Semidefinite approximations of the polynomial abscissa[0]

Roxana Heß[1] Didier Henrion[1,2]
Jean-Bernard Lasserre[1,3] Tiến-Sơn Phạm[4]

Draft of March 18, 2016

Abstract

Given a univariate polynomial, its abscissa is the maximum real part of its roots. The abscissa arises naturally when controlling linear differential equations. As a function of the polynomial coefficients, the abscissa is Hölder continuous, and not locally Lipschitz in general, which is a source of numerical difficulties for designing and optimizing control laws. In this paper we propose simple approximations of the abscissa given by polynomials of fixed degree, and hence controlled complexity. Our approximations are computed by a hierarchy of finite-dimensional convex semidefinite programming problems. When their degree tends to infinity, the polynomial approximations converge in L^1 norm to the abscissa, either from above or from below.

Keywords

Linear systems control, non-convex non-smooth optimization, polynomial approximations, semialgebraic optimization, semidefinite programming.

AMS classification

26C10, 41A10, 90C22, 90C26.

---

[0]A part of this work was done while the fourth author was visiting LAAS-CNRS in April 2015. He would like to thank LAAS-CNRS and J. B. Lasserre for the hospitality and support during his stay.

[1]LAAS-CNRS, Université de Toulouse, CNRS, France.

[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic

[3]Institut de Mathématiques de Toulouse, Université de Toulouse, France.

[4]Department of Mathematics, University of Dalat, 1 Phu Dong Thien Vuong, Dalat, Vietnam.

1 Introduction

Given a univariate polynomial p whose coefficients depend polynomially on parameters, its abscissa maps the parameter vector to the maximum real part of the roots of p for this parameter. When studying linear differential equations, the abscissa of the characteristic polynomial of the equation is used as a measure of the decay or growth rate of the solution. In linear systems control, the abscissa function is typically parametrized by a small number of real parameters (the controller coefficients), and it should be minimized so as to ensure a sufficiently fast decay rate of closed-loop trajectories.

As a function of the polynomial coefficients (expressed in some basis), the abscissa is a Hölder continuous function (with exponent equal to the reciprocal of the polynomial degree), but it is not locally Lipschitz. As a consequence of this low regularity, numerical optimization of the polynomial abscissa is typically a challenge.

For a recent survey on the abscissa function and its applications in systems control, see [6]. A detailed variational analysis of the abscissa was first carried out in [5]. These ideas were exploited in a systems control setup in [4], using randomized techniques of non-convex non-smooth local optimization, however without rigourous convergence guarantees.

In the space of controller parameters, the zero sublevel set of the abscissa function of the characteristic polynomial of a linear system is the so-called stabilizability region, and it is typically non-convex and non-smooth, see [8] where this set is approximated with simpler sets such as balls or ellipsoids. In [7], ellipsoidal approximations of the stabilizability region were generalized to polynomial sublevel set approximations, obtained by replacing negativity of the abscissa function with positive definiteness of the Hermite matrix of the characteristic polynomial.

This paper continues the research efforts of [8] and [7], in the sense that we would like to approximate the complicated geometry of the abscissa function (and its sublevel sets) with a simpler function, namely a low degree polynomial. The level of complexity of the approximation is the degree of the polynomial, to be fixed in advance. Moreover, we would like the quality of the approximation to improve when the degree increases, eventually converging (in some appropriate sense) to the original abscissa function when the degree tends to infinity.

The outline of the paper is as follows. After introducing in Section 2 the abscissa function and some relevant notations, we address in Section 3 the problem of finding an upper approximation of the abscissa. In Section 4, we address the more difficult problem of approximating the abscissa from below, first by using elementary symmetric functions, and second by using the Gauß-Lucas theorem, inspired by [5]. Explicit numerical examples illustrate our findings throughout the text.

2 Preliminaries

Notation and definitions

Let $n \in \mathbb{N}$ and $\mathcal{Q} \subseteq \mathbb{R}^n$ be a compact semi-algebraic set on which a Borel measure with support $\mathcal{Q}$ can be defined and whose moments are easy to compute. For simplicity, in this paper we choose $\mathcal{Q} = [-1, 1]^n = \{q \in \mathbb{R}^n \;:\; 1 - q_1^2 \geq 0, \ldots, 1 - q_n^2 \geq 0\}$.

Let $\mathscr{C}(\mathcal{Q})$ denote the space of continuous functions on $\mathcal{Q}$. Its topological dual is isometrically isomorphic to the vector space $\mathscr{M}(\mathcal{Q})$ of signed Borel measures on $\mathcal{Q}$. By Banach-Alaoglu's theorem [1, 2], the unit ball of $\mathscr{M}(\mathcal{Q})$ is compact (and sequentially compact) in the weak-star topology of $\mathscr{M}(\mathcal{Q})$.

Denote by $\mathbb{R}[q]$ the vector space of real polynomials in the variables $q = (q_1, \ldots, q_n)$ and define $\mathbb{R}[q]_d := \{p \in \mathbb{R}[q] : \deg p \leq d\}$ where $\deg p$ denotes the degree p. Let $\Sigma[q] \subset \mathbb{R}[q]$ be the convex cone of real polynomials that are sums of squares of polynomials and denote by $\Sigma[q]_{2d}$ its subcone of sums of squares of polynomials of degree at most $2d$.

The abscissa function

Consider the monic non-constant polynomial $p \in \mathbb{R}[s]$ defined by

$$p: \; s \mapsto p(q, s) := \sum_{k=0}^{m} p_k(q) s^k$$

with $s \in \mathbb{C}$ complex, $q = (q_1, \ldots, q_n) \in \mathcal{Q}$ and given polynomials $p_k \in \mathbb{R}[q]$ for $k = 0, 1, \ldots, m$ with $p_m(q) \equiv 1$ and $m > 0$. Hence, we have a polynomial whose coefficients depend polynomially on the parameter q.

Denote by $s_r(q),\; r = 1, \ldots, m$, the roots of $p(q, \cdot)$ and by $a\colon \mathcal{Q} \to \mathbb{R}$ (or a_p if it is necessary to clarify the dependence on the polynomial) the abscissa map of p, i.e. the maximal real part of the roots:

$$a(q) := \max_{r=1,\ldots,m} \Re(s_r(q)), \;\; q \in \mathcal{Q}.$$

Equivalently, with $i = \sqrt{-1}$ and $s = x + iy$ write

$$p(q, s) = p_{\Re}(q, x, y) + i p_{\Im}(q, x, y)$$

for two real polynomials $p_{\Re}, p_{\Im} \in \mathbb{R}[q, x, y]$ of total degree m. Then

$$a(q) = \max\{x \in \mathbb{R} : \exists y \in \mathbb{R} : \; p_{\Re}(q, x, y) = p_{\Im}(q, x, y) = 0\}, \;\; q \in \mathcal{Q}.$$

We observe that function $a\colon \mathcal{Q} \to \mathbb{R}$ is semi-algebraic and we define the basic closed semi-algebraic set

$$\mathcal{Z} := \{(q, x, y) \in \mathbb{R}^n \times \mathbb{R}^2 : q \in \mathcal{Q}, \; p_{\Re}(q, x, y) = p_{\Im}(q, x, y) = 0\}.$$

Remark. Set $\mathcal{Z}$ is compact, since $\mathcal{Q}$ is compact and p is monic in s.

Now we can write the abscissa map as

$$a(q) = \max\{x \in \mathbb{R} : \exists y \in \mathbb{R} : \ (q, x, y) \in \mathcal{Z}\}, \ \ q \in \mathcal{Q}.$$

Since p is monic, its abscissa a is continuous, though in general not Lipschitz continuous. For example, for $n = 1$ and $p(q, s) = s^6 + q$ the map $a(q)$ is only Hölder continuous with exponent $\frac{1}{6}$ for small q. To be precise, a is always Hölder continuous by the Łojasiewicz inequality [3], since $\mathcal{Q}$ is compact.

Stability regions for linear systems

For continuous-time dynamical systems described by linear differential equations, stability analysis amounts to studying the location of the roots of the characteristic polynomial obtained by applying the Laplace transform. A polynomial is then called stable if all its roots lie in the open left part of the complex plane, i.e. if its abscissa is negative. In systems control, the characteristic polynomial depend on parameters, which are typically controller coefficients that must be chosen so that the polynomial is stable. For a polynomial with parameterized coefficients, as we consider in this paper, the stability region is then the set of parameters for which the abscissa is negative, that is the zero sublevel set of the abscissa, in our notation $\{q \in \mathcal{Q} : a(q) < 0\}$.

3 Upper abscissa approximation

3.1 Primal and dual formulation

Given a polynomial p defined as above, any function v admissible for the following infinite-dimensional linear programming (LP) problem gives an upper approximation of the abscissa function a_p on $\mathcal{Q}$:

$$\begin{aligned} \rho = \inf_{v \in \mathscr{C}(\mathcal{Q})} & \int_{\mathcal{Q}} v(q)\, dq \\ \text{s.t.} \quad & v(q) - x \geq 0 \text{ for all } (q, x, y) \in \mathcal{Z} \end{aligned} \tag{1}$$

with $\mathscr{C}(\mathcal{Q})$ denoting the space of continuous functions from $\mathcal{Q}$ to $\mathbb{R}$. Note that the value of ρ itself is not an upper bound on the abscissa, in general. The upper approximation is the function $q \mapsto v(q)$ solving LP (1), and it holds $v(q) \geq a(q)$ for all $q \in \mathcal{Q}$.

Remark 3.1. Since the continuous functions defined on compact set $\mathcal{Q}$ can be approximated uniformly by polynomials by the Stone-Weierstraß theorem [13, §16.4.3], we can replace $\mathscr{C}(\mathcal{Q})$ in problem (1) by the ring of polynomials $\mathbb{R}[q]$.

The LP dual to problem (1) can be constructed as described in [2, Chapter IV] and reads

$$\rho^* = \sup_{\mu \in \mathscr{M}^+(\mathcal{Z})} \int_{\mathcal{Z}} x \, d\mu(q,x,y) \tag{2}$$
$$\text{s.t.} \int_{\mathcal{Z}} q^\alpha \, d\mu = \int_{\mathcal{Q}} q^\alpha \, dq, \quad \text{for all } \alpha \in \mathbb{N}^n,$$

where q^α stands for the monomial $q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_n^{\alpha_n}$ and $\mathscr{M}^+(\mathcal{Z})$ denotes the cone of non-negative Borel measures supported on $\mathcal{Z}$, identified with the set of all non-negative continuous linear functionals acting on $\mathscr{C}^+(\mathcal{Z})$, the cone of non-negative continuous functions supported on $\mathcal{Z}$.

Remark 3.2. The constraint $\int_{\mathcal{Z}} q^\alpha \, d\mu = \int_{\mathcal{Q}} q^\alpha \, dq$ for all $\alpha \in \mathbb{N}^n$ implies that the marginal of μ on $\mathcal{Q}$ is the Lebesgue measure on $\mathcal{Q}$, i.e. for every $g \in \mathscr{C}(\mathcal{Q})$ it holds that

$$\int_{\mathcal{Z}} g(q) \, d\mu(q,x,y) = \int_{\mathcal{Q}} g(q) \, dq.$$

In particular this implies that $\|\mu\| = \text{vol}\, \mathcal{Q}$ where $\text{vol}(\cdot)$ denotes the volume or Lebesgue measure.

Lemma 3.1. *The supremum in LP* (2) *is attained, and there is no duality gap between LP (1) and LP (2), i.e.* $\rho = \rho^*$.

Proof. The set of feasible solutions for the dual LP (2) is a bounded subset of $\mathscr{M}^+(\mathcal{Z})$ with $\mathcal{Z}$ compact and therefore it is weak-star compact. Since the objective function is linear, its supremum on this weak-star compact set is attained. For elementary background on weak-star topology, see e.g. [2, Chapter IV].

To prove that there is no duality gap, we apply [2, Theorem IV.7.2]. For this purpose we introduce the notation used in [2] in this context. There, the primal and the dual are written in the following canonical form:

$$\rho^* = \sup_{\mathbf{x} \in E_1} \langle \mathbf{x}, \mathbf{c} \rangle_1 \qquad\qquad \rho = \inf_{\mathbf{y} \in F_2} \langle \mathbf{b}, \mathbf{y} \rangle_2$$
$$\text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in E_1^+ \qquad\qquad \text{s.t. } \mathbf{A}^*\mathbf{y} - \mathbf{c} \in F_1^+$$

So we set $E_1 := \mathscr{M}(\mathcal{Z})$ with its cone $E_1^+ := \mathscr{M}^+(\mathcal{Z})$. Then their (pre-)duals are $F_1 := \mathscr{C}(\mathcal{Z})$ and $F_1^+ := \mathscr{C}^+(\mathcal{Z})$ respectively. Similarly, we define $E_2 := \mathscr{M}(\mathcal{Q})$ and $F_2 := \mathscr{C}(\mathcal{Q})$.

Setting $\mathbf{x} := \mu \in E_1$, $\mathbf{c} := x \in F_1$, $\mathbf{b} \in E_2$ the Lebesgue measure on $\mathcal{Q}$ and $\mathbf{y} := v \in F_2$, the linear operator $\mathbf{A}\colon E_1 \to E_2$ is given by $\mathbf{x} \mapsto \pi_{\mathcal{Q}} \mathbf{x}$ where $\pi_{\mathcal{Q}}$ denotes the projection onto $\mathcal{Q}$, i.e., $\mathbf{A}\mathbf{x}(B) = \mathbf{x}(B \times \mathbb{R}^2)$ for all $B \in \mathcal{B}(\mathcal{Q})$.

According to [2, Theorem IV.7.2] the duality gap is zero if the cone $\{(\mathbf{A}\mathbf{x}, \langle \mathbf{x}, \mathbf{c} \rangle_1) : \mathbf{x} \in E_1^+\}$ is closed in $E_2 \times \mathbb{R}$. This holds in our setup since $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ and $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{c} \rangle_1$ are continuous linear maps and $E_1^+ = \mathscr{M}^+(\mathcal{Z})$ is weak-star closed due to the compactness of $\mathcal{Z}$. So if for some $\mathbf{a} \in E_2$, $\mathbf{A}\mathbf{x_n} \to \mathbf{a}$ as $n \to \infty$ then from the definition of the mapping $\mathbf{A}$ and as $(\mathbf{x_n}) \subset E_1^+$, one has $\|\mathbf{x_n}\| \to \|\mathbf{a}\|$ as $n \to \infty$ (see Remark 3.2). Therefore the sequence $(\mathbf{x_n}) \subset E_1^+$ is bounded and by Banach-Alaoglu's theorem [1, 2], it contains a subsequence $(\mathbf{x_{n_k}}) \subset E_1^+$ that converges to some $\mathbf{x} \in E_1^+$ for the weak-star topology. By continuity of the mappings $\mathbf{A}$ and $\mathbf{c}$, the result follows. □

Remark 3.3. The infimum in LP (1) is not necessarily attained, since the set of feasible solutions is not compact. The reason for this is that we minimize over the L^1 norm of v which implies that the limit of an optimizing sequence $(v_l)_{l\in\mathbb{N}}$ does not necessarily need to be continuous, because on sets with Lebesgue measure zero it does not need to match the abscissa. Nor is the infimum attained when we replace $\mathscr{C}(\mathcal{Q})$ with $\mathbb{R}[q]$, since a is non-Lipschitz, so in particular not polynomial.

However, the infimum is attained if we replace $\mathscr{C}(\mathcal{Q})$ with $\mathbb{R}[q]_d$ for d finite. Then, with $M := \min_{q\in\mathcal{Q}} a(q) > -\infty$ and $\tilde{v}(q) := v(q) - M$ we can rewrite LP (1) as the equivalent problem

$$\inf_{\tilde{v}\in\mathbb{R}[q]_d} \int_{\mathcal{Q}} \tilde{v}(q)dq \quad \text{s.t. } \tilde{v}(q) + M - x \geq 0 \text{ on } \mathcal{Z}.$$

Now, any feasible $\tilde{v}$ is non-negative on $\mathcal{Q}$, so $\int_{\mathcal{Q}} \tilde{v}(q)dq = \|\tilde{v}\|_{L^1} \geq 0$ is a norm because $\mathcal{Q}$ has non-empty interior by assumption, and for every $R \in \mathbb{R}$ the set $\{\tilde{v} \in \mathbb{R}[q]_d : R \geq \int_{\mathcal{Q}} \tilde{v}(q)dq$ and $\tilde{v}(q) + M - x \geq 0$ on $\mathcal{Z}\}$ is closed and bounded in the strong topology, thus compact. Besides, due to the continuity of a, there always exists an $R < \infty$ such that the mentioned set is not empty, hence the infimum is attained.

3.2 SDP hierarchy

Let $d_0 \in \mathbb{N}$ be sufficiently large. As presented in [9], we can write a hierarchy of finite-dimensional convex semidefinite programming (SDP) problems for LP (1) indexed by the parameter $d \in \mathbb{N},\ d \geq d_0$:

$$\begin{aligned} \rho_d = \inf_{v_d,\sigma_0,\sigma_j,\tau_{\Re},\tau_{\Im}} & \int_{\mathcal{Q}} v_d(q)\, dq \\ \text{s.t. } v_d(q) - x = {} & \sigma_0(q,x,y) + \sum_{j=1}^{n} \sigma_j(q,x,y)(1-q_j^2) \\ & + \tau_{\Re}(q,x,y)p_{\Re}(q,x,y) + \tau_{\Im}(q,x,y)p_{\Im}(q,x,y) \end{aligned} \tag{3}$$

for all $(q,x,y) \in \mathbb{R}^n \times \mathbb{R}^2$ and with $v_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q,x,y]_{2d}$, $\sigma_j \in \Sigma[q,x,y]_{2d-2}$ for $j = 1,\dots,n$ and $\tau_{\Re}, \tau_{\Im} \in \mathbb{R}[q,x,y]_{2d-m}$.

Remark 3.4. The quadratic module generated by the polynomials $1-q_1^2,\dots,1-q_n^2$, $\pm p_{\Re}$, $\pm p_{\Im}$ is archimedean by [10, Lemma 3.17], since it contains the polynomial $f(q,x,y) := \sum_{j=1}^{n}(1-q_j^2) - p_{\Re}^2 - p_{\Im}^2$ and the set $\{(q,x,y) \in \mathbb{R}^n \times \mathbb{R}^2 : f(q,x,y) \geq 0\}$ is compact. By [9, Theorem 4.1], this implies that the hierarchy converges, i.e. $\lim_{d\to\infty} \rho_d = \rho$.

Remark 3.5. Note that SDP (3) is not equivalent to LP (1), not even with $\mathscr{C}(\mathcal{Q})$ replaced by $\mathbb{R}[q]$ or $\mathbb{R}[q]_{2d}$ in the latter, but it is a strengthening of it, meaning $\rho_d \geq \rho$. To be more specific, we exchanged non-negativity for a specific certificate of positivity. See [9, Chapter 4.2] for details.

Example 3.1. The infimum in SDP (3) is not necessarily attained, e.g. consider the polynomial $p(q,s) = s^2$. Then $p_{\Re}(q,x,y) = x^2 - y^2$, $p_{\Im}(q,x,y) = 2xy$ and $\mathcal{Z} = \mathcal{Q} \times \{(0,0)\}$. Obviously, the optimal solution to LP (1) is $v \equiv 0$. For SDP (3) we would want

$$v(q) - x = \sigma_0(q,x,y) + \sigma_1(q,x,y)(1-q^2) + \tau_{\Re}(q,x,y)(x^2-y^2) + 2\tau_{\Im}(q,x,y)xy,$$

meaning $0 \equiv v = x + \sigma_0 + \sigma_1(1-q^2) + \tau_{\Re}x^2 - \tau_{\Re}y^2 + 2\tau_{\Im}xy$ with σ_0, σ_1 sums of squares. This is impossible, since it would require the construction of the term $-x$ which in this case is only possible as a summand of σ_0. Then however we would always also produce a constant positive term. Practically this means that the multipliers $\sigma_0, \sigma_1, \tau_{\Re}, \tau_{\Im}$ blow up.

Hence, an optimal solution might not exist, but we always have a near optimal solution. This means we should allow solutions v_d with $\int_{\mathcal{Q}} v_d(q)\,dq \le \rho_d + \frac{1}{d}$, e.g. in the above example we would search for $v \equiv \varepsilon$ for an $\varepsilon > 0$ sufficiently small.

Remark 3.6. The existence of an optimal solution depends on further conditions, such as the ideal generated by the polynomials $1 - q_j^2$, $p_{\Re}$ and $p_{\Im}$ being radical, and goes beyond the scope of this paper. The interested reader is referred to the proof of [7, Lemma 1] for further details.

In the following theorem we prove that the associated sequence of solutions converges:

Theorem 3.2. *Let $v_d \in \mathbb{R}[q]_{2d}$ be near optimal solutions for SDP* (3)*, i.e. $\int_{\mathcal{Q}} v_d(q)\,dq \le \rho_d + \frac{1}{d}$, and consider the associated sequence $(v_d)_{d \ge d_0} \subset L^1(\mathcal{Q})$. Then v_d converges to the abscissa a in L^1 norm on $\mathcal{Q}$ as d tends to infinity, i.e.* $\lim_{d\to\infty} \int_{\mathcal{Q}} |v(q) - a(q)|\,dq = 0$.

Proof. Recall that $\rho^* = \rho$ according to Lemma 3.1. First we show that $\rho = \int_{\mathcal{Q}} a(q)\,dq$. For every $(q,x,y) \in \mathcal{Z}$ we have $x \le a(q)$ and since $\int_{\mathcal{Z}} q^\alpha d\mu = \int_{\mathcal{Q}} q^\alpha dq$ for all $\alpha \in \mathbb{N}^n$ which means that the marginal of μ on $\mathcal{Q}$ is the Lebesgue measure on $\mathcal{Q}$ (see Remark 3.2), it follows that for every feasible solution $\mu \in \mathscr{M}_+(\mathcal{Z})$

$$\int_{\mathcal{Z}} x \; d\mu(q,x,y) \le \int_{\mathcal{Z}} a(q)\,d\mu(q,x,y) = \int_{\mathcal{Q}} a(q)\,dq.$$

Hence $\rho \le \int_{\mathcal{Q}} a(q)\,dq$. On the other hand, for every $q \in \mathcal{Q}$ there exists $(q, x_q, y_q) \in \mathcal{Z}$ such that $a(q) = x_q$. Let μ^* be the Borel measure concentrated on (q, x_q, y_q) for all $q \in \mathcal{Q}$, i.e. for $\mathcal{A}$ in the Borel sigma algebra of $\mathcal{Z}$ it holds

$$\mu^*(\mathcal{A}) := \mathbf{1}_{\mathcal{A}}(q, x_q, y_q).$$

Then μ^* is feasible for problem (2) with value

$$\int_{\mathcal{Z}} x \; d\mu^*(q,x,y) = \int_{\mathcal{Q}} a(q)\,dq,$$

which proves that $\rho \ge \int_{\mathcal{Q}} a(q)\,dq$, hence $\rho = \int_{\mathcal{Q}} a(q)\,dq$.

Next we show convergence in L^1. Since the abscissa a is continuous on the compact set $\mathcal{Q}$, by the Stone-Weierstraß theorem [13, §16.4.3] it holds that for every $\varepsilon > 0$ there exists a polynomial $h_\varepsilon \in \mathbb{R}[q]$ such that

$$\sup_{q\in\mathcal{Q}} |h_\varepsilon(q) - a(q)| < \frac{\varepsilon}{2}.$$

Hence, the polynomial $v_\varepsilon := h_\varepsilon + \varepsilon$ satisfies $v_\varepsilon - a > 0$ on $\mathcal{Q}$ and we have $v_\varepsilon(q) - x > 0$ on $\mathcal{Z}$. Since the corresponding quadratic module is archimedean (see Remark 3.4), by Putinar's

Positivstellensatz [9, Theorem 2.5] there exist $\sigma_0^\varepsilon, \sigma_j^\varepsilon \in \Sigma[q,x,y]$, $\tau_\Re^\varepsilon, \tau_\Im^\varepsilon \in \mathbb{R}[q,x,y]$ such that for all $(q,x,y) \in \mathbb{R}^n \times \mathbb{R}^2$ we can write

$$v_\varepsilon(q) - x = \sigma_0^\varepsilon(q,x,y) + \sum_{j=1}^n \sigma_j^\varepsilon(q,x,y)(1-q_j^2) \\ + \tau_\Re^\varepsilon(q,x,y)p_\Re(q,x,y) + \tau_\Im^\varepsilon(q,x,y)p_\Im(q,x,y).$$

Therefore, for $d \geq d_\varepsilon := \lceil \frac{\deg v_\varepsilon}{2} \rceil$ the tuple $(v_\varepsilon, \sigma_0^\varepsilon, \sigma_j^\varepsilon, \tau_\Re^\varepsilon, \tau_\Im^\varepsilon)$ is a feasible solution for SDP (3) satisfying

$$0 \leq \int_\mathcal{Q} (v_\varepsilon(q) - a(q))\, dq \leq \frac{3\varepsilon}{2} \int_\mathcal{Q} dq.$$

Together with $\int_\mathcal{Q} a(q)\, dq = \rho \leq \rho_d$ which is due to the first part of the proof and ρ_d being a strengthening of ρ, it follows that whenever $d \geq d_\varepsilon$

$$0 \leq \rho_d - \int_\mathcal{Q} a(q)\, dq \leq \int_\mathcal{Q} (v_\varepsilon(q) - a(q))\, dq \leq \frac{3\varepsilon}{2} \int_\mathcal{Q} dq.$$

As $\varepsilon > 0$ was arbitrary, we obtain $\lim_{d\to\infty} \rho_d = \int_\mathcal{Q} a(q)\, dq$ and since $a \leq v_d$ for all d, this is the same as convergence in L^1:

$$0 \leq \lim_{d\to\infty} \|v_d - a\|_1 = \lim_{d\to\infty} \int_\mathcal{Q} |v_d(q) - a(q)|\, dq \\ = \lim_{d\to\infty} \int_\mathcal{Q} (v_d(q) - a(q))\, dq \leq \lim_{d\to\infty} \left(\rho_d + \frac{1}{d}\right) - \int_\mathcal{Q} a(q)\, dq = 0.$$

□

As mentioned in the Preliminaries, the abscissa function is important when studying stability of linear differential equations, and the zero sublevel set of the abscissa of the characteristic polynomial $\{q \in \mathcal{Q} : a(q) < 0\}$ is called the stability region. The following statement on polynomial inner approximations of this set follows immediately from the L^1 convergence result of Theorem 3.2, see also [7].

Corollary 3.3. *Let $v_d \in \mathbb{R}[q]_{2d}$ denote, as in Theorem 3.2, a near optimal solution for SDP (3). Then $\{q \in \mathcal{Q} : v_d(q) < 0\} \subset \{q \in \mathcal{Q} : a(q) < 0\}$ and $\lim_{d\to\infty} \operatorname{vol}\{q \in \mathcal{Q} : v_d(q) < 0\} = \operatorname{vol}\{q \in \mathcal{Q} : a(q) < 0\}$.*

3.3 Examples

While approximating the abscissa function from above we also get an inner approximation of the stability region. The authors of [7] surveyed a different approach. They described the stability region via the eigenvalues of the Hermite matrix of the polynomial and approximated it using an SDP hierarchy.

As stated in Corollary 3.3, while approximating the abscissa function from above we also get an inner approximation of the stability region. The authors of [7] surveyed a different approach. They described the stability region via the eigenvalues of the Hermite matrix of the polynomial and approximated it using an SDP hierarchy. In the following examples we compare the two different methods and highlight the specific advantages of our abscissa approximation.

In this section and in the remainder of the paper, all examples are modelled by Yalmip [12] and solved by Mosek 7 [11] under the Matlab R2014a environment, unless indicated otherwise. We ran the experiments on an HP EliteBook 840 G1 with 16GB RAM memory and an Intel® Core™ i5-4300U processor under a Windows 7 Professional 64-bit operating system.

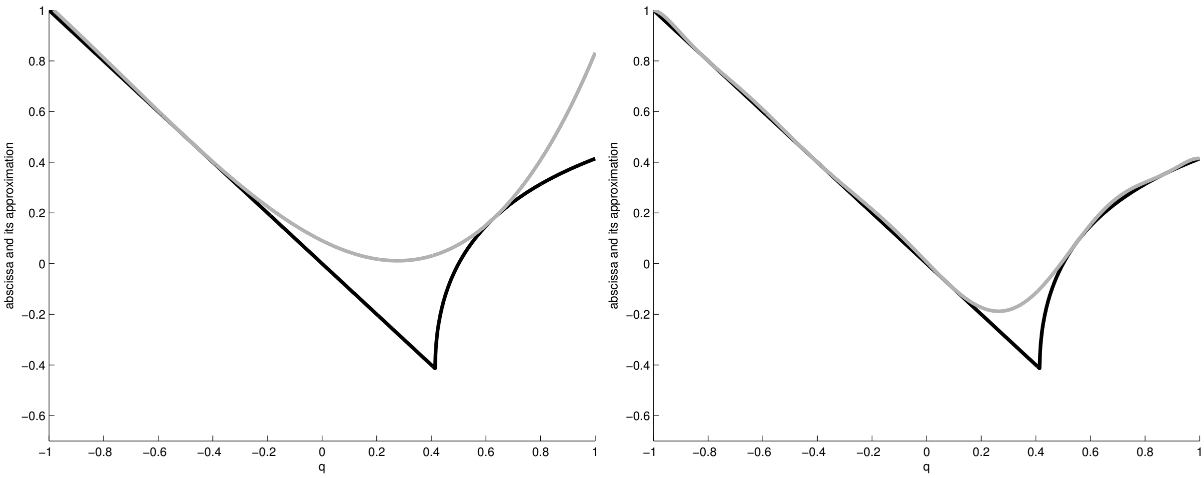

Figure 3.1: Abscissa $a(q)$ (black) and its polynomial upper approximations $v_d(q)$ of degree $d = 4$ (left, gray) and $d = 10$ (right, gray) for Example 3.2. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz.

Example 3.2 (The damped oscillator [6])**.** Consider the second degree polynomial depending on $n = 1$ parameter $q \in \mathcal{Q} = [-1, 1]$:

$$p: \ s \mapsto p(q, s) = s^2 + 2qs + 1 - 2q.$$

Then $\mathcal{Z} = \{(q, x, y) \in [-1, 1] \times \mathbb{R}^2 : x^2 - y^2 + 2qx + 1 - 2q = 2xy + 2qy = 0\}$ and the corresponding hierarchy of SDP problems (3) reads

$$\begin{aligned} \rho_d = \inf_{v_d, \sigma_0, \sigma_1, \tau_{\Re}, \tau_{\Im}} & \int_{-1}^{1} v_d(q)\, dq \\ \text{s.t. } & v_d(q) - x = \sigma_0(q, x, y) + \sigma_1(q, x, y)(1 - q^2) \\ & \qquad + \tau_{\Re}(q, x, y)(x^2 - y^2 + 2qx + 1 - 2q) + \tau_{\Im}(q, x, y)(2xy + 2qy) \end{aligned}$$

for all $(q, x, y) \in \mathbb{R}^3$ and with $v_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, x, y]_{2d}$, $\sigma_1 \in \Sigma[q, x, y]_{2d-2}$ and $\tau_{\Re}, \tau_{\Im} \in \mathbb{R}[q, x, y]_{2d-2}$. Apart from that, we only need the moments of the Lebesgue measure on $[-1, 1]$ for a successful implementation. These are readily given by

$$z_\alpha = \int_{-1}^{1} q^\alpha \, dq = \frac{1 - (-1)^{\alpha+1}}{\alpha + 1},$$

meaning that $\int_{-1}^{1} v_d(q)\,dq = \sum_{\alpha=1}^{d} v_{d\alpha} z_\alpha$ with $v_{d\alpha}$ denoting the coefficient of the monomial q^α of v_d. See Figure 3.1 for the graphs of the degree 4 and 10 polynomial upper approximations of the abscissa.

For the Hermite approximation we compute the Hermite matrix H of p (see [8] for details)

$$H(q) = \begin{pmatrix} 4q - 8q^2 & 0 \\ 0 & 4q \end{pmatrix}$$

and write the hierarchy of optimization problems as presented in [7]:

$$\begin{aligned} &\max_{g_d,\sigma_0,\sigma_1,\tau} \int_{-1}^{1} g_d(q)\,dq \\ &\text{s.t. } u^T H(q) u - g_d(q) = \sigma_0(q,u) + \sigma_1(q,u)(1-q^2) + \tau(q,u)(1-u^T u) \end{aligned}$$

for all $(q,u) \in [-1,1] \times \mathbb{R}^2$ and with $g_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q,u]_{2d}$, $\sigma_1 \in \Sigma[q,u]_{2d-2}$ and $\tau \in \mathbb{R}[q,u]_{2d-2}$. Already for $d = 6$ we observe a close match between the genuine stability region, which is $\{q \in [-1,1] : a(q) < 0\} = (0, \frac{1}{2})$, the Hermite inner approximation $\{q \in [-1,1] : -g_6(q) < 0\}$, and the polynomial upper approximation $\{q \in [-1,1] : v_{10}(q) < 0\}$. These three intervals are visually indistinguishable, so we do not represent them graphically.

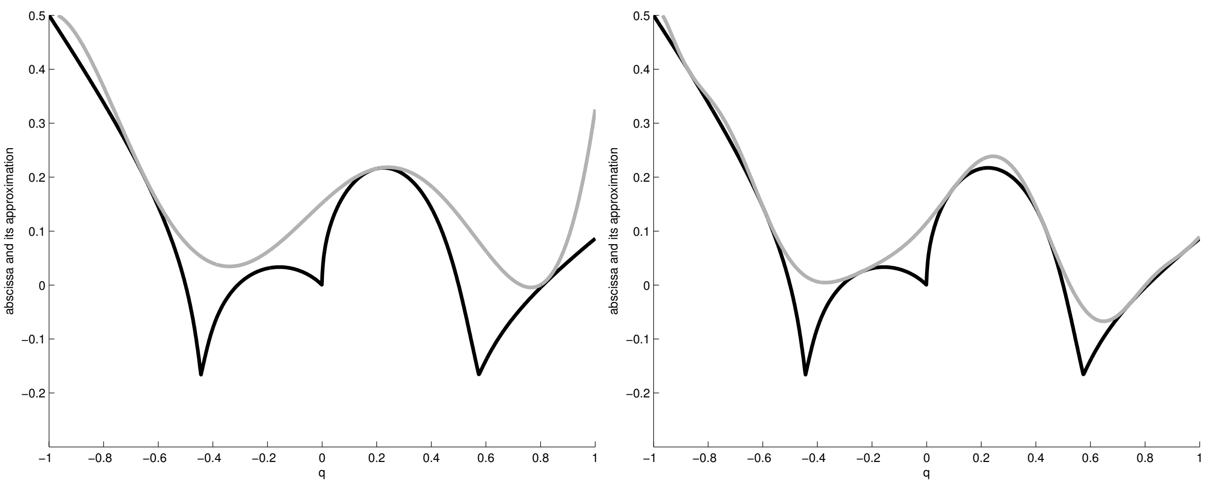

Figure 3.2: Abscissa $a(q)$ (black) and its polynomial upper approximations $v_d(q)$ of degree $d = 6$ (gray, left) and $d = 12$ (gray, right) for Example 3.3. The quality of the approximation deteriorates near the points of non-differentiability of the abscissa.

Example 3.3. Consider the polynomial

$$p: \; s \mapsto p(q,s) = s^3 + \tfrac{1}{2}s^2 + q^2 s + (q - \tfrac{1}{2})q(q + \tfrac{1}{2})$$

for $q \in \mathcal{Q} = [-1,1]$. The abscissa function $a(q)$ of p is not differentiable at three points and therefore it is rather hard to approximate in their neighborhoods. In Figure 3.2 we see the abscissa and its polynomial upper approximations of degrees 6 and 12. Comparing the genuine stability region $\{q \in [-1,1] : a(q) < 0\}$, the polynomial inner approximation $\{q \in [-1,1] : v_{12}(q) < 0\}$ and the Hermite inner approximation $\{q \in [-1,1] : -g_{10}(q) < 0\}$, we observe, maybe surprisingly, that the approximations are very similar and miss the same parts of the stability region. These are not reproduced graphically.

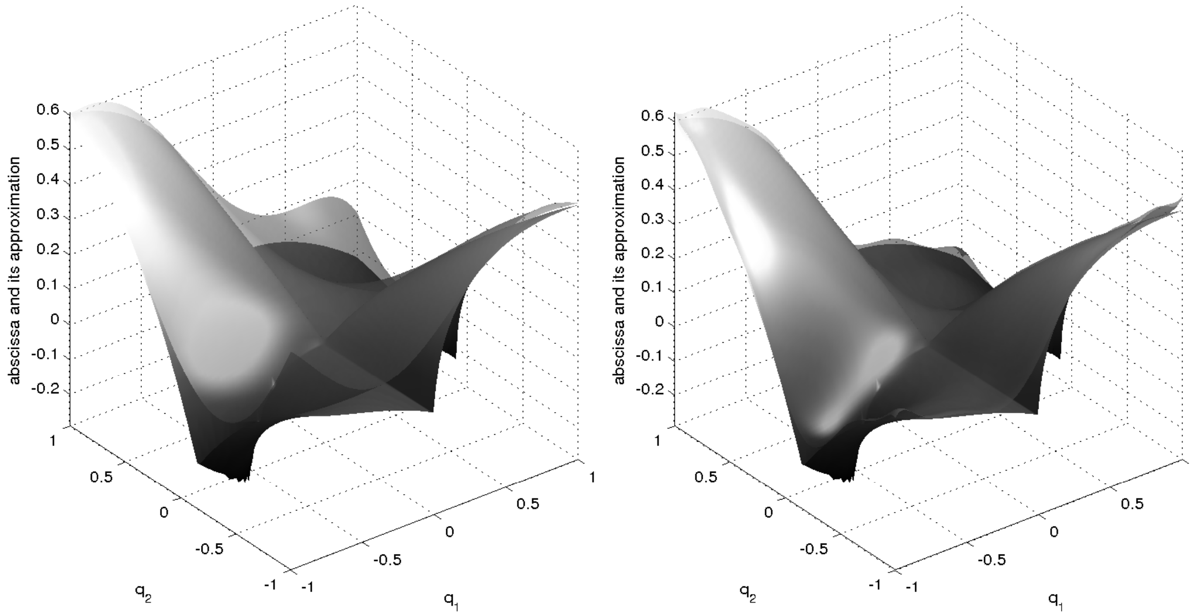

Figure 3.3: Abscissa $a(q)$ (dark, below) and its polynomial upper approximations $v_d(q)$ of degrees $d = 6$ (left, transparent) and $d = 10$ (right, transparent) for Example 3.4. We observe that the approximation deteriorates near the regions of non-differentiability of the abscissa.

Remark 3.7. The approach via the Hermite matrix does not tell us anything about the abscissa function itself, but just approximates its zero sublevel set, the so-called stability region defined in the Preliminaries. As an illustration consider a polynomial of the form $p(q,s) = s^2 + p_0(q)$ for $n = 1$. Then $p(q,\cdot)$ has either 0 as a multiple root, two real roots (of which one is positive) or only imaginary roots. Since these are all possible cases it follows that the stability region of p is empty and its Hermite matrix $H(q)$ is zero. Therefore its eigenvalues and their approximation g_d are also zero for every d. In contrast, the upper abscissa approximation v_d always gives a suitable approximation for the abscissa function.

On the other hand, practical experiments (not reported here) reveal that computing the abscissa approximation is typically more challenging numerically than computing the Hermite approximation. For instance, computing the upper abscissa approximation may fail for polynomials with large coefficients, while the Hermite approximation continues to provide a proper inner approximation of the stability region.

Example 3.4. Consider the polynomial

$$p: \; s \mapsto p(q,s) = s^3 + (q_1 + \tfrac{3}{2})s^2 + q_1^2 s + q_1 q_2$$

depending on $n = 2$ parameters $q \in \mathcal{Q} = [-1,1]^2$. Then $\mathcal{Z} = \{(q,x,y) \in [-1,1]^2 \times \mathbb{R}^2 : x^3 - 3xy^2 + (q_1 + \frac{3}{2})x^2 - (q_1 + \frac{3}{2})y^2 + q_1^2 x + q_1 q_2 = -y^3 + 3x^2 y + 2(q_1 + \frac{3}{2})xy + q_1^2 y = 0\}$. In Figure 3.3 we represent the graphs of the abscissa a and its polynomial approximations v_6 and v_{10}. In Figure 3.4 we represent the stabilizability region, i.e. the zero sublevel set of the abscissa $\{q \in [-1,1]^2 : a(q) < 0\}$ (dark gray region), the degree 8 Hermite sublevel set

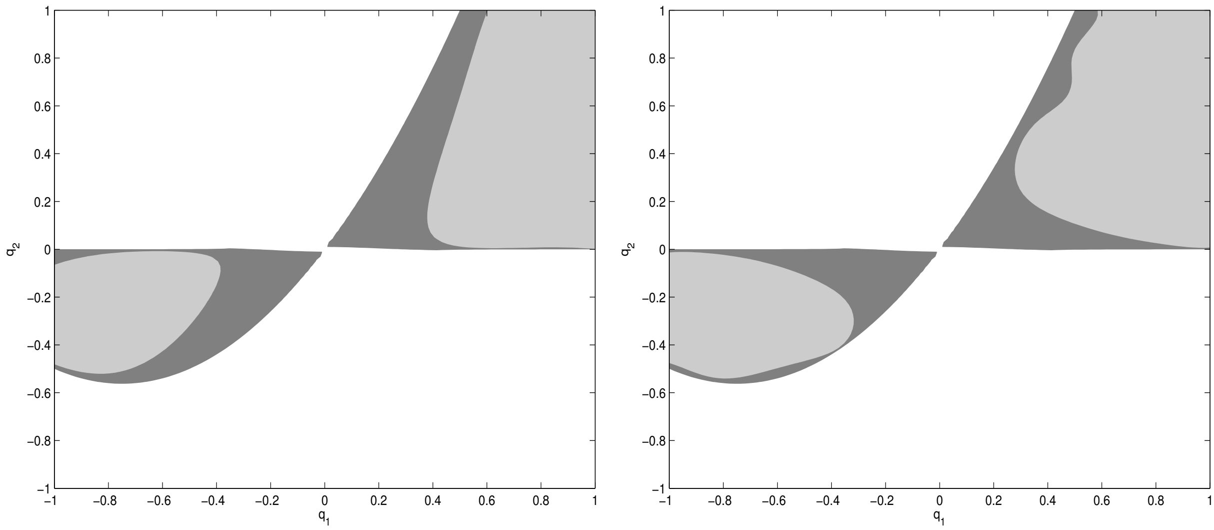

Figure 3.4: Stabilizability region (dark gray region) and its inner approximations with degree 8 Hermite (light gray region, left) and degree 10 upper polynomial approximation (light gray region, right) for Example 3.4 (compare with Figure 3.3).

$\{q \in [-1,1]^2 : -g_8(q) < 0\}$ (light gray region, left) and the degree 10 polynomial sublevel set $\{q \in [-1,1]^2 : v_{10}(q) < 0\}$ (light gray region, right).

Remark 3.8. In the examples we always chose lower degrees for the Hermite approximation than for the upper abscissa approximation. The Hermite approximation converges relatively fast making it unnecessary to consider higher degrees, especially since they require much more time. On the contrary, the upper abscissa approximation usually needs higher degrees to provide a useful approximation, but it is faster to compute.

4 Lower abscissa approximation

At first thought, finding a lower approximation for the abscissa map might sound like a straightforward task, since one is tempted to just solve the analogue of LP (1):

$$\begin{aligned} &\sup_{w \in \mathscr{C}(\mathcal{Q})} \int_{\mathcal{Q}} w(q)\, dq \\ &\text{s.t.} \quad x - w(q) \geq 0 \text{ for all } (q,x,y) \in \mathcal{Z}. \end{aligned} \tag{4}$$

This, indeed, gives a valid lower bound on the abscissa function, however in general a very bad one since it is not approximating the abscissa but the minimal real part of the roots of p. To understand the reason we recall that

$$\mathcal{Z} = \{(q,x,y) \in \mathbb{R}^n \times \mathbb{R}^2 : q \in \mathcal{Q},\ p_{\Re}(q,x,y) = p_{\Im}(q,x,y) = 0\}$$

and therefore this set contains all roots of p and not only those with maximal real part.

Example 4.1. On the left of Figure 4.1 we show the degree 12 solution to the SDP hierarchy corresponding to LP (4) for the polynomial $p(q,s) = s^2 + 2qs + 1 - 2q$ of Example 3.2, which gives a tight lower approximation to the abscissa only in the left part of the domain, corresponding to a pair of complex conjugate roots. We remark that the SDP solver Mosek does not return a correct answer for this particular problem, and we had to use the SDP solver SeDuMi instead in this case. On the right of Figure 4.1 we show the degree 12 solution to the SDP hierarchy corresponding to LP (4) for the polynomial $p(q,s) = s^3 + \frac{1}{2}s^2 + q^2 s + (q - \frac{1}{2})q(q + \frac{1}{2})$ of Example 3.3. The lower approximation is nowhere tight, due to the presence of roots with real parts smaller than the abscissa.

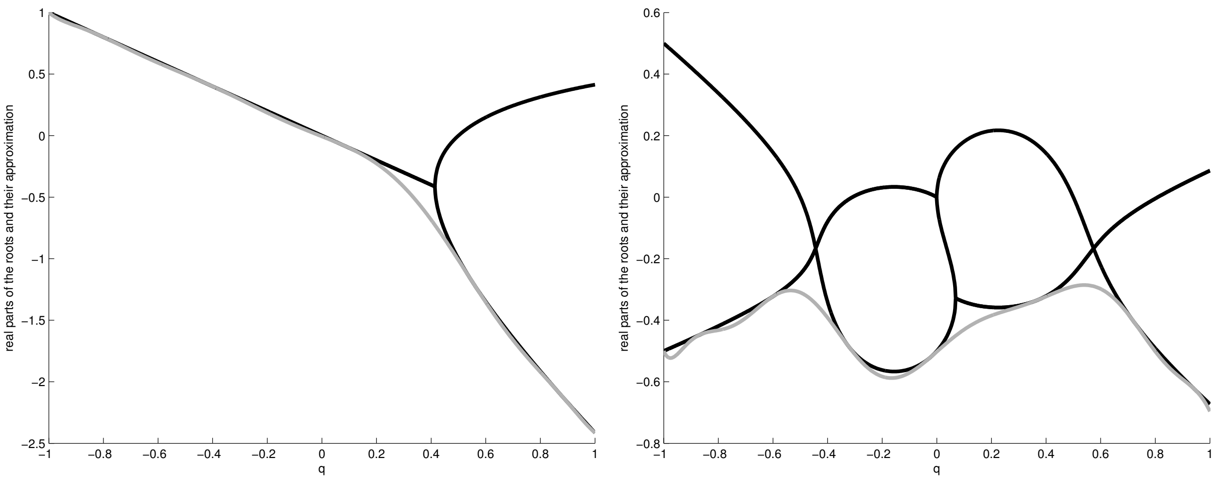

Figure 4.1: Real parts of the roots (black) and degree 12 polynomial lower approximations (gray) for the second degree polynomial (left) and third degree polynomial (right) of Example 4.1.

To find a tighter approximation for the abscissa map from below we pursue two different approaches:

- First, we reformulate the set $\mathcal{Z}$ with the help of elementary symmetric functions, in order to have access to the roots directly. This is a very neat way with options for variation, such as approximating the second largest real part of the roots from above or below, but it also includes many additional variables and it is therefore not very efficient when implemented. However, it can be useful for small problems.
- Second, we restrict LP (4) further using the Gauß-Lucas theorem, i.e. instead of $\mathcal{Z}$ we use a subset of $\mathcal{Z}$ which contains only the roots with the abscissa as its real parts. This approach is much more complicated, relies on assumptions and one needs to solve two optimization problems in order to get the lower approximation. Nevertheless, the implementation is much faster, so it can be used for bigger problems.

4.1 Lower approximation via elementary symmetric functions

4.1.1 Problem formulation

Let us derive another description of the set of roots of p which allows us to pick single roots according to the size of their real part. For this purpose let us recall the definition of our polynomial:

$$p: \ s \mapsto p(q,s) := \sum_{k=0}^{m} p_k(q)s^k \quad \text{with} \quad p_m(q) \equiv 1.$$

Following the notation of the previous sections, we denote the roots of $p(q,\cdot)$ by $s_r(q),\ r = 1,\dots,m$ and split them up into their real and imaginary parts, $s_r(q) = x_r(q) + iy_r(q)$ with $x_r(q), y_r(q) \in \mathbb{R}$. To simplify notation we omit the dependence on q whenever it is clear and write only $s_r,\ x_r$ and y_r.

Now we write the coefficients of the polynomial as elementary symmetric functions of its roots:

$$p_{m-k}(q) = (-1)^k \sum_{1\le l_1<l_2<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k}, \quad k = 1,\dots m.$$

This allows us to define the set of roots of p in the following way, where we can order the roots according to the size of their real part:

$$\begin{aligned}\mathcal{Z}'_o := \{(q, x_1,\dots,x_m, y_1,\dots,y_m) \in \mathcal{Q}\times\mathbb{R}^m\times\mathbb{R}^m : x_r \le x_m,\ r = 1,\dots,m-1,\\ p_{m-k}(q) = (-1)^k \sum_{1\le l_1<l_2<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k}, \quad k = 1,\dots m\}.\end{aligned}$$

To avoid the complex variables s_{l_k} in the description of the set, we could replace them by $s_{l_k} = x_{l_k} + iy_{l_k}$ and split the sum $\sum_{1\le l_1<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k}$ into its real and imaginary parts. The latter would be zero, since all $p_{m-k}(q)$ are real. In the sequel we omit this procedure, since it would only complicate notation.

For illustrative reasons let us fix q for a moment. Then the set $\mathcal{Z}'_o$ contains only one element $(q, x_1,\dots,x_m, y_1,\dots,y_m)$. From this it holds that $x_m = a(q)$ and the points $(q, x_r, y_r),\ r = 1,\dots,m$, are exactly the elements of $\mathcal{Z}$.

Remark 4.1. One could order the roots further by adding more conditions, such as $x_r \le x_{m-1},\ r = 1,\cdots,m-2$. Then one could also access the root with the second largest real part. Of course, this would imply another $m-2$ constraints in an implementation and therefore this would slow down further the solution process.

In theory, m variables suffice to characterize the roots of a real polynomial via the elementary symmetric functions, but since we need all variables x_k explicitly in order to identify the maximal one, we can only eliminate $\left\lfloor \frac{m}{2} \right\rfloor := \max\{z \in \mathbb{Z} \mid z \le \frac{m}{2}\}$ variables. We set

$$y_{r-1} = -y_r, \quad r = 2,\dots,m \qquad \text{if } m \text{ is even}$$

meaning we decide which roots will be pairs in case they are complex. If m is odd, we cannot assign the pairs of conjugate roots as easily as we did in the even case, because it is necessary to keep y_m, since we defined x_m as the abscissa. Furthermore, we need to consider that we do not know whether s_m is real or not. In fact $s_m(q)$ can be real for some q and complex for others. So we set

$$y_{r-1} = -y_r, \quad r = 2, \dots, m-3 \text{ and } y_{m-2} = -y_{m-1} - y_m \qquad \text{if } m \text{ is odd}$$

where the latter assignment comes from the imaginary part of the constraint $p_{m-1}(q) = -s_1 - s_2 - \cdots - s_m$ which reads $0 = -y_1 - y_2 - \cdots - y_m$, and means that at least one of the roots s_{m-2}, s_{m-1}, s_m is real, but we do not specify which.

Remark 4.2. Even though we know for m odd that one root must be real, we cannot eliminate $\left\lceil \frac{m}{2} \right\rceil$ variables, since it might happen that $s_m(q)$ is the single real root for some q while it is complex for other q.

Now we can write the set of roots with fewer variables and fewer constraints. As above we keep the variables s_r in the description of the set for readability reasons, but remark that with the reduced amount of y variables the constraints $0 = \Im(\sum_{1\le l_1 < \cdots < l_k \le m} s_{l_1} s_{l_2} \cdots s_{l_k})$ for $k = 1, \dots, \left\lfloor \frac{m}{2} \right\rfloor$ are superfluous. We have

$$\begin{aligned} \mathcal{Z}_o :=\{&(q, x_1, \dots, x_m, y_2, y_4, \dots, y_{2\lfloor \frac{m}{2} \rfloor}, y_m) \in \mathcal{Q} \times \mathbb{R}^m \times \mathbb{R}^{\lceil \frac{m}{2} \rceil} : \\ &x_r \le x_m, \ r = 1, \dots, m-1, \\ &p_{m-k}(q) = (-1)^k \sum_{1 \le l_1 < l_2 < \cdots < l_k \le m} s_{l_1} s_{l_2} \cdots s_{l_k}, \ k = 1, \dots m\}. \end{aligned}$$

Example 4.2. For $m = 3$ the set $\mathcal{Z}_o$ is given by

$$\begin{aligned} \mathcal{Z}_o =\{&(q, x_1, x_2, x_3, y_2, y_3) \in \mathcal{Q} \times \mathbb{R}^3 \times \mathbb{R}^2 : x_1 \le x_3, \ x_2 \le x_3, \\ &-p_2(q) = x_1 + x_2 + x_3, \\ &p_1(q) = x_1x_2 + x_1x_3 + x_2x_3 + y_2^2 + y_2y_3 + y_3^2, \\ &-p_0(q) = x_1x_2x_3 + (-x_1 + x_2 + x_3)y_2y_3 + x_2y_3^2 + x_3y_2^2, \\ &0 = (x_1 - x_2)y_2 + (x_1 - x_3)y_3, \\ &0 = (x_1 - x_2)x_3y_2 + (x_1 - x_3)x_2y_3 + y_2^2y_3 + y_2y_3^2\}. \end{aligned}$$

To clarify the formula also for m even, we write $\mathcal{Z}_o$ explicitly for $m = 4$

$$\begin{aligned} \mathcal{Z}_o =\{&(q, x_1, x_2, x_3, x_4, y_2, y_4) \in \mathcal{Q} \times \mathbb{R}^4 \times \mathbb{R}^2 : x_1 \le x_4, \ x_2 \le x_4, \ x_3 \le x_4, \\ &-p_3(q) = x_1 + x_2 + x_3 + x_4, \\ &p_2(q) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 + y_2^2 + y_4^2, \\ &-p_1(q) = x_1x_2(x_3 + x_4) + (x_1 + x_2)x_3x_4 + (x_1 + x_2)y_4^2 + (x_3 + x_4)y_2^2, \\ &p_0(q) = x_1x_2x_3x_4 + (x_1 - x_2)(x_4 - x_3)y_2y_4 + x_1x_2y_4^2 + x_3x_4y_2^2 + y_2^2y_4^2, \\ &0 = (x_1 - x_2)(x_3 + x_4)y_2 + (x_1 + x_2)(x_3 - x_4)y_4, \\ &0 = (x_1 - x_2)(x_3x_4 + y_4^2)y_2 + (x_3 - x_4)(x_1x_2 + y_2^2)y_4\}. \end{aligned}$$

Here we have set $y_1 = -y_2$ and $y_3 = -y_4$, so the constraint $0 = \Im(\sum_{1\le l_1<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k})$ for $k=1$ is obviously superfluous, because it reduces to $0=0$. The second superfluous constraint is the one for $k=2$, that is $0 = (x_1-x_2)y_2 + (x_3-x_4)y_4$, since we have $x_1 = x_2$, respectively $x_3 = x_4$, in the case s_2, respectively s_4, is complex.

Finally, we can reformulate LP (4) in such a way that it provides a proper approximation of the abscissa function from below:

$$\begin{aligned} \vartheta = &\sup_{w\in\mathscr{C}(\mathcal{Q})} \int_{\mathcal{Q}} w(q)\,dq \\ &\text{s.t.}\quad x_m - w(q) \ge 0 \text{ for all } (q, x_1, \dots, x_m, y_2, y_4, \dots, y_{2\lfloor \frac{m}{2}\rfloor}, y_m) \in \mathcal{Z}_o. \end{aligned} \tag{5}$$

With the notation of Section 3.1 its dual LP reads

$$\begin{aligned} \vartheta^* = &\inf_{\mu\in\mathscr{M}^+(\mathcal{Z}_o)} \int_{\mathcal{Z}_o} x_m\, d\mu(q, x_1, \dots, x_m, y_2, y_4, \dots, y_{2\lfloor \frac{m}{2}\rfloor}, y_m) \\ &\text{s.t.} \int_{\mathcal{Z}_o} q^\alpha\, d\mu = \int_{\mathcal{Q}} q^\alpha\, dq, \text{ for all } \alpha\in\mathbb{N}^n. \end{aligned} \tag{6}$$

In analogy with the upper approximation we have no duality gap and the infimum is attained:

Lemma 4.1. *The infimum in LP (6) is attained, and there is no duality gap between LP (5) and LP (6), i.e.* $\vartheta = \vartheta^*$.

Since $\mathcal{Z}_o$ is compact, the proof is identical to that of Lemma 3.1.

Remark 4.3. For the same reasons as for the upper approximation (1), the supremum in (5) is not attained for $\mathscr{C}(\mathcal{Q})$ or $\mathbb{R}[q]$, but it is attained for $\mathbb{R}[q]_d$ with d finite. See Remark 3.3 with $M := \min_{q\in\mathcal{Q}} a(q) - N$ for $N\in\mathbb{N}$ sufficiently large, and $R := \int_{\mathcal{Q}} (a(q) - M)\, dq$.

4.1.2 SDP hierarchy

Let $d_0 \in \mathbb{N}$ be sufficiently large. Then for $d\in\mathbb{N},\ d\ge d_0$ the corresponding hierarchy of SDP problems reads

$$\begin{aligned} \vartheta_d = &\sup_{w_d,\sigma_0,\sigma_j,\sigma_{x_r},\tau_{\Re,k},\tau_{\Im,k}} \int_{\mathcal{Q}} w_d(q)\,dq \\ &\text{s.t. } x_m - w_d(q) = \sigma_0 + \sum_{j=1}^{n} \sigma_j(1-q_j^2) + \sum_{r=1}^{m-1} \sigma_{x_r}(x_m - x_r) \\ &\qquad + \sum_{k=1}^{m} \tau_{\Re,k}\left((-1)^k p_{m-k}(q) - \Re\left(\sum_{1\le l_1<l_2<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k}\right)\right) \\ &\qquad + \sum_{k=\lfloor \frac{m}{2}\rfloor}^{m} \tau_{\Im,k}\Im\left(\sum_{1\le l_1<l_2<\cdots<l_k\le m} s_{l_1}s_{l_2}\cdots s_{l_k}\right) \end{aligned} \tag{7}$$

for all $(q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_{2\lfloor \frac{m}{2} \rfloor}, y_m) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{\lceil \frac{m}{2} \rceil}$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0, \sigma_{x_r} \in \Sigma[q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_m]_{2d}$ for $r = 1, \ldots, m-1$, $\sigma_j \in \Sigma[q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_m]_{2d-2}$ for $k = 1, \ldots, n$, $\tau_{\Re,k} \in \mathbb{R}[q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_m]_{2d-k}$ for $k = 1, \ldots, m$ and $\tau_{\Im,k}$ for $k = \left\lfloor \frac{m}{2} \right\rfloor, \ldots, m$.

Remark 4.4. As in Remark 3.5, SDP (7) is a strengthening of LP (5), meaning $\vartheta_d \leq \vartheta$. Also as in Remark 3.4, the quadratic module corresponding to $\mathcal{Z}_o$ is archimedean, i.e. $\lim_{d \to \infty} \vartheta_d = \vartheta$.

We conclude the section with the following result:

Theorem 4.2. *Let $w_d \in \mathbb{R}[q]_{2d}$ be a near optimal solution for SDP* (7)*, i.e. $\int_{\mathcal{Q}} w_d(q)\, dq \geq \vartheta_d - \frac{1}{d}$ and consider the associated sequence $(w_d)_{d \geq d_0} \subset L^1(\mathcal{Q})$. Then w_d converges to a in L^1 norm in $\mathcal{Q}$.*

Unsurprisingly, one can prove this result in exactly the same way as Theorem 3.2, so we do not detail the proof here. We remark that the first part of the proof can be shortened, since $\int_{\mathcal{Z}_o} x_m \, d\mu(q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_m) = \int_{\mathcal{Z}_o} a(q)\, d\mu(q, x_1, \ldots, x_m, y_2, y_4, \ldots, y_m)$.

4.1.3 Examples

Just as the upper abscissa approximation automatically approximates the stability region from inside, the lower approximation gives, as a side effect, an outer approximation. In this section we will examine similar examples as for the upper approximation.

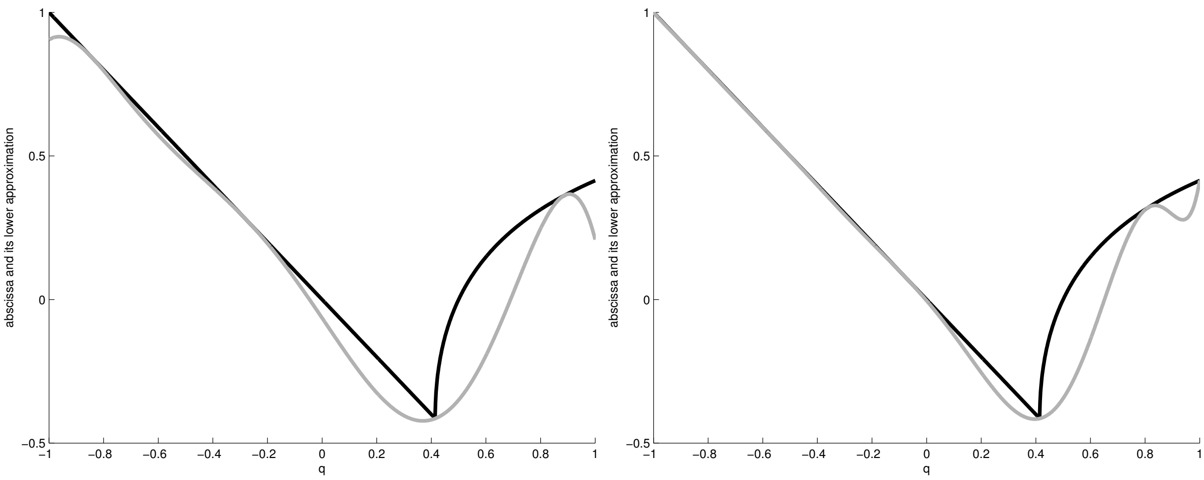

Figure 4.2: Abscissa $a(q)$ (black) and its polynomial lower approximations $w_d(q)$ of degree $d = 6$ (gray, left) and $d = 10$ (gray, right) for Example 4.3. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz, compare with Figure 3.1.

Example 4.3. As in Example 3.2 consider the polynomial

$$p: \; s \mapsto p(q, s) = s^2 + 2qs + 1 - 2q.$$

We have $y_1 = -y_2$, so $\mathcal{Z}_o := \{(q, x_1, x_2, y_2) \in \mathcal{Q} \times \mathbb{R}^3 : x_1 \le x_2,\ -2q = x_1 + x_2,\ 1 - 2q = x_1 x_2 + y_2^2,\ 0 = (x_1 - x_2) y_2\}$. In Figure 4.2 we see the graphs of the degree 6 and 10 polynomial lower approximations obtained by solving SDP (7). As in Example 4.1, we observe that the SDP solver Mosek does not return a correct degree 10 polynomial, and we had to use the SDP solver SeDuMi instead in this case. Due to the rather large number of variables and constraints, computing the degree 10 solution is already relatively expensive, with a few seconds of CPU time.

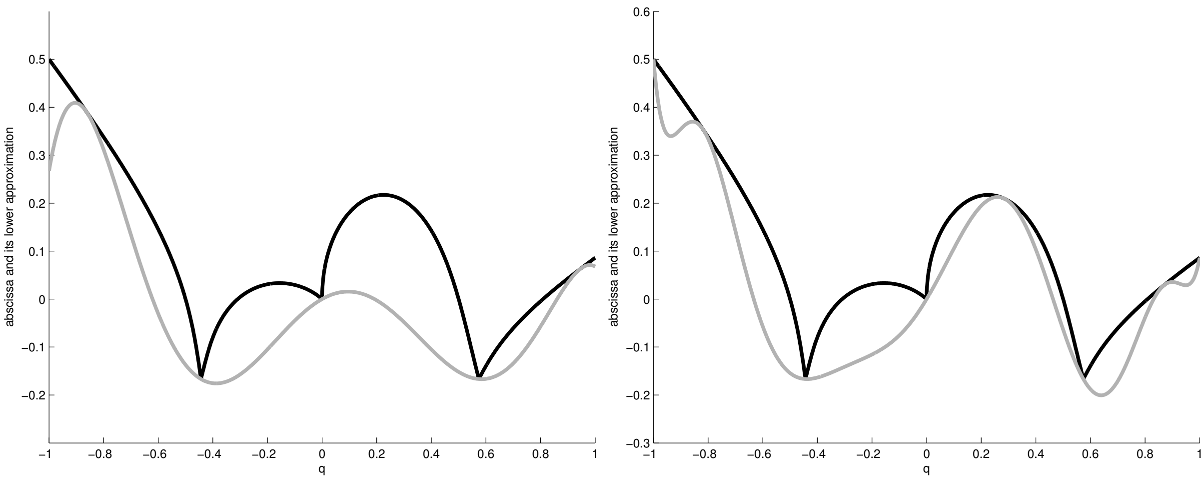

Figure 4.3: Abscissa $a(q)$ (black) and its polynomial lower approximations $w_d(q)$ of degree $d = 6$ (gray, left) and $d = 10$ (gray, right) for Example 4.4. The quality of the approximation deteriorates near the minimum, where the abscissa is not differentiable, compare with Figure 3.2.

Example 4.4. As in Example 3.3 consider the polynomial

$$p:\ s \mapsto p(q,s) = s^3 + \tfrac{1}{2}s^2 + q^2 s + (q - \tfrac{1}{2})q(q + \tfrac{1}{2}).$$

With $y_1 = -y_2 - y_3$ we calculate $\mathcal{Z}_o$ as in Example 4.2. In Figure 4.3 we see the graphs of the degree 6 and 10 polynomial lower approximations obtained by solving SDP (7). The computation time to get the degree 10 solution is around 15 minutes, which is arguably unreasonable given the quality of the approximation.

Remark 4.5. As for the upper abscissa approximation, we observe practically that the implementation for the lower approximation is rather sensitive to polynomials with large coefficients.

Example 4.5. As in Example 3.4, consider the polynomial

$$p:\ s \mapsto p(q,s) = s^3 + (q_1 + \tfrac{3}{2})s^2 + q_1^2 s + q_1 q_2.$$

Since we have $m = 3$, the set $\mathcal{Z}_o$ is again given in Example 4.2. In Figure 4.4 we see the outer approximation of degrees 6 and 8 obtained by solving SDP (7). In the lower half of the picture we notice that the approximation of the stability region is rather bad near $q_1 = 0$, even for degree 8. This is due to a being zero and non-smooth for $q_1 = 0$, meaning $a(0, q_2) = 0$ for all $q_2 \in [-1, 1]$ which makes the abscissa especially hard to approximate in this region. This phenomenon also prevents w_8 from getting closer to a for $q_2 > 0$ than we observe in the upper half of the picture.

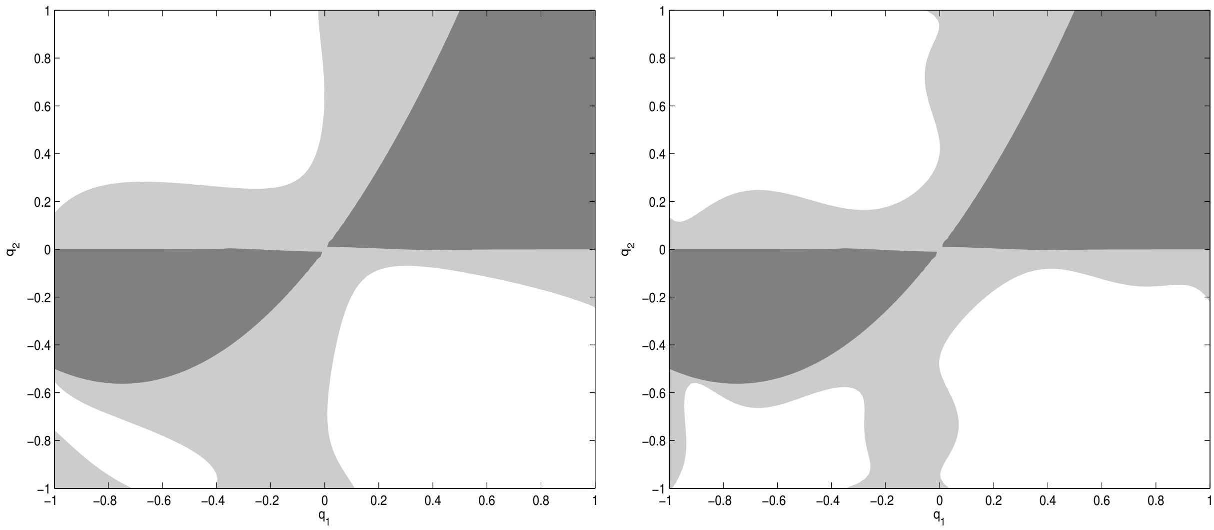

Figure 4.4: Stabilizability region (dark gray region) and its degree 6 outer approximation (light gray region, left) and degree 8 outer approximation (light gray region, right) for Example 4.5. Compare with Figure 3.4.

4.2 Lower approximation via Gauß-Lucas

4.2.1 Problem formulation

As indicated above, we want to find a semi-algebraic subset of $\mathcal{Z}$ which contains only those roots of p whose real part is maximal. In contrast to the approach of Section 4.1, we will not redefine $\mathcal{Z}$, but formulate further constraints.

In order to do this we must distinguish between the roots of $p(q,\cdot)$ according to the size of their real parts. For this purpose we use the following result:

Theorem 4.3 (Gauß-Lucas)**.** *The critical points of a non-constant polynomial lie in the convex hull of its roots.*

We refer to [5] for further information and a proof. Let us denote the derivative of $p(q,s)$ with respect to s by $p'(q,s)$. By Theorem 4.3, the roots of $p'(q,\cdot)$ are contained in the convex hull of the roots of $p(q,\cdot)$. It follows readily that the abscissa $a_{p'}$ of p' lies below the abscissa a_p of p:

$$a_{p'}(q) \leq a_p(q) \text{ for all } q \in \mathcal{Q}.$$

However, p may have some roots with real part strictly smaller than a_p and strictly bigger than $a_{p'}$, meaning that the root whose real part is the abscissa is not the only one whose real part lies above $a_{p'}$. Of course, this cannot happen for real polynomials $\mathbb{R} \to \mathbb{R}$ because of monotonicity, and neither can it for complex polynomials of degree 2. But, for example, for $n = 1$ the polynomial $p(q,s) = s^4 + (q^2+1)s + q$ has two roots with different real parts greater than $a_{p'}$ for $q \in [-1, -0.4]$.

To prevent the lower abscissa approximation from converging to the real part of a root smaller than the abscissa, we make the following assumption:

Assumption 1. *None of the real parts of any root of p that differs from the abscissa coincides with $a_{p'}$ or lies strictly between a_p and $a_{p'}$, i.e. $x \notin [a_{p'}(q), a_p(q)[$ for all $(q, x, y) \in \mathcal{Z}$.*

Remark 4.6. Unfortunately, we do not know how restrictive this assumption is. For $n = 1$ it was rather difficult to find examples that violate it.

Now let $\hat{v} \in \mathscr{C}(\mathcal{Q})$ be a near optimal solution to LP (1) for the polynomial p', meaning $\int_{\mathcal{Q}} \hat{v}(q)\, dq \leq \rho + \varepsilon$ for an $\varepsilon > 0$. Then, $\hat{v}$ is an upper approximation of the abscissa $a_{p'}$ of p'. We define the following subset:

$$\hat{\mathcal{Z}} := \{(q, x, y) \in \mathcal{Z} : x - \hat{v}(q) \geq 0\}.$$

In order to see where we are going, let us suppose for a moment that $\hat{v}$ is actually an optimal solution. Then, under Assumption 1, the set $\hat{\mathcal{Z}}$ would contain exactly the points $(q, a_p(q), y_q)$ (with y_q denoting the imaginary part of the root of $p(q, \cdot)$) with maximal real part. Hence, the solution to the following LP would give a lower approximation of the abscissa function a_p of p:

$$\begin{aligned} &\sup_{w \in \mathscr{C}[q]} \int_{\mathcal{Q}} w(q)\, dq \\ &\text{s.t.} \quad x - w(q) \geq 0 \text{ for all } (q, x, y) \in \hat{\mathcal{Z}}. \end{aligned} \tag{8}$$

Since $\hat{v}$ might not be optimal, the projection of $\hat{\mathcal{Z}}$ onto $\mathcal{Q}$ might not be $\mathcal{Q}$ as required, but differ from it on a set of volume ε. As a consequence, w might not be a valid lower bound of the abscissa on this set.

Taking this into account, we build an SDP hierarchy for LP (8) in the next section. The issue is that we have to consider the hierarchy for the upper approximation of $a_{p'}$ first and the solution to it might interfere with a_p.

4.2.2 SDP hierarchy

For $d'_0 \in \mathbb{N}$ sufficiently large we denote by $\hat{v}_{d'}$, $d' \geq d'_0$, the solutions to SDP (3) for the polynomial p'. Thus, the $\hat{v}_{d'}$ are polynomials in $\mathbb{R}[q]_{2d'}$ and by Theorem 3.2 the sequence $(\hat{v}_{d'})_{d' \in \mathbb{N}}$ converges to $a_{p'}$ from above in L^1 norm.

Next, we want to describe the set $\hat{\mathcal{Z}}$ via the polynomials $\hat{v}_{d'}$ in order to have an implementable problem, i.e. we define

$$\hat{\mathcal{Z}}_{d'} := \{(q, x, y) \in \mathcal{Z} : x - \hat{v}_{d'}(q) \geq 0\}.$$

Of course, the set $\hat{\mathcal{Z}}_{d'}$ is highly dependent on the quality of $\hat{v}_{d'}$ and hence on the choice of d'. Evidently, $\hat{\mathcal{Z}}_{d'}$ is a subset of $\hat{\mathcal{Z}}$, possibly strictly. To ensure that $\hat{\mathcal{Z}}_{d'}$ contains all roots of p with the abscissa as their real parts we need $\hat{v}_{d'} \leq a_p$. However, in practice this is impossible in some cases:

Example 4.6. The abscissa a_p of $p(q,s) = (s^3 + q)^2$ and the abscissa $a_{p'}$ of p' coincide and have a point of non-differentiability at $q = 0$. As another example consider the polynomial $p(q,s) = s^4 + qs$ for which both a_p and $a_{p'}$ are not differentiable at $q = 0$ and $a_p(0) = a_{p'}(0) = 0$.

For these examples we cannot achieve $\hat{v}_{d'} \leq a_p$ with d' finite, since $\hat{v}_{d'}$ is a polynomial and therefore differentiable everywhere.

As a consequence, we formulate another assumption. In general, the points that may cause problems are the ones where a_p and $a_{p'}$ coincide, i.e. the points of the set

$$\mathcal{D} := \{q \in \mathcal{Q} : a_p(q) = a_{p'}(q)\}.$$

On this set the polynomial $\hat{v}_{d'}$ should approximate $a_{p'}$ perfectly for a finite d', meaning $\hat{v}_{d'}(q) = a_{p'}(q)$ for all $q \in \mathcal{D}$. Calling a solution $\hat{v}_{d'}$ near optimal if it satisfies $\int_{\mathcal{Q}} \hat{v}_{d'}(q)\,dq \leq \rho_{d'} + \frac{1}{d'}$, we assume:

Assumption 2. *There is a near optimal solution $\hat{v}_{d'}$ to SDP* (3) *for the polynomial p' with d' finite such that $\hat{v}_{d'}$ and $a_{p'}$ coincide on $\mathcal{D}$.*

Remark 4.7. A sufficient condition for a violation of Assumption 2 is the existence of a value of q for which $a_{p'}$ is not differentiable and $a_p(q) = a_{p'}(q)$. This is the case for the examples given above. Note also that they are of degenerate nature.

To face another issue, we denote the projection of $\hat{\mathcal{Z}}_{d'}$ onto the set $\mathcal{Q}$ by $\pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'})$, i.e.

$$\pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'}) = \{q \in \mathcal{Q} : \exists x, y \in \mathbb{R} : (q, x, y) \in \hat{\mathcal{Z}}_{d'}\}.$$

Since $\hat{v}_{d'}$ converges to a in L^1, but not necessarily uniformly, it might have spikes or similar irregularities, meaning that the set $\mathcal{Q} \setminus \pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'})$ is not empty. However, the L^1 convergence of $\hat{v}_{d'}$, or more precisely the convergence in measure, implies that there is a subsequence $(\hat{v}_{d'_l})_{l \in \mathbb{N}}$ which converges to $a_{p'}$ almost uniformly (see e.g. [1, Theorem 2.5.3]). In other words, for all $\delta > 0$, there exists a set $\mathcal{A}_\delta$ in the Borel sigma algebra of $\mathcal{Q}$ such that $\int_{\mathcal{A}_\delta} dq < \delta$ and $\hat{v}_{d'_l}$ converges uniformly on $\mathcal{A}_\delta^C$ to $a_{p'}$ when $l \to \infty$, where $\mathcal{A}_\delta^C$ is the set-theoretic complement of $\mathcal{A}_\delta$ in $\mathcal{Q}$. With this notation we have

$$\pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'}) \subseteq \mathcal{A}_\delta^C \subseteq \mathcal{Q}.$$

Lemma 4.4. *Let Assumption 2 hold. Then, for every $\delta > 0$ there exists a finite $d' \in \mathbb{N}$ and a set $\mathcal{A}_\delta$ in the Borel sigma algebra of $\mathcal{Q}$ with $\int_{\mathcal{A}_\delta} dq < \delta$ such that $\hat{v}_{d'} \leq a_p$ on $\mathcal{A}_\delta^C$.*

Proof. Fix $\delta > 0$. As discussed above there exists a set $\mathcal{A}_\delta$ in the Borel sigma algebra of $\mathcal{Q}$ such that $\int_{\mathcal{A}_\delta} dq < \delta$ and $\hat{v}_{d'_l}$ converges uniformly to $a_{p'}$ on $\mathcal{A}_\delta^C$ as $l \to \infty$. Obviously we want

$$0 \leq a_p(q) - \hat{v}_{d'}(q) = a_p(q) - a_{p'}(q) + a_{p'}(q) - \hat{v}_{d'}(q) \tag{9}$$

for every $q \in \mathcal{A}_\delta^C \subseteq \mathcal{Q}$. By Theorem 4.3, we have $a_p(q) - a_{p'}(q) \geq 0$ for all $q \in \mathcal{Q}$. In contrary, the difference $a_{p'}(q) - \hat{v}_{d'}(q)$ is negative by construction, but due to Theorem 3.2

we find a subsequence $\hat{v}_{d'_l}$ converging uniformly to $a_{p'}$ on $\mathcal{A}_\delta^C$. Hence, there is a finite d'_{l^*} such that (9) is fulfilled for all $q \in \{q \in \mathcal{A}_\delta^C : a_p(q) > a_{p'}(q)\}$. Because of Assumption 2 there is also a finite d' such that $a_{p'}(q) - \hat{v}_{d'}(q)$ vanishes on $\{q \in \mathcal{A}_\delta^C : a_p(q) = a_{p'}(q)\} \subseteq \mathcal{D}$. Taking $d'_{l'} \geq d'$ with $l' \geq l^*$ completes the proof. □

Remark 4.8. Choosing d' according to Lemma 4.4 implies $\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'}) = \mathcal{A}_\delta^C$.

Under Assumption 1 and 2 and with an appropriate choice of d' (depending on δ) the solution to the following LP gives a lower approximation of the abscissa function a_p of p on the set $\mathcal{A}_\delta^C \subseteq \mathcal{Q}$:

$$\begin{aligned} \vartheta_{d'} = \sup_{w \in \mathscr{C}(\mathcal{Q})} & \int_{\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'})} w(q)\, dq \\ \text{s.t.} \quad & x - w(q) \geq 0 \text{ for all } (q, x, y) \in \hat{\mathcal{Z}}_{d'}. \end{aligned} \tag{10}$$

Remark 4.9. Note that under Assumption 1, LP (10) always provides a proper approximation for the abscissa a_p from below on $\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'})$, but this might not be very useful, since for bad $\hat{v}_{d'}$ this set may have big holes or even be empty. To achieve suitable results on $\mathcal{A}_\delta^C$ we need Assumption 2 and an appropriate d', meaning a sufficiently good $\hat{v}_{d'}$ ensuring $\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'}) = \mathcal{A}_\delta^C$.

In analogy with (2), the dual LP reads

$$\begin{aligned} \vartheta^*_{d'} = \inf_{\mu \in \mathscr{M}^+(\hat{\mathcal{Z}}_{d'})} & \int_{\hat{\mathcal{Z}}_{d'}} x \; d\mu(q, x, y) \\ \text{s.t.} & \int_{\hat{\mathcal{Z}}_{d'}} q^\alpha \, d\mu = \int_{\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'})} q^\alpha \, dq, \text{ for all } \alpha \in \mathbb{N}^n \end{aligned} \tag{11}$$

with the notation of Section 3.1.

Lemma 4.5. *The infimum in LP* (11) *is attained, and there is no duality gap between LP* (10) *and LP* (11), *i.e.* $\vartheta_{d'} = \vartheta^*_{d'}$.

Since $\hat{\mathcal{Z}}_{d'}$ is a compact subset of $\mathcal{Z}$, we can mimic the proof of Lemma 3.1 in order to obtain a proof of Lemma 4.5.

Remark 4.10. As in Remark 3.3, the supremum in LP (10) is not attained for $\mathscr{C}(\mathcal{Q})$ or $\mathbb{R}[q]$, but it is attained for $\mathbb{R}[q]_d$ with d finite. To adjust the proof of Remark 3.3, set $M := \min_{q \in \mathcal{Q}} a(q) - N$ for an $N \in \mathbb{N}$ sufficiently large, and $R := \int_\mathcal{Q} (a(q) - M)\, dq$ as in Remark 4.3.

Finally, for d' as in Lemma 4.4 and $d_0 \geq d'$ sufficiently large we can write an SDP hierarchy indexed by $d \in \mathbb{N},\ d \geq d_0$:

$$\begin{aligned} \vartheta_{d',d} = \sup_{w_d, \sigma_0, \sigma_j, \sigma_{\hat{v}}, \tau_\Re, \tau_\Im} & \int_{\pi_\mathcal{Q}(\hat{\mathcal{Z}}_{d'})} w_d(q)\, dq \\ \text{s.t. } x - w_d(q) = {} & \sigma_0(q,x,y) + \sum_{j=1}^n \sigma_j(q,x,y)(1 - q_j^2) + \sigma_{\hat{v}}(q,x,y)(x - \hat{v}_{d'}(q)) \\ & + \tau_\Re(q,x,y) p_\Re(q,x,y) + \tau_\Im(q,x,y) p_\Im(q,x,y) \end{aligned} \tag{12}$$

for all $(q,x,y) \in \mathbb{R}^n \times \mathbb{R}^2$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q,x,y]_{2d}$, $\sigma_j \in \Sigma[q,x,y]_{2d-2}$ for $j = 1,\dots,n$, $\sigma_{\hat{v}} \in \Sigma[q,x,y]_{2d-d'}$ and $\tau_{\Re}, \tau_{\Im} \in \mathbb{R}[q,x,y]_{2d-m}$.

Remark 4.11. As in section 3.2, SDP (12) is a strengthening of LP (10), meaning $\vartheta_{d',d} \le \vartheta_{d'}$. Besides, the archimedean quadratic module corresponding to the set $\mathcal{Z}$ is contained in the quadratic module corresponding to $\hat{\mathcal{Z}}_{d'}$. Hence, the latter is also archimedean, i.e. $\lim_{d\to\infty} \vartheta_{d',d} = \vartheta_{d'} = \vartheta^*_{d'}$.

Remark 4.12. For numerical applications one can assume that $\mathcal{A}_\delta$ is empty and substitute $\pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'})$ by $\mathcal{Q}$.

The associated sequence converges:

Theorem 4.6. *Let Assumptions 1 and 2 hold and let $\mathcal{A}_\delta^C$ and d' be as in Lemma 4.4. Let $w_d \in \mathbb{R}[q]_{2d}$ be a near optimal solution for SDP (12), i.e. $\int_{\mathcal{Q}} w_d(q)\,dq \ge \vartheta_{d,d'} - \frac{1}{d}$. Consider the associated sequence $(w_d)_{d\ge d_0} \subset L^1(\mathcal{Q})$. Then w_d is a valid lower bound of a_p on $\mathcal{A}_\delta^C$ and it converges to a_p in L^1 norm on $\mathcal{A}_\delta^C$.*

The proof of this result is very similar to the proof of Theorem 3.2, so we omit it. Note that by Lemma 4.4 every feasible solution to SDP (12) is a valid lower bound of a_p on $\mathcal{A}_\delta^C$ and that we have $\pi_{\mathcal{Q}}(\hat{\mathcal{Z}}_{d'}) = \mathcal{A}_\delta^C$ due to our choice of d'. As for the proof of Theorem 4.2, the first part can be shortened, since for every $(q,x,y) \in \hat{\mathcal{Z}}_{d'}$ it holds that $x = a(q)$.

4.2.3 Examples

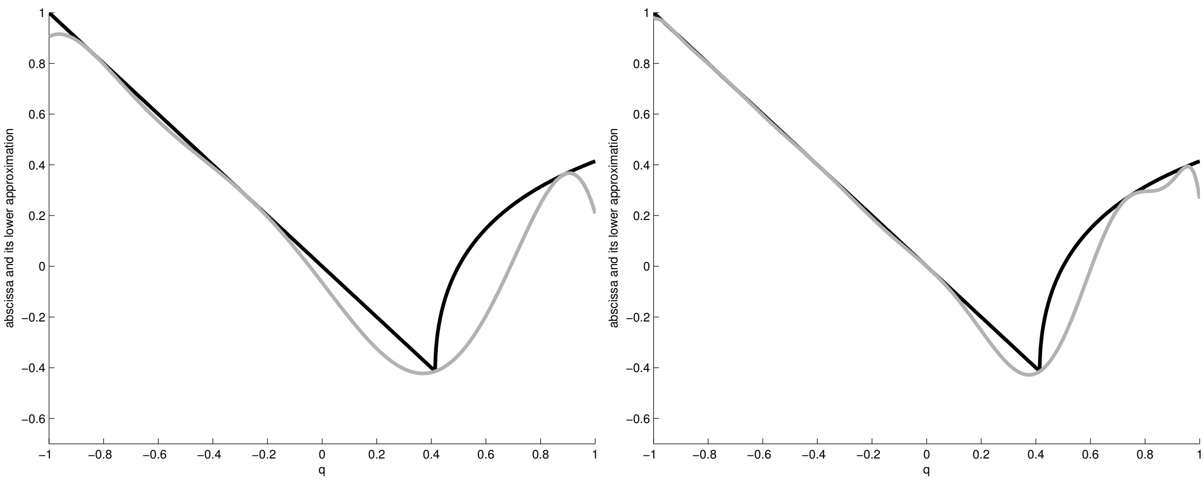

Figure 4.5: Abscissa $a_p(q)$ (black) and its polynomial lower approximations $w_d(q)$ of degree $d = 6$ (gray, left) and $d = 12$ (gray, right) for Example 4.7. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz, compare with Figures 3.1 and 4.2 .

Example 4.7. As in Examples 3.2 and 4.3 consider

$$p: \; s \mapsto p(q,s) = s^2 + 2qs + 1 - 2q.$$

Assumption 1 is naturally fulfilled, since p is of degree 2. In the same way, Assumption 2 is fulfilled, since $a_{p'}(q) = -q$ is polynomial. We have $\hat{\mathcal{Z}}_{d'} = \{(q,x,y) \in [-1,1] \times \mathbb{R}^2 : x - \hat{v}_{d'}(q) \geq 0,\ x^2 - y^2 + 2qx + 1 - 2q = 2xy + 2qy = 0\}$ and the corresponding SDP (3) reads

$$\begin{aligned} \vartheta_{d',d} = \sup_{w_d,\sigma_0,\sigma_1,\sigma_{\hat{v}},\tau_{\Re},\tau_{\Im}} & \int_{-1}^{1} w_d(q)\, dq \\ \text{s.t. } x - w_d(q) &= \sigma_0(q,x,y) + \sigma_1(q,x,y)(1-q^2) + \sigma_{\hat{v}}(q,x,y)(x - \hat{v}_{d'}(q)) \\ &+ \tau_{\Re}(q,x,y)(x^2 - y^2 + 2qx + 1 - 2q) + \tau_{\Im}(q,x,y)(2xy + 2qy) \end{aligned}$$

for all $(q,x,y) \in \mathbb{R}^3$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q,x,y]_{2d}$, $\sigma_1 \in \Sigma[q,x,y]_{2d-2}$, $\sigma_{\hat{v}} \in \Sigma[q,x,y]_{2d-d'}$ and $\tau_{\Re}, \tau_{\Im} \in \mathbb{R}[q,x,y]_{2d-2}$. Due to the simplicity of $a_{p'}$ it suffices to choose $d' = 2$. We see the degree 6 and 12 polynomial lower approximations in Figure 4.5. They are both computed in less than 2 seconds.

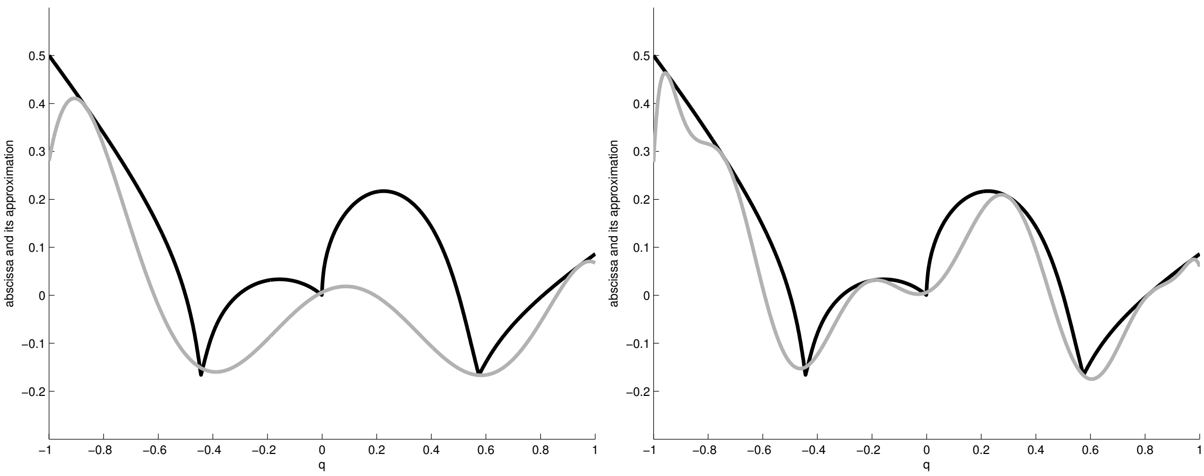

Figure 4.6: Abscissa $a_p(q)$ (black) and its polynomial lower approximations $w_d(q)$ of degree $d = 6$ (gray, left) and $d = 12$ (gray, right) for Example 4.8. We observe that the approximations are not valid near $q = -0.5$ and $q = 0$, as Assumption 2 is violated.

Example 4.8. As in Examples 3.3 and 4.4 consider

$$p: \ s \mapsto p(q,s) = s^3 + \tfrac{1}{2}s^2 + q^2 s + (q - \tfrac{1}{2})q(q + \tfrac{1}{2}).$$

The abscissa $a_{p'}$ of p' is not differentiable in two points, hence it is not a polynomial and it cannot be described perfectly by $\hat{v}_{d'}$ for finite d'. Let us choose $d' = 8$ and $d = 6$ (resp. $d = 12$). We observe in Figure 4.6 that w_6 (resp. w_{12}) is not everywhere a valid lower bound. Indeed, the set $\mathcal{D} = \{q \in \mathcal{Q} : a_p(q) = a_{p'}(q)\}$ contains three points and for two of these (near $q = -0.5$ and $q = 0$), the approximation $\hat{v}_8$ is not tight enough to ensure $\pi_{\mathcal{Q}}(\mathcal{Z}_{r,8}) = \mathcal{Q}$. Consequently, Assumption 2 is violated.

Example 4.9. In order to discuss another example for which $\mathcal{D}$ is a non-empty interval, consider the polynomial

$$p: \ s \mapsto p(q,s) = s^2 + (20q^2 - 1)s + q + \tfrac{1}{2}.$$

Here $a_{p'}(q) = -10q^2 + \frac{1}{2}$ is a quadratic polynomial. Thus, Assumption 2 is fulfilled, in particular $\hat{v}_2 = a_{p'}$, and the lower approximations are valid, see Figure 4.7.

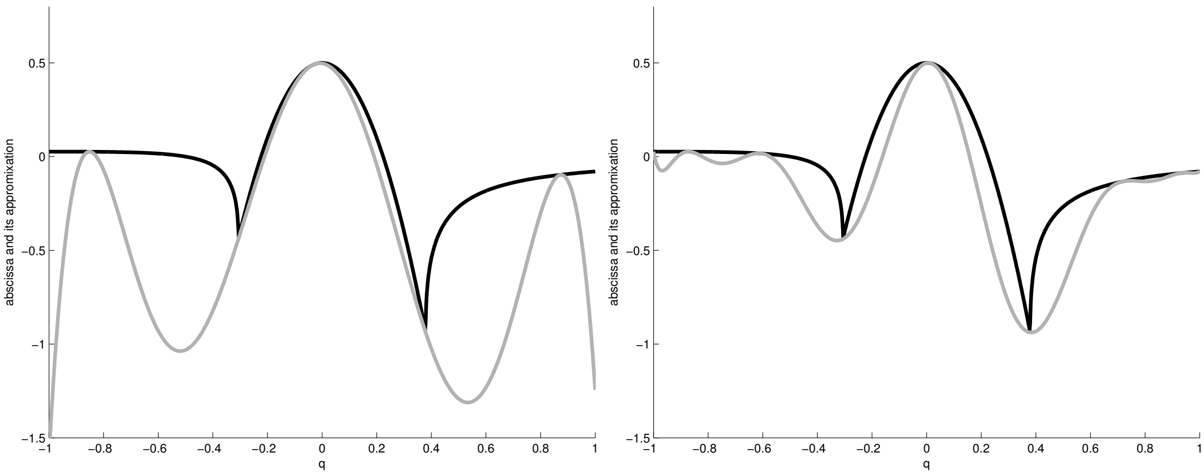

Figure 4.7: Abscissa $a_p(q)$ (black) and its polynomial lower approximations $w_d(q)$ of degree $d = 6$ (gray, left) and $d = 12$ (gray, right) for Example 4.9 ($d' = 2$).

Example 4.10. As in Examples 3.4 and 4.5 consider the polynomial

$$p: \ s \mapsto p(q,s) = s^3 + (q_1 + \tfrac{3}{2})s^2 + q_1^2 s + q_1 q_2.$$

We have $\hat{\mathcal{Z}}_{d'} := \{(q,x,y) \in \mathcal{Z} : x - \hat{v}_{d'}(q) \geq 0\}$ with $\mathcal{Z}$ given in Example 3.4. In Figure 4.8 we see the outer approximations of degree $d = 8$ (resp. $d = 12$) of the stabilizability region obtained for the choice $d' = 8$. A careful examination reveals that Assumption 2 is slightly violated here, yet this has no effect on the validity of the zero sublevel set approximation. Computing the degree 12 approximation takes a few minutes.

5 Conclusion

In this paper we continued our long haul research program consisting of developing and applying semidefinite programming hierarchies for approximating potentially complicated objects arising in optimization and control with simple objects, namely polynomials of given degrees. The complicated object of interest here was the polynomial abscissa, which has low regularity, while being ubiquitous in linear systems control.

Note that we focused exclusively on the polynomial abscissa, but our techniques readily extend to the polynomial radius (defined as the maximum modulus of the roots), or to any semialgebraic function of the polynomial roots. By semialgebraic function, we mean any function whose graph can be described by finitely many intersections and unions of polynomial sublevel sets or level sets (see e.g. [3]).

In section 3 we described how to construct polynomial upper approximations to the abscissa with guarantees of L^1 convergence (or equivalently almost uniform convergence) on compact sets. Constructing polynomial lower approximations with similar convergence guarantees has proved to be much more challenging. We proposed a first approach in Section 4.1

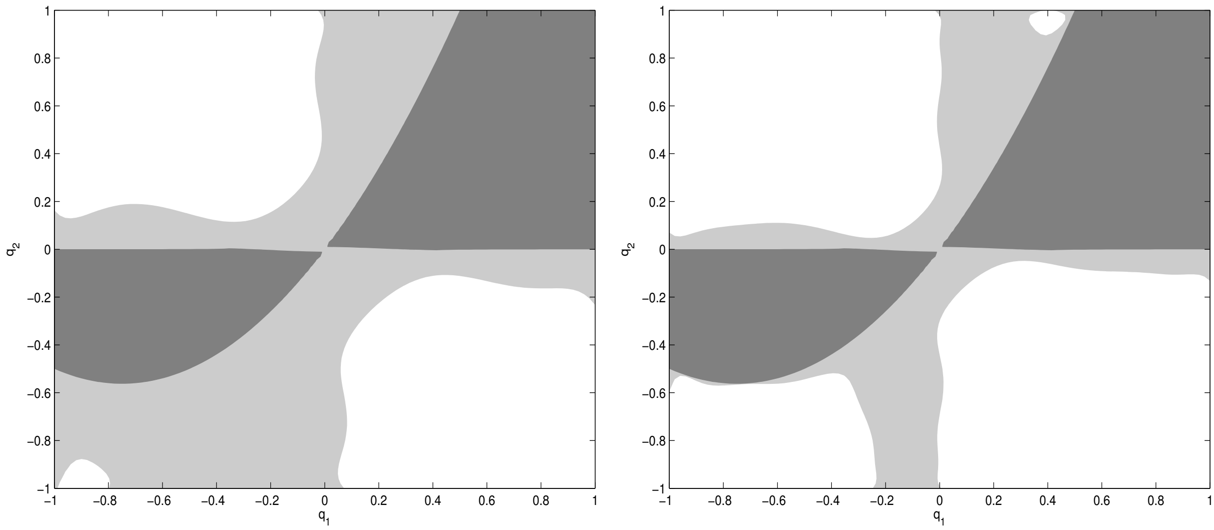

Figure 4.8: Stabilizability region (dark gray region) and its degree 8 outer approximation (light gray region, left) and degree 12 outer approximation (light gray region, right) for Example 4.10 ($d' = 8$). Compare with Figure 4.4.

using elementary symmetric functions which is quite general but also computationally challenging due to the introduction of many lifting variables. This motivated the study of a second approach in Section 4.2 using the Gauß-Lucas theorem which is less computationally demanding, but unfortunately much more involved and subject to working assumptions.

As illustrated by our numerical examples, a shortcoming of our methods is that they can be computationally demanding and not applicable when the degree of the polynomial p and/or the number of parameters in q is large. Moreover it is unknown which level in the semidefinite programming hierarchy can guarantee an a priori given precision level on the abscissa. In terms of complexity, whereas interior-point algorithms can provably solve semidefinite programming problems at given accuracy in polynomial time, it turns out that the number of variables is exponential in the degree of the polynomial p and the number of parameters q, and only the first levels of the semidefinite programming hierarchy can be solved in practice in a reasonable amount of time.

An interesting theoretical question that would deserve careful investigation is whether our L^1 convergence guarantees can be strengthened to L^∞, i.e. to uniform convergence, since we know that the polynomial abscissa is continuous, and hence that it can be uniformly approximated by polynomials on compact sets. For this the semidefinite programming hierarchy should be modified accordingly.

References

[1] R. B. Ash. Probability and measure theory. 2nd edition. Academic Press, San Diego, USA, 2000.

[2] A. Barvinok. A course in convexity. American Mathematical Society, Providence, USA, 2002.

[3] J. Bochnak, M. Coste, M.-F. Roy. Real algebraic geometry. Springer, Berlin, 1998.

[4] J. V. Burke, D. Henrion, A. S. Lewis, M. L. Overton. Stabilization via nonsmooth, nonconvex optimization. IEEE Transactions on Automatic Control 51(11):1760-1769, 2006.

[5] J. V. Burke, A. S. Lewis, M. L. Overton. Variational analysis of the abscissa mapping for polynomials via the Gauss-Lucas theorem. Journal of Global Optimization 28:259-268, 2004.

[6] J. A. Cross. Spectral abscissa optimization using polynomial stability conditions. PhD thesis, University of Washington, Seattle, 2010.

[7] D. Henrion, J. B. Lasserre. Inner approximations for polynomial matrix inequalities and robust stability regions. IEEE Transactions on Automatic Control 57(6):1456-1467, 2012.

[8] D. Henrion, D. Peaucelle, D. Arzelier, M. Šebek. Ellipsoidal approximation of the stability domain of a polynomial. IEEE Transactions on Automatic Control 48(12):2255-2259, 2003.

[9] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, London, UK, 2010.

[10] M. Laurent. Sums of squares, moment matrices and polynomial optimization. In M. Putinar, S. Sullivan (eds.). Emerging applications of algebraic geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications, Springer, Berlin, 2009.

[11] Mosek ApS, Copenhagen, Denmark. `www.mosek.com`

[12] J. Löfberg. YALMIP: A Toolbox for Modeling and Optimization in MATLAB. IEEE CACSD Conference, Taipei, Taiwan, 2004. `users.isy.liu.se/johanl/yalmip/`

[13] V. A. Zorich. Mathematical analysis II. Springer, Berlin, 2004.