

# Linear regressive realizations of LTI state space models

Karel J. Keesman and Nurulhuda Khairudin\*

\* *Systems and Control Group, Wageningen University, The Netherlands*

---

**Abstract:** In this paper, linear regressive model structures are deduced from single input - single output (SISO) deterministic LTI systems in state space form. The resulting linear regression and predictor are exact representations of a discrete-time LTI state space model  $\Sigma_d$  belonging to a certain class  $\mathcal{S}$ . As an illustrative example, a process with first-order reaction and advective transport is approximated by a discrete-time state space model  $\Sigma_d$ , where the one-dimensional space is divided into  $n$  compartments and is reparametrized such that  $\Sigma_d \subseteq \mathcal{S}$ . The system output can then be explicitly predicted by  $y(t) = \theta^T \phi(t-1) - \tilde{\gamma}(t-1)$  as a function of the number of compartments  $n$ , the sensor position  $j^*$ , the parameter vector  $\theta = \xi(\vartheta)$  with  $\vartheta$  a vector with original (physical) parameters, and input-output data. The method is attractive for experimental design, direct parameter estimation and prediction when prior (physical) parameter knowledge is to be preserved.

Keywords: LTI systems, distributed parameter systems, linear regression, sensitivity analysis, sensor location, prediction

---

## 1. INTRODUCTION

Estimation of physical parameters in continuous-time linear time-invariant (LTI) state space models generally leads to nonlinear estimation problems (see *e.g.* [1, 2]). It is well known that nonlinear estimation problems frequently lead to local minima solutions. Furthermore, solving these problems can be very (computer) time consuming, especially when multi-start procedures are used. Since the global solution is not known beforehand, no characterization of the systematic error in the estimates can be given.

Our overall goal is to uniquely estimate model parameters in state space model structures while preserving the (original) physical model structure (see also [3],[4],[5]). For continuous-time LTI systems, our approach consists of two steps, *i.e.* first the system is explicitly discretized and subsequently a linear regressive realization of the dynamical system is found. Unlike data based methods, such as subspace identification [6, 7], we will conserve the physical model structure.

In particular, we will consider the class of finite LTI state space systems  $\mathcal{S}(\Sigma)$ . Under some mild assumptions, we are able to find a linear regressive realization of  $\Sigma$  which is suited for linear estimation and prediction. It will be shown later on, that the resolvent of the system matrix ( $R(A)$ ) of the discrete-time system, plays a key role in this. In specific cases, the discrete-time system matrix ( $I + A(\vartheta)$ ), with  $\vartheta$  a vector with physical parameters, becomes a bidiagonal matrix or a symmetric tridiagonal matrix. For the latter case, explicit expressions of  $R(A)$  are known (see *e.g.* [8, 9, 10, 11]). In particular, we want to unravel the structure of  $R(A)$  in such a way that we can write the LTI system in state space form as a linear regressive set of equations:  $\theta^T \phi = \gamma$ . Herein,  $\theta = \xi(\vartheta)$

contains known reparametrization functions that are not confounded with coefficients that originate from either the discretization step or from constants in  $A$ . From here, it is rather straightforward to arrive at a direct least-squares type of estimate  $\hat{\theta}$  and at an explicit expression for the output at time instant  $k$ .

The key objective of the paper is twofold: to show the derivation of linear regressive model structures from deterministic LTI state space models, while conserving the physical model structure, and to illustrate this to an advection-reaction system.

First, we formulate the problem in section 2. In section 3, we illustrate the methodology to arrive at an exact linear regressive realization suitable for linear estimation and prediction for the class  $\mathcal{S}(\Sigma)$ . Further, it will be shown that explicit expressions of the linear regressive realization can be found without the use of explicit matrix inversion. In section 4, we work out the derivation of an explicit linear regression and predictor for an infinite-dimensional advection-reaction process with boundary input at one side of the 1D-dimensional spatial plane. Furthermore, the results of a sensitivity analysis are presented. In section 5, the proposed method is shortly discussed. Finally, in section 6, conclusions are drawn and a future outline is given.

## 2. PROBLEM STATEMENT

Let us start by describing the problem in some more detail. In this contribution, we study the deduction of a physically-based linear regressive output predictor  $y(t|\theta; Z^-)$  from a deterministic LTI system in state space form belonging to the class  $\mathcal{S}(\Sigma)$ , with  $Z = [u(1), y(1), \dots, u(t-1), y(t-1), u(t), y(t)]$ , the noise-free

input-output data set, and  $Z^- = Z \setminus y(t)$ . In particular, we would like (i) to find the linear regression  $\theta^T \phi = \gamma$  and associated reparametrized vector function  $\theta = \xi(\vartheta)$  and (ii) to derive an expression for the predictor  $y(t|\theta; Z^-)$  in terms of the number of compartments  $n$  and the sensor location  $j^*$ .

We consider the following structured parametrization state space model structure as given,

$$\Sigma(A, B, C, D) := \begin{cases} \dot{x}(t) &= A(\vartheta)x(t) + B(\vartheta)u(t) \\ y(t) &= C(\vartheta)x(t) + D(\vartheta)u(t) \end{cases} \quad (1)$$

with  $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^s$  and matrices  $A, B, C$  and  $D$  of appropriate dimensions and with  $\vartheta$  the physical parameter vector. Hence, the problem is to find the reparametrized vector function  $\theta = \xi(\vartheta)$  and the transformation of  $\Sigma$  to the linear regressive system,

$$\tilde{\Sigma}_d(\phi, \gamma) := \begin{cases} \theta^T \phi(Z) = \gamma(Z) \\ y(t) = g(\theta; Z^-) \end{cases} \quad (2)$$

In the following, we will restrict ourselves to the SISO case with  $C$  a constant vector and  $D$  a constant.

### 3. REALIZATION OF A LINEAR REGRESSIVE PARAMETRIC SYSTEM REPRESENTATION

#### 3.1 Discretization

The first step towards a linear regressive realization of the continuous-time model (1) is to approximate it by a discrete-time model with time instant  $k$ . For ease of notation, we will write the time instant  $k \in \{0, 1, \dots, N\}$  as an index. Typically, an Euler forward scheme, with  $\Delta t = 1$  as a result of appropriate scaling guaranteeing a stable approximation, leads to

$$\Sigma_d : \begin{cases} x_{k+1} &= (I + A(\vartheta))x_k + B(\vartheta)u_k \\ y_k &= Cx_k + Du_k \end{cases} \quad (3)$$

where  $x \in \mathbb{R}^n$  and  $u, y \in \mathbb{R}$ . In what follows, it is assumed that  $A(\vartheta)$  and  $B(\vartheta)$  can be written linearly in  $\vartheta$  as

$$A(\vartheta) = \bar{A} + \sum \tilde{A}_i \vartheta_i \quad (4)$$

$$B(\vartheta) = \bar{B} + \sum \tilde{B}_i \vartheta_i \quad (5)$$

with  $\vartheta$  a lumped combination of original physical parameters such that (3)-(5) holds. For instance, for an advection-reaction process, we will see later on that the original parameters are the velocity, (first-order) reaction constant and the time and spatial discretization constants and that  $\vartheta$  contains lumped combinations of these.

Alternatively, applying the trapezium rule in the numerical scheme leads to the following rational system,

$$\Sigma_d^* : \begin{cases} x_{k+1} &= \left(I - \frac{A(\vartheta)}{2}\right)^{-1} \left(I + \frac{A(\vartheta)}{2}\right) x_k \\ &+ \left(I - \frac{A(\vartheta)}{2}\right)^{-1} \frac{B(\vartheta)}{2} (u_{k+1} + u_k) \\ y_k &= Cx_k + Du_k \end{cases} \quad (6)$$

*Remark 1.* (a) Notice from the first step in the discretization of (1) using the trapezium rule, *i.e.*

$$x_{k+1} - x_k = \frac{A(\vartheta)}{2}(x_{k+1} + x_k) + \frac{B(\vartheta)}{2}(u_{k+1} + u_k)$$

that an implicit numerical scheme inevitably leads to a (mixed) *error-in-variables* (EIV) estimation problem if noise would have been considered (see [4],[5] for estimation of rational systems). However, as yet we only focus on the realization problem from (1) to (2).

(b) Recall that (1), with  $C$  and  $D$  constant, has the recursive solution

$$x_{k+1} = e^{A(\vartheta)\Delta t} x_k + \int_t^{t+\Delta t} e^{A(\vartheta)[t+\Delta t-s]} B(\vartheta)u(s) ds$$

$$y_k = Cx_k + Du_k$$

and thus for each time index  $k$  the output can be accurately calculated. Hence, the approximation error is smaller than in the case of an explicit or implicit scheme. However, notice that this approach still does not allow a linear regression realization of (1), since the parameters in  $A(\vartheta)$  appear nonlinearly in the expression.

In the sequel, the arguments of  $A$  and  $B$  will be omitted for reasons of eligibility. Introducing the forward-shift operator  $q$  for manipulating the linear difference equations of (3) and (6) gives for the explicit scheme

$$G(\vartheta, q) = C(qI - (I + A))^{-1}B + D \quad (7)$$

and for the implicit scheme

$$G(\vartheta, q) = C\left(qI - \left(I - \frac{A}{2}\right)^{-1}\left(I + \frac{A}{2}\right)\right)^{-1}B + D \quad (8)$$

Hence, introducing the forward-shift operator leads directly to a discrete-time transfer function of (3)-(6), which will appear to be an effective step towards a linear regressive realization  $\tilde{\Sigma}_d$  (see Eqn. 2). In the next section, we will consider the transfer function  $G(\vartheta, q) = C(qI - (I + A(\vartheta)))^{-1}B(\vartheta) + D$ , obtained from an explicit scheme (see Eqn. 7), as a rational function of polynomials in the forward shift operator  $q$  and the parameter(s)  $\vartheta$ .

#### 3.2 Linear regression by reparametrization

As yet, our focus is on the classical Euler forward approximation of the deterministic state space model (1). Splitting the rational transfer function  $G(\vartheta, q)$  in a numerator  $\tilde{N}(\vartheta, q)$  and a denominator  $\tilde{M}(\vartheta, q)$  gives

$$\tilde{M}(\vartheta, q)y_k = \tilde{N}(\vartheta, q)u_k \quad (9)$$

with  $\tilde{M}$  and  $\tilde{N}$  polynomials in  $\vartheta$  and  $q$ . For instance, Pintelon *et al.* (see [12]) split  $\tilde{M}$  and  $\tilde{N}$  in functions of the polynomial variable  $q$  (or the Laplace variable  $s$ ) and the parameter vector  $\vartheta$ . Hence, they consider  $M(\vartheta)m(q)y_k = N(\vartheta)n(q)u_k$ . Consequently, the entries of the row vectors  $M(\vartheta)$  and  $N(\vartheta)$  are considered as black-box parameters for further estimation and prediction. As a consequence, the link to the underlying model parameters  $\vartheta$  will soon be lost as the polynomial order of  $M$  and/or  $N$  increases. However, to prevent this loss we introduce the so-called *pqr*-decomposition, by decomposing  $\tilde{M}$  and  $\tilde{N}$  not only in a shift operator ( $q$ ) dependent, but also

in a parameter ( $p$ ) dependent part so to obtain regressor weighting matrices ( $R$ ). More specifically,

$$\tilde{M} = \tilde{p}_M^T M \psi(q), \quad \tilde{N} = \tilde{p}_N^T N \psi(q) \quad (10)$$

with  $\tilde{p}_M \in \mathbb{R}^{p_1}$ ,  $\tilde{p}_N \in \mathbb{R}^{p_2}$  and  $\psi(q) = (1 \ q \ \dots \ q^{n-1} \ q^n)^T$ . Notice that the maximum order of  $q$  in  $\psi(q)$  equals the dimension of the state vector  $x$ . For each parameter  $\tilde{p}_{\bullet i}$ , a polynomial function  $\xi_i(\vartheta)$  for  $i = 1, 2, \dots$  will be found. The polynomial coefficient matrices are defined as,

$$M = \begin{pmatrix} \bar{M} & a_M \\ b_M & c_M \end{pmatrix} \in \mathbb{R}^{p_1 \times (n+1)} \quad (11)$$

$$N = \begin{pmatrix} \bar{N} & a_N \\ b_N & c_N \end{pmatrix} \in \mathbb{R}^{p_2 \times (n+1)} \quad (12)$$

with  $a_M \in \mathbb{R}^{(p_1-1) \times 1}$  and  $a_N \in \mathbb{R}^{(p_2-1) \times 1}$ . Both  $b_M$ ,  $b_N \in \mathbb{R}^{1 \times n}$  and  $c_M$ ,  $c_N \in \mathbb{R}$ .

Furthermore, define

$$Z_k = (U_k^T \mid Y_k^T)^T = \left( \bar{U}_k^T \ u_{k+n} \mid \bar{Y}_k^T \ y_{k+n} \right)^T \quad (13)$$

where

$$\bar{U}_k = (u_k \ u_{k+1} \ \dots \ u_{k+n-1})^T$$

$$\bar{Y}_k = (y_k \ y_{k+1} \ \dots \ y_{k+n-1})^T$$

Substituting (10)–(12) in (9), a transfer function representation of  $\Sigma_d$  is given by

$$\tilde{p}_M^T M \psi(q) y_k = \tilde{p}_N^T N \psi(q) u_k \quad (14)$$

Let the polynomial  $\tilde{p}_M^T M \psi(q)$  be monic. Then, we choose  $M$  such that  $\tilde{p}_M = (\tilde{p}_M^T \ 1)^T$ . In a similar way, but now because of constants in matrix  $A$ , we define  $\tilde{p}_N = (\tilde{p}_N^T \ 1)^T$ . Hence, we obtain

$$\begin{pmatrix} \tilde{p}_N^T \ 1 \mid \tilde{p}_M^T \ 1 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} \bar{N} & a_N \\ b_N & c_N \end{pmatrix} & \mathbf{O}^{p_2 \times (n+1)} \\ \mathbf{O}^{p_1 \times (n+1)} & \begin{pmatrix} -\bar{M} & -a_M \\ -b_M & -c_M \end{pmatrix} \end{pmatrix} \begin{pmatrix} U_k^T \\ Y_k^T \end{pmatrix} = 0 \quad (15)$$

Notice that we have decomposed the transfer function representation into a part with only physical parameters and a part that contains pre-filtered input-output data. Correspondingly, (15) covers the classical linear regression case with black-box parameters that are linear combinations of the original physical parameters. Expanding (15) gives

$$\begin{pmatrix} \tilde{p}_N^T \bar{N} + b_N \mid \tilde{p}_N^T a_N + c_N \mid -\tilde{p}_M^T \bar{M} - b_M \mid \\ -\tilde{p}_M^T a_M - c_M \end{pmatrix} \begin{pmatrix} \bar{U}_k \\ u_{k+n} \\ \bar{Y}_k \\ y_{k+n} \end{pmatrix} = 0 \quad (16)$$

Notice that  $a_M \neq 0$  naturally leads to an error-in-variables problem. Collecting the physical parameters gives

$$\begin{pmatrix} \tilde{p}_N^T \mid \tilde{p}_M^T \end{pmatrix} \begin{pmatrix} \bar{N} & a_N \mid \mathbf{O} \\ \mathbf{O} & -\bar{M} & -a_M \end{pmatrix} Z_k + b_N \bar{U}_k + c_N u_{k+n} - b_M \bar{Y}_k - c_M y_{k+n} = 0 \quad (17)$$

Consequently, we arrive at the linear regression

$$\theta^T \phi_k = \gamma_k \quad (18)$$

with  $\theta^T = (\tilde{p}_N^T \ \tilde{p}_M^T)$ ,  $\gamma_k = (-b_N \ -c_N \ b_M \ c_M) Z_k$  and

$$\phi_k = \begin{pmatrix} \bar{N} & a_N \mid \mathbf{O} \\ \mathbf{O} & -\bar{M} & -a_M \end{pmatrix} Z_k \quad (19)$$

In addition to this, the linear predictor can be written as

$$y_{k+n} = c_M^{-1} (\theta^T \phi_k - \tilde{\gamma}_k), \quad c_M \neq 0 \quad (20)$$

with  $\tilde{\gamma}_k = b_N \bar{U}_k + c_N u_{k+n} - b_M \bar{Y}_k$ . After multiplication of (14) by  $q^{-n}$  (*i.e.* introducing a backward time shift), we may write (18)–(20) as:

$$\tilde{\Sigma}_d : \begin{cases} \theta^T \phi_{k-n} & = \gamma_{k-n} \\ y_k & = g(\theta, \tilde{Z}_{k-n}) \end{cases} \quad (21)$$

with  $g(\cdot, \cdot)$  a function in  $\theta$  and  $\tilde{Z}_{k-n} = (U_{k-n}^T \ \bar{Y}_{k-n}^T)^T$ . This leads to the following proposition, for  $D = 0$  (without loss of generality).

**Proposition 1.** Given system  $\Sigma_d(A, B, C)$  as in (3). Then,

- (i) exact explicit expressions of  $M$  and  $N$  as a function of  $n$  exist.
- (ii)  $\Sigma_d$  can be written in the form of  $\tilde{\Sigma}_d$  with  $\theta = \xi(\vartheta)$ , a polynomial vector function in  $\vartheta$ .

Due to space limitations only a sketch of the proof is given.

**Proof.**

- (i) Let  $r(A, q) = (qI - \tilde{A})$  with  $\tilde{A}$  as in (7) or (8). Given that  $G = \frac{\tilde{N}}{\tilde{M}} = CRB$  with  $R = r^{-1}$ , the resolvent of  $\tilde{A}$ , it follows that:  $R = \text{adj}(r)/|r|$ , iff  $r$  non-singular. Hence,

$$\tilde{M} = |r| \text{ and } \tilde{N} = C \text{adj}(r)B \quad (22)$$

Since we consider only SISO systems:

$$\tilde{N} = \text{adj}(r)_{i^*j^*} = (-1)^{i^*+j^*} |r_{j_i^*}| \quad (23)$$

with  $i^*$  and  $j^*$  the non-zero entry indices of  $B$  and  $C$ , respectively. The matrix  $r_{j_i^*}$  is the sub-matrix of  $r$ , resulting from the deletion of row  $i^*$  and column  $j^*$ . From induction, and given (4)–(5), it follows that  $|r|$  and  $\text{adj}(r)$  will be a polynomials in  $q$  and  $\vartheta$  with maximal order  $n$ . Since  $G = CRB$  and  $A$ ,  $B$  linear in  $\vartheta$ , it follows that  $\tilde{M}(\vartheta, q)$  and  $\tilde{N}(\vartheta, q)$  can be decomposed as in (10)–(12), with polynomial coefficient matrices  $M$  and  $N$  being functions of  $n$ .

- (ii) Given  $G(\vartheta, q) = \frac{\tilde{N}(\vartheta, q)}{\tilde{M}(\vartheta, q)}$ , the transfer function of  $\Sigma_d$ .

Then,  $\tilde{M}(\vartheta, q) y_k = \tilde{N}(\vartheta, q) u_k$  and via direct algebra (see (10)–(18)) and the proof of part (i) one readily obtains  $\theta^T \phi_k = \gamma_k$  with  $\theta = \xi(\vartheta)$ , a polynomial vector function. Notice that  $\gamma_k$  contains  $y_{k+n}$  and we can write the equivalent form  $\tilde{\Sigma}_d$  (21) after multiplication with  $q^{-n}$ . □

**Remark 2.** From the proof of Proposition 1 it directly follows that,

- (a) The polynomial order of  $\tilde{N}(\vartheta, q)$  in  $q$  and  $\vartheta$  is determined by  $B$ ,  $C$  and  $\text{adj}(r)$ .  $\tilde{N}$  is a polynomial of maximal order  $n$  in  $\vartheta_i$  and maximal order  $n-1$  in  $q$ , since  $D = 0$ .
- (b) Explicit expressions of  $\tilde{M}(\vartheta, q)$  are much easier to find than those for  $\tilde{N}(\vartheta, q)$ , since  $\tilde{N}$  also depends on  $B$  and  $C$ , *i.e.* the non-zero entry indices  $i^*$  and  $j^*$ , respectively.

(c) In general,  $B(\vartheta)$  will lead to  $p_N \neq p_M$ , while a constant  $B$  allows  $p_N = p_M$ .

Before continuing, let us first demonstrate the  $pqR$ -decomposition on a simple example with constant  $B$ .

*Example 1.* Given a discrete-time LTI state space system  $\Sigma_d(A, B, C)$  with

$$A = \begin{pmatrix} 1 - \vartheta_2 & 1 \\ 1 & 1 - \vartheta_1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad C = (1 \ 0)$$

Then, the transfer function of this system reads  $G(\vartheta, q) = (q - 1 - \vartheta_2)/(q^2 - (\vartheta_1 + \vartheta_2 + 2)q + \vartheta_1 + \vartheta_2 + \vartheta_1\vartheta_2)$  with the cross term  $\vartheta_1\vartheta_2$ . After defining  $p_N^T = (\vartheta_2 \ 1)$ ,  $p_M^T = (\vartheta_1\vartheta_2 \ \vartheta_1 \ \vartheta_2 \ 1)$  and  $\psi(q)^T = (q^0 \ q^1 \ q^2)$ , we obtain the coefficient matrices  $M$  and  $N$ :

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

□

Alternatively, we could have chosen  $p_N = p_M$ , which would add two zero row vectors in  $N$ .

### 3.3 Sensitivity analysis

Recall that so far the model structure representation  $\tilde{\Sigma}_d(\phi, \gamma)$ , which contains the parameter vector  $\theta$  as a function of the physical parameter vector  $\vartheta$ , has been derived. Let us now, on the basis of this, investigate its sensitivities. It can be easily verified that the parameter sensitivities in a linear regression structure  $\frac{d\gamma_k}{d\theta}$  are completely defined by  $\phi_k$ . If we substitute  $Y_k$  by its convolution sum  $HU_k$ , with  $H \in \mathbb{R}^{n_H \times n_H}$ , the sensitivities only become dependent on the input  $U_k$ . The entries of the Hankel matrix  $H$  contain the *Markov* parameters and can be directly found by an impulse response as a function of the real parameters as  $\theta_0$ . Consequently, after substitution of  $Y_k = HU_k$  in  $\phi_k$  the second derivative  $\gamma_{\theta U} := \frac{d\gamma_{\theta}}{dU}$  only depends on the Hankel matrix and the coefficient matrices  $M$  and  $N$ . Since the sensor location only affects the entries in  $N$  via  $C$ , for good estimation conditions evaluation of  $N$  suffices, providing that no pole-zero cancelation occurs.

A full demonstration of the linear regression realization approach to an advection-reaction system is given in the next section.

## 4. RESULTS

### 4.1 An advection-reaction process

Consider the infinite dimensional system  $\Sigma^e$  on  $[0, \infty) \times [0, \infty)$ ,

$$\Sigma^e : \begin{cases} \frac{\partial w}{\partial t}(z, t) + \nu \frac{\partial w}{\partial z}(z, t) = -\kappa w(z, t), \\ w(z, 0) = w_0(z), \quad w(0, t) = u(t) \\ y(t) = w(z^*, t) \end{cases} \quad (24)$$

where  $w(z, t)$  is the concentration of some solute,  $\nu$  a constant flow velocity and  $\kappa$  a constant reaction constant. Furthermore, we assume that  $w_0(z) \in L_2(0, \infty)$ ,  $z^* \in [0, \infty)$  and  $u(t) \in \mathbb{R}$ .

Typically, the unknown parameters  $\nu$  and  $\kappa$  are estimated from data via non-linear estimation techniques using a numerical solution of (24). As mentioned before, this may easily lead to local solutions and a high computational effort. Following the line of this paper, we seek an approximation of (24),  $\Sigma_d^e$ , such that the linear regressive realization  $\tilde{\Sigma}_d$  as in (21) can be constructed. In the sequel, we will follow the same name conventions for the subsections as in section 3.

### 4.2 Discretization and shift-operator calculus

After applying an Euler forward scheme in space and time, we get the discrete-time compartmental system  $\Sigma_d^e$  in the form of (3) by defining the lumped physical parameters,

$$\vartheta_1 = (\kappa + \frac{\nu}{\Delta z})\Delta t, \quad \vartheta_2 = \frac{\nu\Delta t}{\Delta z}$$

where  $\Delta t$  and  $\Delta z$  denote the time and space differences, respectively. Notice that, for this specific case,  $A$  becomes a bi-diagonal matrix with  $A(\vartheta) = -\vartheta_1 I + \vartheta_2 I_{-1}$ . Furthermore,  $B(\vartheta) = [\vartheta_2 \ 0 \ \dots \ 0]^T$ , and thus  $i^* = 1$ . The observation  $y(t) = w(t, z^*)$  is approximated by  $y_k = Cw_k$ , with  $C$  mapping a ‘point’ observation at the  $j^*$ -th compartment (*i.e.*  $c_{j^*} = 1$  and  $c_j = 0$  for  $j \neq j^*$ ). Finally, from (24) we directly derive that  $D = 0$ . Notice that a grid with  $n$  points directly leads to  $n$  states, because in the distributed parameter system  $\Sigma^e$  we consider only one state variable in one spatial direction.

After forward-shift operations on  $\Sigma_d^e$  we obtain  $G(\vartheta, q) = CR(A, q)B$ . To obtain further insight into this, let us focus on the structure of the matrix  $R$ .

### 4.3 Linear regression realization

Recall that  $R(A, q) = r(A, q)^{-1} = (qI - (I + A))^{-1}$ . For our example,  $r(A, q)$  is given by

$$r(A, q) = \begin{pmatrix} q - 1 + \vartheta_1 & 0 & \dots & 0 \\ -\vartheta_2 & q - 1 + \vartheta_1 & 0 & \dots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \dots & -\vartheta_2 & q - 1 + \vartheta_1 \end{pmatrix} \quad (25)$$

We can decompose the numerator  $\tilde{N}$  and the denominator  $\tilde{M}$  of the transfer function  $G$  as in (10), given  $B$  and  $C$ . For a bi-diagonal system matrix  $A$ , as in the advection-reaction case, with unknown diagonal elements  $\vartheta_1$  and sub-diagonal elements  $\vartheta_2$  we obtain the following.

*Proposition 2.* Given  $\Sigma_d^e(A, B, C)$ , with  $A$ ,  $B$  and  $C$  as defined above and  $i^* = 1$ , then exact explicit expressions for  $\tilde{M}$ ,  $\tilde{N}$  and their components are given by

$$\begin{aligned}\tilde{p}_M^T &= (\vartheta_1^{j^*} \dots \vartheta_1 \ 1) \\ \tilde{p}_N^T &= (\vartheta_2^{j^*}) \\ \psi(q) &= (1 \ q \dots \ q^{j^*})^T \\ M = (m_{kl}) &= \begin{cases} (-1)^{-k+l} \binom{j^*}{j^* - k + 1} \binom{k-1}{l-1}, & k \geq l \\ 0 & \text{elsewhere} \end{cases} \\ N = (n_{kl}) &= \begin{cases} 1 & \text{for } k = l = 1 \\ 0 & \text{elsewhere} \end{cases}\end{aligned}$$

where  $l = 1, \dots, j^* + 1$ .

**Proof.** Via direct application of (22) and (23), we obtain

$$\begin{aligned}\tilde{M} &= |r(A, q)| = (q - 1 + \vartheta_1)^n \\ \tilde{N} &= (-1)^{i^*+j^*} |r(A, q)_{j^* i^*} B_{i^*}| \\ &= \begin{cases} 0 & i^* > j^* \\ (q - 1 + \vartheta_1)^{n+i^*-j^*-1} \vartheta_2^{j^*-i^*+1} & i^* \leq j^* \end{cases}\end{aligned}$$

with  $i^* = 1$  and  $j^*$  free. Hence,  $\tilde{N} = (q - 1 + \vartheta_1)^{n-j^*} \vartheta_2^{j^*}$ ,  $j^* \geq 1$ . Pole-zero cancelation leads to

$$\frac{\tilde{N}}{\tilde{M}} = \frac{\vartheta_2^{j^*}}{(q - 1 + \vartheta_1)^{j^*}}$$

Decomposition of the polynomials  $\tilde{M}$  and  $\tilde{N}$ , using direct algebra, gives  $\tilde{p}_N$ ,  $\tilde{p}_M$ ,  $\psi$  and the matrices  $M$  and  $N$ .  $\square$

*Remark 3.* From the proof we deduce the following:

- There is no need for full matrix inversion. The calculation of relevant minors suffices.
- For  $n$ -compartment models, the transfer function contains  $2n+1$  parameters. However, taking the structure into account, will usually lead to less parameters. As an illustration of this, consider the case with constant  $B$ . Then, as presented before in Example 1,  $p_N = p_M$ , and only  $n + 1$  parameters result.
- Identifiability can be easily checked from the matrices  $M$  and  $N$ . The parameter (combinations) in  $\tilde{p}_M$  and  $\tilde{p}_N$  are identifiable iff  $(\overline{N} \mid -\overline{M})$  (see Eqns 11–12) is a full row rank matrix, where  $\overline{N}$  and  $\overline{M}$  may contain zero rows to accomplish appropriate matrix dimensions.

From (18 – 20) and Proposition 2 for  $k = j^* + 1$  the predictor of the output ( $\hat{y}_{k+n}$ ), which can be based on the *unique* estimate of  $\theta$ , is given by

$$\hat{y}_{k+n} = \theta^T \phi_k - (b_N \overline{U} - b_M \overline{Y}) \quad (26)$$

with

$$b_N = \mathbf{0}$$

$$b_M = (-1)^{-j^*-1+l} \binom{j^*}{l-1}, \quad l = 1, \dots, j^*$$

#### 4.4 Sensitivity analysis results

As mentioned before, the data-independent sensitivity ( $\gamma_{\theta U}$ ) is a function of the given Hankel matrix and the coefficient matrices  $M$  and  $N$ . From the proof of Proposition 2, it follows that both  $\tilde{M}$  and  $\tilde{N}$  depend on the

sensor location  $j^*$ . If we fix the actuator location, as before at  $i^* = 1$ , then we are able to evaluate the polynomial quotient  $\tilde{N}/\tilde{M}$  only as a function of  $j^*$ . Since in this advection-reaction process the polynomials  $\tilde{M}$  and  $\tilde{N}$  do not depend on the number of compartments  $n$ ,  $j^*$  loses its interpretation of sensor location index. Hence, it reduces to a parameter that defines the model order. Table 1 presents the determinant and condition number of  $(\overline{N} \mid -\overline{M})$ , as a function of  $j^*$ . Notice that the number of parameters that have to be estimated from the filtered input-output data is equal to  $j^* + 1$ . On the basis of this, we suggest to take  $j^*$  small, preferably  $j^* = 1$ , because that will keep the estimation of the original physical parameters simple. However, an appropriate selection of  $j^*$  can only be done if the input-output data set is also taken into account.

Table 1. Determinant and condition number  $C$  of  $(\overline{N} \mid -\overline{M})$  as a function of  $j^*$ .

$j^*$	1	2	5	10	20
$\det(\overline{N} \mid -\overline{M})$	1	2	2500	$3.10^{16}$	$8.10^{73}$
$C(\overline{N} \mid -\overline{M})$	1	2	10	252	184756

## 5. DISCUSSION

As expected, for  $n$ -compartment models as derived from (24), the quality of the (yet deterministic) predictor (26) directly depends on  $n$ , which follows from the discretization grid in the spatial domain. Hence, there is an approximation of the solution of the original LTI infinite dimensional system caused by the finite difference scheme and characterized by  $n$ , which needs further numerical analysis.

Also, we may numerically approximate a given LTI infinite dimensional system in the spatial domain by other methods such as finite element or finite volume methods [11]. As shown by equation (6), the application of other numerical methods effects the entries of  $A$ , and thus  $M$  and  $N$ .

When dealing with the identification of infinite dimensional systems, it is a common approach to determine a minimal basis such that the estimation problem can be solved in a finite-dimensional space, *e.g.* [13, 14] for the application of Galerkin approximation schemes, [15] for rational approximations, [16] for collocation methods and [17] for minimal finite element approximations. Our approach, as illustrated to an advection-reaction system, is to handle the parameter estimation problem via a linear regressive realization of the approximate discrete-time system in order to obtain unique estimates.

## 6. CONCLUDING REMARKS

The proposed procedure allows the conservation of the underlying physical model structure in combination with linear regressive parameter estimation. The realization of a linear regressive system  $\tilde{\Sigma}_d$  from a state space system  $\Sigma_d$  is based on linearity of the system and linearity in the parameter  $\vartheta$  of  $A(\vartheta)$  and  $B(\vartheta)$ . An *a priori* sensitivity analysis is performed on the basis of the regression weighting matrices  $M$  and  $N$ .

In practice, however, there will be measurement noise deteriorating both the inputs and outputs of the linear regressive system  $\tilde{\Sigma}_d$ . This will naturally lead to a (mixed) errors-in-variables problem and thus the need for *e.g.* total least-squares solutions, which is a subject for further research.

[17] D. Coca and S.A. Billings, 2002. Identification of finite dimensional models of infinite dimensional dynamical systems, *Automatica*, 38, 1851–1865.

## REFERENCES

- [1] L. Ljung, 1987. *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall.
- [2] J. P. Norton, 1986. *An Introduction to Identification*. Academic Press.
- [3] D. Vries, 2008. *Estimation and Prediction of Convection-Diffusion-Reaction Systems from Point Measurements*. PhD Thesis, Wageningen University, pp. 170.
- [4] T.G. Doeswijk and K.J. Keesman, 2009. Linear Parameter Estimation of Rational Biokinetic Functions, *Water Research*, 43(1), 107–116.
- [5] K.J. Keesman and T.G. Doeswijk, 2009. Direct Least-squares Estimation and Prediction of Rational Systems: Application to Food Storage, *J. Process Control*, 19, 340–348.
- [6] M. Verhaegen and P. Dewilde, 1992. Subspace model identification. Part 1: The output-error state-space model identification class of algorithms. *Int. J. Control*, 56, 1187–1210.
- [7] P. van Overschee, Choice of State-Space Basis in Combined Deterministic-Stochastic Subspace Identification, *Automatica*, 1995, vol 31:12, pp 1877–1883.
- [8] G.Y. Hu and R.F. O’Connell, 1996. Analytical inversion of symmetric tridiagonal matrices, *J. Phys. A: Math. Gen.*, 29, 1511–1513.
- [9] Y. Huang and W.F. McColl, 1997. Analytical inversion of general tridiagonal matrices, *J. Phys. A: Math. Gen.*, 30, 7919–7933.
- [10] C.M. de Fonseca and J. Petronilho, 2001. Explicit inverses of some tridiagonal matrices, *Linear Algebra and its Applications*, 325, 7–21.
- [11] Q. Fang, T. Tsuchiya and T. Yamamoto, 2002. Finite difference, finite element and finite volume methods applied to two-point boundary value problems. *J. Comput. Appl. Math.*, 139, 9–19.
- [12] R. Pintelon, P. Guillaume, G. Vandersteen and Y. Rolain, 1998. Analysis, development and applications of TLS algorithms in frequency domain system identification, *SIAM J. Matrix Anal. Appl.*, 19(4), 983–1004.
- [13] H.T. Banks and K. Kunisch, 1989. *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston.
- [14] J. Baumeister, W. Scondo, M.A. Demetriou and I.G. Rosen, 1997. On line parameter estimation for infinite-dimensional dynamical systems, *SIAM J. Control Optim.*, 35(2), 678–713.
- [15] R. Pintelon, 1998. Identification of Linear Time Invariant Diffusion Phenomena. *IEEE Transactions on instrumentation and measurement*, 47(5), 1053–1055.
- [16] T. Söderström and B. Bhikkaji, 2000. Reduced Order Models for Diffusion Systems via Collocation Methods, in *IFAC System Identification*, Santa Barbara, California, June 21-23.