MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Parametric Polynomial Spectral Factorization Using the Sum of Roots and Its Application to a Control Design Problem

Hirokazu ANAI, Shinji HARA, Masaaki KANNO,
and Kazuhiro YOKOYAMA

(Communicated by Kazuo MUROTA)

METR 2008–04 January 2008

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t-u-tokyo.ac.jp/research/techrep/index.html

Parametric Polynomial Spectral Factorization Using the Sum of Roots and Its Application to a Control Design Problem

Hirokazu ANAI*, Shinji HARA†, Masaaki KANNO‡, and Kazuhiro YOKOYAMA§

January 25th, 2007

Abstract

This report presents an algebraic approach to polynomial spectral factorization, an important mathematical tool in signal processing and control. The approach exploits an intriguing relationship between the theory of Gröbner bases and polynomial spectral factorization which can be observed through the sum of roots, and allows us to perform polynomial spectral factorization in the presence of real parameters. It is discussed that parametric polynomial spectral factorization enables us to express quantities such as the optimal cost in terms of parameters and the sum of roots. Furthermore an optimization method over parameters is suggested that makes use of the results from parametric polynomial spectral factorization and also employs quantifier elimination. The proposed approach is demonstrated on a numerical example of a particular control problem.

Keywords: Sum of roots, parametric polynomial spectral factorization, parametric optimization, Gröbner basis, quantifier elimination, $\mathcal{H}_2$ control

1 Introduction

In various fields of science and engineering, it is often desired to keep several crucial parameters as variables in the course of analysis and design so that the effect of such parameters may be directly observed. Algebraic computation tools have

*Fujitsu Laboratories Ltd, 4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki, 211-8588, Japan. E-mail: anai@jp.fujitsu.com

†Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan. E-mail: Shinji_Hara@ipc.i.u-tokyo.ac.jp

‡CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi-shi, Saitama, 332-0012, Japan. E-mail: M.Kanno.99@cantab.net

§Department of Mathematics, Rikkyo University, 3-34-1 Nishi Ikebukuro, Toshima-ku, Tokyo, 171-8501, Japan. E-mail: yokoyama@rkmath.rikkyo.ac.jp

been satisfying such desires and been proven of significant use [1]. Not only the capability of exact symbolic manipulation but also sophisticated algebraic methods have started finding their ways in the solution of complicated problems of practical significance that cannot be reliably solved by ordinary numerical approaches. Algebraic approaches are now perceived as effective and promising means.

Spectral factorization is an important mathematical tool in signal processing and control for finite-dimensional linear systems. A number of approaches have been proposed for the solution of spectral factorization. Methods proposed so far are based almost exclusively on standard numerical routines designed for high speed floating point arithmetic. While numerically reliable routines are now available, those approaches cannot deal with systems with parameters, and spectral factorization have been a hindrance for the analysis/design of systems with parameters. It is recently pointed out that the notion of the 'sum of roots' allows us to observe an intriguing relationship between polynomial spectral factorization and the theory of Gröbner bases [2]. This approach is expected to have the potential for the parametric case since the required computation is all algebraic.

This report explores this potential and devises an algebraic algorithm that uses the sum of roots and can deal with parametric polynomial spectral factorization. The result indicates that many analysis/design problems (e.g., optimal design) in signal processing and control can be solved in the presence of parameters, thus allowing engineers to carry out optimization, leaving parameters as they are. Hence postoptimal analysis become doable by means of various kinds of approaches. This report also suggests an optimization method based on quantifier elimination (QE), making use of a nice property of the sum of roots, that chooses the most suited values of parameters. As a demonstration, a particular control problem is employed. A difficulty in such a problem lies in the fact that the optimal cost cannot in general be expressed in closed form in terms of parameters. Instead of trying to find an explicit expression for the cost in parameters alone, the approach employs the sum of roots and finds an algebraic relationship between the plant parameters and the sum of roots, and moreover the approach computes an expression for the cost in terms of plant parameters *and* the sum of roots. The crucial point here is that one quantity (sum of roots) is added only in the expression and that this may make the obtained expression amenable to analysis/optimization that follows. Moreover we suggest a particular QE-based optimization approach that utilizes the obtained algebraic expressions in order to find parameter values that maximize/minimize the optimal cost. In short this report proposes

- a parametric optimization method which expresses the optimal cost in terms of parameters and the sum of roots (i.e., to find the 'best' in the presence of parameters); and
- an algebraic approach that employs the obtained expressions for further optimization over parameters (i.e., to find the 'best of the best').

The sum of roots, initially introduced in [3] as merely an *index of average sta-*

bility, is shown to be an essential quantity in signal processing and control that can directly express performance indices and also that can be utilized for computation. The two algebraic tools, namely, the Gröbner basis and quantifier elimination, have proven to be crucial to visualize the relationship between the sum of roots and spectral factorization and also to achieve mathematically rigorous optimality for the optimization problem over parameters. This revelation may only be made with the help of algebraic geometry, and the computational aspect can only be exploited with the aid of algebraic algorithms.

This report is a full version of [4]. This report includes a number of proofs and detailed explanations which are omitted in [2] and [4] due to space limitation. In addition an extensive discussion on the parametric case is given, which is one of the contributions of the report. Several numerical examples are provided to elucidate the results presented in this report.

The rest of the report is organized as follows. Section 2 reviews the solution of the polynomial spectral factorization problem by means of the sum of roots and further extends it to the parametric case. Section 3 then proposes an optimization algorithm that utilizes the results from parametric polynomial spectral factorization and employs quantifier elimination. In Section 4, the development is summarized in the form of algorithms. Moreover, in order to demonstrate the suggested approach, an $\mathcal{H}_2$ control problem is considered and a numerical example is solved. Some concluding remarks are made in Section 5.

2 Parametric Polynomial Spectral Factorization

In ordinary numerical approaches in control, it is common practice to use the so-called state-space representation of a system and the algebraic Riccati equation is fundamental to analysis and design of systems. Once the solution to a Riccati equation is obtained, the optimal controller etc. can be computed in a straightforward manner using simple matrix arithmetic. In a similar vein, when the transfer function representation of a system is employed, the essential step of analysis/design is the execution of spectral factorization and the rest of the computation is direct [5]. Indeed performing polynomial spectral factorization is another way of solving a Riccati equation and also one of numerical solution approaches to spectral factorization is via the solution of an algebraic Riccati equation [6]. As is pointed out in [2] and also stated in the following, polynomial spectral factorization exhibits an intriguing as well as useful connection to the Gröbner basis theory. We exploit the property to solve problems for systems with parameters in this report.

Firstly we review the problem formulation of polynomial spectral factorization in Subsection 2.1. Then Subsections 2.2 through 2.4 review the solution approach via the sum of roots for the non-parametric case, which is reported in [2] but without proofs/detailed explanations. In this report ideal theoretical fundamentals are given and furthermore proofs and in-depth explanations are provided for completeness. Subsection 2.2 reviews the sum of roots and defines some polynomials that have the sum of roots as one of their roots. In Subsection 2.3, we discuss what we call the ideal of spectral factorization and observe an intriguing relationship between the Gröbner basis theory and polynomial spectral factorization. Subsection 2.4 then investigates characterization of the Gröbner basis of the ideal of spectral factorization. Finally, in Subsection 2.5, we extend the results to the parametric case and gives a detailed discussion on the computation for parametric polynomial spectral factorization. It is noted here that, in the actual algorithms implemented on a computer, every polynomial computation shall be carried out over the rational numbers $\mathbb{Q}$, but that the exposition below assumes computation over the real numbers $\mathbb{R}$. This is because we consider real parameters and also for generality. Readers unfamiliar with the Gröbner basis theory are referred to [7, 8].

2.1 Polynomial Spectral Factorization

Consider the following even polynomial of degree $2n$ in $\mathbb{R}[x]$:

$$f(x) = a_{2n}x^{2n} + a_{2n-2}x^{2n-2} + \cdots + a_2x^2 + a_0 \ , \tag{1}$$

where $a_{2k} \in \mathbb{R}$ for $k = 0, 1, \ldots, n$. It can be assumed without loss of generality that $a_{2n} > 0$. Assume that $f(x)$ has no roots on the imaginary axis. (This assumption naturally arises from the formulation of a wide class of control problems and thus is relevant to a broad range of practical applications. See Subsection 4.2.) If α is a root of $f(x)$, then so is $-\alpha$ because $f(x)$ is an even polynomial. Since $f(x)$ has no imaginary axis root, there are exactly n roots in the open left half plane and n roots in the open right half plane. The task in the polynomial spectral factorization problem is to decompose $f(x)$ into two real polynomials: one that captures all the left half plane roots and its 'mirror image'.

Definition 1 *The* spectral factorization *of $f(x)$ in* (1) *is a decomposition of $f(x)$ of the following form:*

$$a_{2n}f(x) = (-1)^n g(x)g(-x) \ , \tag{2}$$

where

$$g(x) = b_nx^n + b_{n-1}x^{n-1} + \cdots + b_1x + b_0 \in \mathbb{R}[x] \ , \tag{3}$$
$$b_n = a_{2n} \ ,$$

and $g(x)$ has roots in the open left *half plane only. The polynomial $g(x)$ is called the* spectral factor *of $f(x)$.*

2.2 Sum of Roots

This subsection reviews the notion of the sum of roots (SoR), and the results that reveal the relationship between the SoR and polynomial spectral factorization are provided from here through Subsection 2.4. A solution approach to polynomial spectral factorization is given based on this relationship. [2] already reported most of the results in this subsection, but few explanations/proofs are provided. In this report we present proofs and some detailed accounts.

Firstly the relationship between the roots of $f(x)$ and the coefficients of the spectral factor $g(x)$ is investigated. Let $\alpha_1, \dots, \alpha_n$ be the n roots of $f(x)$ in the open *left* half plane. The n roots in the open *right* half plane can be written as $-\alpha_1, \dots, -\alpha_n$. Then, $f(x)$ and $g(x)$ can be expressed as

$$\begin{aligned} f(x) &= a_{2n} \prod_{i=1}^{n} (x - \alpha_i)(x + \alpha_i) = a_{2n} \prod_{i=1}^{n} (x^2 - \alpha_i^2) \ , \\ g(x) &= a_{2n} \prod_{i=1}^{n} (x - \alpha_i) \ , \end{aligned} \tag{4}$$

respectively.

Now the *sum of roots* is defined as the following quantity:

$$\sigma = -(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \ . \tag{5}$$

The name derives from the fact that $-\sigma$ is (literally) the sum of roots of the spectral factor $g(x)$. Since $\mathrm{Re}\,(-\alpha_i) > 0$ and moreover, for each non-real root of $f(x)$, its complex conjugate has the same real part, the following fact is immediate.

Fact 2 ([3]) *The quantity σ is real and positive.*

The quantity σ can in principle be found by computing each individual α_i, $i = 1, \dots, n$. Such an approach is not attractive in that it cannot deal with the parametric case. A way to get some expression for σ without explicitly computing α_i's is sought. Also, by expanding the right hand side of (4) and comparing it with the right hand side of (3), we can immediately see that

$$b_{n-1} = a_{2n}\sigma \ . \tag{6}$$

The question is then whether there is a method to find simple relationships between σ and other coefficients. The results presented in the subsections to follow help us to obtain a polynomial that has σ as one of its roots and also to express other coefficients b_i of the spectral factor $g(x)$ in terms of σ. Until Subsection 2.4, we focus on the case without parameters in the coefficients of $f(x)$. The discussion of the parametric case is deferred until Subsection 2.5. We first define several polynomials which have σ as one of their roots.

Definition 3 ([3]) *Let $\mathcal{P} = \big\{(\epsilon_1, \dots, \epsilon_n) \mid \epsilon_i \in \{1, -1\}\big\}$, and $C(\epsilon_1, \dots, \epsilon_n) = \epsilon_1\alpha_1 + \cdots + \epsilon_n\alpha_n$ for each $(\epsilon_1, \dots, \epsilon_n)$ in $\mathcal{P}$. The* characteristic polynomial $S_f(z)$ of σ *is defined as*

$$S_f(z) = \prod_{(\epsilon_1,\dots,\epsilon_n)\in\mathcal{P}} \big(z - C(\epsilon_1, \dots, \epsilon_n)\big) \ . \tag{7}$$

Also the minimal polynomial $R_f(z)$ of σ *is defined as the square-free part of S_f.*

Using the 'C' notation, σ can be written as

$$\sigma = C(-1, -1, \ldots, -1) \ . \tag{8}$$

Note that S_f belongs to $\mathbb{R}[z]$ and moreover that R_f belongs to $\mathbb{R}[z]$, as $R_f = S_f/\gcd(S_f, \frac{dS_f}{dz})$. Furthermore, in the case of $f(x) \in \mathbb{Q}[x]$, S_f and R_f belong to $\mathbb{Q}[z]$, as well (while $g(x) \notin \mathbb{Q}[x]$ in general).

The following lemma states a characterization of the SoR σ as a root of $R_f(z)$ (or $S_f(z)$).

Lemma 4 ([2]) *The SoR σ defined in* (5) *coincides with the largest real root of $R_f(z)$ (or $S_f(z)$). Moreover, under the assumption that $f(x)$ does not have an imaginary axis root, σ is always a simple root.*

Here we give a proof which is omitted in [2] due to space limitation.

Proof: Since $\operatorname{Re} \alpha_i < 0$, if ϵ_1, say, in (8) is changed from -1 to 1, the real part of

$$C(1, -1, \ldots, -1) = \alpha_1 - \alpha_2 - \cdots - \alpha_n = \sigma + 2\alpha_1$$

is smaller than that of σ. Any number of sign changes of ϵ_i (from -1 to 1) thus makes the real part of C smaller. Therefore, σ has the largest real part among the roots of $R_f(z)$ (or $S_f(z)$). This along with Fact 2 implies that σ is the largest real roof of $R_f(z)$ (or $S_f(z)$). What is more, since α_i is in the *open* left half plane and the real part of α_i is strictly smaller than 0, σ is always strictly larger than the real part of any C from $\mathcal{P} \setminus (-1, -1, \ldots, -1)$. This implies that σ is a simple root of $R_f(z)$ (or $S_f(z)$). $\square$

This rather simple fact is nevertheless extremely beneficial because it enables us to focus on the largest real root only. A general case would require that all the roots should be found first and those candidates should be examined in order to find the 'true' solution; in contrast we know beforehand which one to find. This property is exploited appropriately in optimization over parameters in Subsection 3.2.

2.3 Ideal of Spectral Factorization

We now investigate the characteristics of the ideal of polynomials in the coefficients b_i of the spectral factor $g(x)$ which are immediately obtained from the formulation of polynomial spectral factorization. Denote $\{b_0, \ldots, b_{n-1}\}$ by $\mathbf{B}$. Firstly the following observation is the seminal point of the whole development.

Lemma 5 ([2]) *Given $f(x)$ and $g(x)$ as in* (1) *and* (3)*, respectively, consider b_i, $i = 0, \ldots, n-1$, as variables. A system of algebraic equations in terms of b_i's is* *obtained by comparing the coefficients of* (2)*:*

$$(-1)b_{n-1}^2 + 2a_{2n}b_{n-2} - a_{2n}a_{2(n-1)} = 0 \ ,$$
$$(-1)^2 b_{n-2}^2 - 2b_{n-1}b_{n-3} + 2a_{2n}b_{n-4} - a_{2n}a_{2(n-2)} = 0 \ ,$$
$$\vdots$$
$$(-1)^k b_{n-k}^2 + \sum_{\substack{1\le i\le 2k-1\\ i\neq k}} (-1)^i b_{n-i}b_{n-2k+i} + 2a_{2n}b_{n-2k} - a_{2n}a_{2(n-k)} = 0$$
$$\textit{for} \ \ 2k \le n \ ,$$
$$\vdots$$
$$(-1)^k b_{n-k}^2 + \sum_{\substack{2k-n\le i\le n\\ i\neq k}} (-1)^i b_{n-i}b_{n-2k+i} - a_{2n}a_{2(n-k)} = 0 \ \ \textit{for} \ \ 2k > n \ ,$$
$$\vdots$$
$$(-1)^{n-1}b_1^2 + (-1)^{n-2}2b_2b_0 - a_{2n}a_2 = 0 \ ,$$
$$(-1)^n b_0^2 - a_{2n}a_0 = 0 \ .$$

Then the set $\mathcal{G}$ of the polynomials obtained from the polynomial parts of the equations (i.e., the left hand sides of the above equations) forms the reduced Gröbner basis (up to sign) of the ideal generated by itself in $\mathbb{R}[\mathbf{B}]$ with respect to the graded reverse lexicographic order $b_{n-1} \succ \cdots \succ b_0$.

The result is reported in [2] without a proof. Here a proof is also provided for completeness.

Proof: We firstly show that $\mathcal{G}$ is a Gröbner basis. Each element g_i in $\mathcal{G}$ is of the form

$$g_i := b_i^2 + \underbrace{\text{(sum of monomials of total degree 2 consisting of distinct } b_k\text{'s)}}_{W}$$
$$+ \text{ (a monomial of total degree 1 and a constant term)}$$

(ignoring the sign of the coefficient of b_i^2), and any monomial in W contains variable $b_\ell \prec b_i$, which implies that b_i^2 is the leading monomial. Also the leading monomials of any pair of polynomials in $\mathcal{G}$ are prime to each other. Thus Buchberger's criterion [8, Section 2-2, Theorem 6] is satisfied and $\mathcal{G}$ is a Gröbner basis of $\langle \mathcal{G} \rangle$.

Now, for any $g_i \in \mathcal{G}$,

$$\langle \mathrm{LT}(\mathcal{G} - \{g_i\}) \rangle = \{b_{n-1}^2, b_{n-2}^2, \ldots, b_0^2\} \setminus \{b_i^2\} \ , \tag{9}$$

where $\mathrm{LT}(\cdot)$ denotes the set of leading terms of elements. It is immediate that no monomial of g_i lies in (9). That is, $\mathcal{G}$ satisfies the definition of the reduced Gröbner basis [8, Section 2-2, Definition 5] and is thus the reduced Gröbner basis, which concludes the proof. $\square$

We call the ideal $\langle \mathcal{G} \rangle$ of $\mathbb{R}[\mathbf{B}]$ the *ideal of spectral factorization*. As the set of the leading monomials of the elements of $\mathcal{G}$ is $\{b_{n-1}^2, b_{n-2}^2, \ldots, b_0^2\}$,

$$\mathcal{LB} := \left\{ b_0^{k_0} b_1^{k_1} \cdots b_{n-1}^{k_{n-1}} \mid k_i \in \{0, 1\} \right\}$$

forms a basis of the residue class ring $\mathbb{R}[\mathbf{B}]/\langle \mathcal{G} \rangle$ as an $\mathbb{R}$-linear space, and $\dim_{\mathbb{R}} \mathbb{R}[\mathbf{B}]/\langle \mathcal{G} \rangle = \#\mathcal{LB} = 2^n$. Moreover each zero $(\beta_{n-1}, \ldots, \beta_0)$ of $\langle \mathcal{G} \rangle$ corresponds to some $C(\epsilon_1, \ldots, \epsilon_n)$ in that $a_{2n}x^n + \beta_{n-1}x^{n-1} + \cdots + \beta_1 x + \beta_0 = a_{2n} \prod_{i=1}^{n} (x - \epsilon_i \alpha_i)$. Thus the following lemma can be deduced.

Lemma 6 ([2]) *The ideal of spectral factorization is* 0 *dimensional and the number of its zeros with multiplicities counted is* 2^n.

If $f(x)$ has no multiple roots, then there are exactly 2^n distinct zeros of $\langle \mathcal{G} \rangle$ and, moreover, $\langle \mathcal{G} \rangle$ is radical. In this situation, there are 2^n different $g(x)$ satisfying (2) (but ignoring the root location requirement). There is however only one 'true' $g(x)$ that meets the requirement, and that particular $g(x)$ corresponds to the largest real root of $R_f(z)$ (or $S_f(z)$); remember that the SoR is the largest real root of $R_f(z)$ (or $S_f(z)$). With regard to the system of equations stated in Lemma 5, what we seek is the solution with the largest real b_{n-1}.

2.4 Shape Basis of the Ideal of Spectral Factorization

We have seen that the formulation of polynomial spectral factorization directly gives a Gröbner basis of the ideal of spectral factorization. Now we are in the position of discussing another Gröbner basis which shows the relationship between the SoR and the coefficients of the spectral factor.

Before going into the main part, some polynomials related to the ideal of spectral factorization and situations differentiating the relationship among these polynomials are defined. The *characteristic polynomial* $\hat{S}_f(y)$ (resp., the *minimal polynomial* $\hat{N}_f(y)$) of b_{n-1} modulo $\langle \mathcal{G} \rangle$ can be defined as the characteristic polynomial (resp., the minimal polynomial) of the linear map derived from the multiplication map [9]:

$$\mathbb{R}[\mathbf{B}]/\langle \mathcal{G} \rangle \ni g \to b_{n-1} g \in \mathbb{R}[\mathbf{B}]/\langle \mathcal{G} \rangle \ .$$

Then, $\hat{S}_f(y)$ has $\hat{N}_f(y)$ as its factor and also their square-free parts coincide with $R_f(y/a_{2n})$ (remember the relationship (6)). Moreover, by considering each root α_i as a variable, we can show that $\hat{S}_f(y)$ coincides with $S_f(y/a_{2n})$; this point will be discussed in detail before Theorem 14.

Definition 7 *Given $f(x)$, when distinct $(\epsilon_1, \ldots, \epsilon_n) \in \mathcal{P}$ give distinct $\epsilon_1\alpha_1 + \cdots + \epsilon_n\alpha_n$, we call the situation a* generic case. *Otherwise it is called a* singular case.

It is noted that almost all $f(x)$ arising from practical applications fall into the generic case.

The notion of 'generic case' is closely related to that of 'separating element modulo ideal'. (See [10] for details on the separating element.)

Definition 8 *Let $\mathcal{I}$ be a 0 dimensional ideal in a polynomial ring $\mathbb{K}[\mathbf{X}]$ in variables $\mathbf{X}$ over a field $\mathbb{K}$. A polynomial $h(\mathbf{X})$ is called a* separating element *if $h(\gamma) \neq h(\gamma')$ for any distinct pair γ, γ' in $V_{\mathbb{L}}(\mathcal{I})$. Here we denote by $V_{\mathbb{L}}(\mathcal{I})$ the affine variety of $\mathcal{I}$ in $\mathbb{L}$, that is, the set of all distinct zeros of $\mathcal{I}$ in an algebraically closed field $\mathbb{L}$ containing $\mathbb{K}$. (In our setting, if $\mathbb{L}$ can be inferred trivially, we use V instead of $V_{\mathbb{L}}$.)*

Because, in the generic case, S_f is square-free and thus $S_f = R_f$, which also implies $\hat{S}_f(y) = \hat{N}_f(y) = S_f(y/a_{2n})$. Then it is immediate that b_{n-1} is a separating element, as the ideal of spectral factorization has at most 2^n distinct zeros. Moreover in this case the ideal is radical as f cannot have multiple roots.

Remark 9 *If $f(x)$ has multiple roots and thus the number of distinct zeros of $\langle \mathcal{G} \rangle$ is strictly less than 2^n, then it is a singular case. However the converse is not true. That is, a singular case may occur even if $f(x)$ does not have multiple roots. This is because the generic/singular case deals with the roots of S_f only (i.e., deals with b_{n-1} only), while the zeros of the ideal $\langle \mathcal{G} \rangle$ considers the n-tuple $(\beta_{n-1}, \ldots, \beta_0)$ (i.e., considers all b_i). This difference is illustrated in the numerical example at the end of this subsection.*

Due to the facts that $\langle \mathcal{G} \rangle$ is 0 dimensional and radical and that b_{n-1} is a separating element, we can get a special Gröbner basis called the *shape basis*. More formally:

Theorem 10 ([2]) *In the generic case the ideal of spectral factorization has a Gröbner basis of so-called* shape form *with respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$:*

$$\mathcal{F} := \left\{ \hat{S}_f(b_{n-1}),\ b_{n-2} - \hat{h}_{n-2}(b_{n-1}),\ \ldots,\ b_0 - \hat{h}_0(b_{n-1}) \right\},$$

where $\hat{S}_f$ is a polynomial of degree exactly 2^n and $\hat{h}_i$'s are polynomials of degree strictly less than 2^n.

With respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$, the minimal polynomial $\hat{N}_f(b_{n-1})$ appears as the first element of the reduced Gröbner basis. In the generic case, since $\hat{S}_f = \hat{N}_f$, $\hat{S}_f$ appears as the first element.

The theorem states that all coefficients b_i of the spectral factor $g(x)$ can thus be described as polynomials in b_{n-1} and therefore that the polynomial spectral factorization problem reduces to finding the largest real root of $\hat{S}_f(y)$. This result along with (6) implies that there is a polynomial of degree 2^n that defines the SoR σ and that each coefficient of $g(x)$ is described as a polynomial in σ:

$$S_f(\sigma) = 0\ ,\ \ b_{n-1} = a_{2n}\sigma\ ,\ \ b_{n-2} = h_{n-2}(\sigma)\ ,\ \ \ldots\ ,\ \ b_0 = h_0(\sigma)\ ,$$

where $S_f(\sigma) := \hat{S}_f(a_{2n}\sigma)$ and $h_i(\sigma) := \hat{h}_i(a_{2n}\sigma)$. Also, due to the symmetricity, $S_f(\sigma)$ turns out to be a polynomial in σ^2 in the generic case. In general we can efficiently compute a shape basis from the set $\mathcal{G}$ of polynomials by means of the basis conversion (change-of-order) technique [10].

The *singular case*, where $\hat{S}_f$ has multiple roots, happens when, for instance, $f(x)$ has multiple roots. In such a case, the Gröbner basis of shape form may not immediately be computable. However, by adding the 'simple part' of $\hat{S}_f$, we can have a polynomial set of shape form. Let $\hat{T}_f(y)$ be the factor of $\hat{S}_f(y)$ (or $\hat{N}_f(y)$) obtained as the product of $y - \gamma_i$ for all simple roots γ_i's of $\hat{S}_f(y)$. Then the polynomial $\hat{T}_f$ can be computed via square-free factorization. Indeed, letting $\hat{U}_1(y) = \gcd(\hat{S}_f(y), d\hat{S}_f/dy)$ and $\hat{U}_2(y) = \gcd(\hat{U}_1(y), d\hat{U}_1(y)/dy)$, we have $\hat{T}_f = \hat{S}_f\hat{U}_2/\hat{U}_1^2$. Here we call $\hat{T}_f$ the *simple part* of $\hat{S}_f$.

Consider the ideal $\mathcal{J} = \langle \mathcal{G}, \hat{T}_f(b_{n-1})\rangle$. For each root γ of $\hat{T}_f$, γ is a simple root of $\hat{S}_f$. This implies that the system of equations in Lemma 5 with $b_{n-1} = \gamma$ has a unique solution corresponding to $g(x) = a_{2n}\prod_{i=1}^{n}(x - \epsilon_i\alpha_{j_i})$ with $\gamma = a_{2n}(\epsilon_1\alpha_1 + \cdots + \epsilon_n\alpha_n)$. Then it follows that b_{n-1} is again a separating element with respect to the ideal $\mathcal{J}$. Since $\hat{T}_f(b_{n-1})$ is the minimal polynomial of b_{n-1} modulo $\mathcal{J}$, $\mathcal{J}$ is radical and has a Gröbner basis of shape form with respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$. Also remember that, under the assumption that there is no imaginary axis roots in $f(x)$, the SoR σ is always a simple root of S_f and $a_{2n}\sigma$ is a simple root of $\hat{S}_f$ (Lemma 4). Therefore, $\hat{T}_f$ has $a_{2n}\sigma$ as its root and $\langle \mathcal{G} \cup \{\hat{T}_f(b_{n-1})\}\rangle$ has the zero yielding the true spectral factor of f. The discussion here is summarized in the form theorem.

Theorem 11 *The ideal $\langle \mathcal{G} \cup \{\hat{T}_f(b_{n-1})\}\rangle$ has a Gröbner basis of shape form with respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$:*

$$\left\{\hat{T}_f(b_{n-1}),\ b_{n-2} - \bar{h}_{n-2}(b_{n-1}),\ \ldots,\ b_0 - \bar{h}_0(b_{n-1})\right\}\ , \tag{10}$$

where $\bar{h}_i$'s are polynomials of degree strictly less than that of $\hat{T}_f$. Moreover the ideal $\langle \mathcal{G} \cup \{\hat{T}_f(b_{n-1})\}\rangle$ has the zero yielding the true spectral factor of f.

Remark 12 *Even if f has imaginary axis roots, Theorems 10 and 11 still hold except for the condition that those ideals have the zero yielding the spectral factorization.*

Here a numerical example is presented to illustrate some points made so far. Consider the following even polynomial:

$$\begin{aligned} f(x) &= 2x^6 - 28x^4 + 98x^2 - 72 \\ &= 2(x+1)(x-1)(x+2)(x-2)(x+3)(x-3)\ . \end{aligned}$$

This polynomial does not have an imaginary axis root and thus polynomial spectral factorization is possible (in fact it has an obvious spectral factor). Also notice that there is no multiple roots in $f(x)$. We employ the approach developed above. First write the spectral factor $g(x)$ as

$$g(x) = 2x^3 + b_2x^2 + b_1x + b_0 \ .$$

Comparing the coefficients of the both sides of (2), we get a set of polynomial equations:

$$\begin{cases} b_0^2 - 144 = 0 \ , \\ b_1^2 - 2b_0b_2 - 196 = 0 \ , \\ b_2^2 - 4b_1 - 56 = 0 \ . \end{cases}$$

Lemma 5 states that the polynomial parts $\mathcal{G}$ (i.e., the left hand sides of the above equations) forms the reduced Gröbner basis of the ideal generated by itself with respect to the graded reverse lexicographic order $b_2 \succ b_1 \succ b_0$. Since $f(x)$ has no multiple roots, there are exactly $2^3 = 8$ distinct $g(x)$ satisfying (2) (but not necessarily fulfilling the root location requirement). Since each $g(x)$ corresponds to a zero of the ideal $\langle \mathcal{G} \rangle$, $\langle \mathcal{G} \rangle$ has $2^3 = 8$ distinct zeros:

$$\begin{aligned}(b_2, b_1, b_0) =&(0, -14, 12),\ (0, -14, -12),\ (4, -10, -12),\ (-4, -10, 12),\\ &(8, 2, -12),\ (-8, 2, 12),\ (12, 22, 12),\ (-12, 22, -12).\end{aligned} \tag{11}$$

We can see that any of the zero is simple.

If we simply compute a Gröbner basis with respect to, e.g., the pure lexicographic order $b_0 \succ b_1 \succ b_2$, we get

$$\begin{aligned}\{&b_2^7 - 224b_2^5 + 12544b_2^3 - 147456b_2,\\ &\qquad b_1 - \tfrac{1}{4}b_2^2 + 14, b_0b_2 - \tfrac{1}{32}b_2^4 + \tfrac{7}{2}b_2^2, b_0^2 - 144\} \ .\end{aligned} \tag{12}$$

Note the degree of the first polynomial (in b_2 only); it is 7, which is smaller than $2^3 = 8$. This happens because two different 3-tuples of $(\epsilon_1, \epsilon_2, \epsilon_3)$ give an identical value of C:

$$\begin{aligned}C(1, 1, -1) &= (-1) + (-2) - (-3) = 0\\ &= -(-1) - (-2) + (-3) = C(-1, -1, 1) \ ,\end{aligned}$$

where $\alpha_1 = -1, \alpha_2 = -2, \alpha_3 = -3$. Indeed this 'singularity' appears in the zeros of $\langle \mathcal{G} \rangle$ in (11); $b_2 = 0$ is contained in the two zeros. Computing the characteristic polynomial of b_2, we get

$$\begin{aligned}\hat{S}_f(b_2) &= b_2^8 - 224b_2^6 + 12544b_2^4 - 147456b_2^2\\ &= b_2^2(b_2 - 4)(b_2 + 4)(b_2 - 8)(b_2 + 8)(b_2 - 12)(b_2 + 12) \ .\end{aligned}$$

Adding to $\langle \mathcal{G} \rangle$ a polynomial $\hat{T}_f(b_2)$ computed from the simple roots of $\hat{S}_f(b_2)$,

$$\begin{aligned} \hat{T}_f(b_2) &= b_2^6 - 224b_2^4 + 12544b_2^2 - 147456 \\ &= (b_2-4)(b_2+4)(b_2-8)(b_2+8)(b_2-12)(b_2+12) \ , \end{aligned}$$

and computing the shape basis of $\langle \mathcal{G}, \hat{T}_f(b_2) \rangle$, we can get a Gröbner basis of shape form:

$$\left\{ b_2^6 - 224b_2^4 + 12544b_2^2 - 147456, b_1 - \tfrac{1}{4}b_2^2 + 14, b_0 - \tfrac{1}{32}b_2^3 + \tfrac{7}{2}b_2 \right\} . \quad (13)$$

The largest real root of $\hat{T}_f(b_2)$ is $b_2 = 12$, and the corresponding zero of (13) is

$$(b_2, b_1, b_0) = (12, 22, 12) \ ,$$

which gives the (correct) spectral factor

$$g(x) = 2x^3 + 12x^2 + 22x + 12 \ .$$

2.5 Parametric Case

This subsection deals with *our main target where each coefficient a_{2k} is some polynomial in parameters $\mathbf{q} = (q_1, \ldots, q_m)$ over $\mathbb{Q}$*. Even in the parametric case, it often happens that the ideal of spectral factorization is generic for almost all combinations of parameter values. Nevertheless we need to pay special attention to singular situations so that analysis/optimization that follows may be carried out thoroughly. It is shown here that such singularities can also be dealt with. To do so, the notion of 'comprehensive Gröbner system' is crucial and we can apply several techniques for its computation [11, 12, 13, 14]. However our situation is very special, as what we need to do is 'basis conversion' only, that is, transforming a Gröbner basis to another one with respect to a different order. In such computation, computation over the *rational function field in parameters* is in general straightforward and effective.

From now on we regard the even polynomial $f(x)$ as a multivariate polynomial $f(x, \mathbf{q})$ in $\mathbb{Q}[x, \mathbf{q}]$. For each element $\mathbf{c} = (c_1, \ldots, c_m)$ in $\mathbb{R}^m$, we denote by $\varphi_{\mathbf{c}}$ the ring homomorphism from $\mathbb{Q}[\mathbf{q}][\mathbf{B}]$ to $\mathbb{R}[\mathbf{B}]$ obtained by substitution of $\mathbf{q}$ with $\mathbf{c}$. Moreover, for simplicity, we denote by $f_{\mathbf{c}}(x)$ the polynomial $\varphi_{\mathbf{c}}(f)$ which is obtained from $f(x, \mathbf{q})$ by substituting the parameters $\mathbf{q}$ with $\mathbf{c} \in \mathbb{R}^m$.

2.5.1 Regular Region

To perform spectral factorization, we consider a semi-algebraic set $\mathtt{C} \subset \mathbb{R}^m$ such that, for any $\mathbf{c} \in \mathtt{C}$, $a_{2n}(\mathbf{c}) \neq 0$ and there exist no imaginary axis roots (equivalently, the number of roots of $f_{\mathbf{c}}(x)$ with positive real parts is n).

Definition 13 *A semi-algebraic set $\mathtt{C} \subset \mathbb{R}^m$ is called a* regular region *if, for any $\mathbf{c} \in \mathtt{C}$, $a_{2n}(\mathbf{c}) \neq 0$ and there exist no imaginary axis roots in $f_{\mathbf{c}}(x)$.*

The condition that f has no imaginary axis roots can be computed by the *quantifier elimination technique* or *real root counting* methods (e.g., the Sturm-Habicht sequence [15]). Consider the polynomial $\tilde{f}(x, \mathbf{q}) = f(\sqrt{-1}\, x, \mathbf{q}) \in \mathbb{Q}[x, \mathbf{q}]$. For $\mathbf{c} \in \mathbb{R}^m$, the number of imaginary axis roots of $f_\mathbf{c}$ coincides with the number of real roots of $\varphi_\mathbf{c}(\tilde{f})$. Thus the condition on $\mathbf{c}$ such that $f_\mathbf{c}$ has no such roots is equivalent to the condition that $\varphi_\mathbf{c}(\tilde{f})$ is positive for all $x \in \mathbb{R}$.

As an example, we consider

$$f(x) = a_6x^6 + a_4x^4 + a_2x^2 + a_0 ,$$

where a_6, a_4, a_2, a_0 are parameters and $a_6 > 0$. By the quantifier elimination technique, the condition that f has no imaginary axis roots is derived to be

$$a_0 < 0 \ \wedge \ \big(27a_6^2a_0^2 - 18a_6a_4a_2a_0 + 4a_4^3a_0 + 4a_6a_2^3 - a_4^2a_2^2 > 0 \\ \vee \ (a_4 < 0 \ \wedge \ a_2 > 0) \big) .$$

2.5.2 Parametric Basis Conversion

Now assume that $\mathtt{C}$ is a regular region for $f(x, \mathbf{q})$. We can compute the polynomial set $\mathcal{G}$ from (2) and (3), where all the polynomials are treated as ones over $\mathbb{Q}[\mathbf{q}] \subset \mathbb{Q}(\mathbf{q})$. Also, for each $\mathbf{c} \in \mathtt{C}$, we can compute the polynomial set $\mathcal{G}_\mathbf{c}$ from (2) and (3) with parameters $\mathbf{q}$ substituted by $\mathbf{c}$. Then, $\mathcal{G}_\mathbf{c} = \varphi_\mathbf{c}(\mathcal{G})$. Thus, $\mathcal{G}$ is a unique component of the comprehensive Gröbner systems with respect to the graded reverse lexicographic order $b_{n-1} \succ \cdots \succ b_0$.

We now consider the ideal $\langle \mathcal{G} \rangle$ of spectral factorization in $\mathbb{Q}(\mathbf{q})[\mathbf{B}]$ and the ideal $\langle \mathcal{G}_\mathbf{c} \rangle$ of spectral factorization in $\mathbb{R}[\mathbf{B}]$. We note that all arguments in the previous subsections can be applied to the ideals in $\mathbb{Q}(\mathbf{q})[\mathbf{B}]$, as $\mathbb{Q}(\mathbf{q})$ is a field.

Let us consider the characteristic polynomial $\hat{S}_f(y)$ of b_{n-1}, which shall be the first element of the shape basis. Since all the leading coefficients of the elements of $\mathcal{G}$ are constant, $\mathcal{LB} = \{b_0^{k_0} \cdots b_{n-1}^{k_{n-1}} \mid k_i \in \{0, 1\}\}$ is still a linear basis for $\mathbb{Q}(\mathbf{q})[\mathbf{B}]/\langle \mathcal{G} \rangle$. Considering the linear map derived from the multiplication map $\mathbb{Q}(\mathbf{q})[\mathbf{B}]/\langle \mathcal{G} \rangle \ni g \to b_{n-1}g \in \mathbb{Q}(\mathbf{q})[\mathbf{B}]/\langle \mathcal{G} \rangle$, we can show that the matrix representation $M_\mathbf{q}$ of the linear map with respect to $\mathcal{LB}$ is a matrix over $\mathbb{Q}[\mathbf{q}]$. Since the characteristic polynomial $\hat{S}_f$ is the determinant of $yE - M_\mathbf{q}$, where E denotes the identity matrix, $\hat{S}_f$ is a polynomial in y over $\mathbb{Q}[\mathbf{q}]$. In the same manner, for each $\mathbf{c} \in \mathtt{C}$, we can compute the characteristic polynomial $\hat{S}_{f_\mathbf{c}}$ as the characteristic polynomial of the matrix $M_\mathbf{c}$ derived from the linear map. Then, $M_\mathbf{c}$ coincides with the matrix obtained from $M_\mathbf{q}$ by substituting $\mathbf{q}$ with $\mathbf{c}$, and thus $\varphi_\mathbf{c}(\hat{S}_f) = \hat{S}_{f_\mathbf{c}}$. The above discussion leads to the following theorem.

Theorem 14 *The characteristic polynomial $\hat{S}_f$ is a monic polynomial over $\mathbb{Q}[\mathbf{q}]$, and, for each $\mathbf{c} \in \mathtt{C}$, the characteristic polynomial $\hat{S}_{f_\mathbf{c}}$ can be computed by $\hat{S}_{f_\mathbf{c}}(b_{n-1}) = \varphi_\mathbf{c}\big(\hat{S}_f(b_{n-1})\big)$.*

Again there are the *generic case* and the *singular case*. In the generic case, $\hat{S}_f(y)$ is square-free over $\mathbb{Q}(\mathbf{q})$ and the ideal $\langle \mathcal{G} \rangle$ is radical. This situation corresponds to Definition 7 in Subsection 2.4. In this case the ideal $\langle \mathcal{G} \rangle$ in $\mathbb{Q}(\mathbf{q})[\mathbf{B}]$ has a shape basis $\mathcal{F}$ with respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$, as in Theorem 10:

$$\mathcal{F} = \left\{\hat{S}_f(b_{n-1}),\ b_{n-2} - \hat{h}_{n-2}(b_{n-1}),\ \ldots,\ b_0 - \hat{h}_0(b_{n-1})\right\} ,$$

where $\hat{h}_i \in \mathbb{Q}(\mathbf{q})[y]$ for $i = 0, 1, \ldots, n-2$.

For the singular case, we compute the simple part $\hat{T}_f$ of $\hat{S}_f$ by using GCD computation, and consider the ideal $\langle \mathcal{G} \cup \{\hat{T}_f\} \rangle$. Then Theorem 11 implies that it has a shape basis $\bar{\mathcal{F}}$ with respect to any elimination ordering $\{b_0, \ldots, b_{n-2}\} \succ\succ b_{n-1}$:

$$\bar{\mathcal{F}} = \left\{\hat{T}_f(b_{n-1}),\ b_{n-2} - \bar{h}_{n-2}(b_{n-1}),\ \ldots,\ b_0 - \bar{h}_0(b_{n-1})\right\} ,$$

where $\bar{h}_i \in \mathbb{Q}(\mathbf{q})[y]$ for $i = 0, 1, \ldots, n-2$. Moreover, for any $\mathbf{c} \in \mathtt{C}$, one of the roots of $\varphi_{\mathbf{c}}(\hat{T}_f)$ is $\varphi_{\mathbf{c}}(a_{2n}) \times \sigma$ (where σ is the SoR of $f_{\mathbf{c}}$) since $\varphi_{\mathbf{c}}(a_{2n}) \times \sigma$ is a simple root of $\hat{S}_{f_{\mathbf{c}}} = \varphi_{\mathbf{c}}(\hat{S}_f)$. Thus, $\varphi_{\mathbf{c}}(\mathcal{G} \cup \{\hat{T}_f\})$ has the zero yielding the spectral factor of $f_{\mathbf{c}}$.

Now we investigate the generic case in detail. Let $\hat{d}_i(\mathbf{q})$ be the denominator of $\hat{h}_i$. Then both $\mathcal{G}$ and $\mathcal{F}$ are Gröbner bases, and by using the argument in [8, Section 6-3, Exercises 7], the following theorem can be shown.

Theorem 15 *For $\mathbf{c} = (c_1, \ldots, c_m) \in \mathtt{C}$, if $\varphi_{\mathbf{c}}(d_i) \neq 0$ for any i, then there is the zero in $V(\varphi_{\mathbf{c}}(\mathcal{F}))$ which yields the spectral factor of $f_{\mathbf{c}}$. Moreover, in this case, $\varphi_{\mathbf{c}}(\mathcal{F})$ forms a Gröbner basis of $\langle \mathcal{G}_{\mathbf{c}} \rangle$.*

Proof: We show that $\varphi_{\mathbf{c}}(\mathcal{F})$ forms a Gröbner basis of $\langle \mathcal{G}_{\mathbf{c}} \rangle$, from which the first statement directly follows. As $\mathcal{F}$ is a Gröbner basis, each g_i in $\mathcal{G}$ has its standard representation:

$$g_i = \sum_{j=1}^{n} r_{i,j} f_j ,$$

where f_i denotes the i-th element of $\mathcal{F}$. As $r_{i,j}$ can be computed by sequential reduction of f_j's, the denominator of $r_{i,j}$ is a factor of some products of $\hat{d}_0, \ldots, \hat{d}_{n-2}$ and thus it does not vanish at any $\mathbf{c}$ in $\mathtt{C} \setminus V\left(\prod_{i=1}^{n-2} \hat{d}_i\right) = \mathtt{C} \setminus \left(V(\hat{d}_0) \cup \cdots \cup V(\hat{d}_{n-2})\right)$. This implies that

$$\mathcal{G}_{\mathbf{c}} = \varphi_{\mathbf{c}}(\mathcal{G}) \subset \langle \varphi_{\mathbf{c}}(\mathcal{F}) \rangle .$$

Also, repeating the same argument, we can show that

$$\varphi_{\mathbf{c}}(\mathcal{F}) \subset \langle \mathcal{G}_{\mathbf{c}} \rangle .$$

Thus we have $\langle \mathcal{G}_{\mathbf{c}} \rangle = \langle \varphi_{\mathbf{c}}(\mathcal{F}) \rangle$, which concludes the proof. □

Next we deal with the singular case. Let $\bar{d}_i(\mathbf{q})$ be the denominator of $\bar{h}_i$. In this case, $\mathcal{G} \cup \{\hat{T}_f\}$ is not a Gröbner basis. However what we need is the zero of $\mathcal{G}_\mathbf{c}$ that corresponds to the spectral factor of $f_\mathbf{c}$.

Theorem 16 *For $\mathbf{c} = (c_1, \ldots, c_m) \in \mathtt{C}$, if $\varphi_\mathbf{c}(\bar{d}_i) \neq 0$ for any i, then there is the zero in $V\big(\varphi_\mathbf{c}(\bar{\mathcal{F}})\big)$ which gives the spectral factor of $f_\mathbf{c}$. Moreover, if $\varphi_\mathbf{c}(\hat{T}_f)$ is the simple part of $\varphi_\mathbf{c}(\hat{S}_f)$, then $\varphi_\mathbf{c}(\mathcal{F})$ forms a Gröbner basis of $\langle \mathcal{G}_\mathbf{c} \rangle$.*

Proof: From the assumption that $\varphi_\mathbf{c}(\bar{d}_i) \neq 0$ for any i and by the same argument as in the first half of the proof of Theorem 15, it follows that

$$\mathcal{G}_\mathbf{c} \cup \{\varphi_\mathbf{c}(\hat{T}_f)\} = \varphi_\mathbf{c}(\mathcal{G} \cup \{\hat{T}_f\}) \subset \langle \varphi_\mathbf{c}(\bar{\mathcal{F}}) \rangle \ .$$

This implies that $V(\varphi_\mathbf{c}(\mathcal{F})) \subset V(\mathcal{G}_\mathbf{c} \cup \{\varphi_\mathbf{c}(\hat{T}_f)\})$. As the zero yielding the spectral factor of $f_\mathbf{c}$ comes uniquely from the largest real root of $\varphi_\mathbf{c}(\hat{T}_f)$, there is the zero in $V\big(\varphi_\mathbf{c}(\bar{\mathcal{F}})\big)$ which gives the spectral factor of $f_\mathbf{c}$.

Moreover, if $\varphi_\mathbf{c}(\hat{T}_f)$ is the simple part of $\varphi_\mathbf{c}(\hat{S}_f)$, then the ideals $\langle \varphi_\mathbf{c}(\bar{\mathcal{F}}) \rangle$ and $\langle \mathcal{G}_\mathbf{c} \cup \{\varphi_\mathbf{c}(\hat{T}_f)\} \rangle$ are radical and the numbers of those zeros coincide with the degree of $\varphi_\mathbf{c}(\hat{T}_f)$. Hence, $V(\varphi_\mathbf{c}(\mathcal{F})) = V(\mathcal{G}_\mathbf{c} \cup \{\varphi_\mathbf{c}(\hat{T}_f)\})$ and so $\langle \varphi_\mathbf{c}(\mathcal{F}) \rangle = \langle \mathcal{G}_\mathbf{c} \cup \{\varphi_\mathbf{c}(\hat{T}_f)\} \rangle$. This implies that $\varphi_\mathbf{c}(\mathcal{F})$ forms a Gröbner basis of $\langle \mathcal{G}_\mathbf{c} \rangle$. $\square$

Thus we may introduce the following notion.

Definition 17 *For the generic case, the semi-algebraic set $\mathtt{C}' = \mathtt{C} \setminus V\big(\prod_{i=1}^{n-2} \hat{d}_i\big)$ is called the* generic region *for parametric spectral factorization. Also, for the singular case, the semi-algebraic set $\mathtt{C}' = \mathtt{C} \setminus V\big(\prod_{i=1}^{n-2} \bar{d}_i\big)$ is also called the* generic region *for parametric spectral factorization.*

Remark 18 *In the generic case, there is a strong relation between $\prod_{i=0}^{n-2} \hat{d}_i$ and $D(\hat{S}_f)$, where $D(\hat{S}_f)$ denotes the discriminant of $\hat{S}_f$. We note that $D(\hat{S}_f)$ can be computed by the resultant of $\hat{S}_f(y)$ and $d\hat{S}_f(y)/dy$, and thus that $D(\hat{S}_f)$ is a polynomial in $\mathbf{q}$. As $\hat{S}_f$ is assumed to be square-free, $D(\hat{S}_f)$ is non-zero polynomial. Then we can show that each irreducible factor of $\hat{d}_i$ is also a factor of $D(\hat{S}_{f(x,\mathbf{q})})$. Thus, for each irreducible factor w of $\prod_{i=0}^{n-2} \hat{d}_i$, $\hat{S}_f$ is not square-free over the quotient field of $\mathbb{Q}[\mathbf{q}]/\langle w \rangle$.*

This fact can be shown by considering algebraic extensions of $\mathbb{Q}(\mathbf{q})$ by the roots of $\hat{S}_f$. For each b_i, its characteristic polynomial $\hat{U}_i$ is also computed by the multiplication map and it can be shown in the same manner as in Theorem 14 that $\hat{U}_i$ is a monic polynomial in $\mathbf{q}$. Therefore each root γ of $\hat{S}_f$, in the algebraic closure of $\mathbb{Q}(\mathbf{q})$, is integral *over $\mathbb{Q}[\mathbf{q}]$, and so is any root $\hat{h}_k(\gamma)$ of $\hat{U}_k$. Then, using the fact that $\mathbb{Q}[\mathbf{q}]$ is* integrally closed *in $\mathbb{Q}(\mathbf{q})$, we can draw the conclusion that each irreducible factor of $\hat{d}_i$ is also a factor of $D(\hat{S}_f)$.*

Now, as an example, we compute the shape basis for $f = a_6x^6 + a_4x^4 + a_2x^2 + a_0$. As $\hat{S}_f$ is square-free, it is a generic case, and we have the following shape basis:

$$\begin{aligned}
\hat{S}_f(b_2) =& b_2^8 + 4a_4a_6b_2^6 + (-8a_2a_6^3 + 6a_4^2a_6^2)b_2^4 \\
&+ (64a_0a_6^5 - 16a_2a_4a_6^4 + 4a_4^3a_6^3)b_2^2 + 16a_2^2a_6^6 - 8a_2a_4^2a_6^5 + a_4^4a_6^4 \,, \\
\hat{h}_1(b_2) =& \frac{b_2^2 + a_4a_6}{2a_6} \,, \\
\hat{h}_0(b_2) =& \frac{b_2^7 + 4a_4a_6b_2^5 + (-4a_2a_6^3 + 5a_4^2a_6^2)b_2^3 + (64a_0a_6^5 - 8a_2a_4a_6^4 + 2a_4^3a_6^3)b_2}{32a_2a_6^5 - 8a_4^2a_6^4} \,.
\end{aligned}$$

Then irreducible factors of $\hat{d}_0\hat{d}_2$ are a_6 and $4a_2a_6 - a_4^2$. (These factors are also factors of the discriminant $D(\hat{S}_f)$.) Also, due to the condition of $\mathtt{C}$ presented in Subsection 2.5.1, only one factor $4a_2a_6 - a_4^2$ remains.

We then consider the 'singular region' $\mathtt{C} \cap V\left(\prod_{i=0}^{n-2} \hat{d}_i\right)$, or $\mathtt{C} \cap V\left(\prod_{i=0}^{n-2} \bar{d}_i\right)$. To deal with it, the notion of 'comprehensive Gröbner system' is useful and the strategy for its computation can also be applied. However, as pointed out in Theorem 16, we can make good use of Gröbner basis computation over rational function fields.

Here we employ the notation for the generic case. For the singular case, we can execute the same procedure only by replacing the notation. Let $\{w_1, \ldots, w_s\}$ be the set of all irreducible factors of $\prod_{i=0}^{n-2} \hat{d}_i$, and set $\mathtt{C}_i = (\mathtt{C} \cap V(w_i)) \setminus V(\prod_{j>i} w_j)$. Then, $\mathtt{C} \cap V(\prod_{i=0}^{n-2} w_i) = \cup_{i=0}^{n-2} \mathtt{C}_i$. If $\mathtt{C}_i = \emptyset$, then of course we discard such $\mathtt{C}_i$. (For checking this, we can use the Gröbner basis computation and quantifier elimination.)

Based on Theorem 11, we can compute a shape basis in each singular region $\mathtt{C}_i$ as follows:

`Procedure_for_singular_region`

Step I: Compute the simple part $\hat{T}_i(y)$ of $\hat{S}_f(y)$ by using GCD computation over the quotient field of the residue class ring $\mathbb{Q}[\mathbf{q}]/\langle w_i \rangle$. *(As $\langle w_i \rangle$ is a prime ideal, $\mathbb{Q}[\mathbf{q}]/\langle w_i \rangle$ is an integral domain. Computation over the quotient field will be explained in Remark 20.)*

Step II: Perform basis conversion to the ideal $\langle \mathcal{G}, \hat{T}_f(b_{n-1}) \rangle$ over the quotient field of $\mathbb{Q}[\mathbf{q}]/\langle w_i \rangle$ to obtain the shape base $\mathcal{F}_i = \{\hat{T}_i(y), b_{n-2} - \hat{h}_{i,n-2}(y), \ldots, b_0 - \hat{h}_{i,0}(y)\}$. Further let the denominator of $\hat{h}_{i,j}$ be $\hat{d}_{i,j}$.

Step III: Compute the ideal $\mathcal{W}_i = \langle w_i, \prod_{j=0}^{n-2} \hat{d}_{i,j} \rangle$. If $\mathcal{W}_i$ is trivial, we do not need further computation. Otherwise, by prime decomposition, we obtain all its prime divisors $\mathcal{W}_{i,1}, \ldots, \mathcal{W}_{i,t}$. Then we have new singular regions $\mathtt{C}_{i,j} = \left(\mathtt{C} \cap V(\mathcal{W}_{i,j})\right) \setminus \left(V(\prod_{k>j} \mathcal{W}_{i,k}) \cup V(\prod_{k>i} w_j)\right)$ to deal with. (Here, $\mathtt{C}_i = \cup_{j=1}^{t} \mathtt{C}_{i,j}$.)

For actual algorithms for prime decomposition, see [7, 16].

Using Theorem 16, where we consider the ring homomorphism $\varphi_{\mathbf{c}}$ from $(\mathbb{Q}[\mathbf{q}]/\langle w_i \rangle)[\mathbf{B}]$ to $\mathbb{R}[\mathbf{B}]$ for $\mathbf{c} \in \mathtt{C} \setminus V(w_i)$, we have the following.

Theorem 19 *For any* $\mathbf{c} \in \mathtt{C}_i \setminus V\big(\prod_{j=1}^{n-2} \hat{d}_{i,j}\big)$*,* $\varphi_{\mathbf{c}}(\mathcal{F}_i)$ *has the zero giving the spectral factor of* $f_{\mathbf{c}}$*.*

Along with the strategy as in the comprehensive Gröbner system, we deal with a region $(\mathtt{C} \cap V(\mathcal{J}_1)) \setminus V(\mathcal{J}_2)$, where $\mathcal{J}_1$ is a prime ideal in $\mathbb{Q}[\mathbf{q}]$.

Then we apply `Procedure_for_singular_region` with the following modification: in *Step I*, GCD computations are done over the quotient field of the residue class ring $\mathbb{Q}[\mathbf{q}]/\mathcal{J}_1$, and in *Step III*, we compute $\mathcal{J}_1' = \langle \mathcal{J}_1 \cup \{\prod_{j=0}^{n-2} \hat{d}_{i,j}\}\rangle$ and its prime divisors. For the computed shape basis, Theorem 19 holds, where we consider the ring homomorphism $\varphi_{\mathbf{c}}$ from $(\mathbb{Q}[\mathbf{q}]/\mathcal{J}_1)[\mathbf{B}]$ to $\mathbb{R}[\mathbf{B}]$.

Remark 20 *The arithmetic computation over the quotient field of the residue class ring* $\mathbb{Q}[\mathbf{q}]/\mathcal{J}_1$ *can be done in the following manner:*

We take a maximal independent set $\mathbf{q}'$ *in* $\mathbf{q}$ *with respect to the prime ideal* $\mathcal{J}_1$*, which can be efficiently computed by a strongly independent set. (See [7, 16] for the independent set and related topics.) Then the cardinality of* $\mathbf{q}'$ *is the dimension of* $\mathcal{J}_1$*. Consider another polynomial ring* $\mathbb{Q}(\mathbf{q}')[\mathbf{q} \setminus \mathbf{q}']$ *and the extension ideal* $\mathcal{J}_1^e$ *of* $\mathcal{J}_1$*, which is the ideal generated by* $\mathcal{J}_1$ *in* $\mathbb{Q}(\mathbf{q}')[\mathbf{q} \setminus \mathbf{q}']$*. Then,* $\mathcal{J}_1^e$ *is a maximal ideal. Thus the residue class ring* $\mathbb{Q}(\mathbf{q}')[\mathbf{q} \setminus \mathbf{q}']/\mathcal{J}_1^e$ *is a field containing* $\mathbb{Q}[\mathbf{q}]/\mathcal{J}_1$*, and hence it can be identified with the quotient field of* $\mathbb{Q}[\mathbf{q}]/\mathcal{J}_1$*. Thus all arithmetic computation can be done over* $\mathbb{Q}(\mathbf{q}')[\mathbf{q} \setminus \mathbf{q}']/\mathcal{J}_1^e$*.*

Remark 21 *For the computation of* shape bases*, we may ignore the conditions of* $\mathtt{C}$ *except that for* $a_{2n} \neq 0$ *first. This is because, even if* $f_{\mathbf{c}}$ *has imaginary axis roots, Theorems 15, 16, and 19 hold except for the condition that those ideals have the zero yielding the spectral factorization. Then we can pick up all cases which are consistent with the whole condition of* $\mathtt{C}$*.*

Finally we discuss the termination of our whole procedure. In *Step III*, we compute $\mathcal{J}_1' = \langle \mathcal{J}_1 \cup \{\prod_{j=0}^{n-2} \hat{d}_{i,j}\}\rangle$. If, in every non-empty singular region, $\mathcal{J}_1'$ becomes larger than $\mathcal{J}$, the termination is guaranteed by the noetherian property of $\mathbb{Q}[\mathbf{q}]$.

When $\mathcal{J}_1' = \mathcal{J}$ occurs for a non-empty singular region $\big(\mathtt{C} \cap V(\mathcal{J}_1)\big) \setminus V(\mathcal{J}_2)$, we can apply the following special treatment for it. (Such cases seem unusual.)

First compute the decomposition

$$\big(\mathtt{C} \cap V(\mathcal{J}_1)\big) \setminus V(\mathcal{J}_2) = \cup_{i=1}^{s} \big(\mathtt{C} \cap V(\mathcal{L}_{i,1})\big) \setminus V(\mathcal{L}_{i,2}) ,$$

where, for each i, a factor $\bar{T}_f$ of $\hat{S}_f$ over the quotient field of $\mathbb{Q}[\mathbf{q}]/\mathcal{L}_{i,1}$ is computed and $\varphi_{\mathbf{c}}(\bar{T}_f)$ is still the simple part of $\hat{S}_{f_{\mathbf{c}}}$ for any $\mathbf{c} \in (\mathtt{C} \cap V(\mathcal{L}_{i,1})) \setminus V(\mathcal{L}_{i,2})$. This decomposition can be done by 'parametric square-free factorization' ([14]) based on comprehensive Gröbner system computation.

Then we also apply a comprehensive Gröbner system computation algorithm (e.g., one proposed in [12]) to the 'parametric ideal' $\mathcal{G} \cup \{\bar{T}_f\}$, with condition that $\mathbf{c} \in \mathcal{L}_{i,1}, \mathbf{c} \notin \mathcal{L}_{i,2}$ and the condition of $\mathtt{C}$ for parameters $\mathbf{c}$, in the ring $\mathbb{Q}[\mathbf{q}][\mathbf{B}]$. (We may ignore inequalities on the condition of $\mathcal{C}$; see Remark 21.) For any $\mathbf{c} \in (\mathtt{C} \cap V(\mathcal{L}_{i,1})) \setminus V(\mathcal{L}_{i,2})$, as $\varphi_{\mathbf{c}}(\bar{T}_f)$ is still the simple part of $\hat{S}_{f_{\mathbf{c}}}$, $\langle \mathcal{G}_{\mathbf{c}} \cup \{\varphi_{\mathbf{c}}(\bar{T}_f)\}\rangle$ is radical and has its shape basis; see Theorem 11 and the proof of Theorem 16. Therefore, for each component $\{\mathcal{K}, \mathcal{E}\}$ of the computed comprehensive Gröbner system, where $\mathcal{K}$ is the condition on $\mathbf{c}$ and $\mathcal{E}$ is a finite subset of $\mathbb{Q}[\mathbf{q} \cup \mathbf{B}] \setminus \mathbb{Q}[\mathbf{q}]$, the evaluation $\varphi_{\mathbf{c}}(\mathcal{E})$ forms a Gröbner basis. Moreover, by making it to be reduced in a symbolic way if necessary, the desired shape basis can be obtained. (See [12] for the 'reduced comprehensive Gröbner system'.)

On closing this subsection, we continue our computation for $f = a_6x^6+a_4x^4+a_2x^2+a_0$. The remaining singular region is $\mathtt{C} \cap V(4a_2a_6 - a_4^2)$. Over the quotient field of $\mathbb{Q}[a_6, a_4, a_2, a_0]/\langle 4a_2a_6 - a_4^2\rangle$, the GCD of $\hat{S}_f(b_2)$ and $d\hat{S}_f(b_2)/db_2$ is b_2, and so the simple part $\hat{T}_1$ of $\hat{S}_f(b_2)$ is $b_2^6 + 4a_4a_6b_2^4 + 4a_4^2a_6^2b_2^2 + 64a_0a_6^5$. (We note that $64a_0a_6^5 \neq 0$ from the condition of $\mathtt{C}$.) We add $\hat{T}_1$ to $\mathcal{G}$ and compute the shape basis $\mathcal{F}_1$ with respect to the lexicographical ordering $b_2 \prec b_1 \prec b_0$ to get

$$\begin{aligned}\hat{T}_1(b_2) &= b_2^6 + 4a_4a_6b_2^4 + 4a_4^2a_6^2b_2^2 + 64a_0a_6^5\ ,\\ \hat{h}_{11}(b_2) &= \frac{b_2^2 + a_4a_6}{2a_6}\ ,\\ \hat{h}_{10}(b_2) &= \frac{b_2^3 + 2a_4a_6b_2}{8a_6^2}\ .\end{aligned}$$

Notice that the product $\hat{d}_{10}\hat{d}_{11}$ of the denominators has unique irreducible factor a_6. Nevertheless the factor cannot be consistent with the condition of $\mathtt{C}$ and our computation is just finished. Here all field arithmetic computations were done over $\mathbb{Q}(a_6, a_4, a_0)[a_2]/\langle 4a_2a_6 - a_4^2\rangle^e$.

In the previous subsection, we dealt with a non-parametric polynomial $f(x) = 2x^6 - 28x^4 + 98x^2 - 72$, which is a special case of the polynomial considered here with $a_6 = 2,\ a_4 = -28,\ a_2 = 98$ and $a_0 = -72$. Substituting those values, we have the same shape basis as presented in (13).

3 Optimization over Parameters Using the Sum of Roots

3.1 Optimization via Quantifier Elimination

A variety of problems in signal processing and control boil down to spectral factorization and the solutions to such problems can be expressed explicitly in terms of the spectral factor [5]. More specifically the optimal cost or the optimal controller to be obtained may be written as functions in the coefficients of the polynomial spectral factor. The results presented in the preceding section therefore enable us to express such quantities in terms of the SoR. In particular parametric polynomial spectral factorization allows us to carry out parametric optimization and thus to obtain an explicit expression of the quantity in terms of parameters and the SoR. Using the resulting expression, we can further employ various kinds of polynomial optimization algorithms for optimization over parameters.

Here we indicate that another algebraic method, quantifier elimination, is applicable to the latter task. The QE-based optimization approach has already been proposed to solve possibly non-convex optimization problems under polynomial constraints (e.g. [17]). The novelty of the method proposed in this report is:

- the introduction of the SoR that allows us to link parameters with the quantity to be optimized;
- simple formulation of the original control problem in the algebraic form along with the 'largest real root' condition.

More specifically the results in Section 2 suggest that, if the cost function (i.e., the quantity), which we denote by $J^\star$, is expressed as a polynomial/rational function in parameters and the coefficients of the spectral factor, then it can also be expressed algebraically in terms of parameters and the SoR.

Let $\mathbf{q} = (q_1, q_2, \ldots, q_m)$ be a vector of real parameters and $\mathcal{Q} \subset \mathbb{R}^m$ the *permissible region* of the parameters (i.e., it is required that $\mathbf{q} \in \mathcal{Q}$). Suppose that the set of constraints on parameters (i.e., $\mathbf{q} \in \mathcal{Q}$) can be written as $\varphi(\mathbf{q})$ where $\varphi(\mathbf{q})$ is assumed to consist of a set of algebraic expressions (equalities/inequalities) in parameters. By introducing an intermediate variable η to hold the value of $J^\star$, the optimization problem can be recast as

$$\exists\sigma\ \exists\mathbf{q}\ (\ \eta - J^\star(\mathbf{q}, \sigma) = 0\ \wedge\ [\sigma \text{ is the largest real root of } S_f]\ \wedge\ \varphi(\mathbf{q})\ )\ .$$

The vital thing is to write down the condition '[σ is the largest real root of S_f]' as a set of algebraic expressions, which is dealt with in the following subsection. Once this is done, quantifier elimination can then be applied to eliminate all variables but one (i.e., σ and $\mathbf{q}$, but not η) and return an equivalent condition, which is a set of polynomial inequalities in η only. The resulting condition then reveals the range that $J^\star$ can take when parameters vary in the permissible range $\mathcal{Q}$. In the course of the QE procedure, the combination of parameter values that achieves the maximum/minimum can also be obtained. In this way we can optimize $J^\star$ over parameters.

Remark 22 *In the above approach a general QE algorithm is assumed to be used. Many of QE algorithms such as* QEPCAD B *[18] rely on Cylindrical Algebraic Decomposition (CAD). However it seems possible to tailor an efficient CAD-based algorithm to optimization problems such as the one considered here. Some potential areas for improvement are in order. It is observed that a general CAD algorithm yields projection factors unnecessary for this particular optimization. By constructing a special algorithm that computes projection factors required for finding the optimal $J^\star$, the computation time may be much reduced. Also, in most* *problems derived from practical engineering problems, the range of $J^\star$ is an interval (or a semi-interval) (rather than a set of disjoint intervals). Hence it is not necessary to examine all the candidates of sample points of $J^\star$ during the lifting phase, unlike a general QE algorithm. That is, starting from the value of $J^\star$ for the nominal values of parameters, for instance, one would have to check adjacent values of sample points of $J^\star$ until one finds the infeasible value(s). A similar idea is exploited in [19]. Lastly, the choice of variable ordering and the way to specify the SoR have a significant effect for the computation time [20]. This point is discussed further in Remark 23. The efficacy of the suggested approach is under investigation and will be reported elsewhere.*

3.2 Specifying the SoR with the Sturm-Habicht Sequence

In order to use the QE technique, all the conditions are to be expressed algebraically. In the particular problem under consideration, the description that σ is the largest real root of S_f needs to be translated into algebraic expressions. Given a polynomial, the fact that a particular number is its largest real root can be described as the condition

- that the value is a root of the polynomial; and also
- that there is no real root between that value and $+\infty$.

The Sturm-Habicht sequence [15] gives an algebraic condition for the number of polynomial roots in an interval on the real axis when a polynomial with real parametric coefficients is provided. This fact directly yields a condition stating that a particular number is the largest real root of the given polynomial. That is, the QE-based optimization approach proposed in Subsection 3.1 is in fact executable. Readers are referred to [15] for an exposition of an algorithm to compute the Sturm-Habicht sequence.

For simplification of the resulting condition, we further exploit the structure of the problem. As is stated in Subsection 2.2, S_f is a polynomial in σ^2 in the generic case. Therefore the condition to be found is that the square of the SoR is the largest real root of S_f seen as polynomial in σ^2 and that $\sigma > 0$. This simplifies the computation of the Sturm-Habicht sequence since it can half the degree of the polynomial for which the sequence is calculated. It is noted here that the number of inequality constraints derived from this approach grows exponentially with the degree of S_f. Simplification of these constraints is crucial when solving high order cases.

Remark 23 *The language of Extended Tarski formulas accepted in* QEPCAD B *[18] allows indexed roots of polynomials, e.g., an expression like $x_k = root_j\, f(x_1, \dots, x_k)$, to be specified. This may be used instead of the conditions derived from the Sturm-Habicht sequence. The restriction in* QEPCAD B *is that the variables need to be ordered as $x_1 \prec x_2 \prec \cdots \prec x_k$. In the case* *of the problem considered in this report, this variable ordering requires that the SoR σ should be eliminated first. Nevertheless some empirical results indicate that, for some cases, the computation terminates faster if the conditions from the Sturm-Habicht sequence are used and σ is eliminated last (i.e., all parameters are eliminated before σ), than using the indexed root and eliminating σ first; for other cases, using the indexed root yields shorter computation time. Variable ordering and the method of specifying the SoR are two of potential sources of improvement and their effect is to be further investigated.*

4 Optimization Approach and Application to a Control Problem

4.1 Algorithms

In this subsection we summarize in the algorithm form the development in the preceding sections. The first algorithm, corresponding to Section 2, is to solve parametric optimization and to express, for instance, the optimal cost in terms of parameters and the SoR. The second algorithm, corresponding to Section 3, is to carry out optimization over parameters using the results from the first algorithm. It is implicitly assumed that numbers provided in the input to the algorithms are all in $\mathbb{Q}$.

Algorithm 1 *For a problem that reduces to polynomial spectral factorization and that seeks a quantity which can be expressed in terms of the input data and the coefficients of the spectral factor.*

Input: *Coefficients of polynomial $f(x)$ to be decomposed, in the polynomial form in parameters $\mathbf{q}$.*

Output: *Polynomial $S_f(\sigma)$ relating $\mathbf{q}$ and the SoR σ; expression for the quantity in $\mathbf{q}$ and σ.*

Step I: Convert the original problem into a polynomial spectral factorization problem. More specifically, obtain expressions for the coefficients of $f(x)$ in the polynomial form in $\mathbf{q}$.

Step II: Carry out polynomial spectral factorization according to Section 2. Get a polynomial $S_f(\sigma)$ relating $\mathbf{q}$ and σ and also polynomial/rational expressions for the coefficients b_i of $g(x)$ in terms of $\mathbf{q}$ and σ.

Step III: Compute an expression for the pursued quantity in terms of $\mathbf{q}$ and σ. (This step is totally dependent on the problem that the user wants to solve.)

Algorithm 2 *For a problem that has the quantity to be optimized in polynomial/rational form in $\mathbf{q}$ and σ and that specifies the ranges of $\mathbf{q}$ as algebraic constraints.*

Input: *Polynomial/rational expression for the quantity in $\mathbf{q}$ and σ; polynomial relating $\mathbf{q}$ and σ; algebraic constraints on $\mathbf{q}$.*

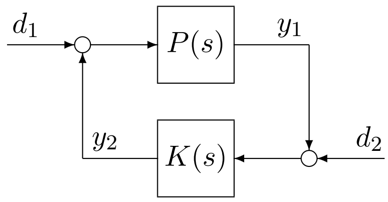

Figure 1: Standard feedback configuration.

Output: *Maximum/minimum values of the quantity and parameter values that achieve them.*

Step I: Rewrite the problem as in Subsection 3.1, using the Sturm-Habicht sequence as in Subsection 3.2.

Step II: Execute quantifier elimination.

Step III: Find out maximum/minimum values of the quantity and also identify optimizers.

4.2 A Control Problem: LQG Control

In order to demonstrate the approach proposed in this report, this and the following subsections consider a particular $\mathcal{H}_2$ control problem, which is called the normalized linear quadratic Gaussian (LQG) control [21, 22]. The problem is formulated as follows. Consider the feedback configuration depicted in Fig. 1. Suppose that the plant (i.e., the system to be controlled) is a single-input-single-output continuous-time, linear, time-invariant plant and that its transfer function is given as an n-th order, strictly proper $P(s)$[1]. The task is then to design a controller (denoted by its transfer function $K(s)$) which stabilizes the closed-loop system and minimizes the $\mathcal{H}_2$-norm of the transfer function matrix $T_{w\to z}(s)$ from $w = (d_1\ d_2)^T$ to $z = (y_1\ y_2)^T$:

$$T_{w\to z}(s) = \frac{1}{1 - PK}\begin{pmatrix} P & PK \\ PK & K \end{pmatrix} .$$

Namely we are interested in the optimal performance level

$$J^\star := \min_{K \text{ stabilizing}} \left\| T_{w\to z}(s) \right\|_2^2 ,$$

and the controller $K_{\text{opt}}(s)$ that achieves $J^\star$.

Here we briefly review the $\mathcal{H}_2$-norm. The $\mathcal{H}_2$-norm of the transfer function matrix $G(s)$ of a system is defined as

$$\left\|G(s)\right\|_2 := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{tr}\left\{\overline{G^T(\sqrt{-1}\,\omega)}G(\sqrt{-1}\,\omega)\right\} d\omega\right)^{\frac{1}{2}} ,$$

[1] A transfer function is a rational function in the Laplace variable s. Its degree is defined as the degree of the denominator, and it is called strictly proper if the degree of the denominator is strictly greater than that of the numerator.

and the square of the $\mathcal{H}_2$-norm of a system is equal to the energy of the system output when an impulse signal is applied to the input [23]. The $\mathcal{H}_2$-norm therefore indicates how promptly the system attenuates impulse-type disturbance. Once the transfer function of G is given, computation of $\|G\|_2^2$ in essence requires solution of a set of linear equations and the resulting expression will be a rational function in the coefficient of G [2, 23].

Given a particular plant, the optimal controller and the optimal performance level can be found by means of spectral factorization [2, 5]. However, in practice, there is some freedom in designing a plant, e.g., the mass of a weight may be chosen from a certain range. Parametric polynomial spectral factorization can thus help engineers to exploit the freedom and to decide parameter values so that they may construct a desirable plant which is easier to control. It is emphasized that, at this point, we are focusing on the design of the *plant* (because we are deciding plant parameters) rather than the *controller*, on the assumption that the optimal controller can be found once the plant is fixed.

Suppose that there is a set $\mathbf{q}$ of parameters in the plant P, and write P as $P(s, \mathbf{q})$ to explicitly express the existence of the parameters. Given a particular value for each parameter, one can compute the optimal cost for the normalized LQG control, which we denote by $J^\star(\mathbf{q})$, again to explicitly express the dependence on the parameter values. Write the permissible region of the parameters as $\mathcal{Q}$. Our task is

- to derive an expression of $J^\star(\mathbf{q})$ in terms of $\mathbf{q}$ and σ; and further
- to find the best choice of parameter values that achieves the smallest value of $J^\star(\mathbf{q})$ from the permissible region, i.e., to find

$$\min_{\mathbf{q}\in\mathcal{Q}} J^\star(\mathbf{q}) = \min_{\mathbf{q}\in\mathcal{Q}} \min_{K \text{ stabilizing}} \left\| T_{w\to z}(s) \right\|_2^2$$

  and $\mathbf{q}$ that achieves the above.

A solution approach to this problem by means of polynomial spectral factorization may be stated in the following way [24]. Note that the approach presented here mainly corresponds to *Step III* of Algorithm 1 in the preceding subsection. Write the transfer function of the plant as

$$P(s) = \frac{P_N(s)}{P_D(s)}\ , \tag{14}$$

where P_N and P_D are polynomials in s. Also let P_D be monic. In the parametric case, we suppose that the coefficients of P_N and P_D are polynomials in $\mathbf{q}$, and further assume that P_N and P_D are coprime for all $\mathbf{q} \in \mathcal{Q}$ (see the comment at the end of this section for the legitimacy of this assumption). If the plant is n-th order, then the degree of P_D is n where seen as a polynomial in s and that of P_N is strictly less than n due to the strict properness of P. Firstly polynomial spectral factorization is carried out for the following even polynomial constructed from P_N and P_D:

$$P_N(-s)P_N(s) + P_D(-s)P_D(s) \; . \tag{15}$$

Its polynomial spectral factor $M_D(s)$ in this case is an n-th order monic polynomial.

Next, in order to find the optimal controller, polynomials $V_N(s)$, $U_N(s)$ of degrees n and (at most) $n-1$, respectively, are found that satisfy

$$P_D(s)V_N(s) - P_N(s)U_N(s) = \{M_D(s)\}^2 \; . \tag{16}$$

By comparing the coefficients on the both sides, a set of linear equations in terms of the coefficients of V_N and U_N can be obtained and the coefficients of V_N and U_N satisfying (16) can be computed by solving the set of equations. The optimal controller is then given as

$$K_{\text{opt}}(s) = \frac{U_N(s)}{V_N(s)} \; . \tag{17}$$

Note that, due to the coprimeness of P_N and P_D, V_N and U_N are uniquely determined. It is further pointed out that the coefficients of V_N and U_N are polynomials in the coefficient b_i of M_D and rational functions in the parameters $\mathbf{q}$.

Once K_{opt} is found, $J^\star$ can be computed in a straightforward manner. The transfer function matrix in that case can be written as

$$T_{w\to z}(P, K_{\text{opt}}) = -\frac{1}{M_D^2}\begin{pmatrix} P_N V_N & P_N U_N \\ P_N U_N & P_D U_N \end{pmatrix} =: \begin{pmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{pmatrix} \; .$$

Since

$$\big\|T_{w\to z}(P, K_{\text{opt}})\big\|_2^2 = \big\|T_{11}\big\|_2^2 + 2\big\|T_{12}\big\|_2^2 + \big\|T_{22}\big\|_2^2 \; ,$$

what we have to do is to compute $\big\|T_{11}\big\|_2^2$ etc. separately. If we write T_{11} as

$$T_{11}(s)\left(= -\frac{P_N(s)V_N(s)}{\{M_D(s)\}^2}\right) = \frac{\beta_{2n-1}s^{2n-1} + \beta_{2n-2}s^{2n-2} + \cdots + \beta_0}{s^{2n} + \alpha_{2n-1}s^{2n-1} + \alpha_{2n-2}s^{2n-2} + \cdots + \alpha_0} \; ,$$

then its $\mathcal{H}_2$-norm may be obtained by solving a Lyapunov equation. Write T_{11} in state-space form, e.g., in the observer canonical form:

$$T_{11}(s) = \left[\begin{array}{ccccc|c} -\alpha_{2n-1} & 1 & 0 & \cdots & 0 & \beta_{2n-1} \\ -\alpha_{2n-2} & 0 & 1 & \cdots & 0 & \beta_{2n-2} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ -\alpha_1 & 0 & 0 & \cdots & 1 & \beta_1 \\ -\alpha_0 & 0 & 0 & \cdots & 0 & \beta_0 \\ \hline 1 & 0 & 0 & \cdots & 0 & 0 \end{array}\right] =: \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \; .$$

Then, by using the (unique) solution L_o (which is a symmetric matrix) to the Lyapunov equation

$$A^* L_o + L_o A + C^* C = 0 \ , \tag{18}$$

$\|T_{11}\|_2^2$ may be computed as

$$\|T_{11}\|_2^2 = \mathrm{tr}\left\{B^* L_o B\right\} \ .$$

The remaining $\|T_{12}\|_2^2$ and $\|T_{22}\|_2^2$ can be computed in the same way. Adding them all, we can get an expression for $J^\star$. Notice that (18) yields a set of linear equations in terms of the elements ℓ_{jk} of L_o. This implies that ℓ_{jk} are expressed as rational functions in b_i, concluding that $J^\star$ is a rational function in b_i and $\mathbf{q}$.

The above solution approach thus indicates that this problem falls into the category of problems tackled by the algorithms in the previous subsections. It is mentioned that a wide range of $\mathcal{H}_2$ control problems can be dealt with in a similar manner (see e.g. [25]).

On closing this subsection, we mention how the assumption that the polynomial to be decomposed has no roots on the imaginary axis arises from the formulation of the control problem. When the transfer function of the plant is written as in (14), control theory requires that P_N and P_D should be coprime so that the plant is controllable/observable [23]. Polynomial spectral factorization is executed for (15), but, due to the coprimeness, the polynomial (15) has no roots on the imaginary axis. In the case where the plant has parameters, i.e., P_N and P_D contain parameters, the coprimeness of P_N and P_D under parameter variations needs to be examined beforehand since the structure (the degree) of the plant may change and there can be some singularity. That is, special attention must be paid before the design of a controller. Therefore we can safely assume the non-existence of imaginary axis roots.

Remark 24 *The quantifier elimination technique have been successfully applied to some control problems (see .e.g. [26, 27]). Most problems assume a fixed plant, and are solved by parametrizing the controller/control strategy and finding feasible/optimal parameters. The problem considered here is different in that parameters in the* system to be controlled *are dealt with and optimization is executed over those parameters, assuming that the optimal controller can always be employed. The SoR and parametric polynomial spectral factorization are instrumental to this approach.*

4.3 Numerical Example

As a demonstration of the algorithms stated in Subsection 4.1, we employ the following numerical example for the control problem formulated in Subsection 4.2:

$$P(s, q_1, q_2) = \frac{s - q_1}{s(s + \frac{1}{10} q_2)} =: \frac{P_N(s, \mathbf{q})}{P_D(s, \mathbf{q})} \ ,$$

where $\mathbf{q} = (q_1, q_2)$ is the vector of tuning parameters restricted to the permissible region

$$\mathcal{Q} = \left\{ \mathbf{q} = (q_1, q_2) \,\middle|\, q_1 \in [\tfrac{1}{5}, 1], q_2 \in [\tfrac{9}{10}, \tfrac{11}{10}] \right\} .$$

First note that the numerator and the denominator are always coprime in the permissible region.

We follow the steps of Algorithms 1 and 2.

Step 1-I: The even polynomial (15) to be decomposed is computed:

$$P_N(-s, \mathbf{q})P_N(s, \mathbf{q}) + P_D(-s, \mathbf{q})P_D(s, \mathbf{q}) = s^4 - \left(\tfrac{1}{100}q_2^2 + 1\right)s^2 + q_1^2 . \quad (19)$$

It is easy to see that (19) does not have an imaginary axis root unless $q_1 = 0$. So the regular region is $\mathtt{C} = \{q_1 \neq 0\}$. Notice that $q_1 = 0$ yields a common factor s in the numerator and the denominator of the transfer function, and that this singularity agrees with the observation of P. Since the permissible region $\mathcal{Q}$ is contained in $\mathtt{C}$, the method stated in Subsection 2.5 is applicable.

Step 1-II: We carry out polynomial spectral factorization for (19). Write its spectral factor as

$$M_D(s) = s^2 + b_1 s + b_0 .$$

Comparing the coefficients of (19) and those of $M_D(s)M_D(-s)$, we get the following set of algebraic equations:

$$\begin{cases} b_1^2 - 2b_0 - \frac{1}{100}q_2^2 - 1 = 0 , \\ b_0^2 - q_1^2 = 0 . \end{cases}$$

As is stated in Lemma 5, the left hand sides of the equations forms the reduced Gröbner basis with respect to the graded reverse lexicographic order $b_1 \succ b_0$. By means of 'parametric' basis conversion, a shape basis is obtained, and the following relationship is obtained:

$$\begin{cases} S_f(\sigma) = \sigma^4 - \left(\frac{1}{50}q_2^2 + 2\right)\sigma^2 - 4q_1^2 + \frac{1}{10000}q_2^4 + \frac{1}{50}q_2^2 + 1 , \\ b_1 = \sigma , \\ b_0 = \frac{1}{2}\sigma^2 - \frac{1}{200}q_2^2 - \frac{1}{2} . \end{cases} \quad (20)$$

Step 1-III: First we compute the optimal controller. According to the degree requirements, write V_N and U_N as

$$V_N(s) = s^2 + v_1 s + v_0 ,$$
$$U_N(s) = u_1 s + u_0 .$$

Comparing the coefficients of the both sides of (16), we can get a set of linear equations in terms of v_i and u_j. By solving it, v_i and u_j are obtained and the optimal controller can be written as in (17), where

$$V_N(s) = s^2 + \big(2b_1 - \tfrac{1}{10}q_2\big)s + \frac{b_1\big(100b_1^2 + 200q_1b_1 - 20q_1q_2 - q_2^2 - 100\big)}{10\big(10q_1 + q_2\big)} ,$$

$$U_N(s) = \frac{\big(100b_1^3 - 20q_2b_1^2 + q_2^2b_1 - 100b_1 + 10q_2 + 100q_1\big)}{10(10q_1 + q_2)} s + q_1 .$$

Notice that the coefficients of V_N and U_N are polynomials in the coefficient b_i of M_D and rational functions in the parameters $\mathbf{q}$.

Now, $J^\star$ is computed. As is suggested in Subsection 4.2, three transfer functions, namely, T_{11}, T_{12}, and T_{22}, are computed from P_N, P_D, V_N, and U_N. Then the square of the $\mathcal{H}_2$-norm is calculated for each transfer function by solving the Lyapunov equation (18). We then get an expression for $J^\star$ as a rational function in b_i and $\mathbf{q}$. Using the relationship (20), we can introduce σ to eliminate b_1 and b_0. Moreover, finding the inverse of the denominator using S_f, we can convert the expression for $J^\star$ into an expression polynomial in σ:

$$J^\star = \frac{\begin{array}{c}(1500q_2^2 + 50000)\sigma^3 + (100000q_1 - 200q_2^3 - 10000q_2)\sigma^2 \\ + (150000q_1^2 - 10000q_1q_2 - 5q_2^4 - 1500q_2^2 - 50000)\sigma \\ - 40000q_1^2q_2 + q_2^5 + 200q_2^3 + 10000q_2\end{array}}{250\big(100q_1^2 + 20q_1q_2 + q_2^2\big)} .$$

This concludes Algorithm 1. Notice that $J^\star$ is in $\mathbb{Q}(q_1, q_2)[\sigma]$ and is suited for the input to Algorithm 2.

We carry on Algorithm 2.

Step 2-I: We first specify the SoR σ in S_f using the Sturm-Habicht sequence. The Sturm-Habicht sequence for S_f (when seen as a polynomial in σ^2) is

$$\big[16q_1^2,\ 2\sigma^2 - 2 - \tfrac{1}{50}q_2,\ S_f(\sigma)\big] . \tag{21}$$

The signs in (21) at $\sigma = +\infty$ are

$$\big[+, +, +\big] .$$

When σ becomes the true SoR, the sign of the first element of (21) is again $+$ and the last element becomes 0. Since no change of sign is allowed, the permissible signs in the sequence is

$$\big[+, +, 0\big] .$$

The condition that σ is the largest real root of S_f is thus equivalent to

$$S_f = 0 \ \wedge\ 2\sigma^2 - 2 - \tfrac{1}{50}q_2 > 0 \ \wedge\ \sigma > 0 .$$

By using this, optimization of $J^\star$ over parameters $\mathbf{q}$ can be formulated as the following QE problem:

$$\begin{array}{l}\exists\sigma\, \exists q_1\, \exists q_2\, (\ \eta - J^\star = 0 \ \ \wedge\ S_f = 0 \ \wedge\ 2\sigma^2 - 2 - \tfrac{1}{50}q_2 > 0 \ \wedge\ \sigma > 0 \\ \qquad \wedge\ \tfrac{2}{10} \le q_1 \le 1 \ \wedge\ \tfrac{9}{10} \le q_2 \le \tfrac{11}{10}\ ) ,\end{array} \tag{22}$$

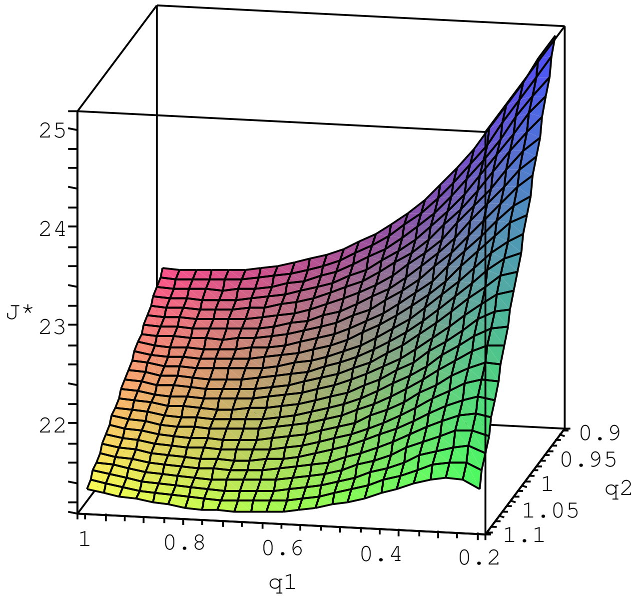

Figure 2: $J^\star$ drawn from exact expression.

where η is a new variable that is assigned to $J^\star$.
Step 2-II: QEPCAD B is then applied to the first-order formula (22), and we obtain the following quantifier-free formula in η:

$$\eta \geq \eta_1 \ \wedge \ \chi(\eta) := 105125000\eta^2 - 2299018951\eta - 8523088490 \leq 0 \,, \qquad (23)$$

where η_1 is the second root of

$$1375000\eta^3 - 58990000\eta^2 + 635959650\eta - 71455341 \,,$$

which is the unique root between $\frac{21675}{1024}$ and $\frac{5419}{256}$ ($\eta_1 \simeq 21.1672$). The formula (23) thus implies the feasible range of $J^\star$: $J^\star \in [\eta_1, \eta_2]$ where η_2 is the unique root of $\chi(\eta)$ between $\frac{25701}{1024}$ and $\frac{12851}{512}$ ($\eta_2 \simeq 25.0996$).
Step 2-III: Tracing down the CAD tree created during the QE process, we can know the parameter values achieving the minimum/maximum explicitly as algebraic numbers. For this example, it is found that the minimum and maximum values are achieved at $(q_1, q_2) \simeq (0.7096, \frac{11}{10})$ and $(q_1, q_2) = (\frac{1}{5}, \frac{9}{10})$, respectively.

Remark 25 *Since S_f in* (20) *is 4-th order, the exact expression for σ can in fact be obtained. Using this expression, we can find a closed-form expression for $J^\star$ in terms of q_1 and q_2, and $J^\star$ is plotted against q_1 and q_2 in Fig. 2. The optimization* *result shown above agrees with the plot, and it can be confirmed that the algorithms successfully found the optimal values.*

It is noted however that the approach finding an exact expression cannot be applied for general high order S_f. Also the obtained expression for $J^\star$ of this example contains many nested radicals and is judged too complicated to apply general optimization methods for. Moreover observe the non-convexity of $J^\star$, which may make it difficult for an ordinary optimization problem to find the (exact) global optimum. By solving the parametric polynomial spectral factorization problem by way of the SoR and carrying out optimization over parameters using QE, the true optimal value can in principle be obtained without failure.

5 Conclusion

This report has exploited the relationship between the sum of roots and polynomial spectral factorization and devised an algebraic approach to the parametric polynomial spectral factorization problem. It has also been shown that, based on the result and quantifier elimination, optimization over parameters can be performed for problems in signal processing and control. The effectiveness of the proposed approach is demonstrated on a numerical example of a control problem. Further efforts are to be made to find out and exploit more structural properties so that the proposed approach may become truly useful for practical applications.

Acknowledgements

The authors are indebted to Prof. Chris W. Brown for his invaluable help with the QE computation using QEPCAD B, and to Prof. Thomas Sturm for constructive suggestions. The authors would also like to thank the anonymous referees of their ISSAC 2007 paper [4], on which this report is based, for useful comments/suggestions and encouragement.

References

[1] A. M. Cohen, editor. *Computer Algebra in Industry: Problem Solving in Practice*. John Wiley & Sons, Chichester, 1993.

[2] M. Kanno, H. Anai, and K. Yokoyama. On the relationship between the sum of roots with positive real parts and polynomial spectral factorization. In T. Boyanov et al., editors, *Numerical Methods and Applications — 6th International Conference, NMA 2006, Borovets, Bulgaria, August, 2006, Revised Papers*, volume 4310 of *Lecture Notes in Computer Science*, pages 320–328. Springer-Verlag, Heidelberg, 2007.

[3] H. Anai, S. Hara, and K. Yokoyama. Sum of roots with positive real parts. In *Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation, ISSAC2005*, pages 21–28, 2005.

[4] M. Kanno, K. Yokoyama, H. Anai, and S. Hara. Parametric optimization in control using the sum of roots for parametric polynomial spectral factorization. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC 2007*, pages 211–218, Waterloo, Ontario, Canada, July-August 2007.

[5] V. Kučera. A tutorial on H_2 control theory: The continuous time case. In M. J. Grimble and V. Kučera, editors, *Polynomial Methods for Control Systems Design*, pages 1–55. Springer, London, 1996.

[6] D. J. Clements and B. D. O. Anderson. Polynomial factorization via the Riccati equation. *SIAM Journal on Applied Mathematics*, 31(1):179–205, July 1976.

[7] T. Becker and V. Weispfenning. *Gröbner Bases: A Computational Approach to Commutative Algebra*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, NY, 1993.

[8] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer, New York, NY, 3rd edition, 2007.

[9] K. Yokoyama, M. Noro, and T. Takeshima. Solutions of systems of algebraic equations and linear maps on residue class rings. *Journal of Symbolic Computation*, 14(4):399–417, October 1992.

[10] M. Noro and K. Yokoyama. A modular method to compute the rational univariate representation of zero-dimensional ideals. *Journal of Symbolic Computation*, 28(1-2):243–264, July 1999.

[11] A. Montes. A new algorithm for discussing Gröbner bases with parameters. *Journal of Symbolic Computation*, 33(2):183–208, February 2002.

[12] A. Suzuki and Y. Sato. A simple algorithm to compute comprehensive Gröbner bases using Gröbner bases. In *Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation, ISSAC2006*, pages 326–331. ACM Press, 2006.

[13] V. Weispfenning. A new approach to quantifier elimination for real algebra. In B.F. Caviness and J.R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, Texts and Monographs in Symbolic Computation, pages 376–392. Springer, Wien, 1998.

[14] K. Yokoyama. Stability of parametric decomposition. In N. Takayama and A. Iglesias, editors, *Proceedings of Second International Congress on Mathematical Software ICMS 2006*, volume 4151 of *Lecture Notes in Computer Science*, pages 391–402. Springer-Verlag, 2006.

[15] L. González-Vega, T. Recio, H. Lombardi, and M.-F. Roy. Sturm-Habicht sequences determinants and real roots of univariate polynomials. In B.F. Caviness and J.R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, Texts and Monographs in Symbolic Computation, pages 300–316. Springer, Wien, New York, 1998.

[16] T. Shimoyama and K. Yokoyama. Localization and primary decomposition of polynomial ideals. *Journal of Symbolic Computation*, 22(3):247–277, March 1996.

[17] V. Weispfenning. Simulation and optimization by quantifier elimination. *Journal of Symbolic Computation*, 24(2):189–208, August 1997.

[18] C. W. Brown. QEPCAD B: A program for computing with semi-algebraic sets using CADs. *ACM SIGSAM Bulletin*, 37(4):97–108, December 2003.

[19] H. Anai and P. A. Parrilo. Convex quantifier elimination for semidefinite programming. In *Proceedings of the 6th International Workshop on Computer Algebra in Scientific Computing (CASC) 2003*, pages 3–11, 2003.

[20] A. Dolzmann, A. Seidl, and T. Sturm. Efficient projection orders for CAD. In *Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation, ISSAC2004*, pages 111–118. ACM Press, 2004.

[21] E. A. Jonckheere and L. M. Silverman. A new set of invariants for linear systems — application to reduced order compensator design. *IEEE Transactions on Automatic Control*, AC-28(10):953–964, October 1983.

[22] D. Mustafa and K. Glover. *Minimum Entropy H_∞ Control*, volume 146 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin Heidelberg, 1990.

[23] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.

[24] M. Kanno and M. C. Smith. Guaranteed accuracy computations in systems and control. In *Proceedings of the European Control Conference ECC 2003*, Cambridge, U.K., September 2003.

[25] M. Kanno, S. Hara, H. Anai, and K. Yokoyama. Sum of roots, polynomial spectral factorization, and control performance limitations. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 2968–2973, New Orleans, Louisiana USA, December 2007.

[26] H. Anai and S. Hara. A parameter space approach to fixed-order robust controller synthesis by quantifier elimination. *International Journal of Control*, 79(11):1321–1330, November 2006.

[27] I. A. Fotiou, P. Rostalski, P. A. Parrilo, and M. Morari. Parametric optimization and optimal control using algebraic geometry methods. *International Journal of Control*, 79(11):1340–1358, November 2006.