The Problem of Matrix Padé Approximation

Xu Guo-liang, A. Bultheel

Report TW 116, November 1988,
This versiion Revised March 1989



Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

The Problem of Matrix Padé Approximation

Xu Guo-liang, A. Bultheel

Report TW 116, November 1988,
This versiion Revised March 1989

Department of Computer Science, K.U.Leuven

Abstract

Some proposals are made to give a general definition of matrix Padé approximants. Depending on the normalization of the denominator we define type I (constant term is the unit matrix) or type II (by conditions on the leading coefficient) approximants. Existence and uniqueness are considered, determinant expressions are given and relation among type I/II and left/right approximants are considered. Also some ideas about the computation are included.

The Problem of Matrix Padé Approximation

Xu Guo-liang A. Bultheel

March 1989

Abstract

Some proposals are made to give a general definition of matrix Padé approximants. Depending on the normalization of the denominator we define type I (constant term is the unit matrix) or type II (by conditions on the leading coefficient) approximants. Existence and uniqueness are considered, determinant expressions are given and relation among type I/II and left/right approximants are considered. Also some ideas about the computation are included.

Contents

§1. Introduction

§2. The Definition of Matrix Padé Approximation

§3. The Existence of MPAs

§4. Determinant Expressions for MPAs

§5. Relations Among the Different MPAs

§6. Properties of MPAs

§7. Duality of MPAs

§8. Uniqueness of MPAs

§9. Direct Computation

§10. Recursive Computation

§11. Conclusion

1 Introduction

The problem of classical (scalar) Padé approximation is well established and a number of textbooks appeared on the subject. E.g., the books from Baker and Graves-Morris [1] are among the more recent ones. However, when the given power series has matrix coefficients, it is not at all clear how the notion of Padé approximant should be generalized. In this paper we shall concentrate on possible definitions in the matrix case. To avoid a duplication of well known results, we shall suppose that the reader is sufficiently familiar with the problem of scalar Padé approximation.

There are already many papers that generalize the scalar Padé approximation problem in one way or another, but it is our conviction that the problem has never been considered in all its generality. There are e.g., many papers on the vector case, or on the square matrix case or even approximants in a general non-commutative algebra we studied. However, the non commutativity of the matrix multiplication is not so essential a complication, if this were the only one. Indeed some papers did appear for the case of a square matrix function where the scalar normality condition

immediately generalizes to the matrix case. Under this normality condition, all the recursive computations etc. from the scalar case do generalize (see e.g. [3, 4]). The more general problem of Padé approximation in a non commutative algebra has been considered by a lot of authors. A. Draux has contributions in this area and has compiled a commented bibliography of about 300 references on this topic [8]. However in the rectangular matrix case, the problem is more complicated since not only is $AB \neq BA$ in general, but it can be that BA does not even exist, even if AB does.

Also the vector case has attracted some attention in the literature. Different vector-Padé interpolants were introduced by P.R. Graves-Morris (see e.g., [10]). The problem here is to approximate a number of functions by rational functions with a common denominator. Because the denominator is a scalar polynomial, the problem is relatively simple. There is e.g., no problem in its normalization or in pinning down its degree. In the general case, the denominator is a square matrix polynomial for which the problem of defining a normalization or defining its degree leaves a lot of possibilities.

There is also an extensive literature in linear system theory on the problem of minimal partial realization which is a problem that in a certain sense is equivalent with Padé approximation (see e.g. [6]). If the system has many inputs and many outputs, then one gets a matrix approximation problem which is closely related to a matrix Padé problem. The main difference is that the given power series there is a series in z^{-1} while the numerator and denominator of the rational approximant are polynomials in z. Moreover, minimality is an important issue in those applications. In the matrix case, this means that the degree of the determinant of the denominator has to be minimal. An effort has been done to translate these results into Padé results (see [5]). Therefore a reformulation of the Padé problem as a minimal Padé problem is given, and it is the latter that is generalized to the matrix case. In this text we shall refer to it as the problem mPA or MmPA in the matrix case.

In a number of papers, the first author has considered a general matrix Padé problem for a rectangular matrix series (see e.g., [16],[17]). It appeared that the solutions of the MmPA problem were matrix Padé approximants that were not covered by this theory. This motivated a more general approach to the problem of defining matrix Padé approximants.

2 The Definition of Matrix Padé Approximation

Let $f(z)$ be a given power series with matrix coefficients, i.e.,

$$f(z) = \sum_{k=0}^{\infty} c_k z^k, \qquad c_k \in \mathbf{C}^{p \times m}, \tag{2.1}$$

where $\mathbf{C}^{p \times m}$ consists of all $p \times m$ matrices with their elements in the complex plane $\mathbf{C}$. We want to determine a (right) rational approximating function $N(z)M(z)^{-1}$ to f by the following equation:

$$f(z)M(z) - N(z) = R(z). \tag{2.2}$$

Where $N(z)$ and $M(z)$ are $p \times m$ and $m \times m$ matrices respectively with polynomials as their elements. We call $R(z)$ the residual of the approximant. We shall impose some lower bounds on the orders of the entries in $R(z)$ and some upper bounds on the degrees of the entries in $M(z)$ and $N(z)$.

This seems to be a fair proposal to generalize the definition of Padé approximation (PA) to the matrix case. However, depending on the purposes which one has in mind, this leaves us with still many ways of defining the matrix Padé approximants (MPAs) as we shall see later. Whatever the definition is, since we refer to them as matrix Padé approximants, the MPAs should have some relations or connections with classical (i.e., scalar) PAs. Therefore, there are some principles in defining the approximants which we should follow. Some of these are listed below.

(i) (matrix) MPAs should be a generalization of (scalar) PAs. This means that when $p = m = 1$ the matrix Padé approximants should coincide with the scalar Padé approximants.

(ii) Solvability. In general the defining equations should have a unique solution. By this we mean that the number of unknowns i.e., the number of free parameters to be determined in the approximants, is equal to the number of approximation conditions, i.e., the number of equations which determine these unknowns. We say "in general" since it may happen that in some cases the problem has none or many solutions.

(iii) The MPAs thus defined should have some nice properties just as scalar PAs have. E.g., invariance under linear fractional transformation (see section 6).

In the theory of linear systems with many inputs and many outputs, people have defined a minimal partial realization problem which is closely related to the problem of Padé approximation. This relation is fully established in the scalar case, but because of the lack of a good definition of matrix Padé approximants, this has never been given explicitly for the matrix case. The report [5] by A. Bultheel and M. Van Barel and the Ph. D. thesis [15] seem to be exceptions. Based on a number of reports on the minimal partial realization problem (see [14]), the authors have in [5] defined a kind of matrix Padé approximants which they called Minimal Padé Approximants (mPAs) because they have minimality properties in the degrees of numerator and denominator. Because it is a non standard definition, and because we shall need it in the sequel, we repeat here the definition of the mPA problem (see [5]) :

Definition 2.1 (mPA Problem) *Given f(z), some nonnegative integer ω and an integer $\delta \in [-\omega, \omega]$, find polynomials $M(z)$ and $N(z)$ such that*

1. $f(z)M(z) - N(z) = O(z^{\omega+1}), \qquad z \to \infty$
2. $\eta_j \leq \mu_j - \delta$ *for* $j = 1, \ldots, m$. *Here η_j and μ_j are the degrees of the j-th column of $N(z)$ and $M(z)$ respectively.*
3. $\sum_{j=1}^{m} \mu_j$ *as low as possible.*
4. $M(z)$ *is $C1$ canonical. This is a kind of normalization which we shall define later on when we come to this problem.*

One can see from the definition of the mPA problem that, even in the scalar case it is different from the of the PA problem. But mPA does have a close relationship with PA. In fact, any solution of the scalar mPA problem is one of the PAs in the Padé table (see [5]). It is one of the main advantages of considering problem mPA instead of problem PA that we can define a table of minimal Padé approximants which is parametrized in only two scalar parameters ω and δ, *even in the matrix case*. In the scalar case, this ω-δ–table is essentially the same as the classical Padé table (see [7]). Therefore we shall also try to define Padé approximants in the matrix case such that they are solutions of an MmPA problem. The last problem setting was the one which motivated this work.

In recent year many extentions and generalizations of the scalar Padé problem have been established. Matrix Padé approximants were, in our opinion, never treated in a sufficiently general form. So we can ask why there has not been a good definition of MPAs so far. What makes this problem so much more complicated than its scalar counterpart and what can we do about it ? We give some possible reasons below and we shall try to propose a solution.

(a) **Non-commutativity :** The multiplication of matrices is not commutative in general.

(b) **Choice of degrees and orders :** There are many possibilities in choosing the degrees of the numerator and denominator and the order of the residual.

(c) **Normalization :** There are many ways of normalizing a solution to the problem. We shall consider two possibilities and refer to them as type I and type II approximants.

(a) **Non-commutativity :** As a consequence of this non-commutative multiplication we have to define right MPAs and left MPAs separately. The relation (2.2) is for a right approximant. A left approximant $\tilde{M}(z)^{-1}\tilde{N}(z)$ would, instead of (2.2), satisfy an equation like

$$\tilde{M}(z)f(z) - \tilde{N}(z) = \tilde{R}(z)$$

where now $\tilde{N}(z)$ and $\tilde{M}(z)$ are $p \times m$ and $p \times p$ polynomial matrices respectively.

(b) **Choice of degrees and orders :** Concerning this point we try to start with the most general situation, in which we choose a degree for each entry of the numerator and a degree for each entry of the denominator. Also with each entry of the residual we associate an order. To this purpose, we introduce the following notations:

$$\mathbf{H}_k := \{p(z) : p(z) = \sum_{i=0}^{k} a_i z^i, \quad a_i \in \mathbf{C}\},$$

and

$$\mathbf{E}_k := \{e(z) : e(z) = \sum_{i=k+1}^{\infty} a_i z^i, \quad a_i \in \mathbf{C}\}.$$

The nonnegative integers are denoted by $\mathbf{Z}_+$ and $\mathbf{Z}_+^{p\times m}$ denotes all the $p \times m$ matrices with elements in $\mathbf{Z}_+$. For $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, we introduce the notations

$$\mathbf{H}_V^{p\times m} = \{P(z) = (p_{i,j}(z))_{i,j=1}^{p,m} : p_{ij}(z) \in \mathbf{H}_{v_{ij}}\},$$

and

$$\mathbf{E}_V^{p\times m} = \{R(z) = (r_{i,j}(z))_{i,j=1}^{p,m} : r_{ij}(z) \in \mathbf{E}_{v_{ij}}\}.$$

Sometimes we shall express the fact that $R(z) \in \mathbf{E}_V^{p\times m}$ as $R(z) = O(z^{V+1})$. It will happen frequently that we use column degrees (orders) or row degrees (orders). We shall denote this as follows: If $V = (v_j) \in \mathbf{Z}_+^{1\times m}$, then the notation $\mathbf{H}_V^{p\times m}$ shall denote the same set as before but all the entries in the j-th column now have the same degree v_j. These v_j-s are the column degrees. Similarly, if $V = (v_i)^T \in \mathbf{Z}_+^{p\times 1}$, then all the entries in the i-th row of the elements in $\mathbf{H}_V^{p\times m}$ have the same degree v_i: the row degree. If all entries of the elements in $\mathbf{H}_V^{p\times m}$ have the same degree $V \in \mathbf{Z}_+$, we shall denote this as $\mathbf{H}_V^{p\times m}$. Similar conventions will also be used in relation with the orders of the entries of the elements of $\mathbf{E}_V^{p\times m}$.

Having the previous conventions in mind, we define, in addition to the usual operations of matrices, the following operations applicable to the degree-matrices or order-matrices.

(i) If $V \in \mathbf{Z}_+^{s\times t}$ and $W \in \mathbf{Z}_+$, then $V + W = (v_{ij} + W)_{i,j=1}^{s,t}$.
(ii) If $V \in \mathbf{Z}_+^{s\times 1}$ and $W \in \mathbf{Z}_+^{1\times t}$, then $V + W = (v_i + w_j)_{i,j=1}^{s,t}$.

It may be somewhat confusing to add matrices of different dimensions but if you just remember that a row matrix should be expanded to a full matrix of appropriate dimension with identical elements on every row, a column matrix should be expanded to a full matrix with identical column elements and a scalar is expanded to a full matrix with all identical elements, then this operation is not so unusual. An expression $W - V = U$ with W and V in $\mathbf{Z}_+^{s\times t}$ and $U \in \mathbf{Z}_+$ is a short way of telling you that the difference between the corresponding elements in W and V is the constant value U for each entry. Along the same line a relation $V \geq 0$ for some $V \in \mathbf{Z}_+^{s\times t}$ has to be understood on an elementwise basis. It means that all the entries of V are nonnegative.

(c) **Normalization :** For the normalization of $N(z)M(z)^{-1}$, we consider two possible types. The first type is to let

$$M(0) = I. \tag{2.3}$$

If $M(z)$ is an element from $\mathbf{H}_U^{m\times m}$ for some $U=(u_{ij})\in\mathbf{Z}_+^{m\times m}$, then, under the normalizing condition (2.3), the j-th column of $M(z)$ has $\sum_{k=1}^m u_{kj}$ unkowns to be determined by approximation conditions.

In the scalar case, a popular alternative for the previous normalization (which is called a comonic normalization) is to normalize the denominator as a monic polynomial, by choosing the leading coefficient as 1. For the matrix case, it is not clear what we should take as the leading coefficient, since each element has its own degree, hence its own leading coefficient. It can be shown (see [13]) that by using unimodular transformations, any square matrix polynomial can be brought into a form such that its row degrees are equal to its column degrees. The leading row (or column) coefficient matrix M^{hr} (or M^{hc}) is then the matrix formed by the leading row (or column) coefficients. The second method for normalization we shall use in this paper is to impose conditions on M^{hc} and M^{hr}. We shall then say that $M(z)$ is brought into a canonical form. What we mean by $C1$ (or $C2$) canonical is defined below (see[5]).

Definition 2.2 ($C1/2$ canonical) *We say that $M(z)\in\mathbf{H}_V^{m\times m}$ is $C1$ (or $C2$) canonical iff*

(i) There is a vector $U\in\mathbf{Z}_+^{m\times 1}$ such that $M(z)\in\mathbf{H}_U^{m\times m}\cap\mathbf{H}_{U^T}^{m\times m}$. This means that row and column degrees are the same.

(ii) The leading row (or column) coefficient matrix M^{hr} (or M^{hc}) of $M(z)$ is the unit matrix.

(iii) The leading column (or row) coefficient matrix M^{hc} (or M^{hr}) of $M(z)$ is unit upper (or lower) triangular.

Note that in the scalar case $C1$ and $C2$ canonical means the same thing viz. $M(z)$ is monic. Note also that if $M(z)$ is $C1$ canonical with degree $U=(u_i)_{i=1}^m$, then there are

$$\sum_{\substack{u_i>u_j\\1\le i\le m}} u_j+\sum_{\substack{u_i\le u_j\\1\le i\le m}} u_i+\sum_{\substack{u_i>u_j\\1\le i<j}} 1$$

unknowns for determining the j-th column of $M(z)$.

Now we are ready to give the following definitions of type I and type II right MPAs. Type I refers to the normalization (2.3) and type II refers to a $C1$ canonical normalization.

Definition 2.3 (Type I MPA) *Let $f(z)\in\mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V=(v_{ij})\in\mathbf{Z}_+^{p\times m}$, $U=(u_{ij})\in\mathbf{Z}_+^{m\times m}$ and $W=(w_{ij})\in\mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^p w_{kj}=\sum_{k=1}^p v_{kj}+\sum_{k=1}^m u_{kj},\quad j=1,2,\ldots,m, \tag{2.4}$$

and $W-V\ge 0$. Then the right MPA problem of the first type is denoted by ${}^R(V,U,W)_I^f$ and consists in finding polynomials $N(z)\in\mathbf{H}_V^{p\times m}$ and $M(z)\in\mathbf{H}_U^{m\times m}$ such that

1. $f(z)M(z)-N(z)\in\mathbf{E}_W^{p\times m}$.
2. $M(0)=I$.

The collection of all rational functions $N(z)M(z)^{-1}$ formed by the solutions of ${}^R(V,U,W)_I^f$ is denoted by ${}^R[V,U,W]_I^f$.

Definition 2.4 (Type II MPA) *Let $f(z)\in\mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V=(v_{ij})\in\mathbf{Z}_+^{p\times m}$, $U=(u_j)\in\mathbf{Z}_+^{m\times 1}$ and $W=(w_{ij})\in\mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^p w_{kj}=\sum_{k=1}^p v_{kj}+\sum_{\substack{u_i>u_j\\1\le i\le m}} u_j+\sum_{\substack{u_i\le u_j\\1\le i\le m}} u_i+\sum_{\substack{u_i>u_j\\1\le i<j}} 1,\quad j=1,2,\ldots,m, \tag{2.5}$$

and $W-V\ge 0$. Then the right MPA problem of the second type is denoted by ${}^R(V,U,W)_{II}^f$ and consists in finding polynomials $N(z)\in\mathbf{H}_V^{p\times m}$ and $M(z)\in\mathbf{H}_U^{m\times m}$ such that

1. $f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}$.

2. $M(z)$ *is* $C1$ *canonical with degree* U.

The solution set of ${}^R(V,U,W)_{II}^f$ is denoted by ${}^R[V,U,W]_{II}^f$.

Conditions (2.4) and (2.5) make the number of (scalar) coefficients to be found equal to the numbers of (scalar) equations. Therefore we shall call them *solvability equalities*.

Besides these two possibilities for the normalization of the denominator, there are of course many other choices that could be made. Each of them can give another type of Padé approximant and it is not sure, as we shall show for the type I and type II approximants that the different normalizations are equivalent in all situations. They give in many cases completely different approximants.

There are some special cases of the above definitions which we should mention here, because each of them can be regarded as a different definition of MPA. We shall make different choices for the degrees and orders and see what (dis)advantages this may have. Before we do this, we introduce the following lemma.

Lemma 2.1 *Let* $U \in \mathbf{Z}_+^{m\times 1}$. *Then*

$$\sum_{i=1}^{m} u_i = \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1\le i\le m}} u_i + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j = 1,2,\ldots,m, \tag{2.6}$$

if and only if

$$\begin{cases} |u_i - u_j| \le 1, \quad for \quad \forall i,j \\ u_1 \ge u_2 \ge \cdots \ge u_m. \end{cases} \tag{2.7}$$

Proof. (2.6) holds if and only if

$$\sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_i = \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j = 1,2,\ldots,m,$$

i.e.,

$$\sum_{\substack{u_i > u_j \\ 1\le i\le m}} (u_i - u_j) = \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j = 1,2,\ldots,m,$$

or

$$\sum_{\substack{u_i > u_j \\ 1\le i< j}} (u_i - u_j) + \sum_{\substack{u_i > u_j \\ j\le i\le m}} (u_i - u_j) = \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j = 1,2,\ldots,m.$$

It follows that

$$u_i - u_j = 1, \quad \text{for} \quad u_i > u_j,\ 1 \le i < j, \quad \text{and} \quad u_i - u_j \le 0 \quad \text{for} \quad j \le i \le m.$$

This is (2.7). □

Note that (2.7) means that all the numbers u_i are equal to each other except for a possible drop by 1 at a certain moment, after which all the numbers stay at this value.

Now we list some special cases.

D1. V, U and $W \in \mathbf{Z}_+$.

If $p = m$, this is the most popular case. Most of the literature on matrix Padé approximation in the past decades is about this situation (see [8] for references). If $p \neq m$, the corresponding problem was discussed in [16]. In this case, both (2.4) and (2.5) lead to the same solvability equality

$$pw = pv + mu.$$

This means that p should be a divisor of mu. This may be a severe limitation in practice.

D2. $V \in \mathbf{Z}_+^{1\times m}$, $U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{1\times m}$.

In this case, (2.4) becomes

$$pw_j = pv_j + mu_j, \quad j = 1, 2, \ldots, m.$$

This implies that p should divide mu_j for every j. Condition (2.5) leads to

$$pw_j = pv_j + \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

This means that the total number represented by the last three terms of the right hand side should be divisible by p.

D3. $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{p\times 1}$.

In this case, (2.4) becomes

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_i + mu_j, \quad j = 1, 2, \ldots, m.$$

This implies that

$$u_1 = u_2 = \cdots = u_m. \tag{2.8}$$

Equation (2.5) leads to (2.7).

D4. $V \in \mathbf{Z}_+^{1\times m}$, $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+^{1\times m}$.

In this case, (2.4) becomes

$$pw_j = pv_j + \sum_{i=1}^{m} u_i, \quad j = 1, 2, \ldots, m.$$

This means that $\sum_{i=1}^{m} u_i$ should be divisible by p and $W - V \in \mathbf{Z}_+$, i.e., $w_j - v_j$ does not depend on j. For type II approximants (2.5) becomes

$$pw_j = pv_j + \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

The situation is the same as for D2.

D5. $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{1\times m}$.

In this case, (2.4) becomes

$$pw_j = \sum_{k=1}^{p} v_i + mu_j, \quad j = 1, 2, \ldots, m,$$

and (2.5) becomes

$$pw_j = \sum_{k=1}^{p} v_i + \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

These are not always possible.

D6. $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+^{m\times 1}$.

Now (2.4) becomes

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k + \sum_{k=1}^{m} u_k, \tag{2.9}$$

and this is always possible. Hence this case is important in this paper. Condition (2.5) becomes

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

This holds only for U satisfying (2.7).

There are also other possibilities by giving row and column degrees and orders. In fact, if we confine V, U and W to be row or column vectors separately, there are in total eight possibilities that can be considered.

By the same approach, the left MPA can be defined, and we use the notation ${}^L(\tilde{V}, \tilde{U}, \tilde{W})_I^f$ and ${}^L(\tilde{V}, \tilde{U}, \tilde{W})_{II}^f$ to denote the first type and second type left problems, and ${}^L[\tilde{V}, \tilde{U}, \tilde{W}]_I^f$ and ${}^L[\tilde{V}, \tilde{U}, \tilde{W}]_{II}^f$ to denote their solutions sets. Writing out all the details is an easy task which we gladly leave to the reader. We should only mention that, in the definition of the second type left MPA, the denominator should be $C2$ canonical instead of $C1$ canonical.

3 The Existence of Matrix PAs

It is well known that in the scalar case the PA problem leads to a linear system of equations, which are called Padé equations. There is an analogue to this in the matrix case. In order to express the problems ${}^R(V, U, W)_I^f$ and ${}^R(V, U, W)_{II}^f$ as systems of linear equations, we introduce the following notations. For

$$g(z) = \sum_{k=0}^{\infty} a_k z^k, \quad a_k \in \mathbf{C}^{s\times t},$$

let

$$T_{mn}^l(g) = \begin{bmatrix} a_l & a_{l-1} & \cdots & a_{l-n+1} \\ a_{l+1} & a_l & \cdots & a_{l-n+2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{l+m-1} & a_{l+m-2} & \cdots & a_{l+m-n} \end{bmatrix} \in \mathbf{C}^{ms\times nt}$$

be a block Toeplitz matrix. We adopt the usual convention that $a_k = 0$ for $k < 0$. If $s = t = 1$, then T^l_{mn} is an ordinary Toeplitz matrix. Let $f = (f_{ij}) \in \mathbf{C}^{p\times m}[[z]]$ be given. For the problem ${}^R(V,U,W)^f_I$ we introduce the matrices $E = (e_{ij}) = W - V$ and

$$ {}^RH^j_I(V,U,W) = \begin{bmatrix} T^{v_{1j}}_{e_{1j},u_{1j}}(f_{11}) & T^{v_{1j}}_{e_{1j},u_{2j}}(f_{12}) & \cdots & T^{v_{1j}}_{e_{1j},u_{mj}}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_{pj}}_{e_{pj},u_{1j}}(f_{p1}) & T^{v_{pj}}_{e_{pj},u_{2j}}(f_{p2}) & \cdots & T^{v_{pj}}_{e_{pj},u_{mj}}(f_{pm}) \end{bmatrix}, $$

and

$$ {}^RB^j_I(V,U,W) = \begin{bmatrix} T^{v_{1j}+1}_{e_{1j},1}(f_{1j}) \\ T^{v_{2j}+1}_{e_{2j},1}(f_{1j}) \\ \vdots \\ T^{v_{pj}+1}_{e_{pj},1}(f_{pj}) \end{bmatrix}. $$

Using these notations, the problem of determining the j-th column of the denominator $M(z)$ in ${}^R(V,U,W)^f_I$ can be expressed by the following equations.

$$ {}^RH^j_I(V,U,W)\ X = -\ {}^RB^j_I(V,U,W). \tag{3.10} $$

If the j-th column $M_j(z)$ of $M(z)$ is determined, the j-th column $N_j(z)$ of the numerator $N(z)$ can be easily found from

$$ N_j(z) = (fM_j)^{(V_j)}(z), $$

where V_j is the degree vector of $N_j(z)$ and $(fM_j)^{(V_j)}$ is $f(z)M_j(z) \bmod z^{V_j}$. Hence the solvability of ${}^R(V,U,W)^f_I$ is equivalent with the solvability of equation (3.10). Therefore we have

Theorem 3.1 *The problem ${}^R(V,U,W)^f_I$ is solvable if and only if*

$$ \operatorname{rank} {}^RH^j_I(V,U,W) = \operatorname{rank} \left[{}^RB^j_I(V,U,W) \quad {}^RH^j_I(V,U,W) \right], \quad j = 1,2,\ldots,m. \tag{3.11} $$

In order to establish a similar theorem for the problem ${}^L(\tilde V,\tilde U,\tilde W)^f_I$ we can transform it by transposing it into the problem ${}^R(\tilde V^T,\tilde U^T,\tilde W^T)^{f^T}_I$ and then use the above theorem. To emphasize the dependence on f, we denote ${}^RH^j_I(V,U,W)$ and ${}^RB^j_I(V,U,W)$ by ${}^RH^j_I(V,U,W,f)$ and ${}^RB^j_I(V,U,W,f)$ respectively. Then for the solvability of ${}^R(\tilde V^T,\tilde U^T,\tilde W^T)^{f^T}_I$, condition (3.11) is changed into

$$ \operatorname{rank} {}^RH^i_I(\tilde V^T,\tilde U^T,\tilde W^T,f^T) = \operatorname{rank} \left[{}^RB^i_I(\tilde V^T,\tilde U^T,\tilde W^T,f^T) \quad {}^RH^i_I(\tilde V^T,\tilde U^T,\tilde W^T,f^T) \right], $$

$$ i = 1,2,\ldots,p. $$

Let $J_n = (\delta_{i,n+1-j})^n_{i,j=1}$, where

$$ \delta_{ij} = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{for } i \neq j \end{cases} $$

then we have

$$ \left[J_m T^l_{mn}(g) J_n \right]^T = T^{l+m-n}_{nm}. $$

Hence, if we set

$$ J_1 := \operatorname{diag}(J_{\tilde u_{i1}},\ldots,J_{\tilde u_{ip}}) \quad \text{and} \quad J_2 := \operatorname{diag}(J_{\tilde w_{i1}-\tilde u_{i1}},\ldots,J_{\tilde w_{im}-\tilde u_{im}}), $$

then

$$ \left[J_1 \quad {}^RH^i_I(\tilde V^T,\tilde U^T,\tilde W^T,f^T) J_2 \right]^T = {}^LH^i_I(\tilde V,\tilde U,\tilde W,f), $$

and

$$\left\{J_1 \left[{}^{R}B_I^i(\tilde{V}^T, \tilde{U}^T, \tilde{W}^T, f^T) \quad {}^{R}H_I^i(\tilde{V}^T, \tilde{U}^T, \tilde{W}^T, f^T) \right] \mathrm{diag}(1, J_2)\right\}^T$$

$$= \left[\begin{array}{c} {}^{L}B_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) \\ {}^{L}H_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) \end{array} \right],$$

where

$${}^{L}H_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) = \left[\begin{array}{ccc} T^{\tilde{w}_{i1}-\tilde{u}_{i1}}_{\tilde{u}_{i1},\tilde{e}_{i1}}(f_{11}) & \cdots & T^{\tilde{w}_{im}-\tilde{u}_{i1}}_{\tilde{u}_{i1},\tilde{e}_{im}}(f_{1m}) \\ \cdots & \cdots & \cdots \\ T^{\tilde{w}_{i1}-\tilde{u}_{ip}}_{\tilde{u}_{ip},\tilde{e}_{i1}}(f_{p1}) & \cdots & T^{\tilde{w}_{im}-\tilde{u}_{ip}}_{\tilde{u}_{ip},\tilde{e}_{im}}(f_{pm}) \end{array} \right],$$

$${}^{L}B_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) = \left[\begin{array}{ccc} T^{\tilde{w}_{i1}}_{1,\tilde{e}_{i1}}(f_{i1}) & \cdots & T^{\tilde{w}_{im}}_{1,\tilde{e}_{im}}(f_{im}) \end{array} \right],$$

and

$$\tilde{E} = (\tilde{e}_{ij})_{i,j=1}^{p,m} = \tilde{W} - \tilde{V}.$$

Thus the solvability condition for the left problem ${}^{L}(\tilde{V}, \tilde{U}, \tilde{W})_I^f$ can now be formulated as in the next corollary :

Corollary 3.2 *The problem ${}^{L}(\tilde{V}, \tilde{U}, \tilde{W})_I^f$ is solvable if and only if*

$$\mathrm{rank}\, {}^{L}H_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) = \mathrm{rank} \left[\begin{array}{c} {}^{L}B_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) \\ {}^{L}H_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) \end{array} \right], \quad i = 1, 2, \ldots, p. \tag{3.12}$$

In the special case where we consider row degrees for numerator and denominator, and where the order is constant for all entries, we get the following corollary :

Corollary 3.3 *Suppose*

$$V \in \mathbf{Z}_+^{p\times 1}, \quad U \in \mathbf{Z}_+^{m\times 1}, \quad W \in \mathbf{Z}_+,$$

and

$$\tilde{V} = (W - U)^T, \ \tilde{U} = (W - V)^T, \quad \tilde{W} = W.$$

Define

$${}^{R}H_I(V, U, W) = \left[\begin{array}{cccc} T^{v_1}_{\tilde{u}_1,u_1}(f_{11}) & T^{v_1}_{\tilde{u}_1,u_2}(f_{12}) & \cdots & T^{v_1}_{\tilde{u}_1,u_m}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_p}_{\tilde{u}_p,u_1}(f_{p1}) & T^{v_p}_{\tilde{u}_p,u_2}(f_{p2}) & \cdots & T^{v_p}_{\tilde{u}_p,u_m}(f_{pm}) \end{array} \right], \tag{3.13}$$

$${}^{R}B_I^j(V, U, W) = \left[\begin{array}{c} T^{v_1+1}_{\tilde{u}_1,1}(f_{1j}) \\ T^{v_2+1}_{\tilde{u}_2,1}(f_{1j}) \\ \vdots \\ T^{v_p+1}_{\tilde{u}_p,1}(f_{pj}) \end{array} \right], \tag{3.14}$$

and

$${}^{L}B_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f) = \left[\begin{array}{ccc} T^{w}_{1,u_1}(f_{i1}) & \cdots & T^{w}_{1,u_m}(f_{im}) \end{array} \right]. \tag{3.15}$$

Then

(i) ${}^R(V,U,W)_I^f$ *is solvable if and only if*

$$\operatorname{rank}{}^R H_I(V,U,W) = \operatorname{rank}[{}^R B_I^j(V,U,W)\ {}^R H_I(V,U,W)], \quad j=1,2,\ldots,m; \tag{3.16}$$

(ii) ${}^L(\tilde V,\tilde U,\tilde W)_I^f$ *is solvable if and only if*

$$\operatorname{rank}{}^R H_I(V,U,W) = \operatorname{rank}\begin{bmatrix} {}^L B_I^i(\tilde V,\tilde U,\tilde W,f) \\ {}^R H_I(V,U,W)\end{bmatrix} \quad i=1,2,\ldots,p; \tag{3.17}$$

(iii) If the matrix ${}^R H_I(V,U,W)$ *is nonsingular, then both* ${}^R(V,U,W)_I^f$ *and* ${}^L(\tilde V,\tilde U,\tilde W)_I^f$ *are solvable.*

We note that in the case of Corollary 3.3, the existence of both sides MPA is related to the same matrix ${}^R H_I(V,U,W)$, which does not depend on j.

Now we consider the solvability of the type II problem ${}^R(V,U,W)_{II}^f$. Here $V,W\in \mathbf{Z}_+^{p\times m}$ and $U\in \mathbf{Z}_+^{m\times 1}$. To the sequence $(u_i)_1^m$, we associate the matrices

$$C(U)=(c_{ij})_{i,j=1}^m, \quad T(U)=(t_{ij})_{i,j=1}^m \quad \text{and} \quad S(U)=(s_{ij})_{i,j=1}^m, \tag{3.18}$$

where

$$c_{ij}=\begin{cases} 1, & \text{for } i\ge j, \text{ or } u_i\le u_j \text{ and } i<j \\ 0, & \text{otherwise ,}\end{cases} \tag{3.19}$$

$$t_{ij}=\min\{u_i,u_j\}-c_{ij}, \quad \text{and } s_{ij}=t_{ij}+1. \tag{3.20}$$

Assume matrix $M(z)$ is $C1$ canonical with degree U. Then the meaning of the above three matrices is the following:

a. $c_{ij}=1$ means that, in order to make $M(z)$ $C1$ canonical, there is one extra condition at the position (i,j) in determining $M(z)$, otherwise, if $c_{ij}=0$, there is no condition.

b. s_{ij} stands for the number of unknowns in determining $M(z)$ at position (i,j).

c. If $i\ne j$, t_{ij} denotes the degree of $M_{ij}(z)$. If $i=j$, the degree of $M_{ij}(z)$ is $t_{ij}+1$.

Using these notations, the problem for determining the j-th column of the denominator $M(z)$ in the problem ${}^R(V,U,W)_{II}^f$ can be expressed by the following set of equations.

$${}^R H_{II}^j(V,U,W)\ X = -\ {}^R B_{II}^j(V,U,W), \tag{3.21}$$

where

$${}^R H_{II}^j(V,U,W)=\begin{bmatrix} T_{e_{1j},s_{1j}}^{v_{1j}+1}(f_{11}) & T_{e_{1j},s_{2j}}^{v_{1j}+1}(f_{12}) & \cdots & T_{e_{1j},s_{mj}}^{v_{1j}+1}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{e_{pj},s_{1j}}^{v_{pj}+1}(f_{p1}) & T_{e_{pj},s_{2j}}^{v_{pj}+1}(f_{p2}) & \cdots & T_{e_{pj},s_{mj}}^{v_{pj}+1}(f_{pm})\end{bmatrix},$$

$${}^R B_{II}^j(V,U,W)=\begin{bmatrix} T_{e_{1j},1}^{v_{1j}-u_j+1}(f_{1j}) \\ T_{e_{2j},1}^{v_{2j}-u_j+1}(f_{1j}) \\ \vdots \\ T_{e_{pj},1}^{v_{pj}-u_j+1}(f_{pj})\end{bmatrix},$$

and $E=(e_{ij})=W-V$. Therefore we have

Theorem 3.4 *The problem* ${}^{R}(V,U,W)_{II}^{f}$ *is solvable if and only if*

$$\operatorname{rank}{}^{R}H_{II}^{j}(V,U,W) = \operatorname{rank}[{}^{R}B_{II}^{j}(V,U,W)\ {}^{R}H_{II}^{j}(V,U,W)], \quad j = 1,2,\ldots m. \tag{3.22}$$

A similar conclusion can be established for ${}^{L}(\tilde{V},\tilde{U},\tilde{W})_{II}^{f}$.
One special, but important case is the following

Corollary 3.5 *If*

$$V \in \mathbf{Z}_{+}^{p\times 1},\ U \in \mathbf{Z}_{+}^{m\times 1},\ \ W \in \mathbf{Z}_{+},$$

and

$$u_1 \geq u_2 \geq \cdots \geq u_m, \quad |u_1 - u_m| \leq 1, \tag{3.23}$$

and

$$v_1 \leq v_2 \leq \cdots \leq v_p, \quad |v_p - v_1| \leq 1, \tag{3.24}$$

let

$$\tilde{V} = (W-U)^T,\ \tilde{U} = (W-V)^T,\ \ \tilde{W} = W,$$

and define

$${}^{R}H_{II}(V,U,W) = \begin{bmatrix} T_{\tilde{u}_1,u_1}^{v_1+1}(f_{11}) & T_{\tilde{u}_1,u_2}^{v_1+1}(f_{12}) & \cdots & T_{\tilde{u}_1,u_m}^{v_1+1}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{\tilde{u}_p,u_1}^{v_p+1}(f_{p1}) & T_{\tilde{u}_p,u_2}^{v_p+1}(f_{p2}) & \cdots & T_{\tilde{u}_p,u_m}^{v_p+1}(f_{pm}) \end{bmatrix}. \tag{3.25}$$

Then

(i) ${}^{R}(V,U,W)_{II}^{f}$ *is solvable if and only if*

$$\operatorname{rank}{}^{R}H_{II}(V,U,W) = \operatorname{rank}{}^{R}H_{II}(V,U+1,W), \tag{3.26}$$

(ii) ${}^{L}(\tilde{V},\tilde{U},\tilde{W})_{II}^{f}$ *is solvable if and only if*

$$\operatorname{rank}{}^{R}H_{II}(V,U,W) = \operatorname{rank}{}^{R}H_{II}(V-1,U,W), \tag{3.27}$$

(iii) If the matrix ${}^{R}H_{II}(V,U,W)$ *is nonsingular, then both* ${}^{R}(V,U,W)_{II}^{f}$ *and* ${}^{L}(\tilde{V},\tilde{U},\tilde{W})_{II}^{f}$ *are solvable.*

Proof. (i) Under condition (3.23), $s_{ij} = u_i$. This implies

$${}^{R}H_{II}^{j}(V,U,W) = {}^{R}H_{II}(V,U,W).$$

On the other hand, (3.22) implies

$$\begin{aligned} \operatorname{rank}{}^{R}H_{II}(V,U,W) &= \operatorname{rank}[{}^{R}B_{II}^{1}(V,U,W)\ \cdots {}^{R}B_{II}^{m}(V,U,W)\ {}^{R}H_{II}^{j}(V,U,W)] \\ &= \operatorname{rank}{}^{R}H_{II}(V,U+1,W). \end{aligned}$$

This proves (i).
(ii) can be proved by an approach very similar to the one used in Corallary 3.2.
(iii) follows from (i) and (ii). □

4 Determinant Expressions for MPAs

In this section we want to give determinant expressions for the numerator and denominator of the elements in ${}^R[V,U,W]_I^f$ and ${}^R[V,U,W]_{II}^f$. We shall use the notations introduced in the above section. For a given power series

$$g(z) = \sum_{i=0}^{\infty} a_i z^i \quad a_i \in \mathbf{C},$$

let

$$g^{(k)}(z) = \sum_{i=0}^{k} a_i z^i, \quad a_i \in \mathbf{C},$$

and

$$\bar{g}^{(k)}(z) = \sum_{i=k+1}^{\infty} a_i z^i, \quad a_i \in \mathbf{C}.$$

If $g(z) = (g_{ij}(z))_{i,j=1}^{s,t}$ and $V \in \mathbf{Z}_+^{s\times t}$, then we denote $(g_{ij}^{(v_{ij})}(z))_{i,j=1}^{s,t}$ and $(\bar{g}_{ij}^{(v_{ij})}(z))_{i,j=1}^{s,t}$ by $g^{(V)}(z)$ and $\bar{g}^{(V)}(z)$ respectively. The zero vector of k entries is denoted as :

$$\theta_k = [0,0,\ldots,0] \in \mathbf{C}^{1\times k}.$$

Then we have

Theorem 4.1 *Let $N(z)M(z)^{-1} \in {}^R[V,U,W]_I^f$. Then, if the matrix ${}^RH_I^j(V,U,W)$ is nonsingular,*

$$M_{ij}(z) = \frac{1}{\det {}^RH_I^j(V,U,W)} \det \begin{bmatrix} \delta_{ij} & \lambda_I(i,j) \\ {}^RB_I^j(V,U,W) & {}^RH_I^j(V,U,W) \end{bmatrix}, \tag{4.28}$$

$$N_{ij}(z) = \frac{1}{\det {}^RH_I^j(V,U,W)} \det \begin{bmatrix} f_{ij}^{(v_{ij})}(z) & \omega_I(i,j) \\ {}^RB_I^j(V,U,W) & {}^RH_I^j(V,U,W) \end{bmatrix}, \tag{4.29}$$

where

$$\lambda_I(i,j) = [\theta_{u_{1j}}, \ldots, \ \theta_{u_{i-1,j}}, \ z, \ldots, z^{u_{ij}}, \ \theta_{u_{i+1,j}} \ , \ldots, \theta_{u_{m,j}}],$$

$$\omega_I(i,j) = [\delta_I(i,j,1), \ldots, \delta_I(i,j,m)],$$

and

$$\delta_I(i,j,l) = [z f_{il}^{(v_{ij}-1)}(z), \ldots, z^{u_{lj}} f_{il}^{(v_{ij}-u_{lj})}(z)].$$

Proof. The elements $M_{ij}(z)$ and $N_{ij}(z)$ as expressed by the determinants above have the correct degrees u_{ij} and v_{ij} respectively. Also the normalization condition $M(0) = I$ is easily checked. We shall prove that the (i,j)-th element $R_{ij}(z)$ of $R(z) = f(z)M(z) - N(z)$ with the entries of $M(z)$ and $N(z)$ as defined in (4.28) and (4.29) has the correct order w_{ij}. Then we have proved the theorem. As a side result we also have a determinant expression for the residual entries $R_{ij}(z)$.

$$\begin{aligned} R_{ij}(z) &= \sum_{k=1}^{m} f_{ik}(z) M_{kj}(z) - N_{ij}(z) \\ &= \frac{1}{\det {}^RH_I^j(V,U,W)} \det \begin{bmatrix} f_{ij}(z) - f_{ij}^{(v_{ij})}(z) & \sum_{k=1}^{m} f_{ik}(z)\lambda_I(k,j) - \omega_I(i,j) \\ {}^RB_I^j(V,U,W) & {}^RH_I^j(V,U,W) \end{bmatrix} \\ &= \frac{1}{\det {}^RH_I^j(V,U,W)} \det \begin{bmatrix} \bar{f}_{ij}^{(v_{ij})}(z) & z\bar{f}_{i1}^{(v_{ij}-1)} \cdots z^{u_{1j}} \bar{f}_{i1}^{(v_{ij}-u_{1j})}(z) \ z\bar{f}_{i2}^{(v_{ij}-1)} \cdots \\ {}^RB_I^j(V,U,W) & {}^RH_I^j(V,U,W) \end{bmatrix} \end{aligned}$$

We can eliminate a number of terms from the first row by subtracting multiples of the next rows in the above determinant:

$$R_{ij}(z) = \frac{1}{\det {}^R H_I^j(V,U,W)} \det \begin{bmatrix} \bar{f}_{ij}^{(w_{ij})}(z) & z\bar{f}_{i1}^{(w_{ij}-1)} \cdots z^{u_{1j}} \bar{f}_{i1}^{(w_{ij}-u_{1j})}(z) \; z\bar{f}_{i2}^{(w_{ij}-1)} \cdots \\ {}^R B_I^j(V,U,W) & {}^R H_I^j(V,U,W) \end{bmatrix}$$

$$= O(z^{w_{ij}+1}).$$

Now the theorem is proved. □

For the second type MPA, similar formulas can be established.

Theorem 4.2 *Let $N(z)M(z)^{-1} \in {}^R[V,U,W]_{II}^f$. Then, if the matrix ${}^R H_{II}^j(V,U,W)$ is nonsingular,*

$$M_{ij}(z) = \frac{1}{\det {}^R H_{II}^j(V,U,W)} \det \begin{bmatrix} \delta_{ij} z^{s_{jj}} & \lambda_{II}(i,j) \\ {}^R B_{II}^j(V,U,W) & {}^R H_{II}^j(V,U,W) \end{bmatrix}, \tag{4.30}$$

$$N_{ij}(z) = \frac{1}{\det {}^R H_{II}^j(V,U,W)} \det \begin{bmatrix} z^{s_{jj}} f_{ij}^{(v_{ij}-s_{jj})}(z) & \omega_{II}(i,j) \\ {}^R B_{II}^j(V,U,W) & {}^R H_{II}^j(V,U,W) \end{bmatrix}, \tag{4.31}$$

where

$$\lambda_{II}(i,j) = [\theta_{s_{1j}}, \ldots, \theta_{s_{i-1,j}}, 1, \; z \; , \ldots, z^{t_{ij}}, \; \theta_{s_{i+1,j}}, \ldots, \theta_{s_{m,j}}],$$

$$\omega_{II}(i,j) = [\delta_{II}(i,j,1), \ldots, \delta_{II}(i,j,m)],$$

and

$$\delta_{II}(i,j,l) = [f_{il}^{(v_{ij})}(z), \ldots, z^{t_{lj}} f_{il}^{(v_{ij}-t_{lj})}(z)].$$

s_{ij} and t_{ij} are as defined in (3.18)-(3.20).

Proof. The proof is very similar to the previous one. First note that from (4.30) we can see that the degree of $M_{jj}(z)$ is given by

$$\partial M_{jj} = s_{jj} = u_j, \quad \text{and the coefficient of } z^{u_j} \text{ is one.}$$

For $i \neq j$ the degrees of the M_{ij} satisfy

$$\partial M_{ij} \leq t_{ij}.$$

It follows by the construction of $T(U)$ and $S(U)$ (see section 3) that $M(z)$ is $C1$ canonical. Also the degrees for the $N_{ij}(z)$ entries is bounded by v_{ij}. Thus only the order of the residual entries remains to be checked. Calculate the (i,j)-th element $R_{ij}(z)$ of $R(z) = f(z)M(z) - N(z)$

$$R_{ij}(z) = \sum_{k=1}^{m} f_{ik}(z) M_{kj}(z) - N_{ij}(z)$$

$$= \frac{1}{\det {}^R H_{II}^j(V,U,W)} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(v_{ij}-s_{jj})}(z) & \sum_{k=1}^{m} f_{ik}(z)\lambda_{II}(k,j) - \omega_{II}(i,j) \\ {}^R B_{II}^j(V,U,W) & {}^R H_{II}^j(V,U,W) \end{bmatrix}$$

$$= \frac{1}{\det {}^R H_{II}^j(V,U,W)} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(v_{ij}-s_{jj})}(z) & \bar{f}_{i1}^{(v_{ij})} \cdots z^{t_{1j}} \bar{f}_{i1}^{(v_{ij}-t_{1j})}(z) & \bar{f}_{i2}^{(v_{ij})} \cdots \\ {}^R B_{II}^j(V,U,W) & {}^R H_{II}^j(V,U,W) \end{bmatrix}$$

Again we find that by row combinations we can eliminate some of the terms in the first row to find

$$R_{ij}(z) = \frac{1}{\det {}^R H_{II}^j(V,U,W)} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(w_{ij}-s_{jj})}(z) & \bar{f}_{i1}^{(w_{ij})} \cdots z^{t_{1j}} \bar{f}_{i1}^{(w_{ij}-t_{1j})}(z) & \bar{f}_{i2}^{(w_{ij})} \cdots \\ {}^R B_{II}^j(V,U,W) & {}^R H_{II}^j(V,U,W) \end{bmatrix}$$

$$= O(z^{w_{ij}+1}).$$

Now the theorem is proved. □

For the left MPA, similar formulas can be obtained.

5 Relations Among the Different MPAs

The problem we consider in this section is to find some relations among the different definitions of MPA. We try to answer the questions : do the two types of MPAs define the same approximant? and is the solution of problem MmPA a solution of problem ${}^R(V,U,W)_I^f$ or ${}^R(V,U,W)_{II}^f$? We shall show that, in general, the different definitions lead to different approximants. Only in some special cases, they determine the same one.

Theorem 5.1 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$ be given. Then*

(i) if solvability equality (2.5) holds and

$$|u_i - u_j| \le 1, \quad \forall (i,j), \tag{5.32}$$

$$u_1 \ge u_2 \ge \cdots \ge u_m, \tag{5.33}$$

then

$$\{NM^{-1} \in {}^R[V,U,W]_{II}^f : M(0) \text{ is nonsingular}\} \subset {}^R[V,U,W]_I^f.$$

(ii) if solvability equality (2.4) holds, then

$$\{NM^{-1} \in {}^R[V,U,W]_I^f : \exists\, Q \in \mathbf{C}^{m\times m} \text{ s.t. } MQ \text{ is } C1 \text{ canonical with degree } U\}$$

$$\subset {}^R[V',U,W]_{II}^f,$$

where (V',U,W) satisfies (2.5) and $V' \ge V$.

Proof. (i). Under the conditions (5.32) and (5.33), it follows from Lemma 2.1 that the solvability equality (2.5) becomes (2.4). On the other hand, note that we consider row degrees for N and M and row orders for $fM - N$, and that these do not change when the matrices are multiplied from the right with $M(0)^{-1}$. Therefore conclusion (i) holds.

(ii). For the same V, U and W, the right hand side of (2.4) is larger than or equal to the right side of (2.5). Hence if (2.4) holds, then (2.5) holds for (V',U,W) with $V' \ge V$. Therefore (ii) is true. □

The next theorem shows that in some special cases, the second type problem can be transformed into the first one.

Theorem 5.2 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+$ be given such that solvability equality (2.4) holds. (Note that in this case (2.4) is the same as (2.5).) For the given $f(z)$, define*

$$g(z) = (g_{ij})_{i,j=1}^{p,m},$$

where

$$g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1}).$$

Suppose furthermore that there exists a type II solution, i.e.,

$${}^R[V,U,W]_{II}^f \ne \emptyset.$$

Then we have the following :

(i) Any type II solution of f

$$N_2(z)M_2(z)^{-1} \in {}^R[V,U,W]^f_{II},$$

can be transformed into a type I solution for g. More precisely : if we define

$$M_1(z) = M_2(z^{-1})z^u,$$

and

$$N_1(z) = g(z)M_2(z^{-1})z^u - \mathrm{diag}[z^{w_1},\ldots,z^{w_p}]N_2(z^{-1})z^u,$$

then

$$N_1(z)M_1(z)^{-1} \in {}^R[U-1,U,W+U-V-1]^g_I.$$

(ii) Any type I solution

$$N_1(z)M_1(z)^{-1} \in {}^R[U-1,U,W+U-V-1]^g_I,$$

for g can be transformed into a type II solution for f, i.e., if we set

$$M_2(z) = M_1(z^{-1})z^u,$$

and

$$N_2(z) = f^{(W)}(z)M_1(z^{-1})z^u - \mathrm{diag}[z^{w_1},\ldots,z^{w_p}]N_1(z^{-1})z^u,$$

then

$$N_2(z)M_2(z)^{-1} \in {}^R[V,U,W]^f_{II}.$$

Proof. (i). Since the problem ${}^R(V,U,W)^f_{II}$ is independent of $\bar{f}^{(w_i)}_{ij}(z)$, we may consider it as being zero and replace $f_{ij}(z)$ by $f^{(w_i)}_{ij}(z)$. Let $N_2(z)M_2(z)^{-1} \in {}^R[V,U,W]^f_{II}$. Then

$$f(z)M_2(z) - N_2(z) =: R(z) \in \mathbf{E}^{p\times m}_W \cap \mathbf{H}^{p\times m}_{U+W}. \tag{5.34}$$

Let $z = t^{-1}$. Multiplying by $\mathrm{diag}[t^{w_1},\ldots,t^{w_p}]$ on the left and by t^u on the right, we get

$$g(t)M_2(t^{-1})t^u - \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u = N_1(t), \tag{5.35}$$

where

$$\begin{aligned} N_1(t) &= \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]R(t^{-1})t^u \\ &= g(t)M_2(t^{-1})t^u - \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u. \end{aligned} \tag{5.36}$$

Let

$$M_1(t) = M_2(t^{-1})t^u.$$

Then (5.35) can be written as

$$g(t)M_1(t) - N_1(t) = \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u. \tag{5.37}$$

Since $M_2(z)$ is $C1$ canonical it follows that $M_1(0) = I$. From (5.34) and (5.37) we also find that

$$N_1(t) \in \mathbf{H}_{U-1}^{p\times m}$$

and

$$g(t)M_1(t) - N_1(t) \in \mathbf{E}_{W+U-V-1}^{p\times m}.$$

Finally we can note that because of (2.4)

$$\begin{aligned}\sum_{i=1}^{p}(w_i + u_i - v_i - 1) &= \sum_{i=1}^{p} w_i + p(u-1) - \sum_{i=1}^{p} v_i \\ &= mu + p(u-1).\end{aligned}$$

i.e., the solvablity equality holds for the problem ${}^R(U-1, U, W+U-V-1)_I^g$. Hence

$$N_1(t)M_1(t)^{-1} \in {}^R[U-1, U, W+U-V-1]_I^g.$$

(ii). Conclusion (ii) can be proved by a similar approach. □

Now we turn our interests to finding some relations between our definitions of MPAs and problem MmPA in the matrix case. We recall the definition of Kronecker and dual Kronecker indices first.

Definition 5.1 (Kronecker indices) *Given a sequence $a = \{a_0, a_1, \ldots, a_\mu\}$ of elements from $\mathbf{C}^{p\times m}$. Consider for $j = 0, 1, \ldots, \mu$ the matrices*

$$A^{(j)} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{\mu-j} \\ a_1 & a_2 & \cdots & a_{\mu-j+1} \\ \cdots & \cdots & \cdots & \cdots \\ a_j & a_{j+1} & \cdots & a_\mu \end{bmatrix}$$

The dual Kronecker indices of the sequence a are defined for $i = 1, \ldots, p$ by

$$\tilde{u}_i = \begin{cases} \min\{j : \textit{row } jp+i \textit{ is lin. dep. on the previous rows of } A^{(j)}\} \\ \mu + 1 \qquad \textit{if there is no linear dependency for any } j \end{cases}$$

The m Kronecker indices are similarly defined expressing linear dependencies for the columns of the matrices

$$\tilde{A}^{(j)} = \begin{bmatrix} a_0 & a_1 & \cdots & a_j \\ a_1 & a_2 & \cdots & a_{j+1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{\mu-j} & a_{\mu-j+1} & \cdots & a_\mu \end{bmatrix}.$$

The Kronecker and dual Kronecker indices for the empty sequence are supposed to be zero.

Now we can formulate the next theorem :

Theorem 5.3 *Let $f(z) = \sum_{k=0}^{\infty} c_k z^k \in \mathbf{C}^{p\times m}[[z]]$ be given, as well as the integer $\omega \in \mathbf{Z}_+$. Let $U = (u_i)_{i=1}^m \in \mathbf{Z}^{1\times m}$ and $\tilde{U} = (\tilde{u}_i)_{i=1}^p \in \mathbf{Z}^{1\times p}$ be the Kronecker and dual Kronecker indices respectively associate with the sequence $\{c_\omega, c_{\omega-1}, \ldots, c_{-\delta+1}\}$. Finally suppose that NM^{-1} denotes a solution of problem MmPA for the given ω and some $\delta \in \mathbf{Z} \cap [-\omega, \omega]$. Then the following is true :*

(i) If

$$V = \omega - \tilde{U} \quad \textit{and} \quad W = \omega, \tag{5.38}$$

then the solvability equality (2.4) holds for V, U and W. Moreover NM^{-1} is a type I MPA, i.e.,

$$NM^{-1} \in {}^{R}[V, U, W]_{I}^{f},$$

provided $M(0)$ is nonsingular.

(ii) For any $V, W \in \mathbf{Z}_{+}^{p\times 1}$, such that

$$\omega - \tilde{U} \le V \le W \le \omega, \tag{5.39}$$

suppose that the solvability equality (2.5) holds for V, U, W. In that case the MmPA solution NM^{-1} will be a type II MPA i.e.,

$$NM^{-1} \in {}^{R}[V, U, W]_{II}^{f}.$$

Proof. (i). Since (see [5])

$$\sum_{i=1}^{p} \tilde{u}_i = \sum_{i=1}^{m} u_i,$$

$$\sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i = p\omega - \sum_{i=1}^{p} \tilde{u}_i + \sum_{i=1}^{m} u_i = p\omega.$$

i.e., (2.4) holds, it follows from [5] that the degree of the i-th row of N is less than or equal to $\omega - \tilde{u}_i = v_i$. Then $(N, M) \in \mathbf{H}_{V}^{p\times m} \times \mathbf{H}_{U}^{m\times m}$ and $(NM(0)^{-1}, MM(0)^{-1}) \in \mathbf{H}_{V}^{p\times m} \times \mathbf{H}_{U}^{m\times m}$. Hence (i) is valid.

(ii). We have mentioned before that, for the same V and U, the right hand side of (2.4) is not less than the right hand side of (2.5). From (i) we know that, for V and W satisfying (5.38) the equality (2.4) holds for V, U and W. Therefore in order to make (2.5) to become true, we take V and W as in (5.39). This is always possible. Because M is $C1$ canonical, $NM^{-1} \in {}^{R}[V, U, W]_{II}^{f}$. □

Because the type II MPAs and the solutions of the problem MmPA both have a canonical normalization, one can ask if these two problems are equivalent in a certain sense. In the next theorem we shall show that this is true under a normality condition of the function.

In what follows $\lfloor a \rfloor$ will denote the largest integer not exceeding a and $\lceil a \rceil$ the smallest integer larger than or equal to a.

We start with the definition of a normal power series.

Definition 5.2 (normal) *Let $f(z) = \sum_{k=0}^{\infty} c_k z^k$ be given. We say that $f(z)$ is normal if all the matrices H_k^{ω} are nonsingular for any $\omega \in \mathbf{Z}_{+}$ and $k > 0$, satisfying*

$$\lceil k/m \rceil \le \omega, \quad \lceil k/p \rceil \le \omega,$$

where H_k^{ω} is a $k \times k$ matrix defined by

$$H_k^{\omega} = \begin{bmatrix} c_\omega & \cdots & c_{\omega-q_1+1} & c_{\omega-q_1}(p,s) \\ \cdots & \cdots & \cdots & \cdots \\ c_{\omega-q_2+1} & \cdots & c_{\omega-q_1-q_2+2} & c_{\omega-q_1-q_2-1}(p,s) \\ c_{\omega-q_2}(t,m) & \cdots & c_{\omega-q_1-q_2+1}(t,m) & c_{\omega-q_1-q_2}(t,s) \end{bmatrix},$$

with

$$k = q_1 m + s, \quad s < m,$$

$$k = q_2 p + t, \quad t < p,$$

and $c_l(i,j)$ denotes a matrix formed from the first i rows and the first j columns of c_l.

Note that in the scalar case for $p = m = 1$ the above definition reduces to the classical notion of normality.

Theorem 5.4 *For the given (ω, δ) as in the MmPA problem, let*

$$\lambda = \max\{0, -\delta + 1\},$$

$$\mu = \lfloor \frac{p(\omega - \lambda + 1) - m}{p + m} \rfloor,$$

$$\tilde{\mu} = \lfloor \frac{m(\omega - \lambda + 1) - p}{p + m} \rfloor,$$

$$s = \max\{0, p(\omega - \lambda + 1) - (\mu + 1)(p + m)\},$$

and

$$t = \max\{0, m(\omega - \lambda + 1) - (\tilde{\mu} + 1)(p + m)\}.$$

Assume $f(z)$ is normal. Then the problem MmPA is equivalent to the problem ${}^R(V, U, W)^f_{II}$, where

$$V \in \mathbf{Z}_+^{p\times 1}, \quad U \in \mathbf{Z}_+^{m\times 1}, \quad \textit{and} \quad W \in \mathbf{Z}_+$$

are defined by

$$v_i = \omega - \tilde{u}_i, \quad i = 1, 2, \ldots, p,$$

$$u_i = \begin{cases} \mu + 2, & i = 1, 2, \ldots, s, \\ \mu + 1, & i = s + 1, \ldots, m, \end{cases}$$

$$\tilde{u}_i = \begin{cases} \tilde{\mu} + 2, & i = 1, 2, \ldots, t, \\ \tilde{\mu} + 1, & i = t + 1, \ldots, p, \end{cases}$$

and

$$W = \omega.$$

Proof. If the definitions of u_i and $\tilde{u}_i$ have to make sense, we should first show that $s < m$ and $t < p$. Since

$$p(\omega - \lambda + 1) - m = \mu(p + m) + r, \quad r < p + m,$$

it follows that

$$p(\omega - \lambda + 1) - (\mu + 1)(p + m) = r - p < m.$$

Hence $s < m$. Similarly, $t < p$ can be shown.

It is easy to see that (see [5]), under the normality condition of f, $(u_i)_{i=1}^m$ and $(\tilde{u}_i)_{i=1}^p$ thus defined are Kronecker indices and dual Kronecker indices of $c_\omega, \ldots, c_{-\delta+1}$. Hence, it follows as in the proof of Theorem 5.3 that solvability equality (2.4) holds for V, U and W as defined above. It follows then from Lemma 2.1 that (2.5) also holds. Therefore the solution of problem MmPA is in ${}^R[V, U, W]^f_{II}$.

On the other hand, by row exchange and column exchange, the matrix ${}^RH^j_{II}(V, U, W)$, which is the coefficient matrix of the linear system of equations determining the denominator of ${}^R[V, U, W]^f_{II}$, can be transformed into H^ω_k, where $k = \sum_{i=1}^m u_i$. Hence, under the assumption of normality, the solution of ${}^R(V, U, W)^f_{II}$ always exists uniquely. Then ${}^R[V, U, W]^f_{II}$ is the solution of problem MmPA.

This completes the proof of the equivalence. □

6 Properties of MPAs

In this section we derive some invariance properties of MPAs under linear franctional transformations. The first two are very simple to prove and are only mentioned for completeness without including the proof.

Property 6.1 *Let $A \in \mathbf{C}^{p\times p}$ be nonsingular. If $V, W \in \mathbf{Z}_+^{1\times m}$ and $U \in \mathbf{Z}_+^{m\times m}$, then*

$$ {}^R[V,U,W]_I^{Af} = A\ {}^R[V,U,W]_I^f,$$

and

$$ {}^R[V,U,W]_{II}^{Af} = A\ {}^R[V,U,W]_{II}^f.$$

Property 6.2 *Let $A \in \mathbf{C}^{p\times p}$ and $B \in \mathbf{C}^{m\times m}$ be nonsingular. If V, U and W are all in $\mathbf{Z}_+$, then*

$$ {}^R[V,U,W]_I^{AfB} = A\ \ {}^R[V,U,W]_I^f\ B,$$

and

$$ {}^R[V,U,W]_{II}^{AfB} = A\ \ {}^R[V,U,W]_{II}^f\ B.$$

In the following property we consider a linear fractional transformation of the variable z for MPAs of type I.

Property 6.3 *Let*

$$ z = \frac{\alpha y}{1+\beta y}, \quad \alpha \neq 0, \quad \textit{and} \quad g(y) = f(z).$$

Then if $V = U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{1\times m}$,

$$ {}^R[V,U,W]_I^g(y) = \ {}^R[V,U,W]_I^f(z). \tag{6.40}$$

Proof. Let $N(z)M(z)^{-1} \in\ {}^R[V,U,W]_I^f$. Then

$$ f(\frac{\alpha y}{1+\beta y})M(\frac{\alpha y}{1+\beta y}) - N(\frac{\alpha y}{1+\beta y}) = O\left((\frac{\alpha y}{1+\beta y})^{W+1}\right),$$

$$ M(0) = I.$$

Multiplying the first equality on the right by $D(U) = \mathrm{diag}[(1+\beta y)^{u_1}, \ldots, (1+\beta y)^{u_m}]$, we have

$$ g(y)q(y) - p(y) = O(y^{W+1}),$$

$$ q(0) = I,$$

where

$$ q(y) = M(\frac{\alpha y}{1+\beta y})D(U), \quad p(y) = N(\frac{\alpha y}{1+\beta y})D(U).$$

Hence

$$ p(y)q(y)^{-1} = N(z)M(z)^{-1} \in\ {}^R[V,U,W]_I^g,$$

and then

$$ {}^R[V,U,W]_I^f(z) \subset\ {}^R[V,U,W]_I^g(y).$$

Similarly, we can prove

$$ {}^R[V,U,W]_I^g(y) \subset\ {}^R[V,U,W]_I^f(z).$$

Therefore (6.40) holds. □

The following property says that if a polynomial of a sufficiently low degree is added to f, then the type I or II MPA also gets this polynomial added.

Property 6.4 *Let $V, W, E \in \mathbf{Z}_+^{p\times m}$ and $U \in \mathbf{Z}_+^{m\times m}$. If*

$$\max_{1\le k\le m}\{e_{ik}+u_{kj}\} \le v_{ij}, \quad i=1,\dots,p; \quad j=1,\dots,m, \tag{6.41}$$

then for any $B(z) \in \mathbf{H}_E^{p\times m}$, we have

$$^R[V,U,W]_{I,II}^f + B(z) = {}^R[V,U,W]_{I,II}^g, \tag{6.42}$$

where

$$g(z) = f(z) + B(z).$$

Proof. Let $N(z)M(z)^{-1} \in {}^R[V,U,W]_{I,II}^f$. Then

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}.$$

Hence

$$[f(z)+B(z)]M(z) - [N(z)+B(z)M(z)] \in \mathbf{E}_W^{p\times m}.$$

It follows from (6.41) that $N(z)+B(z)M(z) \in \mathbf{H}_V^{p\times m}$. Therefore

$$[N(z)+B(z)M(z)]M(z)^{-1} = N(z)M(z)^{-1} + B(z) \in {}^R[V,U,W]_{I,II}^g.$$

i.e.,

$$^R[V,U,W]_{I,II}^f + B(z) \subset {}^R[V,U,W]_{I,II}^g.$$

The opposite direction can similarly be proved. □

The next property shows how MPAs for successive parts of the expansion of f can be combined to give a MPA for the composed series. A property like this forms the basis of the divide and conquer strategy for the fast algorithm as proposed in [2] for the scalar case.

Property 6.5 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+\setminus\{0\}$ and suppose solvability equality (2.4) or (2.5) holds. Let $\tilde{U} = (\tilde{u}_i) \in \mathbf{Z}_+^{p\times 1}$,*

$$g(z) = \mathrm{diag}(z^{-\tilde{u}_1},\dots,z^{-\tilde{u}_p})[f(z) - f^{(\tilde{U}-1)}(z)],$$

where

$$\tilde{u}_i = \begin{cases} v_i - u + 1, & v_i \ge u, \\ 0, & v_i < u. \end{cases}$$

Then

$$^R[V,U,W]_{I,II}^f = f^{(\tilde{U}-1)}(z) + \mathrm{diag}(z^{\tilde{u}_1},\dots,z^{\tilde{u}_p})\,{}^R[V-\tilde{U},U,W-\tilde{U}]_{I,II}^g.$$

Proof. It is obvious that the solvability equality holds for the problem ${}^R(V-\tilde{U},U,W-\tilde{U})_{I,II}^g$. Let $N_1(z)M_1(z)^{-1} \in {}^R[V-\tilde{U},U,W-\tilde{U}]_{I,II}^g$. Since

$$g(z)M_1(z) - N_1(z) \in \mathbf{E}_{W-\tilde{U}}^{p\times m},$$

we have

$$\mathrm{diag}(z^{\tilde{u}_1},\dots,z^{\tilde{u}_p})[g(z)M_1(z) - N_1(z)] \in \mathbf{E}_W^{p\times m}.$$

Therefore

$$f(z)M_1(z) - [f^{(\tilde{U}-1)}(z)M_1(z) + \mathrm{diag}(z^{\tilde{u}_1},\dots,z^{\tilde{u}_p})N_1(z)] \in \mathbf{E}_W^{p\times m}.$$

By the definition of $\tilde{U}$, we have

$$f^{(\tilde{U}-1)}(z)M_1(z) + \mathrm{diag}(z^{\tilde{u}_1}, \ldots, z^{\tilde{u}_p})N_1(z) \in \mathbf{H}_V^{p\times m}.$$

Therefore

$$f^{(\tilde{U}-1)}(z) + \mathrm{diag}(z^{\tilde{u}_1}, \ldots, z^{\tilde{u}_p})\ {}^R[V-\tilde{U}, U, W-\tilde{U}]_{I,II}^g \subset\ {}^R[V,U,W]_{I,II}^f.$$

For the inverse side of the above relation, assume that $N(z)M(z)^{-1} \in\ {}^R[V,U,W]_{I,II}^f$. Since

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m},$$

also

$$[f(z) - f^{(\tilde{U}-1)}(z)]M(z) - [N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{E}_W^{p\times m}.$$

Then

$$\mathrm{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})\left[[f(z) - f^{(\tilde{U}-1)}(z)]M(z) - [N(z) - f^{(\tilde{U}-1)}(z)M(z)]\right] \in \mathbf{E}_{W-\tilde{U}}^{p\times m},$$

i.e.,

$$g(z)M(z) - \mathrm{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{E}_{W-\tilde{U}}^{p\times m}.$$

Note that

$$N(z) - f^{(\tilde{U}-1)}(z)M(z) \in \mathbf{E}_{\tilde{U}-1}^{p\times m} \cap \mathbf{H}_V^{p\times m},$$

so that we have

$$\mathrm{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{H}_{V-\tilde{U}}^{p\times m},$$

i.e.,

$$\left[\mathrm{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)]\right] M(z)^{-1} \in\ {}^R[V-\tilde{U}, U, W-\tilde{U}]_{I,II}^g.$$

Hence

$${}^R[V,U,W]_{I,II}^f \subset f^{(\tilde{U}-1)}(z) + \mathrm{diag}(z^{\tilde{u}_1}, \ldots, z^{\tilde{u}_p})\ {}^R[V-\tilde{U}, U, W-\tilde{U}]_{I,II}^g.$$

So the proof is completed. □

From the discussions of Sections 2–3, we should note that the problem ${}^R(V,U,W)_{I,II}^f$ can be decoupled column-wise into m sub-problems, i.e., each column of $N(z)$ and $M(z)$ can be determined independently. Now we denote the subproblem for the j-th column by ${}_j^R(V_j,U_j,W_j)_{I,II}^f$, and the set of its solutions (N_j, M_j) by ${}_j^R[V_j,U_j,W_j]_{I,II}^f$, where V_j, U_j and W_j are the j-th columns of V, U and W respectively. Note however that here the solution set is a collection of column polynomial pairs rather than a collection of rational functions. With this notation we can formulate the next property :

Property 6.6 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$ be such that*

$$\sum_{i=1}^{p} w_i = \sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i.$$

Let $(N(z), M(z)) \in\ {}_j^R[V,U,W]_I^f$. If V^*, U^* are the degree vectors of $N(z), M(z)$ and W^* is the highest order of $fM - N$, then for any $V', W' \in \mathbf{Z}_+^{p\times 1}$ and $U' \in \mathbf{Z}_+^{m\times 1}$ satifying

$$V' \geq V^*, \quad U' \geq U^*,$$

$$W' \leq W^*, \quad V' \leq W',$$

$$\sum_{i=1}^{p} w_i' = \sum_{i=1}^{p} v_i' + \sum_{i=1}^{m} u_i',$$

we have

$$(N(z), M(z)) \in\ {}_j^R[V', U', W']_I^f.$$

The proof of the property is simple and we shall ommit it here. □

This property implies that there is something like a block property of the MPA table, but here, in the matrix case, it is not easy to characterize the structure of the blocks because

(i) The MPA table is multi-dimensional.

(ii) The set ${}^{R}_{j}[V,U,W]^{f}_{I}$ may have many elements (see §8). For the different elements in this set, the degrees V^*, U^* and the orders W^* may be different. So the block may have many "levels".

We close this section with one final property.

Property 6.7 *Let $V, W \in \mathbf{Z}_{+}^{p\times 1}$ and $U \in \mathbf{Z}_{+}^{m\times 1}$ such that*

$$\sum_{i=1}^{p} w_i = \sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i.$$

Let $(N(z), M(z)) \in {}^{R}_{j}[V,U,W]^{f}_{I}$, and

$$\|V\| = \max_{1\le i\le p} v_i \le \min_{1\le i\le p} w_i.$$

If for some $A \in \mathbf{C}^{1\times p}$, it happens that

$$Af \in \mathbf{E}^{1\times m}_{\|V\|}, \tag{6.43}$$

then $AN = 0$.

The proof of the property is simple and ommitted □

This property implies that if the rows of f are linearly dependent in the sense of (6.43), then the rows of N inherit the same linear dependency.

7 The Duality of MPA

The duality considered in this section has to be understood in the following sense : To a given right MPA problem, we want to associate a left MPA problem which has the same solutions as the right one. The left problem will be called the dual problem of the right one and conversely. The precise definition is given below.

Definition 7.1 (dual type I problems) *Consider the right and left MPA problems of type I : ${}^{R}(V,U,W)^{f}_{I}$ and ${}^{L}(\tilde{V},\tilde{U},\tilde{W})^{f}_{I}$. If for any $(N,M) \in \mathbf{H}_V \times \mathbf{H}_U$ and $(\tilde{N},\tilde{M}) \in \mathbf{H}_{\tilde{V}} \times \mathbf{H}_{\tilde{U}}$ satisfying*

$$f(z)M(z) - N(z) \in \mathbf{E}^{p\times m}_{W},$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) \in \mathbf{E}^{p\times m}_{\tilde{W}},$$

one has

$$\tilde{N}(z)M(z) = \tilde{M}(z)N(z), \tag{7.44}$$

then we say that these first type left and right problems (and their solution sets) are each others dual.

Definition 7.2 (dual type II problems) *Consider the type II problems* ${}^{R}(V,U,W)^{f}_{II}$ *and* ${}^{L}(\tilde{V},\tilde{U},\tilde{W})^{f}_{II}$. *Define* $U'=(u'_{ij})\in\mathbf{Z}_{+}^{m\times m}$ *and* $\tilde{U}'=(\tilde{u}'_{ij})\in\mathbf{Z}_{+}^{p\times p}$ *by*

$$u'_{ij}=\min\{u_i,u_j\},\quad \tilde{u}'_{ij}=\min\{\tilde{u}_i,\tilde{u}_j\}.$$

If for any $(N,M)\in\mathbf{H}_V\times\mathbf{H}_{U'}$ *and* $(\tilde{N},\tilde{M})\in\mathbf{H}_{\tilde{V}}\times\mathbf{H}_{\tilde{U}'}$ *satisfying*

$$f(z)M(z)-N(z)\in\mathbf{E}_{W}^{p\times m},$$

and

$$\tilde{M}(z)f(z)-\tilde{N}(z)\in\mathbf{E}_{\tilde{W}}^{p\times m},$$

one has

$$\tilde{N}(z)M(z)=\tilde{M}(z)N(z),$$

then we say that these second type left and right problems (and their solution sets) are dual.

Similarly, we can also define mixed dualilty, i.e., when the left and right problems are of different type.

First we shall consider the duality of the type I MPAs, i.e., the duality of definition 7.1.

Lemma 7.1 *Let* $V,W\in\mathbf{Z}_{+}^{p\times m}$, $U\in\mathbf{Z}_{+}^{m\times m}$, $\tilde{V},\tilde{W}\in\mathbf{Z}_{+}^{p\times m}$ *and* $\tilde{U}\in\mathbf{Z}_{+}^{p\times p}$. *Then, under the conditions*

$$\sum_{k=1}^{p}w_{kj}=\sum_{k=1}^{p}v_{kj}+\sum_{k=1}^{m}u_{kj},\quad j=1,2,\ldots,m,\tag{7.45}$$

and

$$\sum_{k=1}^{m}\tilde{w}_{ik}=\sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{k=1}^{p}\tilde{u}_{ik},\quad i=1,2,\ldots,p,\tag{7.46}$$

we have, for $i=1,\ldots,p$ *and* $j=1,\ldots,m$,

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_{ik}+u_{kj}\},\max_{1\le k\le p}\{\tilde{u}_{ik}+v_{kj}\}\right\}\le\min\left\{\min_{1\le k\le m}\{\tilde{w}_{ik}\},\min_{1\le k\le p}\{w_{kj}\}\right\},\tag{7.47}$$

if and only if

1. $w_{ij}=\tilde{w}_{st}=w$, *for any* (i,j) *and* (s,t),
2. $u_{ij}=u_i,\quad j=1,2,\ldots,m,$
3. $v_{ij}=v_i,\quad j=1,2,\ldots,m,$
4. $\tilde{u}_{ij}=\tilde{u}_j,\quad i=1,2,\ldots,p,$
5. $\tilde{v}_{ij}=\tilde{v}_j,\quad i=1,2,\ldots,p,$
6. $u_i+\tilde{v}_i=w,\quad i=1,2,\ldots,m,$
7. $v_i+\tilde{u}_i=w,\quad i=1,2,\ldots,p.$

Proof. The proof of sufficiency is simple. We only prove the necessity part.
(a) Suppose

$$\min_{1\le k\le m}\{\tilde{w}_{ik}\} \ge \min_{1\le k\le p}\{w_{kj}\}, \quad \text{for some } (i,j) \text{ fixed}, \tag{7.48}$$

then

$$\max_{1\le k\le m}\{\tilde{v}_{ik}+u_{kj}\} \le \min_{1\le k\le p}\{w_{kj}\}, \tag{7.49}$$

$$\max_{1\le k\le p}\{\tilde{u}_{ik}+v_{kj}\} \le \min_{1\le k\le p}\{w_{kj}\}. \tag{7.50}$$

Hence

$$\tilde{v}_{ik}+u_{kj} \le \min_{1\le k\le p}\{w_{kj}\}, \quad k=1,2,\ldots,m,$$

$$\tilde{u}_{ik}+v_{kj} \le \min_{1\le k\le p}\{w_{kj}\}, \quad k=1,2,\ldots,p,$$

and then

$$\sum_{k=1}^{m}(\tilde{v}_{ik}+u_{kj}) \le m \min_{1\le k\le p}\{w_{kj}\},$$

$$\sum_{k=1}^{p}(\tilde{u}_{ik}+v_{kj}) \le p \min_{1\le k\le p}\{w_{kj}\}.$$

Using (7.45) and (7.46), we have

$$\begin{aligned}\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj} &\le \sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{k=1}^{m}u_{kj}+p\min_{1\le k\le p}\{w_{kj}\}\\ &= \sum_{k=1}^{m}(\tilde{v}_{ik}+u_{kj})+p\min_{1\le k\le p}\{w_{kj}\}. \end{aligned}\tag{7.8}$$

This implies

$$\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj} \le (m+p)\min_{1\le k\le p}\{w_{kj}\}. \tag{7.9}$$

Therefore we have from (7.48) that

$$\tilde{w}_{ik}=\min_{1\le k\le p}\{w_{kj}\}=w, \ k=1,2,\ldots,m, \tag{7.10}$$

$$w_{kj}=\min_{1\le k\le p}\{w_{kj}\}=w, \ k=1,2,\ldots,p. \tag{7.11}$$

(b) If

$$\min_{1\le k\le m}\{\tilde{w}_{ik}\} < \min_{1\le k\le p}\{w_{kj}\}, \quad \text{for some } (i,j) \text{ fixed}, \tag{7.12}$$

then we have

$$\max_{1\le k\le m}\{\tilde{v}_{ik}+u_{kj}\} \le \min_{1\le k\le m}\{\tilde{w}_{ik}\},$$

and

$$\max_{1\le k\le p}\{\tilde{u}_{ik}+v_{kj}\} \le \min_{1\le k\le m}\{\tilde{w}_{ik}\}.$$

Parallel to the proof of (7.9), we have

$$\sum_{k=1}^{m} \tilde{w}_{ik} + \sum_{k=1}^{p} w_{kj} \leq (m+p) \min_{1\leq k\leq m} \{\tilde{w}_{ik}\}.$$

This contradicts (7.12). Thus (7.48) is true for all (i,j). This in turn implies that (7.10) and (7.11) are valid for all (i,j). This proves that relation 1 of the lemma holds.

It follows from (7.49) and (7.8) that

$$\tilde{v}_{ik} + u_{kj} \leq w, \quad \forall (i,j,k),$$

$$mw \leq \sum_{k=1}^{m} (\tilde{v}_{ik} + u_{kj}).$$

Hence

$$\tilde{v}_{ik} + u_{kj} = w, \quad \forall (i,j,k). \tag{7.13}$$

These relations imply the relations 2, 5 and 6 of the lemma.

Summing up (7.45) and (7.46) and then using (7.13), we have

$$(p+m)w = mw + \sum_{k=1}^{p} (v_{kj} + \tilde{u}_{ik}),$$

and

$$\sum_{k=1}^{p} (v_{kj} + \tilde{u}_{ik}) = pw, \quad \forall (i,j). \tag{7.14}$$

It follows from (7.47) that

$$v_{kj} + \tilde{u}_{ik} \leq w, \quad \forall (i,j,k). \tag{7.15}$$

By (7.14) one has

$$v_{kj} + \tilde{u}_{ik} = w, \quad \forall (i,j,k).$$

Relations 3, 4 and 7 follow from these equalities. □

We are now ready to state the following duality property for type I MPAs.

Theorem 7.2 *Assume $V, W \in \mathbf{Z}_+^{p\times m}$, $U \in \mathbf{Z}_+^{m\times m}$, $\tilde{V}, \tilde{W} \in \mathbf{Z}_+^{p\times m}$ and $\tilde{U} \in \mathbf{Z}_+^{p\times p}$. If relations 1–7 hold, then the problems ${}^{R}(V,U,W)_I^f$ and ${}^{L}(\tilde{V},\tilde{U},\tilde{W})_I^f$ are dual problems in the sense of definition 7.1.*

Proof. Assume $(N,M) \in \mathbf{H}_V \times \mathbf{H}_U$ and $(\tilde{N},\tilde{M}) \in \mathbf{H}_{\tilde{V}} \times \mathbf{H}_{\tilde{U}}$ satisfy

$$f(z)M(z) - N(z) = R(z) \in \mathbf{E}_W^{p\times m}, \tag{7.16}$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) = R_1(z) \in \mathbf{E}_{\tilde{W}}^{p\times m}. \tag{7.17}$$

Multiplying (7.16) from the left by $\tilde{M}(z)$ and (7.17) from the right by $M(z)$ and then subtracting, we get

$$\tilde{N}(z)M(z) - \tilde{M}(z)N(z) = \tilde{M}(z)R(z) - R_1(z)M(z). \tag{7.18}$$

The degree of the (i,j)-th element of the left hand side polynomial of the above equality is less than or equal to

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_{ik}+u_{kj}\},\max_{1\le k\le p}\{\tilde{u}_{ik}+v_{kj}\}\right\}$$

and the order of the (i,j)-th element on the right hand side is greater than or equal to

$$\min\left\{\min_{1\le k\le m}\{\tilde{w}_{ik}\},\min_{1\le k\le p}\{w_{kj}\}\right\}.$$

It follows from Lemma 7.1 that if relations 1–7 hold,

$$\tilde{N}(z)M(z)-\tilde{M}(z)N(z)=0.$$

Thus the theorem is proved. □

Now we consider the duality of the second type MPAs.

Assume we have type II problems ${}^{R}(V,U,W)^{f}_{II}$ and ${}^{L}(\tilde{V},\tilde{U},\tilde{W})^{f}_{II}$. For any $(N,M)\in\mathbf{H}_V\times\mathbf{H}_{U'}$ and $(\tilde{N},\tilde{M})\in\mathbf{H}_{\tilde{V}}\times\mathbf{H}_{\tilde{U}'}$ satisfying

$$f(z)M(z)-N(z)\in\mathbf{E}_W^{p\times m}$$

and

$$\tilde{M}(z)f(z)-\tilde{N}(z)\in\mathbf{E}_{\tilde{W}}^{p\times m},$$

we want the equality

$$\tilde{N}(z)M(z)=\tilde{M}(z)N(z),$$

to be true. Similar to (7.18), we have

$$A\tilde{N}(z)M(z)B-A\tilde{M}(z)N(z)B=A\tilde{M}(z)R(z)B-AR_1(z)M(z)B,$$

where $R\in\mathbf{E}_W^{p\times m}$, $R_1\in\mathbf{E}_{\tilde{W}}^{p\times m}$, $A\in\mathbf{C}^{p\times p}$ and $B\in\mathbf{C}^{m\times m}$. In order to make the right hand side of above equality zero, it is sufficient to have

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_k+u_k\},\max_{1\le k\le p}\{\tilde{u}_k+v_k\}\right\}\le\min_{i,j,p,q}\{\tilde{w}_{ij},w_{pq}\}\equiv w,$$

where

$$v_k=\max_i v_{ki}\quad\text{and}\quad\tilde{v}_k=\max_i\tilde{v}_{ik}.$$

Then

$$\max_{1\le k\le p}\tilde{v}_{kj}+u_j\le w,\quad j=1,2,\ldots,m,$$

$$\tilde{u}_i+\max_{1\le k\le m}v_{ik}\le w,\quad i=1,2,\ldots,p.$$

Summing up the two solvablities:

$$\sum_{k=1}^{m}\tilde{w}_{ik}=\sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{\substack{\tilde{u}_k>\tilde{u}_i\\1\le k\le p}}\tilde{u}_i+\sum_{\substack{\tilde{u}_i\ge\tilde{u}_k\\1\le k\le p}}\tilde{u}_k+\sum_{\substack{\tilde{u}_k>\tilde{u}_i\\1\le k<i}}1,\quad i=1,2,\ldots,p,$$

and

$$\sum_{k=1}^{p}w_{kj}=\sum_{k=1}^{p}v_{kj}+\sum_{\substack{u_i>u_j\\1\le i\le m}}u_j+\sum_{\substack{u_i\le u_j\\1\le i\le m}}u_i+\sum_{\substack{u_i>u_j\\1\le i<j}}1,\quad j=1,2,\ldots,m,$$

we get

$$\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj}=\sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{\substack{\tilde{u}_k>\tilde{u}_i\\1\le k\le p}}\tilde{u}_i+\sum_{\substack{\tilde{u}_i\ge\tilde{u}_k\\1\le k\le p}}\tilde{u}_k+\sum_{\substack{\tilde{u}_k>\tilde{u}_i\\1\le k<i}}1$$

$$+\sum_{k=1}^{p}v_{kj}+\sum_{\substack{u_i>u_j\\1\le i\le m}}u_j+\sum_{\substack{u_i\le u_j\\1\le i\le m}}u_i+\sum_{\substack{u_i>u_j\\1\le i<j}}1$$

$$\le\sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{k=1}^{p}\tilde{u}_k+\sum_{k=1}^{p}v_{kj}+\sum_{k=1}^{m}u_k$$

$$\le\sum_{k=1}^{m}(\max_i\tilde{v}_{ik}+u_k)+\sum_{k=1}^{p}(\max_j v_{kj}+\tilde{u}_k)$$

$$\le(p+m)w.$$

On the other hand,

$$\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj}\ge(p+m)w.$$

Therefore all the inequalities in the estimation above should be equalities. This leads to

a. $w_{ij}=\tilde{w}_{st}=w$, for any (i,j) and (s,t),

b. $u_{ij}=u_i,\quad j=1,2,\ldots,m$, and condition (2.7) in Lemma 2.1 is satisfied.

c. $v_{ij}=v_i,\quad j=1,2,\ldots,m$,

d. $\tilde{u}_{ij}=\tilde{u}_j,\quad i=1,2,\ldots,p$, and condition (2.7) in Lemma 2.1 is satisfied when U is substituted by $\tilde{U}$.

e. $\tilde{v}_{ij}=\tilde{v}_j,\quad i=1,2,\ldots,p$,

f. $u_i+\tilde{v}_i=w,\quad i=1,2,\ldots,m$,

g. $v_i+\tilde{u}_i=w,\quad i=1,2,\ldots,p$.

Therefore we have proved the following

Theorem 7.3 *Let $V,W\in\mathbf{Z}_+^{p\times m}$, $U\in\mathbf{Z}_+^{m\times m}$, and $\tilde{V},\tilde{W}\in\mathbf{Z}_+^{p\times m}$, $\tilde{U}\in\mathbf{Z}_+^{p\times p}$. If relations a–g hold. Then the type II problems ${}^R(V,U,W)_{II}^f$ and ${}^L(\tilde{V},\tilde{U},\tilde{W})_{II}^f$ are dual to each other.*

From the deductions of Lemma 7.1 and Theorem 7.3, we can get mixed duality as in the following corollary.

Corollary 7.4 *Let $V,W\in\mathbf{Z}_+^{p\times m}$, $U\in\mathbf{Z}_+^{m\times m}$, and $\tilde{V},\tilde{W}\in\mathbf{Z}_+^{p\times m}$, $\tilde{U}\in\mathbf{Z}_+^{p\times p}$.*

(i) If relation b in the conditions a–g is replaced by relation 2 in theorem 7.2. Then the problems ${}^R(V,U,W)_I^f$ and ${}^L(\tilde{V},\tilde{U},\tilde{W})_{II}^f$ are dual.

(ii) If relation d in the conditions a–g is replaced by relation 4 in theorem 7.2. Then the problems ${}^R(V,U,W)_{II}^f$ and ${}^L(\tilde{V},\tilde{U},\tilde{W})_I^f$ are dual.

8 Uniqueness of MPA

We begin this section with the discussion of the uniqueness of the solution of the problem ${}^R(V,U,W)^f_I$. The second type problem will be considered later. The left MPA problem can be discussed in a parallel manner, so we do not treat it here. In fact, the uniqueness problem can be solved only in the case where ${}^R(V,U,W)^f_I$ has a dual problem. We know from (7.44) that if ${}^R(V,U,W)^f_I$ and its dual problem have solutions, then both solutions are unique. This section will mainly answer a question like: If a left or a right solution exists, what conclusions can we draw about the uniqueness ?

Let $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+$, and assume relation (2.4) holds. Set further

$$\tilde{V} = (W-U)^T, \quad \tilde{U} = (W-V)^T \text{ and } \tilde{W} = W.$$

We introduce the following sets.

$${}^L R(\tilde{V},\tilde{U},\tilde{W})^f = \{(\tilde{N},\tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}}^{p\times p} \setminus \{0\} : \tilde{M}f - \tilde{N} \in \mathbf{E}_{\tilde{W}}^{p\times m}\}, \tag{8.19}$$

$${}^R R(V,U,W)^f = \{(N,M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m} \setminus \{0\} : fM - N \in \mathbf{E}_W^{p\times m}\} \tag{8.20}$$

and

$${}^R R^*(V,U,W)^f = \{f^*M - N : NM^{-1} \in {}^R[V,U,W]^f_I\}, \tag{8.21}$$

where $f^* = N^*M^{*-1} \in {}^R[V,U,W]^f_I$. Here we suppose ${}^R[V,U,W]^f_I \neq \emptyset$.

Recall also the definition of the matrices ${}^R H(V,U,W)$ of (3.13).

Lemma 8.1 *Suppose f^* has the formal expansion $f^*(z) = \sum_{k=0}^{\infty} c_k^* z^k$. Then for any $R \in {}^R R^*(V,U,W)^f$, there exist $\beta_i \in \mathbf{C}^{|u|}$, $(|u| = \sum_{i=1}^m u_i)$ $i = 1,2,\ldots,m$ such that*

$${}^R H_I(V,U,W)\beta_i = 0, \quad i = 1,\ldots,m, \tag{8.22}$$

and

$$R(z) = z^{w+1} \sum_{k=0}^{\infty} E_k z^k, \tag{8.23}$$

where

$$E_k = G_k[\beta_1,\ldots,\beta_m],$$

with

$$G_k = \begin{bmatrix} T_{1,u_1}^{w+k}(f_{11}^*) & T_{1,u_2}^{w+k}(f_{12}^*) & \cdots & T_{1,u_m}^{w+k}(f_{1m}^*) \\ \cdots & \cdots & \cdots & \cdots \\ T_{1,u_1}^{w+k}(f_{p1}^*) & T_{1,u_2}^{w+k}(f_{p2}^*) & \cdots & T_{1,u_m}^{w+k}(f_{pm}^*) \end{bmatrix} \tag{8.24}$$

Proof. The $R \in {}^R R^*(V,U,W)^f$ can be written as

$$R(z) = f^*(z)M(z) - N(z),$$

for some N and M. If f^* is defined by $f^* = N^*M^{*-1}$, we define ΔM and ΔN by

$$\Delta M = M - M^* = \begin{bmatrix} \sum_{i=1}^{u_1} \beta_{11}^{(i)} z^i & \cdots & \sum_{i=1}^{u_1} \beta_{1m}^{(i)} z^i \\ \cdots & \cdots & \cdots \\ \sum_{i=1}^{u_m} \beta_{m1}^{(i)} z^i & \cdots & \sum_{i=1}^{u_m} \beta_{mm}^{(i)} z^i \end{bmatrix},$$

$$\Delta N = N - N^*.$$

So that we have

$$R(z) = f^*(z)\Delta M(z) - \Delta N(z).$$

The β-vectors can now be defined with the coefficients of the ΔM matrix as follows :

$$\beta_i = [\beta_{1i}^{(1)}, \ldots, \beta_{1i}^{(u_1)}; \beta_{2i}^{(1)}, \ldots, \beta_{2i}^{(u_2)}; \cdots; \beta_{mi}^{(1)}, \ldots, \beta_{mi}^{(u_m)}]^T.$$

We indeed have for these β_i that

$${}^{R}H_I(V, U, W)\beta_i = 0, \quad i = 1, \ldots, m.$$

Since $c_i^* = c_i, \quad$ for $i = 0, 1, \ldots, w$, it follows that

$$\begin{aligned} R(z) &= (\sum_{i=0}^{\infty} c_i^* z^i)\Delta M - \Delta N \\ &= z^{w+1} \sum_{k=0}^{\infty} G_k[\beta_1, \ldots, \beta_m] z^k \\ &= z^{w+1} \sum_{k=0}^{\infty} E_k z^k. \end{aligned}$$

This concludes the proof of the lemma. □

At this moment we should note that

$$G_0 = \begin{bmatrix} T_{1,u_1}^w(f_{11}) & T_{1,u_2}^w(f_{12}) & \cdots & T_{1,u_m}^w(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{1,u_1}^w(f_{p1}) & T_{1,u_2}^w(f_{p2}) & \cdots & T_{1,u_m}^w(f_{pm}) \end{bmatrix}. \tag{8.25}$$

The i-th row of G_0 is just ${}^{L}B_I^i(\tilde{V}, \tilde{U}, \tilde{W}, f)$ (see (3.15)). This fact will be used in the proof of Theorem 8.4.

Let $R_i = f^* M_i - N_i \in {}^{R}R^*(V, U, W)^f$ for $i = 1, 2$. Then by Lemma 8.1, there exist matrices $E_k^{(1)}$ and $E_k^{(2)}$ such that

$$R_i(z) = z^{w+1} \sum_{k=0}^{\infty} E_k^{(i)} z^k, \ i = 1, 2.$$

If $E_0^{(1)} \neq E_0^{(2)}$, then by

$$\begin{aligned} N_1 M_1^{-1} - N_2 M_2^{-1} &= (f^* - N_2 M_2^{-1}) - (f^* - N_1 M_1^{-1}) \\ &= R_2 M_2^{-1} - R_1 M_1^{-1} \\ &= z^{w+1} \left[(\sum_{k=0}^{\infty} E_k^{(2)} z^k) M_2^{-1} - (\sum_{k=0}^{\infty} E_k^{(1)} z^k) M_1^{-1} \right], \end{aligned}$$

we have

$$N_1 M_1^{-1} \neq N_2 M_2^{-1}.$$

Therefore by characterizing the set ${}^{R}R^*(V, U, W)^f$ we may achieve a result about the the uniqueness of type I MPAs.

Lemma 8.2 *Let us denote the null space of a matrix $A \in \mathbf{C}^{p \times s}$ by $\mathbf{N}(A)$, i.e.,*

$$\mathbf{N}(A) = \{x \in \mathbf{C}^s : \ Ax = 0\}.$$

Suppose we have a matrix $A \in \mathbf{C}^{p \times s}$ and a matrix $B \in \mathbf{C}^{q \times s}$. We shall then denote the range of B when restricted to the nullspace of A by

$$\text{range}(B, A) = \{y \in \mathbf{C}^q : \ y = Bx, \ x \in \mathbf{N}(A)\}.$$

Then the following result holds

$$\dim \operatorname{range}(B, A) = \operatorname{rank} C - \operatorname{rank} A,$$

where $\dim \operatorname{range}(B, A)$ *denotes the dimension of the linear space* $\operatorname{range}(B, A)$*, and*

$$C = \begin{bmatrix} A \\ B \end{bmatrix}.$$

The proof of this lemma can be found in [16]. □

Now we are ready to prove the following theorem :

Theorem 8.3 *Let* $f^*(z) = \sum_{k=0}^{\infty} c_k^* z^k$*. Then* ${}^R R^*(V, U, W)^f$ *is a linear space and*

$$\dim\ {}^R R^*(V, U, W)^f = m \sum_{k=0}^{\infty} (\operatorname{rank} H_{k+1} - \operatorname{rank} H_k),$$

where

$$H_k = \begin{bmatrix} T^{v_1}_{e_1+k,u_1}(f^*_{11}) & T^{v_1}_{e_1+k,u_2}(f^*_{12}) & \cdots & T^{v_1}_{e_1+k,u_m}(f^*_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_p}_{e_p+k,u_1}(f^*_{p1}) & T^{v_p}_{e_p+k,u_2}(f^*_{p2}) & \cdots & T^{v_p}_{e_p+k,u_m}(f^*_{pm}) \end{bmatrix},$$

and $E = W - V \in \mathbf{Z}_+^{p \times 1}$.

Proof. (i) First we note that

$$H_0 = {}^R H_I(V, U, W).$$

From Lemma 8.1 we have

$${}^R R^*(V, U, W)^f = \{R(z) = z^{w+1} \sum_{k=0}^{\infty} G_k B z^k :\ B \in \mathbf{C}^{|u| \times m}, H_0 B = 0\}, \tag{8.26}$$

where $|u| = \sum_{i=1}^{m} u_i$. Therefore ${}^R R^*(V, U, W)^f$ is a linear space.

(ii). Consider the space

$$S_k = \{E = G_k B :\ B \in \mathbf{C}^{|u| \times m}, H_k B = 0\},$$

and suppose

$$E_{k,1}, E_{k,2}, \ldots, E_{k,i_k}$$

is the basis for this space S_k (i_k may be zero) and associate with this basis the matrices

$$B_{k,1}, B_{k,2}, \ldots, B_{k,i_k} \in \{B \in \mathbf{C}^{|u| \times m} :\ H_k B = 0\},$$

such that

$$E_{k,j} = G_k B_{k,j}, \quad j = 1, 2, \ldots, i_k.$$

Then for any

$$R(z) = z^{w+1} \sum_{k=0}^{\infty} E_k z^k = z^{w+1} \sum_{k=0}^{\infty} G_k B z^k \ \in\ {}^R R^*(V, U, W)^f$$

there exist constants α_{ij} such that

$$E_k = G_k \sum_{l=0}^{\infty} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}. \tag{8.27}$$

In fact, we can prove by induction that there exist α_{ij} such that

$$E_k = G_k \sum_{l=0}^{k} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}. \tag{8.28}$$

For $k = 0$, since E_{0j} is a base of S_0, there exist α_{0j} such that

$$E_0 = \sum_{j=1}^{i_0} \alpha_{0j} E_{0j} = G_0 \sum_{j=1}^{i_0} \alpha_{0j} B_{0j}.$$

Suppose α_{ij} have been determined for $i = 0, 1, \ldots, k-1$. Note here that the rows of H_k are the rows of H_{k-1} extended with the rows of G_{k-1} for any $k = 1, 2, \ldots$. This implies

$$\begin{aligned} E_k &= E_k - G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} + G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} \\ &= G_k (B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) + G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} \end{aligned}$$

and

$$G_t (B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = E_t - E_t = 0, \quad t = 0, 1, \ldots, k-1,$$

so that $H_k (B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = 0$. i.e.,

$$G_k (B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) \in S_k.$$

Thus there exist α_{kj} such that

$$G_k (B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = \sum_{j=1}^{i_k} \alpha_{kj} E_{kj} = G_k \sum_{j=1}^{i_k} \alpha_{kj} B_{kj}.$$

Hence (8.28) is proved. Since G_k is formed by some rows of $H_{k'}$ if $k' > k$,

$$G_k B = 0, \quad \text{if} \quad H_{k'} B = 0, \quad \text{and } k' > k.$$

This fact and (8.28) imply (8.27).

On the other hand, it is easy to show that

$$R_{lj}(z) := z^{W+1} \sum_{k=0}^{\infty} G_k B_{lj} z^k \in {}^{R}R^*(V, U, W)^f, \qquad l = 0, 1, \ldots, \infty; j = 1, 2, \ldots, i_l$$

are linearly independent. Then from (8.27) we have that $\{R_{lj}(z)\}$ is a basis of ${}^{R}R^*(V, U, W)^f$. Therefore

$$\dim \ {}^{R}R^*(V, U, W)^f = \sum_{k=0}^{\infty} i_k = \sum_{k=0}^{\infty} \dim S_k.$$

So we need only consider the dimension of the space S_k.

Now we note that

$$X^k = \{[x_1, \ldots, x_k] \in \mathbf{C}^{d \times k} : \ x_i \in X \subset \mathbf{C}^d \quad i = 1, \ldots, k\}$$

is a $(k\cdot\dim X)$-dimensional linear subspace of $\mathbf{C}^{d\times k}$ provided X is a linear subspace of $\mathbf{C}^d$. Therefore, it follows from

$$\begin{aligned} S_k &= \{y : y = G_k x,\ x \in \mathbf{N}(H_k)\}^m \\ &= \{\text{range}(G_k, H_k)\}^m \end{aligned}$$

and Lemma 8.2, that

$$\dim S_k = m[\text{rank}\, H_{k+1} - \text{rank}\, H_k].$$

This concludes the proof of theorem 8.3. □

Now we are all set up to establish the uniqueness theorem for type I MPAs.

Theorem 8.4 *Let $V \in \mathbf{Z}_+^{p\times 1}$ $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+$. Suppose ${}^R[V,U,W]_I^f \neq \emptyset$. Then the following statements are equivalent.*

(i) ${}^R[V,U,W]_I^f$ *is unique.*

(ii) ${}^L[W-U^T, W-V^T, W]_I^f \neq \emptyset$.

(iii) There exists a $(\tilde N, \tilde M) \in {}^L R(W-U^T, W-V^T, W)^f$ *such that* $\det \tilde M \neq 0$.

(iv) For any $(N,M) \in {}^R R(V,U,W)^f$,

$$ {}^R[V,U,W]_I^f M - N = \{0\}. \tag{8.29}$$

Proof. a). (i) $\Rightarrow$ (ii). If ${}^R[V,U,W]_I^f$ is unique, then by Theorem 8.3, we have $\dim {}^R R^*(V,U,W)^f = 0$ and therefore

$$\text{rank}\, H_1 - \text{rank}\, H_0 = 0.$$

Since $H_0 = {}^R H_I(V,U,W)$ and H_1 is an extension of H_0 obtained by adding the rows of G_0 (see (8.25)), the above relation implies that the condition (3.17) holds for $i = 1, 2, \ldots, p$. From this we conclude that the left MPA exists.

b). (ii) $\Rightarrow$ (iii). Suppose $\tilde M^{-1}\tilde N \in {}^L[W-U^T, W-V^T, W]_I^f$. Then $(\tilde N, \tilde M) \in {}^L R(W-U^T, W-V^T, W)^f$ and $\det \tilde M \neq 0$.

c). (iii) $\Rightarrow$ (iv). Suppose $(\tilde N, \tilde M) \in {}^L R(W-U^T, W-V^T, W)^f$ and $\det \tilde M \neq 0$. Then for any $(N,M) \in {}^R R(V,U,W)^f$, we have by the duality that for any $N^*M^{*-1} \in {}^R[V,U,W]_I^f$,

$$\begin{aligned} \tilde M(N^*M^{*-1}M - N) &= \tilde N M^* M^{*-1} M - \tilde M N \\ &= \tilde N M - \tilde N M \\ &= 0. \end{aligned}$$

Hence (8.29) holds.

d). (iv) $\Rightarrow$ (i). From (8.29) one has ${}^R R^*(V,U,W)^f = \{0\}$. It then immediately follows that ${}^R[V,U,W]_I^f$ is unique. □

We shall now consider the uniqueness problem for the type II MPAs. Let $V, W \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+$ and solvability equality (2.5) hold. It follows from Theorem 5.2 that

$$ {}^R[V,U,W]_{II}^f = f^{(W)}(z) - \text{diag}[z^{w_1}, \ldots, z^{w_p}]\ {}^R[U-1, U, U+W-V-1]_I^g(z^{-1}), \tag{8.30}$$

where $g(z) = (g_{ij})_{i,j=1}^{p,m}$ and $g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1})$.

Relation (8.30) implies that there is a one to one map between ${}^R[V,U,W]^f_{II}$ and ${}^R[U-1,U,U+W-V-1]^g_I$. Therefore the uniqueness problem of ${}^R[V,U,W]^f_{II}$ can be solved by considering the uniqueness problem of ${}^R[U-1,U,U+W-V-1]^g_I$. In order to use Theorem 8.4 for the problem ${}^R(U-1,U,U+W-V-1)^g_I$, we should have

$$U+W-V-1 \in \mathbf{Z}_+.$$

i.e., $w_i - v_i = c$ (a constant) for $i = 1,2,\ldots,p$. This means by (2.5) that p should divide mu. From Theorem 8.4 and Theorem 5.2, we have the following

Corollary 8.5 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $W-V, U \in \mathbf{Z}_+$. Let*

$$g(z) = (g_{ij})_{i,j=1}^{p,m},$$

with

$$g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1}).$$

Suppose ${}^R[V,U,W]^f_{II} \neq \emptyset$. Then the following statements are equivalent.

(i) ${}^R[V,U,W]^f_{II}$ *is unique.*

(ii) ${}^R[U-1,U,U+W-V-1]^g_I$ *is unique.*

(iii) ${}^L[W-V-1,W-V,U+W-V-1]^g_I \neq \emptyset$.

(iv) There exists a $(\tilde{N},\tilde{M}) \in {}^LR(W-V-1,W-V,U+W-V-1)^g$ *such that* $\det \tilde{M} \neq 0$.

(v) For any $(N,M) \in {}^RR(U-1,U,U+W-V-1)^g$,

$${}^R[U-1,U,U+W-V-1]^g_I M - N = \{0\}.$$

Based on the duality principle (see Corollary 7.4), we have

Corollary 8.6 *Let $V \in \mathbf{Z}_+^{p\times 1}$, $W \in \mathbf{Z}_+$, and $U \in \mathbf{Z}_+^{m\times 1}$ satisfy condition (2.7). If*

$${}^R[V,U,W]^f_I \neq \emptyset \quad \textit{and} \quad {}^R[V,U,W]^f_{II} \neq \emptyset,$$

then the uniqueness of ${}^R[V,U,W]^f_I$ implies the uniqueness of ${}^R[V,U,W]^f_{II}$

Proof. If ${}^R[V,U,W]^f_I$ exists uniquely, then by Theorem 8.4 we know that ${}^L[W-U^T,W-V^T,W]^f_I$ exists. Since ${}^L(W-U^T,W-V^T,W)^f_I$ is just the mixed dual problem of ${}^R(V,U,W)^f_{II}$, ${}^R[V,U,W]^f_{II}$ is unique. □

The condition in this Corollary is slightly more general than the conditions in Corollary 8.5, but here we only get a sufficient condition for the uniqueness. What conclusions can be drawn from the uniqueness of ${}^R[V,U,W]^f_{II}$ is a problem to be solved. However, if ${}^R(V,U,W)^f_{II}$ has a dual problem ${}^L(\tilde{V},\tilde{U},\tilde{W})^f_{II}$, then we have the following results.

Theorem 8.7 *Let $V \in \mathbf{Z}_+^{p\times 1}$ $W \in \mathbf{Z}_+$, $U \in \mathbf{Z}_+^{m\times 1}$ and let conditions (3.23), (3.24) be satisfied. If ${}^R[V,U,W]^f_{II} \neq \emptyset$, then the uniqueness of ${}^R[V,U,W]^f_{II}$ is equivalent to the existence of ${}^L[(W-U)^T,(W-V)^T,W]^f_{II}$.*

Proof. From Theorem 7.3, we know that the problems ${}^R(V,U,W)^f_{II}$ and ${}^L((W-U)^T,(W-V)^T,W)^f_{II}$ are dual. Hence by the definition of duality, the uniqueness of ${}^R[V,U,W]^f_{II}$ follows from the existence of ${}^L[(W-U)^T,(W-V)^T,W]^f_{II}$. Now we prove the inverse side. Consider the matrix

$$G = \begin{bmatrix} T^{v_1}_{1,u_1}(f_{11}) & T^{v_1}_{1,u_2}(f_{12}) & \cdots & T^{v_1}_{1,u_m}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_p}_{1,u_1}(f_{p1}) & T^{v_p}_{1,u_2}(f_{p2}) & \cdots & T^{v_p}_{1,u_m}(f_{pm}) \end{bmatrix}.$$

First we show that

$$\text{range}(G,\ {}^RH_{II}(V,U,W)) = \{0\}. \tag{8.31}$$

Let $N^*M^{*-1} \in {}^R(V,U,W)^f_{II}$. Then for (N,M) satisfying $NM^{-1} \in {}^R(V,U,W)^f_{II}$, we have

$$R(z) := N^*M^{*-1}M - N = N^*M^{*-1}\Delta M - \Delta N, \tag{8.32}$$

where

$$\Delta M = M - M^*, \quad \Delta N = N - N^*.$$

Since M^* (and M) is $C1$ canonical, it can be written as

$$M^*(z) = \text{diag}[z^{u_1},\ldots,z^{u_m}]\tilde{M}^*(z^{-1}), \quad \tilde{M}^*(0) = I.$$

Hence

$$\begin{aligned} M^*(z)^{-1} &= \tilde{M}^*(z^{-1})^{-1}\,\text{diag}[z^{-u_1},\ldots,z^{-u_m}] \\ &= \sum_{i=0}^{\infty} d_i z^{-i}\,\text{diag}[z^{-u_1},\ldots,z^{-u_m}], \end{aligned}$$

and

$$R(z) = N^*\sum_{i=0}^{\infty} d_i z^{-i}[\tilde{M}(z^{-1}) - \tilde{M}^*(z^{-1})] - \Delta N.$$

It follows from the uniqueness that

$$\Delta N = N^*\sum_{i=0}^{\infty} d_i z^{-i}[\tilde{M}(z^{-1}) - \tilde{M}^*(z^{-1})].$$

Comparing the leading coefficient (the highest degree is V) of both sides, we find that

$$\begin{bmatrix} T^{v_j}_{1,u_1}(f_{11}) & \cdots & T^{v_j}_{1,u_m}(f_{1m}) \end{bmatrix} X = 0, \quad \text{for} \quad {}^RH(V,U,W)X = 0, \quad j = 1,\ldots,m,$$

and thus (8.31) holds.

From (8.31) and Lemma 8.2, it follows that

$$\text{rank}\ {}^RH_{II}(V,U,W) = \text{rank}\begin{bmatrix} G \\ {}^RH_{II}(V,U,W) \end{bmatrix} = \text{rank}\ {}^RH_{II}(V-1,U,W).$$

Therefore, by Corollary 3.5, ${}^L[(W-U)^T,(W-V)^T,W]^f_{II}$ exists. □

We should mention here that although the three results concerning the uniqueness of the second type MPA seem to be similar in forms, but each one treats a different case. So they do not overlap.

9 Direct Computation

9.1 Near Toeplitz Structure

In Section 3, we already noted that the numerator is quite easy to determine when the denominator is known. Hence the problem of computing a MPA is reduced to the problem of computing the denominator. By the term "Direct Computation" we mean that we solve the system of linear equations (see (3.10))

$$ {}^{R}H^{j}_{I\ or\ II}(V,U,W)X = -\ {}^{R}B^{j}_{I\ or\ II}(V,U,W), \tag{9.33} $$

directly. Since the computation of the left MPA is essentially the same as the computation of the right MPA, we shall only consider the computation of the right MPA. If $V, W \in \mathbf{Z}_{+}^{p\times 1}$, $U \in \mathbf{Z}_{+}^{m\times 1}$, then the matrices ${}^{R}H^{j}_{I}(V,U,W)$ $(:= {}^{R}H_{I}(V,U,W))$ do not depend on j. This is an important feature in the computation since we then only need to (implicitly) invert one matrix instead of m. For the second type MPA, this feature will be kept also, if U satisfies condition (2.7). This is of course a special case, but it is an important one, because when f is normal, the MmPA problem is equivalent to such a problem.

From the definition of ${}^{R}H^{j}_{I\ or\ II}(V,U,W)$, we know a nice property of this matrix. That is, it has near–Toeplitz structure or low displacement rank (see [9]). For an $n\times n$ matrix A, its displacement rank is defined by

$$ \alpha(A) = \mathrm{rank}(A - U_n A U_n^T), \tag{9.34} $$

where

$$ U_n = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}. $$

In fact we have the following results

Lemma 9.1 *(i) The displacement rank of the matrix ${}^{R}H^{j}_{I}(V,U,W)$ is at most*

$$ \sum_{i=1}^{m}\delta(u_{ij}) + \sum_{i=1}^{p}\delta(w_{ij}-v_{ij}); $$

(ii) The displacement rank of the matrix ${}^{R}H^{j}_{II}(V,U,W)$ is at most

$$ \sum_{i=1}^{m}\delta(s_{ij}) + \sum_{i=1}^{p}\delta(w_{ij}-v_{ij}), $$

where

$$ \delta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0. \end{cases} $$

Proof. From the definition of ${}^{R}H^{j}_{I}(V,U,W)$ it can be seen that if $u_{lj} > 0$ and $w_{kj} - v_{kj} > 0$, the block at the k-th row and the l-th column in the block matrix ${}^{R}H^{j}_{I}(V,U,W)$ is $T^{v_{kj}}_{w_{kj}-v_{kj},u_{lj}}(f_{kl})$ and this is a Toeplitz block. Otherwise $T^{v_{kj}}_{w_{kj}-v_{kj},u_{lj}}(f_{kl})$ is an empty matrix. Noting this fact, we can get conclusion (i) easily. (ii) is similar. □

From the lemma we have immediately that

$$ \alpha\left({}^{R}H^{j}_{I\ or\ II}(V,U,W)\right) \le p+m. $$

Since the coefficient matrices of equations (9.33) have low displacement rank, the algorithms developed in [9] can be used to invert the matrices ${}^{R}H^{j}_{I\ or\ II}(V,U,W)$. These algorithms need only

$$O\left(\alpha({}^{R}H^{j}_{I}(V,U,W))(\sum_{i=1}^{m}u_{ij})^2\right)$$

or

$$O\left(\alpha({}^{R}H^{j}_{II}(V,U,W))(\sum_{i=1}^{m}s_{ij})^2\right)$$

operations instead of $O\left((\sum_{i=1}^{m}u_{ij})^3\right)$ or $O\left((\sum_{i=1}^{m}s_{ij})^3\right)$.

These are somewhat more complex than the Hankel or Toeplitz algorithms that can be used is the scalar case (see [11], [12]) but they are still below the complexity of a general system. Possibly some speed-up can be implemented in the style of [2].

9.2 Two Terms Computation in the U Direction

Now we solve the problem ${}^{R}_{j}(V,U,W)^{f}_{I\ or\ II}$ recursively, for $V,W\in\mathbf{Z}_{+}^{p\times 1}$ and $U\in\mathbf{Z}_{+}^{m\times 1}$. In fact, we compute the inverses of the type I matrices ${}^{R}H^{j}_{I}(V,U,W)$ successively. The second type problem can be treated similarly but we shall not do it here. By two terms computation in the U direction we mean that we have already computed ${}^{R}H^{j}_{I}(V,U,W)^{-1}$ and then want to compute ${}^{R}H^{j}_{I}(V,U',W')^{-1}$ with

$$0\le U'-U\le 1,\quad 0\le W'-W\le 1. \tag{9.35}$$

At each step, the solvability equality should be satisfied. In this way we can get a sequence of MPAs.

We shall solve this problem for two different cases :

(A). (i) $U'=(u'_i)_{i=1}^{m}$, $\quad u'_i=u_i+\delta(u_i)$.
(ii) $0\le W'-W\le 1$.
(iii) solvability equality (2.4) holds.
(iv) $w'_i=w_i$, if $w_i=v_i$.

(B). (i) $0\le U'-U$ and $|U'-U|=1$.
(ii) $0\le W'-W$ and $|W'-W|=1$.

In case (A), several rows and columns are added to ${}^{R}H^{j}_{I}(V,U,W)$, while in case (B), only one row and one column are added.

We shall start with case (A). For the sake of simplicity of notation, we assume all $u_i>0$ and all $w_i>v_i$. Let $i_1<i_2<\cdots<i_m$ be such that $w'_{i_k}>w_{i_k}, k=1,\ldots,m$. Then ${}^{R}H^{j}_{I}(V,U',W')$ is an extension of ${}^{R}H^{j}_{I}(V,U,W)$ in the sense that if we delete the $\sum_{i=1}^{j}u'_i$–th columns for $j=1,\ldots,m$, and the $\sum_{j=1}^{i_k}(w'_j-v_j)$–th rows for $k=1,\ldots,m$ from ${}^{R}H^{j}_{I}(V,U',W')$, then we get ${}^{R}H^{j}_{I}(V,U,W)$. Define the following numbers

$$\alpha'_k=\sum_{j=1}^{i_k}(w'_j-v_j),\quad k=1,2,\ldots,m,$$

$$\beta'_k=\sum_{j=1}^{i_k-1}(w'_j-v_j)+1,\quad k=1,2,\ldots,m,$$

$$\alpha_k = \sum_{j=1}^{i_k} (w_j - v_j), \quad k = 1, 2, \ldots, m,$$

$$\beta_k = \sum_{j=1}^{i_k - 1} (w_j - v_j) + 1, \quad k = 1, 2, \ldots, m.$$

We first compute the β'_k–th column q'_k and α'_k–th column r'_k of ${}^R H_I^j(V, U', W')^{-1}$ for $k = 1, \ldots, m$. Let $\bar{q}_k$ and $\bar{r}_k$ $(k = 1, \ldots, m)$ be the β_k–th column and α_k–th column of ${}^R H_I^j(V, U, W)^{-1}$ respectively. Let q_k and r_k $(k = 1, \ldots, m)$ be the extensions of $\bar{q}_k$ and $\bar{r}_k$ obtained by inserting m zeros such that for each $j = 1, \ldots, m$, the $\sum_{i=1}^{j} u'_i$–th element of q_k is an inserted zero. For r_k, the inserted zeros are found at the positions $\sum_{i=1}^{j-1} u'_i + 1$ for $j = 1, \ldots, m$. Then we have

$$h_i q_k = \begin{cases} 1, & i = \beta'_k, \\ h_{\alpha'_j} q_k, & i = \alpha'_j, \quad j = 1, \ldots, m, \\ 0, & \text{otherwise}, \end{cases} \tag{9.36}$$

and

$$h_i r_k = \begin{cases} 1, & i = \alpha'_k, \\ h_{\beta'_j} r_k, & i = \beta'_j, \quad j = 1, \ldots, m, \\ 0, & \text{otherwise}, \end{cases} \tag{9.37}$$

where h_i is the i-th row of ${}^R H_I^j(V, U', W')$. We show that q'_k and r'_k can be written as

$$q'_k = \sum_{i=1}^{m} x_{ik} q_i + \sum_{i=1}^{m} y_{ik} r_i, \tag{9.38}$$

and

$$r'_k = \sum_{i=1}^{m} a_{ik} q_i + \sum_{i=1}^{m} b_{ik} r_i. \tag{9.39}$$

The matrices X, Y, A and B with the coefficients of these relations can now be found as follows : Because

$${}^R H_I^j(V, U', W') q'_j = e_{\beta'_j}, \quad j = 1, \ldots, m, \tag{9.40}$$

and

$${}^R H_I^j(V, U', W') r'_j = e_{\alpha'_j}, \quad j = 1, \ldots, m, \tag{9.41}$$

where $e_k \in \mathbf{C}^{|U'|}$ is the k-th unit vector, we have

$$HZ = I_{2m}, \tag{9.42}$$

where

$$H = \begin{bmatrix} 1 & 0 & \cdots & 0 & h_{\beta'_1} r_1 & h_{\beta'_1} r_2 & \cdots & h_{\beta'_1} r_m \\ 0 & 1 & \cdots & 0 & h_{\beta'_2} r_1 & h_{\beta'_2} r_2 & \cdots & h_{\beta'_2} r_m \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & h_{\beta'_m} r_1 & h_{\beta'_m} r_2 & \cdots & h_{\beta'_m} r_m \\ h_{\alpha'_1} q_1 & h_{\alpha'_1} q_2 & \cdots & h_{\alpha'_1} q_m & 1 & 0 & \cdots & 0 \\ h_{\alpha'_2} q_1 & h_{\alpha'_2} q_2 & \cdots & h_{\alpha'_2} q_m & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{\alpha'_m} q_1 & h_{\alpha'_m} q_2 & \cdots & h_{\alpha'_m} q_m & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{9.43}$$

$$= \begin{bmatrix} I_m & R \\ Q & I_m \end{bmatrix},$$

and

$$Z = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} & a_{11} & a_{12} & \cdots & a_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} & a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{m1} & y_{m2} & \cdots & y_{mm} & a_{m1} & a_{m2} & \cdots & a_{mm} \\ y_{11} & y_{12} & \cdots & y_{1m} & b_{11} & b_{12} & \cdots & b_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} & b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{m1} & y_{m2} & \cdots & y_{mm} & b_{m1} & b_{m2} & \cdots & b_{mm} \end{bmatrix} \tag{9.44}$$

$$= \begin{bmatrix} X & A \\ Y & B \end{bmatrix}.$$

This implies the following equalities

$$X + RY = I_m, \quad QX + Y = 0,$$

$$A + RB = 0, \qquad QA + B = I_m.$$

Hence the coefficient matrices are found as follows

$$Y = (R - Q^{-1})^{-1},$$

$$X = -Q^{-1}Y,$$

$$B = -YQ^{-1},$$

$$A = Q^{-1}(I - B).$$

Now r'_k and q'_k can be computed by (9.38) and (9.39).

Secondly we compute the other columns of ${}^{R}H_I^j(V, U', W')^{-1}$. Suppose we want to compute the the t–th column s'_t with $t \neq \alpha'_k, \beta'_k$, $k = 1, \ldots, m$. Then there are three possibilities :

(i) $t < \beta'_1$, or

(ii) there exists an i such that $t = \alpha'_m + i$, or

(iii) there exist a k and a j such that $t = \beta'_k + j$ and $j < \alpha'_k - \beta'_k$.

Now, define

$$t_1 = \begin{cases} t, & \text{for case (i)}, \\ \alpha_m + i, & \text{for case (ii)}, \\ \beta_k + j, & \text{for case (iii)}, \end{cases}$$

and set

$$s'_t = s_{t_1} + x,$$

for some vector x to be determined, where s_{t_1} is the t_1-th column of ${}^{R}H_I^j(V, U, W)^{-1}$ but extended with zeros in the same manner as $\bar{q}_k$ was extended to q_k. Then by

$${}^{R}H_I^j(V, U', W')s'_t = e_t$$

it is easily seen that

$$x = -\sum_{i=1}^{m}(h_{\alpha'_i} s_{t_1})r'_i.$$

Hence the new column s'_t is given by

$$s'_t = s_{t_1} - \sum_{i=1}^{m} (h_{\alpha'_i} s_{t_1}) r'_i. \tag{9.45}$$

This formula shows that the remaining columns of ${}^{R}H_I^j(V, U', W')^{-1}$ can be computed from the columns r'_k for $k = 1, \ldots, m$. Therefore if q_k and r_k are known for $k = 1, \ldots, m$, it is sufficient to compute q'_k and r'_k for $k = 1, \ldots, m$ and then go to the next step, provided we do not need all the columns of the inverse matrix. This would reduce the cost of operations.

For case (B), there may be no Toeplitz property that can be utilized, since the newly added row and column may not have any relation with the previous matrix ${}^{R}H_I^j(V, U, W)$. So we treat this case separately.

By row exchanges and column exchanges, the added row and column can be moved to the right and bottom of the matrix. So we may assume the added row and column are at this position. We now use a classical bordering technique to invert the matrix. Set

$${}^{R}H_I^j(V, U', W') = \begin{bmatrix} {}^{R}H_I^j(V, U, W) & a \\ b & c \end{bmatrix},$$

and let its inverse be given by

$${}^{R}H_I^j(V, U', W')^{-1} = \begin{bmatrix} X & y \\ z & t \end{bmatrix}.$$

Then the invesion formula gives

$$t = \frac{1}{c - b\ {}^{R}H_I^j(V, U, W)^{-1} a},$$

$$y = {}^{R}H_I^j(V, U, W)^{-1} a t,$$

$$z = -b\ {}^{R}H_I^j(V, U, W)^{-1} t,$$

$$X = {}^{R}H_I^j(V, U, W)^{-1}(I - az).$$

This concludes the computation of the MPAs for case (B).

To clarify this general treatment, we now give an example to illustrate the usage of the two cases to compute a series of MPAs. Let $p = 4, m = 2$. In the last step we want to get ${}^{R}_{j}[V, U, W]$ with

$$V = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 2 \end{bmatrix}, \quad U = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad W = \begin{bmatrix} 4 \\ 6 \\ 6 \\ 5 \end{bmatrix}.$$

It is easy to see that solvability equality (2.4) holds. Starting from $U_0 = 0$ and $W_0 = V$, we go on as in the following scheme:

$$\text{Step 0.} \qquad U_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad W_0 = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 2 \end{bmatrix}. \qquad \text{Step 1.} \quad U_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad W_1 = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 2 \end{bmatrix}.$$

Step 2. $U_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $W_2 = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 2 \end{bmatrix}$. Step 3. $U_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $W_3 = \begin{bmatrix} 2 \\ 4 \\ 6 \\ 2 \end{bmatrix}$.

Step 4. $U_4 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $W_4 = \begin{bmatrix} 2 \\ 4 \\ 6 \\ 3 \end{bmatrix}$. Step 5. $U_5 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $W_5 = \begin{bmatrix} 3 \\ 5 \\ 6 \\ 3 \end{bmatrix}$.

Step 6. $U_6 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, $W_6 = \begin{bmatrix} 3 \\ 6 \\ 6 \\ 4 \end{bmatrix}$. Step 7. $U_7 = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, $W_7 = \begin{bmatrix} 4 \\ 6 \\ 6 \\ 5 \end{bmatrix}$.

The first four steps use case (B), the final three steps use case (A). It should be noted that the size of the matrix is small at the beginning. In this example, the matrix size is $i \times i$ at step $i-1$ for $i \leq 5$.

9.3 Two Terms Computation in the V Direction

Suppose we know ${}^R H_I^j(V,U,W)^{-1}$ and want to compute ${}^R H_I^j(V',U,W')^{-1}$ with

$$V' \geq V,\ |V'-V| = 1, \quad \text{and} \quad W' \geq W,\ |W'-W| = 1.$$

This problem is quite easy to handle. Because the updated matrix is just the previous one with one row added and one row deleted. In general, the following steps are easy to understand. Given a matrix A, suppose B is obtained from A by deleting the s–th row and adding one row at the t–th line. Assume $s \leq t$ without loss of generality. Then

$$B^{-1} = [h_1 \cdots h_{s-1}\ h_{s+1} \cdots h_{t-1}\ (b_t h_s)^{-1} h_s\ h_{t+1} \cdots h_n],$$

where h_i is the i–th column of A^{-1}, b_t is the t–th row of B.

10 Recursive Computation

In this section we assume once and for all that $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$, so that we don't have to mention this condition every time. On the other hand, we mainly consider the computation of type I right MPAs. The computation of the second type will be shown to be equivalent to the computation of the first type under some conditions. So the left superscript R and right subscript I or II in the notations will be deleted if there is no confusion. Consider the set

$$[V,U,W] = \{(N,M) : NM^{-1} \in {}^R[V,U,W]^f\}.$$

The problem we want to solve in this part is to compute $[V,U,W]$ from two "previous" ones. We shall need some normality condition for f which is different from the one we have defined earlier. We call the new concept I-normality. It is defined as follows :

Definition 10.1 (I-normal) *If for any V, U and W which satisfy condition (2.4), the matrix $H(V,U,W)$ is nonsingular. Then we say that f is I-normal.*

For $(N,M) \in [V,U,W]$, we introduce the following notations for the elements of the numerator, denominator and residual

$$N_{ij}(z) = \sum_{k=0}^{v_i} n_{ij}^{(k)} z^k, \quad M_{ij}(z) = \sum_{k=0}^{u_i} m_{ij}^{(k)} z^k,$$

and

$$(fM - N)_{ij}(z) = \sum_{k=0}^{\infty} e_{ij}^{(k)} z^k.$$

We introduce the following matrices:

$$N(V,U,W,V') = \begin{bmatrix} N_1^{(v_1'+1)}(V,U,W) \\ \vdots \\ N_1^{(v_1)}(V,U,W) \\ \vdots \\ N_p^{(v_p'+1)}(V,U,W) \\ \vdots \\ N_p^{(v_p)}(V,U,W) \end{bmatrix}, \quad M(V,U,W) = \begin{bmatrix} M_1^{(0)}(V,U,W) \\ \vdots \\ M_1^{(u_1)}(V,U,W) \\ \vdots \\ M_m^{(0)}(V,U,W) \\ \vdots \\ M_m^{(u_m)}(V,U,W) \end{bmatrix},$$

and

$$E(V,U,W,W') = \begin{bmatrix} E_1^{(w_1+1)}(V,U,W) \\ \vdots \\ E_1^{(w_1')}(V,U,W) \\ \vdots \\ E_p^{(w_p+1)}(V,U,W) \\ \vdots \\ E_p^{(w_p')}(V,U,W) \end{bmatrix},$$

where

$$N_i^{(k)}(V,U,W) = [n_{i1}^{(k)} \ n_{i2}^{(k)} \ \cdots \ n_{im}^{(k)}],$$

$$M_i^{(k)}(V,U,W) = [m_{i1}^{(k)} \ m_{i2}^{(k)} \ \cdots \ m_{im}^{(k)}],$$

and

$$E_i^{(k)}(V,U,W) = [e_{i1}^{(k)} \ e_{i2}^{(k)} \ \cdots \ e_{im}^{(k)}].$$

The degrees $V'+1$ and orders W' satisfy

$$V'+1 \in \mathbf{Z}_+^{p\times 1} \quad \text{and} \quad V' \le V, \ \ W' \ge W.$$

It can be seen that $N(V,U,W,-1)$ is the coefficient matrix of N, $M(V,U,W)$ is the coefficient matrix of M, and $E(V,U,W,\infty)$ is the coefficient matrix of $R = fM - N$. Hence from the definition of MPA, we have

$$H(V'+1,U+1,V+1)M(V,U,W) = N(V,U,W,V'), \tag{10.46}$$

$$H(V+1,U+1,W+1)M(V,U,W) = 0, \tag{10.47}$$

$$H(W+1,U+1,W'+1)M(V,U,W) = E(V,U,W,W'). \tag{10.48}$$

Lemma 10.1 *Let f be I–normal. Then for any V,U,W satisfying (2.4), we have*

(i) The matrix

$$\begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix} \text{ is nonsingular,} \tag{10.49}$$

where

$$-1 \le V' \le V, \quad W' \ge W \text{ and } |V-V'| + |W'-W| = m. \tag{10.50}$$

(ii) The leading coefficient matrix of the denominator M

$$M^{hc}(V,U,W) = \begin{bmatrix} M_1^{(u_1)}(V,U,W) \\ \vdots \\ M_m^{(u_m)}(V,U,W) \end{bmatrix}$$

is nonsingular.

(iii) The matrix

$$\begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix} - \begin{bmatrix} N(V,U+1,W'',V') \\ 0 \end{bmatrix} \text{ is nonsingular,} \tag{10.51}$$

where V', W' satisfy (10.50), $W'' \ge W'$ and $(V,\ U+1,\ W'')$ satisfy (2.4).

Proof. (i) For $N \in \mathbf{Z}_+^{p\times 1}$, define

$$\theta_N^{(j)} = \begin{bmatrix} \theta_{n_j,n_1} & \theta_{n_j,n_2} & \cdots & \theta_{n_j,n_p} \end{bmatrix},$$

and

$$I_N^{(j)} = \begin{bmatrix} \theta_{n_j,n_1} & \cdots & \theta_{n_j,n_{j-1}} & I_{n_j} & \theta_{n_j,n_{j+1}} & \cdots & \theta_{n_j,n_p} \end{bmatrix},$$

where $\theta_{q,k} \in \mathbf{C}^{q\times k}$ is a zero matrix. For $N_i \in \mathbf{Z}_+^{p\times 1}$, $i = 0, 1, \ldots, k$, and $N_0 \le N_1 \le \cdots \le N_k$, let

$$L^{(j)}(N_0, N_1, \ldots, N_k) = \begin{bmatrix} I^{(j)}_{N_1-N_0} & \theta^{(j)}_{N_2-N_1} & \cdots & \theta^{(j)}_{N_k-N_{k-1}} \\ \theta^{(j)}_{N_1-N_0} & I^{(j)}_{N_2-N_1} & \cdots & \theta^{(j)}_{N_k-N_{k-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \theta^{(j)}_{N_1-N_0} & \theta^{(j)}_{N_2-N_1} & \cdots & I^{(j)}_{N_k-N_{k-1}} \end{bmatrix},$$

and

$$L(N_0, N_1, \ldots, N_k) = \begin{bmatrix} L^{(1)}(N_0, N_1, \ldots, N_k) \\ L^{(2)}(N_0, N_1, \ldots, N_k) \\ \vdots \\ L^{(p)}(N_0, N_1, \ldots, N_k) \end{bmatrix}.$$

Then

$$H(V'+1, U+1, W'+1) = L(V', V, W, W') \begin{bmatrix} H(V'+1, U+1, V+1) \\ H(V+1, U+1, W+1) \\ H(W+1, U+1, W'+1) \end{bmatrix}. \tag{10.52}$$

Hence

$$\operatorname{rank} H(V'+1, U+1, W'+1) = \operatorname{rank} \{H(V'+1, U+1, W'+1)\ [I(U)\ M(V,U,W)]\}$$

$$= \operatorname{rank} \begin{bmatrix} H(V', U, V) & H(V'+1, U+1, V+1) M(V, U, W) \\ H(V, U, W) & 0 \\ H(W, U, W') & H(W+1, U+1, W'+1) M(V, U, W) \end{bmatrix}$$

$$= \operatorname{rank} H(V, U, W) + \operatorname{rank} \begin{bmatrix} N(V, U, W, V') \\ E(V, U, W, W') \end{bmatrix}, \tag{10.53}$$

where

$$I(U) = \begin{bmatrix} \theta_U \\ I_U^{(1)} \\ \theta_U \\ I_U^{(2)} \\ \vdots \\ I_U^{(m)} \end{bmatrix}$$

and $\theta_U \in \mathbf{C}^{1 \times |U|}$ is a zero vector. (10.49) follows from (10.53) and the I-normality of f.

(ii) Let

$$\bar{M}(V, U, W) = \begin{bmatrix} M_1^{(0)}(V, U, W) \\ \vdots \\ M_1^{(u_1 - 1)}(V, U, W) \\ \vdots \\ M_m^{(0)}(V, U, W) \\ \vdots \\ M_m^{(u_m - 1)}(V, U, W) \end{bmatrix}.$$

Then

$$\operatorname{rank} H(V+1, U, W+1) = \operatorname{rank} \{ H(V+1, U, W+1) \, [I(U-1) \;\; \bar{M}(V, U, W)] \}$$

$$= \operatorname{rank} [H(V, U-1, W) \quad H(V+1, U, W+1) \bar{M}(V, U, W)]$$

$$= \operatorname{rank} \left[H(V, U-1, W) \quad - \left[C^1(V, U, W) \cdots C^m(V, U, W) \right] M^{hc}(V, U, W) \right]$$

$$= \operatorname{rank} \left\{ \left[H(V, U-1, W) \; C^1(V, U, W) \cdots C^m(V, U, W) \right] \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V, U, W) \end{bmatrix} \right\}$$

$$= \operatorname{rank} \left\{ H(V, U, W) \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V, U, W) \end{bmatrix} \right\}$$

$$= \operatorname{rank} \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V, U, W) \end{bmatrix},$$

where

$$C^j(V, U, W) = \begin{bmatrix} T_{w_1 - v_1, 1}^{v_1 - u_j + 1}(f_{1j}) \\ T_{w_2 - v_2, 1}^{v_2 - u_j + 1}(f_{2j}) \\ \vdots \\ T_{w_p - v_p, 1}^{v_p - u_j + 1}(f_{pj}) \end{bmatrix}.$$

Hence conclusion (ii) holds.

(iii) Let

$$M_1(V,U,W) = \begin{bmatrix} M_1^{(1)}(V,U,W) - M_1^{(1)}(V,U+1,W'') \\ \vdots \\ M_1^{(u_1)}(V,U,W) - M_1^{(u_1)}(V,U+1,W'') \\ M_1^{(u_1+1)}(V,U+1,W'') \\ \vdots \\ M_m^{(1)}(V,U,W) - M_m^{(1)}(V,U+1,W'') \\ \vdots \\ M_m^{(u_m)}(V,U,W) - M_m^{(u_m)}(V,U+1,W'') \\ M_m^{(u_m+1)}(V,U+1,W'') \end{bmatrix}.$$

Noting that $M(0) = I$ and that the leading coefficient matrix $M^{hc}(V,U+1,W'')$ is nonsingular, we have by (10.52) that

$$\begin{aligned} &\operatorname{rank} H(V',U+1,W') \\ &\quad = \operatorname{rank}\{H(V',U+1,W')\,[M_1(V,U,W) \;\; J(U)]\} \\ &\quad = \operatorname{rank}\begin{bmatrix} N(V,U,W,V') - N(V,U+1,W'',V') & H(V',U,V)) \\ 0 & H(V,U,W) \\ E(V,U,W,W') & H(W,U,W') \end{bmatrix} \\ &\quad = \operatorname{rank}\left\{\begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix} - \begin{bmatrix} N(V,U+1,W'',V') \\ 0 \end{bmatrix}\right\} + \operatorname{rank} H(V,U,W), \end{aligned}$$

where

$$J(U) = \begin{bmatrix} I_U^{(1)} \\ \theta_U \\ I_U^{(2)} \\ \theta_U \\ \vdots \\ I_U^{(m)} \\ \theta_U \end{bmatrix}.$$

Thus conclusion (iii) holds. □

As a consequence of this we have

Corollary 10.2 *Suppose f is I-normal. Then*

(i) $E(V,U,W,W')$ is nonsingular, when $W' \geq W$ and $|W' - W| = m$.

(ii) $N(V,U,W,V')$ is nonsingular, when $-1 \leq V' \leq V$ and $|V - V'| = m$.

These are two special cases of the conclusion (i) of Lemma 10.1.

Corollary 10.3 *Suppose f is I-normal. If U satisfies condition (2.7), then the two types of MPA problems are equivalent.*

Proof. From conclusion (ii) of Lemma 10.1, it follows that the leading coefficient matrix M^{hc} of $M(z)$ is nonsingular. Multiply (N, M) from the right with M^{hc} and the leading coefficient matrix of M will become the unit matrix. Then the denominator is made $C1$ canonical. On the other hand, $[V, U, W]_{II}$ exist uniquely, since $H_{II}(V, U, W) = H_I(V + 1, U, W + 1)$. Then the equivalency is proved. □

This is the reason why we don't have to consider the computation of MPAs of type II.

Theorem 10.4 *Suppose $m \geq p$ and f is I-normal. Then, for given V, U, W satisfying (2.4), and $W' \geq W$, $|W' - W| = m$,*

(i) there exists a $Q(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V + 1, U + 1, W' + 1] = z[V, U, W]Q(V, U, W, W') + [V + 1, U, W + 1], \tag{10.54}$$

and

$$Q(V, U, W, W') = -E(V, U, W, W')^{-1}E(V + 1, U, W + 1, W' + 1) \tag{10.55}$$

is nonsinguar.

(ii) If $W' \geq W + 1$, there exists a $D(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V + 1, U + 1, W' + 1] = z[V, U, W]D(V, U, W, W') + [V, U + 1, W'], \tag{10.56}$$

and

$$D(V, U, W, W') = -E(V, U, W, W')^{-1}H(W + 2, U + 2, W' + 2)M(V, U + 1, W') \tag{10.57}$$

has rank p.

(iii) If $W' \geq W + 1$, there exists a $T(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V - 1, U + 1, W' - 1] = [V, U, W]T(V, U, W, W') + [V, U + 1, W'](I - T(V, U, W, W')), \tag{10.58}$$

and

$$T(V, U, W, W') = -\begin{bmatrix} N(V, U, W, V - 1) - N(V, U + 1, W', V - 1) \\ E(V, U, W, W' - 1) \end{bmatrix}^{-1} \begin{bmatrix} N(V, U + 1, W', V - 1) \\ 0 \end{bmatrix} \tag{10.59}$$

has rank p.

(iv) If $W' \geq W + 1$, there exists a $S(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V - 1, U + 1, W' - 1] = [V, U, W] + z[V - 1, U, W - 1]S(V, U, W, W'), \tag{10.60}$$

and

$$S(V, U, W, W') = -\begin{bmatrix} N(V - 1, U, W - 1, V - 2) \\ E(V - 1, U, W - 1, W' - 2) \end{bmatrix}^{-1} \begin{bmatrix} N(V, U, W, V - 1) \\ E(V, U, W, W' - 1) \end{bmatrix} \tag{10.61}$$

is nonsingular.

Proof. (i) Let

$$(N_1, M_1) = [V, U, W], \quad (N_2, M_2) = [V + 1, U, W + 1], \quad (N_3, M_3) = [V + 1, U + 1, W' + 1].$$

Then by

$$fM_1 - N_1 = O(z^{W+1}), \quad fM_2 - N_2 = O(z^{W+2}) \quad \text{and} \quad fM_3 - N_3 = O(z^{W'+2}),$$

we have that $Q(V, U, W, W')$ satisfies

$$E(V, U, W, W')Q(V, U, W, W') + E(V + 1, U, W + 1, W' + 1) = 0.$$

From Lemma 10.1, it follows that (10.54) and (10.55) hold and $Q(V,U,W,W')$ is nonsingular.

(ii) Let $(N_1,M_1)=[V,U,W]$, $(N_2,M_2)=[V,U+1,W']$ and $(N_3,M_3)=[V+1,U+1,W'+1]$. Then by

$$fM_1-N_1=O(z^{W+1}),\quad fM_2-N_2=O(z^{W'+1})\quad\text{and}\quad fM_3-N_3=O(z^{W'+2}),$$

we have that $D(V,U,W,W')$ satisfies

$$E(V,U,W,W')D(V,U,W'W')+H(W+2,U+2,W'+2)M(V,U+1,W')=0.$$

From Lemma 10.1, it follows that (10.56) and (10.57) hold, and

$$\begin{aligned}\operatorname{rank}D(V,U,W,W') &= \operatorname{rank}[H(W+2,U+2,W'+2)M(V,U+1,W')]\\ &= \operatorname{rank}\left\{\begin{bmatrix} H(W+2,U+2,W'+1)\\ H(W'+1,U+2,W'+2)\end{bmatrix}M(V,U+1,W')\right\}\\ &= \operatorname{rank}\begin{bmatrix}0\\ H(W+2,U+2,W'+1)\end{bmatrix}=p.\end{aligned}$$

(iii) Let $(N_1,M_1)=[V,U,W]$, $(N_2,M_2)=[V,U+1,W']$ and $(N_3,M_3)=[V-1,U+1,W'-1]$. Then by their definitions

$$fM_1-N_1=O(z^{W+1}),\quad fM_2-N_2=O(z^{W'+1})\quad\text{and}\quad fM_3-N_3=O(z^{W'}),$$

and thus we have that $T(V,U,W,W')$ satisfies

$$N(V,U,W,V-1)T(V,U,W,W')+N(V,U+1,W',V-1)(I-T(V,U,W,W'))=0,$$

$$E(V,U,W,W'-1)T(V,U,W,W')=0.$$

From Lemma 10.1, it follows that (10.58) and (10.59) hold and that

$$\operatorname{rank}T(V,U,W,W')=\operatorname{rank}N(V,U+1,W',V-1)=p.$$

(iv) Let $(N_1,M_1)=[V,U,W]$, $(N_2,M_2)=[V-1,U,W-1]$ and $(N_3,M_3)=[V-1,U+1,W'-1]$. Then by

$$fM_1-N_1=O(z^{W+1}),\quad fM_2-N_2=O(z^{W})\quad\text{and}\quad fM_3-N_3=O(z^{W'}),$$

we have that $S(V,U,W,W')$ satisfies

$$N(V,U,W,V-1)+N(V-1,U,W-1,V-2)S(V,U,W,W')=0,$$

and

$$E(V,U,W,W'-1)+E(V-1,U,W-1,W'-2)S(V,U,W,W')=0.$$

From Lemma 10.1, it follows that (10.60) and (10.61) hold and that $S(V,U,W,W')$ is nonsingular. The proof of the theorem is now completed. □

Formulas (10.54),(10.56), (10.58) and (10.60) can be written in matrix form. These matrices are the coefficient matrices of numerator and denominator. For example, (10.54) can be written as

$$N(V+1,U+1,W'+1,-1)=\downarrow N(V,U,W,-1)Q(V,U,W,W')+N(V+1,U,W+1,-1),$$

$$M(V+1,U+1,W'+1)=\downarrow M(V,U,W)Q(V,U,W,W')+M(V+1,U,W+1),$$

where

$$\downarrow N(V,U,W,-1)=\begin{bmatrix}\theta_m\\ N_1^{(0)}(V,U,W)\\ \vdots\\ N_1^{(v_1)}(V,U,W)\\ \vdots\\ \theta_m\\ N_p^{(0)}(V,U,W)\\ \vdots\\ N_p^{(v_p)}(V,U,W)\end{bmatrix}\quad\text{and}\quad\downarrow M(V,U,W)=\begin{bmatrix}\theta_m\\ M_1^{(0)}(V,U,W)\\ \vdots\\ M_1^{(u_1)}(V,U,W)\\ \vdots\\ \theta_m\\ M_m^{(0)}(V,U,W)\\ \vdots\\ M_m^{(u_m)}(V,U,W)\end{bmatrix}.$$

For the coefficient matrix of the residual, there is a similar formula. Therefore, formulas (10.54)–(10.61) can be used to compute the MPAs recursively, provided two approximants $[V,U,W]$ and $[V+1,U,W+1]$ or $[V,U,W]$ and $[V,U+1,W']$ are known. These two approximants can be evaluated by the methods given in Section 9. If all the numerator degrees are the same, i.e., $U\in\mathbf{Z}_+$, we can start from $[V,0,V]=(f^{(V)}(z),\ I_m)$ and $[V+1,0,V+1]=(f^{(V+1)}(z),\ I_m)$.

From Theorem 10.4 we know that, if $p=m$ all the parameters Q,D,T,S are nonsingular.

Now we want to establish some recursive relations among the parameters Q,D,T,S.

Theorem 10.5 *Suppose $m\geq p$ and f is I–normal. Then, for given V,U,W satisfying (2.4), and W' satisfying*

$$W'\geq W+1,\quad |W'-W|=m,$$

we have for the matrices appearing in the previous theorem, the following relations :

(i)

$$D(V,U,W,W')=Q(V,U,W,W')-S(V+1,U,W+1,W'+1),\tag{10.62}$$

(ii)

$$T(V+1,U,W+1,W'+1)=Q^{-1}(V,U,W,W')D(V,U,W,W'),\tag{10.63}$$

(iii) For $W''\leq W$ and $|W+1-W''|=m$,

$$S(V+2,U,W+2,W'+2)=Q(V,U,W,W')-D(V+1,U-1,W'',W+1),\tag{10.64}$$

(iv) For $W''\leq W$ and $|W+1-W''|=m$,

$$D(V+1,U-1,W'',W+1)=T(V+1,U-1,W'',W+1)Q(V,U,W,W').\tag{10.65}$$

Proof. (i) In formula (10.60), changing the parameter (V,U,W,W') to $(V+1,U,W+1,W'+1)$ and subtracting with formula (10.54), we get

$$\begin{aligned}[V+1,U+1,W'+1]-[V,U+1,W'] &= z[V,U,W]\,[Q(V,U,W,W')\\ &\quad -S(V+1,U,W+1,W'+1)].\end{aligned}$$

It follows from (10.56) that the right hand side of the above equality is just $z[V,U,W]D(V,U,W,W')$. Therefore conclusion (i) holds.

(ii) Eliminating $z[V,U,W]$ from (10.54) and (10.56), we get

$$\begin{aligned}&[V+1,U+1,W'+1]-[V+1,U+1,W'+1]Q^{-1}(V,U,W,W')D(V,U,W,W')\\ &\quad=[V,U+1,W']-[V+1,U,W+1]Q^{-1}(V,U,W,W')D(V,U,W,W'),\end{aligned}$$

i.e.,

$$[V,U+1,W'] = [V+1,U,W+1]Q^{-1}(V,U,W,W')D(V,U,W,W')$$
$$+[V+1,U+1,W'+1]\left[I-Q^{-1}(V,U,W,W')D(V,U,W,W')\right].$$

Comparing this with formula (10.58), we get (ii).

(iii) Since

$$\sum_{i=1}^{p}(v_i+1)+\sum_{i=1}^{m}(u_i-1)=\sum_{i=1}^{p}w_i+p-m=\sum_{i=1}^{p}w_i'',$$

(2.4) holds for $(V+1,U-1,W'')$. In formula (10.58), take the parameter (V,U,W,W') to be $(V+1,U-1,W'',W+1)$, then the conditions on conclusion (iii) of the theorem hold. (10.58) now becomes

$$[V,U,W]=[V+1,U-1,W'']T(V+1,U-1,W'',W+1)$$
$$+[V+1,U,W+1](I-T(V+1,U-1,W'',W+1)).$$

Substituting it into (10.54), we have

$$[V+1,U+1,W'+1]$$
$$=z[V+1,U-1,W'']T(V+1,U-1,W'',W+1)Q(V,U,W,W')$$
$$+z[V+1,U,W+1]\left[I-T(V+1,U-1,W'',W+1)\right]Q(V,U,W,W')$$
$$+[V+1,U,W+1]. \tag{10.66}$$

In Formula (10.60), changing the parameter (V,U,W,W') to $(V+2,U,W+2,W'+2)$, we have

$$[V+1,U+1,W'+1]=[V+2,U,W+2]$$
$$+z[V+1,U,W+1]S(V+2,U,W+2,W'+2). \tag{10.67}$$

Comparing the coefficient of z of the denominator on the right hand side of (10.66) and (10.67), we get

$$Q(V,U,W,W')+M^{(1)}(V+1,U,W+1)=M^{(1)}(V+2,U,W+2)$$
$$+S(V+2,U,W+2,W'+2), \tag{10.68}$$

where $M^{(1)}(V,U,W)$ denotes the coefficient matrix of z in the denominator. In formula (10.56), take the parameter (V,U,W,W') as $(V+1,U-1,W'',W+1)$, then one has

$$[V+2,U,W+2]=z[V+1,U-1,W'']D(V+1,U-1,W'',W+1)+[V+1,U,W+1].$$

It follows that

$$M^{(1)}(V+2,U,W+2)=D(V+1,U-1,W'',W+1)+M^{(1)}(V+1,U,W+1).$$

Combining it with (10.68), we get (10.64).

(iv) Comparing the leading coefficient of the denominator in (10.66) and (10.67), we have

$$M^{hc}(V+1,U,W+1)\left[I-T(V+1,U-1,W'',W+1)\right]Q(V,U,W,W')$$
$$=M^{hc}(V+1,U,W+1)S(V+2,U,W+2,W'+2).$$

Since $M^{hc}(V+1,U,W+1)$ is nonsingular,

$$S(V+2,U,W+2,W'+2) = [I - T(V+1,U-1,W'',W+1)]\, Q(V,U,W,W').$$

Using (10.64), we get (10.65). The proof is completed. □

If $p = m$ and $V, U, W, W', W'' \in \mathbf{Z}_+$, i.e.,

$$V = v,\ U = u,\ W = w,\ W' = w',\ W'' = w'',$$

then

$$w = v + u, \quad w' = w + 1, \quad w'' = w.$$

Let

$$U_u^v = S(V+1,U-1,W,W+1),$$

$$T_u^v = T(V+1,U-1,W,W+1),$$

$$E_u^v = -D(V-1,U-1,W-2,W-1),$$

$$Q_u^v = -Q(V-1,U-1,W-2,W-1).$$

Then we have the following relations for these matrices

Corollary 10.6 *Suppose f is normal, then with the notation just introduced, we get the following qd-like relations :*

$$E_u^v = Q_u^v + U_u^{v-1}, \qquad v \geq 1, u \geq 1, \tag{10.69}$$

$$U_u^{v-1} = E_{u-1}^v - Q_u^{v-1}, \qquad v \geq 1, u \geq 1, \tag{10.70}$$

$$Q_{u+1}^v = (T_u^{v-1})^{-1} E_u^{v+1}, \quad v \geq 1, u \geq 1, \tag{10.71}$$

$$T_u^{v-1} = (Q_u^v)^{-1} E_u^v, \qquad v \geq 1, u \geq 1, \tag{10.72}$$

where

$$E_0^v = 0, \qquad v = 1, 2, \ldots, \tag{10.73}$$

$$Q_1^v = c_v^{-1} c_{v+1}, \quad v = 0, 1, \ldots, \tag{10.74}$$

$$Q_u^0 = 0, \qquad u = 2, 3, \ldots. \tag{10.75}$$

Proof. All the formulas in (10.69)–(10.72) are special cases of Theorem 10.5 except for $v = 1$ or $u = 1$ in (10.70), because the parameters Q_u^0 and E_0^v are not defined in Theorem 10.5. Therefore, the only thing to be verified is that the initial parameters given in (10.73)–(10.75) are correct indeed. From the definition of Q_1^v, we have

$$\begin{aligned} Q_1^v &= -Q(v-1,0,v-1,v) \\ &= E(v-1,0,v-1,v)^{-1} E(v,0,v,v+1) \\ &= c_v^{-1} c_{v+1}. \end{aligned}$$

i.e., (10.74) holds.

If $u = 1$ and $v \geq 1$, we get from the definition of U_1^{v-1},

$$\begin{aligned} U_1^{v-1} &= S(v, 0, v, v+1) \\ &= -N(v-1, 0, v-1, v-2)^{-1} N(v, 0, v, v-1) \\ &= -c_{v-1}^{-1} c_v. \end{aligned}$$

i.e., $U_1^{v-1} = -Q_1^{v-1}$. Therefore, (10.73) is correct.

If $v = 1, u \geq 2$, from (10.56) and (10.60) we have

$$[1, u-1, u] = -z[0, u-2, u-2]E_{u-1}^1 + [0, u-1, u-1],$$

$$[0, u, u] = [1, u-1, u] + z[0, u-1, u-1]U_u^0.$$

Hence

$$\begin{aligned} [0, u, u] = &-z[0, u-2, u-2]E_{u-1}^1 + [0, u-1, u-1] \\ &+z[0, u-1, u-1]U_u^0. \end{aligned}$$

Equating the coefficient of z in the denomimator, we have

$$U_u^0 = E_{u-1}^1 + [M^{(1)}(0, u, u) - M^{(1)}(0, u-1, u-1)].$$

Since $M^{(1)}(0, u, u) - M^{(1)}(0, u-1, u-1) = 0$, (10.75) is correct. □

Using the formulas given in the corollary, all the parameters can be evaluated recursively from the initial parameters.

From the previous corollary, we immediately get the following one

Corollary 10.7 *Suppose f is normal, then*

$$E_u^v = Q_u^v + E_{u-1}^v - Q_u^{v-1}, \quad v \geq 1, u \geq 1,$$

$$Q_{u+1}^v = (E_u^v)^{-1} Q_u^v, E_u^{v+1}, \quad v \geq 1, u \geq 1,$$

where

$$E_0^v = 0, \qquad v = 1, 2, \ldots,$$

$$Q_1^v = c_v^{-1} c_{v+1}, \quad v = 0, 1, \ldots,$$

$$Q_u^0 = 0, \qquad u = 2, 3, \ldots.$$

This is the analogue of the qd algorithm for the classical Padé approximants.

11 Conclusion

We have considered some possible definitions of matrix Padé approximants for a power series with rectangular matrix coefficients. Given the problems that occur when we try to generalize the scalar Padé approximants to the matrix case, which are :

(a) The non commutativity of the matrix multiplication,

(b) The many possibilities to fix the degrees and the orders,

(c) The many possibilities of normalization for the approximant,

we had to consider left and right approximants (a), but also type I and type II approximants (c) The choice of the degrees and orders is kept as general as possible, but most interesting properties can only hold when these are restricted to e.g. row or column degrees and sometimes even a restriction to the same order for all the entries was necessary.

We succeeded in deriving a number of properties of these approximants under appropriate restrictions which generalize related properties of scalar approximants.

Also some indications are included on how to compute the approximants.

References

[1] G.A. Baker Jr., P.R. Graves-Morris, *Padé Approximants*, (2 volumes), Addison-Wesley, Reading, Massachusetts, 1981.

[2] R.P. Brent, F.G. Gustavson, D.Y.Y. Yun, Fast solution of Toeplitz systems of equations and computation of Padé approximants, J. Algorithms, 1(1980) 259–295

[3] A. Bultheel, Recursive Relations for Block Hankel and Toeplitz systems, Part I–II, J. Comp. Appl. Math., 10(1984), 301–354.

[4] A. Bultheel, Recursive algorithm for the matrix Padé tabe, Math. Comp., 35(1980), 875–892.

[5] A. Bultheel, M. Van Barel, A matrix Euclidean algorithm and matrix Padé approximations, Report TW 104, January 1988, Department of Computer Science, K.U.Leuven.

[6] A. Bultheel, M. Van Barel, Padé techniques for model reduction in linear system theory: a survey, J. Comp. Appl. Math., 14(1986), 401–438.

[7] A. Bultheel, G. L. Xu, The structure of Minimal Padé Approximation table, Technical Report TW117, November 1988, Department of Computer Science, K.U.Leuven.

[8] A. Draux, Bibliography - Index, Report ANO - 145, Université de Sciences et Techniques de Lille, November 1984.

[9] B. Friedlander, T. Kailath, L. Ljung, M. Morf, New inversion formulas for matrices classified in terms of their distance from Toeplitz- matrices, Lin. Algebra Appl. 27, 1979, pp.31–60.

[10] P.R. Graves-Morris, Vector valued rational approximants II, IMA J. Numer. Anal., 4(1984), 209–224.

[11] G. Heinig, K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Akademie–Verlag, Berlin, 1984.

[12] I.S. Iohvidov, *Hankel and Toeplitz Matrices and Forms: Algebraic Theory*, Birkhäuser, 1982.

[13] T. Kailath, *Linear systems*, Prentice Hall, Englewood Cliffs, N.J., 1980.

[14] M. Van Barel, A.Bultheel, A Minimal partial algorithm for MIMO systems, I–V, 1988, Technical reports TW79, TW91, TW93, TW94, TW100, Department of Computer Science, K.U.Leuven.

[15] M. Van Barel, Nested minimal partial realizations and related matrix rational approximants, Ph. D. Thesis, Dept. Computer Sc., K.U. Leuven, January 1989.

[16] G.L. Xu, J.K.Li, Generalized Matrix Padé Approximants, submitted.

[17] G.L.Xu, Existence and Uniqueness of Matrix Padé Approximants, J. Comp. Math. , to appear.