generated by the linear constraints $a - Ax$ and $1 - B^T z$ by $\mathrm{QM}_t(P, Q)$. The Putinar (or sos) relaxation of



(a) P (grey area) in Q_1 (doted). (b) P (grey area) in Q_2 (dashed).

Figure 4.2.1: Two non-symmetric examples as defined in Example 4.2.8.

problem (4.1.1) reads as

$$\mu(t) = \inf \left\{ \mu \mid \mu - x^T z \in \mathrm{QM}_t(P, Q) \right\}. \tag{4.3.1}$$

Denote the ith constraint defining $P \times Q^\circ$ by g_i. Let $\mu - x^T z = \sigma_0 + \sum_{i=1}^{k+l} \sigma_i g_i$ be an sos decomposition. Assume $t = 1$. Then monomials of degree at most 2 appear, i.e., $\deg(\sigma_0) \in \{0, 2\}$ and $\deg(\sigma_i g_i) \leq 2$. Since $\deg(g_i) = 1$ and σ_i is sos, σ_i must be constant (otherwise $\deg(\sigma_i) = 2$ and $\deg(\sigma_i g_i) = 3$, i.e., monomials of degree greater than 2 appear). Thus $\deg(\sum_i \sigma_i g_i) \leq 1$. Moreover, if $\deg(\sigma_0) = 2$, then purely quadratic terms like x_j^2 or z_j^2 appear for some j. Thus σ_0 is constant as well. As a consequence, the first relaxation order making sense is $t = 2$. We call $t = 2$ the *initial relaxation order*.

As for the application of Handelman's Positivstellensatz, asymptotic convergence of the relaxation in the general case and finite convergence in the strong containment case, i.e., certification of containment in finitely many steps, follow from the general theory. We have the following analog of Theorem 4.2.1.

Theorem 4.3.1. *Let P be an $\mathcal{H}$-polytope and Q be a $\mathcal{V}$-polytope with $0 \in \operatorname{int} Q$.*

(1) *If $\mu(t) \leq 1$ for some integer $t \geq 2$, then $P \subseteq Q$.*

(2) *The relaxation* (4.3.1) *converges asymptotically from above to the optimal value μ^* of problem* (4.1.1).

(3) *If P is strongly contained in Q (i.e., $P \subseteq Q$ and $\partial P \cap \partial Q = \emptyset$), then the hierarchy* (4.3.1) *certifies containment in finitely many steps, i.e., there exists an integer $t \geq 2$ such that $\mu(t) \leq 1$.*

Proof. The first statement is clear by construction of the relaxation.

Consider the third statement. Since all constraints are linear in x, z and the feasible region is bounded, the quadratic module generated by the constraints of problem (4.1.1) is Archimedean, see Corollary 2.4.4, and thus contains all polynomials $f(x, z) \in \mathbb{R}[x, z]$ strongly positive on $P \times Q^\circ$ by Putinar's Positivstellensatz 2.4.3.

Let P be strongly contained in Q. Then the optimal value μ^* of problem (4.1.1) is less than one by part (2) of Proposition 4.1.1. Thus the polynomial $1 - x^T z$ is positive on $P \times Q^\circ$ and, by the above, has a sos-representation of certain degree. This proves part (3) of the statement.

The second statement follows by blowing up Q such that $P \subseteq Q$ and $\partial P \cap \partial Q = \emptyset$. □

A priori it is not clear whether in the non-strong case finite convergence holds. In fact, for general polynomials, there are examples where finite convergence is not possible. We have a deeper look at this in Section 4.3.1 where we prove an extension of Theorem 4.3.1.

Similar to Section 4.2, we deduce in Theorem 4.3.5 a necessary condition for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the tth relaxation step of Putinar's hierarchy. Here, the criterion will employ the projection of a spectrahedron. To prepare for this criterion, we show several properties of Putinar's hierarchy for the Polytope Containment problem in Lemmas 4.3.2–4.3.4, which are the semidefinite analogs to Lemmas 4.2.2–4.2.3.

First, we see that in our situation the moment relaxation is invariant under redundant constraints, i.e., redundant inequalities in the $\mathcal{H}$-representation of P or redundant points in the $\mathcal{V}$-representation of Q. Note that for a general semialgebraic constraint set, this is not always true, even in the case of optimizing a linear function over it; see [Hen08, Section 5.2] for a well-known example (cf. also [AGV11]).

Lemma 4.3.2 (Redundant constraints)**.** *Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ and $Q_B = \operatorname{conv}(B)$ be nonempty polytopes with $a \in \mathbb{R}^{k+1}$, $A \in \mathbb{R}^{(k+1)\times d}$, and $B \in \mathbb{R}^{d\times(l+1)}$.*

(1) *If $(a - Ax)_{k+1} \geq 0$ is a redundant inequality in the $\mathcal{H}$-representation of P_A, then it is also redundant in the sos representation* (4.3.1)*, i.e., the inclusion of P_A in Q_B is certified by a certain step of the hierarchy if and only if it is certified by the same step considering $P_{A\setminus A_{k+1}}$ instead.*

(2) *If $b^{(l+1)}$ is a redundant point in the $\mathcal{V}$-representation of Q_B, then it is also redundant in the sos representation* (4.3.1)*, i.e., $P_A \subseteq Q_B$ is certified by a certain step of the hierarchy if and only if it is certified by the same step considering $Q_{B\setminus b^{(l+1)}}$ instead.*

Proof. We only prove statement (1), the proof of part (2) is analogous. Consider an sos representation of $\mu(t) - x^T z$ for some $t \geq 2$,

$$\mu(t) - x^T z = \sigma_0 + \sum_{i=1}^{k+1} \sigma_i \left(a - Ax\right)_i + \sum_{i=1}^{l} \sigma_{k+1+i} \left(1 - B^T z\right)_i \in \mathrm{QM}(A, B)\,,$$

where $\sigma_0, \ldots, \sigma_{k+l+1} \in \Sigma[x, z]$ are sos polynomials with $\deg \sigma_0 \leq 2t$ and $\deg \sigma_i \leq 2t - 2$ for $i \in \{1, \ldots, k + l + 1\}$. Since $(a - Ax)_{k+1}$ is redundant in the description of P_A, we can write it as a conic combination of the remaining linear polynomials,

$$(a - Ax)_{k+1} = \lambda_0 + \lambda^T (a - Ax),\ \lambda \in \mathbb{R}^k_+,\ \lambda_0 \in \mathbb{R}_+.$$

Replacing $\sigma_{k+1}(a - Ax)_{k+1}$ in the sos representation by

$$\sigma_{k+1} \left(a - Ax\right)_{k+1} = \lambda_0 \sigma_{k+1} + \sum_{i=1}^{k} \lambda_i \sigma_{k+1} \left(a - Ax\right)_i$$

yields an sos representation of the form

$$\mu(t) - x^T z = \sigma_0' + \sum_{i=1}^{k} \sigma_i' \left(a - Ax\right)_i + \sum_{i=1}^{l} \sigma_{k+i} \left(1 - B^T z\right)_i \in \mathrm{QM}(A\setminus A_{k+1}, B),$$

where $\sigma_i' = \lambda_i\sigma_{k+1} + \sigma_i \in \Sigma[x,z]$ with degree $\deg(\sigma_i') = \max\{\deg(\lambda_i\sigma_{k+1}), \deg(\sigma_i)\} \leq 2t-2$ for $i \in \{i,\ldots,k\}$. $\square$

Lemma 4.3.3 (Monotonicity)**.** *Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ be a polytope and let $a_{k+1} \in \mathbb{R}$, $A_{k+1} \in \mathbb{R}^{1\times d}$ such that $P' := P_A \cap \{x \in \mathbb{R}^d \mid (a-Ax)_{k+1} \geq 0\}$ is a proper subset of P_A. If for a certain relaxation order $t \geq 2$ relaxation* (4.3.1) *with respect to P_A has an optimal value of at most one, then this holds when considering P', i.e., if the relaxation certifies containment of P_A in a $\mathcal{V}$-polytope Q, then containment of P' in Q is certified as well.*

Proof. Given an sos representation of $\mu(t) - x^Tz$ w.r.t. P_A, by setting the additional sos polynomial σ_{k+1} to the zero-polynomial this yields an sos decomposition w.r.t. $P_{[A,A_{k+1}]}$. $\square$

Lemma 4.3.4 (Transitivity)**.**

(1) *Given a $\mathcal{V}$-polytope Q and $\mathcal{H}$-polytopes P and P' such that $P' \subseteq P \subseteq Q$. If for a certain relaxation order $t \geq 2$ relaxation* (4.3.1) *certifies containment of P in Q, then it also certifies containment of P' in Q.*

(2) *Given $\mathcal{V}$-polytopes Q and Q', and an $\mathcal{H}$-polytope P such that $P \subseteq Q \subseteq Q'$. If for a certain relaxation order $t \geq 2$ relaxation* (4.3.1) *certifies containment of P in Q, then it also certifies containment of P in Q'.*

Proof. Starting with P, incorporate the defining inequalities of P' into the representation of P step-by-step. By Lemma 4.3.3, in every step the lower bound of the optimal value in (4.3.1) cannot increase. At the end of this process the defining inequalities of P are all redundant (since $P' \subseteq P$) and thus can be dropped by Lemma 4.3.2. This proves part (1) of the statement. The proof of (2) is analog. $\square$

Similar to Theorem 4.2.4 for the Handelman situation, we can now state a necessary characterization for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the tth relaxation step of Putinar's hierarchy.

Recall from (4.2.2) that $C_d(\bar{x})$ denotes the (formal) natural $\mathcal{H}$-representation of a point $\bar{x}$ considered as d-dimensional cube with edge length 0. For $t \geq 2$ and a matrix $B \in \mathbb{R}^{d\times l}$, let $Q = \operatorname{conv}(B)$ and define the set

$$\mathcal{S}_B^t = \left\{\bar{x} \in \mathbb{R}^d \mid 1 \geq \inf\{\mu \mid \mu - x^Tz \in \mathrm{QM}_t(C_d(\bar{x}), Q)\}\right\}.$$

Clearly, $\mathcal{S}_B^t$ is a subset of Q for each $t \geq 2$. Moreover, $\mathcal{S}_B^t$ is the projection of a set defined by semidefinite conditions, i.e., the projection of a spectrahedron.

Theorem 4.3.5. *Let $Q = \operatorname{conv}(B)$ be a fixed $\mathcal{V}$-polytope. If the containment of an $\mathcal{H}$-polytope P in Q is certified by the tth relaxation step of Putinar's hierarchy* (4.3.1)*, then P is a subset of $\mathcal{S}_B^t$.*

Proof. Assume that the inclusion $P \subseteq Q$ is certified by the tth relaxation step, i.e.,

$$1 \geq \inf\left\{\mu \mid \mu - x^Tz \in \mathrm{QM}_t(P,Q)\right\}, \tag{4.3.2}$$

and assume $P \not\subseteq S_B^t$. Then there exists $\bar{x} \in P \backslash S_B^t$. Considering $\bar{x}$ as fixed, we have

$$1 < \alpha := \inf\left\{\mu \mid \mu - x^Tz \in \mathrm{QM}_t(C_d(\bar{x}), Q)\right\}.$$

But, by Lemma 4.3.4, this implies a contradiction to (4.3.2). □

4.3.1 Finite Convergence

By Theorem 4.3.1, there exists a Putinar representation of the polynomial $1-x^T z$ w.r.t. $P \times Q^\circ$ whenever P is strongly contained in Q. This is a severely limited case. It does not take into account that P and Q may have common boundary points (implying $\sup\{x^T z \mid (x,z) \in P \times Q^\circ\} \geq 1$) or P is not contained in Q. In this section, we prove a partial extension of Theorem 4.3.1 to the case where the bilinear optimization problem (4.1.1) has only finitely many optimal solutions (as characterized in Lemma 4.1.5).

Theorem 4.3.6. *Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ be an $\mathcal{H}$-polytope with nonempty interior and let $Q_B = \operatorname{conv}(B)$ be a $\mathcal{V}$-polytope containing the origin in its interior. Assume that one of the equivalent statements in Lemma 4.1.5 holds (e.g., there are only finitely many optimal solutions to problem* (4.1.1)*). Then $\mu^* - x^T z \in \operatorname{QM}(P_A, Q_B)$, and thus relaxation* (4.3.1) *converges in finitely many steps to the optimal value of* (4.1.1)*.*

To prove the theorem, we introduce a sufficient convergence condition by Marshall (see [Mar08]), which is based on a boundary Hessian condition.

Given $g_1, \ldots, g_k \in \mathbb{R}[x]$ and a boundary point $\bar{x}$ of $S = \{x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$. We assume that (say, by an application of the inverse function theorem), there exists a local parameterization for $\bar{x}$ in the following sense: There exist open sets $U, V \subseteq \mathbb{R}^d$ such that $\bar{x} \in U$, $\phi : U \to V$, $x \mapsto t := (t_1, \ldots, t_d)$ is bijective, the inverse $\phi^{-1} : V \to U$ is a continuously differentiable function on V, and the region R defined by $t_1 \geq 0, \ldots, t_r \geq 0$ (for some $r \in \{1, \ldots, d\}$) equals the set $S \cap U$. Given a polynomial $f \in \mathbb{R}[x]$, denote by f_1 and f_2 the linear and quadratic part of f in the localizing parameters $t_1, \ldots, t_d$, respectively.

Condition 4.3.7 (Boundary Hessian condition, BHC)**.** *If the linear form $f_1 = c_1 t_1 + \cdots + c_r t_r$ has only positive coefficients and the quadratic form $f_2(0, \ldots, 0, t_{r+1}, \ldots, t_d)$ is negative definite, then the restriction $f_{|R}$ has a local maximum in $\bar{x}$.*

Using this condition, the following generalization of Putinar's Theorem can be stated.

Proposition 4.3.8 ([Mar08, Theorem 9.5.3])**.** *Let $f, g_1, \ldots, g_k \in \mathbb{R}[x]$, and suppose that the quadratic module $\operatorname{QM}(G)$ generated by $G = \{g_1, \ldots, g_k\}$ is Archimedean. Further assume that for each global maximizer $\bar{x}$ of f over $S = \{x \in \mathbb{R}^d \mid g(x) \geq 0 \ \forall g \in G\}$ there exists an index set $I \subseteq \{1, \ldots, k\}$ such that (after renaming the variables w.r.t. the indices in I and w.r.t. the indices not in I) f satisfies BHC at $\bar{x}$. Denote by $f_{\max}$ the global maximum of f on S. In this situation, $f_{\max} - f \in \operatorname{QM}(G)$.*

Our goal is to show that under the assumptions of Theorem 4.3.6 the boundary Hessian condition holds. We will use the following version of the Karush-Kuhn-Tucker conditions adapted to the bilinear situation.

Lemma 4.3.9. *Let $f(x,z) \in \mathbb{R}[x,z]$ be a continuously differentiable function and let $\mathbb{P} := P_A \times P_B = \{(x,z) \in \mathbb{R}^{2d} \mid a - Ax \geq 0, \ b - Bz \geq 0\}$ be the product of two nonempty polytopes. If f attains a local maximum in $(\bar{x}, \bar{z})$ on $\mathbb{P}$, i.e., there exists $\varepsilon > 0$ such that for all $(x,z) \in \mathbb{P} \cap U_\varepsilon(\bar{x}, \bar{z})$ the relation $f(\bar{x}, \bar{z}) \geq f(x,z)$ holds, then there exists (α, β) such that*

$$
\begin{aligned}
\nabla f(\bar{x}, \bar{z}) &= \begin{bmatrix} A^T & 0 \\ 0 & B^T \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\
0 &= \alpha_i (a - A\bar{x})_i = \beta_j (b - B\bar{z})_j \,, \quad i \in [k], \ j \in [l] \qquad (4.3.3) \\
\alpha &\geq 0, \ \beta \geq 0,
\end{aligned}
$$

and the positive multipliers correspond to linearly independent columns in A and B, respectively.

In the lemma, only multipliers corresponding to active constraints can be positive since otherwise one of the equations (4.3.3) is violated.

Proof. Denote by I and J the index sets of the active constraints in $\bar{x}$ and $\bar{z}$, respectively. We only have to show the linear independence statement since all other parts of the lemma are the Karush-Kuhn-Tucker conditions together with the well-known fact that in case of linear constraints a constraint qualification in the KKT Theorem is not required (see, e.g., [BSS06, Section 5.1]).

By Carathéodory's Theorem [Sch86, Corollary 7.1l], there exist subsets of (I, J) such that the corresponding columns are linearly independent and $\nabla f(\bar{x}, \bar{z})$ is a strictly positive combination of these columns. Hence, (α, β) can be chosen in this way. $\square$

We are now able to prove Theorem 4.3.6. In a more general setting, Nie used the Karush-Kuhn-Tucker optimality conditions to certify the BHC; see [Nie12]. Because of the special structure of problem (4.1.1), we do not need the whole machinery used by Nie. In particular, the local parameterization needed for the BHC (see the paragraph before Condition 4.3.7) comes from an affine variable transformation. As a consequence, for POLYTOPE CONTAINMENT, our direct approach allows to prove a stronger result than we would obtain just by applying Nie's Theorem. Specifically, we obtain a geometric characterization of the degenerate situations as given in Theorem 4.3.6.

Proof (of Theorem 4.3.6). Let $(\bar{x}, \bar{z}) \in P_A \times Q_B^\circ$ be an arbitrary but fixed optimal solution. By Lemma 4.3.9, there exists $(\alpha, \beta) \in \mathbb{R}^{k+l}$ such that

$$\begin{aligned}(\bar{z}, \bar{x}) &= (A^T\alpha, B\beta) \\ 0 &= \alpha_i(a - A\bar{x})_i = \beta_j(1 - B^T\bar{z})_j\,, \quad i \in [k],\ j \in [l] \\ \alpha &\geq 0,\ \beta \geq 0,\end{aligned} \tag{4.3.4}$$

and the set of positive multipliers corresponds to linearly independent rows in A and B^T, respectively. As mentioned before, only multipliers corresponding to active constraints can be positive. Denote by I and J the index sets of linearly independent, active constraints in $\bar{x}$ and $\bar{z}$, respectively. Then $|I| \leq d$ and $|J| \leq d$.

Assume $|I| < d$. Since $\bar{z} \in \operatorname{cone}\{A_i^T\alpha \mid i \in I\}$, $\bar{z}$ lies in the outer normal cone of an at least one-dimensional face F of P_A containing $\bar{x}$. Thus $x^T\bar{z} = \bar{x}^T\bar{z}$ for all $x \in F$, in contradiction to the assumption of the theorem and Lemma 4.1.5. By a symmetric argument, $|J| < d$ is not possible either.

We apply the affine variable transformation $\phi:\ \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ defined by

$$\phi(x, z) = \begin{bmatrix} (a - Ax)_I \\ (\mathbb{1}_l - B^Tz)_J \end{bmatrix}$$

and denote the new variables by $(s, t) := (s_1, \ldots, s_d, t_1, \ldots, t_d) = (\phi_1(x, z), \ldots, \phi_{2d}(x, z))$. It is apparent that ϕ is a local parameterization at $(\bar{x}, \bar{z})$ in the sense of Condition 4.3.7. The inverse of ϕ is given by

$$(s, t) \mapsto \begin{bmatrix} A_I^{-1}(s - a_I) \\ (B_J^T)^{-1}(t - \mathbb{1}_J) \end{bmatrix}.$$

$d \backslash t$	2	3	4	5
2	0.7071	0.9937	0.9994	0.9999
3	0.5774	0.8819	0.9949	0.9994
4	0.5000	0.7906	0.9461	
5	0.4472	0.7211		

Table 4.2: Computational test of containment of r-scaled $\mathcal{H}$-unit cube in $\mathcal{V}$-unit cube as described in Example 4.3.11. The entries denote the maximal r (rounded to four decimal places) such that containment in dimension d is certified by the respective relaxation order t.

Setting $M := B_J^{-1}A_I^{-1}$, the objective $x^T z$ has the form

$$f(s,t) := \left(A_I^{-1}(s-a_I)\right)^T \left((B_J^T)^{-1}(t-\mathbb{1}_J)\right) = s^T M^T t - s^T M^T \mathbb{1}_J - a_I^T M^T t + a_I^T M^T \mathbb{1}_J$$

in the local parameterization space. Denote by f_1 the homogeneous part of degree 1. Then $(\bar{x}, \bar{z}) = \phi^{-1}(0) = (-A_I^{-1}a_I, -(B_J^T)^{-1}\mathbb{1}_J)$ implies

$$\nabla_{s,t} f_1(0) = (-\mathbb{1}_J^T B_J^{-1} A_I^{-1}, -B_J^{-1}A_I^{-1}a_I) = (\bar{z}^T A_I^{-1}, B_J^{-1}\bar{x}) = (\alpha_I, \beta_J),$$

where the last equation follows from the first identity in (4.3.4). Thus the first part of Condition 4.3.7 is satisfied. Since $|I| + |J| = r = 2d$ (where r is from Condition 4.3.7), the second assumption in the condition is obsolete. Therefore, by Proposition 4.3.8, $\mu^* - x^T z \in \mathrm{QM}(P_A, Q_B)$. □

4.3.2 Examples

We apply the semidefinite hierarchy to the examples in Section 4.2.2.

Example 4.3.10. *Let $P = \{-1 \le x_i \le 1,\ i \in [d]\}$ be the d-dimensional unit cube and let $Q^\circ = \{-1 \le ez_i \le 1,\ i \in [d]\}$, i.e., Q is the d-dimensional unit cross polytope scaled by a positive integer e. Clearly, $P \subseteq Q$ if and only if $e \ge d$.*

Consider the Putinar representation of order $t = 2$

$$\begin{aligned}\frac{d}{e} - x^T z = {} & \frac{1}{8e}\sum_{i=1}^{d}\Big[(1-x_i)[(1+x_i)^2 + (1+ez_i)^2] + (1+x_i)[(1-x_i)^2 + (1-ez_i)^2]\Big] \\ & + \frac{1}{8e}\sum_{i=1}^{d}\Big[(1-ez_i)[(1+x_i)^2 + (1+ez_i)^2] + (1+ez_i)[(1-x_i)^2 + (1-ez_i)^2]\Big].\end{aligned}$$

If $e \ge d$, then $1 - x^T z \ge \frac{d}{e} - x^T z \ge 0$, certifying the containment $P \subseteq Q$ (with strong containment if $e > d$). If $e < d$, then $1 - x^T z < \frac{d}{e} - x^T z$. This is not a certificate for non-containment, since there might be a different sos-representation. However, in this case this is not possible since $e \ge d$ is a necessary condition for containment.

Note that like in the application of Handelman's Positivstellensatz, Example 4.2.6, the necessary relaxation order is low and the number of terms is linear in the dimension. □

Example 4.3.11. *Consider the d-dimensional r-scaled unit cube rP in $\mathcal{H}$-representation and the unit cube Q in $\mathcal{V}$-representation as defined in Example 4.2.7. Again, we are interested in the maximal r such that the containment $rP \subseteq Q$ is certified by a certain relaxation degree t. On the other hand, we could ask for the minimal t such that $P_A = 1P_A \subseteq Q_B$ is certified. Note that, for $r = 1$, a priori the existence of such a t is not clear since neither Theorem 4.3.1 nor Theorem 4.3.6 applies.*

Modifying the Handelman representation, we get the same bound for Putinar's hierarchy. Recall the notation from Example 4.2.7. Then, for $r \leq 1/d$, $rP \subseteq Q$ is certified by the Putinar representation

$$\begin{aligned} dr - x^T z = {} & \frac{1}{2^{d+1}} \sum_{i=1}^{d} \sum_{(\circ,*) \in \{+,-\}^d} (r \circ x_i) \left(1 \circ^{-1} z_i * z_{j \neq i}\right)^2 \\ & + \frac{dr}{2^{d+1}} \sum_{* \in \{+,-\}^d} (1 *_1 z_1 \cdots *_d z_d) \left(1 *_1^{-1} z_1 \cdots *_d^{-1} z_d\right)^2 \in \mathrm{QM}_2(P, Q). \end{aligned}$$

We are not aware of a more compact Putinar representation. Numerically, for $t = 2$ and $d \leq 5$, we get $r(d) = \sqrt{d}/d$; see Table 4.2. Comparing the table with Table 4.1, we see that in this situation, the initial Putinar relaxation is strictly better than the initial Handelman relaxation. On the other hand, while the $(d+1)$st Handelman relaxation is exact (see (4.2.4)), it is not clear whether this is true for Putinar's relaxation. $\square$

Example 4.3.12. *Consider the $\mathcal{H}$-polytope $P = \{x \in \mathbb{R}^2 \mid \mathbb{1}_4 - Ax \geq 0\}$ and the $\mathcal{V}$-polytopes $Q_1 = \operatorname{conv} B_1$ and $Q_2 = \operatorname{conv} B_2$ as defined in Example 4.2.8. See also Figure 4.2.1.*

Recall that P is contained in both Q_1 and Q_2. Q_1 and P share infinitely many boundary points. Q_2 and P intersect in a single vertex. Thus in both examples, we are not in the situation of Theorem 4.3.6.

As for Handelman's relaxation, the first problem $P \subseteq Q_1$ is not certified with the initial Putinar relaxation $t = 2$. If we scale P, the maximum scaling factor for which containment is certified is $r = 0.9271$, which is better than for Handelman's relaxation ($r = 0.8409$). While for the second problem $P \subseteq Q_2$, the initial Handelman relaxation certifies containment, the first Putinar relaxation does only for a scalarization up to $r = 0.9996$. Thus, in contrast to Example 4.3.11, the Handelman relaxation outperforms the Putinar relaxation. $\square$

5 Positivstellensatz Certificates for the Spectrahedron Containment Problem

This chapter deals with the Spectrahedron Containment problem. Starting from a geometric viewpoint, we state connections to real algebraic geometry (leading to an asymptotically convergent hierarchy of sufficient semidefinite programs) and to operator theory. Usage of this triad allows to state several exactness and finite convergence results, partially yielding explicit certificates for containment.

While the containment problem concerning only $\mathcal{H}$-polyhedra is solvable in polynomial time, the Spectrahedron Containment problem is a co-NP-hard problem (cf. Chapter 3). Starting from the $\mathcal{H}$-in-$\mathcal{H}$ containment problem, we deduce a sufficient semidefinite criterion for the Spectrahedron Containment problem (Theorem 5.1.3). Unfortunately, necessity of the criterion fails in general. Thus two principal questions arise. First, under which additional assumptions necessity holds, and, second, whether there is a better criterion in the sense that necessity can be achieved.

To tackle the second question, we formulate the Spectrahedron Containment problem in terms of a polynomial feasibility problem, yielding a hierarchy of sufficient semidefinite criteria to decide containment. Contrary to the scalar case, finite convergence in the strong containment case is not an immediate consequence of the general theory. However, if the inner set is a spectratope and some mild additional assumption holds, then finite convergence for strong containment can be achieved by applying Hol-Scherer's Positivstellensatz (Theorem 5.1.8). From that asymptotic convergence in the general case can be deduced.

The 0th relaxation step of the hierarchy turns out to be the semidefinite feasibility criterion coming from the geometric approach as described above. Thus answering the first question (regarding necessity of the solitary criterion) is equivalent to stating Hol-Scherer certificates of degree zero.

In Section 5.2 we take a look at the Spectrahedron Containment problem involving a polyhedron. Containment of a spectrahedron in an $\mathcal{H}$-polyhedron is certified in the 0th step of the hierarchy if the spectrahedron has an interior point and the polyhedron is given in normal form (Theorem 5.2.3). We show the effectiveness of the approaches by providing (partially explicit) certificates for some structured cases. As a consequence of the connection between the Spectrahedron Containment Problem and positivity of linear maps, we get finite convergence for a special family of 2-dimensional spectrahedra (Theorem 5.4.10).

5.1 Hol-Scherer Certificates for the Spectrahedron Containment Problem

Motivating the Spectrahedron Containment problem by developing a necessary and sufficient condition for containment of $\mathcal{H}$-polyhedra (and $\mathcal{H}$-polytopes), we deduce a sufficient condition to decide the Spectrahedron Containment problem. We then introduce a hierarchy of semidefinite feasibility problems sufficient for Spectrahedron Containment coming from a sums of squares approach.

5.1.1 From the Polyhedral Case to the Spectrahedral Case

Our point of departure is the containment problem for pairs of $\mathcal{H}$-polyhedra, which by Proposition 3.1.1 can be decided in polynomial time. Theorem 5.1.1 below is a slight extension of a statement in [KTT13]. (Namely, here we drop the condition $b = \mathbb{1}_l$ as well as the boundedness condition.)

To be consistent with the notation of a linear matrix inequality (and contrary to Chapter 4), throughout this chapter $\mathcal{H}$-polyhedra (2.2.1) are denoted by inequalities $a + Ax \geq 0$ (instead of $a - Ax \geq 0$). Given an $\mathcal{H}$-representation of a polyhedron $P_A = \{x \in \mathbb{R}^d \mid a + Ax \geq 0\}$ with coefficient matrix $A \in \mathbb{R}^{k \times d}$ and $a \in \mathbb{R}^k$, we call

$$\widehat{a} + \widehat{A}x := \begin{pmatrix} 1 \\ a \end{pmatrix} + \begin{bmatrix} 0_{1\times d} \\ A \end{bmatrix} x \tag{5.1.1}$$

the *extended representation* of the polyhedron P_A.

Theorem 5.1.1. *Let $P_A = \{x \in \mathbb{R}^d \mid a + Ax \geq 0\} \neq \emptyset$ and $P_B = \{x \in \mathbb{R}^d \mid b + Bx \geq 0\}$ be polyhedra.*

(1) *P_A is contained in P_B if and only if there exists a nonnegative matrix $C \in \mathbb{R}^{l\times(k+1)}$ with $b = C\widehat{a}$ and $B = C\widehat{A}$.*

(2) *Let P_A be a polytope that is not a singleton. P_A is contained in P_B if and only if there exists a nonnegative matrix $C \in \mathbb{R}^{l\times k}$ with $b = Ca$ and $B = CA$.*

Testing whether P_A is a singleton is easy as one has to check that the system of equalities $a + Ax = 0$ has a single solution. Certainly, in this situation, checking containment is trivial as $P_A \subseteq P_B$ is equivalent to test whether a single point has nonnegative entries. The precondition in part (2) of Theorem 5.1.1, however, cannot be removed in general; see part (1) of Example 5.1.2.

For unbounded polyhedra, the extended representation of $a + Ax$ is required in order for the criterion to be exact. Without it, already in the simple case of two half spaces defined by two parallel hyperplanes, the restriction of the condition in part (1) of Theorem 5.1.1 to part (2) can fail to be feasible; see part (2) of Example 5.1.2.

Example 5.1.2.

(1) Consider the polytopes P_A and P_B given by the systems of linear inequalities

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} x \geq 0 \text{ and } \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & -1 \\ -1 & 1 \end{bmatrix} x \geq 0,$$

respectively. P_A is the singleton $\{(1,0)\}$ and P_B is a simplex containing P_A. There is no matrix C satisfying the conditions in part (2) of Theorem 5.1.1. Indeed, $b = Ca$ implies $0 = C_{11} - C_{12}$ and $B = CA$ implies $(1,0) = (-C_{11} + C_{12}, -C_{11})$, a contradiction.

Moreover, it is easy to see, that for any $P_B = \{x \in \mathbb{R}^2 \mid b + Bx \geq 0\}$ containing P_A containment is certified if and only if B has the form $B = [-b, c]$, where c is a nonpositive vector.

(2) Consider the half space given by the linear polynomial $a(x) = 1 - x_1 - x_2$. Let $b(x) = b + [B_1, B_2]x$ be any half space. The condition in part (2) of the Theorem 5.1.1 is satisfied if and only if $b = c$, $B_1 = -c$, $B_2 = -c$ and $c \geq 0$. Thus either $b(x) \equiv 0$ or $b(x)$ is a positive multiple of $a(x)$. □

For preparing related statements in more general contexts below, we review the proof of Theorem 5.1.1 which uses the affine version of Farkas' Lemma 2.2.2.

Proof. If $B = CA$ and $b = Ca$ (or $B = C\widehat{A}$ and $b = C\widehat{a}$) with a nonnegative matrix C, for any $x \in P_A$ we have
$$b + Bx = C\,(a + Ax) \geq 0,$$
i.e., $P_A \subseteq P_B$.

Conversely, if $P_A \subseteq P_B$, then any of the linear polynomials $(b+Bx)_i$, $i \in [l]$, is nonnegative on P_A. Hence, by Lemma 2.2.2, $(b + Bx)_i$ can be written as a linear combination

$$(b + Bx)_i \;=\; c'_{i0} + \sum_{j=1}^{k} c'_{ij}(a + Ax)_j \;=\; \sum_{j=0}^{k} c'_{ij}(\widehat{a} + \widehat{A}x)_j$$

with nonnegative coefficients c'_{ij}. Comparing coefficients yields $b_i = c'_{i0} + \sum_{j=1}^{k} c'_{ij} a_j$ for $i \in [l]$, implying part (1) of the statement.

To prove the second part, first translate both P_A and P_B to the origin. By assumption, there exists $\bar{x} \in P_A$. Define $\bar{a} := a + A\bar{x}$ and $\bar{b} := b + B\bar{x}$. Then $\bar{a} \geq 0,\ 0 \in \{x \in \mathbb{R}^d \mid \bar{a} + Ax \geq 0\}$ and

$$\bar{b} = C\bar{a},\ B = CA \iff b + B\bar{x} = Ca + CA\bar{x},\ B = CA \iff b = Ca,\ B = CA.$$

Thus w.l.o.g. let $a \geq 0$.

Since P_A is a polytope, Lemma 2.2.3 implies $\{0\} = \{x \in \mathbb{R}^d \mid Ax \geq 0\} = \{x \in \mathbb{R}^d \mid x^T A^T \geq 0\}$. By Stiemke's Transposition Theorem (cf. Lemma 2.2.4) there exists $\lambda > 0$ such that $A^T\lambda = 0$. For that λ we have
$$\lambda^T(a + Ax) = \lambda^T a > 0$$
whenever $a \neq 0$. We can scale λ such that $\lambda^T(a+Ax) = \lambda^T a = 1$. By multiplying that equation with c'_{i0} from above, we obtain nonnegative c''_{ij} with $\sum_{j=1}^{k} c''_{ij}(a + Ax)_j = c'_{i0}$, yielding

$$(b + Bx)_i \;=\; \sum_{j=1}^{k}(c'_{ij} + c''_{ij})(a + Ax)_j.$$

Hence, $C = (c_{ij})_{i,j=1}^{k}$ with $c_{ij} := c'_{ij} + c''_{ij}$ is a nonnegative matrix with $B = CA$ and $(Ca)_i = \sum_{j=1}^{k}(c'_{ij} + c''_{ij})a_j = b_i - c'_{i0} + c'_{i0}\,\lambda^T a = b_i$ for every $i \in [l]$. □

It is apparent from the proof that instead of using the extended representation (5.1.1), part (1) of Theorem 5.1.1 can be stated as the existence of a matrix $C \in \mathbb{R}_+^{l\times k}$ and a vector $c_0 \in \mathbb{R}_+^l$ such that $b = c_0 + Ca$ and $B = CA$.

The sufficiency part from Theorem 5.1.1 can be extended to the case of spectrahedra in a natural way. The normal form of a polyhedron P_A as a spectrahedron, as defined in (2.3.2), is given by
$$P_A = \left\{ x \in \mathbb{R}^d \mid A(x) = \bigoplus_{i=1}^{k} a_i(x) \succeq 0 \right\},$$
where $a_i(x)$ is the ith entry of the vector $a + Ax$. Then, as in the definition of a linear pencil (2.3.1), A_p is the diagonal $k \times k$-matrix $\operatorname{diag}(A_{:,p})$ of the pth column of A. Proceed in the same way with P_B. Now define a $kl \times kl$-matrix $\widetilde{C}$ by writing the entries of C on the

diagonal, i.e., $\tilde{C} = \operatorname{diag}(c_{11}, \dots, c_{l1}, \dots, c_{1k}, \dots, c_{lk})$. Then the condition from Theorem 5.1.1 translates to the existence of a diagonal matrix C with

$$C = (C_{ij})_{i,j=1}^k \succeq 0, \ \forall p \in [0, d] : \ B_p = \sum_{i=1}^k (A_p)_{ii} C_{ii},$$

where $C_{ij} \in \mathbb{R}^{l \times l}$ and $(A_p)_{ij}$ is the (i, j)th entry of A_p.

Theorem 5.1.3 below tells us that C does not need to be diagonal in order to yield a sufficient condition for the Spectrahedron Containment problem. Subsequently, the indeterminate matrix $C = (C_{ij})_{i,j=1}^k$ is a symmetric $kl \times kl$-matrix, where the C_{ij} are $l \times l$-blocks.

As introduced in the preliminary Section 2.4.3, denote by $\mathcal{S}^k$ the set of all real symmetric $k \times k$-matrices and by $\mathcal{S}^k[x]$ the set of symmetric $k \times k$-matrices with polynomial entries in $x = (x_1, \dots, x_d)$. As for polyhedra, given a linear pencil $A(x) \in \mathcal{S}^k[x]$, we call

$$\widehat{A}(x) := \begin{bmatrix} 1 & 0 \\ 0 & A(x) \end{bmatrix} \in \mathcal{S}^{k+1}[x] \tag{5.1.2}$$

the *extended linear pencil* of the spectrahedron $S_A = S_{\widehat{A}}$. (The spectrahedra coincide, since the 1 we add for technical reasons is redundant.) The coefficient matrices of the pencil $\widehat{A}(x) = \widehat{A}_0 + \sum_{p=1}^d x_p \widehat{A}_p$ are denoted by $\widehat{A}_p$ for $p \in [0, d]$, as usual.

Theorem 5.1.3. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. If there exists a symmetric matrix $C = (C_{ij})_{i,j=1}^k \in \mathcal{S}^{kl}$ such that*

$$C \succeq 0, \ B_p = \sum_{i,j=1}^k (A_p)_{ij} C_{ij} \ \textit{for } p \in [0, d], \tag{5.1.3}$$

then S_A is contained in S_B.

The criterion in Theorem 5.1.3 can also be stated for the extended representation $\widehat{A}$. But, since the representation plays a role only in the necessity part, we drop this here for simplicity.

Helton, Klep, and McCullough proved Theorem 5.1.3 by very different methods in the setting of so-called "matricial spectrahedra" (or, free spectrahedra); see [HKM13, Theorem 3.5]. We present a more streamlined proof here.

Proof. From the linear constraints, we get

$$\begin{aligned} B(x) = B_0 + \sum_{p=1}^d x_p B_p \ &= \ \sum_{i,j=1}^k (A_0)_{ij} \, C_{ij} + \sum_{p=1}^d \sum_{i,j=1}^k x_p \, (A_p)_{ij} \, C_{ij} \\ &= \ \sum_{i,j=1}^k (A(x))_{ij} \, C_{ij} \ = \ \mathbb{I}^T \left((A(x))_{ij} \, C_{ij} \right)_{i,j=1}^k \mathbb{I} \end{aligned}$$

with $\mathbb{I} = [I_l, \dots, I_l]^T \in \mathbb{R}^{kl \times l}$. Thus to prove the claim, it suffices to show that the matrix $\left((A(x))_{ij} \, C_{ij} \right)_{i,j=1}^k$ is positive semidefinite. Indeed, then

$$\begin{aligned} v^T B(x) v &= v^T \left(\mathbb{I}^T \ \left((A(x))_{ij} \, C_{ij} \right)_{i,j=1}^k \ \mathbb{I} \right) v \\ &= \left(v^T, \dots, v^T \right) \left((A(x))_{ij} \, C_{ij} \right)_{i,j=1}^k (v, \dots, v)^T \geq 0 \end{aligned}$$

for every $v \in \mathbb{R}^l$.

Let $x \in S_A$. Since $A(x)$ and C are positive semidefinite, the Kronecker product $A(x) \otimes C$ is positive semidefinite as well, see (2.1.1). As a consequence, all principal submatrices of $A(x) \otimes C$ are positive semidefinite. $A(x) \otimes C$ is a $k^2 l \times k^2 l$-matrix with k^2-blocks of the form

$$(A(x))_{ij} C = \begin{bmatrix} (A(x))_{ij} C_{11} & \cdots & (A(x))_{ij} C_{1k} \\ \vdots & (A(x))_{ij} C_{ij} & \vdots \\ (A(x))_{ij} C_{k1} & \cdots & (A(x))_{ij} C_{kk} \end{bmatrix} \in \mathcal{S}^{kl}.$$

(Remember that $(A(x))_{ij}$ is a scalar). Consider the principal submatrix where we take the (i,j)th subblock of every (i,j)th block $(A(x))_{ij} C$, i.e., $(A(x))_{ij} C_{ij}$. Since $((A(x))_{ij} C_{ij})_{ij=1}^k$ is a principal submatrix of $A(x) \otimes C$, $B(x)$ is positive semidefinite as well. □

Combining the above Theorems 5.1.1 and 5.1.3, we get the next corollary.

Corollary 5.1.4. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be normal forms of polyhedra* (2.3.2)*.*

(1) *S_A is contained in S_B if and only if system* (5.1.3) *has a solution with respect to $\widehat{A}(x)$ and $B(x)$.*

(2) *Let S_A be a polytope that is not a singleton. $S_A \subseteq S_B$ if and only system* (5.1.3) *has a solution with respect to $A(x)$ and $B(x)$.*

In our approach it becomes apparent that we can relax the criterion by replacing the linear constraint on the constant matrices in (5.1.3) with semidefinite constraints,

$$C = (C_{ij})_{i,j=1}^k \succeq 0, \quad B_0 - \sum_{i,j=1}^k (A_0)_{ij} C_{ij} \succeq 0, \quad \forall p \in [d]: \ B_p = \sum_{i,j=1}^k (A_p)_{ij} C_{ij}. \qquad (5.1.4)$$

If S_A is contained in the positive orthant, we can give a stronger version of the criterion introduced in Theorem 5.1.3.

Corollary 5.1.5. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils and let S_A be contained in the positive orthant. If the following system is feasible, then $S_A \subseteq S_B$.*

$$C = (C_{ij})_{i,j=1}^k \succeq 0, \quad B_0 - \sum_{i,j=1}^k (A_0)_{ij} C_{ij} \succeq 0, \quad \forall p \in [d]: \ B_p - \sum_{i,j=1}^k (A_p)_{ij} C_{ij} \succeq 0. \quad (5.1.5)$$

Proof. The proof is along the lines of the proof of Theorem 5.1.3. Indeed, since S_A lies in the positive orthant, we have $x \geq 0$ for all $x \in S_A$ and hence,

$$B(x) = B_0 + \sum_{p=1}^n x_p B_p \ \succeq \ \sum_{i,j=1}^k (A_0)_{ij} \, C_{ij} + \sum_{p=1}^n \sum_{i,j=1}^k x_p \, (A_p)_{ij} \, C_{ij} \ = \ \sum_{i,j=1}^k (A(x))_{ij} \, C_{ij}. \ \square$$

By relaxing system (5.1.3) to (5.1.5) the number of scalar variables remains $\frac{1}{2} kl(kl+1)$, whereas the $\frac{1}{2}(n+1)l(l+1)$ linear constraints are replaced by $n+1$ semidefinite constraints of size $l \times l$.

If containment restricted to the positive orthant implies containment everywhere else, criterion (5.1.5) can be applied, even if the spectrahedron is not completely contained in the positive orthant. To make use of this fact, we have to premise a certain structure of the spectrahedra. We give an example for this in the following corollary.

Corollary 5.1.6. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils defining spectrahedra with a reflection symmetry with respect to all coordinate hyperplanes. If system* (5.1.5) *is feasible, then $S_A \subseteq S_B$.*

Criterion (5.1.3) is not necessary for Spectrahedron Containment in general, even when a constraint qualification (e.g., as in Lemma 2.3.10 or Lemma 2.3.9) is satisfied. Subsequently, we review an example from [HKM13, Example 3.1, 3.4] which shows that the containment criterion is not exact in general.

Example 5.1.7. *Consider the monic linear pencils $A(x) = I_3 + x_1(E_{1,3} + E_{3,1}) + x_2(E_{2,3} + E_{3,2}) \in \mathcal{S}^3[x]$ and*

$$B(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + x_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Clearly, both define the unit disc, i.e., $S_A = S_B$. First, we look into the inclusion $S_B \subseteq S_A$.

Claim. The containment question $S_B \subseteq S_A$ is certified by criterion (5.1.3).

Note that the roles of A and B in (5.1.3) *have to be interchanged. Criterion* (5.1.3) *is satisfied if and only if there exist $c_1, c_2, c_3 \in \mathbb{R}$ such that*

$$C = \left[\begin{array}{ccc|ccc} \frac{1}{2} & 0 & \frac{1}{2} & 0 & c_1 & c_2 \\ 0 & \frac{1}{2} & 0 & -c_1 & 0 & c_3 \\ \frac{1}{2} & 0 & \frac{1}{2} & -c_2 & 1-c_3 & 0 \\ \hline 0 & -c_1 & -c_2 & \frac{1}{2} & 0 & -\frac{1}{2} \\ c_1 & 0 & 1-c_3 & 0 & \frac{1}{2} & 0 \\ c_2 & c_3 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \in \mathbb{R}^{6\times 6}$$

is positive semidefinite. Since the 2×2-block in the top left corner is positive definite, the matrix C is positive semidefinite if and only if the Schur complement (see, e.g., [dK02, Theorem A.9]) with respect to this block is positive semidefinite. One can easily check that this is the case if and only if $c_1 = c_3 = \frac{1}{2}$ and $c_2 = 0$.

Claim. The reverse containment question $S_A \subseteq S_B$ is not certified by criterion (5.1.3).

$S_A \subseteq S_B$ is certified by (5.1.3) *if and only if there exist $c_1, \ldots, c_{12} \in \mathbb{R}$ such that*

$$C = \left[\begin{array}{cc|cc|cc} c_1 & c_2 & c_9 & c_{10} & \frac{1}{2} & c_7 \\ c_2 & c_3 & c_{11} & c_{12} & -c_7 & -\frac{1}{2} \\ \hline c_9 & c_{11} & c_4 & c_5 & 0 & c_8 \\ c_{10} & c_{12} & c_5 & c_6 & 1-c_8 & 0 \\ \hline \frac{1}{2} & -c_7 & 0 & 1-c_8 & 1-c_1-c_4 & -c_2-c_5 \\ c_7 & -\frac{1}{2} & c_8 & 0 & -c_2-c_5 & 1-c_3-c_6 \end{array}\right] \in \mathbb{R}^{6\times 6}$$

is positive semidefinite. We show the infeasibility of system (5.1.3).

Assume that C is positive semidefinite. Then all principal minors are nonnegative. In particular, the diagonal minors are nonnegative. Consider the (leading) principal minor

$$\begin{vmatrix} c_1 & \frac{1}{2} \\ \frac{1}{2} & 1-c_1-c_4 \end{vmatrix} = c_1(1-c_1-c_4) - \frac{1}{4} = \left[c_1(1-c_1) - \frac{1}{4}\right] - c_1c_4.$$

Since the expression in the brackets as well as the second term are always less than or equal to zero, the minor is nonpositive. Therefore, $c_1(1-c_1) - \frac{1}{4} = 0$ and $c_1c_4 = 0$, or equivalently, $c_1 = \frac{1}{2}$ and $c_4 = 0$. Recall that whenever a diagonal element of a positive semidefinite matrix *is zero, the corresponding row is the zero vector, that is* $c_5 = c_8 = c_9 = c_{11} = 0$. *Now, we get a contradiction since nonnegativity of the (leading) principal minor*

$$\begin{vmatrix} c_6 & 1-c_8 \\ 1-c_8 & 1-c_1-c_4 \end{vmatrix} = \begin{vmatrix} c_6 & 1 \\ 1 & \frac{1}{2} \end{vmatrix} = \frac{1}{2}c_6 - 1$$

implies that $c_6 \geq 2$ *and thus* $1-c_3-c_6 \leq -1-c_3$. *Therefore either* $c_3 \leq -1$ *or* $1-c_3-c_6 < 0$, *which both contradicts positive semidefiniteness. This proves the claim.*

This example is not only a counterexample for the necessity of criterion (5.1.3) *for containment but serves also as a counterexample for necessity under validity of a constraint qualification. Indeed, both constraint qualifications, Lemma 2.3.10 and Lemma 2.3.9, are satisfied by the pencil* $A(x)$ *but, as seen above, certification of containment fails.*

In Example 5.3.5, we contrast this phenomenon by showing that for this example there exists a scaling factor r *for one of the spectrahedra so that the containment criterion is satisfied after this scaling.* □

5.1.2 Hol-Scherer Certificates for Spectrahedron Containment

In the subsequent sections, we study a sum of squares relaxation for Spectrahedron Containment based on a quantified polynomial (in fact, semidefinite) optimization problem. The hierarchy of semidefinite feasibility problems coming out of this approach is at least as powerful as the solitary semidefinite criterion (5.1.3).

Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils (2.3.1). For nonnegative integers $t \geq 0$, consider the hierarchy of semidefinite feasibility problems

$$B(x) = S_0(x) + \langle S(x), A(x)\rangle_l \in \mathrm{QM}_t^l(A),\ S_0 \in \Sigma_t^l[x],\ S \in \Sigma_t^{kl}[x], \tag{5.1.6}$$

where $\mathrm{QM}_t^l(A)$ is the *truncated quadratic module* as defined in (2.4.8) and $\langle\cdot,\cdot\rangle_l$ is the lth scalar product (2.4.6). Clearly, if $B(x)$ has such a representation, then S_A is contained in S_B. Applying Hol-Scherer's Positivstellensatz 2.4.8, the converse holds for the strong containment case under mild additional assumptions. Recall that a bounded spectrahedron is called a spectratope.

Theorem 5.1.8. *Let* $A(x) \in \mathcal{S}^k[x]$ *be a linear pencil.*

(1) *If* $B(x) \in \mathrm{QM}_t^l(A)$ *for some* $t \geq 0$, *then* $S_A \subseteq S_B$.

(2) *Let* $B(x) \in \mathrm{QM}_t^l(A)$. *If* $S_0(x) \succ 0$ *for all* $x \in S_A$, *then* $S_A \subseteq \operatorname{int} S_B$.

(3) *Let* S_A *be a spectratope and* $B(x) \in \mathcal{S}^l[x]$ *be a reduced linear pencil (cf. Proposition 2.3.2). If* S_A *is strongly contained in* S_B *(i.e.,* $S_A \subseteq S_B$ *and* $\partial S_A \cap \partial S_B = \emptyset$*), then* $B(x) \in \mathrm{QM}^l(A)$.

Proof. The first statement is clear by construction of the relaxation and by definition of the lth scalar product (2.4.6).

If $B(x) = S_0(x) + \langle S(x), A(x)\rangle_l \in \mathrm{QM}_t^l(A)$ with $S_0(x) \succ 0$ for all $x \in S_A$, then $B(x) \succ 0$ for all $x \in S_A$. Thus $S_A \subseteq \{x \in \mathbb{R}^d \mid B(x) \succ 0\} \subseteq \operatorname{int} S_B$.

Consider the third statement. By Proposition 2.4.7, boundedness of S_A implies Archimedeanness of $\mathrm{QM}^l(A)$. Since strong containment and reducedness of $B(x)$ implies $B(x) \succ 0$ on S_A, the statement follows from Hol-Scherer's Positivstellensatz 2.4.8. □

For computational reasons it is interesting to have an optimization version of the hierarchy (5.1.6).

Proposition 5.1.9. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. Consider the quantified polynomial optimization problem*

$$\begin{aligned} \mu^* \;=\; & \sup \;\; \mu \\ & \text{s.t.} \;\; B(x) - \mu I_l \succeq 0 \;\forall x \in S_A. \end{aligned} \tag{5.1.7}$$

(1) *$\mu^* \geq 0$ if and only if $S_A \subseteq S_B$.*

(2) *$\mu^* > 0$ implies $S_A \subseteq \operatorname{int} S_B$. If $B(x)$ is reduced (cf. Proposition 2.3.2), then the converse holds true.*

Proof. By definition, $S_A \subseteq S_B$ if and only if $B(x) \succeq 0$ for all $x \in S_A$. Thus the first statement is clear. If $\mu^* > 0$, then $B(x) \succeq \mu^* I_l \succ 0$ for all $x \in S_A$, implying $B(x) \succ 0$ on S_A.

Let $B(x)$ be reduced. Assume $\mu^* \leq 0$. Clearly, if $\mu^* < 0$, then S_A is not contained in S_B. If $\mu^* = 0$, then for every $\varepsilon > 0$ there exists $x \in S_A$ such that $B(x) - \varepsilon I_l$ has a negative eigenvalue. By tending $\varepsilon \downarrow 0$, continuity implies that $B(x)$ has zero as an eigenvalue. □

By replacing the positive semidefiniteness condition $B(x) - \mu I_l \succeq 0$ on S_A with the membership in the truncated quadratic module $\mathrm{QM}_t^l(A)$ for some $t \geq 0$, Problem (5.1.7) leads to the hierarchy

$$\begin{aligned} \mu_{\mathrm{sos}}(t) = & \sup \;\; \mu \\ & \text{s.t.} \;\; B(x) - \mu I_l = S_0(x) + \langle S(x), A(x) \rangle_l \\ & \qquad S_0(x) \in \Sigma_t^l[x], \; S(x) = (S_{i,j}(x))_{i,j=1}^l \in \Sigma_t^{kl}[x] \end{aligned} \tag{5.1.8}$$

for $t \geq 0$. We refer to the 0th step of the hierarchy or, equivalently, to the 0th feasibility/membership problem (5.1.6) as the *initial relaxation step* of the Hol-Scherer relaxation.

The statements for the feasibility problem (5.1.6) can be brought forward to the optimization version (5.1.8). Clearly, if $\mu_{\mathrm{sos}}(t) \geq 0$ for some integer $t \geq 0$, then $S_A \subseteq S_B$. Applying Hol-Scherer's Positivstellensatz, finite convergence in the case of strong containment (under the reducedness assumption) and asymptotic convergence in the general (bounded) case can be achieved.

Corollary 5.1.10. *Let S_A be a spectratope given by the linear pencil $A(x) \in \mathcal{S}^k[x]$.*

(1) *The optimal value of the sos-relaxation* (5.1.8) *converges asymptotically from below to the optimal value μ^* of the quantified semidefinite optimization problem* (5.1.7).

(2) *Let $B(x) \in \mathcal{S}^l[x]$ be a reduced linear pencil (cf. Proposition 2.3.2). If S_A is strongly contained in S_B, then hierarchy* (5.1.8) *certifies containment in finitely many steps, i.e., there exists an integer $t \geq 0$ such that $\mu_{\mathrm{sos}}(t) \geq 0$.*

Proof. By [HS06, Theorem 1], relaxation (5.1.8) converges asymptotically to the optimal value of the quantified optimization problem (5.1.7) if the quadratic module is Archimedean. By Proposition 2.4.7, the latter holds if S_A is a spectratope.

Since $B(x)$ is reduced, strong containment implies $B(x) \succ 0$ on S_A. By Theorem 5.1.8, $B(x) \in \mathrm{QM}^l(A)$. □

5.1.3 The Initial Step of the Hol-Scherer Hierarchy

We start by a fundamental result that stresses the importance of the initial relaxation step of the Hol-Scherer hierarchy (5.1.6). Namely, the initial step of the hierarchy coincides with the semidefinite criterion coming from the geometric approach in Section 5.1.1.

It is easy to see that a matrix polynomial $S \in \mathcal{S}^{kl}[x]$ is sos if and only if there exists a positive semidefinite matrix Z such that $S(x) = (I_{kl} \otimes [x]_t)^T Z (I_{kl} \otimes [x]_t)$, where $[x]_t$ denotes the truncated monomial basis as in Section 2.4; see [HS06, Lemma 1]. Analogously, $S_0(x) = (I_l \otimes [x]_t)^T Z^0 (I_l \otimes [x]_t)$ for some positive semidefinite matrix Z^0 of appropriate size. By an easy computation, one can see that $S(x) = \langle Z, [x]_t[x]_t^T\rangle_{kl}$ and $S_0(x) = \langle Z^0, [x]_t[x]_t^T\rangle$. Write

$$[x]_t[x]_t^T = \sum_{|\alpha|\leq 2t} x^\alpha P_\alpha^0,$$

where P_α^0 are $\binom{d+t}{t} \times \binom{d+t}{t}$ -matrices with entries indexed by α equal to one and zero elsewhere. Similarly,

$$(I_k \otimes [x]_t)\, A(x)\, (I_k \otimes [x]_t)^T = A(x) \otimes [x]_t[x]_t^T = \sum_{|\alpha|\leq 2t} x^\alpha P_\alpha$$

and $B(x) = B_0 + \sum_{i=1}^d x_i B_i + \sum_{2\leq|\alpha|\leq 2t} x^\alpha \cdot 0$. The membership problem $B(x) \in \mathrm{QM}_t^l(A)$ (cf. (5.1.6)) for some $t \geq 0$ is then equivalent to the semidefinite feasibility problem

$$\exists Z^0, Z \succeq 0 : \ \left\langle Z^0, P_\alpha^0 \right\rangle_l + \langle Z, P_\alpha\rangle_l = \begin{cases} B_0 & |\alpha| = 0, \\ B_p & |\alpha| = 1,\ \alpha_p = 1,\ p \in [d], \\ 0 & \text{else.} \end{cases} \qquad \forall |\alpha| \leq 2t$$

For more details we refer to [HS06, Section 5]. We obtain the following statement.

Theorem 5.1.11. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. The following are equivalent.*

(1) $B(x) \in \mathrm{QM}_0^l(A)$

(2) $\exists C' \in \mathcal{S}_+^{kl},\ C_0' \in \mathcal{S}_+^l :\ B_0 = C_0' + \langle A_0, C'\rangle_l\,,\ B_p = \langle A_p, C'\rangle_l \ \forall p \in [d]$

(3) $\exists C = (C_{ij})_{i,j=0}^k \in \mathcal{S}_+^{(k+1)l} :\ B_0 = \sum_{i,j=0}^k (\widehat{A}_0)_{ij} C_{ij},\ B_p = \sum_{i,j=0}^k (\widehat{A}_p)_{ij} C_{ij} \ \forall p \in [d]$

where $\widehat{A}(x)$ denotes the extended linear pencil (5.1.2).

In particular, the initial relaxation step (5.1.6) *certifies containment if and only if the semidefinite feasibility criterion* (5.1.3) *does when applied to the extended linear pencil $\widehat{A}(x)$.*

Moreover, the initial relaxation step (5.1.6) *certifies containment with $S_0 = 0$ if and only if the semidefinite feasibility criterion* (5.1.3) *does when applied to $A(x)$.*

Part (3) of the theorem can equivalently be stated as

$$\exists C = (C_{ij})_{i,j=0}^k \in \mathcal{S}_+^{(k+1)l} :\ B_0 = C_{00} + \sum_{i,j=1}^k (A_0)_{ij} C_{ij},\ B_p = \sum_{i,j=1}^k (A_p)_{ij} C_{ij} \ \forall p \in [d],$$

where C_{00} becomes obsolete, if applied to $A(x)$ instead of $\widehat{A}(x)$.

Proof. Consider the first equivalence. Let $t = 0$ and set $x_0 \equiv 1$. Then $[x]_0 = 1$ and $[x]_0[x]_0^T = 1 = 1x_0 + \sum_{p=1}^{d} 0x_p$. Thus the coefficient matrices P_p^0 (as introduced before Theorem 5.1.11) are scalars with $P_0^0 = 1$ and $P_p^0 = 0$ for $p \in [d]$. For the term involving $A(x)$, we get

$$A(x) \otimes [x]_0[x]_0^T = A(x) = \sum_{p=0}^{d} A_p x_p.$$

Thus $Z \in \mathcal{S}^{kl}$ and $Z^0 \in \mathcal{S}^l$. The constraints in the semidefinite feasibility system read as

$$B_p - \langle Z, A_p \rangle_l = \left\langle Z^0, P_p^0 \right\rangle_l = P_p^0 Z^0 = \begin{cases} Z^0 & p = 0 \\ 0 & p \neq 0 \end{cases}.$$

Thus $B_p = \langle Z, A_p \rangle_l$ for $p \in [d]$ and $B_0 = Z^0 + \langle Z, A_0 \rangle$. Rewriting $\widehat{A}_0 = 1 \oplus A_0,\ \widehat{A}_p = 0 \oplus A_p$, and defining $\widehat{Z} = (\widehat{Z}_{ij})_{i,j}$ by $\widehat{Z}_{s,t} = Z_{s,t}^0 \oplus Z_{s,t} \in \mathcal{S}^{k+1}$, we get that the initial step of the hierarchy (5.1.6) certifies containment if and only if there exists a positive semidefinite matrix $\widehat{Z} \in \mathcal{S}_+^{(k+1)l}$ such that

$$B_0 = \left\langle \widehat{Z}, \widehat{A}_0 \right\rangle_l \text{ and } B_p = \left\langle \widehat{Z}, \widehat{A}_p \right\rangle_l \text{ for } p \in [d].$$

To prove the equivalence of (2) and (3), consider a matrix $C = (C_{ij})_{i,j=1}^k$. We define C' by permuting rows and columns simultaneously, i.e., $C' = (C'_{st})_{s,t=1}^l$. It is a well-known fact that this operation preserves positive semidefiniteness since any permutation matrix is invertible. Hence, $C \succeq 0$ if and only if $C' \succeq 0$. Moreover,

$$(B_p)_{s,t} = \sum_{i,j=1}^{k} (A_p)_{ij} (C'_{st})_{i,j} = \langle A_p, C'_{st} \rangle$$

for $p \in [0, d]$. This proves the claim. □

In order to state certificates for some structured examples and important cases in the forthcoming sections, we have to develop some auxiliary results on the behavior of the initial Hol-Scherer relaxation (5.1.6) and criterion (5.1.3), which are also of independent interest. While we state all statements only for the initial Hol-Scherer relaxation, by Theorem 5.1.11 they are also valid for the solitary criterion (5.1.3) and its refinements (5.1.4) and (5.1.5). Throughout the chapter, we mostly use the reformulations of the initial step given in Theorem 5.1.11 when we refer to the initial Hol-Scherer relaxation.

Theorem 5.1.12. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. The initial Hol-Scherer relaxation* (5.1.6) *is*

(1) *invariant under translation.*

(2) *invariant under non-singular congruence transformation of $A(x)$.*

(3) *invariant under non-singular congruence transformation of $B(x)$.*

(4) *transitive.*

In particular, statements (2) and (3) include invariance under orthogonal transformation.

For reasons of clarity, we prove the statements in the theorem separately.

The subsequent statement shows that the initial Hol-Scherer relaxation (5.1.6) is invariant under translation. Let S_A be a spectrahedron defined by the linear pencil $A(x) = A_0 +$

$\sum_{p=1}^{d} x_p A_p$. To translate S_A by a vector $v = (v_1, \ldots, v_d)$ we substitute $x - v$ into the pencil $A(x)$,

$$A(x-v) = A_0 - \sum_{p=1}^{d} v_p A_p + \sum_{p=1}^{d} x_p A_p.$$

Lemma 5.1.13 (Translation symmetry)**.** *The initial Hol-Scherer relaxation* (5.1.6) *is invariant under translation.*

Proof. Given linear pencils $A(x)$ and $B(x)$, let C be a solution to system (5.1.3). Then it is also a solution for the translated pencils $A(x-v)$ and $B(x-v)$ for any $v \in \mathbb{R}^d$. Since the translation only has an impact on the constant matrix, we only have to show

$$B_0 - \sum_{p=1}^{d} v_p B_p - \left(\sum_{i,j=1}^{k} \left((A_0)_{ij} - \sum_{p=1}^{d} v_p (A_p)_{ij} \right) C_{ij} \right) = 0. \tag{5.1.9}$$

Since $B_p = \sum_{i,j=1}^{k} (A_p)_{ij} C_{ij}$ for all $p \in [d]$, (5.1.9) is equivalent to $B_0 - \sum_{i,j=1}^{k} (A_0)_{ij} C_{ij} = 0$, which is the condition on the constant matrices before translating. □

Lemma 5.1.14 (Non-singular congruence invariance)**.** *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. The initial Hol-Scherer relaxation is invariant under non-singular congruence transformation of $A(x)$ respectively $B(x)$. To be more precise,*

(1) *$B(x) \in \mathrm{QM}_0^l(A)$ if and only if $B(x) \in \mathrm{QM}_0^l\left(M^T A(x) M\right)$ for any $M \in \mathrm{GL}_k(\mathbb{R})$.*

(2) *$B(x) \in \mathrm{QM}_0^l(A)$ if and only if $M^T B(x) M \in \mathrm{QM}_0^l(A)$ for any $M \in \mathrm{GL}_l(\mathbb{R})$*

Proof.

To (1): Let $M \in \mathrm{GL}_k(\mathbb{R})$. Then the linear condition in part (2) of Theorem 5.1.11 reads as

$$B_p = \left(\left\langle M A_p M^T, C'_{st} \right\rangle \right)_{s,t=1} = \left(\left\langle A_p, M^T C'_{st} M \right\rangle \right).$$

Set $C'' = (M^T C'_{st} M)_{s,t=1}^{l} = (M \oplus \cdots \oplus M)^T C' M \oplus \cdots \oplus M$. Since $M^T \oplus \cdots \oplus M^T \in \mathrm{GL}_{kl}(\mathbb{R})$, C'' is positive semidefinite if and only if C' is positive semidefinite, implying the claim.

To (2): Let $M \in \mathrm{GL}_l(\mathbb{R})$. Then the linear condition in system (5.1.3) reads as

$$M^T B_p M = \sum_{i,j=1}^{k} (A_p)_{ij} C_{ij} \iff B_p = \sum_{i,j=1}^{k} (A_p)_{ij} (M^T)^{-1} C_{ij} M^{-1}.$$

Define $C'' = ((M^T)^{-1} C_{ij} M^{-1})_{i,j=1}^{k}$. The claim follows as in part (1). □

Next we discuss the transitivity behavior of the Hol-Scherer relaxation. This completes the proof of Theorem 5.1.12.

Theorem 5.1.15 (Transitivity)**.** *Let $E \in \mathcal{S}^e[x]$, $F \in \mathcal{S}^f[x]$, and $G \in \mathcal{S}^g[x]$ be linear pencils such that $S_E \subseteq S_F \subseteq S_G$ holds. Assume $F \in \mathrm{QM}_s^f(E)$ and $G \in \mathrm{QM}_t^g(F)$ for some nonnegative integers s, t.*

(1) *$G \in \mathrm{QM}_u^g(E)$ for some $s + t \geq u \geq \max\{s,t\}$.*

(2) *$u = \max\{s,t\}$ if and only if $F \in \mathrm{QM}_0^f(E)$ or $G \in \mathrm{QM}_0^g(F)$.*

(3) *If $F \in \mathrm{QM}_0^f(E)$ and $G \in \mathrm{QM}_0^g(F)$, then $G \in \mathrm{QM}_0^g(E)$.*

Proof. Consider the sos decompositions $F = S_0 + \langle S, E\rangle_f$ and $G = T_0 + \langle T, F\rangle_g$ with $S_0 \in \Sigma_s^f[x], S = (S_{kl})_{k,l=1}^f \in \Sigma_s^{ef}[x]$ and $T_0 \in \Sigma_t^g[x], T = (T_{ij})_{i,j=1}^g \in \Sigma_t^{fg}[x]$. Define $U = (U_{ij})_{i,j=1}^g$ by $U_{ij} = \sum_{k,l=1}^f S_{kl}(T_{ij})_{k,l}$ and $U_0 = T_0 + \langle T, S_0\rangle_g$.

We show $U \in \Sigma^{eg}[x]$. Note that after simultaneous permutation of rows and columns, T can be written as $T' = (T'_{kl})_{k,l=1}^f \in \Sigma^{fg}[x]$ with $(T'_{kl})_{i,j} = (T_{ij})_{k,l}$. Thus $U = \sum_{k,l=1}^f S_{kl} \otimes T'_{kl}$. Since $S \in \Sigma^{ef}[x]$, there exists a polynomial matrix $V \in \mathbb{R}^{ef\times m}[x]$ for some positive integer m such that $S = VV^T$. Clearly, the entries in V have degree at most $\lfloor s/2\rfloor$. Decompose V into blocks $V_1, \ldots, V_f \in \mathbb{R}^{e\times m}[x]$. Then $S_{kl} = V_k V_l^T$. Similarly, $T' \in \Sigma^{fg}[x]$ implies $T' = WW^T$ with $T'_{kl} = W_k W_l^T$ for some polynomial matrix $W = [W_1, \ldots, W_f]^T \in \mathbb{R}^{fg\times n}[x]$ and some positive integer n. Then

$$U = \sum_{k,l=1}^f S_{kl} \otimes T'_{kl} = \sum_{k,l=1}^f V_k V_l^T \otimes W_k W_l^T = \sum_{k,l=1}^f (V_k \otimes W_k)(V_l \otimes W_l)^T$$

is a sum of squares $eg\times eg$-matrix, where we used the mixed-product property of the Kronecker product [Lau05, Theorem 13.3].

From the construction it follows immediately that the degree of U is bounded by

$$\deg(S) + \deg(T) = s + t \geq \deg(U) \geq \max\{s,t\} = \max\{\deg(S), \deg(T)\}.$$

If S or T is of degree one or higher, i.e., at least one entry is of degree one or higher, then by the definition of U every block in U has degree at least one. Obviously, if S and T are constant (i.e., $s = t = 0$), then $u = 0 = \max\{s,t\}$.

$U_0 \in \Sigma^g[x]$ can be proved in a similar way.

Using the sos decomposition of F, we get

$$\begin{aligned}
U_0 + \langle U, E\rangle_g &= U_0 + \left(\langle U_{ij}, E\rangle\right)_{i,j=1}^g \\
&= U_0 + \left(\sum_{k,l=1}^f (T_{ij})_{k,l}\, \langle S_{kl}, E\rangle\right)_{i,j=1}^g \\
&= T_0 + \langle T, S_0\rangle_g + \left(\left\langle T_{ij}, \langle S, E\rangle_f\right\rangle\right)_{i,j=1}^g \\
&= T_0 + \langle T, S_0\rangle_g + \langle T, F - S_0\rangle_g \\
G &= T_0 + \langle T, F\rangle_g .
\end{aligned}$$

□

If a pencil has a block structure, we can use this to reduce the number of variables in the initial Hol-Scherer relaxation (5.1.6).

Proposition 5.1.16 ([HKM13, Propositions 4.1 and 4.2])**.** *Let $A(x) \in \mathcal{S}^k[x]$, $B(x) \in \mathcal{S}^l[x]$ and $G^q(x) \in \mathcal{S}_{d_q}[x]$ be linear pencils with $G^q(x) = G_0^q + \sum_{p=1}^d x_p G_p^q$ for $q \in [m]$.*

(1) *Assume $B(x) = \bigoplus_{q=1}^m G^q(x)$ is the direct sum with $l = \sum_{q=1}^m d_q$. Then $B(x) \in \mathrm{QM}_0^l(A)$ if and only if $G^q(x) \in \mathrm{QM}_0^{d_q}(A)$ for $q \in [m]$, i.e.,*

$$\exists C^q = (C_{ij}^q)_{i,j=1}^k \succeq 0 : \quad G_p^q = \sum_{i,j=1}^k (A_p)_{ij} C_{ij}^q \ \forall p \in [0,d]. \tag{5.1.10}$$

(2) *Assume $A(x) = \bigoplus_{q=1}^m G^q(x)$ is the direct sum with $k = \sum_{q=1}^m d_q$. Then $B(x) \in \mathrm{QM}_0^l(A)$*

if and only if $B(x) \in \mathrm{QM}_0^l(G^q)$ *for* $q \in [m]$, *i.e.,*

$$\exists C^q = (C^q_{ij})_{i,j=1}^{d_q} \succeq 0: \quad B_p = \sum_{q=1}^{m} \sum_{i,j=1}^{d_q} (G^q_p)_{ij} C^q_{ij} \ \forall p \in [0,d]. \tag{5.1.11}$$

(Note that $C = \bigoplus_{q=1}^m C^q \succeq 0$ *is a block diagonal matrix and each* C^q_{ij} *is of size* $l \times l$*.)*

Since [HKM13] does not contain a proof of statement (1), we provide a short one. The proof of statement (2) is given in [HKM13, Proposition 4.1].

Proof. Let $C^1, \dots, C^m$ be solutions to (5.1.10), and set $C = \bigoplus_{q=1}^m C^q$. Define C' as the direct sum of blocks of C, $C'_{ij} = \bigoplus_{q=1}^m C^q_{ij}$. Then C' is a solution to (5.1.3). C' results by simultaneously permuting rows and columns of C and is thus positive semidefinite. We have $B_p = \bigoplus_{q=1}^m D^q_p = \bigoplus_{q=1}^m \sum_{i,j=1}^k (A_p)_{ij} C^q_{ij} = \sum_{i,j=1}^k (A_p)_{ij} C'_{ij}$.

Conversely, let C' be a solution to (5.1.3). We are interested in the m diagonal submatrices of each block C'_{ij}, defined as follows. For $q \in [m]$, let C'^q_{ij} be the $d_q \times d_q$-submatrix of C'_{ij} with row and column indices $\{\sum_{r=1}^{q-1} d_r + 1, \dots, \sum_{r=1}^{q} d_r\}$. Now the submatrix $C^q = (C'^q_{ij})_{i,j=1}^k$ consisting of the qth diagonal blocks of each matrix C'_{ij} is a solution to (5.1.10). C^q is a principal submatrix of C' and thus positive semidefinite. The equations in (5.1.10) are a subset of the equations in (5.1.3) and remain valid. □

5.2 Certificates for Containment of $\mathcal{H}$-Polyhedra and Spectrahedra

Specializing the Spectrahedron Containment problem to the case where one of the linear pencils involved defines an $\mathcal{H}$-polyhedron, allows to reduce the complexity of the initial Hol-Scherer relaxation (5.1.6) in the sense that the number of variables in the semidefinite feasibility system is reduced. If the outer set is an $\mathcal{H}$-polyhedron and some mild assumptions are fulfilled, the containment question can be decided with the initial relaxation step.

If a family of symmetric matrices $B_0, B_1, \dots, B_d \in \mathcal{S}^l$ commutes (pairwise), then the matrices are simultaneously diagonalizable, i.e., there exists an orthogonal matrix $Q \in O_l(\mathbb{R})$ such that for $p \in [0,d]$ there is some diagonal matrix $D_p \in \mathcal{S}^l$ with $B_p = Q^T D_p Q$. Since $O_l(\mathbb{R}) \subset \mathrm{GL}_l(\mathbb{R})$, this implies that the spectrahedron S_B is a polyhedron. By the invariance Theorem 5.1.12, Corollary 5.1.4 on the containment of $\mathcal{H}$-polyhedra can be generalized from pencils in normal form to pencils with coefficient matrices simultaneously congruent to a diagonal matrix.

Lemma 5.2.1. *Let* $A(x) \in \mathcal{S}^k[x]$ *and* $B(x) \in \mathcal{S}^l[x]$ *be linear pencils. Assume* $A_0, \dots, A_d$ *and* $B_0, \dots, B_d$ *are two families of matrices simultaneously congruent to a diagonal matrix (i.e.,* S_A *and* S_B *are polyhedra).*

(1) $S_A \subseteq S_B$ *if and only if* $B(x) \in \mathrm{QM}_0^l(A)$.

(2) *Assume, in addition,* S_A *is bounded and not a singleton.* $S_A \subseteq S_B$ *if and only if* $B(x) \in \mathrm{QM}_0^l(A)$ *with* $S_0 = 0$.

In particular, if all coefficients are diagonal matrices, membership of $B(x)$ *in the truncated quadratic module* $\mathrm{QM}_0^l(A)$ *is equivalent to the inclusion* $S_A \subseteq S_B$.

Proof. Since $A_0, \dots, A_d$ are simultaneously congruent to a diagonal matrix, there exists a non-singular matrix $Q \in \mathrm{GL}_k(\mathbb{R})$ and diagonal matrices $D_0, \dots, D_d$ such that $Q^T A_i Q = D_i$

for $i \in [0,d]$. Applying the same to $B(x)$, yielding a diagonal pencil $E(x) := \sum_{p=1}^d x_p E_p$, Lemma 5.1.14 implies the equivalence of $B(x) \in \mathrm{QM}_0^l(A)$ and $E(x) \in \mathrm{QM}_0^l(D)$, where $D(x) := D_0 + \sum_{p=1}^d x_p D_p$. Thus the two statements reduce to diagonal pencils and we are in situation of Corollary 5.1.4, implying the claims. □

As we have seen in Section 5.1.1, for polyhedra in normal form (2.3.2) there is a diagonal solution to (5.1.6). Thus it is sufficient to check the feasibility of the restriction of (5.1.6) to the diagonal and checking inclusion of polyhedra reduces to a linear program. Note that for rational polyhedra in normal form, the certificate is again rational. For unbounded polyhedra, the term $S_0(x)$ (resp. the extended normal form of $A(x)$) is required in order for the criterion to be exact; see Example 5.1.2.

As every symmetric matrix is diagonalizable, an immediate consequence of Lemma 5.2.1 is the somewhat artificial statement on univariate linear pencils defining unbounded intervals.

Corollary 5.2.2. *Let $A(x) = x_1 A_1 \in \mathcal{S}^k[x]$ and $B(x) = x_1 B_1 \in \mathcal{S}^l[x]$. Then $S_A \subseteq S_B$ if and only if $B(x) \in \mathrm{QM}_0^l(A)$.*

5.2.1 Spectrahedron in $\mathcal{H}$-Polyhedron

Helton, Klep, and McCullough [HKM13, Proposition 5.3] showed that the containment criterion (5.1.3) is exact in an important case, namely if S_B is the cube given by the monic linear pencil

$$B(x) = I_{2d} + \frac{1}{r} \sum_{p=1}^{d} x_p \left(E_{p,p} - E_{d+p,d+p}\right). \qquad (5.2.1)$$

The goal of this section is to generalize this to all polyhedra S_B given by a linear pencil with coefficient matrices simultaneously congruent to a diagonal matrix, particularly for polyhedra in normal form (2.3.2).

Theorem 5.2.3. *Let $A(x) \in \mathcal{S}^k[x]$ be a strictly feasible linear pencil and let the coefficients of the linear pencil $B(x) \in \mathcal{S}^l[x]$ be simultaneously congruent to a diagonal matrix.*

(1) *$S_A \subseteq S_B$ if and only if $B(x) \in \mathrm{QM}_0^l(A)$.*

(2) *Assume S_B is a polytope with nonempty interior. Then $S_A \subseteq S_B$ if and only if $B(x) \in \mathrm{QM}_0^l(A)$ with $S_0 = 0$.*

In particular, the statements hold for a diagonal linear pencil $B(x)$, i.e., a polyhedron in normal form (2.3.2).

This theorem can also be deduced from results of Klep and Schweighofer in [KS11]. A linear scalar-valued polynomial is positive on a spectrahedron if and only if it is positive on the matricial version of the spectrahedron.

In order to prove this statement (where the sufficiency-parts are clear from the construction of the hierarchy (5.1.6)), we use some of the auxiliary results on the behavior of the criterion with regard to block diagonalization and transitivity proved in Section 5.1.3.

Proof. Suppose $B_0, B_1, \ldots, B_d \in \mathcal{S}^l$ are simultaneously congruent to a diagonal matrix, i.e., there exists a non-singular matrix $Q \in \mathrm{GL}_l(\mathbb{R})$ such that $Q^T B_p Q = D_p$ is a diagonal matrix for all $p \in [0,d]$. Since the linear pencils $B(x)$ and $Q^T B(x) Q$ have the same positivity domain, the spectrahedron is polyhedral. By Theorem 5.1.12, the initial Hol-Scherer step (5.1.3) certifies containment with respect to $B(x)$ and $A(x)$ if and only if it does when considering the pencil $D(x)$ instead of $B(x)$. Thus w.l.o.g. let $B(x)$ be diagonal.

Let $B(x) = \bigoplus_{q=1}^{l} b^q(x) \in \mathcal{S}^l[x]$ with $b^q(x) = b_0^q + x^T b^q$ for $q \in [l]$ be the normal form of a polyhedron (2.3.2). Assume w.l.o.g. that the $\mathcal{H}$-representation of S_B is reduced, i.e., S_B has l facets and each b^q is an inner normal vector of one of the facets. Denote by $b_0^q, b_1^q, \ldots, b_d^q$ the coefficients of the linear form $b^q(x) = (b_0 + Bx)_q$. Set $b^q := (b_1^q, \ldots, b_d^q)$.

Proposition 5.1.16 implies that the initial Hol-Scherer relaxation is feasible if and only if the system

$$C^q = (C_{ij})_{i,j=1}^k \succeq 0, \quad \forall p \in [0,d]: \; b_p^q = \sum_{i,j=1}^k (A_p)_{ij} C_{ij}^q$$

is feasible for all $q \in [l]$. Note that C^q is in $\mathcal{S}^k$. Hence, the system has the form

$$C^q = (C_{ij})_{i,j=1}^k \succeq 0, \quad \forall p \in [0,d]: \; b_p^q = \langle A_p, C^q \rangle . \tag{5.2.2}$$

We show the existence of a solution by duality theory of semidefinite programming and transitivity of the criterion; see Theorem 5.1.15. For $q \in [l]$, consider the semidefinite program

$$\begin{aligned} r^q := \max \; & \langle -b^q, x \rangle \\ \text{s.t. } & A(x) \succeq 0. \end{aligned} \tag{5.2.3}$$

By assumption, the primal problem (5.2.3) is strictly feasible and the optimal value is finite. Hence, by Proposition 2.3.15, the dual problem

$$\begin{aligned} \min \; & \langle A_0, Y^q \rangle \\ \text{s.t. } & \langle A_p, Y^q \rangle = b_p^q \quad \text{for } p \in [d] \\ & Y^q \succeq 0 \end{aligned} \tag{5.2.4}$$

has the same optimal value r^q and attains it. (Note that by duality $\langle -A_p, Y^q \rangle = -b_p^q$.) Consequently, for all $q \in [l]$ there exists a symmetric $k \times k$-matrix C^q such that

$$C^q \succeq 0, \quad r^q = \langle A_0, C^q \rangle , \quad b_p^q = \langle A_p, C^q \rangle .$$

As mentioned before (5.2.2), the matrices C^q certify the inclusion $S_A \subseteq S_{B'}$, where $B'(x)$ is defined as the linear pencil

$$B'(x) = \bigoplus_{q=1}^{l} \left(r^q + \sum_{p=1}^{d} b_p^q x_p \right) .$$

Now we have to distinguish between the two cases in the statement of the theorem.

First consider the case that S_B is a polytope containing the origin in its interior. Since $B(x)$ is in normal form, we have $\max_{x \in S_B} \langle -b^q, x \rangle = b_0^q$. Further, since $S_A \subseteq S_B$, the definition of r^q implies $r^q \leq b_0^q$ and hence $S_{B'} \subseteq S_B$. By transitivity and by exactness of the criterion for polytopes, see Theorem 5.1.15 and Lemma 5.2.1, respectively, there is a initial Hol-Scherer certificate for the inclusion $S_A \subseteq S_B$.

To prove the unbounded case in the theorem, we construct a certificate (5.1.6) for the inclusion $S_{\widehat{A}} \subseteq S_{\widehat{B'}}$, where $\widehat{B'}(x) = 1 \oplus B'(x)$ denotes the extended normal form (5.1.2). Then the claim follows by Lemma 5.2.1 and Theorem 5.1.15 as above.

First note that $S_{\widehat{A}} \subseteq S_B$ is equivalent to $S_A \subseteq S_B$. Denote by C' the matrix that certifies the inclusion $S_A \subseteq S_{B'}$. Then the symmetric $(k+1)(l+1) \times (k+1)(l+1)$-matrix

$$\widehat{C} := E_{11} \oplus \begin{bmatrix} 0 & 0 \\ 0 & C'_{ij} \end{bmatrix}_{i,j=1}^{k},$$

where E_{11} and the blocks $\begin{bmatrix} 0 & 0 \\ 0 & C'_{ij} \end{bmatrix}_{i,j=1}^{k}$ are of size $(l+1) \times (l+1)$, certifies the inclusion $S_{\widehat{A}} \subseteq S_{\widehat{B'}}$. Indeed, adding zero-columns and zero-rows simultaneously preserves positive semidefiniteness and, clearly, the sum of the diagonal blocks of $\widehat{C}$ equals the extended constant term $1 \oplus B_0$. Since in every $\widehat{A}_p$ the first column and the first row are the zero vector, we get

$$\sum_{i,j=0}^{k} (\widehat{A}_p)_{ij}\, \widehat{C}_{ij} = 0 \cdot E_{11} + \begin{bmatrix} 0 & 0 \\ 0 & \sum_{i,j=1}^{k} a_{ij}^{p} C'_{ij} \end{bmatrix} = \widehat{B'}_p. \qquad \square$$

If the description of the polyhedron in Theorem 5.2.3 cannot be transformed to a pencil in normal form, the situation changes. Recall that there exist representations of polyhedra, called $\mathcal{S}$-representations (cf. Proposition 2.3.1), for which it is not possible to achieve a normal form by simultaneous congruence transformation. As mentioned in Section 2.3, the polyhedrality recognition problem (PRP) is an NPH problem. Therefore, we cannot expect to deduce a normal form representation from a general $\mathcal{S}$-representation of a polyhedron. However, if polyhedrality of the set S_B is known a priori, Proposition 2.3.1 allows to reduce the problem to a blockdiagonal pencil. Application of Proposition 5.1.16 yields then a smaller problem.

5.2.2 $\mathcal{H}$-Polyhedron in Spectrahedron – A Note on the Polytope Recognition Problem

Let $A(x) \in \mathcal{S}^k$ be a linear pencil such that S_A is a polyhedron containing the origin. By Proposition 2.3.1, there exists a non-singular matrix $M \in \mathrm{GL}_k(\mathbb{R})$ and linear pencils $D(x) \in \mathcal{S}^m$ and $S(x) \in \mathcal{S}^{k-m}$ for some $k \geq m \geq 1$ such that $D(x)$ is diagonal and

$$M^T A(x) M = D(x) \oplus S(x) \text{ and } S_A = S_D.$$

Thus the pencil $D(x)$ is another representation of the spectrahedron S_A. Consider another pencil $B(x) \in \mathcal{S}^l$. Using Proposition 5.1.16 and Lemma 5.1.14, system (5.1.3) can be reduced to test whether

$$C^1, \ldots, C^m \in \mathcal{S}_+^l,\ \forall p \in [0,d]:\ B_p = \sum_{q=1}^{m} D_p^q C^q$$

has a solution, where $D^q(x) = D_0^q + \sum_{p=1}^{d} x_p D_p^q$ with $D_p^q \in \mathbb{R}$ – a decrease in the number of variables from $\frac{1}{2}kl(kl+1)$ to $\frac{1}{2}ml(l+1)$. Note that this can only be done if a priori the polyhedrality of S_A is known.

However, even if polyhedrality of S_A is unknown, the approach based on Proposition (2.3.1) still yields a reduction to the SDFP

$$C^1, \ldots, C^m \in \mathcal{S}_+^l,\ C^{m+1} \in \mathcal{S}_+^{(k-m)l},$$
$$\forall p \in [0,d]:\ B_p = \sum_{q=1}^{m} D_p^q C^q + \sum_{i,j=1}^{k-m} (S_p)_{ij} C^{m+1}.$$

The number of variables decreases from $\frac{1}{2}kl(kl+1)$ to $\frac{1}{2}l\left(l(k-m)^2+ml+k\right)$.

If a priori polyhedrality of S_A is unknown and one wants to solve the polytope recognition problem (PRP), the problem can be simplified by using Proposition (2.3.1). S_A is a polyhedron if and only if $S_D \subseteq S_S$. (Note that the reverse inclusion is clear by construction; see [BRS11].) Applying Proposition 5.1.16, the semidefinite feasibility problem

$$\exists C^1,\ldots,C^m \in \mathcal{S}_+^{k-m}:\ S_p=\sum_{q=1}^{m} D_p^q C^q \quad \forall p\in[0,d]$$

yields a sufficient criterion for polyhedrality of S_A.

5.3 Containment Certificates for Some Structured Classes

It turns out that the initial Hol-Scherer relaxation (5.1.6) even provide containment certificates in several structured cases. Detailed statements of these results and their proofs will be given in Lemma 5.3.2, Lemma 5.3.4, and in Section 5.4. A summary is given in Theorem 5.3.1.

For ease of notation, most statements in this section are given for monic pencils. While the statements are only stated for the initial Hol-Scherer relaxation, all the results are valid for the solitary criterion (5.1.3) and its refinement (5.1.4). In order to prove the statements, we use the equivalent descriptions of the initial step stated in Theorem 5.1.11.

Theorem 5.3.1. *Let $A(x)\in\mathcal{S}^k[x]$ and $B(x)\in\mathcal{S}^l[x]$ be monic linear pencils. In the following cases, containment $S_A\subseteq S_B$ is certified by the initial relaxation step* (5.1.6)*.*

(1) *$A(x)$ and $B(x)$ are normal forms of ellipsoids* (2.3.14) *(both centrally symmetric with axis-aligned semi-axes).*

(2) *$A(x)$ is the normal form of an ellipsoid and the coefficients of $B(x)$ are simultaneously congruent to a diagonal matrix.*

(3) *$A(x)=A_0+x_1A_1+x_2A_2\in\mathcal{S}^2[x]$ with $\operatorname{span}\{A_0,A_1,A_2\}=\mathcal{S}^2$, and S_A is a nonempty spectratope.*

In this section, we provide the proofs of (1) and (2), where the sufficiency parts follow from the construction of the Hol-Scherer relaxation and by Theorem 5.1.3. Case (3) is treated in Section 5.4.

5.3.1 Some Exact Cases

The following statement on ellipsoids uses the normal form (2.3.14).

Lemma 5.3.2. *Let two ellipsoids S_A and S_B be given by the normal forms*

$$A(x)=I_{d+1}+\sum_{p=1}^{d}\frac{x_p}{a_p}(E_{p,d+1}+E_{d+1,p}) \text{ and } B(x)=I_{d+1}+\sum_{p=1}^{d}\frac{x_p}{b_p}(E_{p,d+1}+E_{d+1,p}),$$

respectively. Here $(a_1,\ldots,a_d)>0$ and $(b_1,\ldots,b_d)>0$ are the vectors of the length of the semi-axes. Then $S_A\subseteq S_B$ if and only if $B(x)\in\mathrm{QM}_0^l(A)$.

Proof. Note first that $k = l = d + 1$. It is obvious that $S_A \subseteq S_B$ if and only if $b_p - a_p \geq 0$ for every $p \in [d]$. The matrices underlying the matrix pencils $A(x)$ and $B(x)$ are

$$A_p = \frac{1}{a_p}(E_{p,d+1} + E_{d+1,p}) \text{ and } B_p = \frac{1}{b_p}(E_{p,d+1} + E_{d+1,p})$$

for all $p \in [d]$. Now define an $(d+1)^2 \times (d+1)^2$-block matrix C by

$$(C_{i,j})_{s,t} = \begin{cases} 1 & i = j = s = t, \\ \frac{a_j}{b_j} & i = s = d+1,\ j = t \leq d, \\ \frac{a_i}{b_i} & i = s \leq d,\ j = t = d+1, \\ \frac{a_i\, a_j}{b_i\, b_j} & i = s \leq d,\ j = t \leq d,\ i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

We show that C is a solution to (5.1.3). Decompose $x \in \mathbb{R}^{(d+1)^2}$ in blocks of length $d+1$ and write $x_{i,j}$ for the jth entry in the ith block. The matrix C is positive semidefinite since

$$\begin{aligned} x^T C x &= \sum_{i=1}^{d+1} x_{i,i}^2 + 2 \sum_{i<j\leq d} \frac{a_i\, a_j}{b_i\, b_j} x_{i,i} x_{j,j} + 2 \sum_{i=1}^{d} \frac{a_i}{b_i}\, x_{i,i}\, x_{d+1,d+1} \\ &= \left(\sum_{i=1}^{d} \frac{a_i}{b_i}\, x_{i,i} + x_{d+1,d+1} \right)^2 + \sum_{i=1}^{d} \left(1 - \frac{a_i^2}{b_i^2} \right) x_{i,i}^2 \geq 0 \end{aligned}$$

for all $x \in \mathbb{R}^{(d+1)^2}$. Clearly, the sum of the diagonal blocks is the identity matrix I_{d+1}. Since every A_p has only two non-zero entries, every B_p is a linear combination of only two blocks of C,

$$B_p = \frac{1}{a_p} C_{d+1,p} + \frac{1}{a_p} C_{p,d+1}.$$

This equality is true by the definition of C. $\square$

Remark 5.3.3. *Using the square matrices $E_{i,j}$ of size $(d+1)\times(d+1)$ introduced in Chapter 2, the matrix C in the proof of Lemma 5.3.2 has the form*

$$\begin{bmatrix} E_{1,1} & f_{1,2}E_{1,2} & \cdots & f_{1,d}E_{1,d} & \frac{a_1}{b_1}E_{1,d+1} \\ f_{2,1}E_{2,1} & E_{2,2} & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & f_{d-1,d}E_{d-1,d} & \vdots \\ f_{d,1}E_{d,1} & \cdots & f_{d,d-1}E_{d,d-1} & E_{d,d} & \frac{a_d}{b_d}E_{d,d+1} \\ \frac{a_1}{b_1}E_{d+1,1} & \frac{a_2}{b_2}E_{d+1,2} & \cdots & \frac{a_d}{b_d}E_{d+1,d} & E_{d+1,d+1} \end{bmatrix} \in \mathcal{S}^{(d+1)^2},$$

where $f_{i,j} := \frac{a_i\, a_j}{b_i\, b_j}$.

Now we prove exactness of the initial Hol-Scherer relaxation (5.1.6) for the containment of an ellipsoid in an $\mathcal{H}$-polyhedron. Our result slightly generalizes [KTT13, Lemma 4.13].

Lemma 5.3.4. *Let S_A be an ellipsoid given by the linear pencil*

$$A(x) = \begin{bmatrix} A & 0 \\ 0 & r \end{bmatrix} + \sum_{p=1}^{d} x_p (E_{p,d+1} + E_{d+1,p})$$

with $A \in \mathcal{S}_{++}^d$ and $r > 0$. Let $B(x)$ be a strictly feasible linear pencil whose coefficient *matrices are simultaneously congruent to a diagonal matrix (i.e., S_B is an $\mathcal{H}$-polyhedron). Then $S_A \subseteq S_B$ if and only if $B(x) \in \mathrm{QM}_0^l(A)$.*

Proof. We first reduce the problem using Theorem 5.1.12 to the case of monic linear pencils.

Since $B_0, \dots, B_d$ are simultaneously congruent to a diagonal matrix, i.e., there exist a non-singular matrix $Q \in GL_l(\mathbb{R})$ and diagonal matrices $D_0, \dots, D_d$ such that $Q^T B_i Q = D_i$ for $i \in [0, d]$, Theorem 5.1.12 implies that $B(x) \in \mathrm{QM}_0^l(A)$ is equivalent to $D(x) \in \mathrm{QM}_0^l(A)$, where $D(x) := D_0 + \sum_{p=1}^d x_p D_p$. As $B(x)$ is strictly feasible, (possibly after an affine translation) $D_0 \succ 0$. Thus w.l.o.g. assume $B(x)$ to be the normal form of an $\mathcal{H}$-polyhedron (2.3.2) with $B_0 = I_l$.

Since $A \in \mathcal{S}_{++}^d$, there exists a non-singular matrix $Q \in \mathrm{GL}_d(\mathbb{R})$ such that $QAQ^T = I_d$. By applying Theorem 5.1.12, the membership problem $B(x) \in \mathrm{QM}_0^l(A)$ can be reduced to $B(x) \in \mathrm{QM}_0^l(A')$, where

$$A'(x) = \begin{bmatrix} I_d & S^{-1}Qx \\ (S^{-1}Qx)^T & r \end{bmatrix}.$$

After a coordinate transformation, $x = S^{-1}Qx$, we end up with the question whether the r-scaled unit ball is contained in the polyhedron P_B (in the new coordinate system). Therefore, assume $A(x)$ to be the normal form of a ball (2.3.14) of radius r, i.e., $S_A = \mathbb{B}_r(0)$.

Since $B(x)$ is monic, the linear polynomials describing S_B are of the form $b_i(x) = 1 + \sum_{p=1}^d b_{i,p} x_p$ for $i \in [l]$. If $S_{A'} \subseteq S_B$, we have $S_B^\circ \subseteq S_{A'}^\circ = \mathbb{B}_{1/r}(0)$. As $-b_i \in S_B^\circ$ for $i \in [l]$, we have $b_i^T b_i = \| - b_i\|_2^2 \le 1/r^2$.

We give a feasible matrix C to system (5.1.3) to show exactness of the criterion. In this case, C is an $(d+1)l \times (d+1)l$-block matrix defined as follows:

$$(C_{i,j})_{s,t} = \begin{cases} \frac{r^2 b_{s,i}^2}{2} & i = j < k,\ s = t, \\ 1 - \frac{r^2}{2}\sum_{p=1}^d b_{s,i}^2 & i = j = k,\ s = t, \\ \frac{r^2 b_{s,i} b_{s,j}}{2} & i < k,\ j < k,\ i \neq j,\ s = t, \\ \frac{r\, b_{s,j}}{2} & i = k,\ j < k,\ s = t, \\ \frac{r\, b_{s,i}}{2} & j = k,\ i < k,\ s = t, \\ 0 & \text{otherwise.} \end{cases}$$

To show positive semidefiniteness of C, consider a vector $x \in \mathbb{R}^{(d+1)l}$. Decompose x into blocks of length l, and we write $x_{i,j}$ for the jth entry in the ith block. Now C is positive semidefinite since

$$\begin{aligned} x^T C x = \sum_{s=1}^l \Bigg[& \sum_{i=j<k} x_{i,s}^2 \frac{r^2 b_{s,i}^2}{2} + x_{k,s}^2 \left(1 - \frac{r^2}{2}\sum_{p=1}^d b_{s,p}^2\right) \\ & + 2 \sum_{i<j<k} x_{i,s} x_{j,s} \frac{r^2 b_{s,i} b_{s,j}}{2} + 2 \sum_{i<k} x_{i,s} x_{k,s} \frac{r\, b_{s,i}}{2} \Bigg] \\ = \sum_{s=1}^l \Bigg[& \left(\sum_{i=j<k} x_{i,s} \frac{r\, b_{s,i}}{\sqrt{2}} + \frac{x_{k,s}}{\sqrt{2}} \right)^2 + \frac{1}{2} x_{k,s}^2 \left(1 - r^2 \sum_{p=1}^d b_{s,p}^2\right) \Bigg] \ge 0 \end{aligned}$$

for all $x \in \mathbb{R}^{(d+1)l}$. The term $1 - r^2 \sum_{p=1}^d b_{s,p}^2$ is nonnegative since the ball of radius r is contained in S_B and therefore $\frac{1}{r^2} \ge \sum_{p=1}^d b_{s,p}^2$. By construction, the sum of the diagonal blocks is the identity matrix I_l. Every B_p is a linear combination of only two blocks of C,

$$B_p = \frac{1}{r} C_{d+1,p} + \frac{1}{r} C_{p,d+1}. \qquad \square$$

Observe that in Lemmas 5.3.2 and 5.3.4, for rational input, C is rational as well.

5.3.2 Containment of Scaled Spectrahedra

Surprisingly, certifying containment can always be achieved by an appropriate scaling of one of the spectrahedra involved, leading to an optimization version of the initial step (5.1.6) and the solitary criterion (5.1.3). We start by revisiting Example 5.1.7.

Example 5.3.5 (Example 5.1.7 revisited)**.** *Let $A(x)$ and $B(x)$ be the linear pencil representations of the 2-dimensional unit ball as in Example* (5.1.7)*. Generalizing $A(x)$, let $A^r(x)$ be the linear pencil of the ball with radius $(1 >)r > 0$ in normal form. With regard to the containment question $S_{A^r} = rS_A \subseteq S_B$, we show the feasibility of system* (5.1.3) *for r sufficiently small.*

Claim. The containment problem $rS_A \subseteq S_B$ is certified by criterion (5.1.3) and by the initial Hol-Scherer relaxation (5.1.6) for $0 < r \leq \frac{1}{2}\sqrt{2}$.

Consider the matrix

$$C = \left[\begin{array}{cc|cc|cc} c & 0 & 0 & c & \frac{r}{2} & 0 \\ 0 & c & -c & 0 & 0 & -\frac{r}{2} \\ \hline 0 & -c & c & 0 & 0 & \frac{r}{2} \\ c & 0 & 0 & c & \frac{r}{2} & 0 \\ \hline \frac{r}{2} & 0 & 0 & \frac{r}{2} & 1-2c & 0 \\ 0 & -\frac{r}{2} & \frac{r}{2} & 0 & 0 & 1-2c \end{array}\right] \in \mathbb{R}^{6\times 6}.$$

Obviously the equality constraints in (5.1.3) *are fulfilled. If one of the diagonal entries is zero, i.e., $c = 0$ or $1 - 2c = 0$, then $r = 0$. Therefore, $0 < c < \frac{1}{2}$ and the 2×2-block in the top left corner C_{11} is positive definite. Thus the matrix C is positive semidefinite if and only if the Schur complement (see, e.g., [dK02, Theorem A.9]) with respect to C_{11} is positive semidefinite. This is the case if and only if*

$$1 - 2c - \frac{r^2}{4c} \geq 0 \;\Leftrightarrow\; f(c) := 8c^2 - 4c + r^2 \leq 0.$$

Assume $r > \frac{1}{2}\sqrt{2}$. Then $f(c) > 0$ for all c since f has no real roots and the constant term $f(0) = r^2$ is positive. Otherwise, $f(\frac{1}{4}) = -\frac{1}{2} + r^2 \leq 0$. Hence, system (5.1.3) *is feasible for $0 < r \leq \frac{1}{2}\sqrt{2}$.*

Alternatively, we can prove the positive semidefiniteness of C in the same manner as in the proofs of Lemmas 5.3.2 and 5.3.4. To do this, set $2\sqrt{c(1-2c)} = r$. Since $0 < r$ by assumption, the equation has real solutions in c if and only if $0 < r \leq \frac{1}{2}\sqrt{2}$. Equivalently, $c \in (0, \frac{1}{2})$. Let $x \in \mathbb{R}^6$. Then

$$\begin{aligned} x^T C x &= c\, x_1^2 + c\, x_4^2 + (1-2c)\, x_5^2 + 2c\, x_1 x_4 + r\, x_1 x_5 + r\, x_4 x_5 \\ &\quad + c\, x_2^2 + c\, x_3^2 + (1-2c)\, x_6^2 - 2c\, x_2 x_3 - r\, x_2 x_6 + r\, x_3 x_6 \\ &= \left(\sqrt{c}\, x_1 + \sqrt{c}\, x_4 + \sqrt{1-2c}\, x_5\right)^2 + \left(-\sqrt{c}\, x_2 + \sqrt{c}\, x_3 + \sqrt{1-2c}\, x_6\right)^2 \geq 0 \end{aligned}$$

for $c \in (0, \frac{1}{2})$.

The problem of maximizing r *such that containment is certified by* (5.1.3) *and* (5.1.6) *can be formulated as a semidefinite program. A numerical computation yields an optimal value of* $0.707 \approx \frac{1}{2}\sqrt{2}$. *Note that we are in the situation of Corollary 5.1.6. For the relaxed version* (5.1.5), *a numerical computation gives the optimal value of* $0.950 \approx \frac{19}{20}$. *In particular, this shows that the relaxed criterion* (5.1.5) *can be satisfied in cases where the non-relaxed criterion* (5.1.3) *does not certify an inclusion. It is an open research question to establish a quantitative relationship comparing criterion* (5.1.3) *to* (5.1.5) *in the general case.* □

Generalizing the observation from Example 5.3.5, we show that for two spectrahedra S_A and S_B containing the origin in their interior, there always exists some scaling factor ν such that the initial Hol-Scherer relaxation (5.1.6) and criterion (5.1.3) (and (5.1.4)) certify the inclusion $\nu S_A \subseteq S_B$. This extends the following result of Ben-Tal and Nemirovski, who treated containment of a cube in a spectrahedron.

For a linear pencil $A(x) \in \mathcal{S}^k[x]$ with positive definite constant term $A_0 \succ 0$ and a positive constant $\nu > 0$, define the *ν-scaled (linear) pencil* as

$$A^{\nu}(x) := A\left(\frac{x}{\nu}\right) = A_0 + \frac{1}{\nu}\sum_{p=1}^{d} x_p A_p. \tag{5.3.1}$$

Similarly, we denote by $\nu S_A := \{x \in \mathbb{R}^d \mid A^{\nu}(x) \succeq 0\}$ the corresponding *ν-scaled spectrahedron.*

Proposition 5.3.6 ([BTN02, Thm. 2.1])**.** *Let* S_A *be the cube in normal form* (5.2.1) *with edge length* $r > 0$ *and consider a linear pencil* $B(x)$ *with* $B_0 \succ 0$. *Let* $\mu = \max\{\operatorname{rank} B_p \mid p \in [d]\}$ *be the maximum rank of the coefficient matrices of* $B(x)$. *If* $S_A \subseteq S_B$, *then* $B(x) \in \mathrm{QM}_0^l(A^{\nu(\mu)})$ *with regard to the* $\nu(\mu)$*-scaled pencil* $A^{\nu(\mu)}$, *where* $\nu(\mu)$ *is given by*

$$\nu(\mu) = \min_{y \in \mathbb{R}^{\mu}, \|y\|_1 = 1} \left\{ \int_{\mathbb{R}^{\mu}} \left| \sum_{i=1}^{\mu} y_i u_i^2 \right| \left(\frac{1}{2\pi}\right)^{\frac{\mu}{2}} \exp\left(-\frac{u^T u}{2}\right) du \right\}$$

and $\|y\|_1 = \sum_{i=1}^{\mu} |y_i|$. *For all* $\mu > 0$ *the bound* $\nu(\mu) \geq \frac{2}{\pi\sqrt{\mu}}$ *holds.*

Recently, Helton, Klep, McCullough, and Schweighofer presented a new proof of Ben-Tal-Nemirovski's scaling result based on a new dilation theorem and "very subtle non-trivial properties of Binomial and Beta distributions"; see [HKMS14, Theorems 5.8 and 5.9]. This allows to derive a more explicit description of the constant $\nu(\mu)$ if μ is even, namely

$$\nu(\mu) = \frac{\sqrt{\pi}\,\Gamma\left(1 + \frac{\mu}{4}\right)}{\Gamma\left(\frac{1}{2} + \frac{\mu}{4}\right)},$$

where Γ denotes the Euler gamma function; see [HKMS14, Theorem 13.1]. Moreover, the approach of Helton et al. allows them to improve the statement in the sense that the given bound $\nu(\mu)$ is the best possible for the criterion (5.1.3), and thus for the initial Hol-Scherer relaxation (5.1.6).

Helton et al. also get a bound for all centrally symmetric matricial spectrahedra [HKMS14, Proposition 8.1]. (Note that so far it is unknown whether for a centrally symmetric spectrahedron, the matricial version is again centrally symmetric.) A quantitative result is not known for the general case. However, combining Proposition 5.3.6 with our results from Sections 5.1

and 5.2.1 we get that for spectrahedra with nonempty interior, there is always a scaling factor such that system (5.1.3) and thus also system (5.1.4) hold.

Proposition 5.3.7. *Let $A(x)$ and $B(x)$ be linear pencils with $A_0 \succ 0$ and $B_0 \succ 0$. Assume that S_A is bounded. Then there exists a constant $\nu > 0$ such that for the scaled spectrahedron νS_A the inclusion $\nu S_A \subseteq S_B$ is certified by the initial Hol-Scherer relaxation* (5.1.6) *and by system* (5.1.3).

We provide a proof based on the framework established in the previous sections. Alternatively, for monic linear pencils, the result can be deduced from statements about the matricial relaxation of criterion (5.1.3) given in the work by Helton and McCullough [HM04], see also [HKM13]. Criterion (5.1.3) is satisfied for monic linear pencils $A^\nu(x)$ and $B(x)$ if and only if the matricial version of νS_A is contained in the matricial version of S_B.

Proof. Denote by S_D the cube, defined by the monic linear pencil (5.2.1), with the minimal edge length such that S_A is contained in it. Since $B_0 \succ 0$, there is an open subset around the origin contained in S_B. Thus there is a scaling factor $\nu_1 > 0$ so that $\nu_1 S_A \subseteq \nu_1 S_D \subseteq S_B$.

By Proposition 5.3.6, there exists a constant $\nu_2 > 0$ such that $B(x) \in \mathrm{QM}_0^l(D^\nu)$ with $\nu = \nu_1\nu_2$. By Theorem 5.2.3, $D^\nu(x) \in \mathrm{QM}_0^{2d}(A)$. Finally, Theorem 5.1.15 implies $B(x) \in \mathrm{QM}_0^l(A^\nu)$, certifying the inclusion $\nu S_A \subseteq S_B$. $\square$

In the proof of Proposition 5.3.7, we scale the spectrahedron S_A by a certain factor ν. Since $\nu S_A \subseteq S_B$ is equivalent to $S_A \subseteq \frac{1}{\nu} S_B$, the initial Hol-Scherer relaxation and criterion (5.1.3) remain a positive semidefinite condition even in the presence of the factor ν. Moreover, we can optimize for ν such that the criterion remains satisfied. Proposition 5.3.7 implies that for bounded spectrahedra containing the origin, the maximization problem for ν always has a positive optimal value.

This yields a natural framework for the approximation of smallest enclosing spectrahedra and largest enclosed spectrahedra. In [HKM13, Section 4], an example of computing a bound for the norm of the elements of a spectrahedron S_A (represented by a monic linear pencil) is provided. This can be achieved by choosing S_B to be the ball centered at the origin, see (2.3.14). For the criterion (5.1.3), we obtain a particularly nice representation, it reduces to the semidefinite system

$$
\begin{aligned}
C = (C_{ij})_{i,j=1}^k &\succeq 0, \\
I_{d+1} &= \sum_{i=1}^k C_{ii}, \\
\forall p \in [d]\ \forall\, (s,t) \in [d+1]^2 &: \\
\left(\sum_{i,j=1}^k (A_p)_{ij} C_{ij} \right)_{st} &= \begin{cases} \frac{1}{r} & \text{if } (s,t) \in \{(p,d+1),\, (d+1,p)\}, \\ 0 & \text{else.} \end{cases}
\end{aligned}
\tag{5.3.2}
$$

5.4 Containment of Spectrahedra and Positive Linear Maps

As we will see in this section, checking positivity of a linear map on matrix subspaces is equivalent to checking containment for spectrahedra. We can thus apply our hierarchy for the containment question. On the other hand, the theory of positive linear maps gives some interesting insights into the Spectrahedron Containment problem. Most notably, we prove necessity of the initial Hol-Scherer relaxation (5.1.6) for a family of 2-dimensional spectrahedra where the inner set is a spectratope with nonempty interior given by a linear pencil of size 2×2; see Theorem 5.4.10.

After reviewing the general theory of (completely) positive linear maps (on subspaces) in Section 5.4.1, we study the containment criterion based on complete positivity of operators that was studied in [HKM13].

5.4.1 Positive Linear Maps

The concepts discussed in this subsection can be defined in a much more general setting, using the language of operator theory. For an introduction to positive and completely positive maps on C^*-algebras (and, in particular, Hermitian matrix algebras), we refer to [Pau03]. For our needs, we restrict ourselves to linear maps between real (symmetric) matrix spaces.

Definition 5.4.1. *Let $\mathcal{A} \subseteq \mathbb{R}^{k\times k}$ and $\mathcal{B} \subseteq \mathbb{R}^{l\times l}$ be linear subspaces and consider a linear map $\Phi : \mathcal{A} \to \mathcal{B}$.*

(1) *Φ is called* positive *if the image of any positive semidefinite matrix in $\mathcal{A}$ under Φ is positive semidefinite in $\mathcal{B}$ as well, i.e., $\Phi(\mathcal{A} \cap \mathcal{S}_+^k) \subseteq \mathcal{B} \cap \mathcal{S}_+^l$.*

(2) *Φ is called r*-positive *if the map $\Phi_r : \mathbb{R}^{r\times r} \otimes \mathcal{A} \to \mathbb{R}^{r\times r} \otimes \mathcal{B}$ is positive, i.e., for all $(A_{ij})_{i,j=1}^r \in (\mathbb{R}^{r\times r} \otimes \mathcal{A}) \cap \mathcal{S}_+^{rk}$ we have $(\Phi(A_{ij}))_{i,j=1}^r \in (\mathbb{R}^{r\times r} \otimes \mathcal{B}) \cap \mathcal{S}_+^{rl}$.*

(3) *If Φ is r-positive for all $r \in \mathbb{N}$, then Φ is called* completely positive.

Naturally, every d-positive map is e-positive for all positive integers $e \leq d$. While complete positivity implies positivity, the converse is not always true. We now provide a condition under which those two notions are equivalent.

Proposition 5.4.2. *Let $\Phi : \mathcal{S}^k \to \mathcal{S}^l$ be a linear map. Assume* $\min\{k, l\} \leq 2$. *Then Φ is positive if and only if Φ is completely positive.*

Proof. Let $k = 1$. Then Φ is completely positive if and only if $\Phi_k = \Phi_1 : \mathbb{R}^{1\times 1} \otimes \mathbb{R} \to \mathbb{R}^{1\times 1} \otimes \mathcal{S}^l$ is positive. Clearly, $\Phi_1 = \Phi$.

Let $l = 1$. Then Φ is completely positive if and only if $y^T(\Phi(A_{ij}))_{i,j} y \geq 0$ for all $y \in \mathbb{R}^k$ and $(A_{ij})_{i,j} \in \mathcal{S}_+^k$. Define $\mathbb{I} = (I_k, \ldots, I_k)^T$. Then

$$y^T \left(\Phi(A_{ij})\right)_{i,j} y = \sum_{i,j=1}^{k} y_i y_j \Phi\left(A_{ij}\right) = \Phi\left(\sum_{i,j=1}^{k} y_i y_j A_{ij}\right) = \Phi\left(\mathbb{I}^T (y_i y_j A_{ij})_{i,j} \mathbb{I}\right)$$

is nonnegative since $(y_i y_j A_{ij})_{i,j}$ is a principal submatrix of the positive semidefinite matrix $yy^T \otimes (A_{ij})_{i,j}$ and Φ is a positive map.

The case $\min\{k, l\} = 2$ is more involved. We refer to [Cho75a, Theorem 7]. □

Choi showed the connection between positivity of linear maps and nonnegativity of biquadratic forms as well as complete positivity and a biquadratic form being a sum of squares of bilinear forms.

Proposition 5.4.3 ([Cho75b, Section 3])**.** *A linear map $\Phi : \mathcal{S}^k \to \mathcal{S}^l$ is a positive map if and only if the biquadratic form $F(x, y) = y^T \Phi(xx^T) y : \mathbb{R}^k \otimes \mathbb{R}^l \to \mathbb{R}$ is nonnegative.*

A positive linear map $\Phi : \mathcal{S}^k \to \mathcal{S}^l$ is completely positive, i.e., $\Phi(A) = \sum_s V_s^T A V_s$ for some matrices $V_s \in \mathbb{R}^{k\times l}$, if and only if the corresponding biquadratic form $F(x, y)$ is a sum of squares of bilinear forms, $F(x, y) = \sum_s (x^T V_s y)^2$.

As a corollary, we get an (incomplete) complexity classification of deciding (complete) positivity of linear maps.

Corollary 5.4.4. *Let $\Phi: \mathcal{S}^k \to \mathcal{S}^l$ be a linear map. Deciding positivity of Φ is an NP-hard problem. Deciding complete positivity of Φ can be formulated as an SDFP* (2.3.13) *whose size is polynomial in the description size of the input data.*

Proof. The first follows from the fact that deciding nonnegativity of a polynomial is NP-hard. Since deciding whether a polynomial is a sum of squares is an SDFP (2.3.13), the second statement follows. □

For the cases $\min\{k, l\} > 2$, there exist examples of positive maps which are not completely positive; see [Ter39, Satz 9] for Terpstra's example. However, since deciding positivity of linear maps is NP-hard, only few examples are known that are not completely positive. Stinespring and Arveson stated examples of positive but not completely positive maps; but all these examples fail to be 2-positive. Choi [Cho75a] stated the first example of a 2-positive map that is not completely positive (for $k = l = 3$); see also Example 5.4.11 below. See [Cho72] for a generalization to $(k-1)$-positive maps which are not k-positive; and [CKL92] for a generalization of "Choi's map" (for $k = l = 3$) to a parametric family of (2-)positive maps which are not completely positive. The structure of positive maps on higher dimensional spaces is not completely understood; see, e.g., [SSŻ09, Stø63].

Provided that $\mathcal{A}$ contains a positive definite matrix, complete positivity of Φ is equivalent to k-positivity. Interestingly, in this situation every completely positive map does have a completely positive extension to the full matrix space and can therefore be represented by a positive semidefinite matrix. This is well-known in the general setting of C^*-algebras, and persists in our real setting.

Proposition 5.4.5 ([Pau03, Theorems 6.1 and 6.2.])**.** *Let $\mathcal{A} \subseteq \mathbb{R}^{k\times k}$ be a linear subspace containing a positive definite matrix.*

(1) *A linear map $\Phi: \mathcal{A} \to \mathcal{S}^l$ is completely positive if and only if it is k-positive.*

(2) *Each completely positive map $\Phi: \mathcal{A} \to \mathcal{S}^l$ has an extension to a completely positive map $\tilde{\Phi}: \mathbb{R}^{k\times k} \to \mathbb{R}^{l\times l}$. Moreover, complete positivity of the map $\tilde{\Phi}$ is equivalent to positive semidefiniteness of the matrix $C = (C_{ij})_{i,j=1}^k = (\tilde{\Phi}(E_{ij}))_{i,j=1}^k \in \mathcal{S}^{kl}$.*

A significant implication of Proposition 5.4.5 is the following. Given a linear subspace $\mathcal{A}$ containing a positive definite matrix, a linear map $\Phi: \mathcal{A} \to \mathbb{R}^{l\times l}$ is completely positive if and only if at least one of all possible extensions of Φ to the whole matrix space is completely positive. The set of extensions is determined by linear equations, fixing some (but not all) of the entries in the matrix C. Testing the partially indeterminate matrix C for a positive semidefinite extension is a semidefinite feasibility problem (SDFP). Recall from the preliminary Section 2.3 that while the computational complexity of solving SDFPs is open, in practice it can be done efficiently by semidefinite programming. Surprisingly, the same is not true for positive maps. Positive maps on subspaces do not always have a positive extension to the full space, even if the subspace contains a positive definite matrix; see, e.g., [Stø86, Example 3.16].

5.4.2 Containment of 2-dimensional Spectrahedra

Our main goal is to prove necessity of the initial Hol-Scherer relaxation (5.1.6) for a special family of 2-dimensional spectrahedra. For the convenience of the reader, we first collect the relevant connections between the containment problem and (complete) positivity of maps on matrix subspaces.

By showing the equivalence of containment of the so-called *matricial relaxations* (also called *free spectrahedra*) of two spectrahedra S_A, S_B given by monic linear pencils and the existence of a completely positive unital linear map

$$\Phi: \ \mathrm{span}\{I_k, A_1, \ldots, A_d\} \to \mathrm{span}\{I_l, B_1, \ldots, B_d\}, \ A_p \mapsto B_p,$$

the authors of [HKM12, HKM13] proved that the system

$$C = (C_{ij})_{i,j=1}^k \in \mathcal{S}_+^{kl}, \quad I_l = \sum_{i=1}^k C_{ii}, \quad \forall p \in [d]: \ B_p = \sum_{i,j=1}^k (A_p)_{ij} C_{ij}$$

has a solution if and only if the matricial relaxation of S_A is contained in the one of S_B. In this case $S_A \subseteq S_B$. Clearly, this is the sufficient semidefinite containment criterion (5.1.3) stated in Section 5.1.1 when restricted to monic linear pencils. Thus Theorem 5.1.3 serves as an extension and a more streamlined proof of the Helton-Klep-McCullough criterion. Moreover, in our approach it becomes apparent that we can relax the criterion given by Helton, Klep and McCullough by replacing the linear constraints on the constant matrices in (5.1.3) with semidefinite constraints.

Given the linear pencils $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$, define the corresponding linear subspaces

$$\begin{aligned}
\mathcal{A} &= \mathrm{span}(A_0, A_1, \ldots, A_d) \subseteq \mathcal{S}^k, \\
\widehat{\mathcal{A}} &= \mathrm{span}(1 \oplus A_0, 0 \oplus A_1, \ldots, 0 \oplus A_d) \subseteq \mathcal{S}^{k+1}, \text{ and} \\
\mathcal{B} &= \mathrm{span}(B_0, B_1, \ldots, B_d) \subseteq \mathcal{S}^l.
\end{aligned}$$

Recall from Section 5.1.1 the definition of the *extended linear pencil* $\widehat{A}(x) = 1 \oplus A(x)$.

For linearly independent $A_1, \ldots, A_d$, let $\widehat{\Phi}_{AB}: \ \widehat{\mathcal{A}} \to \mathcal{B}$ be the linear map defined by

$$\widehat{\Phi}_{AB}(1 \oplus A_0) = B_0, \ \forall p \in [d]: \ \widehat{\Phi}_{AB}(0 \oplus A_p) = B_p.$$

Note that since every linear combination $0 = \lambda_0(1 \oplus A_0) + \sum_{p=1}^d \lambda_p(0 \oplus A_p)$ for real scalars $\lambda_0, \ldots, \lambda_d$ yields $\lambda_0 = 0$, it suffices to assume the linear independence of the coefficient matrices $A_1, \ldots, A_d$ to ensure that $\widehat{\Phi}_{AB}$ is well-defined. To obtain linear independence, the lineality space can be treated separately, as described in Proposition 2.3.3. Note that the lineality space for the extended pencil is the same as for the actual pencil.

If additionally, $A_0, A_1, \ldots, A_d$ are linearly independent, we can retreat to the simpler map $\Phi_{AB}: \ \mathcal{A} \to \mathcal{B}$ defined by

$$\Phi_{AB}: \ A_p \mapsto B_p \quad \forall p \in [d].$$

In [HKM13, Theorem 3.5] the authors state the relationship between r-positive maps and the question of containment of (bounded) matricial positivity domains. For $r = 1$, this contains the case of spectrahedra. Their proof is based on operator algebra. We give a more streamlined proof concerning only positive maps and spectrahedra.

Proposition 5.4.6. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils.*

(1) *If Φ_{AB} or $\widehat{\Phi}_{AB}$ is positive, then $S_A \subseteq S_B$.*

(2) *If $S_A \neq \emptyset$, then $S_A \subseteq S_B$ implies positivity of $\widehat{\Phi}$.*

(3) *If $S_A \neq \emptyset$ is bounded, then $S_A \subseteq S_B$ implies positivity of Φ.*

Proof. Set $\widehat{A}(x_0, x) := x_0(1 \oplus A_0) + \sum_{p=1}^{d} x_p(0 \oplus A_p) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$.

To (1): Let Φ_{AB} be positive. Fix $x \in S_A$. Then $B(x) = \Phi(A(x)) \in \mathcal{S}_+^l \cap \mathcal{B}$ and hence $x \in S_B$. If $\widehat{\Phi}_{AB}$ is positive, the proof is verbatim the same.

To (2): Since the spectrahedra defined by $A(x)$ and $\widehat{A}(x)$ coincide, we have $S_{\widehat{A}} \subseteq S_B$. Let $\widehat{A}(x_0, x) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$. Then $x_0 \geq 0$.

Case $x_0 > 0$. Scaling of the linear pencil with $1/x_0$ preserves positive semidefiniteness. Thus, $\widehat{A}(1, x/x_0) = 1/x_0 \cdot \widehat{A}(x_0, x) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$ and $x/x_0 \in S_A \subseteq S_B$. Scaling $B(x/x_0)$ by x_0 yields $\widehat{\Phi}(\widehat{A}(x_0, x)) = x_0 B_0 + \sum_{p=1}^{d} x_p B_p = x_0 B(x/x_0) \in \mathcal{S}_+^l \cap \mathcal{B}$.

Case $x_0 = 0$. If $(x, x_0) = (0, 0)$, the statement is obvious. Let $x \neq 0$. By assumption, S_A is nonempty. Fix a point $\bar{x} \in S_A$. Then $\widehat{A}(1, \bar{x} + tx) = \widehat{A}(1, \bar{x}) + \widehat{A}(0, tx) \succeq 0$ for all $t > 0$, implying $\bar{x} + tx \in S_A \subseteq S_B$ for all $t > 0$. Thus x is a point of the recession cone of S_A which clearly is contained in the recession cone of S_B. Indeed, $\frac{1}{t} B(1, \bar{x}) + B(0, x) = \frac{1}{t} B(\bar{x} + tx) \succeq 0$ for all $t > 0$. By closedness of the cone of positive semidefinite matrices, we get $B(0, x) \succeq 0$. Hence, $\widehat{\Phi}(\widehat{A}(x_0, x)) = \widehat{\Phi}(\widehat{A}(0, x)) = B(0, x) \succeq 0$.

To (3): Let $A(x_0, x) = x_0 A_0 + \sum_{p=1}^{d} x_p A_p$ be in $\mathcal{S}_+^k \cap \mathcal{A}$. We distinguish between the following two cases.

Case $x_0 > 0$. This case follows by a similar scaling argument as in part (2). (Note that $A_0 \neq 0$ by linear independence.)

Case $x_0 \leq 0$. Since S_A is nonempty, there exists $\bar{x} \in S_A$ and hence

$$\begin{aligned} A(0, x + |x_0|\bar{x}) &= A(0, x) + A(0, |x_0|\bar{x}) \\ &\succeq |x_0| A_0 + A(0, |x_0|\bar{x}) = |x_0| \cdot A(1, \bar{x}) \succeq 0. \end{aligned}$$

For $A(0, x + |x_0|\bar{x}) \neq 0$, one has an improving ray of the spectrahedron S_A, in contradiction to the boundedness of S_A. For $A(0, x + |x_0|\bar{x}) = 0$, linear independence of $A_0, \ldots, A_d$ implies $x + |x_0|\bar{x} = 0$. But then $x_0 A(1, \bar{x}) = A(x_0, x) \succeq 0$ together with $x_0 \leq 0$ and $A(1, \bar{x}) \succeq 0$ imply either $A(1, \bar{x}) = 0$, in contradiction to linear independence, or $(x_0, x) = 0$. Clearly, in this case, $\Phi_{AB}(0) = 0$. ☐

The assumptions in parts (2) and (3) of Proposition 5.4.6 may not be omitted in general, as the next examples show.

Example 5.4.7.

(1) Consider the two linear pencils

$$A(x) = \begin{bmatrix} -3 + x_1 + x_2 & 0 & 0 \\ 0 & -1 + x_1 & 0 \\ 0 & 0 & -1 + x_2 \end{bmatrix} \text{ and } B(x) = \begin{bmatrix} -1 + x_1 + x_2 & 0 & 0 \\ 0 & x_1 & 0 \\ 0 & 0 & x_2 \end{bmatrix},$$

defining unbounded, nonempty polyhedra in $\mathbb{R}^2$. It is easy to see that the coefficient matrices are linearly independent and that S_A does not contain the origin.

While S_A is contained in S_B, the linear map Φ_{AB} is not positive. Indeed, the homogeneous pencil $A(x_0, x)$ evaluated at the point $(x_0, x_1, x_2) = (-1, -1/2, -1/2)$ is positive definite while $B(x_0, x)$ is indefinite.

Therefore, the boundedness assumption in part (3) of Proposition 5.4.6 cannot be omitted in general. Using the extended linear pencil $\widehat{A}(x) = 1 \oplus A(x)$ *instead of* $A(x)$*, the resulting constraint* $x_0 \geq 0$ *yields the positivity of* $\widehat{\Phi}_{AB}$*. In fact,* $\widehat{\Phi}_{AB}$ *is completely positive, which can be checked by the SDFP* (5.1.3)*.*

(2) Consider the two linear pencils

$$A(x) = \begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} \text{ and } B(x) = \begin{bmatrix} 1 & -x \\ -x & 1 \end{bmatrix}.$$

with linearly independent coefficient matrices. The corresponding spectrahedra are the empty set $S_A = \emptyset$ *and the interval* $S_B = [-1, 1]$*. Thus* $S_A \subseteq S_B$*. However, the linear map* Φ_{AB} *is not positive since the homogeneous pencil* $A(x_0, x_1)$ *is positive semidefinite at* $(x_0, x_1) = (0, 1)$ *but* $B(x_0, x_1)$ *is not. Note that this holds for the extended pencil as well. Thus nonemptyness of the inner spectrahedron cannot be dropped.* □

If our setting is changed from the case of linear subspaces to the case of affine subspaces, with a natural adaption of the notion of positivity to affine maps, Proposition 5.4.6 has a slightly easier formulation and proof.

Lemma 5.4.8. *Let* $A(x) \in \mathcal{S}^k[x]$ *and* $B(x) \in \mathcal{S}^l[x]$ *be linear pencils. Define the affine subspaces* $\bar{\mathcal{A}} = \frac{1}{d}A_0 + \operatorname{span}(A_1, \ldots, A_d)$ *and* $\bar{\mathcal{B}} = \frac{1}{d}B_0 + \operatorname{span}(B_1, \ldots, B_d)$ *for linearly independent* $A_1, \ldots, A_d$*. Then* $S_A \subseteq S_B$ *if and only if the affine function* $\bar{\Phi}_{AB} : \bar{\mathcal{A}} \to \bar{\mathcal{B}}$ *defined by* $\frac{1}{d}A_0 + A_p \mapsto \frac{1}{d}B_0 + B_p$ *for* $i \in [d]$ *is positive.*

Proof. First, let $\bar{\Phi}_{AB}$ be positive and let $x \in S_A$. Since $\bar{\Phi}_{AB}$ is positive, we have $B(x) = \bar{\Phi}_{AB}(A(x)) \succeq 0$, thus $x \in S_B$. Conversely, let $\frac{1}{d}A_0 + \sum_{p=1}^d x_p A_p \in \bar{\mathcal{A}} \cap \mathcal{S}_+^k$. Then $dx \in S_A \subseteq S_B$ and hence $\bar{\Phi}_{AB}\left(\frac{1}{d}A_0 + \sum_{p=1}^d x_p A_p\right) = \frac{1}{d}B_0 + \sum_{p=1}^d x_p B_p \succeq 0$. □

Example 5.4.9. *Consider part (2) of Example 5.4.7. In the affine case, we get*

$$\bar{\mathcal{A}} = \left\{ A \mid A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + x \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\} \text{ and } \bar{\mathcal{B}} = \left\{ B \mid B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \right\}$$

with $\bar{\mathcal{A}} \cap \mathcal{S}_+^2 = \emptyset$ *and* $\bar{\Phi}_{AB}(\emptyset) = \emptyset \subseteq \mathcal{B} \cap \mathcal{S}_+^2$*.* □

Unfortunately, checking if a given linear map between operator systems is positive is not an easy task in general. Showing that it has the stronger property of being completely positive can be done using the semidefinite feasibility problem (5.1.3) (cf. Proposition 5.4.5 and Theorem 5.1.3). As a consequence of Proposition 5.4.2 and Proposition 5.4.6, we get the following statement on the Spectrahedron Containment problem concerning 2-dimensional spectrahedra.

Theorem 5.4.10. *Let* $A(x) = A_0 + x_1 A_1 + x_2 A_2 \in \mathcal{S}^2[x]$ *be strictly feasible and let* $B(x) \in \mathcal{S}^l[x]$ *be a linear pencil. Assume* $\mathcal{A} = \mathcal{S}^2$ *and that* S_A *is bounded. Then* $S_A \subseteq S_B$ *if and only if* $B(x) \in \mathrm{QM}_0^l[x]$*.*

Proof. By the assumptions, S_A is contained in S_B if and only if the map $\Phi_{AB} : \mathcal{S}^2 \to \mathcal{S}^l$ is positive. In this case, positivity of Φ_{AB} is equivalent to complete positivity; see Proposition 5.4.2. As complete positivity is equivalent to the feasibility of the SDFP (5.1.3) and thus to the initial Hol-Scherer relaxation (5.1.6), the claim follows by the equivalence of the Spectrahedron Containment problem and positivity of linear maps, Proposition 5.4.6. □

Note that the first part of Example 5.1.7 belongs to the situation of Theorem 5.4.10.

To close the section, we apply our hierarchy to a well-known example of a 2-positive but not completely positive linear map.

Example 5.4.11. *Consider the linear map*

$$\Phi:\ \mathcal{S}^3 \to \mathcal{S}^3,\ A \mapsto 2\begin{bmatrix} A_{11}+A_{22} & & \\ & A_{22}+A_{33} & \\ & & A_{33}+A_{11}\end{bmatrix} - A.$$

Due to Choi [Cho75b], the map Φ is (1- and 2-)positive but not completely positive. Indeed, neither the SDFP (5.1.3) *nor the initial Hol-Scherer relaxation* (5.1.6) *are feasible but for $t=1$ the relaxation yields a small positive value implying positivity of Φ.* □

5.5 A Bilinear Programming Approach to the Spectrahedron Containment Problem

Besides the relaxation approach based on Hol-Scherer's Positivstellensatz, one could wonder whether it is possible to apply a bilinear programming approach using the polar of the outer spectrahedron, as used in Chapter 4. The problem to overcome here is that the best known representation of the polar spectrahedron is as (the closure of) a projected spectrahedron.

Recall from Section 2.3 that given a linear pencil $B(x) \in \mathcal{S}^l[x]$ with $B_0 \succeq 0$, the polar of the spectrahedron S_B is

$$S_B^\circ = \operatorname{cl}\left\{z \in \mathbb{R}^d \mid \exists Y \in \mathcal{S}_+^l,\ y_0 \in \mathbb{R}_+:\ z_p = -\langle B_p, Y\rangle,\ \langle B_0, Y\rangle + y_0 = 1\right\}.$$

Our starting point is the following reformulation of the Spectrahedron Containment problem as a bilinear feasibility problem.

Proposition 5.5.1. *Let $A(x) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils such that S_A is nonempty and that S_B contains the origin, i.e., $B_0 \succeq 0$. S_A is contained in S_B if and only if*

$$x^T z \le 1 \quad \text{for all } (x,z) \in S_A \times S_B^\circ.$$

Proof. If $S_A \subseteq S_B$, then for any $x \in S_A$ we have $x^T z \le 1$ for all $z \in S_B^\circ$. Conversely, if $x^T z \le 1$ holds for all $z \in S_B^\circ$, then for any $x \in S_A$ we have $S_A \subseteq S_B^{\circ\circ} = S_B$. □

The problem occurring with this formulation is, for one thing, the fact that the polar S_B° is not closed in general, for another thing, that the number of variables increases significantly as one has to consider the optimization variables $z = (z_1, \ldots, z_d)$ as well as the projection variables $Y = (Y_{ij})_{i,j=1}^l$ and y_0. However, both problems can easily be tackled. By assuming $0 \in \operatorname{int}(S_B)$, the first problem can be avoided. (Note that the computational complexity of testing whether the origin (or any point) is contained in the interior of a spectrahedron has not been classified so far.) A way around the second problem, yielding a substantial reduction on the number of additional variables, is explained in the following.

Denote by $\tilde{B}(x) = \sum_{i=1}^d x_i B_i$ the pure-linear part of $B(x)$. By Proposition 5.5.1, $S_A \subseteq S_B$ if and only if

$$0 \le \inf\left\{1 - x^T z \mid A(x) \succeq 0,\ z_p = -\langle B_p, Y\rangle,\ \langle B_0, Y\rangle + y_0 = 1, Y \in \mathcal{S}_+^l,\ y_0 \in \mathbb{R}_+\right\}.$$

Substituting 1 and z in the objective, and using the fact that $\left\langle x, L^T(y)\right\rangle = \langle L(x), y\rangle$ for a linear operator L, yields

$$S_A \subseteq S_B \;\Leftrightarrow\; 0 \geq \sup\left\{\left\langle -\tilde{B}(x), Y\right\rangle - \langle B_0, Y\rangle - y_0 \mid A(x) \succeq 0,\ Y \in \mathcal{S}_+^l,\ y_0 \in \mathbb{R}_+\right\}.$$

Clearly, y_0 is superfluous and thus $S_A \subseteq S_B$ if and only if

$$0 \leq \inf\left\{\langle B(x), Y\rangle \mid A(x) \succeq 0,\ Y \in \mathcal{S}_+^l\right\}.$$

Since the set of rank-1 matrices generates the cone of positive semidefinite matrices, this can equivalently stated as

$$0 \leq \inf\left\{z^T B(x) z \mid A(x) \succeq 0,\ z \in \mathbb{R}^l\right\}. \tag{5.5.1}$$

In fact, this last expression can easily be seen from the definition of containment and positive semidefiniteness. Using Putinar's Positivstellensatz, one gets another hierarchy of semidefinite programs to decide containment of spectrahedra. This is the basis for the optimization approach in Trabandt's PhD thesis [Tra14]. See [KTT15] for the dual relaxation based on Lasserre's moment approach for matrix polynomials [HL06].

The bilinear approach seems to be the scalarized version of the Hol-Scherer based approach in Section 5.1.2 as the following theorem indicates.

Theorem 5.5.2 ([KTT15, Theorem 4.9])**.** *Let $A(x) \in \mathcal{S}^k[x]$ be a linear pencil. If $B(x) \in \mathrm{QM}_0^l(A)$, then the infimum of the initial relaxation step of the Putinar based relaxation is nonnegative (and thus certifies containment).*

The converse is not true in general. Let $A(x) \in \mathcal{S}^{16}[x]$ be the normal form of the ball in $\mathbb{R}^{15}$ with radius $r = 1/2$ centered at the origin and let $B(x) \in \mathcal{S}^6[x]$ be the normal form of the elliptope. Then the initial Putinar relaxation of problem (5.5.1) *certifies containment while the initial Hol-Scherer relaxation* (5.1.6) *does not.*

Due to Theorem 5.5.2, all convergence results stated in this chapter (in particular, Theorem 5.3.1) can be brought forward to the scalarized approach.

A crucial point in the polynomial optimization approach based on Putinar's Positivstellensatz is the introduction of additional variables $z = (z_1, \ldots, z_l)$ already in the original, unrelaxed polynomial formulation. While in the quantified semidefinite program (5.1.7) no additional variables $z = (z_1, \ldots, z_l)$ are needed, the number of unknowns of the relaxation grows not only in the number of variables d and the relaxation order t, i.e., half the degree of the entries in $S(x)$, but also in the size of both the outer pencil l and the inner pencil k. To be more precise, using the approach of Hol and Scherer, the number of unknowns in the SDP coming from the sos-relaxation is generically

$$1 + \frac{1}{2}\binom{d+t}{t} \cdot \left[k^2 l^2 \binom{d+t}{t} + l^2 \binom{d+t}{t} + kl + l\right] - ml(l+1),$$

where m denotes the number of affine equation constraints arising in the Hol-Scherer formulation; see [HS06, Section 5]. In Putinar's approach there are

$$\frac{1}{2}\binom{d+l+t}{t}\left[\binom{d+l+t}{t} - 1\right]$$

variables. On the other hand, in certain situations with small t (i.e., $t \in \{0,1\}$), the Hol-Scherer based approach may lead to SDPs with a simpler structure than the Putinar based relaxation. But if $t \in \{0,1\}$ does not give an answer to the containment question, in our computational examples the Hol-Scherer method with $t \in \{2,3\}$ is significantly slower than the Putinar based approach. Nevertheless, in almost all examples we were able to compute the sum of squares decomposition in moderate time, we got the same optimal value.

For more details on the relaxation based on Putinar's Positivstellensatz see [KTT15, Tra14].

6 The Projection Containment Problems

For containment problems concerning $\mathcal{H}$-polyhedra and spectrahedra, including the Spectrahedron Containment problem, clean formulations as polynomial feasibility problems are available. In this chapter, we show that this persists if the inner set is the projection of an $\mathcal{H}$-polyhedron or a spectrahedron. If the outer set is a projection, then the containment problem is more involved (already in the case $\mathcal{H}$-in-$\pi\mathcal{H}$). An aim of this chapter is to give a starting and motivation point for the study of Positivstellensätze on projected convex sets, such as $\pi\mathcal{H}$-polyhedra and $\pi\mathcal{S}$-spectrahedra.

The $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment problem can be formulated as a bilinear feasibility problem (Theorem 6.1.1). Continuing with the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem, a polynomial formulation lacks in the fact that projected spectrahedra are not closed in general. However, under an additional assumption, which is common in semidefinite programming, the statement holds (Theorem 6.2.1).

While for the general case, the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem, we hit on problems, like the lack of a clean Farkas type Lemma for cones as well as the absence of a sophisticated Positivstellensatz, retreating to the $\pi\mathcal{H}$-in-$\mathcal{H}$ and $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problems allows to bring forward results from the non-projected case. More precisely, for the $\pi\mathcal{H}$-in-$\mathcal{H}$ containment problem, we establish an analog statement as for the $\mathcal{H}$-in-$\mathcal{H}$ containment problem (Theorem 6.3.1). Among others, this serves as an algorithmic proof of Theorem 3.2.3. From that we deduce a sufficient semidefinite criterion for the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem (Theorem 6.3.3).

In principle, the application of Hol-Scherer's Positivstellensatz to the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem is possible. The drawback of this approach is that it relies on the geometry of the inner spectrahedron rather than its projection, namely boundedness of the spectrahedron as well as the appearance of the projection variables in the quadratic module. To address this, we establish a refinement of Hol-Scherer's Positivstellensatz that allows to reduce the complexity of the problem significantly. Particularly, the projection variables do not appear in the quadratic module or the relaxation anymore (Theorem 6.3.5).

This chapter is structured as follows. First, we have a look at the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment in Section 6.1. We then extend our results to the case of projected spectrahedra; see Section 6.2. Section 6.3 deals with the $\pi\mathcal{H}$-in-$\mathcal{H}$ and $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problems. We close the chapter with discussing a possible extension of the connection between positive linear maps and containment of spectrahedra (as shown in Section 5.4) to the projected case.

6.1 A Bilinear Formulation of the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ Containment Problem

For $a \in \mathbb{R}^k,\ A \in \mathbb{R}^{k\times d},\ A' \in \mathbb{R}^{k\times m}$ and $b \in \mathbb{R}^l,\ B \in \mathbb{R}^{l\times d},\ B' \in \mathbb{R}^{l\times n}$, let

$$\begin{aligned} \pi(P_A) &= \left\{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m :\ a + Ax + A'y \geq 0\right\} \\ \text{and}\quad \pi(P_B) &= \left\{x \in \mathbb{R}^d \mid \exists y' \in \mathbb{R}^n :\ b + Bx + B'y' \geq 0\right\} \end{aligned} \tag{6.1.1}$$

be projections of the $\mathcal{H}$-polyhedra P_A and P_B, respectively. Note that both $\pi(P_A)$ and $\pi(P_B)$ are $\mathcal{H}$-polyhedra themselves (and thus closed sets). A quantifier-free $\mathcal{H}$-description however can be exponential in the input size (d, m, k) respectively (d, n, l); cf. Section 2.2.

Our starting point is the formulation of the containment problem as a bilinear feasibility problem. Interestingly, the projection variables y' of the outer polyhedron do not appear in the feasibility system (or the optimization version below) only the corresponding coefficients B'.

Theorem 6.1.1. *Let $\pi(P_A)$ and $\pi(P_B)$ be as defined in (6.1.1) and $\pi(P_A)$ be nonempty.*

(1) *$\pi(P_A)$ is contained in $\pi(P_B)$ if and only if*

$$z^T(b + Bx) \geq 0 \quad \text{on} \quad \pi(P_A) \times \left(\ker(B'^T) \cap \mathbb{R}^l_+\right).$$

(2) *Assume $\ker(B'^T) \cap \mathbb{R}^l_+ = \operatorname{span}(B')^\perp \cap \mathbb{R}^l_+ \neq \{0\}$. Then $\pi(P_A) \subseteq \pi(P_B)$ if and only if*

$$z^T(b + Bx) \geq 0 \quad \text{on} \quad \pi(P_A) \times \left(\ker(B'^T) \cap \Delta^l\right),$$

where $\Delta^l = \{z \in \mathbb{R}^l \mid \mathbb{1}_l^T z = 1,\ z \geq 0\}$ is the l-simplex (2.2.5).

The additional assumption on the kernel of B'^T seems to be somewhat artificial, however, if the projection of P_B to the x-coordinates is bounded, then the condition holds. The two main advantages of part (2) in Theorem 6.1.1 are the boundedness of the z variables and that the condition $z^T(b + Bx) \geq 0$ is indeed an inequality. (Note that in part (1), containment is equivalent to $z^T(b + Bx) \equiv 0$ on $\pi(P_A) \times (\operatorname{span}(B')^\perp \cap \mathbb{R}^l_+)$, as $(x, z) = (x, 0)$ is a feasible solution for all $x \in \pi(P_A)$.) The next lemma serves as a first step in a geometric interpretation of this precondition.

Lemma 6.1.2. *Let $\pi(P_B)$ be as in (6.1.1). Then $\ker(B'^T) \cap \mathbb{R}^l_+ = \operatorname{span}(B')^\perp \cap \mathbb{R}^l_+ = \{0\}$ if and only if $\operatorname{span}(B') \cap \mathbb{R}^l_{++} \neq \emptyset$. In this case, $\pi(P_B) = \mathbb{R}^d$.*

In particular, if $\pi(P_B)$ is bounded, then $\ker(B'^T) \cap \mathbb{R}^l_+ = \operatorname{span}(B')^\perp \cap \mathbb{R}^l_+ \neq \{0\}$.

Proof. The equivalence $\ker(B'^T) \cap \mathbb{R}^l_+ = \operatorname{span}(B')^\perp \cap \mathbb{R}^l_+ = \{0\} \iff \operatorname{span}(B') \cap \mathbb{R}^l_{++} \neq \emptyset$ is easy to see. If so, then there exists $y' \in \mathbb{R}^l$ such that $B'y' > 0$. Thus, for every $x \in \mathbb{R}^d$, there exists $t > 0$ sufficiently large such that $b + Bx + B'(ty') \geq 0$. This implies $\pi(P_B) = \mathbb{R}^d$. Thus, for bounded $\pi(P_B)$, $\ker(B'^T) \cap \mathbb{R}^l_+ \neq \{0\}$. $\square$

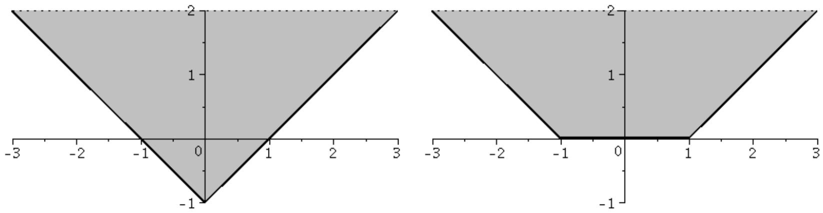
Before proving Theorem 6.1.1, we observe that neither the implication "$\operatorname{span}(B') \cap \mathbb{R}^l_{++} \neq \emptyset \implies \pi(P_B) = \mathbb{R}^d$" nor the implication "$\pi(P_B)$ is bounded $\implies \ker(B'^T) \cap \mathbb{R}^l_+ \neq \{0\}$" in Lemma 6.1.2 is an equivalence. Example 6.1.3 also shows that the precondition in part (2) of Theorem 6.1.1 cannot be dropped.

Example 6.1.3.

(1) Consider the polyhedron

$$P_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \end{pmatrix} y \geq 0 \right\}.$$

P_1 is a pointed polyhedral cone; see Figure 6.1.1 (a). We have $\operatorname{span}(B') \cap \mathbb{R}^2_{++} \neq \emptyset$ and thus the intersection of $\ker(B'^T) = \ker(1, 1)$ and the nonnegative real numbers is zero-dimensional,



(a) P_1 as defined in Example 6.1.3. (b) P_2 as defined in Example 6.1.3.

Figure 6.1.1

i.e., $\ker(B'^T) \cap \mathbb{R}^2_+ = \{0\}$. *Moreover, in this case, the restriction to the* 1*-simplex as in part (2) of Theorem 6.1.1 is not possible.*

(2) Consider the polyhedron

$$P_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} y \geq 0 \right\},$$

which is unbounded but not a cone; see Figure 6.1.1 (b). We have $\operatorname{span}(B') \cap \mathbb{R}^3_{++} = \emptyset$ *and thus, by Lemma 6.1.2,* $\ker(B'^T) \cap \mathbb{R}^3_+ \neq \{0\}$. *Indeed, for every* $t \geq 0$, *we have* $(0,0,t) \in \ker(B'^T) \cap \mathbb{R}^3_+ \neq \{0\}$. *On the other hand,* $\pi(P_B) = \mathbb{R}$ *shows that the reverse of the other (and above mentioned) implications in Lemma 6.1.2 are not equivalences.* □

Proof of Theorem 6.1.1.

To (1): $\pi(P_A) \not\subseteq \pi(P_B)$ if and only if there exists a point $x \in \pi(P_A) \backslash \pi(P_B)$, i.e., for $x \in \pi(P_A)$ there exists no $y' \in \mathbb{R}^n$ with $b + Bx + B'y' \geq 0$. By Farkas' Lemma 2.2.1 this is equivalent to the existence of a point $z \in \mathbb{R}^l_+$ with $z^T B' = 0$ such that $z^T(b + Bx) < 0$ holds. Equivalently, there exists $(x, z) \in \pi(P_A) \times (\ker(B'^T) \cap \mathbb{R}^l_+)$ such that $z^T(b + Bx) < 0$.

To (2): If there exists $(x, z) \in \pi(P_A) \times \left(\ker(B'^T) \cap \mathbb{R}^l_+\right)$ such that $z^T(b+Bx) < 0$, then $z \neq 0$ and thus $z'^T(b + Bx) < 0$ for $z' = \frac{z}{|z|} \geq 0$ with $|z'| = \sum_{i=1}^l z'_i = \frac{1}{|z|} \sum_{i=1}^l z'_i = 1$.

Assume $z^T(b + Bx) \geq 0$ holds for all $(x, z) \in \pi(P_A) \times \left(\ker(B'^T) \cap \mathbb{R}^l_+\right)$. By assumption, there exists $0 \neq z \in \ker(B'^T) \cap \mathbb{R}^l_+$. Applying the same scaling as above yields $z^T(b + Bx) \geq 0$ for every $z' \in \Delta^l$, implying the claim. □

Note that the set $\ker(B'^T) \cap \Delta^l_+$ is intrinsically linked to the polar of $\pi(P_B)$. Namely, it is the set of convex combinations of the columns in B'^T that are equal to the origin, i.e., $0 = B'^T z$ with $1 = \mathbb{1}_l^T z$ and $z \geq 0$. Thus $\ker(B'^T) \cap \Delta^l_+$ is a polytope.

Recall Konno's result on disjointly constrained bilinear programs (cf. Proposition 4.1.3). As it only depends on the structure of the problem rather than on the explicit representation of the polytopes involved, it is possible to state a similar result for $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ as for the Polytope Containment problem in Chapter 4. To do so, consider the optimization version of Theorem 6.1.1

$$\begin{aligned}\inf\ & z^T(b+Bx) \\ \text{s.t.}\ & (x,y,z) \in P_A \times Q_B,\end{aligned} \tag{6.1.2}$$

where we define $Q_B := \ker(B'^T) \cap \Delta_+^l = \{z \in \mathbb{R}^l \mid B'^T z = 0,\ \mathbb{1}_l^T z = 1,\ z \geq 0\}$. Assuming nonemptyness of Q_B, Theorem 6.1.1 implies that $\pi(P_A) \subseteq \pi(P_B)$ if and only if the infimum is nonnegative.

Proposition 6.1.4. *Let P_A be a nonempty polytope and assume $\ker(B'^T) \cap \Delta^l \neq \emptyset$ (e.g., P_B is bounded). The infimum of problem (6.1.2) is finite and attained at a pair of vertices of P_A and Q_B.*

Clearly, by application of Putinar's Positivstellensatz (also allowing equality constraints) the sum of squares relaxation converges asymptotically to the optimal value of problem (6.1.2). Unfortunately, so far it is not clear whether finite convergence in the strong containment case occurs.

6.2 A Bilinear Formulation of the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Containment Problem

The $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem is slightly more involved than the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ problem as the natural extension of Theorem 6.1.1 fails in general. However, under an additional assumption the statement holds.

Throughout the section, let

$$A(x,y) = A_0 + \sum_{i=1}^d A_i x_i + \sum_{j=1}^m A'_j y_j \in \mathcal{S}^k[x,y] \text{ and } B(x,y') = B_0 + \sum_{i=1}^d B_i x_i + \sum_{j=1}^n B'_j y'_j \in \mathcal{S}^l[x,y']$$

be linear pencils with $y = (y_1,\ldots,y_m)$ and $y' = (y'_1,\ldots,y'_n)$ for $n \geq 1$. Denote the projection of the corresponding spectrahedra onto the x-variables by

$$\begin{aligned}\pi(S_A) &= \left\{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m:\ A(x,y) \succeq 0\right\} \\ \text{and}\quad \pi(S_B) &= \left\{x \in \mathbb{R}^d \mid \exists y' \in \mathbb{R}^n:\ B(x,y') \succeq 0\right\}.\end{aligned}$$

Recall from Section 2.3 that the projection of a spectrahedron is not necessarily closed and thus, in general, not a spectrahedron itself.

Define $\bar{\mathcal{B}} = \operatorname{span}\{B'_1,\ldots,B'_n\}$ and recall the equivalence

$$\langle B'_i, Z\rangle = 0\ \forall i \in [n] \iff Z \in \bar{\mathcal{B}}^\perp.$$

We state the following generalization of Theorem 6.1.1.

Theorem 6.2.1. *Let $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x,y') \in \mathcal{S}^l[x,y']$ be linear pencils such that $\pi(S_A) \neq \emptyset$. Define $\mathcal{Z} = \bar{\mathcal{B}}^\perp \cap \mathcal{S}_+^l$.*

(1) *If $\pi(S_A)$ is contained in $\pi(S_B)$, then $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$.*

(2) *If $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$, then $\pi(S_A) \subseteq \operatorname{cl}\pi(S_B)$.*

(3) *Assume the constraint qualification in Lemma 2.3.10 holds for the y'-part of $B(x,y')$ (i.e., whenever $\sum_{i=1}^n B'_i y_i \succeq 0$, then $\sum_{i=1}^n B'_i y_i = 0$). Then $\pi(S_A) \subseteq \pi(S_B)$ if and only if $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$.*

As in the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ problem, the projection variables y' of the outer spectrahedron do not appear in the polynomial formulation, only the corresponding coefficient matrices $B'_1, \ldots, B'_n$.

Proof.

To (1): Assume $\pi(S_A) \subseteq \pi(S_B)$. Let $x \in \pi(S_A) \subseteq \pi(S_B)$. Then there exists $y' \in \mathbb{R}^n$ such that $B(x, y') \succeq 0$. For all $Z \in \mathcal{Z}$ it holds that

$$\langle B(x,0), Z\rangle = \langle B(x,y'), Z\rangle \geq 0.$$

Since $x \in \pi(S_A)$ is arbitrary, $\langle B(x,0), Z\rangle$ is nonnegative on $\pi(S_A) \times \mathcal{Z}$.

To (2): Assume $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$. Let $x \in \pi(S_A)$ be fixed but arbitrary and set $B'_0 = B(x,0)$. By Farkas' Lemma 2.3.8, there exist $B''_0 \in \mathcal{S}^l$ and $y' \in \mathbb{R}^n$ such that $B''_0 + \sum_{i=1}^n B'_i y'_i \in \mathcal{S}^l_+$ and $\|B'_0 - B''_0\| < \varepsilon$ for all $\varepsilon > 0$. By letting ε tend to zero, there exists a sequence $(y'_\varepsilon)_\varepsilon \subseteq \mathbb{R}^n$ such that $\lim_{\varepsilon \to 0} B(x, y'_\varepsilon) \succeq 0$. As $x \in \pi(S_A)$ is arbitrary, the claim follows.

To (3): Assume $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$. Let $x \in \pi(S_A)$ be fixed but arbitrary. By Lemma 2.3.10 (i.e., Farkas' Lemma for cones under the constraint qualification stated in part (3) of the statement), the spectrahedron $\{y' \in \mathbb{R}^n \mid B'_0 + \sum_{i=1}^n B'_i y'_i \succeq 0\}$ is nonempty. Thus there exists $y' \in \mathbb{R}^n$ such that $B(x, y') \succeq 0$. □

Unfortunately, the reverse implication in part (1) of Theorem 6.2.1, or, equivalently, part (2) without taking the closure of $\pi(S_B)$ is generally not true as the next example shows.

Example 6.2.2. *Consider the linear pencil*

$$B(x,y') = \begin{bmatrix} -y'_1 & x & 0 \\ x & 1-y'_2 & 0 \\ 0 & 0 & -x+y'_2 \end{bmatrix} = \begin{bmatrix} 0 & x & 0 \\ x & 1 & 0 \\ 0 & 0 & -x \end{bmatrix} + y'_1 \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + y'_2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and let $A(x)$ be the univariate linear pencil

$$A(x) = \begin{bmatrix} 1-x & 0 \\ 0 & 1+x \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

describing the interval $S_A = [-1,1]$. By inspecting the principal minors of B, the spectrahedron S_B has the form $\{(x,y') \in \mathbb{R}^3 \mid y'_1 \leq 0,\ x \leq y'_2 \leq 1,\ y'_1(1-y'_2)+x^2 \leq 0\}$. For $x = 1$, the second condition implies $y'_2 = 1$ and thus the third condition reads as $x^2 \leq 0$, a contradiction. Thus $S_A \not\subseteq \pi(S_B)$.

For every $Z \in \bar{\mathcal{B}}^\perp \cap \mathcal{S}^3_+$ it holds that

$$0 = \langle Z, B'_1\rangle = -Z_{11} \Longrightarrow Z_{12} = 0, \quad 0 = \langle Z, B'_2\rangle = Z_{33} - Z_{22}$$

implying $\langle B(x,0), Z\rangle = Z_{22} + x(-Z_{33} + 2Z_{12}) = Z_{22}(1-x) \geq 0$ for all $x \in S_A$.

It should not be surprising that the constraint qualification on the pencil $B(x,y')$ is not satisfied. Indeed, for $(y'_1, y'_2) = (y'_1, 0)$ with $y'_1 < 0$,

$$B'_1 y'_1 + B'_2 y'_2 = \begin{bmatrix} -y'_1 & x & 0 \\ x & -y'_2 & 0 \\ 0 & 0 & -x+y'_2 \end{bmatrix} = \begin{bmatrix} -y'_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is positive semidefinite but not identically zero. □

An issue when considering the practical utility of Theorem 6.2.1 is the unboundedness of the set $\mathcal{Z}$. Under an analog condition as in Theorem 6.1.1, $\mathcal{Z}$ can be replaced by the spectrahedral analog of the simplex. Recall from the preliminary Section 2.3 that $\mathbb{T}^l = \{Z \in \mathcal{S}^l_+ \mid \langle I_l, Z\rangle = 1\}$ is a nonempty spectratope, called spectraplex (2.3.16).

Corollary 6.2.3. *Let $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x,y') \in \mathcal{S}^l[x,y']$ be linear pencils such that $\pi(S_A) \neq \emptyset$. Assume $\mathcal{Z} = \bar{\mathcal{B}}^\perp \cap \mathcal{S}^l_+ \neq \{0\}$ and let $\mathbb{T}^l = \{Z \in \mathcal{S}^l_+ \mid \langle I_l, Z\rangle = 1\}$. Then $\pi(S_A) \subseteq \pi(S_B)$ if and only if $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times (\bar{\mathcal{B}}^\perp \cap \mathbb{T})$.*

Proof. Since $\bar{\mathcal{B}}^\perp \cap \mathbb{T} \subseteq \mathcal{Z}$, the “only if”-part follows from Theorem 6.2.1.

For the converse, first suppose there exists $(x, Z) \in \pi(S_A) \times \mathcal{Z}$ such that $\langle B(x,0), Z\rangle < 0$. Then $0 \neq Z \in \mathcal{S}^l_+$ and thus $\mathrm{tr}(Z) = \langle I_l, Z\rangle > 0$. This implies $\langle B(x,0), Z'\rangle < 0$ for $Z' = \frac{Z}{\mathrm{tr}(Z)}$ with $\mathrm{tr}(Z') = \langle I_l, Z'\rangle = \frac{1}{\mathrm{tr}(Z)}\langle I_l, Z\rangle = 1$.

Assume $\langle B(x,0), Z\rangle \geq 0$ on $\pi(S_A) \times \mathcal{Z}$. By assumption, there exists $0 \neq Z \in \mathcal{Z} = \bar{\mathcal{B}}^\perp \cap \mathcal{S}^l_+$. Applying the above scaling, the claim follows. ☐

Clearly, as for the polynomial formulation of the $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment problem, an optimization version of Theorem 6.2.1 respectively Corollary 6.2.3 is available.

Restricting the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem to the special case $\pi\mathcal{S}$-in-$\pi\mathcal{H}$ allows to state improved versions of Theorem 6.2.1 and Corollary 6.2.3.

Proposition 6.2.4. *Let $\pi(P_B)$ be as in* (6.1.1) *and let $A(x,y) \in \mathcal{S}^l[x,y]$ be a linear pencil.*

(1) *$\pi(S_A) \subseteq \pi(P_B)$ if and only if $z^T(b + Bx) \geq 0$ on $\pi(S_A) \times (\ker(B'^T) \cap \mathbb{R}^l_+)$.*

(2) *Assume $\ker(B'^T) \cap \mathbb{R}^l_+ \neq \{0\}$. Then $\pi(S_A) \subseteq \pi(P_B)$ if and only if $z^T(b + Bx) \geq 0$ on $\pi(S_A) \times (\ker(B'^T) \cap \Delta^l)$.*

Proof. $\pi(S_A) \not\subseteq \pi(P_B)$ if and only if there exists $x \in \pi(S_A)$ such that $\nexists y' \in \mathbb{R}^n : \ b + Bp + B'y' \geq 0$. By Farkas' Lemma 2.2.1, this is equivalent to the existence of a $z \in \mathbb{R}^l_+$ with $z^T B' = 0$ and $z^T(b + Bp) < 0$. The claim follows as in the proofs of Theorem 6.2.1 and Corollary 6.2.3. ☐

6.2.1 An Alternative Approach to the $\pi\mathcal{S}$-in-$\pi\mathcal{S}$ Containment Problem

Inspecting Example 6.2.2, the cause of failure in the sufficiency part of the polynomial feasibility criterion in Theorem 6.2.1 is the absence of a clean Farkas' Lemma for linear pencils. However, as seen in Chapter 2, the gap in Farkas' Lemma can be closed. Using Ramana's Lemma 2.3.11, we get the following polynomial reformulation of the containment problem for projected spectrahedra.

Proposition 6.2.5. *Let $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x,y') \in \mathcal{S}^l[x,y']$ be linear pencils such that $\pi(S_A) \neq \emptyset$. Let $\mathcal{W}_l$ be as defined in Ramana's Lemma 2.3.11. Then $\pi(S_A)$ is contained in $\pi(S_B)$ if and only if*

$$\langle B(x,0), U + W\rangle \geq 0 \text{ on } \pi(S_A) \times \mathcal{S}^l_+ \times \mathcal{W}_l \text{ with } \langle B'_i, U + W\rangle = 0 \ \forall\, i \in [n].$$

Proof. $\pi(S_A) \not\subseteq \pi(S_B)$ if and only if there there exists $x \in \pi(S_A)$ such that there is no $y' \in \mathbb{R}^n$ with $B(x,y') \succeq 0$. By Ramana's Lemma 2.3.11, for fixed but arbitrary $x \in \pi(S_A)$, this is equivalent to the existence of a tuple $(U, W) \in \mathcal{S}^l_+ \times \mathcal{W}_l$ with $\langle B'_i, U + W\rangle = 0$ for $i \in [n]$ and $\langle B(x,0), U + W\rangle < 0$, implying the claim. ☐

From a computational point of view, the approach based on Ramana's Lemma needs a too high amount of extra variables even in the polynomial formulation of the problem. Alternatively, one could use the recent approach of Liu and Pataki [LP14] based on "elementary reformulations“ (of the dual spectrahedron).

6.3 Positivstellensatz Certificates for the $\pi\mathcal{S}$-in-$\mathcal{S}$ Containment Problem

Retreating to the special cases $\pi\mathcal{H}$-in-$\mathcal{H}$ and $\pi\mathcal{S}$-in-$\mathcal{S}$ allows to bring forward several results from the non-projected case. We start with the first problem. Afterwards, we state and prove a Positivstellensatz for the second problem.

6.3.1 From the $\pi\mathcal{H}$-in-$\mathcal{H}$ to the $\pi\mathcal{S}$-in-$\mathcal{S}$ Containment Problem

As for the non-projected case (Section 5.1.1), we start with the polyhedral situation in Theorem 6.3.1. It also serves as an algorithmic proof of Theorem 3.2.3, classifying the $\pi\mathcal{H}$-in-$\mathcal{H}$ containment problem as decidable in polynomial time. As the proofs of the statements in this section are very similar to the ones given in Section 5.1.1, we only stress the emerging differences in the proofs.

Theorem 6.3.1. *Consider $P_A = \{(x,y) \in \mathbb{R}^{d+m} \mid a + Ax + A'y \geq 0\}$ and $P_B = \{x \in \mathbb{R}^d \mid b + Bx \geq 0\}$. Let P_A be nonempty. Denote by $\widehat{a} + \widehat{A}x + \widehat{A}'y$ the extended representation of P_A; see* (5.1.1).

(1) *$\pi(P_A)$ is contained in P_B if and only if there exists a nonnegative matrix $C \in \mathbb{R}^{l\times(k+1)}$ with $b = C\widehat{a}$, $B = C\widehat{A}$, and $0 = C\widehat{A}'$.*

(2) *Let P_A be a polytope that is not a singleton. Then $\pi(P_A)$ is contained in P_B if and only if there exists a nonnegative matrix $C \in \mathbb{R}^{l\times k}$ with $b = Ca$, $B = CA$, and $0 = CA'$.*

If specialized to polyhedra (i.e., $m = 0$), Theorem 6.3.1 coincides with Theorem 5.1.1. Recall from Example 5.1.2 that the preconditions in part (2) cannot be dropped.

Proof. If $B = CA$, $0 = CA'$, and $b = Ca$ (or $B = C\widehat{A}$, $0 = C\widehat{A'}$, and $b = C\widehat{a}$) with a nonnegative matrix C, for any $x \in \pi(P_A)$ we have

$$b + Bx + 0y = C\left(a + Ax + A'y\right) \geq 0,$$

i.e., $\pi(P_A) \subseteq P_B$.

Conversely, if $\pi(P_A) \subseteq P_B$, then any of the linear polynomials $(b + Bx + 0y)_i$, $i \in [l]$, is nonnegative on P_A. Hence, by Lemma 2.2.2, $(b + Bx + 0y)_i$ can be written as a linear combination

$$(b + Bx + 0y)_i \ = \ c'_{i0} + \sum_{j=1}^{k} c'_{ij}(a + Ax + A'y)_j \ = \ \sum_{j=0}^{k} c'_{ij}(\widehat{a} + \widehat{A}x + \widehat{A}'y)_j$$

with nonnegative coefficients c'_{ij}. Comparing coefficients yields $b_i = c'_{i0} + \sum_{j=1}^{k} c'_{ij}$ for $i \in [l]$, implying part (1) of the statement.

To prove the second part, recall from the proof of Theorem 5.1.1 that after an appropriate translation we can w.l.o.g. assume that $a \geq 0$. Then Stiemke's Transposition Theorem (cf.

Lemma 2.2.4) implies the existence of a $\lambda > 0$ such that $[A^T, A'^T]\lambda = 0$, and thus

$$\lambda^T(a + Ax + A'y) = \lambda^T a = 1$$

after an appropriate rescaling. By multiplying that equation with c'_{i0} from above, we obtain nonnegative c''_{ij} with $\sum_{j=1}^k c''_{ij}(a + Ax + A'y)_j = c'_{i0}$, yielding

$$(b + Bx)_i \;=\; \sum_{j=1}^{k}(c'_{ij} + c''_{ij})(a + Ax + A'y)_j.$$

Hence, $C = (c_{ij})_{i,j=1}^k$ with $c_{ij} := c'_{ij} + c''_{ij}$ is a nonnegative matrix with $B = CA$, $0 = CA'$, and $(Ca)_i = \sum_{j=1}^k (c'_{ij} + c''_{ij})a_j = b_i - c'_{i0} + c'_{i0}\,\lambda^T a = b_i$ for every $i \in [l]$. □

As in the non-projected case, the sufficiency part of Theorem 6.3.1 can be extended to the case of projected spectrahedra via the normal form of a (projected) polytope P_A as a (projected) spectrahedron,

$$\pi(P_A) \;=\; \left\{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^m : \; A(x, y) = \operatorname{diag}(a_1(x, y), \ldots, a_k(x, y)) \succeq 0\right\},$$

where $a_i(x, y)$ is the ith entry of the vector $a + Ax + A'y$. This yields the following analog of Corollary 5.1.4

Corollary 6.3.2. *Let $A(x, y) \in \mathcal{S}^k[x, y]$ and $B(x) \in \mathcal{S}^l[x]$ be the normal form of polyhedra* (2.3.2)*.*

(1) *$\pi(S_A)$ is contained in S_B if and only if the system*

$$C \succeq 0, \;\forall p \in [0, d] : \; B_p = \sum_{i=1}^{k}(A_p)_{ii}C_{ii}, \;\forall p \in [m] : \; 0 = \sum_{i=1}^{k}(A'_p)_{ii}C_{ii}, \tag{6.3.1}$$

has a solution with respect to $\widehat{A}(x, y)$ and $B(x)$.

(2) *Let S_A be a polytope that is not a singleton. $\pi(S_A) \subseteq S_B$ if and only if system* (6.3.1) *has a solution with respect to $A(x)$ and $B(x)$.*

We get an analog of Theorem 5.1.3 for the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem.

Theorem 6.3.3. *Let $A(x, y) \in \mathcal{S}^k[x, y]$ and $B(x) \in \mathcal{S}^l[x]$ be a linear pencils. Denote by $\pi(S_A)$ the coordinate projection of the spectrahedron S_A. $\pi(S_A)$ is contained in S_B if there exists a symmetric matrix $C = (C_{ij})_{i,j=1}^k \in \mathcal{S}^{kl}$ such that*

$$C \succeq 0, \; B_p = \sum_{i,j=1}^{k}(A_p)_{ij}C_{ij} \textit{ for } p \in [0, d], \; \sum_{i,j=1}^{k}(A'_p)_{ij}C_{ij} = 0 \textit{ for } p \in [m]. \tag{6.3.2}$$

Proof. As in the proof of Theorem 5.1.3, we get

$$\begin{aligned} B(x) &= B_0 + \sum_{p=1}^{d} x_p B_p + \sum_{p=1}^{m} y_p 0 \\ &= \sum_{i,j=1}^{k} (A(x, y))_{ij}\, C_{ij} = \mathbb{I}^T \left((A(x, y))_{ij}\, C_{ij}\right)_{i,j=1}^{k} \mathbb{I} \end{aligned} \tag{6.3.3}$$

with $\mathbb{I} = [I_l, \dots, I_l]^T$.

Let $x \in \pi(S_A)$. By definition, there exists $y \in \mathbb{R}^m$ such that $A(x,y) \succeq 0$. Thus, as in the proof of Theorem 5.1.3, $A(x,y) \otimes C$ is positive semidefinite. Since $((A(x,y))_{ij} C_{ij})_{i,j=1}^k$ is a principal submatrix of $A(x,y) \otimes C$, $B(x)$ is positive semidefinite as well. □

6.3.2 The Naive Way

Consider the linear pencils $A(x,y) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$. Then $\pi(S_A)$ is contained in S_B if and only if $B(x) \succeq 0$ on $\pi(S_A)$. Equivalently, the quantified polynomial optimization problem

$$\begin{aligned} \mu^* \;=\; & \sup \;\; \mu \\ & \text{s.t.} \;\; B(x) - \mu I_l \succeq 0 \;\forall x \in \pi(S_A) \end{aligned} \tag{6.3.4}$$

has a nonnegative optimal value.

In order to relax the above optimization problem, consider, for $l \in \mathbb{N}$, the quadratic module $\mathrm{QM}^l(A)$ associated to $A(x,y)$,

$$\mathrm{QM}^l(A) = \left\{ S_0 + \langle S, A(x,y) \rangle_l \;\mid\; S_0 \in \Sigma^l[x,y],\; S \in \Sigma^{kl}[x,y] \right\},$$

as defined in (2.4.7). Clearly, if $B(x) \in \mathrm{QM}^l(A)$ for a linear pencil $B(x) \in \mathcal{S}^l[x]$, then $\operatorname{cl}\pi(S_A) \subseteq S_B$. Thus Statements 5.1.8 and 5.1.10 for the Spectrahedron Containment problem remain valid in this case.

The drawback of this approach to the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem is that it relies on the geometry of the spectrahedron S_A rather than its projection, namely the boundedness assumption on S_A in the above mentioned statements as well as the appearance of the projection variables y in the quadratic module. In the next subsection, we address this by developing a refinement of Hol-Scherer's Positivstellensatz allowing to reduce the complexity of the problem significantly. Particularly, we can eliminate the variables y in the sense that they neither appear in the quadratic module nor in the relaxation.

Before that, we point to the fact that the bilinear approach and thus the Putinar based relaxation, as discussed in Section 5.5, can also be used in this case.

6.3.3 A more Sophisticated Positivstellensatz

Gouveia and Netzer [GN11] derived a Positivstellensatz for polynomials positive on the closure of a projected spectrahedron.

Proposition 6.3.4 ([GN11, Theorem 5.1])**.** *Let $A(x,y) \in \mathcal{S}^k[x,y]$ be a strictly feasible linear pencil. Define the quadratic module*

$$\mathrm{QM}(\pi, A) = \left\{ s_0 + \langle A(x,0), S \rangle \;\mid\; \langle A_i', S \rangle = 0 \text{ for } i \in [m],\; s_0 \in \Sigma[x],\; S \in \Sigma^k[x] \right\}.$$

If $\pi(S_A)$ is bounded, then $\mathrm{QM}(\pi, A)$ is Archimedean and thus contains all polynomials positive on the closure of $\pi(S_A)$.

So far neither a version of Proposition 6.3.4 for linear pencils positive definite on $\operatorname{cl}\pi(S_A)$ nor a version without preconditions, namely strict feasibility, is known. Subsequently, we state and proof an extension to linear pencils positive definite on a projected spectrahedron.

Define the quadratic module

$$\mathrm{QM}^l(\pi, A) = \left\{ S_0 + \langle S, A(x,0)\rangle_l \;\middle|\; \langle S, A_i'\rangle_l = 0 \;\forall i \in [m],\; S_0 \in \Sigma^l[x],\; S \in \Sigma^{kl}[x] \right\}. \quad (6.3.5)$$

It is easy to see that $\mathrm{QM}^l(\pi, A)$ is in fact a quadratic module. Note that $\mathrm{QM}^l(\pi, A)$ does not have to be finitely generated; see [GN11, Section 5]. Clearly, every element of $\mathrm{QM}^l(\pi, A)$ is positive semidefinite on the closure of $\pi(S_A)$.

Theorem 6.3.5. *Let $A(x, y) \in \mathcal{S}^k[x, y]$ be a strictly feasible linear pencil such that $\pi(S_A)$ is bounded. For $l \in \mathbb{N}$, the quadratic module $\mathrm{QM}^l(\pi, A)$ is Archimedean and thus contains every matrix polynomial positive definite on $\operatorname{cl} \pi(S_A)$.*

Proof. By boundedness of $\pi(S_A)$, there exists $N \in \mathbb{N}$ sufficiently large such that $N \pm x_i$ is nonnegative on $\pi(S_A)$ for all $i \in [d]$. If $N \pm x_i \in \mathrm{QM}^l(\pi, A)$ for all $i \in [d]$, then by Definition 2.4.2 the quadratic module is Archimedean. We show that under the preconditions in the theorem $\mathrm{QM}^l(\pi, A)$ contains every linear polynomial nonnegative on $\pi(S_A)$.

Let $b(x) = b_0 + b^T x \in \mathbb{R}[x]_1$ be a fixed but arbitrary affine linear polynomial nonnegative on $\pi(S_A)$. Consider the following primal-dual pair of SDPs.

$$\begin{aligned} p^* := \inf \;& b(x) \\ \text{s.t. }\; & A(x,y) \succeq 0 \end{aligned} \qquad\qquad \begin{aligned} \sup \;& \langle -A_0, Z\rangle \\ \text{s.t. }\; & \langle A_i, Z\rangle = b_i \;\forall i \in [d] \\ & \langle A_i', Z\rangle = 0 \;\forall i \in [m] \\ & Z \in \mathcal{S}^k_+ \end{aligned}$$

Since $A(x, y)$ is strictly feasible by assumption, the dual problem (on the right-hand side) has the optimal value $p^* - b_0$ and attains it; see Proposition 2.3.15. Since $b(x) \geq 0$ on $\pi(S_A)$, we have $-b_0 \leq p^* - b_0$ and thus

$$b_0 - z_0 = \langle A_0, Z\rangle\,, \;\langle A_i, Z\rangle = b_i \;\forall i \in [d],\; \langle A_i', Z\rangle = 0 \;\forall i \in [m]$$

for some $Z \in \mathcal{S}^k_+$ and $z_0 \geq 0$. Define $S(x)$ as the blockdiagonal $kl \times kl$-matrix with l copies of Z on its diagonal, i.e., $S(x) = \oplus_{j=1}^l Z$, and $S_0(x) = z_0 I_l$. Then

$$\begin{aligned} S_0(x) + \langle S(x), A(x,0)\rangle_l &= z_0 I_l + \bigoplus_{j=1}^{l} \langle Z, A(x,0)\rangle_l \\ &= z_0 I_l + \bigoplus_{j=1}^{l} \left(b_0 - z_0 + \sum_{i=1}^{d} b_i x_i \right) = b(x) I_l \end{aligned}$$

and $\langle S(x), A_i'\rangle_l = \oplus_{j=1}^l \langle Z, A_i'\rangle = 0$ for $i \in [m]$. This implies $b(x) \in \mathrm{QM}^l(\pi, A)$. By Hol-Scherer's Theorem, every matrix polynomial positive definite on $\operatorname{cl} \pi(S_A)$ is contained in $\mathrm{QM}^l(\pi, A)$. □

Theorem 6.3.5 leads to the following hierarchy for the $\pi\mathcal{S}$-in-$\mathcal{S}$ containment problem.

Proposition 6.3.6. *Let $A(x, y) \in \mathcal{S}^l[x, y]$ be a strictly feasible linear pencil such that $\pi(S_A)$ is bounded, and let $B(x) \in \mathcal{S}^l[x]$ be a linear pencil. Consider the truncated quadratic module*

$$\mathrm{QM}^l_t(\pi, A) = \left\{ S_0 + \langle S, A(x,0)\rangle_l \;\middle|\; \langle S, A_i'\rangle_l = 0 \;\forall i \in [m],\; S_0 \in \Sigma^l_t[x],\; S \in \Sigma^{kl}_t[x] \right\}. \quad (6.3.6)$$

If $\pi(S_A) \subseteq S_B$, then the optimal value of the problem

$$\begin{aligned} \mu(t) \ = \ \sup \ & \mu \\ \text{s.t.} \ & B(x) - \mu I_l \in \mathrm{QM}_t^l(\pi, A) \end{aligned}$$

converges to the optimal value of problem (6.3.4).

It is evident from the definition of the quadratic modules in Sections 6.3.2 and 6.3.3 that the latter approach is preferable to the naive way from the theoretical viewpoint (provided that $A(x, y)$ is strictly feasible). So far, it is not clear whether there is a connection between $B(x)$ being positive definite on $\operatorname{cl}\pi(S_A)$ and strong containment.

As in the non-projected case, the 0-th step of the hierarchy based on (6.3.5) is exactly the containment criterion in Theorem 6.3.3 when applied to the extended linear pencil (5.1.2).

Proposition 6.3.7. *Let $A(x, y) \in \mathcal{S}^k[x, y]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. Assume $A(x, y)$ is strictly feasible. The following are equivalent.*

(1) $B(x) \in \mathrm{QM}_0^l(\pi, A)$.

(2) *There exist $C' \in \mathcal{S}_+^{kl}$ and $C_0' \in \mathcal{S}_+^l$ such that*

$$B_0 = C_0' + \langle A_0, C' \rangle_l, \ B_p = \langle A_p, C' \rangle_l \ \forall p \in [d], \ 0 = \left\langle A_q', C' \right\rangle_l \ \forall q \in [m].$$

(3) *There exists $C \in \mathcal{S}_+^{(k+1)l}$ such that*

$$B_0 = \sum_{i,j=0}^{k} (\widehat{A}_0)_{ij} C_{ij}, \ B_p = \sum_{i,j=0}^{k} (\widehat{A}_p)_{ij} C_{ij} \ \forall p \in [d], \ 0 = \sum_{i,j=0}^{k} (\widehat{A}_q')_{ij} C_{ij} \ \forall q \in [m],$$

where $\widehat{A}(x, y)$ denotes the extended linear pencil (5.1.2).

In particular, the initial relaxation step (6.3.6) *certifies containment if and only if the semidefinite feasibility criterion* (6.3.2) *does when applied to the extended linear pencil $\widehat{A}(x)$.*

We skip the proof of Proposition 6.3.7 as it is very similar to the one for Theorem 5.1.11.

The proof of Theorem 6.3.5 evidently yields an analog of Theorem 5.2.3. This, in particular, shows the (theoretical) effectiveness of the approach based on Theorem 6.3.5.

Theorem 6.3.8. *Let $A(x, y) \in \mathcal{S}^k[x, y]$ be a strictly feasible linear pencil and let the coefficients of the linear pencil $B(x) \in \mathcal{S}^l[x]$ be simultaneously congruent to a diagonal matrix.*

(1) $\pi(S_A) \subseteq S_B$ *if and only if* $B(x) \in \mathrm{QM}_0^l(\pi, A)$.

(2) *Assume S_B is a polytope with nonempty interior. Then $\pi(S_A) \subseteq S_B$ if and only if $B(x) \in \mathrm{QM}_0^l(\pi, A)$ with $S_0 = 0$.*

In particular, the statements (1) and (2) hold for a diagonal linear pencil $B(x)$, i.e., a polyhedron in normal form (2.3.2).

In order to prove Theorem 6.3.8, we use natural adaptions of the auxiliary results on the behavior of the initial Hol-Scherer relaxation 5.1.6 with regard to block diagonalization and transitivity. Using Proposition 6.3.7, it is easy to verify the validity of these statements.

Proof. As in the proof of Theorem 5.2.3, we can retreat to the normal form (2.3.2) $B(x) = \bigoplus_{q=1}^l b^q(x) \in \mathcal{S}^l[x]$ with $b^q(x) = b_0^q + x^T b^q$ for $q \in [l]$. Denote by $b_0^q, b_1^q, \ldots, b_d^q$ the coefficients of the linear form $b^q(x) = (b_0 + Bx)_q$. Set $b^q := (b_1^q, \ldots, b_d^q)$.

The proof of Theorem 6.3.5 yields certificates

$$b_0^q - z_0^q = \langle A_0, Z^q \rangle \,, \; \langle A_i, Z^q \rangle = b_i^q \; \forall i \in [d], \; \langle A_i', Z^q \rangle = 0 \; \forall i \in [m]$$

for some $Z^q \in \mathcal{S}_+^k$ and $z_0^q \geq 0$. Setting $S(x) = \bigoplus_{q=1}^l Z^q$ and $S_0(x) = \bigoplus_{q=1}^l z_0^q$, this implies part (1) of the statement.

To prove the second part, let $S(x)$ as before and set $S_0(x)$ to be zero. Then

$$\langle S(x), A(x,0) \rangle_l = \bigoplus_{q=1}^l \langle A(x,0), Z^q \rangle_l = \bigoplus_{q=1}^l \left(f_0 - z_0^q + \sum_{i=1}^d b_i^q x_i \right)$$

certifies the containment $\pi(S_A) \subseteq S_{B'}$, where $B'(x)$ is defined as

$$B'(x) = \bigoplus_{q=1}^l \left(r^q + \sum_{p=1}^d x_p \right).$$

As in the proof of Theorem 5.2.3, assuming that S_B is a polytope, we have $S_{B'} \subseteq S_B$ and thus, by transitivity and exactness of the initial Hol-Scherer step for polytopes, there is a certificate for the containment question $\pi(S_A) \subseteq S_B$ of degree zero with $S_0(x) = 0$. □

As a special case of Theorem 6.3.5, we gain a Positivstellensatz for polynomials on projected polyhedra having boundedness as its only precondition.

Proposition 6.3.9. *Let $P_A = \{(x,y) \in \mathbb{R}^{d+m} \mid a + Ax + A'y \geq 0\}$ be a nonempty polyhedron such that $\pi(P_A)$ is bounded. The quadratic module*

$$\mathrm{QM}^1(\pi, A) = \left\{ s_0 + \sum_{i=1}^k s_i(x)(a+Ax)_i \mid \sum_{i=1}^k s_i(x)(A'_{i,j}) = 0 \; \forall j \in [m], \; s_0, \ldots, s_k \in \Sigma[x] \right\}$$

is Archimedean and thus contains every polynomial positive on $\pi(P_A)$.

Proof. The proof follows from the proof of Theorem 6.3.5 by retreating to diagonal pencils and the fact that strong duality holds for linear programming (cf. Proposition 2.2.5). □

6.3.4 Examples

We discuss some academic examples for the hierarchy stated in Proposition 6.3.6.

In the tables, "time (sec)" states the time in seconds for setting up the problem in YALMIP and solving it with Mosek (cf. Sections 2.3 and 2.4). All computations are made on a desktop computer with Intel Core i3-2100 @ 3.10 GHz and 4 GB of RAM.

Example 6.3.10. *The so-called TV screen (see, e.g., [BPT13, Section 6.3.1]) is the projection of the spectrahedron*

$$S_A = \left\{ (x,y) \in \mathbb{R}^{2+2} \mid A(x,y) = \begin{bmatrix} 1+y_1 & y_2 \\ y_2 & 1-y_1 \end{bmatrix} \oplus \begin{bmatrix} 1 & x_1 \\ x_1 & y_1 \end{bmatrix} \oplus \begin{bmatrix} 1 & x_2 \\ x_2 & y_2 \end{bmatrix} \succeq 0 \right\},$$

onto the x variables; see Figure 6.3.1a.

Note that while the TV screen is centrally symmetric (as its boundary equals the variety defined by the polynomial $1 - x_1^4 - x_2^4$), its defining spectrahedron is not (as the point $(1,0,1,0)$ is contained in S_A but its negative $(-1,0,-1,0)$ is not).

$\pi(S_A)$	rS_B	r	t	time (sec)	$\mu(t)$
TV screen	2-ball	1.18	0	0.8514	-0.0078
		$\sqrt[4]{2}$	0	2.0564	$-1.3621 \cdot 10^{-08}$
		1.19	0	0.9964	$6.6628 \cdot 10^{-04}$
		1.2	0	0.9854	0.0090
S_A	rS_B				
	4-ball	1.55	0	1.1036	-0.0024
		$\sqrt{\sqrt{2}+1}$	0	0.9373	$3.1037 \cdot 10^{-09}$
		1.56	0	1.0723	0.0040

Table 6.1: Computational test of containment for the TV screen in the 2-ball of radius r and for its defining spectrahedron in the 4-ball of radius r. See Example 6.3.10.

As one can see in Table 6.1, the circumradius of the TV screen is at most $\sqrt[4]{2} \approx 1.1892$*, while the "centrally symmetric circumradius" of the defining spectrahedron S_A is at most* $\sqrt{\sqrt{2}+1} \approx 1.5538$*. Here we used the normal form of a ball* (2.3.14)*.*

Actually, the computed values for the (centrally symmetric) circumradius are exact. For $p = \left(\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \in S_A$ *we have* $\|\pi(p)\|_2 = \sqrt[4]{2}$ *and* $\|p\|_2 = \sqrt{\sqrt{2}+1}$*, implying that the circumradius of the TV screen is at least* $\sqrt[4]{2}$ *and that* $\sqrt{\sqrt{2}+1}$ *is the smallest possible radius of a ball (centered at the origin) containing S_A.* ☐

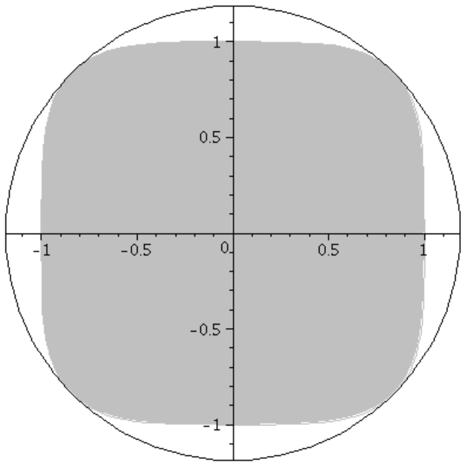
$\pi(S_A)$	rS_B	r	t	time (sec)	$\mu(t)$
two disks	2-ball	1.99	0	0.7978	-0.0050
		2	0	0.8215	$5.9978 \cdot 10^{-08}$
		2.01	0	0.9173	0.0050
S_A	rS_B				
	3-ball	2.23	0	0.8690	-0.0027
		2.2361	0	0.7470	$1.4339 \cdot 10^{-05}$
		2.24	0	0.8803	0.0018

Table 6.2: Computational test of containment as described in Example 6.3.11.

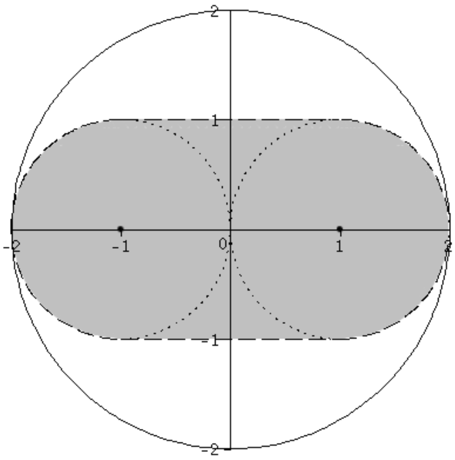
Example 6.3.11. *Let M be the convex hull of the shifted unit disks D_1 and D_2 defined by the identities* $1-(x_1+1)^2-x_2^2=0$ *and* $1-(x_1-1)^2-x_2^2=0$*, respectively; see Figure 6.3.1b. M is a projected spectrahedron. Indeed, considering only D_1 and shifting it along the segment* $[-1,1]\times\{0\}$ *yields M. Thus* $M = \{x \in \mathbb{R}^2 \mid \exists y \in \mathbb{R}:\ 1-(x-y)^2-x_2^2 \geq 0,\ -1 \leq y \leq 1\}$ *is the projection of the 3-dimensional cylinder, see Figure 6.3.1c, defined by the linear pencil*

$$A(x,y) = \begin{bmatrix} 1-x_2 & x_1-y \\ x_1-y & 1+x_2 \end{bmatrix} \oplus \begin{bmatrix} 1-y & 0 \\ 0 & 1+y \end{bmatrix}.$$

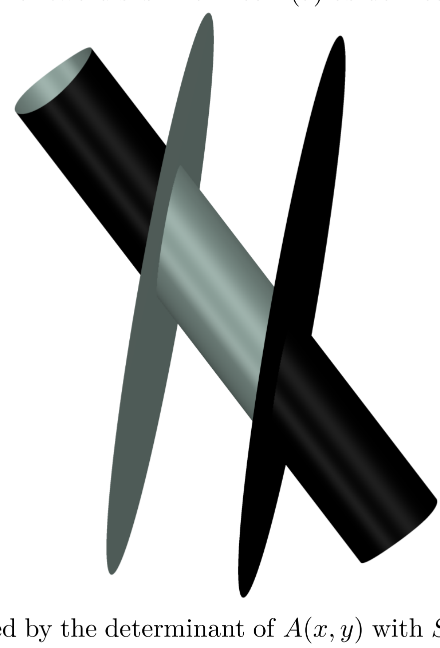
Consider the 2-ball of radius $r>0$. It follows from the construction of $M=\pi(S_A)$ that it is centrally symmetric and that its circumradius is 2. Up to numerical accuracy, this value is computed by our approach; see Table 6.2. ☐



(a) TV screen in a $\sqrt[4]{2}$-ball as defined in Example 6.3.10.



(b) The convex hull of two disks in a 2-ball (a) as defined in Example 6.3.11.



(c) The algebraic surface defined by the determinant of $A(x, y)$ with S_A being the grey cylinder in the middle of the picture as defined in Example 6.3.11.

Figure 6.3.1

6.4 Containment of Projected Spectrahedra and Positive Linear Maps

In this section, we discuss possible extentions of the concept of positive linear maps (as introduced in Section 5.4) to projected spectrahedra.

Given two linear pencils $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x,y') \in \mathcal{S}^l[x,y']$ with $y = (y_1,\dots,y_m)$ and $y' = (y'_1,\dots,y'_n)$, define the linear subspaces

$$\begin{aligned}
\mathcal{A} &= \operatorname{span}\{A_0,\dots,A_d,A'_1,\dots,A'_m\},\\
\mathcal{B} &= \operatorname{span}\{B_0,\dots,B_d\}, \text{ and } \bar{\mathcal{B}} = \operatorname{span}\{B'_1,\dots,B'_n\}.
\end{aligned}$$

Every element in $\mathcal{A}$ can be associated to a homogeneous linear pencil $A(x_0,x,y) \in \mathcal{S}^k[x_0,x,y]$. As introduced in (5.1.2), the linear pencil

$$\begin{aligned}
\widehat{A}(x_0,x,y) &:= x_0\widehat{A}_0 + \sum_{p=1}^{d} x_p\widehat{A}_p + \sum_{q=1}^{m} y_q\widehat{A}'_q\\
&:= x_0(1\oplus A_0) + \sum_{p=1}^{d} x_p(0\oplus A_p) + \sum_{q=1}^{m} y_q(0\oplus A'_q)
\end{aligned}$$

is called the *extended linear pencil* associated to $A(x_0,x,y)$. The associated linear subspace is $\widehat{\mathcal{A}} = \operatorname{span}\{\widehat{A}_0,\widehat{A}_1,\dots,\widehat{A}_d,\widehat{A}'_1,\dots,\widehat{A}'_m\}$.

For linearly independent $A_1,\dots,A_d,A'_1,\dots,A'_m$, let $\widehat{\Phi}_{AB} : \widehat{\mathcal{A}} \to \mathcal{B}$ be the linear map defined by

$$\widehat{\Phi}_{AB}(1\oplus A_0) = B_0,\ \forall p\in[d]:\ \widehat{\Phi}_{AB}(0\oplus A_p) = B_p \text{ and } \forall p\in[m]:\ \widehat{\Phi}_{AB}(0\oplus A'_p) = 0.$$

Note that since every linear combination $0 = \lambda_0(1\oplus A_0) + \sum_{p=1}^{d}\lambda_p(0\oplus A_p) + \sum_{q=1}^{m}\lambda_{d+q}(0\oplus A'_q)$ for real scalars $\lambda_0,\dots,\lambda_{d+m}$ yields $\lambda_0 = 0$, it suffices to assume the linear independence of the coefficient matrices $A_1,\dots,A_d,A'_1,\dots,A'_m$ to ensure that $\widehat{\Phi}_{AB}$ is well-defined. Recall the discussion on the lineality space from Section 2.3 (see also Section 5.4).

$\pi(S_A)$ is contained in $\pi(S_B)$ if and only if for all $(x_0,x,y) \in \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^m$ with $\widehat{A}(x_0,x,y) \in \widehat{\mathcal{A}} \cap \mathcal{S}_+^{k+1}$ the set $\left(\widehat{\Phi}(A(x_0,x,y)) + \bar{\mathcal{B}}\right) \cap \mathcal{S}_+^l = \left(B(x_0,x,0) + \bar{\mathcal{B}}\right) \cap \mathcal{S}_+^l$ is nonempty. The last statement says that the spectrahedron $S = \{y' \in \mathbb{R}^n \mid B(x_0,x;y') \succeq 0\}$ is nonempty for any fixed but arbitrary (x_0,x,y) with $(x,y) \in S_A$. A necessary condition for this is the non-existence of a matrix $Z \in \bar{\mathcal{B}}^\perp \cap \mathcal{S}_+^l$ with $\langle B(x_0,x,0), Z\rangle < 0$. Unfortunately, the reverse implication is not true as seen in Example 6.2.2 above.

However, if we restrict the outer set to be a spectrahedron (instead of a projection of one), then $\pi(S_A) \subseteq S_B$ if and only if for all $(x_0,x,y) \in \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^m$ with $A(x_0,x,y) \in \mathcal{A}\cap\mathcal{S}_+^k$ the single matrix $\Phi(A(x_0,x,y)) = B(x_0,x)$ is positive semidefinite. We formalize this in Theorem 6.4.1 below.

Given the linear pencils $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x) \in \mathcal{S}^l[x]$, let

$$\widehat{A}(x,y) = 1\oplus A(x) = 1\oplus A_0 + \sum_{p=1}^{d} x_p(0\oplus A_p) + \sum_{q=1}^{m} y_q(0\oplus A'_q)$$

be the extended linear pencil of $A(x,y)$. As in Section 5.4, if $A_0,A_1,\dots,A_{d+m}$ are linearly independent, then we can retreat to the simpler map $\Phi_{AB}: \mathcal{A} \to \mathcal{B}$ defined by

$$\forall p \in [d]: \ \Phi_{AB}: \ A_p \mapsto B_p \text{ and } \forall q \in [m]: \ \widehat{\Phi}_{AB}(A'_q) = 0.$$

We have the following extension of Proposition 5.4.6.

Theorem 6.4.1. *Let $A(x,y) \in \mathcal{S}^k[x,y]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils.*

(1) *If Φ_{AB} or $\widehat{\Phi}_{AB}$ is positive, then $\pi(S_A) \subseteq S_B$.*

(2) *If $\pi(S_A) \neq \emptyset$, then $\pi(S_A) \subseteq S_B$ implies positivity of $\widehat{\Phi}$.*

(3) *If $\pi(S_A) \neq \emptyset$ and S_A is bounded, then $\pi(S_A) \subseteq S_B$ implies positivity of Φ.*

The proof follows along the same lines as the proof of Proposition 5.4.6.

Proof. Set $\widehat{A}(x_0,x,y) := x_0(1 \oplus A_0) + \sum_{p=1}^d x_p(0 \oplus A_p) + \sum_{q=1}^m y_q(0 \oplus A'_q) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$.

To (1): Let Φ_{AB} be positive. For every point $x \in \pi(S_A)$ there exists $y \in \mathbb{R}^m$ such that $A(x,y) \succeq 0$, i.e., $A(x,y) \in \mathcal{S}_+^k \cap \mathcal{A}$. Then $B(x) = B(x) + \sum_{q=1}^m y_q \Phi(A'_q) = \Phi(A(x,y)) \in \mathcal{S}_+^l \cap \mathcal{B}$ and hence $x \in S_B$. There is no difference in the proof if $\widehat{\Phi}_{AB}$ is positive.

To (2): Since the spectrahedra defined by $A(x,y)$ and $\widehat{A}(x,y)$ coincide, their projections equal, and hence we have $\pi\left(S_{\widehat{A}}\right) \subseteq S_B$. Let $\widehat{A}(x_0,x,y) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$. Then $x_0 \geq 0$.

Case $x_0 > 0$. By scaling the linear pencil with $1/x_0$ the positive semidefiniteness is preserved. Thus, $\widehat{A}(1, x/x_0, y/x_0) = \widehat{A}(1, x/x_0, y/x_0) \in \mathcal{S}_+^{k+1} \cap \widehat{\mathcal{A}}$ and $x/x_0 \in \pi(S_A) \subseteq S_B$. Scaling $B(x/x_0)$ by x_0 yields $\widehat{\Phi}(\widehat{A}(x_0,x,y)) = x_0 B_0 + \sum_{p=1}^d x_p B_p = x_0 B(x/x_0) \in \mathcal{S}_+^l \cap \mathcal{B}$.

Case $x_0 = 0$. If $(x, x_0) = (0,0)$, the statement is obvious. Let $x \neq 0$. By assumption, $\pi(S_A)$ is nonempty. Fix a point $\bar{x} \in \pi(S_A)$. Then, for some $\bar{y}, y \in \mathbb{R}^m$, $\widehat{A}(1, \bar{x}+tx, \bar{y}+ty) = \widehat{A}(1,\bar{x},\bar{y}) + \widehat{A}(0,tx,ty) \succeq 0$ for all $t > 0$, implying $\bar{x}+tx \in \pi(S_A) \subseteq S_B$ for all $t > 0$. Thus x is a point of the recession cone of $\pi(S_A)$ which clearly is contained in the recession cone of S_B. Indeed, $\frac{1}{t}B(1,\bar{x}) + B(0,x) = \frac{1}{t}B(\bar{x}+tx) \succeq 0$ for all $t > 0$. By closedness of the cone of positive semidefinite matrices, we get $B(0,x) \succeq 0$. Hence, $\widehat{\Phi}(\widehat{A}(x_0,x,y)) = \widehat{\Phi}(\widehat{A}(0,x,y)) = B(0,x) \succeq 0$.

To (3): Let $A(x_0,x,y) = x_0 A_0 + \sum_{p=1}^d x_p A_p + \sum_{q=1}^m y_q A'_q$ be in $\mathcal{S}_+^k \cap \mathcal{A}$. We distinguish between the following two cases.

Case $x_0 > 0$. This case follows by a similar scaling argument as in part (2).

Case $x_0 \leq 0$. Since $\pi(S_A)$ is nonempty, there exists $\bar{x} \in \pi(S_A)$ and hence, for some $\bar{y} \in \mathbb{R}^m$,

$$\begin{aligned} A(0, x+|x_0|\bar{x}, y+|x_0|\bar{y}) &= A(0,x,y) + A(0, |x_0|\bar{x}, |x_0|\bar{y}) \\ &\succeq |x_0| A_0 + A(0,|x_0|\bar{x},|x_0|\bar{y}) = |x_0| \cdot A(1,\bar{x},\bar{y}) \succeq 0. \end{aligned}$$

For $A(0, x+|x_0|\bar{x}, y+|x_0|\bar{y}) \neq 0$, one has an improving ray of the spectrahedron S_A, in contradiction to the boundedness of S_A. For $A(0,x+|x_0|\bar{x},y+|x_0|\bar{y}) = 0$, linear independence of $A_0, \ldots, A_{d+m}$ implies $(x+|x_0|\bar{x}, y+|x_0|\bar{y}) = (0,0)$. But then $x_0 A(1,\bar{x},\bar{y}) = A(x_0,x,y) \succeq 0$ together with $x_0 \leq 0$ and $A(1,\bar{x},\bar{y}) \succeq 0$ imply either $A(1,\bar{x},\bar{y}) = 0$, in contradiction to linear independence, or $(x_0, x) = 0$. Clearly,in this case, $\Phi_{AB}(0) = 0$. □

If, as in Lemma 5.4.8, our setting is changed from the case of linear subspaces to the case of affine subspaces, Theorem 6.4.1 has a slightly easier formulation. As the proof is very similar to the one given for Lemma 5.4.8, we omit it.

Lemma 6.4.2. *Let $A(x,y) \in \mathcal{S}^k[x]$ and $B(x) \in \mathcal{S}^l[x]$ be linear pencils. Define the affine subspaces $\bar{\mathcal{A}} = \frac{1}{d}A_0 + \operatorname{span}(A_1, \ldots, A_d, A'_1, \ldots, A'_m)$ and $\bar{\mathcal{B}} = \frac{1}{d}B_0 + \operatorname{span}(B_1, \ldots, B_d)$ for linearly independent $A_1, \ldots, A_d, A'_1, \ldots, A'_m$. Then $\pi(S_A) \subseteq S_B$ if and only if the affine function $\bar{\Phi}_{AB} : \bar{\mathcal{A}} \to \bar{\mathcal{B}}$ defined by $\frac{1}{d}A_0 + A_p \mapsto \frac{1}{d}B_0 + B_p$ for $i \in [d]$ and $\frac{1}{d}A_0 + A'_p \mapsto \frac{1}{d}B_0$ for $i \in [m]$ is positive.*

7 Final Remarks and Open Questions

We close with a short discussion of open questions related to the containment problems studied in the previous chapters.

We studied Handelman certificates and Putinar certificates for Polytope Containment. In Theorem 4.3.6, we saw that the Putinar relaxation always finitely converges (under mild preconditions). Does the Handelman relaxation always finitely converge for Polytope Containment? Note that by Theorem 4.1.3, the optima of the bilinear programming formulation are always attained at the boundary, which would allow for a positive answer to the question. (Recall that the Handelman hierarchy cannot finitely converge whenever there exists an optimizer in the interior of the feasible region).

For the Polytope Containment problem, can the structure of the certificates be better characterized? Such as, what are improved degree bounds with regard to Polytope Containment or, somewhat more general, with regard to general bilinear programming problems? How is Fourier-Motzkin-elimination (as an $\mathcal{H}$-in-$\mathcal{V}$ conversion algorithm) related to the Handelman certificates for Polytope Containment?

The treatment of the $\mathcal{H}$-in-$\mathcal{V}$ and $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment problems show that disjointly constraint bilinear programming is intrinsically linked to geometric problems. On this basis, it is natural to ask for a Positivstellensatz for bilinear polynomials on linearly constrained sets. In this context, can the finite convergence result for the Putinar relaxation in the $\mathcal{H}$-in-$\mathcal{V}$ setting be extended to $\pi\mathcal{H}$-in-$\pi\mathcal{H}$ containment, or more generally, to bilinear programming problems with disjointly constrained sets?

We stated an asymptotically convergent hierarchy to decide the Spectrahedron Containment problem based on Hol-Scherer's Positivstellensatz. If the outer set is given by a diagonal linear pencil (describing an $\mathcal{H}$-polyhedron) and some mild and common preconditions on the inner linear pencil hold, the initial relaxation step decides containment. As the proof is intrinsically based on Farkas' Lemma, we ask for a more complete statement without preconditions using Ramana's Lemma. For that, the sum of squares formulation of SDP duality by Klep and Schweighofer [KS11, KS13] might be interesting.

Concerning the general Spectrahedron Containment problem, can the convergence behavior of the Hol-Scherer relaxation be characterized, in particular, with regard to the initial relaxation step? Such as, are there properties of the linear pencils serving as sufficient conditions for the initial relaxation step to be exact? Moreover, regarding the connection of Spectrahedron Containment and positivity of linear maps, we ask for the relation between the exactness of the tth relaxation step and $(k-t)$-positivity of the particular linear map.

Regarding the $\mathcal{S}$-in-$\pi\mathcal{S}$ containment problem, we showed a refinement of Hol-Scherer's Positivstellensatz under the additional assumption of strict feasibility. Can the precondition in Theorem 6.3.5 be dropped? Note that this question is related to the above discussion of exactness for the $\mathcal{S}$-in-$\mathcal{H}$ containment problem using Ramana's Lemma. The questions from the non-projected case can also be stated here. Moreover, (under what preconditions) does strong containment imply the existence of a Hol-Scherer certificate?

Bibliography

[AA00] E. Andersen and K. Andersen. The MOSEK optimization software. *EKA Consulting ApS, Denmark*, 2000.

[ABS97] D. Avis, D. Bremner, and R. Seidel. How good are convex hull algorithms? *Comput. Geom.*, 7(5-6):265–301, 1997. 11th ACM Symposium on Computational Geometry (Vancouver, BC, 1995).

[AGV11] M.F. Anjos, B. Ghaddar, and J.C. Vera. An iterative scheme for valid polynomial inequality generation in binary polynomial programming. In O. Günlük and G.J. Woeginger, editors, *Integer Programming and Combinatorial Optimization*, volume 6655 of *Lecture Notes in Comput. Sci.*, pages 207–222. Springer, Heidelberg, 2011.

[Ali93] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5:13–51, 1993.

[Ave13] G. Averkov. Constructive proofs of some Positivstellensätze for compact semialgebraic subsets of $\mathbb{R}^d$. *J. Optim. Theory Appl.*, 158(2):410–418, 2013.

[Bar12] A. Barvinok. Approximations of convex bodies by polytopes and by projections of spectrahedra. Preprint, arXiv:1204.0471, 2012.

[BBE$^+$08] E. Boros, K. Borys, K. Elbassioni, V. Gurvich, and L. Khachiyan. Generating all vertices of a polyhedron is hard. *Discrete Comput. Geom.*, 39(1-3):174–190, 2008.

[BCS99] E. Bannai, J. Conway, and N.J.A. Sloane. *Sphere Packings, Lattices and Groups.* A series of comprehensive studies in mathematics. Springer, 1999.

[Bix12] R.E. Bixby. A brief history of linear and mixed-integer programming computation. In Martin Grötschel, editor, *Optimization stories*, Doc. Math., J. DMV, extra vol. 2012.

[BK01] C. Böhm and H.-P. Kriegel. Determining the convex hull in large multidimensional databases. In Y. Kambayashi, W. Winiwarter, and M. Arikawa, editors, *Data Warehousing and Knowledge Discovery*, volume 2114, pages 294–306. Springer, 2001.

[Ble12] G. Blekherman. Nonnegative polynomials and sums of squares. *J. Amer. Math. Soc.*, 25(3):617–635, 2012.

[BPT13] G. Blekherman, P.A. Parrilo, and R. Thomas. *Semidefinite Optimization and Convex Algebraic Geometry.* SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2013.

[BRS11] A. Bhardwaj, P. Rostalski, and R. Sanyal. Deciding polyhedrality of spectrahedra. Preprint, arXiv:1102.4367, 2011.

[BSS06] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming.* Wiley-Interscience, Hoboken, NJ, third edition, 2006.

[BTN02] A. Ben-Tal and A. Nemirovski. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM J. Optim.*, 12(3):811–833, 2002.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, New York, NY, USA, 2004.

[Cho72] M.-D. Choi. Positive linear maps on C*-algebras. *Can. J. Math.*, 24(2):520–529, 1972.

[Cho75a] M.-D. Choi. Positive linear maps on complex matrices. *Linear Algebra Appl.*, 10:285–290, 1975.

[Cho75b] M.-D. Choi. Positive semidefinite biquadratic forms. *Linear Algebra Appl.*, 12(2):95 – 100, 1975.

[CKL92] S.J. Cho, S.-H. Kye, and S.G. Lee. Generalized choi maps in three-dimensional matrix algebra. *Linear Algebra Appl.*, 171:213–224, 1992.

[CT14] L. Công-Trình. Some Positivstellensätze for polynomial matrices. *Positivity*, pages 1–16, 2014. Online Version.

[dK02] E. de Klerk. *Aspects of Semidefinite Programming*, volume 65 of *Applied Optimization.* Kluwer Academic Publishers, Dordrecht, 2002.

[FO85] R.M. Freund and J.B. Orlin. On the complexity of four polyhedral set containment problems. *Math. Program.*, 33(2):139–145, 1985.

[GGK+13] J. Gouveia, R. Grappe, V. Kaibel, K. Pashkovich, R. Robinson, and R. Thomas. Which nonnegative matrices are slack matrices? *Linear Algebra Appl.*, 439:2921–2933, 2013.

[GK89] P. Gritzmann and V. Klee. On the 0–1-maximization of positive definite quadratic forms. In *Operations Research Proceedings 1988*, pages 222–227. Springer, 1989.

[GK92] P. Gritzmann and V. Klee. Inner and outer j-radii of convex bodies in finite-dimensional normed spaces. *Discrete Comput. Geom.*, 7(1):255–280, 1992.

[GK93] P. Gritzmann and V. Klee. Computational complexity of inner and outer j-radii of polytopes in finite-dimensional normed spaces. *Math. Program.*, 59(2, Ser. A):163–213, 1993.

[GK94] P. Gritzmann and V. Klee. On the complexity of some basic problems in computational convexity. I. Containment problems. *Discrete Math.*, 136(1-3):129–174, 1994.

[GKL95] P. Gritzmann, V. Klee, and D. Larman. Largest j-simplices in n-polytopes. *Discrete Comput. Geom.*, 13(3-4):477–515, 1995.

[GM07] B. Gärtner and J. Matousek. *Understanding and Using Linear Programming.* Springer, 2007.

[GM12] B. Gärtner and J. Matousek. *Approximation Algorithms and Semidefinite Programming.* Springer, 2012.

[GN11] J. Gouveia and T. Netzer. Positive polynomials and projections of spectrahedra. *SIAM J. Optim.*, 21(3):960–976, 2011.

[GPT10] J. Gouveia, P.A. Parrilo, and R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20(4):2097–2118, 2010.

[GR95] A.J. Goldman and M.V. Ramana. Some geometric results in semidefinite programming. *J. Global Optim.*, 7(1):33–50, 1995.

[GU77] G. Gallo and A. Ülkücü. Bilinear programming: An exact algorithm. *Math. Program.*, 12(1):173–194, 1977.

[GVL96] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

[GW95] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.

[Han88] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.*, 132(1):35–62, 1988.

[Hen08] D. Henrion. On semidefinite representations of plane quartics. Technical Report 08444, LAAS-CNRS, 2008.

[HJ94] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[HKM12] J.W. Helton, I. Klep, and S. McCullough. The convex Positivstellensatz in a free algebra. *Adv. Math.*, 231(1):516 – 534, 2012.

[HKM13] J.W. Helton, I. Klep, and S. McCullough. The matricial relaxation of a linear matrix inequality. *Math. Program.*, 138(Ser. A):401–445, 2013.

[HKMS14] J.W. Helton, I. Klep, S. McCullough, and M. Schweighofer. Inclusion of spectrahedra, dilations, the matrix cube problem and coin tossing. Preprint, arXiv:1412.1481, 2014.

[HL06] D. Henrion and J.-B. Lasserre. Convergent relaxations of polynomial matrix inequalities and static output feedback. *IEEE Trans. Autom. Control*, 51(2):192–202, 2006.

[HM04] J.W. Helton and S. McCullough. A Positivstellensatz for non-commutative polynomials. *Trans. Amer. Math. Soc.*, 356(9):3721–3737, 2004.

[HN10] J.W. Helton and J. Nie. Semidefinite representation of convex sets. *Math. Program.*, 122(1, Ser. A):21–64, 2010.

[HN12] J.W. Helton and J. Nie. Semidefinite representation of convex sets and convex hulls. In M.F. Anjos and J.-B. Lasserre, editors, *Handbook of Semidefinite, Conic and Polynomial Programming*, pages 77–112. Springer, New York, 2012.

[HS04] C.W.J. Hol and C.W. Scherer. Sum of squares relaxations for polynomial semidefinite programming. *Proc. Symp. on Mathematical Theory*, pages 1–10, 2004.

[HS06] C.W.J. Hol and C.W. Scherer. Matrix sum-of-squares relaxations for robust semidefinite programs. *Math. Program.*, 107(1-2):189–211, 2006.

[HV07] J.W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.*, 60(5):654–674, 2007.

[JZ04] M. Joswig and G.M. Ziegler. Convex hulls, oracles, and homology. *J. Symb. Comp.*, 38(4):1247 – 1259, 2004. Symbolic Computation in Algebra and Geometry.

[Kha80] L.G. Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20(1):53–72, 1980.

[Kon76] H. Konno. A cutting plane algorithm for solving bilinear programs. *Math. Program.*, 11(1):14–27, 1976.

[KP97] L. Khachiyan and L. Porkolab. On the complexity of semidefinite programs. *J. Global Optim.*, 10(4):351–365, 1997.

[KP03] V. Kaibel and M.E. Pfetsch. Some algorithmic problems in polytope theory. In M. Joswig and N. Takayama, editors, *Algebra, Geometry, and Software Systems*, pages 23–47. Springer, Berlin, 2003.

[KS10] I. Klep and M. Schweighofer. Pure states, positive matrix polynomials and sums of hermitian squares. *Indiana Univ. Math. J.*, 59(3):857–874, 2010.

[KS11] I. Klep and M. Schweighofer. Infeasibility certificates for linear matrix inequalities. Oberwolfach Preprints 28, 2011.

[KS13] I. Klep and M. Schweighofer. An exact duality theory for semidefinite programming based on sums of squares. *Math. Oper. Res.*, 38(3):569–590, 2013.

[KT14] K. Kellner and T. Theobald. Sum of Squares certificates for containment of $\mathcal{H}$-polytopes in $\mathcal{V}$-polytopes. Preprint, arXiv:1409.5008, 2014.

[KTT13] K. Kellner, T. Theobald, and C. Trabandt. Containment problems for polytopes and spectrahedra. *SIAM J. Optim.*, 23(2):1000–1020, 2013.

[KTT15] K. Kellner, T. Theobald, and C. Trabandt. A semidefinite hierarchy for containment of spectrahedra. *SIAM J. Optim.*, 25(2):1013–1033, 2015.

[LÖ4] J. Löfberg. Yalmip: A toolbox for modeling and optimization in MATLAB. In *Proc. CACSD Conference*, Taipei, Taiwan, 2004.

[Las01] J.-B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796, 2001.

[Las02] J.-B. Lasserre. Semidefinite programming vs. LP relaxations for polynomial programming. *Math. Oper. Res.*, 27(2):347–360, 2002.

[Las10] J.-B. Lasserre. *Moments, Positive Polynomials and their Applications*, volume 1. Imperial College Press, London, 2010.

[Lau05] A.J. Laub. *Matrix Analysis for Scientists and Engineers.* Society for Industrial and Applied Mathematics, 2005.

[Lau09] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In M. Putinar and S. Sullivant, editors, *Emerging Applications of Algebraic Geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, New York, 2009.

[Lov79] L. Lovász. On the shannon capacity of a graph. *IEEE Trans. Inf. Theory*, 25(1):1–7, 1979.

[LP14] M. Liu and G. Pataki. Exact duality in semidefinite programming based on elementary reformulations. Preprint, arXiv:1406.7274, 2014.

[Mar08] M. Marshall. *Positive Polynomials and Sums of Squares.* Mathematical surveys and monographs. American Mathematical Society, 2008.

[McM70] P. McMullen. The maximum numbers of faces of a convex polytope. *Mathematika*, 17:179–184, 1970.

[Mit] H.D. Mittelmann. Benchmarks for Optimization Software. `http://plato.la.asu.edu/bench.html`. (accessed 15-August-2014).

[Mit12] H.D. Mittelmann. The state-of-the-art in conic optimization software. In M.F. Anjos and J.-B. Lasserre, editors, *Handbook of Semidefinite, Conic and Polynomial Programming*, pages 671–686. Springer, New York, 2012.

[Nie12] J. Nie. Optimality conditions and finite convergence of Lasserre's hierarchy. *Math. Program.*, pages 1–25, 2012.

[NS07] J. Nie and M. Schweighofer. On the complexity of Putinar's Positivstellensatz. *J. Compl.*, 23(1):135–150, 2007.

[Pac04] A. Packer. Polynomial-time approximation of largest simplices in V-polytopes. *Discrete Appl. Math.*, 134(1-3):213–237, 2004.

[Pau03] V. Paulsen. *Completely Bounded Maps and Operator Algebras.* Cambridge University Press, 2003.

[Pow11] V. Powers. Rational certificates of positivity on compact semialgebraic sets. *Pacific J. Math.*, 251(2):385–391, 2011.

[PR01] V. Powers and B. Reznick. A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *J. Pure Appl. Algebra*, 164(1):221–229, 2001.

[Put93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993.

[Ram97a] M.V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Math. Program.*, 77(1):129–162, 1997.

[Ram97b] M.V. Ramana. Polyhedra, spectrahedra, and semidefinite programming. *Topics in semidefinite and interior-point methods, Fields Institute Communications*, 18:27–38, 1997.

[RTW97] M.V. Ramana, L. Tunçel, and H. Wolkowicz. Strong duality for semidefinite programming. *SIAM J. Optim.*, 7:641–662, 1997.

[Sch86] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986.

[SSŻ09] Ł. Skowronek, E. Størmer, and K. Życzkowski. Cones of positive maps and their duality relations. *J. Math. Phys.*, pages 1–22, 2009.

[Stø63] E. Størmer. Positive linear maps of operator algebras. *Acta Math.*, (19022), 1963.

[Stø86] E. Størmer. Extension of positive maps into $B(\mathcal{H})$. *J. Funct. Anal.*, 66:235–254, 1986.

[Stu00] J. Sturm. Error bounds for linear matrix inequalities. *SIAM J. Optim.*, 10(4):1228–1248, 2000.

[Ter39] F.J. Terpstra. Die Darstellung biquadratischer Formen als Summen von Quadraten mit Anwendung auf die Variationsrechnung. *Math. Ann.*, 116(1):166–180, 1939.

[Tra14] C. Trabandt. *A New Relaxation Technique for Polynomial Optimization and Spectrahedral Geometry Problems.* PhD thesis, Johann Wolfgang Goethe-Universität Frankfurt am Main, 2014.

[TTT99] K. C. Toh, M.J. Todd, and R. H. Tütüncü. SDPT3 – a MATLAB software package for semidefinite programming. *Optim. Methods and Software*, 11:545–581, 1999.

[Tun10] L. Tunçel. *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization.* Fields Institute monographs. American Mathematical Society, 2010.

[Web94] R.J. Webster. *Convexity.* Oxford Science Publications. Oxford University Press, 1994.

[Zie95] G.M. Ziegler. *Lectures on Polytopes.* Graduate Texts in Mathematics. Springer, 1995.