A Convergent Hierarchy of Certificates for Constrained Signomial Positivity

by

Allen Wang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2020

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Optimization is at the heart of many engineering problems. Many optimization problems, however, are computationally intractable. One approach to tackle such intractability is to find a tractable problem whose solution, when found, approximates that of the original problem. Specifically, convex optimization problems are often efficiently solvable, and finding a convex formulation that approximates a nonconvex problem, known as *convex relaxation*, is an effective approach.

This work concerns a particular class of optimization problem, namely constrained signomial optimization. Based on the idea that optimization of a function is equivalent to verifying its positivity, we first study a certificate of signomial positivity over a constrained set, which finds a decomposition of the signomial into sum of parts that are verifiably positive via convex constraints. However, the certificate only provides a sufficient condition for positivity. The main contribution of the work is to show that by multiplying additionally more complex functions, larger subset of signomials that are positive over a compact convex set, and eventually all, may be certified by the above method. The result is analogous to classic positivstellensatz results from algebraic geometry which certifies polynomial positivity by finding its representation with sum of square polynomials.

The result provides a convergent hierarchy of certificate for signomial positivity over a constrained set that is increasingly more complete. The hierarchy of certificate in turn gives a convex relaxation algorithm that computes the lower bounds of constrained signomial optimization problems that are increasingly tighter at the cost of additional computational complexity. At some finite level of the hierarchy, we obtain the optimal solution.

Acknowledgements

I would like to acknowledge the contributions of my supervisor, Professor Yaoliang Yu for introducing me to the world of research and supervising me through graduate studies. This work certainly would have not have been possible without his support and guidance. He is also the person that ignited my curiosity for mathematical optimization with a discovery that lead to the current work. Needless to say, his technical insights had direct impact on the work, as well.

I would also like to thank Professor Pascal Poupart for his direct contribution to this work. One of the most crucial steps in the proof of the main result was made possible due to his insight and suggestion. I'm also grateful for his encouragement at various times during graduate studies.

Many thanks to my lab mates from whom I learned a lot from during graduate studies. Being of Chinese origin and having left the country since young, my interactions with them have caused some reverse culture shock but nevertheless been educational. A special shoutout to Priyank Jaini, a collaborator on this project, and Kaiwen Wu with whom I collaborated with on a separate project and have learned a lot from.

More broadly, I would like to thank the graduate community, including the staff whom I consulted with and other professors with whom I took courses with, for their indirect contributions.

I also would like to thank Riley Murray, a graduate student who studies optimization at a different university, for helpful discussions on the topic.

And of course, thanks to my parents, for their support as always.

Dedication

A age 13 or so, I had a budding dream of becoming a mathematician. To test my potential for original mathematics, having learned to solve linear equations, I decided to tackle solving quadratic equations myself. After spending about a month, I decided to look up the answer. The answer is a series of elementary algebraic manipulations involving "complete the square". A bit dissapointed, I thought a career in mathematics was not for me. Since then, my curiosity took me to anthropology, engineering, computer science and the tech industry. Somehow, I came back to doing mathematics anyways.

The current work studies signomials, which generalizes polynomials, including those of degree 2. Instead of finding its zeros as in quadratic equations, the work studies its positivity over a constrained set. In a sense I think I answered my (one of many) childhood aspirations.

So I dedicate this work to those who aspire, whatever their aspirations may be.

Table of Contents

Chapter 1

Introduction

A *signomial* is a function of the form: $f(\mathbf{x}) = \sum_{j=1}^{\ell} c_j \exp(\mathbf{A}_j \mathbf{x})$, where $\mathbf{c} \in \mathbb{R}, \mathbf{A}_j \in \mathbb{R}^n$ for $j = 1...l$ are fixed. Optimization of such function subject to signomial inequalities and equalities is called signomial programming (SP). Although computationally difficult, SPs have wide range of applications in chemical engineering [31], aeronautics [35], circuit design [17], and communications network optimization [10].

Signomials may be thought of as a generalization of polynonmials. By a change of variable $y_i = \exp x_i$, one has the expression $p(\mathbf{y}) = \sum_{j=1}^{\ell} c_j \prod_{i=1}^{n} y_i^{\alpha_{ij}}$; in polynomials, the exponents are restricted to be integers. Algorithms for polynomial optimization, with its wide applications, have been well studied. The algorithms are based on sum of squares (SOS) certificate of polynomial positivity as described by positivstellensatz results from algebraic geometry and are computationally tied to Semidefinite Programming (SDP) [22, 26]. More recently, the concept of sums of nonnegative circuit polynomials (SONC) has been proposed as a new sufficient condition for positivity suited for sparse polynomials, which may be used to design efficient algorithms that depend on the number of terms in the polynomials and not the degrees [16].

Such methods for polynomial optimization reap the equivalence between global optimization of a function and verification of its positivity. Under this view, Chandrasekaran and Shah proposed the seminal Sums-of-AM/GM Exponential (SAGE) certificate of signomial positivity [8]. Similar to SOS certificate, the SAGE certificate is based on finding a decomposition of signomials into sum of parts, each being positive and is efficiently verifiable. The key insight of such certificate is that for a signomial has at most one negative term, verifying its positivity may be reduced to finding a solution to a sufficiently small set of convex constraints. In 2019, Murray, Chandrasekaran and Wierman generalized SAGE certificates to signomial positivity over a convex set, namely conditional SAGE, derived based on convex duality [23]. The generalization finds that when a signomial has one negative component, certifying its positivity over a convex set may again be reduced to checking a set of convex constraints. The same authors have also adapted SAGE certificate to polynomials and have shown certain equivalence between SAGE and SONC [23].

Common to all of above results is the notion of *hierarchy*. Note that SOS, SONC, SAGE are all relaxations of certificate for positivity; not all positive signomials or polynomials may be certified directly via SOS, SONC or SAGE. However, with additional computational complexity, larger subset of positive functions may be verified as so. The specific way in which computational complexity is increased in the hierarchy depends on individual results. In many results, especially the positivstellensatz results in the polynomial literature, additional computational complexity is imposed by multiplication of an extra function to the one of interest or inclusion of additional terms in the decomposition. The hierarchy is *complete* if all positive functions of a given class may be verified as so at some finite level of the hierarchy.

In this thesis, we first discuss the conditional SAGE certificates for signomial positivity over a convex set. The certificate was derived by my supervisor Yaoliang Yu in 2018 and also by Murray et al. [23], independently. We will then discuss the conditional SAGE and how it may be used to design an algorithm for constrained signomial optimization. Such results have also been independently examined by Murray et al. [23]. Then, we prove a completeness result for conditional SAGE. The result guarantees that after multiplication of a sufficiently large function, every signomial that is positive over a compact convex set may be verified via the conditional SAGE certificate.

Chapter 2

Notations and Background

2.1 Notations

We use bold fonts to denote vectors and matrices. Use $\mathbf{v}_{\setminus i} \in \mathbb{R}^{n-1}$ to denote vector $\mathbf{v} \in \mathbb{R}^n$ with the ith coordinate removed, and v_i to denote its ith coordinate. Given a matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$, use $\mathbf{A}^{(i)} \in \mathbb{R}^m$ to denote its ith row, and use $\mathbf{A}_{\setminus i} \in \mathbb{R}^m$ to denote the submatrix with ith row removed. Given a row vector $\mathbf{u} \in \mathbb{R}^n$ and column vector $\mathbf{v} \in \mathbf{R}^n$, $\mathbf{uv}$ is used to denote the dot product between them. Given a function that maps from real to real (i.e. $f : \mathbb{R} \to \mathbb{R}$), overload is definition by allowing vector value input for which the output is the element wise map of the function. For example, given $\mathbf{x} \in \mathbb{R}^n$, $\exp(\mathbf{x}) = [\exp(x_1), \ldots \exp(x_n)]^\top$.

For brevity, given $\mathbf{c} \in \mathbb{R}^\ell$ and $\mathbf{A} \in \mathbb{R}^{\ell\times n}$,

$$\mathrm{Sig}(\mathbf{c}, \mathbf{A})(\mathbf{x}) = \sum_{j=1}^{\ell} c_j \exp(\mathbf{A}_j \mathbf{x})$$

is the signomial defined by $\mathbf{c}$ and $\mathbf{A}$. Given $\mathbf{A} \in \mathbb{R}^{\ell\times n}$ and a set $\mathcal{X} \subset \mathbf{R}^n$, define the cone of coefficients $\mathbf{c}$ for which $\mathrm{Sig}(\mathbf{c}, \mathbf{A})(\mathbf{x})$ is nonnegative over $\mathcal{X}$.

$$C_{NNS}(\mathbf{c}, \mathbf{A})(\mathbf{x}) = \{\mathbf{c} \in \mathbb{R}^\ell : \mathrm{Sig}(\mathbf{c}, \mathbf{A})(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathcal{X}\}$$

And use the following notation for relative entropy for also known as KL-divergence. Given vectors $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$, define:

$$D(\mathbf{v}, \mathbf{u}) = \sum_{i=1}^{n} v_i \log(\frac{v_i}{u_i})$$

Where the log is base 2. Finally, given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, define the following. For some $p \in \mathbb{Z}_{++}$,

$$E_p(\mathbf{A}) = [\mathbf{A}_1 \dots \mathbf{A}_{\ell^{(p+1)}}]^\top \text{ where } (\mathbf{A}^{(i)})_{i=1\dots\ell^{(p+1)}} =$$
$$\{\mathbf{A} \in \mathbb{R}^n : \mathbf{A} = \sum_{i=1}^{\ell} w_i \mathbf{A}^{(i)} \text{ such that } w_i \in \mathbb{Z}\ \forall i, \text{ and } \sum_{i}^{\ell} w_i \leq p\}$$

The notation is imported from [8]. $E_p(\mathbf{A})$ is the matrix whose rows are the integer combinations of rows in $\mathbf{A}$ up to summation p.

2.2 Background

2.2.1 Certificate of Positivity and Optimization

Certificate for positivity is motivated by its application in optimization. Consider the following constrained optimization problem:

$$f^*_{\mathcal{X}} = \inf_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x})$$

which may be reformulated as a problem of certifying function positivity.

$$f^*_{\mathcal{X}} = \sup \lambda \text{ s.t. } f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathcal{X}$$

In other words, constrained optimization problem may be reduced to checking the positivity of a function over a set. Consider a class of function F, and subset $F_{\mathcal{X}}$ such that $\inf_{\mathcal{X}} f(\mathbf{x}) \geq 0 \iff f(\mathbf{x}) \in F_{\mathcal{X}}$. The optimization problem is solved if membership in $F_{\mathcal{X}}$ can be checked.

$$f^*_{\mathcal{X}} = \sup \lambda \text{ s.t. } f(\mathbf{x}) - \lambda \in F_{\mathcal{X}}$$

Naturally, certifying function positivity is no easier than the optimization problem itself. Of interest is to find a sufficient condition for function positivity; we want to develop a tractable set $G_{\mathcal{X}} \subset F_{\mathcal{X}}$ such that checking $\mathrm{Sig}(\mathbf{c}, \mathbf{A}) \in G$ is sufficiently easy. Then:

$$f_{\mathcal{X}}^{G} = \sup \lambda \ \text{ s.t. } \ f(\mathbf{x}) - \lambda \in G_{\mathcal{X}}$$

is a tractable problem and $f_{\mathcal{X}}^{G} \leq f_{\mathcal{X}}^{*}$.

2.2.2 Polynomial Optimization

The equivalence of optimization and certificate of positivity is well established in polynomial optimization. Consider a general problem formulation of the following form:

$$\begin{aligned} \min \quad & p(\mathbf{x}) \\ s.t. \quad & g_i(\mathbf{x}) \geq 0 \quad \forall i = 1, \dots m \end{aligned}$$

Use $K = \{\mathbf{x} : g_i(\mathbf{x}) \geq 0 \ \forall i = 1, \dots m\}$ to denote the semialgebraic set defined by polynomials. The above is equivalent to solving the following problem.

$$\begin{aligned} \max \quad & \gamma \\ s.t. \quad & p(\mathbf{x}) - \gamma \geq 0 \ \forall \mathbf{x} \in K \end{aligned}$$

which calls for characterization of nonnegative polynomials. One observation is that if a polynomial can be written as a sum of squares, then it is positive. Naturally, one wonders whether all positive polynomials may be written as sum of squares, and Hilbert posed such question in the 17th century. However, it is not difficult to show the existence of counter examples. For example

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

may be shown to be non-negative but cannot be expressed as sum of squares. Hilbert later showed that non-negative homogeneous polynomial in n variables and degree 2d can be represented as sum of squares of other polynomials if and only if $n = 2$, or $2d = 2$, or $n = 3$ and $2d = 4$. [15].

Hilbert's question began the study of positive polynomials and finding their algebraic certificates, known as positivtellensatz. While numerous theorems have been developed, below we highlight a few results.

The following theorem from the 19th century shows that every positive polynomial may be expressed as the sum of squares of rational functions.

Theorem 1 *(Artin's Positivstellensatz [4]) For a globally positive polynomial $p(\mathbf{x})$, there is a nonzero SOS polynomial $q(\mathbf{x})$ such that $q(\mathbf{x})p(\mathbf{x})$ is a sum of squares*

Note that verification of a polynomial as sum of squares, as in above theorem, may be reformulated as an SDP problem. The specifics are omitted from this thesis, readers are directed to the seminal works by Lassere and Perrilo [26, 19]. The hierarchy is induced by limiting the degree of $q(\mathbf{x})$ to search for.

The following characterizes a polynomial positive over a compact semialgebraic set.

Theorem 2 *(Schmudgen's Positivstellensatz [32]) Assume the set $\mathcal{K} = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x})_i \geq 0 \;\; i = 1 \dots \ell\}$, defined by polynomials $(g(\mathbf{x})_i)_{i=1\dots\ell}$ is compact. If polynomial $p(\mathbf{x})$ is positive on $\mathcal{K}$, then*

$$
\begin{aligned}
p(\mathbf{x}) = s_0(\mathbf{x}) + \sum_{i \in [\ell]} s_i(\mathbf{x}) g_i(\mathbf{x}) + & \sum_{(i_1, i_2) \in [l]^2} s_{i_1 i_2}(\mathbf{x}) g_i(\mathbf{x}) g_j(\mathbf{x}) + \dots \\
& \sum_{i_1 \dots i_l \in [l]^\ell} s_{i_1 \dots i_2}(\mathbf{x}) g_{i_1}(\mathbf{x}) g_{i_2}(\mathbf{x}) \dots g_{i_l}(\mathbf{x})
\end{aligned}
$$

where $s(\mathbf{x})_0, s(\mathbf{x})_i, s(\mathbf{x})_{i_1, i_2} \dots s(\mathbf{x})_{i_1 \dots i_l}$ are sum of square polynomials. The above is essentially a search problem over sum of square polynomials, and may again be converted to SDP by finding a representation of SOS polynomials as semidefinite matrices [26]. The problem may also be relaxed by restricting the number of compositions of $g_i(\mathbf{x})$'s (i.e. truncating later summation terms on the RHS). The relaxation yields a converging hierarchy for the problem of verifying polynomial positivity over a semialgebraic set.

The following provides a so called *optimization-free* positivstellensantz. That is, verification of positivity is easy at sufficient level on the hierarchy.

Theorem 3 *(Polya's Positivstellensatz [28]) For a globally positive polynomial $p(\mathbf{x})$ of even form, there exists $p \in \mathbb{Z}_+$ such that $(\sum_{i=1}^{n} \mathbf{x}_i^2)^p p(\mathbf{x})$ has non negative coefficients.*

Recently, Dickinson and Povh developed a new positivstellensatz for polynomial positive over a semialgebraic set, which is closely related to both Schmudgen and Polya's Positivstellensatz.

Theorem 4 *(Dickson's Positivstellensatz [12]) Let $f_0, \ldots\ldots f_m \in \mathbb{R}[\mathbf{x}]$ be homogeneous polynomials on $\mathbb{R}^n$ such that $f_0(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}_+^n \cap \bigcap_{i=1}^{m} f_i^{-1}(\mathbb{R}_+) \backslash \{\mathbf{0}\}$. Then for some $p \in \mathbb{Z}_+$, there exists homogeneous polynomials with nonnegative coefficients $g_0, \ldots\ldots g_m \in \mathbb{R}[\mathbf{x}]$ such that $(\mathbf{x}^\top \mathbf{1})^p f_0(\mathbf{x}) = g_0(\mathbf{x}) + \sum_{i=1}^{m} f_i(\mathbf{x}) g_i(\mathbf{x})$.*

Positivstellensatz is to this day an active subject of study. Also recently, Ahmadi and Hall have proposed an optimization free Positivstellensatz for polynomial positivity over a semialgebraic set, by which positivity can be verified by checking that the multiplication of two polynomials that depend by objective polynomial $p(\mathbf{x})$ and constraint polynomials $(g_i(\mathbf{x}))_{i=1\ldots\ell}$ [3].

Chapter 3

SAGE Relaxation of Signomial Positivity

In this chapter we discuss a certificate of positivity for signomials based on decomposition of the function into sum of parts that are each verifiable positive, and a resulting optimization algorithms due to the certificate. The approach to verification is similar to SOS verification in finding a decomposition, but as opposed to relying on the fact that the square of a function is positive, we first derive an efficiently computable characterization of signomials with at most one negative term and is positive over a convex set. That is, given that a signomial has at most one negative term, it is positive over a convex set if and only if a set of convex constraints depending on the parameters defining the signomial and convex set are satisfied. Given that a signomial with one negative term is not generally a convex function, checking whether it is positive globally or over a convex set is may seem to be intractable. The key insight is to clear the exponential of the negative term so the problem is reduced to checking the minimum value of a signomial with only positive terms, which is indeed a convex function.

A sufficient condition for signomial positivity follows from such characterization. We then discuss a hierarchy of certificates that can verify the positivity of increasingly larger subset of signomials, as well as the resulting algorithm for computing a lower bounds for constrained signomial optimization problem.

3.1 SAGE

SAGE certificate was originally developed as a certificate for *global* positivity by Chandrasekaran et al. in 2015 [8]. The building block of such certificate is a signomial with at most one negative term, which Chandrasekaran et al. called AM-GM exponentials (AGE). The same authors made the observation that the global positivity of such function is reduced to finding the solution to relative entropy program. Then, if a signomial is a summation of those with at most one negative term that are positive, the signomial is positive. Such signomials are referred to as sum of AM/GM-exponentials (SAGE), which forms a subset of globally positive signomials.

Since then, we and Murray et al. have independently generalized SAGE to positivity over arbitrary set $\mathcal{X} \subset \mathbb{R}^n$, which encompasses the case when $\mathcal{X} = \mathbb{R}^n$ [23]. The work is based on the idea that the positivity of a signomial over a convex set may also be reduced to solving a convex program depending on the parameters defining the signomial, if it has at most one negative term. The derivation is based on algebraic operations and an application of convex duality. Summation of such signomial can also be verified to be positive on the same convex set easily, by seeking a decomposition of their coefficients such that the parameters of the decomposed signomial satisfy the convex constraints. Murray et al. presented this work during the study of the thesis and have named the approach Conditional SAGE. Below we discuss key definitions and derivations.

Definition 1 *(Conditional AGE Signomials). $AGE(\mathbf{A}, \mathcal{X}, i)$ is the cone of signomials with at most one negative term at the ith index and is positive over $\mathcal{X}$. Given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$ and $i \in [\ell]$, the ith AGE cone with respect to $\mathbf{A}$ and $\mathcal{X}$ is:*

$$AGE(\mathbf{A}, \mathcal{X}, i) = \{Sig(\mathbf{c}, \mathbf{A}) : \mathbf{c}_{\setminus i} \geq 0 \textit{ and } Sig(\mathbf{c}, \mathbf{A})(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathcal{X}\}$$

By definition, $AGE(\mathbf{A}, \mathcal{X}, i)$ is a cone; it is easy to verify that it is closed under addition and nonnegative scaling. We may also define a similar set for the coefficients.

Definition 2 *(Conditional AGE Cone). $C_{AGE}(\mathbf{A}, \mathcal{X}, i)$ is the cone of coefficients with at most one negative term at the ith index, such that resulting signomial is positive over $\mathcal{X}$. Given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$ and $i \in [\ell]$, the ith AGE coefficients with respect to $\mathcal{X}$ and $\mathbf{A}$ is:*

$$C_{AGE}(\mathbf{A}, \mathcal{X}, i) = \{\mathbf{c} \in \mathbb{R}^{\ell} : \mathbf{c}_{\setminus i} \geq 0 \textit{ and } Sig(\mathbf{c}, \mathbf{A})(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathcal{X}\}$$

Equipped with above definition, we may define the SAGE cone. whereas SOS is summation of squares, SAGE is a summation AGE signomials.

Definition 3 *(Conditional SAGE Signomials). $SAGE(\mathbf{A}, \mathcal{X})$ is Minkowski sum of $AGE(\mathbf{A}, \mathcal{X}, i)$ for $i = 1 \ldots \ell$.*

$$SAGE(\mathbf{A}, \mathcal{X}) = \sum_{i=1}^{\ell} AGE(\mathbf{A}, \mathcal{X}, i)$$

While SOS finds decomposition by terms in the polynomials, SAGE finds the decomposition by coefficients. Such nature of SAGE allows the following the definition:

Definition 4 *(Conditional SAGE Cone). $C_{SAGE}(\mathbf{A}, \mathcal{X})$ is Minkowski sum of $C_{AGE}(\mathbf{A}, \mathcal{X}, i)$ for $i = 1 \ldots \ell$.*

$$C_{SAGE}(\mathbf{A}, \mathcal{X}) = \sum_{i=1}^{\ell} C_{AGE}(\mathbf{A}, \mathcal{X}, i)$$

It follows that $C_{SAGE}(\mathbf{A}, \mathcal{X})$ is also a cone. By definition, $C_{SAGE}(\mathbf{A}, \mathcal{X}) \subset C_{NNS}(\mathbf{A}, \mathcal{X})$. The following theorem shows that $C_{AGE}(\mathbf{A}, \mathcal{X}, i)$ is a tractable set via convex constraints.

Theorem 5 *[23]. Given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$ and $i \in [\ell]$. Let $\sigma_{\mathcal{X}}(\boldsymbol{\lambda}) = \sup_{\mathbf{x} \in \mathcal{X}} \boldsymbol{\lambda}^\top \mathbf{x}$. Then:*

$$\begin{aligned} C_{AGE}(\mathbf{A}, \mathcal{X}, i) = \{\mathbf{c} \in \mathbb{R}^\ell : \exists \mathbf{v} \in \mathbb{R}^{\ell-1} \textit{ and } \boldsymbol{\lambda} \in \mathbb{R}^n \textit{ s.t.} \\ \sigma_{\mathcal{X}}(\boldsymbol{\lambda}) + D(\mathbf{v}, \mathbf{c}_{\setminus i}) - \mathbf{1}^\top \mathbf{v} \leq c_i \\ [\mathbf{A}_{\setminus i} - \mathbf{1}\mathbf{A}_i]^\top \mathbf{v} + \boldsymbol{\lambda} = \mathbf{0} \textit{ and } \mathbf{c}_{\setminus i} \geq \mathbf{0}\} \end{aligned}$$

Proof: Let $\delta_{\mathcal{X}}$ denote the indicator function of $\mathcal{X}$. A vector $\mathbf{c} \in \mathbb{R}^\ell$ is in the $C_{AGE}(\mathbf{A}, \mathcal{X}, i)$ cone if and only if:

$$\begin{aligned} & \sum_{j=1}^{\ell} c_j \exp(\mathbf{A}_j \mathbf{x}) \geq 0 \ \mathbf{x} \in \mathcal{X} \text{ and } \mathbf{c}_{\setminus i} \geq 0 \\ \iff & \sum_{1 \leq j \leq l, j \neq i} c_j \exp((\mathbf{A}_j - \mathbf{A}_i)\mathbf{x}) \geq -c_i \ \forall \mathbf{x} \in \mathcal{X} \text{ and } \mathbf{c}_{\setminus i} \geq 0 \\ \iff & p^* = \inf_{\mathbf{x} \in \mathbb{R}^n} \ \delta_{\mathcal{X}}(\mathbf{x}) + \sum_{1 \leq j \leq l, j \neq i} c_j \exp((\mathbf{A}_j - \mathbf{A}_i)\mathbf{x}) \geq -c_i \text{ and } \mathbf{c}_{\setminus i} \geq 0 \end{aligned}$$

Where we removed the convex constraint with indicator function in the last step. Then we may apply Fenchel duality to the minimization problem. The resulting dual is

$$d^* = \sup_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^n \\ \mathbf{v} \in \mathbb{R}^{m-1} \\ (\mathbf{A}_{\setminus i} - \mathbf{1}\mathbf{A}_i)^\top \mathbf{v} + \boldsymbol{\lambda} = \mathbf{0}}} -\sigma_{\mathcal{X}}(\boldsymbol{\lambda}) - D(\mathbf{v}, \mathbf{c}_{\setminus i}) + \mathbf{v}^\top \mathbf{1}$$

When $\mathcal{X}$ is nonempty, strong duality holds by corollary 3.3.11 of [30]. Consequently:

$$p^* \leq -c_i \text{ and } c_i \geq 0 \iff -d^* \leq c_i \text{ and } \mathbf{c}_{\setminus i} \geq 0$$

and we have the desired result. When $\mathcal{X}$ is empty, $p^* = +\infty$ by definition of indicator function. By choosing $\mathbf{v}, \boldsymbol{\lambda} = \mathbf{0}$, we have $d^* = \infty$ and desired result also follows. ■

As corollary, C_{SAGE} is characterized by the following set of convex constraints:

Corollary 1 *Given* $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$ *and* $i \in [\ell]$. *Let* $\sigma_{\mathcal{X}}(\boldsymbol{\lambda}) = \sup_{\mathbf{x} \in \mathcal{X}} \boldsymbol{\lambda}^\top \mathbf{x}$. *Then:*

$$\begin{aligned} C_{SAGE}(\mathbf{A}, \mathcal{X}) = \{\mathbf{c} \in \mathbb{R}^\ell : \exists c^{(i)} \in \mathbb{R}^\ell, \mathbf{v}^{(i)} \in \mathbb{R}^{\ell-1} \text{ and } \boldsymbol{\lambda}^{(i)} \in \mathbb{R}^n \text{ s.t.} \\ \sum_{i=1}^{\ell} c^{(i)} = \mathbf{c} \text{ and} \\ \sigma_{\mathcal{X}}(\boldsymbol{\lambda}^{(i)}) + D(\mathbf{v}^{(i)}, c_{\setminus i}^{(i)}) - \mathbf{v}^{(i)\top} \mathbf{v}^{(i)} \leq c_i^{(i)} \\ [\mathbf{A}_{\setminus i} - \mathbf{1}\mathbf{A}_i]^\top \mathbf{v}^{(i)} + \boldsymbol{\lambda}^{(i)} = \mathbf{0} \text{ and } c_{\setminus i}^{(i)} \geq \mathbf{0} \text{ for } i = 1 \ldots \ell\} \end{aligned}$$

There are $O(\ell n)$ constraints defined by $O(\ell(\ell + n))$ variables. Although the complexity of the problem does increase depending on both the dimension of input to the signomial and the number of terms, it is highly tractable.

3.2 SAGE Hierarchy

$C_{SAGE}(\mathbf{A}, \mathcal{X})$ is an inner approximation of the $C_{NNS}(\mathbf{A}, \mathcal{X})$ cone. We may however find larger and more accurate inner approximation through what is known as *modulation* in the literature. Define the hierarchy of SAGE cones as:

Definition 5 *Given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$, the pth level SAGE cone with respect to $\mathbf{A}$ and $\mathcal{X}$ is:*

$$C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X}) = \{\mathbf{c} \in \mathbb{R}^\ell : (\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^p Sig(\mathbf{c}, \mathbf{A}) \in SAGE(E_{p+1}(\mathbf{A}), \mathcal{X})\}$$

That is, instead of certifying a given signomial to be positive, we verify whether its product with a positive definite term is positive. Multiplication of a positive definite term does not change positivity, and thus the test is valid. Equivalently, expanding the above definition and paying attention to the coefficients of the larger signomial after multiplication of the extra function gives:

$$\begin{aligned} C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X}) = \{\mathbf{c} \in \mathbb{R}^\ell : \exists \mathbf{c}^{(i)} \in \mathbb{R}^{\ell^{(p+1)}}, \mathbf{v}^{(i)} \in \mathbb{R}^{\ell^{(p+1)}-1}, \tilde{\mathbf{A}} \in \mathbb{R}^{\ell^{(p+1)} \times n} \text{ and } \boldsymbol{\lambda}^{(i)} \in \mathbb{R}^n \text{ s.t.} \\ \tilde{\mathbf{A}}_{j_1 + \ell \cdot j_2 + \ldots + \ell^p \cdot j_{p+1}} = \mathbf{A}_{j_1} + \ldots + \mathbf{A}_{j_\ell} \text{ for } j_1 \ldots j_{p+1} = 1 \ldots \ell \\ c^{(i)}_{k+1\cdot\ell} = c^{(i)}_{k+2\cdot\ell} = \ldots = c^{(i)}_{k+\ell^p} \; \forall k = 1 \ldots \ell, \; \sum_{i=1}^{\ell} c^{(i)}_{1:\ell} = \mathbf{c} \\ \sigma_{\mathcal{X}}(\boldsymbol{\lambda}^{(i)}) + D(\mathbf{v}^{(i)}, \mathbf{c}^{(i)}_{\setminus i}) - \mathbf{v}^{(i)\top}\mathbf{v}^{(i)} \le c^{(i)}_i \\ [\tilde{\mathbf{A}}_{\setminus i} - \mathbf{1}\tilde{\mathbf{A}}_i]^\top \mathbf{v}^{(i)} + \boldsymbol{\lambda}^{(i)} = \mathbf{0} \text{ and } \mathbf{c}^{(i)}_{\setminus i} \ge \mathbf{0} \text{ for } i = 1 \ldots \ell^{(p+1)}\} \end{aligned}$$

A more careful construction can reduce some of the constraints. In particular, the equality constraints on $\mathbf{c}^{(i)}$ are simply to construct a large vector with repeated sub-vectors, and the constraints on $\tilde{\mathbf{A}}$ simply denote that its rows are linear combinations of rows of $\mathbf{A}$. In implementation, we may avoid such constraints by recycling variables. Ignoring such constraints, at the pth level of the hierarchy, there are $O(\ell^{(p+1)}n)$ constraints defined by $O(\ell^{(p+1)}(\ell^{(p+1)} + n))$ variables. Upon setting $p = 0$, we recover $C_{SAGE}(\mathbf{A}, \mathcal{X})$. Moreover, we have the following relations.

Theorem 6 *Given $\mathbf{A} \in \mathbb{R}^{\ell \times n}$, $\mathcal{X} \subset \mathbb{R}^n$, and some $p \in \mathbb{Z}_{++}$,*

$$C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X}) \subseteq C^{(p+1)}_{SAGE}(\mathbf{A}, \mathcal{X}) \subseteq C_{NNS}(\mathbf{A}, \mathcal{X})$$

Proof: Let $\mathbf{c} \in C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X})$. There exists $\mathbf{c}^{(1)} \ldots \mathbf{c}^{(\ell)}$
such that $\mathbf{c} = \sum_{i=1}^{\ell} c^{(i)}$ and $(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^p \mathrm{Sig}(c^{(i)}, \mathbf{A}) \in AGE(E_{p+1}(\mathbf{A}), \mathcal{X}, i)$.

$$(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^{(p+1)} \mathrm{Sig}(c^{(i)}, \mathbf{A}) = \sum_{j=1}^{\ell} \exp(\mathbf{A}_j \mathbf{x})(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^p \mathrm{Sig}(c^{(i)}, \mathbf{A})$$

Observe that $\exp(\mathbf{A}_j\mathbf{x})(\exp(\mathbf{A}\mathbf{x})^\top\mathbf{1})^p\mathrm{Sig}(c^{(i)},\mathbf{A}) \in AGE(E_{p+2}(\mathbf{A}),\mathcal{X},i)$ given $\exp(\mathbf{A}\mathbf{x})^\top\mathbf{1})^p\mathrm{Sig}(c^{(i)},\mathbf{A}) \in AGE(E_{p+1}(\mathbf{A}),\mathcal{X},i)$ since multiplication of a one exponential term does not change positivity, nor the positive coefficients. Since $AGE(E_{p+2}(\mathbf{A}),\mathcal{X},i)$ and $SAGE(E_{p+2}(\mathbf{A}),\mathcal{X})$ are both closed under addition, we have:

$$\begin{aligned}(\exp(\mathbf{A}\mathbf{x})^\top\mathbf{1})^{(p+1)}\mathrm{Sig}(c^{(i)},\mathbf{A}) &\in AGE(E_{p+2}(\mathbf{A}),\mathcal{X},i) \subseteq SAGE(E_{p+2}(\mathbf{A}),\mathcal{X}) \\ &\implies c^{(i)} \in C^{(p+2)}_{SAGE}(\mathbf{A},\mathcal{X}) \implies \mathbf{c} \in C^{(p+2)}_{SAGE}(\mathbf{A},\mathcal{X})\end{aligned}$$

Now let $\mathbf{c} \in C^{(p+1)}_{SAGE}(\mathbf{A},\mathcal{X})$.

$$\begin{aligned}&(\exp(\mathbf{A}\mathbf{x})^\top\mathbf{1})^{(p+1)}\mathrm{Sig}(\mathbf{c},\mathbf{A}) \in SAGE(E_{p+1}(\mathbf{A}),\mathcal{X}) \\ \implies &(\exp(\mathbf{A}\mathbf{x})^\top\mathbf{1})^{(p+1)}\mathrm{Sig}(\mathbf{c},\mathbf{A})(\mathbf{x}) \geq 0\ \forall\mathbf{x}\in\mathcal{X} \implies \mathrm{Sig}(\mathbf{c},\mathbf{A})(\mathbf{x}) \geq 0\ \forall\mathbf{x}\in\mathcal{X} \\ &\implies \mathbf{c} \in C_{NNS}(\mathbf{A},\mathcal{X})\end{aligned}$$

The second to last implication is due to the fact that multiplication of a positive term does not change positivity. ■

The above shows that the hierarchy of SAGE cone provides an increasingly accurate inner approximation of non-negative signomials.

3.3 SAGE Relaxation

Given hierarchy of SAGE cones, we may formulate a hierarchy of relaxations for signomial optimization. Consider signomial $\mathrm{Sig}(\mathbf{c},\mathbf{A})$ and a convex set $\mathcal{X}$, let $f^*_{\mathcal{X}} = \inf_{\mathbf{x}\in\mathcal{X}} \mathrm{Sig}(\mathbf{c},\mathbf{A})(\mathbf{x})$. The relaxation may be formulated as follows.

$$\begin{aligned}f^{(p)}_{SAGE} &= \sup_{\lambda} \lambda \text{ s.t. } \mathrm{Sig}(\mathbf{c},\mathbf{A}) - \lambda \in SAGE^{(p)}(\mathbf{A},\mathcal{X}) \\ &= \sup_{\lambda} \lambda \text{ s.t. } \mathbf{c} - \lambda[1,0,\ldots] \in C^{(p)}_{SAGE}(\mathbf{A},\mathcal{X})\end{aligned}$$

Theorem 6 tells us that the conditional SAGE cones are inner approximations of signomials that positive over a convex set. As result, we have that $f^{(p)}_{SAGE} \leq f^{(p+1)}_{SAGE} \leq f^*_{\mathcal{X}}$.

3.4 Dual SAGE relaxation and Solution Recovery

The primal formulation presented in previous section computes lower bounds for a constrained signomial optimization problem. Fixing the exponents $\mathbf{A}$ and a constrained set $\mathcal{X}$, the set of coefficients for which $\mathrm{Sig}(\mathbf{A},\mathcal{X})$ is a conditional SAGE signomial, $C_{SAGE}(\mathbf{A},\mathcal{X})$ is cone. In this section, we derive a dual formulation which optimizes over the dual cone of $C_{SAGE}(\mathbf{A},\mathcal{X})$. A we will see, the dual SAGE cone may be seen as an outer approximation of the exponentials of $\mathrm{Sig}(\mathbf{A},\mathcal{X})$ (i.e. $\exp(\mathbf{A}\mathbf{x})$), and as such, the formulation can be used to recover a solution to the optimization problem (i.e. some $\mathbf{x}\in\mathcal{X}$), as opposed to just a lower bound to the optimization problem.

We begin by deriving the dual cone of $C_{AGE}(\mathbf{A},\mathcal{X},i)$. Recalling that $C_{SAGE}(\mathbf{A},\mathcal{X})$ is a Minkowski sum of $(C_{AGE}(\mathbf{A},\mathcal{X},i))_{i=1\ldots\ell}$, its dual cone is easily found through a basic theorem from convex analysis.

Theorem 7 *The dual cone of $C_{AGE}(\mathbf{A},\mathcal{X})$ is characterized as follows:*

$$\begin{aligned}C_{AGE}(\mathbf{A},\mathcal{X})^* = cl\{&\mathbf{v}\in\mathbb{R}^l_+ : \exists\alpha_1\ldots\alpha_{n+1}\in\mathbb{R}_+,\ \mathbf{x}_1,\ldots,\mathbf{x}_{n+1}\in\mathcal{X},\ \textit{such that}\\ &v_i=\sum_{k=1}^{l}\alpha_k,\ \forall j\neq i\ v_j\geq\sum_{k=1}^{l}\alpha_k,\exp((\mathbf{A}_j-\mathbf{A}_i)^\top\mathbf{x}_k)\}\end{aligned}$$

Proof: We begin by considering the following expression of $C_{AGE}(\mathbf{A},\mathcal{X},i)$.

$$\begin{aligned}C_{AGE}(\mathbf{A},\mathcal{X},i) &= \{\mathbf{c}\in\mathbb{R}^\ell : \mathbf{c}_{\setminus i}\geq\mathbf{0},\ \text{and}\ \forall\mathbf{x}\in\mathcal{X},\ \mathbf{c}^\top\exp(\mathbf{A}\mathbf{x})\geq 0\}\\ &= \{\mathbf{c}\in\mathbb{R}^\ell : \mathbf{c}_{\setminus i}\geq\mathbf{0},\ \text{and}\ \forall\mathbf{x}\in\mathcal{X},\ \mathbf{c}^\top\exp((\mathbf{A}_{\setminus i}-\mathbf{1}\mathbf{A}_i)^\top\mathbf{x})\geq 0\}\end{aligned}$$

Where in the last step we divided the expression by $\exp(\mathbf{A}_i^\top\mathbf{x})$. Consider the set

$$\begin{aligned}\tilde{C}_{AGE}(\mathbf{A},\mathcal{X},i) &= \{\mathbf{c}\in\mathbb{R}^\ell : \forall\mathbf{x}\in\mathcal{X},\ \mathbf{c}^\top\exp((\mathbf{A}_{\setminus i}-\mathbf{A}_i)^\top\mathbf{x})\geq 0\}\\ &\subseteq C_{AGE}(\mathbf{A},\mathcal{X},i)\end{aligned}$$

That is, we removed the AM-GM constraint that guarantees all but ith coefficient to be non-negative. It follows from the definition of $\tilde{C}_{AGE}(\mathbf{A},\mathcal{X},i)$ that its dual is the convex cone generated by the following set: $\{\exp((\mathbf{A}_{\setminus i}-\mathbf{1}\mathbf{A}_i)^\top\mathbf{x})\ :\mathbf{x}\in\mathcal{X}\}$. By Caratheodory's Theorem, any point in a convex set of dimension n may be expressed by at most $n+1$ points [30], and thus we have the following characterization:

$$\begin{aligned}\tilde{C}_{AGE}(\mathbf{A},\mathcal{X},i)^* = \{&\mathbf{v}\in\mathbf{R}^n : \exists\alpha_1\ldots\alpha_{n+1}\in\mathbb{R}_+,\ \mathbf{x}_1\ldots\mathbf{x}_{n+1}\in\mathcal{X}\\ &\mathbf{v}=\sum_{i=1}^{n+1}\alpha_i\exp((\mathbf{A}_{\setminus i}-\mathbf{1}\mathbf{A}_i)^\top\mathbf{x}\}\end{aligned}$$

Now recall that $C_{AGE}(\mathbf{A},\mathcal{X},i) = \tilde{C}_{AGE}(\mathbf{A},\mathcal{X},i) \cap \{\mathbf{c} : \mathbf{c}_{\setminus i} \geq 0\}$. By elementary rule of convex analysis, $(A \cap B)^* = A^* + B^*$, and the dual cone of $\{\mathbf{c} : \mathbf{c}_{\setminus i} \geq 0\}$ is $(\mathbb{R}_+^{i-1} \times \{\mathbf{0}\} \times \mathbb{R}_+^{l-i-1})$. So we have:

$$C_{AGE}(\mathbf{A},\mathcal{X},i)^* = \mathrm{cl}\{\mathbf{v} \in \mathbb{R}_+^l : \exists \alpha_1 \ldots \alpha_{n+1} \in \mathbb{R}_+,\ \mathbf{x}_1,\ldots,\mathbf{x}_{n+1} \in \mathcal{X},\ \text{such that}$$
$$v_i = \sum_{k=1}^{l} \alpha_k,\ \forall j \neq i\ v_j \geq \sum_{k=1}^{l} \alpha_k, \exp((\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x}_k)\}.$$

■

With the inequalities, the above may be seen as an outer approximation of the set:

$$\{\exp((\mathbf{A}_{\setminus i} - \mathbf{1}\mathbf{A}_i)^\top \mathbf{x})\ : \mathbf{x} \in \mathcal{X}\}$$

which is not tractable via convex constraints, as it essentially characterizes $\mathrm{Sig}(\mathbf{A}_{\setminus i} - \mathbf{1}\mathbf{A}_i, \mathcal{X})$ over $\mathcal{X}$. However, we show below that by the relaxation of such set through outer approximation, $C_{AGE}(\mathbf{A},\mathcal{X},i)^*$ is indeed tractable via convex constraints.

Theorem 8 *The dual cone $C_{AGE}(\mathbf{A},\mathcal{X},i)^*$ is equivalent to*

$$C_{AGE}(\mathbf{A},\mathcal{X},i)^* = cl(\{\mathbf{v} \in \mathbb{R}_+^l : \exists \mathbf{z} \in v_i\mathcal{X} \subseteq \mathbb{R}^n \text{ such that } \forall j, v_i \log \frac{v_i}{v_j} \leq -(\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{z}\})$$

Proof: We prove the equivalence of the terms inside the closure in Theorem 7 and Theorem 8.

Let $\mathbf{v}$ be such that

$$v_i = \sum_{k=1}^{l} \alpha_k, \quad \forall j \neq i,\ v_j \geq \sum_{k=1}^{l} \alpha_k \exp((\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x}_k)$$

for some $\alpha_1 \ldots \alpha_{n+1} \in \mathbb{R}_+$ and $\mathbf{x}_1,\ldots,\mathbf{x}_{n+1} \in \mathcal{X}$. Then,

$$\log \frac{v_j}{v_i} \geq \log \frac{\sum_{k=1}^{l} \alpha_k \exp((\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x}_k)}{\sum_{k=1}^{l} \alpha_i} = \log \sum_{k=1}^{l} \frac{\alpha_k}{\sum_{k=1}^{l} \alpha_k}((\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x}_k)$$
$$\geq \log \exp((\mathbf{A}_j - \mathbf{A}_i)^\top \sum_{k=1}^{l} \frac{\alpha_i}{\sum_{k=1}^{l} \alpha_k} \mathbf{x}_i) = (\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x},$$

for some $\mathbf{x} = \sum_{k=1}^{l} \frac{\alpha_k}{\sum_{k=1}^{l} \alpha_k} \mathbf{x}_k \in \mathcal{X}$. Thus, defining $\mathbf{z} = v_i \mathbf{x}$, we see that $\mathbf{v}$ belongs to the right-hand side of the expression in Theorem 8.

Conversely, let $\mathbf{v}$ be such that

$$\forall j,\ v_i \log \frac{v_i}{v_j} \leq -(\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{z}$$

for some $\alpha_1 \ldots \alpha_{n+1} \in \mathbb{R}_+$ and $\mathbf{z} \in v_i \mathcal{X}$. Then,

$$\log \frac{v_j}{v_i} \geq (\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x} \iff v_j \geq v_i \exp((\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{x}),$$

for some $\mathbf{x} \in \mathcal{X}$. Thus, setting $\alpha_i = \frac{1}{m+1} v_i$ and $\mathbf{x}_i = \mathbf{x}$, we know $\mathbf{v}$ belongs to the right-hand side of the definition in Theorem 7. ■

Recall that $C_{SAGE}(\mathbf{A}, \mathcal{X}) = \sum_i^\ell C_{AGE}(\mathbf{A}, \mathcal{X}, i)$. Thus the dual cone $C_{SAGE}(\mathbf{A}, \mathcal{X})^* = \cap_{i=1}^{\ell} C_{AGE}(\mathbf{A}, \mathcal{X}, i)^*$.

Corollary 2

$$\begin{aligned} C_{SAGE}(\mathbf{A}, \mathcal{X})^* = cl(\{\mathbf{v} \in \mathbb{R}_+^l : \exists \mathbf{z}^{(i)} \in v_i \mathcal{X} \subseteq \mathbb{R}^n \textit{ for } i = 1 \ldots \ell \textit{ s.t.} \\ \forall j, v_i \log \frac{v_i}{v_j} \leq -(\mathbf{A}_j - \mathbf{A}_i)^\top \mathbf{z}^{(i)}\}) \end{aligned}$$

And more generally, following a similar analysis of finding $C_{SAGE}^{(*)}$, we obtain:

Corollary 3

$$\begin{aligned} C_{SAGE}^{(p)}(\mathbf{A}, \mathcal{X})^* = cl(\{\mathbf{v} \in \mathbb{R}_+^\ell : \exists \boldsymbol{\lambda} \in \mathbb{R}_+^{\ell(p+1)},\ &\boldsymbol{\lambda}^{(1)} \ldots \boldsymbol{\lambda}^{(\ell^p)} \in \mathbb{R}_+^\ell,\ \tilde{\mathbf{A}} \in \mathbb{R}^{\ell(p+1) \times n}, \\ &\mathbf{z}^{(i)} \in \lambda_i \mathcal{X} \subseteq \mathbb{R}^n \textit{ for } i = 1 \ldots \ell^{(p+1)} \textit{ s.t.} \\ &\tilde{\mathbf{A}}_{j_1 + \ell \cdot j_2 + \ldots + \ell^p \cdot j_{p+1}} = \mathbf{A}_{j_1} + \ldots + \mathbf{A}_{j_\ell}\ \forall j_1 \ldots j_{p+1} = 1 \ldots \ell \\ &\boldsymbol{\lambda} = [\boldsymbol{\lambda}^{(1)}, \ldots, \boldsymbol{\lambda}^{(\ell^p)}]^\top,\ \sum_{i=1}^{\ell^p} \boldsymbol{\lambda}^{(i)} = \mathbf{v}\ \ \forall k = 1 \ldots \ell \\ &\forall j, \lambda_i \log \frac{\lambda_i}{\lambda_j} \leq -(\tilde{\mathbf{A}}_j - \tilde{\mathbf{A}}_i)^\top \mathbf{z}^{(i)}\}) \end{aligned}$$

With the dual SAGE cone defined, we may find a dual formulation to the SAGE relaxation.

$$
\begin{aligned}
f^{(p)}_{SAGE} &= \sup_{\lambda} \lambda \text{ s.t. } \operatorname{Sig}(\mathbf{c}, \mathbf{A}) - \lambda \in SAGE^{(p)}(\mathbf{A}, \mathcal{X}) && (3.1)\\
&= \sup_{\lambda} \lambda \text{ s.t. } \mathbf{c} - \lambda[1, 0, \ldots] \in C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X}) && (3.2)\\
&= \sup_{\lambda} \inf_{\mathbf{v} \in C^{(p)}_{SAGE}(\mathbf{A},\mathcal{X})^*} \lambda + (\mathbf{c} - \lambda[1, 0, \ldots])^\top \mathbf{v} && (3.3)\\
&= \inf_{\mathbf{v} \in C^{(p)}_{SAGE}(\mathbf{A},\mathcal{X})^*} \sup_{\lambda} \lambda + (\mathbf{c} - \lambda[1, 0, \ldots])^\top \mathbf{v} && (3.4)\\
&= \inf_{\mathbf{v}} \mathbf{c}^\top \mathbf{v} \text{ s.t. } \mathbf{v} \in C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X})^*, \mathbf{v}_1 = 1 && (3.5)\\
& && (3.6)
\end{aligned}
$$

Strong duality holds; $C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X})^*$ is a cone and $\mathbf{v}$ s.t. $\mathbf{v}_1 = 1$ can be found in the relative interior [7].

We have already seen that $C_{AGE}(\mathbf{A}, \mathcal{X}, i)^*$ is an outer approximation of exponential over the constrained set $\{\exp(\mathbf{Ax}) : \mathbf{x} \in \mathcal{X}\}$, and by construction and so is $C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X})^*$. This allows for a method to recover a feasible solution to the original problem $\inf_{\mathbf{x}\in\mathcal{X}} \operatorname{Sig}(\mathbf{A}, \mathcal{X})(\mathbf{x})$. First, it is possible that the solution to the above dual problem is in the feasible region: $\tilde{\mathbf{v}} = \exp(\mathbf{Ax})$ for some $\mathbf{x} \in \mathcal{X} \in \mathcal{X} \iff \log \tilde{\mathbf{v}} = \mathbf{Ax}$. In such case, $\tilde{\mathbf{v}}$ is indeed an optimal solution the original problem. Suppose $\tilde{\mathbf{v}}$ is not in the feasible region. We may project $\tilde{\mathbf{v}}$ to the feasible region by solving $\operatorname{argmin}_{\tilde{\mathbf{x}}\in\mathcal{X}} || \log \mathbf{v}^* - \mathbf{Ax}||$, which is a convex program. The value of the objective function after such projection, $\mathbf{c}^\top \tilde{\mathbf{x}}$ is then an upper bound on the optimal solution to the minimization problem. Together with $\mathbf{c}^\top \tilde{\mathbf{v}}$, which is a lower bound to the original minimization problem, we obtain a range of the possible optimal values.

3.4.1 Applications to Engineering Problems

A problem of the form:

$$
\begin{aligned}
&\min \operatorname{Sig}(\mathbf{c}^{(0)}, \mathbf{A}^{(0)})(\mathbf{x})\\
&\text{s.t. } \operatorname{Sig}(\mathbf{c}^{(i)}, \mathbf{A}^{(i)})(\mathbf{x}) \le 1 \text{ for } i = 1 \ldots m\\
&\qquad \operatorname{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) = 1 \text{ for } j = m+1 \ldots n
\end{aligned}
$$

Where $(\mathrm{Sig}(\mathbf{c}^{(i)}, \mathbf{A}^{(i)}))_{i=1\ldots m}$ are signomials with positive coefficients and $(\mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)}))_{j=m+1\ldots n}$ are monomials is referred to as geometric programming [14]. Although it is not a convex program in the above form due to the equality constraints on the monomials, it may be converted so with a log transform. In particular, by letting $\mathbf{y} = \log(\mathbf{x}) \iff \mathbf{x} = \exp(\mathbf{y})$, and applying log transform of the functions, $(\mathrm{Sig}(\mathbf{c}^{(i)}, \mathbf{A}^{(i)}))_{i=1\ldots m}$, the signomials with positive coefficients become log-sum-exp functions, which are convex, and $(\mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)}))_{j=m+1\ldots n}$, the monomials, become affine. Such formulation, restricting the objective signomials to consist of only positive coefficients, nevertheless can be used to model many physical processes such as communication systems [9] and circuit design. [7]

A generalization of the above formulation are problems of the following form:

$$\begin{aligned} \min\, & \mathrm{Sig}(\mathbf{c}^{(0)}, \mathbf{A}^{(0)})(\mathbf{x}) \\ & \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) \leq 0 \quad \text{for } j = 1 \ldots n \end{aligned}$$

Where $\mathrm{Sig}(\mathbf{c}^{(0)}, \mathbf{A}^{(0)})(\mathbf{x})$, $(\mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}))_{j=1\ldots n}$ are general signomial functions. To see that it is a strict generalization, we may move the nonzero constant over to one side, and express the equality constraint with two inequality constrains. Such formulation encompasses problems in aircraft design [1] [2], in optimizing the structure and geometry of various parts such as wing, vertical tail and fuselage. When any of $\mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)}$ is non-convex, however, the feasible region of the above problem may be non-convex. Indeed, in many real problems, the constraint signomials are not guaranteed to be convex. Murray et al. have proposed to use *partial dualization*, where Lagrangian relaxation to the constrained problem is applied only with the non-convex signomials [23]. Specifically, let:

$$\begin{aligned} I &= \{j : j \geq 1,\ \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)}) \text{ is convex}\} \\ K &= \{j : j \geq 1,\ \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)}) \text{ is non-convex}\} \\ \mathcal{X} &= \{\mathbf{x} : \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) \leq 0 \quad \text{for } j \in I\} \end{aligned}$$

Consider the following formulation:

$$f_{\mathcal{X}}^{K} = \sup_{\gamma}\{\gamma : \lambda_j \geq 0 \text{ for } j \in K,\ \mathrm{Sig}(\mathbf{c}^{(0)}, \mathbf{A}^{(0)})(\mathbf{x}) + \sum_{j\in K} \lambda_j \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) - \gamma \geq 0\ \forall \mathbf{x} \in \mathcal{X}\}$$

Where $\lambda_j \geq 0$ for $j \in K$ are the dual variables. One may check that $f_{\mathcal{X}}^{L} \leq f_{\mathcal{X}}^{*}$ by noting the roles of the dual variables when some $\mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) \leq 0$ is satisfied. We may combine partial dualization with SAGE relaxation as follows:

$$f_{\mathcal{X}}^{(SAGE,K)} = \sup_{\gamma}\{\gamma : \lambda_i \geq 0 \text{ for } i \in K,\ \mathrm{Sig}(\mathbf{c}^{(0)}, \mathbf{A}^{(0)})(\mathbf{x}) + \sum_{j \in K} \lambda_j \mathrm{Sig}(\mathbf{c}^{(j)}, \mathbf{A}^{(j)})(\mathbf{x}) - \gamma \\ \in \mathrm{SAGE}(\mathbf{A}, \mathcal{X})\}$$

Where A is a matrix that includes the rows of all of $(\mathbf{A}^{(j)})_{j=0\ldots n}$. It follows that $f_{\mathcal{X}}^{(SAGE,L)} \leq f_{\mathcal{X}}^{*}$. Such blending of the two methods allow application of SAGE in wider range of optimization problems involving signomials, even when the constraint set is not strictly convex.

Chapter 4

A Completeness Theorem

In this chapter we discuss the main contribution of the thesis, a completeness theorem for conditional SAGE.

4.1 Main Result

Theorem 9 *Let* $\mathbf{A} = [\mathbf{A}_1 \dots \mathbf{A}_\ell]^\top \subset \mathbb{Q}^{\ell \times n}$ *and* $\mathbf{c} \in \mathbb{R}^\ell$*. Consider a signomial* $Sig(\mathbf{c}, \mathbf{A})$ *and a compact convex set* $\mathcal{X} \subset \mathbb{R}^n$*. If* $Sig(\mathbf{c}, \mathbf{A})(\mathbf{x}) > 0 \ \forall \mathbf{x} \in \mathcal{X}$*, then there exists some* $p \in \mathbb{Z}_+$ *s.t.* $(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1}))^p Sig(\mathbf{c}, \mathbf{A}) \in SAGE(E_{p+1}(\mathbf{A}), \mathcal{X})$*.*

Few notable consequences in relation to what we discussed follow from the theorem. First, it shows that the given a fixed set of exponents and a compact convex set, SAGE cone becomes exactly the set of coefficients for the signomial defined by the exponents is positive over the set at finite level of hiearchy. That is,

$$\exists p \in \mathbb{Z}_+ \text{ s.t. } C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X}) = C_{NNS}(\mathbf{A}, \mathcal{X})$$

Similarly, the dual SAGE cone can express the set of exponentials over a convex set at some finite level of the hierarchy.

$$\exists p \in \mathbb{Z}_+ \text{ s.t. } C^{(p)}_{SAGE}(\mathbf{A}, \mathcal{X})^* = \{\exp(\mathbf{A}\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$$

It follows that the conditional SAGE relaxation attains optimal value at some finite level in the hierarchy. That is

$$\exists p \in \mathbb{Z}_+ \text{ s.t. } f_{SAGE}^{(p)} = \inf_{\mathbf{x} \in \mathcal{X}} \text{Sig}(\mathbf{c}, \mathbf{A})(\mathbf{x})$$

4.2 Related Work

The theorem may be thought of as a positivstellensatz for signomials over a compact convex set, similar to Schmundgen's Positivstellensatz (Theorem 2) for polynomilas over a compact semialgebraic set, where the verification is through SAGE as opposed to SOS. Positivstellensatz to this day is an active area of study. Recently, Dressler, Ilima and De Wolff have shown a Schmüdgen type positivstellensatz using SONC polynomials [13]. Also recently, Ahmadi and Hall have proposed an optimization free positivstellensatz for polynomial positivity over a semialgebraic set, by which positivity can be verified by checking that the multiplication of two polynomials that depend by objective polynomial $p(\mathbf{x})$ and constraint polynomials $(g_i(\mathbf{x}))_{i=1\ldots\ell}$ [3], without the procedure for finding sum of square representation through SDP. This work was well received in the optimization community as it posed the possibility of much more efficient polynomial optimization algorithms.

Directly relevant to the current result is the completeness theorem for global positivity of signomials, which was presented by Chandrasekaran and Shah in their introduction of SAGE [8]. The result showed that for any globally positive signomial function $f(\mathbf{x})$ satisfying some mild assumptions, there exists some function $w(\mathbf{x})$ and $p \in \mathbb{Z}_+$ such that $w(\mathbf{x})^p f(\mathbf{x})$ may be certified by SAGE. In fact, the proof essentially shows that $w(\mathbf{x})^p f(\mathbf{x})$ is a signomial with positive coefficients. This may be viewed as an analogue of Polya-type results for signomials [28]. The same article also presented a convergent hierarchy for constrained signomial optimization by appealing to representation theorems. In particular, any signomial positive on a compact semialgebraic set may be verified via SAGE certificate through Lagrangian relaxation with the signomials defining the set.

There are a few key characteristics of the current completeness result. First, unlike previous convergent hierarchy for unconstrained SAGE, it does not require assumptions on the exponents besides rationality. Second, the convergent hierarchy holds for *any arbitrary* *compact convex set*. Thus it is more general than historical results for convergent hierarchies on semialgebraic sets (recall that a semialgebraic set is defined by finite polynomial inequalities), or the recent constrained SAGE hierarchy which assumed the constrained set to be defined by signomials. The proof of the current result uses redundant constraints which may be added due to compactness assumed on the convex set, in order to bypass certain issues. In the proofs for previous constrained SAGE and the recent SONC hierarchies, redundant constrains due to compactness assumption is made to guarantee Archimedean property in the set generated by the functions defining the constrained set, which allows appeal to representation theorems. The proof for the current article does not appeal to a representation theorem but rather a recent positivstellensantz result developed by Dickson, covered in the previous chapter. The connection is made mostly through elementary algebraic operations, as will be clear in the rest of the chapter.

4.3 Proof of Main Result

The proof is structured as follows. We first note that any compact convex set $\mathcal{X}$ may be expressed as intersection of a set of (possibly infinite) rational halfspaces $H_{\mathcal{X}}$. We then apply change of variable $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$ and show that positivity of signomial $\mathrm{Sig}(\mathbf{c}, \mathbf{A})$ over intersection of halfspaces $H_{\mathcal{X}}$ implies the positivity of a corresponding polynomial $p(\mathbf{y})$ over a set $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$, the intersection of positive orthant and a set defined by (possibly infinite) polynomial inequalities. We then make modification to the aforementioned set to show that the positivity of $p(\mathbf{y})$ over such set implies its positivity over $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$, the intersection of nonnegative orthant and a homogeneous semialgebraic set (defined by finite homogeneous polynomials). We then appeal to Dickson's Positivstellensatz to claim that $(\mathbf{y}^\top \mathbf{1})^p p(\mathbf{y})$ for some $p \in \mathbb{Z}_+$ yields a representation by polynomials with only positive coefficients and the homogeneous polynomilas defining $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$. Undoing the variable change, we show through the representation that the original signomial after multiplication of extra term, $(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^p \mathrm{Sig}(\mathbf{c}, \mathbf{A})$ is a SAGE signomial.

Without loss of generality, we may make the following assumptions about the structure of exponent vectors $\mathbf{A} \in \mathbb{Q}^{\ell \times n}$ defining $\mathrm{Sig}(\mathbf{c}, \mathbf{A})$.

(a) the first n rows $(\mathbf{A}_j)_{j=1\ldots n}$ are linearly independent

(b) $\mathbf{A}_{n+1} = \mathbf{0}$

The assumptions are in fact not restrictive. To satisfy the first condition, we may select a set of linearly independent vectors as the first n. The proof is easily generalized to the case when the span of the vectors has dimension less than n. The second condition is not restrictive either, since we may insert a zero vector into the set of exponents. However, it is a variable required for in certain step in the proof.

The proof at heart is a reduction to Dickson's Positivstellensatz. The proof is divided into sections explaining each step in the reduction.

4.3.1 Representation of Compact Convex Set as Rational Halfspaces

In the section we discuss the following result which finds connection between compact convex set and rational half-spaces. It has been shown by Silva and Tuncel [33], and we will paraphrase it below.

Theorem 10 *([33]) A rational halfspace $h \subset \mathbb{R}^n$ is a set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{w}^\top \mathbf{x} \leq \mathbf{b}\}$ where $\mathbf{w} \in \mathbb{Q}^n, \mathbf{b} \in \mathbb{Q}$. Given a compact convex set $\mathcal{X}$, there exists a set of rational halfspaces H such that $\mathcal{X} = \cap_{h \in H} h$*

Proof: The proof is nonconstructive. It suffices to show that for each $\mathbf{y} \in \mathbb{R}^n \backslash \mathcal{X}$, there exists $\mathbf{w} \in \mathbb{Q}^n, b \in \mathbb{Q}$ such that $\mathbf{w}^\top \mathbf{x} \leq b < \mathbf{w}^\top \mathbf{y}\ \forall \mathbf{x} \in \mathcal{X}$. Recall that $\mathbb{Q}$ is dense in $\mathbb{R}$. Thus it suffices to show that for each $\mathbf{y} \in \mathbb{R}^n \backslash \mathcal{X}$, there exists $\mathbf{w} \in \mathbb{Q}^n$ such that $\sigma_{\mathcal{X}}(\mathbf{w}) < \mathbf{w}^\top \mathbf{y}$. Then by denseness, we may find $b \in \mathbb{Q}$ s.t $\sigma_{\mathcal{X}}(\mathbf{w}) < b < \mathbf{w}^\top \mathbf{y}$, which implies the desired result.

Let ξ be the set the intersection of all halfspaces defined by rational vector containing $\mathcal{X}$. $\xi = \{\mathbf{x} : \mathbf{w}^\top \mathbf{x} \leq \sigma_{\mathcal{X}}(\mathbf{w}),\ \forall \mathbf{w} \in \mathbb{Q}^n\}$. Proof is complete if $\mathcal{X} = \xi$, since by definition of ξ, if $\mathbf{y} \in \mathbb{R}^n \backslash \xi$, there exists some $\mathbf{w} \in \mathbb{Q}^n$ such that $\sigma_{\mathcal{X}}(\mathbf{w}) < \mathbf{w}^\top \mathbf{y}$. Indeed, $\mathcal{X} = \xi$ as shown below.

First, suppose $\bar{\mathbf{x}} \in \mathcal{X}$. For any $\mathbf{w} \in \mathbb{Q}^n$, $\mathbf{w}^\top \bar{\mathbf{x}} \leq \sigma_{\mathbb{X}}(\mathbf{w})$ by definition, and thus $\bar{\mathbf{x}} \in \xi$.

$\mathcal{X} \subseteq \xi$ by construction.

Next, suppose $\bar{\mathbf{x}} \in \xi$. Consider any $\bar{\mathbf{w}} \in \mathbb{R}^n$. Let $(\mathbf{w}_k)_{k\in\mathbb{N}} \subset \mathbb{Q}^n$ be a sequence of vectors converging to $\bar{\mathbf{w}}$. Since $\mathcal{X}$ is compact, support function $\sigma_{\mathcal{X}}(\cdot)$ is continuous. Image of convergent sequence under continuous map converges, thus $(\sigma_{\mathcal{X}}(\mathbf{w}_k))_{k\in\mathbb{N}}$ converges to $\sigma_{\mathcal{X}}(\bar{\mathbf{w}})$. By assumption and definition of ξ, $\mathbf{w}_k^\top \bar{\mathbf{x}} \leq \sigma_{\mathcal{X}}(\mathbf{w}_k)$ for all k. The limit preserves inequality, so $\bar{\mathbf{w}}^\top \bar{\mathbf{x}} \leq \sigma_{\mathcal{X}}(\bar{\mathbf{w}})$. This implies $\bar{\mathbf{x}} \in \{\mathbf{x} \in \mathbb{R}^n : \mathbf{w}^\top \mathbf{x} \leq \sigma_{\mathcal{X}}(\mathbf{w}) \ \forall \mathbf{w} \in \mathbb{R}^n\}$. By Theorem 13.1 of [30], a convex set is the intersection of all of its supporting hyperplanes. i.e. $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{w}^\top \mathbf{x} \leq \sigma_{\mathcal{X}}(\mathbf{w}) \ \forall \mathbf{w} \in \mathbb{R}^n\}$ and thus $\bar{\mathbf{x}} \in \mathcal{X}$. ■

Without loss of generality, for the compact convex set $\mathcal{X}$, let $H_{\mathcal{X}}$ denote the set of rational halfspaces defining $\mathcal{X}$. We have shown:

$$\mathbf{x} \in \mathcal{X} \iff \cap_{h\in H_{\mathcal{X}}} h$$

4.3.2 Signomial to Polynomial

We have now reduced the positivity of signomial $\mathrm{Sig}(\mathbf{c}, \mathbf{A})$ over compact convex set $\mathcal{X}$ to its positivity over intersection of rational halfspaces $H_{\mathcal{X}}$. In this section we apply the change of variable $\mathbf{y} = \exp(\mathbf{A}\mathbf{x}) \in \mathbb{R}^\ell_{++}$. We obtain a polynomial after the change of variable $p(\mathbf{y}) = \mathbf{c}^\top \mathbf{y}$ as result. Further, we are interested in the sufficient and necessary conditions on $\mathbf{y} = \exp(\mathbf{A}\mathbf{x}) \in \mathbb{R}^\ell_{++}$ given $\mathbf{x} \in \cap_{h\in H_{\mathcal{X}}} h$.

Representation of Exponentials As Polynomial Constraints

When $\mathbf{A} \in \mathbb{Q}^{\ell\times n}$ has rank less than ℓ, the range does not cover $\mathbb{R}^\ell$, which adds restriction to $\mathbf{y} = \exp(\mathbf{A}\mathbf{x}) \in \mathbb{R}^\ell_{++}$. Since first n exponents $(\mathbf{A}_j)_{j=1\ldots n}$ are linearly independent, the rest of the vectors may be expressed as rational linear combinations of the first n vectors, and thus they are constrained by the first n vectors. For $\mathbf{A}_j$ with $j \geq n+2$, $\mathbf{A}_j = \sum_i^{n+1} w_i^{(j)} \mathbf{a}^{(i)}$ where $\mathbf{w}^{(j)} \in \mathbb{Q}^n$ for all j. Then;

$$y_j = \exp(\mathbf{A}_j\mathbf{x}) = \exp(\sum_i^{n+1} w_i^{(j)} \mathbf{A}^{(i)}\mathbf{x}) = \prod_i^{n+1} \exp(\mathbf{A}_j\mathbf{x})^{w_i^{(j)}} = \prod_i^{n+1} y_i^{w_i^{(j)}} = \prod_i^{n} y_i^{w_i^{(j)}}$$

The last step is from the fact that $\mathbf{A}_n = \mathbf{0}$. $w_i^{(j)} \geq 0 \ \ \forall i$ and since $w_i^{(j)}$'s are rationals, we may raise both sides by the smallest common denominator to clear the fractions. For example, $y_j = y_1^{\frac{1}{2}} y_2^{\frac{1}{4}} \iff y_j^4 = y_1^2 y_2$.

We may apply such operation to y_j for all $j \geq n+2$. The operation is only valid in the positive orthant.

$$y_j = \exp(\mathbf{A}_j \mathbf{x}) = \prod_i^n y_i^{w_i^{(j)}} \ \forall j = n+2...l \iff y_j^{\lambda_j^{(j)}} = \prod_i^n y_i^{\lambda_i^{(j)}} \ \forall j = n+2...l$$
$$\iff y_j^{\lambda_j^{(j)}} - \prod_i^n y_i^{\lambda_i^{(j)}} = 0 \ \forall j = n+2...l$$

Where $\boldsymbol{\lambda}^{(j)}$'s are obtained from the above procedure.

Rational Halfspace to Polynomial Constraint

To characterize the constraint $\mathbf{x} \in \cap_{h \in H_{\mathcal{X}}}$ on $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$, we begin by considering single rational halfspace constraint on $\mathbf{x} \in \mathbb{R}^n$. Let $h = \{\mathbf{x} : \mathbf{w}^\top \mathbf{x} \leq d\} \in H_{\mathcal{X}}$. $\mathbf{w} \in \mathbb{Q}^n$ and $d \in \mathbb{Q}$. Recall that $\mathbf{A} \in \mathbb{Q}^{\ell \times n}$. Given such constraint on $\mathbf{x}$, what can we say about the exponential of rational linear map of $\mathbf{x}$, $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$?

First consider the rational linear map $\mathbf{A}\mathbf{x} \in \mathbb{R}^\ell$ given a rational half-space constraint on $\mathbf{x}$: $\mathbf{w}^\top \mathbf{x} \leq d$, we may find a rational halfspace constraint on the linear map. By assumption we have that $\text{rank}(\mathbf{A}) = n$ which implies columns of $\mathbf{A}$ are linearly independent. Thus there exits a left-inverse $\mathbf{L} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \in \mathbb{Q}^{n \times \ell}$ such that $\mathbf{L}\mathbf{A}\mathbf{x} = \mathbf{x}$. The rationality of $\mathbf{L}$ follows from the rationality of $\mathbf{A}$ and inverse operation which preserves rationality. Letting $\mathbf{b} = \mathbf{w}^\top \mathbf{L} \in \mathbb{Q}^n$, we have that $\mathbf{b}^\top (\mathbf{A}\mathbf{x}) \leq d$.

We then make the observation that exponential of a polyhedron is equivalent to polynomial inequality. Beginning with $\mathbf{y} = \exp(\mathbf{A}\mathbf{x}) \iff \log \mathbf{y} = \mathbf{A}\mathbf{x}$, a series of algebraic operations follows below. In doing so, we assume $\mathbf{y} \in \mathbb{R}^n_{++}$.

$$\begin{aligned} \mathbf{b}(\mathbf{A}\mathbf{x}) \le d \iff \mathbf{b}(\log \mathbf{y}) \le d &\iff \sum_{i=1}^{n} b_i(\log y_i) \le d \\ &\iff \sum_{i=1}^{n}(\log y_i^{b_i}) \le d \iff \log(\prod_{i=1}^{n} y_i^{b_i}) \le d \\ &\iff \prod_{i=1}^{n} y_i^{b_i} \le \exp(d) \iff \prod_{i\,:\,b_i>0} y_i^{b_i} \le \exp(d) \prod_{i\,:\,b_i<0} y_i^{-b_i} \end{aligned}$$

The last step moves terms with negative power by multiplication on both sides. For example; $y_1^2 y_2^{-3} \le 1 \iff y_1 \le y_2^3$. Now since $\mathbf{b} \in \mathbb{Q}^n$ has rational entries, we may raise both sides by a common denominator to clear fraction. Let m be the common denominator of $(b_i)_{i=1\ldots n}$

$$\begin{aligned} &\prod_{i\,:\,b_i>0} y_i^{b_i} \le \exp(d) \prod_{i\,:\,b_i<0} y_i^{-b_i} \\ \iff &\prod_{i\,:\,b_i>0} y_i^{mb_i} \le \exp(md) \prod_{i\,:\,b_i<0} y_i^{-mb_i} \\ \iff &\exp(md) \prod_{i\,:\,b_i<0} y_i^{-mb_i} - \prod_{i\,:\,b_i>0} y_i^{mb_i} \ge 0 \end{aligned}$$

which are indeed polynomial equalities.

Intersection of Rational Halfspaces

In the above, single halfspace constraint has been shown equivalent to a polynomial constraint. To extend the above to intersection of (possibly infinite) halfspaces, we simply take the intersection of the polynomial equations generated from them. For each k, Let $\boldsymbol{\gamma}^{(k)} = m^{(k)}\mathbf{B}_{k,:} \in \mathbb{Z}^\ell$ and $c^{(k)} = \exp(m^{(k)}d) \in \mathbb{R}_{++}$. Let $K_{\mathcal{X}}$ be a set of (possibly infinite) indices depending on the rational halfspaces defining $\mathcal{X}$. We may write as below.

$$c^{(k)} \prod_{i\,:\,\gamma_i^{(k)}<0} y_i^{-\gamma_i^{(k)}} - \prod_{i\,:\,\gamma_i^{(k)}>0} y_i^{\gamma_i^{(k)}} \ge 0 \;\; \forall k \in K_{\mathcal{X}}$$

Now, consider the following set that depends on $\mathbf{A} \in \mathbb{Q}^{\ell\times n}$ and $\mathcal{X} \subset \mathbb{R}^n$. Recall the assumption $\mathbf{A}_{n+1} = \mathbf{0}$, which implies $y_{n+1} = \exp(\mathbf{A}_{n+1}\mathbf{x}) = 1$.

$$\begin{aligned}
\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} &= \{\mathbf{y} = \exp(\mathbf{A}\mathbf{x}) : \mathbf{x} \in \mathcal{X}\} \\
&= \{\mathbf{y} \in \mathbb{R}^n_{++} :\ y_{n+1} = 1,\ y_j^{\lambda_j^{(j)}} - \prod_i^n y_i^{\lambda_i^{(j)}} = 0 \ \ \forall j = n+2...l \\
&\qquad c^{(k)} \prod_{i\,:\,\gamma_i^{(k)}<0} y_i^{-\gamma_i^{(k)}} - \prod_{i\,:\,\gamma_i^{(k)}>0} y_i^{\gamma_i^{(k)}} \geq 0 \ \forall k \in K_{\mathcal{X}}\} \\
&= \{\mathbf{y} \in \mathbb{R}^n_{++} :\ y_{n+1} = 1,\ p_1^{(j)}(\mathbf{y}) - p_2^{(j)}(\mathbf{y}) \geq 0 \ \forall j = n+2...l, \\
&\qquad q_1^{(k)}(\mathbf{y}) - q_2^{(k)}(\mathbf{y}) \geq 0 \ \forall k \in K_{\mathcal{X}}\}
\end{aligned}$$

Where in the last expression we have written the terms abstractly. Each $p_1^{(j)}(\mathbf{y})$, $p_2^{(j)}(\mathbf{y})$, $q_1^{(k)}(\mathbf{y})$, $q_2^{(k)}(\mathbf{y})$ is monomial. By construction, we have that $\mathbf{x} \in \mathcal{X} \iff \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$.

Signomial Positivity to Polynomial Positivity

The change of variable $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$ subject to $\mathbf{x} \in \mathcal{X}$ reduces signomial positivity over a compact convex set to polynomial positivity over intersection of positive orthant and set defined (possible infinite) polynomial inequalities. In summary, we have shown the following:

$$\mathrm{Sig}(\mathbf{c}, \mathbf{A}) > 0 \ \forall \mathbf{x} \in \mathcal{X} \implies \mathbf{c}^\top \mathbf{y} > 0 \ \forall \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$$

4.3.3 Positivity to Positivstellensatz

We recall however that the Dickson's Positivstellensatz assumes polynomial positivity over intersection of *nonnegative* orthant and set defined by finite homogeneous polynomials, excluding the origin. Namely, there are three conditions that $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ does not satisfy.

1. $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ does not include the faces of the nonnegative orthant.
2. $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ is defined by polynomials that are possibly non-homogeneous.
3. $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ is defined by possibly infinite polynomials.

The goal of this section is to describe modifications to the set $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ such that it satisfies the premises of Dickson's Positivstellensatz, while positivity is preserved. That is, there exists a set $\mathcal{T}^{(\mathbf{A},\mathcal{X})*}$ defined as the intersection of the nonnegative orthant and a semialgebraic set defined by homogeneous polynomials such that:

$$\mathbf{c}^\top \mathbf{y} > 0 \ \forall \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} \implies \mathbf{c}^\top \mathbf{y} > 0 \ \forall \mathbf{y} \in (\mathcal{T}^{(\mathbf{A},\mathcal{X})*} \backslash \mathbf{0}).$$

Inclusion of Points on the Faces of Nonnegative Orthant

In this subsection, we make modifications to $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ so that the resulting set is defined at the intersection of the nonnegative orthant and set of polynomial equations, while preserving positivity. Consider the following set that extends $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ to nonnegative orthant.

$$\begin{aligned}\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} = \{\mathbf{y} \in \mathbb{R}^n_{++} :\ & y_{n+1} = 1,\ p_1^{(j)}(\mathbf{y}) - p_2^{(j)}(\mathbf{y}) \geq 0 \ \forall j = n+2 \ldots \ell, \\ & q_1^{(k)}(\mathbf{y}) - q_2^{(k)}(\mathbf{y}) \geq 0 \ \forall k \in K_{\mathcal{X}}\}\end{aligned}$$

However, the positivity of $\mathbf{c}^\top \mathbf{y}$ is not preserved in general. That is:

Proposition 1 *There exists $\mathbf{A} \in \mathbb{Q}^{\ell \times n}$, $\mathbf{c} \in \mathbb{R}^\ell$ and compact convex set $\mathcal{X}$ such that*

$$\mathbf{c}^\top \mathbf{y} > 0 \ \forall \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} \not\Longrightarrow \mathbf{c}^\top \mathbf{y} > 0 \ \forall \mathbf{y} \in (\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \backslash \mathbf{0}).$$

Proof: We prove this by an example. Let $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4]$ where $\mathbf{a}_1 = [1,0]^\top, \mathbf{a}_2 = [0,1]^\top, \mathbf{a}_3 = [0,0]^\top, \mathbf{a}_4 = [0.1, 0.1]^\top$, $\mathbf{c} = [1,1,-1,0]^\top$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^3 : -2x_1 + x_2 \leq 0, x_1 - 2x_2 \leq 0, x_1 + x_2 \leq 1\}$. $\mathcal{X}$ is conveniently defined to be a compact intersection of halfspaces.

Let $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$. We have

$$y_1 = \exp(x_1),\ y_2 = \exp(x_2),\ y_3 = 1,\ y_4^{10} = y_1 y_2$$

Also

$$\begin{aligned}\mathbf{x} \in \mathcal{X} &\iff -2x_1 + x_2 \leq 0, x_1 - 2x_2 \leq 0 \text{ and } x_1 + x_2 \leq 1 \\ &\iff -2\log y_1 + \log y_2 \leq 0, \log y_1 - 2\log y_2 \leq 0 \text{ and } \log y_1 + \log y_2 \leq 1 \\ &\iff \log(y_1^{-2} y_2) \leq 0, \log(y_1 y_2^{-2}) \leq 0 \text{ and } \log(y_1 y_2) \leq 1 \\ &\iff y_1^2 - y_2 \geq 0, y_2^2 - y_1 \geq 0 \text{ and } e - y_1 y_2 \geq 0\end{aligned}$$

Therefore $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} = \{\mathbf{y} \in \mathbb{R}_{++}^4 : y_4^{10} = y_1y_2,\ y_1^2 - y_2 \geq 0,\ y_2^2 - y_1 \geq 0,\ e - y_1y_2 \geq 0\}$ and $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} = \{\mathbf{y} \in \mathbb{R}_{+}^4 : y_4^{10} = y_1y_2,\ y_1^2 - y_2 \geq 0,\ y_2^2 - y_1 \geq 0,\ e - y_1y_2 \geq 0\}$.

First, we may check that $\mathbf{c}^\top \mathbf{y} \geq 1\ \forall \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$, as the constraints $y_1^2 - y_2 \geq 0,\ y_2^2 - y_1 \geq 0$ may only be satisfied when $y_1, y_2 \geq 1$ given $\mathbf{y} > \mathbf{0}$. However, letting $\bar{\mathbf{y}} = [0,0,1,0]^\top \in \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})}$, $\mathbf{c}^\top \bar{\mathbf{y}} = -1 < 0$. ■

While counter-intuitive at first sight, the reason is in fact simple: while $\mathbf{y} = \exp(\mathbf{Ax})$ implicitly implies $\mathbf{y} \in \mathbb{R}_{++}^n$, the resulting polynomial equations after algebraic operations do not. Explicit restriction of $\mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$ to $\mathbb{R}_{++}^n$ retains the constraint, but extending the definition to nonnegative orthant includes "jumps" with some coordinates of $\mathbf{y}$ being zero to satisfy the polynomial constraints.

We may avoid such jumps by adding redundant constraints due to the compactnesss of $\mathcal{X}$. Thus $\mathbf{y} \in \exp(\mathbf{A}\mathcal{X})$ is compact. And $l_i \leq y_i \leq u_i$ for each i. Define the following set.

$$\Upsilon = \{\mathbf{y} : u_i - y_i \geq 0,\ y_i - l_i \geq 0 \quad \forall i = 1 \dots \ell\}$$

Υ, while being polynomial constraints, restricts to positive orthant since $\mathbf{y} \notin \Upsilon$ if $y_i = 0$ for any i. In other words, the set is now restricted to be in the positive orthant through redundant constraints. By definition of ℓ_i and u_i, $\Upsilon \supseteq \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$. Thus $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon = \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})}$. $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$ is indeed an intersection of nonnegative orthant and polynomial equations. The constraints in Υ are also useful for making a certain claim, which is shown in the next section.

Positivity over Non-negative Orthant of Homogeneous Polynomials

In this sub-section, we show a modification to polynomials in the definition of $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$ so that they are homogeneous. We also show that the modification preserves positivity.

$$\begin{aligned} \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon = \{\mathbf{y} \in \mathbb{R}_{++}^n :\ & y_{n+1} = 1,\ p_1^{(j)}(\mathbf{y}) - p_2^{(j)}(\mathbf{y}) \geq 0\ \forall j = n+2...l, \\ & q_1^{(k)}(\mathbf{y}) - q_2^{(k)}(\mathbf{y}) \geq 0\ \forall k \in K_{\mathcal{X}} \\ & u_i - y_i \geq 0, y_i - l_i \geq 0 \quad \forall i = 1 \dots \ell\} \end{aligned}$$

We modify the polynomials to be homogeneous by making the following transformation. Let $[x]_+ = \max(x, 0)$

$$
\begin{aligned}
\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon} = \{\mathbf{y} \in \mathbb{R}^n_+ :& \\
& y_{n+1}^{[\deg(p_2^{(j)})-\deg(p_1^{(j)})]_+} p_1^{(j)}(\mathbf{y}) - y_{n+1}^{[\deg(p_1^{(j)})-\deg(p_2^{(j)})]_+} p_2^{(j)}(\mathbf{y}) = 0 \ \forall j = n+2...l \\
& y_{n+1}^{[\deg(q_2^{(k)})-\deg(q_1^{(k)})]_+} q_1^{(k)}(\mathbf{y}) - y_{n+1}^{[\deg(q_1^{(k)})-\deg(q_2^{(k)})]_+} q_2^{(k)}(\mathbf{y}) \geq 0 \ \forall k \in K_{\mathcal{X}} \\
& u_i y_{n+1} - y_i \geq 0, y_i - v_i y_{n+1} \geq 0 \ \forall i = 1 \ldots \ell\}
\end{aligned}
$$

We have multiplied y_{n+1} to appropriate terms so that the polynomials are homogeneous. We have also removed the condition $y_{n+1} = 1$. The set has been modified considerably. However, we claim the following:

Theorem 11 *Consider a signomial Sig*$(\mathbf{c}, \mathbf{A})$ *and compact convex set* $\mathcal{X}$ *as in Theorem 9. Then*

$$\mathbf{c}^\top \mathbf{y} > 0 \ \ \forall \mathbf{y} \in (\mathcal{T}_+^{(\mathbf{A},\mathcal{X})} \cap \Upsilon) \backslash \mathbf{0} \implies \mathbf{c}^\top \mathbf{y} > 0 \ \ \forall \mathbf{y} \in (\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}) \backslash \mathbf{0}$$

Proof: Consider some $\mathbf{y} \in (\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}) \backslash \mathbf{0}$. First we have that $y_{n+1} \neq 0$ for if otherwise $\mathbf{y} = \mathbf{0}$ by the constraints $u_i y_{n+1} - y_i \geq 0 \quad \forall i$. Let $\widetilde{\mathbf{y}} = \mathbf{y}/y_{n+1}$. Since $\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$ is a semialgebraic set defined by homogeneous polynomials, it is closed under positive scaling thus $\widetilde{\mathbf{y}} \in \widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$. Since $\widetilde{\mathbf{y}}_{n+1} = 1$, the conditions for $\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$ reduces to conditions for $\mathcal{T}_+^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$, and thus $\widetilde{\mathbf{y}} \in \mathcal{T}_+^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$. By assumption $\mathbf{c}^\top \widetilde{\mathbf{y}} > 0$. Recalling that $y_{n+1} > 0$, $\mathbf{c}^\top \mathbf{y} = \mathbf{c}^\top (y_{n+1} \widetilde{\mathbf{y}}) > 0$. ■

Infinite to finite polynomial inequalities

$\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$ is defined by possibly infinite polynomials. In this sub-section, we want to show that positivity over such set implies positivity over a set defined by finite polynomials.

Consider the following theorem, extended from the original statement in [12].

Theorem 12 *Consider a set of homogeneous polynomials* $\{f_0\} \cup \{f_i \mid i \in I\} \subseteq \mathbb{R}[\mathbf{x}]$ *with infinite cardinality. If* $f_0(\mathbf{x}) > 0$ *for all* $\mathbf{x} \in \mathbb{R}^n_+ \cap \bigcap_{i \in I} f_i^{-1}(\mathbb{R}_+) \backslash \{\mathbf{0}\}$*, there exists a subset* $J \subseteq I$ *of finite cardinality such that* $f_0(\mathbf{x}) > 0$ *for all* $\mathbf{x} \in \mathbb{R}^n_+ \cap \bigcap_{i \in J} f_i^{-1}(\mathbb{R}_+) \backslash \{\mathbf{0}\}$.

The proof of the above theorem is left in the appendix, and is adapted from [12] as well.

In Theorem 12, let $\mathbb{R}^n_+ \cap \bigcap_{i\in I}^m f_i^{-1}(\mathbb{R}_+) = \widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$ and $\mathbb{R}^n_+ \cap \bigcap_{i\in J} f_i^{-1}(\mathbb{R}_+) = \mathcal{T}_+^{(\mathbf{A},\mathcal{X})*}$. Then we have that $\mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in (\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon})\backslash\mathbf{0} \implies \mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in \mathcal{T}_+^{(\mathbf{A},\mathcal{X})*}$. Although the proof is non constructive, the the finite polynomials defining $\mathcal{T}_+^{(\mathbf{A},\mathcal{X})*}$ are a subset of ones defining $\widetilde{\mathcal{T}}_+^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$.

With the above, we have completed the reduction as below

$$\begin{aligned}
\mathrm{Sig}(\mathbf{c},\mathbf{A}) > 0 \; \forall \mathbf{x} \in \mathcal{X} &\implies \mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} \\
&\implies \mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon \\
&\implies \mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in (\widetilde{\mathcal{T}}_{+}^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon})\backslash\mathbf{0} \\
&\implies \mathbf{c}^\top \mathbf{y} > 0 \; \forall \mathbf{y} \in \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}\backslash\mathbf{0}.
\end{aligned}$$

where $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$ is an intersection of nonnegative orthant and a semialgebraic set defined by finite homogeneous polynomials, as desired. In addition, we have the following relation.

Theorem 13 *Consider any* $\mathbf{A} \in \mathbb{Q}^{\ell\times n}$ *and* $\mathcal{X}$ *as in Theorem 9. With change of variable* $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$*:*

$$\mathbf{x} \in \mathcal{X} \implies \mathbf{y} \in (\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*})\backslash\mathbf{0}$$

Proof: We have already shown $\mathbf{x} \in \mathcal{X} \iff \mathbf{y} \in \mathcal{T}_{++}^{(\mathbf{A},\mathcal{X})} = \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$. We have $y_{n+1} = \exp(\mathbf{a}_{n+1}^\top \mathbf{x}) = 1$ by assumption on $\mathbf{A}$. By construction of $\widetilde{\mathcal{T}}_{+}^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$, if $\mathbf{y} \in \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})} \cap \Upsilon$ with $y_{n+1} = 1$, then $\mathbf{y} \in \widetilde{\mathcal{T}}_{+}^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon}$, as the modification is null when $y_{n+1} = 1$. $\widetilde{\mathcal{T}}_{+}^{(\mathbf{A},\mathcal{X})} \cap \widetilde{\Upsilon} \subseteq \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$ since the latter is defined by subset of polynomials of those defining the former. So $\mathbf{y} \in \mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$. Lastly, $\mathbf{y} = \exp \mathbf{A}\mathbf{x} \neq \mathbf{0}$, which completes the proof. ■

4.3.4 Positivstellensatz to Conditional SAGE

In this sub-section, we appeal to Dickson's Positivstellensatz to find a representation of the polynomial $p(\mathbf{y})$. Observe that all homogeneous polynomials defining $\mathcal{T}_{+}^{(\mathbf{A},\mathcal{X})*}$ are of the form $m_1(\mathbf{y}) - m_2(\mathbf{y})$; the difference of two monomials. Without loss of generality, we may write: $\mathcal{T}_+^{(\mathbf{A},\mathcal{X})*} = \{\mathbf{y} \in \mathbb{R}_+^n : m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y}) \geq 0,\ j = 1 \dots m\}$. As a consequence of Dickson's Positivstellensatz. (Theorem 4), for some $p \in \mathbb{Z}_+$, there exists homogeneous posynomials $g_1(\mathbf{y}) \dots g_m(\mathbf{y})$ s.t.

$$(\mathbf{y}^\top \mathbf{1})^p \mathbf{c}^\top \mathbf{y} = \sum_{j=1}^{m} g_j(\mathbf{y})(m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y}))$$

We show explicitly below that the the RHS is summation of terms each of which is positive and have at most one negative term. Recalling that $\mathbf{y} = \exp(\mathbf{A}\mathbf{x})$, this then implies that it is a SAGE signomial. A key observation is that the LHS is a homogeneous polynomial, and $m_1^{(j)}(\mathbf{y}) - m_2^{(j)}$ is homogeneous for each j. Thus without loss of generality, for all j, $\deg(g_j(\mathbf{y})) + \deg(m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y})) = p + 1$. That is, we may ignore indices that yield polynomials with degree that is not $p+1$, since they must cancel out. Moreover, for each j, without loss of generality $g_j(\mathbf{y}) = \sum_k^{\ell(j)} h_k^{(j)}(\mathbf{y})$. where $h_k^{(j)}$ is a monomial and $\ell(j)$ is the number of terms in polynomial $g_j(\mathbf{y})$. Then:

$$\begin{aligned}(\mathbf{y}^\top \mathbf{1})^p \mathbf{c}^\top \mathbf{y} &= \sum_{j=1}^{m} g_j(\mathbf{y})(m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y})) \\ &= \sum_{j=1}^{m} \sum_{k=1}^{\ell(j)} h_k^{(j)}(\mathbf{y})(m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y}))\end{aligned}$$

Now:

$$\begin{aligned}\mathbf{x} \in \mathcal{X} &\implies \mathbf{y} = \exp \mathbf{A}\mathbf{x} \in \mathcal{T}_+^{(\mathbf{A},\mathcal{X})*} \\ &\implies m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y}) \geq 0 \ \ \forall j = 1 \dots m \\ &\implies h_k^{(j)}(\mathbf{y})(m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y})) \geq 0 \ \ \forall j, k \\ &\implies h_k^{(j)}(\exp \mathbf{A}\mathbf{x})(m_1^{(j)}(\exp \mathbf{A}\mathbf{x}) - m_2^{(j)}(\exp \mathbf{A}\mathbf{x})) \geq 0 \ \ \forall j, k\end{aligned}$$

Let $o_k^{(j)}(\mathbf{x}) = h_k^{(j)}(\exp \mathbf{A}\mathbf{x})(m_1^{(j)}(\exp \mathbf{A}\mathbf{x}) - m_2^{(j)}(\exp \mathbf{A}\mathbf{x}))\ \forall j, k$. One may verify that it is a signomial in the exponential form. Make the following observations for all j and k:

- $o_k^{(j)}(\mathbf{x}) \geq 0\ \forall \mathbf{x} \in \mathcal{X}$, as implied by above.
- $o_k^{(j)}(\mathbf{x})$ has one negative term, namely $h_k^{(j)}(\exp \mathbf{A}\mathbf{x}) m_2^{(j)}(\exp \mathbf{A}\mathbf{x})$.

- Since $\deg(g_j(\mathbf{y}) + (m_1^{(j)}(\mathbf{y}) - m_2^{(j)}(\mathbf{y}))) = p + 1$, the exponentials of $o_k^{(j)}(\mathbf{x})$ are ones of $E_{p+1}(\mathbf{A})$

Which implies $o_k^{(j)}(\mathbf{x}) \in SAGE(E_{p+1}(\mathbf{A}), \mathcal{X})$. $SAGE(E_{p+1}(\mathbf{A}), \mathcal{X})$ is a cone and thus *closed under addition.*

So $(\exp(\mathbf{A}\mathbf{x})^\top \mathbf{1})^p \sum_{j=1}^{\ell} c_j \exp(\mathbf{A}_j \mathbf{x}) = \sum_{j=1}^{m} \sum_{k=1}^{\ell(j)} o_k^{(j)}(\mathbf{x}) \in SAGE(E_{p+1}(\mathbf{A}), \mathcal{X})$.

Chapter 5

Discussions and Future Work

In this thesis we presented a convergent hierarchy of certificate for signomial positivity. We first described a certificate of signomial positivity derived based from convex duality, and a resulting algorithm for computing the lower bounds of constrained signomial programs.

The main contribution of the thesis shows that the certificate is complete. The proof of the completeness result takes a similar approach as recent positivstellensatz results in reducing positivity of desired form into another [8] [3]. In doing so, we used redundant constraints as done in previous proofs, but rather than appealing to representation theorems, the current proof reduces to Dickson's Positivstellensatz through algebraic operations. The redundant constraints in particular is used to avoid "jumps" when extending the aforementioned set to the intersection of non-negative orthant and polynomial equations, as well as to preserve positivity of polynomial when modifying the set to be homogeneous.

We note a few technical insights in the proof. The proof for convergent hierarchy of unconstrained SP presented by Chandrasekaran et al., is in fact a *optimization-free* positivstellesatz [8]. That is, it shows that given a globally positive signomials satisfying a "mild assumptions" on structure the exponent vectors, multiplication of a function yields a signomials with positive coefficients. In the current proof, the mild assumption is removed by exploiting the compactness assumption. After appealing to Dickson's Positivstellensatz to find a representation of the corresponding polynomial, we exploit the fact that the polynomials defining the semialgebraic set is a difference of two terms to demonstrate a decomposition of the representation into sum of parts that each has one negative term and is positive over the constrained set.

As for future directions, there are open questions resulting from this work. Although the result holds for any compact convex set, we crucially relied on the rationality of the exponents defining the signomials to reduce signomial positivity to polynomial positivity. One direct extension would be to remove such assumption, which would yield a result further unique from the previous polynomial literature. Another direction to pursue is to relax the compactness assumption on the convex set. The difficulty of such extension mainly lies in avoiding jumps when extending the the aforementioned set to nonnegative orthant, which may be done through redundant constraints. We have considered this extension and have found difficulty in developing a general method to construct redundant constraints that avoid all jumps.

References

[1] *The Power of Log Transformation: A Comparison of Geometric and Signomial Programming*. American Institute of Aeronautics and Astronautics, January 2018.

[2] Application of Signomial Programming to Aircraft Design, author=Philippe G. Kirschen and Martin A. York and Berk Ozturk and Warren W. Hoburg. *Journal of Aircraft*, 55(3):965–987, 2018.

[3] Amir Ali Ahmadi and Georgia Hall. On the Construction of Converging Hierarchies for Polynomial Optimization Based on Certificates of Global Positivity. *Mathematics of Operations Research*, 44:1192–1207, 2017.

[4] Emil Artin. Uber die zerlegung deniter funktionen in quadrate. *Abhandlungen aus dem Mathematischen Seminar der Universitat Hamburg*, 5:100–115, 1927.

[5] Jonathan M. Borwein and Adrian S. Lewis. *Convex Analysis and Nonlinear Optimization*. Springer, 2nd edition, 2006.

[6] S. Boyd, S. J. Kim, D. Patil, and M. Horowitz. Digital Circuit Optimization via Geometric Programming. *Operations Research*, 53:899–932, 2005.

[7] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[8] Venkat Chandrasekaran and Parikshit Shah. Relative Entropy Relaxations for Signomial Optimization. *SIAM Journal on Optimization*, 26(2):1147–1173, 2016.

[9] Mung Chiang. Geometric Programming for Communication Systems. *Foundations and Trends in Communications and Information Theory*, 2(1-2):1–154, 2005.

[10] Mung Chiang. Nonconvex Optimization for Communication Networks. In *Advances in Applied Mathematics and Global Optimization:In Honour of Gilbert Strang*, volume 17. Springer, 2009.

[11] Mung Chiang and Stephen Boyd. Geometric Programming Duals of Channel Capacity and Rate Distortion. *IEEE Transactions on Information Theory*, 50(2):245–258, 2004.

[12] Peter J. C. Dickinson and Janez Povh. On an Extension of Polya's Positivstellensatz. *Journal of Global Optimization*, 61(4):615–625, 2014.

[13] Mareike Dressler, Sadik Ilima, and Timo de Wolf. A Positivstellensatz for Sums of Nonnegative Circuit Polynonmials. *SIAM Journal of Applied Algebra and Geometry*, 1:536–555, 2017.

[14] Joseph Ecker. Geometric Programming: Methods, Computations and Applications. *SIAM Review*, 22(3):338–362, 2018.

[15] David Hilbert. Ueber die Darstellung definiter Formen als Summe von Formenquadraten. *Mathematische Annalen*, 32:342–350, 1888.

[16] Sadik Ilimanand and Timo de Wolff. Amoebas, Nonnegative Polynomials and Sums of Squares Supported on Circuits. *Mathematical Sciences*, 3(9):461–475, 2016.

[17] Rabih A. Jabr. Inductor Design using Signomial Programming. *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 26(2):461–475, 2007.

[18] Yun KangSui and Xi ChengWang. Second-order Method of Generalized Geometric Programming for Spatial Frame Optimization. *Computer Methods in Applied Mechanics and Engineering*, 141:117–123, 1997.

[19] Jean B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11:796–817, 2001.

[20] Nathan Linial, Alex Samorodnitsky, and Avi Wigderson. Deterministic Strongly Polynomial Algorithm for Matrix Scaling and Approximate Permanents. *Combinatorica*, 20:545–568, 2000.

[21] Costas D. Maranas and Christodoulos A. Floudas. Global Optimization in Generalized Geometric Programming. *Computers and Chemical Engineering*, 21(4):351–369, 1997.

[22] Murray Marshall. *Positive Polynomials and Sums of Squares*, volume 146. American Mathematical Society, 2008.

[23] Riley Murray, Venkat Chandrasekaran, and Adam Wierman. Signomial and Polynomial Optimization via Relative Entropy and Partial Dualization. *arXiv preprint arXiv:1907.00814*, 2019.

[24] Riley Murray, Venkat Chandrasekaran, and Adam Wierman. Newton Polytopes and Relative Entropy Optimization. *arXiv preprint arXiv:1810.01614*, 2020.

[25] Max M. J. Opgenoord, Brian S. Cohen, and Warren W. Hoburg. Comparison of Algorithms for Including Equality Constraints in Signomial Programming. Technical report, MIT, August 2017.

[26] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.

[27] Elmor L. Peterson. Geometric Programming. *SIAM Review*, 18(1):1–51, 1996.

[28] George Pólya. Über positive Darstellung von Polynomen. 73:141–145, 1928.

[29] Bruce Reznick. Uniform Denominators in Hilbert's 17th Problem. *Mathematische Zeitschrift*, 220:75–97, 1995.

[30] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[31] D. H. Roundtree and A. K. Rigler. A Penalty Treatment of Equality Constraints in Generalized Geometric Programming. *Journal of Optimization Theory and Applications*, 11:169–178, 1982.

[32] Konrad Schmüdgen. The k-moment Problem for Compact Semi-algebraic Sets. *Mathematische Annalen*, 289:203–206, 1991.

[33] Marcel K. De Carli Silva and Levent Tuncel. A Notion of Total Dual Integrality for Convex, Semidefinite, and Extended Formulations. *SIAM Journal on Discrete Mathematics*, 34(1):1147–1173, 2020.

[34] Jie Wang. Nonnegative Polynomials and Circuit Polynnomials. *arXiv preprint arXiv:1804.09455*, 2020.

[35] Martin A. York, Berk Öztürk, Edward Burnell, and Warren W. Hoburg. Efficient Aircraft Multidisciplinary Design Optimization and Sensitivity Analysis via Signomial Programming. *AIAA Journal*, 56(11):4546–4561, 2018.

APPENDICES

Proof of Theorem 12. Consider the set $\Omega = \{\mathbf{x} \in \mathbb{R}_+^n : ||\mathbf{x}||_2 = 1\}$, which is compact. $f_0(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}_+^n \cap \bigcap_{i \in J} f_i^{-1}(\mathbb{R}_+) \backslash \{\mathbf{0}\}$ iff $f_0(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega \cap \mathbb{R}_+^n \cap \bigcap_{i \in J} f_i^{-1}(\mathbb{R}_+)$ since $f_0(\mathbf{x})$ is homogeneous and its positivity is invariant to the scale of $\mathbf{x}$. Thus we may restrict to to the leveled set Ω. Make the following observations:

1. Without loss of generality assume that $\deg(f_i(\mathbf{x})) \geq 1$ and $\max_{\mathbf{x} \in \Omega}\{||\nabla f_i(\mathbf{x})||\} \leq 1 \; \forall i \in I$. The second condition is not restrictive since the theorem only considers positivity of homogeneous polynomials $(f_i(\mathbf{x}))_{i \in I}$, which is invarint under scaling.

2. By mean value theorem, for any $\mathbf{x}, \mathbf{y} \in \Omega$, and any $i \in I$, there exists $\alpha \in [0, 1]$ s.t. $f_i(\mathbf{x}) - f_i(\mathbf{y}) = (\mathbf{x} - \mathbf{y})^\top \nabla f_i(\alpha \mathbf{x})$. Thus

$$\begin{aligned} ||f_i(\mathbf{x}) - f_i(\mathbf{y})||_2 &= ||(\mathbf{x} - \mathbf{y})^\top \nabla f_i(\alpha \mathbf{x})||_2 \\ &\leq ||\mathbf{x} - \mathbf{y}||_2 ||\nabla f_i(\alpha \mathbf{x})||_2 \\ &\leq ||\mathbf{x} - \mathbf{y}||_2 \max_{\mathbf{x} \in \Omega}\{||\nabla f_i(\mathbf{x})||\} \\ &\leq ||\mathbf{x} - \mathbf{y}||_2 \; \forall i \in I \end{aligned}$$

   This implies $\forall i \in I$, $f_i(\mathbf{x})$ is a continuous function.

3. $\forall i \in I$, $||f_i(\mathbf{x})||_2 = ||f_i(\mathbf{x}) - f_i(\mathbf{0})||_2 \leq ||\mathbf{x}||_2 \leq 1 \; \forall \mathbf{x} \in \Omega$.

Now define the following compact sets:

- $\Omega_0 = \Omega \cap f_0^{-1}(-\mathbb{R}_+)$
- $\Omega_j = \Omega \cap f_j^{-1}(\mathbb{R}_+)$

- $\Omega_J = \Omega \cap \bigcap_{i \in J} f_i^{-1}(\mathbb{R}_+) = \bigcap_{i \in J} \Omega_i \quad \forall J \subseteq I$

Observe that $f_0(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^n_+ \cap \bigcap_{i \in I} f_i^{-1}(\mathbb{R}_+) \setminus \{\mathbf{0}\}$ iff $\Omega_0 \cap \Omega_I = 0$. The goal is to show that $\exists J \subseteq I \ \Omega_0 \cap \Omega_J = \emptyset$, which implies then $f_0(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^n_+ \cap \bigcap_{i \in J} f_i^{-1}(\mathbb{R}_+) \setminus \{\mathbf{0}\}$.

Consider the function: $\xi(\mathbf{x}) = \sup\{-f_i(\mathbf{x}) \mid i \in I\} \ \forall \mathbf{x} \in \Omega$. By observation (2), $\zeta(\mathbf{x})$ is a supremum of continuous function and is thus continuous. $\Omega_0 \cap \Omega_I = 0$ by assumption, so $\xi(\mathbf{x}) > 0 \ \forall \mathbf{x} \in \Omega_0$. Moreover, by observation (3), $\xi(\mathbf{x}) \in (0, 1] \ \forall \mathbf{x} \in \Omega_0$. Let $\epsilon = \min_{\mathbf{x} \in \Omega_0} \xi(\mathbf{x})$. Since Ω_0 is compact, and $\xi(\mathbf{x})$ is continuous, by the extreme value theorem, the min is attained in Ω_0. So $\epsilon \in (0, 1]$.

Now consider the following two facts. (a) any $\mathbf{x} \in \Omega_0$, there exists some $i \in I$ s.t. $-f_i(\mathbf{x}) \geq \frac{2}{3}\xi(\mathbf{x}) \geq \frac{2}{3}\epsilon > 0$. (b) for any $\mathbf{y} \in \Omega_0$ s.t. $||\mathbf{x} - \mathbf{y}|| \leq \frac{1}{3}\epsilon$, $f_i(\mathbf{y}) \leq f_i(\mathbf{x}) + ||\mathbf{x} - \mathbf{y}||_2 \leq -\frac{2}{3}\epsilon + \frac{1}{3}\epsilon < 0$. So $\mathbf{y} \notin \Omega_i$. Now consider the algorithm below.

Algorithm 1 Finding $J \subseteq I$ s.t. $\Omega_0 \cap \Omega_J = \emptyset$

1: Let $J = \emptyset$
2: **while** $\exists \mathbf{z} \in \Omega_0 \cap \Omega_J$ **do**
3: If for some $i \in I$, $f_i(\mathbf{z}) \leq -\frac{2}{3}\epsilon$, then $J = J \cup \{i\}$
4: Return J

Suppose $\mathbf{z}_t$ is chosen at t_{th} iteration of while loop. First, $\exists i \in I$ to add to J, by claim (a). Since $f_i(\mathbf{z}_t) \leq -\frac{2}{3} < 0$, by claim (b), for any $\mathbf{z}_{t+1}$ s.t. $||\mathbf{z}_{t+1} - \mathbf{z}_t|| \leq \frac{1}{3}\epsilon$, $\mathbf{z}_{t+1} \notin \Omega_i \implies \mathbf{z}_{t+1} \notin \Omega_{J \cup i}$. Thus in each iteration, $\mathbf{z}_t$ has a distance of at least $\frac{1}{3}\epsilon$ from the previous ones. Since Ω_0 is a compact set, the algorithm terminates in finite time, which implies $\Omega_0 \cap \Omega_J = \emptyset$. The desired $J \subseteq I$ is obtained.