Functiones et Approximatio
42.1 (2010), 37–49

AN ALGORITHM FOR FINDING LOW DEGREE RATIONAL SOLUTIONS TO THE SCHUR COEFFICIENT PROBLEM

Vladimir Bolotnikov

Abstract: We present an algorithm producing all rational functions f with prescribed $n+1$ Taylor coefficients at the origin and such that $\|f\|_\infty \leqslant 1$ and $\deg f \leqslant k$ for every fixed $k \geqslant n$. The case where $k < n$ is also discussed.

Keywords: Schur problem, low degree rational interpolants.

1. Introduction

Let H^∞ be the Banach space of bounded analytic functions on the open unit disk $\mathbb{D}$ with norm $\|f\|_\infty := \sup_{z\in\mathbb{D}} |f(z)| < \infty$. The closed unit ball $\mathcal{S}$ of H^∞ (sometimes called *the Schur class*) thus consists of analytic functions mapping $\mathbb{D}$ into its closure. The classical Schur problem which we will denote by $\mathbf{SP}_n$ consists of finding $f \in \mathcal{S}$ having prescribed $n+1$ Taylor coefficients at the origin.

$\mathbf{SP}_n$: *Given $c_0, \dots, c_n \in \mathbb{C}$, find all functions $f \in \mathcal{S}$ of the form*

$$f(z) = c_0 + c_1 z + \ldots + c_n z^n + O(z^{n+1}). \tag{1.1}$$

The problem has a solution if and only if the Pick matrix of the problem given by

$$P_n = I - \mathcal{T}(c_n, \dots, c_0)\mathcal{T}(c_n, \dots, c_0)^*$$

is positive semidefinite. Here and in what follows, I denotes the identity matrix of the size always clear from the context, and $\mathcal{T}(c_0, \dots, c_n)$ stands for the lower triangular Toeplitz matrix with the bottom row entries indicated in the parentheses:

$$\mathcal{T}(c_n, \dots, c_0) := \begin{bmatrix} c_0 & 0 & 0 & \cdots & 0 \\ c_1 & c_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}. \tag{1.2}$$

Mathematics Subject Classification: 41A05, 41A20, 30E05

If $P_n \geqslant 0$ is singular, then the problem $\mathbf{SP}_n$ has a unique solution which is a finite Blaschke product of degree equal to the rank of P_n. In what follows, we assume that the data set $\{c_0, \dots, c_n\}$ is such that $P_n > 0$ and we will call such a data set *admissible*. For an admissible data set, the parametrization of the solution set of the problem $\mathbf{SP}_n$ was established in [7] via the famous Schur algorithm which we now recall. Starting with $c_0, \dots, c_n$, define the numbers $c_k^{(j)}$ $(j = 1, \dots, n; k = 0, \dots, n-j)$ from the following recursion:

$$\begin{bmatrix} c_0^{(0)} \\ c_1^{(0)} \\ \vdots \\ c_n^{(0)} \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} c_0^{(j+1)} \\ c_1^{(j+1)} \\ \vdots \\ c_{n-j-1}^{(j+1)} \end{bmatrix} = M_j^{-1} \begin{bmatrix} c_1^{(j)} \\ c_2^{(j)} \\ \vdots \\ c_{n-j}^{(j)} \end{bmatrix} \quad (j \geqslant 0), \tag{1.3}$$

where the matrix

$$M_j = \mathcal{T}\left(-\overline{c}_0^{(j)} c_{n-j-1}, \ \dots, \ -\overline{c}_0^{(j)} c_2^{(j)}, \ -\overline{c}_0^{(j)} c_1^{(j)}, 1 - |c_0^{(j)}|^2\right)$$

is defined via formula (1.2). Let

$$\gamma_j = c_0^{(j)} \qquad \text{for } j = 0, \dots, n. \tag{1.4}$$

If $c_0, \dots, c_n$ are the Taylor coefficients of an $f \in \mathcal{S}$, then the numbers γ_i constructed above are the $n+1$ first Schur parameters of f and condition $P_n > 0$ is equivalent to $|\gamma_i| < 1$ for $i = 0, \dots, n$. The Schur algorithm relies on the following fact:

A function f belongs to $\mathcal{S}$ and satisfies (1.1) *if and only if it is of the form*

$$f(z) = \frac{z f_1(z) + c_0}{z \overline{c}_0 f_1(z) + 1} \tag{1.5}$$

for some $f_1 \in \mathcal{S}$ such that $f_1(z) = c_0^{(1)} + c_1^{(1)} z + \ldots + c_{n-1}^{(1)} z^{n-1} + O(z^n)$ where $c_0^{(1)}, \dots, c_{n-1}^{(1)}$ are the numbers defined via (1.3).

Starting with a function $f_0 := f \in \mathcal{S}$ of the form (1.1) and applying recursion (1.5) n times one gets a sequence of Schur class functions satisfying

$$f_j(z) = \frac{z f_{j+1}(z) + c_0^{(j)}}{z \overline{c}_0^{(j)} f_{j+1}(z) + 1} = \frac{z f_{j+1}(z) + \gamma_j}{z \overline{\gamma}_j f_{j+1}(z) + 1} \qquad (j = 0, \dots, n) \tag{1.6}$$

and such that $f_j(z) = c_0^{(j)} + c_1^{(j)} z + \ldots + c_{n-j}^{(j)} z^{n-j} + O(z^{n-j+1})$ where $c_k^{(j)}$ are the numbers defined via (1.3). Upon taking the superposition of linear fractional transformations (1.6) one gets the linear fractional formula

$$f = \mathbf{T}_\Theta[\mathcal{E}] := \frac{A\mathcal{E} + B}{C\mathcal{E} + D} \tag{1.7}$$

which parametrizes all solutions to the $\mathbf{SP}_n$ where the free parameter $\mathcal{E} := f_n$ runs through $\mathcal{S}$ and the coefficient matrix $\Theta = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is given by

$$\Theta(z) = W_0(z)W_1(z)\cdots W_n(z) \qquad \text{where } W_j(z) = \begin{bmatrix} z & \gamma_j \\ z\bar{\gamma}_j & 1 \end{bmatrix}. \tag{1.8}$$

Motivated by engineering applications (where it is desirable for the solution f of an interpolation problem to be rational and of small McMillan degree), the rational coefficient interpolation problem (as well as its multi-point analogs) was considered in [1] with an additional constraint on the degree (complexity) of rational interpolants. In what follows, the polynomials N_f and D_f will denote the numerator and the denominator from the coprime representation $f = N_f/D_f$ of a rational function f. By $\deg f = \max\{\deg N_f, \deg D_f\}$ we mean the McMillan degree of f. The algebra of rational functions will be denoted by $\mathcal{R}$ and we will let

$$\mathcal{R}_k := \{f \in \mathcal{R} :\ \deg f = k\} \quad \text{and} \quad \mathcal{R}_{\leqslant k} := \{f \in \mathcal{R} :\ \deg f \leqslant k\}\,.$$

Being adapted to the single-point case, the problem formulated in [1] is:

$\mathbf{RP}_{n,k}$: *Given* $c_0, \ldots, c_n \in \mathbb{C}$ *and* $k \geqslant 0$, *find all* $f \in \mathcal{R}_{\leqslant k}$ *of the form* (1.1).

The problem was solved in [1] and in [2] (for the matrix-valued case) as follows.

Theorem 1.1. *Let q denote the rank of the Hankel matrix $H = [c_{i+j-1}]_{i,j\geqslant 1}$ constructed from the given numbers c_j (the matrix H is $\frac{n-1}{2} \times \frac{n-1}{2}$ if n is odd or $\frac{n-2}{2} \times \frac{n}{2}$ if n is even). Then*

1. *There is no $f \in \mathcal{R}_k$ satisfying* (1.1) *for every $k < q$ or $q < k \leqslant n - q$.*
2. *There exists at most one function f of complexity $k = q$ subject to* (1.1).
3. *For every $k > n-q$, there are infinitely many solutions of the problem $\mathbf{RP}_{n,k}$ which are parametrized by the formula*

$$f = \mathbf{T}_{\mathfrak{A}}[g] := \frac{\mathfrak{A}_{11}g + \mathfrak{A}_{12}}{\mathfrak{A}_{21}g + \mathfrak{A}_{22}}, \tag{1.9}$$

where the coefficients $\mathfrak{A}_{ij}$ are polynomials explicitly constructed from the data set and such that

$$\deg \begin{bmatrix} \mathfrak{A}_{11} \\ \mathfrak{A}_{21} \end{bmatrix} = q \qquad \text{and} \qquad \deg \begin{bmatrix} \mathfrak{A}_{12} \\ \mathfrak{A}_{22} \end{bmatrix} = n + 1 - q,$$

and where the parameter $g = N_g/D_g \in \mathcal{R}$ is such that

$$\deg N_g \leqslant k - q, \quad \deg D_g \leqslant k + q - n - 1, \quad \mathfrak{A}_{21}(0)N_g(0) + \mathfrak{A}_{22}(0)D_g(0) \neq 0.$$

We refer to [2] for more details. In what follows, we use notation

$$\mathcal{SR} = \mathcal{S} \cap \mathcal{R}, \qquad \mathcal{SR}_k = \mathcal{S} \cap \mathcal{R}_k \qquad \text{and} \qquad \mathcal{SR}_{\leqslant k} = \mathcal{S} \cap \mathcal{R}_{\leqslant k}$$

for the classes of functions in $\mathcal{R}$, $\mathcal{R}_k$ and $\mathcal{R}_{\leqslant k}$ respectively, which are bounded by one in modulus on $\mathbb{D}$. Upon imposing both H^∞-norm and complexity constraints (i.e., upon combining problems $\mathbf{SP}_n$ and $\mathbf{RP}_{n,k}$) we arrive at the following interpolation problem.

$\mathbf{RSP}_{n,k}$: *Given an admissible data set $c_0, \dots, c_n$ and $k \geqslant 0$, find all functions $f \in \mathcal{SR}_{\leqslant k}$ of the form* (1.1).

One may try to treat the latter problem using either formula (1.9) or (1.7). In the first case, the complexity of f is completely controlled by the complexity of the corresponding parameter g and it suffices to pick up all parameters g with $\deg g \leqslant k - q$ leading via formula (1.9) to Schur-class functions f. However, this task is hard, since formula (1.9) does not control $\|\mathbf{T}_{\mathfrak{A}}[g]\|_\infty$ in terms of $\|g\|_\infty$. It may happen that a Schur class parameter g produces $f \notin \mathcal{S}$ and on the other hand, a Schur class function $f \in \mathcal{SR}_{\leqslant k}$ may arise from a non-Schur class parameter g. Although Theorem 1.1 guarantees that there are infinitely many functions $f \in \mathcal{R}_{n+1-q}$ of the form (1.1), it is not known whether or not one of them is of the Schur class. The question about the minimal possible k for which the problem $\mathbf{RSP}_{n,k}$ has a solution, is still open.

It is not even clear from (1.9) that the problem $\mathbf{RSP}_{n,k}$ has solutions for k large enough. On the other hand, the affirmative answer for the latter question is readily seen from parametrization formula (1.7) which in contrast to (1.9), perfectly controls the H^∞-norm of f: every Schur-class rational solution to the problem $\mathbf{SP}_n$ arises via formula (1.7) from some Schur-class rational parameter $\mathcal{E}$. The complexities of interpolants are controlled here to some extent. A straightforward induction argument deduces from (1.8) that the coefficients A, B, C and D in (1.7) are polynomials of respective degrees $\deg A = n + 1$, $\deg B \leqslant n$, $\deg C \leqslant n + 1$, $\deg D \leqslant n$ and therefore,

$$\deg \mathbf{T}_\Theta[\mathcal{E}] \leqslant n + 1 + \deg \mathcal{E}. \tag{1.10}$$

Letting $\mathcal{E}$ in (1.7) to run through the class of constant functions (not exceeding one in modulus), one gets a family of solutions f of the problem $\mathbf{RSP}_{n,n+1}$, but not all the solutions. It turns out that zero cancellations may occur in (1.7) due to which some solutions to the $\mathbf{RSP}_{n,n+1}$ may arise from non-constant parameters. We also observe that the parameter $\mathcal{E} \equiv 0$ leads via (1.7) to the function $\mathbf{T}_\Theta[0] = B/D \in \mathcal{SR}_{\leqslant n}$ which is therefore, a solution to the problem $\mathbf{RSP}_{n,n}$. The next example shows that this function might be *the only* solution to the $\mathbf{RSP}_{n,n}$.

Example 1.2. Let $|c_0| < 1$ and $c_j = 0$ for $j = 1, \dots, n$. With this data, the problem $\mathbf{RSP}_{n,n}$ has only one solution $f \equiv c_0$. This follows from Theorem 1.1 since in this case $q = 0$.

Otherwise (that is, if $c_j \neq 0$ at least for one $j \geqslant 1$ so that $q \geqslant 1$), Theorem 1.1 guarantees the existence of infinitely many functions $f \in \mathcal{R}_{\leqslant n}$ of the form (1.1), at least one of which ($\mathbf{T}_\Theta[0]$) belongs to $\mathcal{SR}_{\leqslant n}$. As was shown in [4]-[6], the set of such functions is infinite and can be parametrized by polynomials σ with $\deg\sigma \leqslant n$ and with all the roots outside $\mathbb{D}$. More precisely, for every such σ, there exists a unique (up to a common unimodular constant factor) pair of polynomials $a(z)$ and $b(z)$, each of degree at most n and such that

1. $|a(z)|^2 - |b(z)|^2 = |\sigma(z)|^2$ for $|z| = 1$ and
2. the function $f = b/a$ (which belongs to $\mathcal{SR}_n$ by part (1)) satisfies (1.1) and therefore, solves the $\mathbf{RSP}_{n,n}$.

The objective of this note is to present an alternative parametrization of the solution set of the problem $\mathbf{RSP}_{n,k}$ (see Theorem 1.3 below) which relies entirely on parametrization formula (1.7). Some elementary analysis of the Schur algorithm will relate complexities of $\mathcal{E}$ and $\deg \mathbf{T}_\Theta[\mathcal{E}]$ more accurately than in (1.10); this in turn, will allow us to describe all parameters $\mathcal{E} \in \mathcal{SR}$ leading via formula (1.7) to solutions f of the problem $\mathbf{RSP}_{n,k}$ (these parameters will be called *admissible*). Explicit construction of these parameters is given below in terms of certain algorithm which seems to be quite efficient and simple from the computational point of view. Here is the **Algorithm**:

Step 1. Given $c_0, \dots, c_n$, compute the numbers $\gamma_0, \gamma_1, \dots, \gamma_n$ by formula (1.4) *using iteration* (1.3).

Step 2. Using the numbers $\gamma_0, \dots, \gamma_n$ compute the polynomials

$$A_n(z) = \sum_{j=0}^{n} a_j z^j \qquad \text{and} \qquad B_n(z) = \sum_{j=0}^{n} b_j z^j \tag{1.11}$$

from the system of recursions

$$\begin{cases} A_0(z) \equiv \gamma_n, \quad B_0(z) \equiv 1, \\ A_{j+1}(z) = zA_j(z) + \gamma_{n-j-1}B_j(z), \\ B_{j+1}(z) = z\overline{\gamma}_{n-j-1}A_j(z) + B_j(z), \end{cases} \quad (j = 0, \dots, n-1). \tag{1.12}$$

It is readily seen that $B_j(0) = 1$ for $j = 0, \dots, n$. In particular, $b_0 = B_n(0) = 1$.

Step 3. Using the coefficients a_j, b_j from (1.11) *define the lower triangular Toeplitz matrices*

$$\mathbf{A} = \begin{bmatrix} a_n & 0 & \cdots & 0 \\ a_{n-1} & a_n & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ a_1 & a_2 & \cdots & a_n \end{bmatrix}, \qquad \widetilde{\mathbf{B}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \overline{b}_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \overline{b}_{n-1} & \overline{b}_{n-2} & \cdots & 1 \end{bmatrix} \tag{1.13}$$

and compute the lower triangular Toeplitz matrix

$$\mathbf{R} = \mathcal{T}(r_1, r_2 \dots, r_n) := \widetilde{\mathbf{B}}^{-1}\mathbf{A}. \tag{1.14}$$

The three first steps are preliminary and can be carried out in finitely many steps. The last step tells which parameters $\mathcal{E}$ in (1.7) should be taken to get solutions to the problem $\mathbf{RSP}_{n,k}$. We first consider the case where $k = n$.

Step 4. For any n-tuple $\{\alpha_1,\dots,\alpha_n\}$ *of complex numbers, compute the function*

$$\mathcal{E}(z)=\frac{\beta_0+\beta_1z+\ldots+\beta_{n-1}z^{n-1}}{\alpha_0+\alpha_1z+\ldots+\alpha_nz^n} \tag{1.15}$$

where $\beta_0,\dots,\beta_{n-1}$ *are defined by*

$$\begin{bmatrix}\beta_{n-1}\\ \vdots\\ \beta_0\end{bmatrix}=-\mathbf{R}\begin{bmatrix}\alpha_n\\ \vdots\\ \alpha_1\end{bmatrix} \tag{1.16}$$

where $\mathbf{R}$ *is given in* (1.14) *and* α_0 *is such that* $\mathcal{E}\in\mathcal{S}$.

The main result of the paper is the following theorem; the proof will be given in Section 2.

Theorem 1.3. *Let* $\mathcal{E}$ *be constructed as in Step 4 and let* Θ *be as in* (1.8). *Then the function* $f=\mathbf{T}_\Theta[\mathcal{E}]$ (1.7) *solves the problem* $\mathbf{RSP}_{n,n}$ *and conversely, all solutions of the* $\mathbf{RSP}_{n,n}$ *arise in this way.*

Remark 1.4. The only relatively uncertain part in Step 4 is the choice of α_0. However, it is readily seen that for any α_0 satisfying $|\alpha_0|\geqslant\sum_{i=1}^{n}(|\alpha_i|+|\beta_{i-1}|)$, the function $\mathcal{E}$ in (1.16) belongs to the Schur class which immediately gives infinitely many solutions of the problem $\mathbf{RSP}_{n,n}$. To be more precise, let us write (1.15) as

$$\mathcal{E}(z)=\frac{P(z)}{\alpha_0+zQ(z)},$$

where $P(z)=\beta_0+\beta_1z+\ldots+\beta_{n-1}z^{n-1}$ and $Q(z)=\alpha_1+\ldots+\alpha_nz^{n-1}$ and let $\mathbb{D}(c,r)$ denote the disk of radius r centered at c. Then the set of all admissible α_0's (for already chosen $\alpha_1,\dots,\alpha_n$ and $\beta_0,\dots,\beta_{n-1}$) is the exterior (complement) of the domain Ω defined as

$$\Omega=\bigcup_{|z|<1}\mathbb{D}(-zQ(z),|P(z)|).$$

Remark 1.5. It follows from (1.15) that a parameter $\mathcal{E}$ leading to a solution of the $\mathbf{RSP}_{n,n}$ has to satisfy $\mathcal{E}(\infty)=0$. Thus, $\mathcal{E}\equiv0$ is the only admissible constant parameter for the problem $\mathbf{RSP}_{n,n}$. Combining this fact with (1.10), we conclude that every other constant function $\mathcal{E}\in\mathcal{S}$ leads via (1.7) to a solution of $\mathbf{RSP}_{n,n+1}$.

As we have already seen, in contrast to the case $n=k$, the existence of infinitely many solutions of the problem $\mathbf{RSP}_{n,k}$ with $k>n$ is immediate. However, the description of all solutions is even somewhat more complicated. We get this description by an appropriate modification of Step 4 as follows.

Step 4′. Let $k > n$ be fixed and let Θ and $\mathbf{R}$ be as above. All solutions f to the problem $\mathbf{RSP}_{n,k}$ are obtained via formula (1.7) where the parameter $\mathcal{E}$ is either any function from $\mathcal{SR}_{\leqslant k-n-1}$ or a function from $\mathcal{SR}_{\leqslant k}$ of the form

$$\mathcal{E}(z) = \frac{\beta_{n-k} + \beta_{n-k+1}z + \ldots + \beta_{n-1}z^{k-1}}{\alpha_{n-k} + \alpha_{n-k+1}z + \ldots + \alpha_n z^k} \tag{1.17}$$

where the coefficients $\alpha_{n-k+1}, \alpha_{n-k+2}, \ldots, \alpha_n$ and $\beta_{n-k}, \beta_{n-k+1}, \ldots, \beta_{-1}$ are picked up arbitrarily, after which the coefficients $\beta_0, \ldots, \beta_{n-1}$ are defined as in (1.16) and where after all, the coefficient α_{n-k} is chosen so that the function $\mathcal{E}$ of the form (1.17) belongs to the Schur class $\mathcal{S}$.

Justification of Step 4′ will be given in Section 2. In Section 3 we will present a version of Step 4 suitably modified for the case where $k < n$. There we will explain the reasons (by means of parametrization formula (1.7)) for which the algorithm is not efficient for $k < n$.

2. Proof of Theorem 1.3

In this section we justify the algorithm presented in the previous section. Let

$$\Theta_k(z) := W_{n-k}(z) \cdots W_n(z) \tag{2.1}$$

where the factors W_j are defined in (1.8). Comparing (2.1) and (1.8) we see that Θ_n equals the coefficient matrix Θ of the transformation (1.7). It is not hard to check by induction that Θ_k is of the form

$$\Theta_k(z) = \begin{bmatrix} zB_k^\sharp(z) & A_k(z) \\ zA_k^\sharp(z) & B_k(z) \end{bmatrix} \tag{2.2}$$

where the polynomials A_k and B_k are constructed from system (1.12) and where $A_k^\sharp$ and $B_k^\sharp$ are defined as follows:

$$A_k^\sharp(z) = z^k\overline{A_k(1/\bar{z})}, \qquad B_k^\sharp(z) = z^k\overline{B_k(1/\bar{z})}. \tag{2.3}$$

Let us take any $\mathcal{E} = \frac{N_\mathcal{E}}{D_\mathcal{E}} \in \mathcal{SR}$ and substitute it together with formula (2.2) for $\Theta_n = \Theta$ into (1.7):

$$f(z) = \frac{zB_n^\sharp(z)N_\mathcal{E}(z) + A_n(z)D_\mathcal{E}(z)}{zA_n^\sharp(z)N_\mathcal{E}(z) + B_n(z)D_\mathcal{E}(z)}. \tag{2.4}$$

Remark 2.1. The numerator and the denominator in (2.4) do not have common zeros and thus,

$$N_f = zB_n^\sharp N_\mathcal{E} + A_n D_\mathcal{E} \qquad \text{and} \qquad D_f = zA_n^\sharp N_\mathcal{E} + B_n D_\mathcal{E}. \tag{2.5}$$

Proof. Taking determinants in (1.8), (2.1) and (2.2) (with $k = n$) gives

$$\begin{aligned} B_n(z)B_n^\sharp(z) - A_n(z)A_n^\sharp(z) &= \frac{1}{z} \cdot \det \Theta_n(z) \\ &= \frac{1}{z} \cdot \prod_{j=0}^{n} \det W_j(z) = z^n \cdot \prod_{j=0}^{n} (1 - |\gamma_j|^2). \end{aligned} \tag{2.6}$$

Therefore, the only possible common zero for the numerator and the denominator in (2.4) is $z = 0$. But if this is the case, we then have $B_n(0)D_{\mathcal{E}}(0) = D_{\mathcal{E}}(0) = 0$ which is impossible since the Schur function $\mathcal{E}$ cannot have a pole at the origin. ■

We shall now compare McMillan degrees of f and f_1 in formula (1.5).

Lemma 2.2. *Let $f \in \mathcal{SR}$ be of the form* (1.5)*. Then* $\deg f - 1 \leqslant \deg f_1 \leqslant \deg f$. *Moreover,*

$$\deg f_1 = \deg f \Longleftrightarrow f_1(\infty) = 0 \Longleftrightarrow f(\infty) \neq 1/\bar{c}_0 \tag{2.7}$$

and

$$\deg f_1 = \deg f - 1 \Longleftrightarrow f_1(\infty) \neq 0 \Longleftrightarrow f(\infty) = 1/\bar{c}_0. \tag{2.8}$$

Proof. Take f_1 in the form $f_1 = N_{f_1}/D_{f_1}$ and rewrite (1.5) as

$$f(z) = \frac{zN_{f_1}(z) + c_0 D_{f_1}(z)}{z\bar{c}_0 N_{f_1}(z) + D_{f_1}(z)} = \frac{F(z)}{G(z)} \tag{2.9}$$

from which we see that $\deg N_f \leqslant \deg f_1 + 1$, $\deg D_f \leqslant \deg f_1 + 1$ and thus, $\deg f \leqslant \deg f_1 + 1$. Now let us take f in the form $f = \frac{N_f}{D_f}$ and solve equation (1.5) for f_1:

$$f_1(z) = \frac{(N_f(z) - c_0 D_f(z))/z}{D_f(z) - \bar{c}_0 N_f(z)} \tag{2.10}$$

Since $c_0 = f(0) = N_f(0)/D_f(0)$, it follows that the numerator in (2.10) is a polynomial of degree not exceeding $\deg f - 1$. Therefore, $\deg N_{f_1} \leqslant \deg f$, $\deg D_{f_1} \leqslant \deg f$ and thus, $\deg f_1 \leqslant \deg f$. This completes the proof of the first statement.

Since there are only two possibilities for the value of $(\deg f - \deg f_1)$, statements (2.7) are equivalent to (2.8). We next observe that the polynomials F and G in (2.9) do not have common zeros (the proof is the same as in Lemma 2.2) and therefore we can conclude from (2.9) that

$$\deg f = \max\{\deg F, \deg G\}. \tag{2.11}$$

Now we verify (2.7) (or (2.8)) separately for the following three cases.

Case 1. Let $\deg D_{f_1} > \deg N_{f_1} + 1$. Then it follows from (2.9) that $\deg f = \deg D_{f_1} = \deg f_1$ and on the other hand, $f_1(\infty) = 0$ and $f(\infty) = c_0 \neq 1/\bar{c}_0$.

Case 2. Let $\deg D_{f_1} < \deg N_{f_1} + 1$. Then $\deg f = \deg N_{f_1} + 1 = \deg f_1 + 1$ and on the other hand, $f_1(\infty) \neq 0$ and $f(\infty) = 1/\bar{c}_0$.

Case 3. Let $\deg D_{f_1} = \deg N_{f_1} + 1$. Let a_0 and b_0 be the leading coefficients of the polynomials N_{f_1} and D_{f_1} respectively. Then the leading coefficients of F and G are $a_0 + c_0 b_0$ and $\bar{c}_0 a_0 + b_0 = 0$, respectively. Assuming that $\deg F < \deg N_{f_1} + 1$ and $\deg G < \deg N_{f_1} + 1$ we have $a_0 + c_0 b_0 = 0$ and $\bar{c}_0 a_0 + b_0 = 0$ which gives $a_0 = b_0 = 0$ which is a contradiction. Therefore, $\max\{\deg F, \deg G\} = \deg N_{f_1} + 1$ and by (2.11), $\deg f = \deg N_{f_1} + 1 = \deg D_{f_1} = \deg f_1$. Finally, since $\deg D_{f_1} = \deg N_{f_1} + 1$, we have $f_1(\infty) = 0$ and it follows from (2.9) that $f(\infty) = \dfrac{a_0 + c_0 b_0}{\bar{c}_0 a_0 + b_0}$ which is not equal to $1/\bar{c}_0$, since $b_0 \neq 0$ and $|c_0| \neq 1$. ■

Let us apply the backward Schur algorithm (1.6) to a function $\mathcal{E} \in \mathcal{SR}_k$ by letting

$$f_{n+1} = \mathcal{E} \qquad \text{and} \qquad f_j(z) = \frac{z f_{j+1}(z) + \gamma_j}{z\overline{\gamma}_j f_{j+1}(z) + 1} \quad \text{for } j = 0, \dots, n. \tag{2.12}$$

Lemma 2.3. *If* $\deg f_i = \deg f_{i+1} + 1$, *then* $\deg f_j = \deg f_{j+1} + 1$ *for every* $j < i$. *If* $f_i(\infty) = 0$, *then* $f_j(\infty) = 0$ *and* $\deg f_j = \deg f_i$ *for every* $j > i$.

Proof. If $\deg f_i = \deg f_{i+1} + 1$, then by virtue of (2.8) (with f, f_1 and c replaced respectively by f_i, f_{i+1} and γ_i) we have $f_i(\infty) = \frac{1}{\overline{\gamma}_i} \neq 0$. Then again by (2.8) (applied to the new triple f_{i-1}, f_i and γ_{i-1}) we get $\deg f_{i-1} = \deg f_i + 1$ and therefore, $f_{i-1}(\infty) = \frac{1}{\overline{\gamma}_{i-1}} \neq 0$. The first statement then follows by induction.

We now assume that $f_i(\infty) = 0$. Since $f_i(\infty) \neq \frac{1}{\overline{\gamma}_i}$, we conclude from (2.7) that $f_{i+1}(\infty) = 0$ and $\deg f_{i+1} = \deg f_i$. The induction argument completes the proof of the second statement. ■

Proof of Theorem 1.3. Let f be a solution to the problem $\mathbf{RSP}_{n,n}$, i.e., f is a rational Schur-class function of degree at most n satisfying equality (1.1). Then f is of the form (1.7) for some rational Schur-class function $\mathcal{E}$ or equivalently, $f = f_0$ is obtained from $\mathcal{E} = f_n$ via recursion (2.12). Then we necessarily have

$$\deg \mathcal{E} \leqslant n \qquad \text{and} \qquad \mathcal{E}(\infty) = 0. \tag{2.13}$$

Indeed, $\deg f \geqslant \deg \mathcal{E}$ by Lemma 2.2 and since $\deg f \leqslant n$ by the assumption, the first relation in (2.13) follows. If we assume that $\mathcal{E}(\infty) \neq 0$, then we get by virtue of (2.8), that $\deg f_{n-1} = \deg \mathcal{E} + 1$ and then we also have $\deg f = \deg \mathcal{E} + n + 1 \geqslant n + 1$ (by the first statement in Lemma 2.3) which contradicts the assumption. Thus, $\mathcal{E}(\infty) = 0$. Due to (2.13) we can take $\mathcal{E}$ in the form (1.15), i.e., we can let

$$N_{\mathcal{E}}(z) = \sum_{j=0}^{n-1} \beta_j z^j \qquad \text{and} \qquad D_{\mathcal{E}}(z) = \sum_{j=0}^{n} \alpha_j z^j. \tag{2.14}$$

It remains to show that the coefficients α_i and β_i are related as in (1.16). Observe, that the polynomials A_n and B_n constructed in (1.12) are of degree at most n;

we take them in the form (1.11) so that the reflected polynomials $A_n^\sharp$ and $B_n^\sharp$ (see (2.3)) are given by

$$A_n^\sharp(z)=\sum_{j=0}^{n}\overline{a}_{n-j}z^j \qquad \text{and} \qquad B_n^\sharp(z)=\sum_{j=0}^{n}\overline{b}_{n-j}z^j. \tag{2.15}$$

Substituting (1.11), (2.14) and (2.15) into (2.5) we get

$$N_f(z)=z^{n+1}\cdot\sum_{\ell=0}^{n-1}\left(\sum_{j=0}^{n-\ell-1}(\overline{b}_{n-\ell-j-1}\beta_{n-j-1}+a_{\ell+j+1}\alpha_{n-j})\right)z^\ell+P_1(z),$$

$$D_f(z)=z^{n+1}\cdot\sum_{\ell=0}^{n-1}\left(\sum_{j=0}^{n-\ell-1}(\overline{a}_{n-\ell-j-1}\beta_{n-j-1}+b_{\ell+j+1}\alpha_{n-j})\right)z^\ell+P_2(z),$$

where P_1 and P_2 are polynomials of degree at most n. The two latter formulas imply that $\deg f\leqslant n$ if and only if

$$\sum_{j=0}^{n-\ell-1}(\overline{b}_{n-\ell-j-1}\beta_{n-j-1}+a_{\ell+j+1}\alpha_{n-j})=0 \qquad (\ell=0,\dots,n-1), \tag{2.16}$$

$$\sum_{j=0}^{n-\ell-1}(\overline{a}_{n-\ell-j-1}\beta_{n-j-1}+b_{\ell+j+1}\alpha_{n-j})=0 \qquad (\ell=0,\dots,n-1). \tag{2.17}$$

Making use of the Toeplitz matrices

$$\begin{array}{ll}\mathbf{A}=\mathcal{T}(a_1,a_2,\dots,a_n), & \mathbf{B}=\mathcal{T}(b_1,b_2,\dots,b_n),\\ \widetilde{\mathbf{A}}=\mathcal{T}(\overline{a}_{n-1},\overline{a}_{n-2},\dots,\overline{a}_0), & \widetilde{\mathbf{B}}=\mathcal{T}(\overline{b}_{n-1},\overline{b}_{n-2},\dots,\overline{b}_0),\end{array} \tag{2.18}$$

and of the vectors

$$\boldsymbol{\alpha}=\begin{bmatrix}\alpha_n\\ \vdots\\ \alpha_1\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\beta}=\begin{bmatrix}\beta_{n-1}\\ \vdots\\ \beta_0\end{bmatrix}, \tag{2.19}$$

one can write equations (2.16) and (2.17) in the matrix form as

$$\widetilde{\mathbf{B}}\boldsymbol{\beta}+\mathbf{A}\boldsymbol{\alpha}=0 \qquad \text{and} \qquad \widetilde{\mathbf{A}}\boldsymbol{\beta}+\mathbf{B}\boldsymbol{\alpha}=0 \tag{2.20}$$

respectively. Since $b_0=B_n(0)=1$, the matrix $\widetilde{\mathbf{B}}$ is invertible. Then we get from the first equation in (2.20)

$$\boldsymbol{\beta}=-\widetilde{\mathbf{B}}^{-1}\mathbf{A}\boldsymbol{\alpha}=-\mathbf{R}\boldsymbol{\alpha} \tag{2.21}$$

which is the same as (1.16). We thus showed that every solution f to the problem $\mathbf{RSP}_{n,n}$ can be obtained via the Schur algorithm from a parameter $\mathcal{E}\in\mathcal{S}$ of the form (1.15), (1.16).

To show that any such parameter is admissible, we have to verify that the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ related as in (2.21) satisfy both equations in (2.20). The first equation is clearly equivalent to (2.21). Substituting (2.21) into the second equation and taking into account that all the matrices in (2.18) commute, we get

$$\widetilde{\mathbf{A}}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\alpha} = -\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}^{-1}\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\alpha} = \widetilde{\mathbf{B}}^{-1}\left(\mathbf{B}\widetilde{\mathbf{B}} - \mathbf{A}\widetilde{\mathbf{A}}\right)\boldsymbol{\alpha}. \tag{2.22}$$

We next substitute formulas (1.11) and (2.15) into (2.6) and examine the coefficients of $z^{2n-\ell}$ for $\ell = 0, \ldots, n-1$ to get equalities

$$\sum_{j=0}^{\ell}(b_{n+j-\ell}\overline{b}_j - a_{n+j-\ell}\overline{a}_j) = 0 \qquad (\ell = 0, \ldots, n-1), \tag{2.23}$$

which can be written in terms of matrices (2.18) as $\mathbf{B}\widetilde{\mathbf{B}} = \mathbf{A}\widetilde{\mathbf{A}}$. We now conclude from (2.22) that the second equation in (2.20) is satisfied. Thus, for every $\mathcal{E} \in \mathcal{S}$ of the form (1.15), (1.16), the coefficients α_i, β_i satisfy equalities (2.16), (2.17) (i.e., equalities (2.20)), which in turn guarantees that the McMillan degree of the function f obtained from $\mathcal{E}$ via the Schur algorithm, does not exceed n. Since this f belongs to $\mathcal{S}$ and satisfies (1.1), it solves the problem $\mathbf{RSP}_{n,n}$. ■

Justification of Step 4′: Let $k > n$ be a fixed integer. Every solution f to the problem $\mathbf{RSP}_{n,k}$ is of the form (1.7) for some rational parameter $\mathcal{E} \in \mathcal{SR}$ with $\deg \mathcal{E} \leqslant k$. We have either $\mathcal{E}(\infty) \neq 0$ or $\mathcal{E}(\infty) = 0$. In the first case, $\deg f = \deg \mathcal{E} + n + 1$ (by Lemmas 2.2 and 2.3) and therefore, $\deg \mathcal{E} \leqslant k - n - 1$. On the other hand, it follows from (1.10) that $\deg \mathbf{T}_{\Theta}[\mathcal{E}] \leqslant k$ for every $\mathcal{E} \in \mathcal{SR}_{\leqslant k-n-1}$. In the second case, we can take $\mathcal{E}$ in the form (1.17), that is to let

$$N_{\mathcal{E}}(z) = \sum_{j=0}^{k-1}\beta_{n-k+j}z^j \qquad \text{and} \qquad D_{\mathcal{E}}(z) = \sum_{j=0}^{k}\alpha_{n-k+j}z^j.$$

Substituting the latter formulas along with (1.11) and (2.15) into (2.5) we get the formulas for N_f and D_f as in the proof of Theorem 1.3 but with the factor z^{k+1} (rather than z^{n+1}) on the left and with polynomials P_1 and P_2 of degree at most k. Then we conclude that $\deg f \leqslant k$ if and only if conditions (2.20) hold which is equivalent to (2.21).

3. Concluding remarks

In conclusion we present a version of the main algorithm for the case where $k < n$. The three first steps are the same as before; the last step describing all admissible parameters in parametrization formula (1.7) is the following.

Step 4″. Let $k < n$ be fixed and let Θ and $\mathbf{R} = \mathcal{T}(r_1, r_2, \ldots, r_n)$ be as above. All solutions f to the problem $\mathbf{RSP}_{n,k}$ are obtained via formula (1.7) *where the* *parameter $\mathcal{E}$ is a Schur-class function of the form*

$$\mathcal{E}(z) = \frac{N_{\mathcal{E}}(z)}{D_{\mathcal{E}}(z)} = \frac{\beta_0 + \beta_1 z + \ldots + \beta_{k-1}z^{k-1}}{\alpha_0 + \alpha_1 z + \ldots + \alpha_k z^k}. \tag{3.1}$$

where the coefficients $\alpha_0, \ldots, \alpha_k$ and $\beta_0, \ldots, \beta_{k-1}$ satisfy the system

$$\begin{bmatrix} \beta_{k-1} \\ \vdots \\ \beta_1 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} r_n & 0 & \cdots & 0 \\ r_{n-1} & r_n & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ r_{n-k+1} & r_{n-k} & \cdots & r_n \end{bmatrix} \begin{bmatrix} \alpha_k \\ \vdots \\ \alpha_2 \\ \alpha_1 \end{bmatrix}, \tag{3.2}$$

$$\begin{bmatrix} r_1 & r_2 & \cdots & r_{k+1} \\ r_2 & r_3 & \cdots & r_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n-k} & r_{n-k+1} & \cdots & r_n \end{bmatrix} \begin{bmatrix} \alpha_k \\ \alpha_{k-1} \\ \vdots \\ \alpha_0 \end{bmatrix} = 0, \tag{3.3}$$

(the matrix in (3.3) *is of Hankel structure).*

Proof. As in the proof of Theorem 1.3 we first observe that every solution f of the problem $\mathbf{RSP}_{n,k}$ is of the form (1.7) for some $\mathcal{E} \in \mathcal{SR}_{\leqslant k}$ subject to $\mathcal{E}(\infty) = 0$. Therefore, $\mathcal{E}$ can be taken in the form (3.1). Substituting (1.11), (2.15) and (3.1) into (2.5) we now get N_f and D_f the polynomials of degree at most $n + k$. Then equating the coefficients of $z^{k+\ell}$ of these polynomials to zero for $j = \ell, \ldots, n-1$, we get necessary and sufficient conditions (similar to (2.16) and (2.17)) for $\deg f = \max\{\deg N_f, \deg D_f\}$ not to exceed k. These conditions are

$$\sum_{j=0}^{\min\{n-\ell-1,k-1\}} \overline{b}_{n-\ell-j-1}\beta_{k-j-1} + \sum_{j=0}^{\min\{n-\ell-1,k\}} a_{\ell+j+1}\alpha_{k-j} = 0,$$

$$\sum_{j=0}^{\min\{n-\ell-1,k-1\}} \overline{a}_{n-\ell-j-1}\beta_{k-j-1} + \sum_{j=0}^{\min\{n-\ell-1,k\}} b_{\ell+j+1}\alpha_{k-j} = 0$$

($\ell = 0, \ldots, n$) and it is not hard to see that they can be written in the matrix form (2.20) as

$$\widetilde{\mathbf{B}}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\alpha} = 0 \qquad \text{and} \qquad \widetilde{\mathbf{A}}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\alpha} = 0 \tag{3.4}$$

respectively where the matrices $\mathbf{A}$, $\mathbf{B}$, $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ are the same as in (2.18) and where now

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_k \\ \vdots \\ \alpha_1 \\ \alpha_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_{k-1} \\ \vdots \\ \beta_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.5}$$

Since $\mathbf{B}\widetilde{\mathbf{B}} = \mathbf{A}\widetilde{\mathbf{A}}$, it follows as in the proof of Theorem 1.3, that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ solve the system (3.4) if and only if they are related as in (2.21). Substituting (3.5) into (2.21) and comparing the k top entries in the obtained equality, we get (3.2); comparison of the $n-k$ bottom entries gives (3.3). ■

Remark 3.1. Although Step 4″ looks very similar to Step 4 in Section 1, in fact it is much less efficient. Let us demonstrate this by the case where $k = n-1$. Then condition (3.3) takes the form

$$r_1\alpha_{n-1} + r_2\alpha_{n-2} + \ldots + r_{n-1}\alpha_1 + r_n\alpha_0 = 0 \tag{3.6}$$

and if $r_n \neq 0$, then a_0 is uniquely determined by $\alpha_1, \ldots, \alpha_{n-1}$. The problem is to describe all the tuples $\{\alpha_1, \ldots, \alpha_{n-1}\}$ (which now are the only free parameters) for which the function

$$\mathcal{E}(z) = \frac{\beta_0 + \beta_1 z + \ldots + \beta_{n-2}z^{n-2}}{\alpha_0 + \alpha_1 z + \ldots + \alpha_{n-1}z^{n-1}} \tag{3.7}$$

with the coefficients $\alpha_0, \beta_0, \ldots, \beta_{n-2}$ determined by formulas (3.6) and (3.2) (with $k = n-1$), belongs to the Schur class. The problem is hard; at the moment we even do not know necessary and sufficient conditions for the existence of at least one such tuple.

References

[1] A. C. Antoulas, B. D. O. Anderson, *On the Scalar Rational Interpolation Problem*, IMA J. Math. Control Inform. **3** (1986), no. 2-3, 61–88.

[2] A. C. Antoulas, J. A. Ball, J. Kang, J. C. Willems, *On the solution of the minimal rational interpolation problem*, Linear Algebra Appl. **137/138** (1990), 511–573.

[3] J. A. Ball, I. Gohberg, L. Rodman. *Interpolation of rational matrix functions*, OT45, Birkhäuser Verlag, 1990.

[4] C. I. Byrnes, A. Lindquist, S. V. Gusev, A. S. Matveev, *A complete parameterization of all positive rational extensions of a covariance sequence*, IEEE Trans. Automat. Control 40 (1995), 18411857.

[5] C. I. Byrnes, S. V. Gusev, A. Lindquist, *A convex optimization approach to the rational covariance extension problem*, SIAM J. Contr. and Optimiz. **37** (1998), 211–229.

[6] T. T. Georgiou, *The interpolation problem with a degree constraint*, IEEE Trans. Automat. Control **44** (1999), no. 3, 631–635.

[7] I. Schur, *Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind*, J.Reine Angew. Math. **147** (1917), 205Â-232.

Address: Department of Mathematics, The College of William and Mary, Williamsburg VA 23187-8795, USA.

E-mail: vladi@math.wm.edu

Received: 12 February 2009; **revised:** 25 July 2009