Proceedings of the 17th International Symposium on Mathematical
Theory of Networks and Systems, Kyoto, Japan, July 24-28, 2006

Recursive computation of the multidimensional MPUM

Eva Zerz
Lehrstuhl D für Mathematik
RWTH Aachen
52062 Aachen, Germany
E-Mail: eva.zerz@math.rwth-aachen.de

MTNS 2006

1 Introduction to multidimensional linear exact modeling

Let a finite number of multivariate, vector-valued, polynomial-exponential functions be given, say $w_1, \ldots, w_N$, where

$$w_l : \quad \mathbb{R}^n \to \mathbb{C}^q, \quad t \mapsto p_l(t) \exp(\lambda_l t)$$

for some $p_l \in \mathbb{C}[t]^q := \mathbb{C}[t_1, \ldots, t_n]^q$, and $\lambda_l \in \mathbb{C}^{1\times n}$.

The goal is to construct a model for these data. The model class considered here consists of all linear, shift-invariant, differential systems $\mathcal{B} \subseteq \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C}^q)$. Thus, we are actually looking for a linear constant-coefficient system of partial differential equations that are satisfied by the data functions, but following the behavioral spirit, we identify a model with the solution set rather than with the equations.

One says that such a model is unfalsified by the data if it contains all w_l. Moreover, we want our equations to be as restrictive as possible, that is, they should not admit more solutions than necessary. An unfalsified model that is contained in any other unfalsified model is called the most powerful unfalsified model (MPUM).

The unique existence of the MPUM in the considered model class was shown in [1, 4] for $n = 1$, and for arbitrary n in [6, 7]: the MPUM is precisely the span (over the complex numbers) of the given functions and all their derivatives (it is clear that this space must be contained in any unfalsified model in the model class, but one has to show that conversely, it belongs to the model class). It is also known that the MPUM is finite-dimensional (as a complex vector space), because the span (over $\mathbb{C}$) of a polynomial-exponential function and all its derivatives is finite-dimensional [5].

In [6, 7], a direct construction method for the MPUM was proposed. Here, we address the question of recursive update, and we discuss some new aspects of the minimality issues raised in [6].

2 Basic idea of recursive update

We introduce the notation $\langle w_1, \ldots, w_N \rangle$ for the MPUM of $w_1, \ldots, w_N$. In other words, $\langle w_1, \ldots, w_N \rangle$ is the span, over $\mathbb{C}$, of all w_l and their derivatives.

Suppose that $\langle w_1, \ldots, w_N \rangle$ has already been constructed. This means that we have found R such that

$$R(\partial)w = 0 \quad \Leftrightarrow \quad w \in \langle w_1, \ldots, w_N \rangle. \tag{1}$$

Here, $R \in \mathbb{C}[s_1, \ldots, s_n]^{g \times q}$ is a polynomial matrix in n variables s_i, and $R(\partial) := R(\partial_1, \ldots, \partial_n)$ results from replacing each indeterminate s_i by the partial differential operator ∂_i. In this way, $R(\partial)$ is a linear constant-coefficient partial differential operator. In the situation of (1), we call R a (kernel) representation of $\langle w_1, \ldots, w_N \rangle$.

Given an additional trajectory w_{N+1}, we would like to adapt the representation accordingly. Thus we will address the following question: How should one modify the matrix R such that the result represents the MPUM of $w_1, \ldots, w_N, w_{N+1}$? In other words, the task is to construct, from R and w_{N+1}, a polynomial matrix R_{new} such that

$$R_{\text{new}}(\partial)w = 0 \quad \Leftrightarrow \quad w \in \langle w_1, \ldots, w_{N+1} \rangle. \tag{2}$$

For this, we define the error signal $e := R(\partial)w_{N+1}$, which is again polynomial-exponential. Let $\Gamma \in \mathbb{C}[s_1, \ldots, s_n]^{h \times g}$ be a representation of the MPUM of e, that is,

$$\Gamma(\partial)v = 0 \quad \Leftrightarrow \quad v \in \langle e \rangle. \tag{3}$$

We will show below that $R_{\text{new}} := \Gamma R$ represents the MPUM of $w_1, \ldots, w_{N+1}$.

Theorem: Let w_l, $1 \le l \le N+1$, be polynomial-exponential functions. Suppose that (1) and (3) hold, where $e = R(\partial)w_{N+1}$. Set $R_{\text{new}} = \Gamma R$. Then we have (2).

Proof: It is easy to verify that $R_{\text{new}}(\partial)w_l = 0$ for $1 \le l \le N+1$, because we have $R(\partial)w_l = 0$ for $1 \le l \le N$ by assumption, and $R_{\text{new}}(\partial)w_{N+1} = \Gamma(\partial)e = 0$ by construction. For the converse, let $R_{\text{new}}(\partial)\xi = 0$. We need to show that $\xi \in \langle w_1, \ldots, w_{N+1}\rangle$. We have $\Gamma(\partial)R(\partial)\xi = 0$ and thus, by the construction of Γ,

$$R(\partial)\xi \in \langle e \rangle,$$

that is,

$$R(\partial)\xi = \textstyle\sum_{\mu\in\mathbb{N}^n} a_\mu \partial^\mu e, \quad \text{where } \partial^\mu := \partial_1^{\mu_1}\cdots\partial_n^{\mu_n},$$

for some $a_\mu \in \mathbb{C}$ which are almost all zero. Then

$$R(\partial)\xi = \textstyle\sum_\mu a_\mu \partial^\mu R(\partial)w_{N+1} = R(\partial)\sum_\mu a_\mu \partial^\mu w_{N+1}$$

and thus

$$R(\partial)(\xi - \textstyle\sum_\mu a_\mu \partial^\mu w_{N+1}) = 0$$

which implies, by the assumption that R represents the MPUM of $w_1, \ldots, w_N$,

$$\xi - \textstyle\sum_\mu a_\mu \partial^\mu w_{N+1} \in \langle w_1, \ldots, w_N\rangle.$$

Thus $\xi \in \langle w_1, \ldots, w_N, w_{N+1}\rangle$ as desired. □

3 Refinement of the recursion

The adaptation scheme from above requires the computation of the MPUM of a single trajectory, namely, the error signal e. This is facilitated by breaking the problem of augmenting $\langle w_1, \ldots, w_N\rangle$ with w_{N+1} into several simpler subproblems, following the procedure proposed in [1] for the one-dimensional case.

Let $w_{N+1} = p\exp_\lambda$ be given, where $p \in \mathbb{C}[t]^q$, $\lambda \in \mathbb{C}^{1\times n}$, and $\exp_\lambda(t) = \exp(\lambda t)$. We first observe that

$$\langle w_1, \ldots, w_N, w_{N+1}\rangle = \langle w_1, \ldots, w_N, \{(\partial-\lambda)^\mu w_{N+1} \mid \mu \in \mathbb{N}^n\}\rangle, \tag{4}$$

where we use the multi-index notation $(\partial-\lambda)^\mu := (\partial_1-\lambda_1)^{\mu_1}\cdots(\partial_n-\lambda_n)^{\mu_n}$. In (4), the inclusion "$\subseteq$" is obvious (take $\mu = 0$), and the inclusion "$\supseteq$" follows from the fact that

$$(\partial-\lambda)^\mu w_{N+1} = \textstyle\sum_{\nu\le\mu}\binom{\mu}{\nu}(-\lambda)^\nu \partial^{\mu-\nu} w_{N+1},$$

where $\nu \leq \mu$ means $\nu_i \leq \mu_i$ for all i, and $\binom{\mu}{\nu} := \binom{\mu_1}{\nu_1} \cdots \binom{\mu_n}{\nu_n}$, $\lambda^\nu := \lambda_1^{\nu_1} \cdots \lambda_n^{\nu_n}$. This is a complex linear combination of w_{N+1} and its derivatives, and thus, it belongs to $\langle w_1, \ldots, w_{N+1} \rangle$. The identity

$$(\partial - \lambda)^\mu p \exp_\lambda = (\partial^\mu p) \cdot \exp_\lambda$$

implies that we can also write

$$\langle w_1, \ldots, w_{N+1} \rangle = \langle w_1, \ldots, w_N, \{(\partial^\mu p) \cdot \exp_\lambda \mid \mu \in \mathbb{N}^n\} \rangle. \tag{5}$$

Since only finitely many $\partial^\mu p$ are non-zero (for instance, it suffices to consider all multi-indices μ with $|\mu| := \mu_1 + \ldots + \mu_n \leq \deg(p)$, where $\deg(\cdot)$ denotes the total degree), the generating set on the right hand side of (5) is actually still finite, and this holds also for (4). By successively computing

$$\langle w_1, \ldots, w_N, \{(\partial^\mu p) \cdot \exp_\lambda \mid |\mu| = d\} \rangle$$

for $d = \deg(p), \deg(p) - 1, \ldots, 0$, the problem can be reduced, without loss of generality, to the situation where $w_{N+1} = p \exp_\lambda$ and for all $1 \leq i \leq n$, we have $(\partial_i p) \cdot \exp_\lambda \in \langle w_1, \ldots, w_N \rangle$ (after modification of N in each step).

The advantage lies in the fact that in this case, the error signal $e = R(\partial) w_{N+1}$ is purely exponential, that is,

$$e = \varepsilon \exp_\lambda \quad \text{for some } \varepsilon \in \mathbb{C}^g.$$

This is because

$$0 = R(\partial)((\partial_i p) \cdot \exp_\lambda) = R(\partial)(\partial_i - \lambda_i) p \exp_\lambda = (\partial_i - \lambda_i) R(\partial) w_{N+1} = (\partial_i - \lambda_i) e$$

for all i, and thus $e = \varepsilon \exp_\lambda$ for some $\varepsilon \in \mathbb{C}^g$.

The MPUM of a purely exponential function $e = \varepsilon \exp_\lambda$ is particularly easy to find: If $\varepsilon = 0$, we set $\Gamma := I$. If $\varepsilon \neq 0$, let $r := \min\{j \mid \varepsilon_j \neq 0\}$. We may assume that $\varepsilon_r = 1$, without loss of generality. Thus

$$\varepsilon = \begin{bmatrix} 0 \\ 1 \\ \varepsilon' \end{bmatrix} \in \mathbb{C}^g, \quad \text{where } \varepsilon' \in \mathbb{C}^{g-r}.$$

Set

$$\Gamma := \begin{bmatrix} I_{r-1} & 0 & 0 \\ 0 & s_1 - \lambda_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & s_n - \lambda_n & 0 \\ 0 & -\varepsilon' & I_{g-r} \end{bmatrix} \in \mathbb{C}[s_1, \ldots, s_n]^{(g-1+n) \times g}. \tag{6}$$

Then it is quite easy to see that $\Gamma(\partial) v = 0$ if and only if $v = ae$ for some $a \in \mathbb{C}$. In other words, Γ represents the MPUM of e.

4 Worked example

In [3], the MPUM of the following four trajectories was computed:

$$w_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; w_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{t_1}, \; w_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{t_2}, \; w_4 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{t_1+t_2}.$$

Since these functions are all purely exponential, the MPUM is simply given by all linear combinations of the four signals themselves, that is,

$$w \in \langle w_1, \ldots, w_4 \rangle \quad \Leftrightarrow \quad w = \textstyle\sum_{l=1}^{4} a_l w_l \text{ for some } a_l \in \mathbb{C}.$$

Using the construction from [6, 7], the MPUM is also characterized by

$$w \in \langle w_1, \ldots, w_4 \rangle \quad \Leftrightarrow \quad \exists x \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{C}^4) : \begin{cases} \partial_1 x &= A_1 x \\ \partial_2 x &= A_2 x \\ w &= Cx \end{cases} \tag{7}$$

which can be rewritten as

$$w \in \langle w_1, \ldots, w_4 \rangle \quad \Leftrightarrow \quad \exists x_0 \in \mathbb{C}^4 : w(t) = C \exp(A_1 t_1 + A_2 t_2) x_0,$$

where

$$C = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 \end{bmatrix}, \; A_1 = \operatorname{diag}(0,1,0,1), \; A_2 = \operatorname{diag}(0,0,1,1).$$

A representation can be computed by applying the fundamental principle [5] to (7). Using the computer algebra system Singular [2], we obtain

$$R = \begin{bmatrix} -s_2^2 + s_2 & 0 \\ 0 & s_2^2 - s_2 \\ -s_1 + 1 & s_2 - 1 \\ s_2 & s_1 \end{bmatrix}$$

which is equivalent to the representation given in [3]. Now let us adapt this representation to the situation where

$$w_5 = \begin{bmatrix} t_1 + t_2 \\ t_1 \end{bmatrix}$$

is an additional data trajectory. Proceeding as in Section 2, one computes the error signal, and constructs a representation of its MPUM as in [6, 7]. We obtain

$$e = R(\partial) w_5 = \begin{bmatrix} 1 \\ 0 \\ -1 + t_2 \\ 2 \end{bmatrix} \quad \text{and} \quad \Gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 \\ s_2 & 0 & 0 & 0 \\ -1 & 0 & s_2 & 0 \\ s_1 & 0 & 0 & 0 \\ 0 & 0 & s_1 & 0 \end{bmatrix}.$$

Thus we set $R_{\text{new}} = \Gamma R$. The iterative approach of Section 3 amounts to adding not only w_5, but also

$$w_6 = \partial_1 w_5 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad w_7 = \partial_2 w_5 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

However, $w_6 = w_1$ and thus can be disregarded. For w_7, we have

$$\hat{e} = R(\partial) w_7 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \hat{\Gamma} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & s_1 & 0 \\ 0 & 0 & s_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where we have used (6), and we obtain

$$\hat{R} = \hat{\Gamma} R = \begin{bmatrix} -s_2^2 + s_2 & 0 \\ 0 & s_2^2 - s_2 \\ s_1(-s_1+1) & s_1(s_2-1) \\ s_2(-s_1+1) & s_2(s_2-1) \\ s_2 & s_1 \end{bmatrix}$$

as an intermediary result. Next, we compute, again using (6),

$$\tilde{e} = \hat{R}(\partial) w_5 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \tilde{\Gamma} = \begin{bmatrix} s_1 & 0 & 0 & 0 & 0 \\ s_2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 & 1 \end{bmatrix}$$

which finally yields $R_{\text{new}} = \tilde{\Gamma} \hat{R}$. The result coincides with the one computed earlier (up to a permutation of the rows). After some streamlining through Gröbner basis computation, we obtain the final result

$$R_{\text{new}} = \begin{bmatrix} s_2^2(s_2-1) & 0 \\ 0 & s_2^2 - s_2 \\ s_2(2s_2-1) & s_1 \\ s_2(s_1-s_2) & 0 \\ s_1^2 - s_1 - s_2^2 + s_2 & 0 \end{bmatrix}$$

as a representation of the MPUM of $w_1, \ldots, w_5$. From this form, one can see (using Gröbner basis theory) that the dimension of the MPUM as a complex vector space equals 6. This is in accordance with straightforward reasoning: We have $\langle w_1, \ldots, w_4 \rangle = \operatorname{span}_{\mathbb{C}}\{w_1, \ldots, w_4\}$ and thus $\dim_{\mathbb{C}} \langle w_1, \ldots, w_4 \rangle = 4$, and similarly, $\langle w_5 \rangle = \operatorname{span}_{\mathbb{C}}\{w_5, w_6, w_7\}$ and thus $\dim_{\mathbb{C}} \langle w_5 \rangle = 3$, but the intersection $\langle w_1, \ldots, w_4 \rangle \cap \langle w_5 \rangle = \langle w_1 \rangle = \langle w_6 \rangle$ has $\mathbb{C}$-dimension 1. Put differently, the MPUM of w_2, w_3, w_4 has dimension 3, the MPUM of w_1, w_5 has dimension 3, and these two MPUMs intersect trivially.

5 Minimality issues

The direct way of computing the MPUM of a single trajectory given in [6, 7] involves the computation of the left kernel of a matrix of the form

$$H = \begin{bmatrix} s_1 I - A_1 \\ \vdots \\ s_n I - A_n \\ C \end{bmatrix} \in \mathbb{C}[s_1, \ldots, s_n]^{(n\delta+q)\times\delta},$$

where $A_i \in \mathbb{C}^{\delta\times\delta}$ are pairwise commuting matrices, and $C \in \mathbb{C}^{q\times\delta}$. The computational cost of this so-called syzygy computation [2] depends crucially on the number δ. One calls δ the size of the representation $(A_1, \ldots, A_n, C)$ of

$$\begin{aligned} \mathcal{B} &= \{w : \mathbb{R}^n \to \mathbb{C}^q | \exists x \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C}^q) : \partial_i x = A_i x \text{ for } 1 \le i \le n \text{ and } w = Cx\} \\ &= \{w : \mathbb{R}^n \to \mathbb{C}^q | \exists x_0 \in \mathbb{C}^\delta \forall t \in \mathbb{R}^n : w(t) = C \exp(A_1 t_1 + \ldots + A_n t_n) x_0\}. \quad (8) \end{aligned}$$

It was shown in [6] that such a representation is minimal (i.e., there exists no representation of strictly smaller size) if and only if

$$\bigcap_{\mu\in\mathbb{N}^n} \ker(CA^\mu) = \{0\}, \quad \text{where } A^\mu := A_1^{\mu_1} \cdots A_n^{\mu_n},$$

which is the natural generalization of observability to the considered model class. However, there is also another characterization of minimality (generalizing the Hautus test) that is actually preferable from the computational point of view in the setting of this paper.

For this, we need the following fact on joint eigenvectors of pairwise commuting matrices, which is without doubt part of the linear algebra folklore, but giving a short proof seems to be easier than finding a reference.

Lemma: Let $A_1, \ldots, A_n \in \mathbb{C}^{\delta\times\delta}$ be pairwise commuting matrices. Define

$$\operatorname{spec}(A_1, \ldots, A_n) := \{\lambda \in \mathbb{C}^{1\times n} \mid \exists 0 \neq z \in \mathbb{C}^\delta : A_i z = \lambda_i z \text{ for all } 1 \le i \le n\}.$$

Then $\operatorname{spec}(A_1, \ldots, A_n)$ is a finite, non-empty set.

Proof: The finiteness is clear, because

$$\operatorname{spec}(A_1, \ldots, A_n) \subseteq \operatorname{spec}(A_1) \times \ldots \times \operatorname{spec}(A_n).$$

Therefore it suffices to show that $\operatorname{spec}(A_1, \ldots, A_n)$ is not empty.

For $n = 1$, there is nothing to prove. For $n = 2$, let $\lambda_1 \in \operatorname{spec}(A_1)$ and let

$$\mathcal{V} := \{z \in \mathbb{C}^\delta \mid A_1 z = \lambda_1 z\}.$$

This is a non-zero subspace of $\mathbb{C}^\delta$, and it is A_2-invariant, because

$$A_1 z = \lambda_1 z \quad \Rightarrow \quad A_1 A_2 z = A_2 A_1 z = \lambda_1 A_2 z.$$

Thus $\mathcal{V}$ contains an eigenvector of A_2, say $0 \neq y \in \mathcal{V}$ with $A_2 y = \lambda_2 y$. Since $y \in \mathcal{V}$, we also have $A_1 y = \lambda_1 y$. Thus $(\lambda_1, \lambda_2) \in \operatorname{spec}(A_1, A_2)$.

Suppose that the statement has been shown for $n-1$. Let $(\lambda_1, \ldots, \lambda_{n-1}) \in \operatorname{spec}(A_1, \ldots, A_{n-1})$ and let

$$\mathcal{V} := \{z \in \mathbb{C}^\delta \mid A_1 z = \lambda_1 z, \ldots, A_{n-1} z = \lambda_{n-1} z\}$$

which is a non-zero A_n-invariant subspace of $\mathbb{C}^\delta$, which consequently contains an eigenvector of A_n. The rest of the argument is analogous to the case where $n = 2$.

$\square$

Now we can give the alternative characterization of minimality mentioned above.

Theorem: Let $A_1, \ldots, A_n \in \mathbb{C}^{\delta \times \delta}$ be pairwise commuting matrices and let $C \in \mathbb{C}^{q \times \delta}$. The following are equivalent:

1. $\bigcap_{\mu \in \mathbb{N}^n} \ker(CA^\mu) = \{0\}$.
2. For all $\lambda \in \operatorname{spec}(A_1, \ldots, A_n)$, we have $\ker(H(\lambda)) = \{0\}$, where

$$H(\lambda) := \begin{bmatrix} \lambda_1 I - A_1 \\ \vdots \\ \lambda_n I - A_n \\ C \end{bmatrix} \in \mathbb{C}^{(n\delta + q) \times \delta}.$$

Proof: If there exists $\lambda \in \operatorname{spec}(A_1, \ldots, A_n)$ and $0 \neq z \in \mathbb{C}^\delta$ such that $Cz = 0$ and $(\lambda_i I - A_i) z = 0$ for all i, then

$$CA^\mu z = C\lambda^\mu z = \lambda^\mu C z = 0,$$

where $\lambda^\mu = \lambda_1^{\mu_1} \cdots \lambda_n^{\mu_n}$.

Conversely, let

$$\mathcal{V} := \bigcap_{\mu \in \mathbb{N}^n} \ker(CA^\mu) \neq \{0\}.$$

Let V be a matrix whose columns are a basis of $\mathcal{V}$, and let $1 \leq i \leq n$. Since $\mathcal{V}$ is A_i-invariant, we have $A_i V = V B_i$ for some B_i. The fact that $A_i A_j = A_j A_i$ for all i, j implies that B_i are also pairwise commuting matrices. Thus there exists

$\lambda \in \operatorname{spec}(B_1, \ldots, B_n)$, that is, $B_i z = \lambda_i z$ for some $0 \neq z \in \mathbb{C}^{\dim(\mathcal{V})}$. However, this implies that $A_i V z = \lambda_i V z$ for all i. Since $0 \neq Vz \in \mathcal{V}$, we have

$$\begin{bmatrix} \lambda_1 I - A_1 \\ \vdots \\ \lambda_n I - A_n \\ C \end{bmatrix} Vz = 0.$$

Moreover, $\lambda \in \operatorname{spec}(A_1, \ldots, A_n)$ which completes the proof. $\square$

The Hautus condition from above leads to a procedure for reducing a given representation to minimality, which can be used as a pre-processing step for the syzygy calculation. The combination of tools from computational linear algebra with computer algebraic techniques (Gröbner bases etc.) can yield a considerable speed-up of the symbolic calculation. Moreover, if the matrices A_i result from the construction method of [6, 7], then $\operatorname{spec}(A_1, \ldots, A_n)$ is actually known, because it consists precisely of the frequency vectors λ that are present among the data trajectories.

Corollary: Let $A_1, \ldots, A_n \in \mathbb{C}^{\delta \times \delta}$ be pairwise commuting matrices and let $C \in \mathbb{C}^{q \times \delta}$. Suppose that there exists $\lambda \in \operatorname{spec}(A_1, \ldots, A_n)$ with

$$r := \operatorname{rank}(H(\lambda)) < \delta.$$

Then there exist pairwise commuting matrices $A_1^{(1)}, \ldots, A_n^{(1)} \in \mathbb{C}^{r \times r}$ and $C^{(1)} \in \mathbb{C}^{q \times r}$ such that $(A_1, \ldots, A_n, C)$ and $(A_1^{(1)}, \ldots, A_n^{(1)}, C^{(1)})$ represent the same system according to (8).

Proof: Select r linearly independent columns in $H(\lambda)$. Without loss of generality, suppose that these are the first r columns of $H(\lambda)$. Then

$$H(\lambda) = [H_1(\lambda), H_1(\lambda)X],$$

where $H_1(\lambda)$ has full column rank and $X \in \mathbb{C}^{r \times (\delta - r)}$. Let $C = [C^{(1)}, C^{(1)}X]$ and

$$A_i = \begin{bmatrix} A_{i11} & A_{i12} \\ A_{i21} & A_{i22} \end{bmatrix}$$

be partitioned accordingly. Set $A_i^{(1)} := A_{i11} + XA_{i21}$ and let

$$T := \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}.$$

By assumption, we have

$$\begin{aligned} (\lambda_i I - A_{i11})X &= -A_{i12} \\ -A_{i21}X &= \lambda_i I - A_{i22}. \end{aligned}$$

Using this, a straightforward computation yields

$$T^{-1}A_iT = \begin{bmatrix} A_i^{(1)} & 0 \\ A_{i21} & \lambda_i I \end{bmatrix} \text{ and } CT = [C^{(1)}, 0],$$

which implies the result. $\square$

Note that there is no guarantee that $A_1^{(1)}, \ldots, A_n^{(1)}$, $C^{(1)}$ is already minimal, that is, it may happen that

$$\operatorname{rank} \begin{bmatrix} \lambda_1 I - A_1^{(1)} \\ \vdots \\ \lambda_n I - A_n^{(1)} \\ C^{(1)} \end{bmatrix} < r.$$

Then the same reduction procedure can be applied to $A_1^{(1)}, \ldots, A_n^{(1)}$, $C^{(1)}$. Since the size of the matrices becomes strictly smaller in each step, we obtain a minimal model after finitely many iterations.

References

[1] A. C. Antoulas, J. C. Willems. A behavioral approach to linear exact modeling. *IEEE Trans. Automatic Control* 38, 1776–1802 (1993).

[2] G.-M. Greuel, G. Pfister, H. Schönemann. Singular 3.0. A Computer Algebra System for Polynomial Computations. Centre for Computer Algebra, University of Kaiserslautern, Germany, 2005. `http://www.singular.uni-kl.de`.

[3] P. Rocha, J. C. Willems, I. Bras. On the Markov property for continuous multidimensional behaviors. *Proceedings of the MTNS 2004*, Leuven, Belgium, 2004.

[4] J. C. Willems. From time series to linear system, II. Exact modelling. *Automatica* 22, 675–694 (1986).

[5] J. Wood, U. Oberst, E. Rogers, D. H. Owens. A behavioral approach to the pole structure of one-dimensional and multidimensional linear systems. *SIAM J. Control Optimization* 38, 627–661 (2000).

[6] E. Zerz. Characteristic frequencies, polynomial-exponential trajectories, and linear exact modeling with multidimensional behaviors. *SIAM J. Control Optimization* 44, 1148–1163 (2005).

[7] E. Zerz. Poles, polynomial-exponential trajectories, and the MPUM. *Proceedings of the MTNS 2004*, Leuven, Belgium, 2004.