Proceedings 5th MATHMOD Vienna, February 2006 (I.Troch, F.Breitenecker, eds.)

LINEAR REGRESSIVE MODEL STRUCTURES FOR ESTIMATION AND PREDICTION OF COMPARTMENTAL DIFFUSIVE SYSTEMS

Corresponding author: D. Vries
*Dept. of Agrotechnology and Food Sciences, Systems and Control Group,
Wageningen University, The Netherlands
Phone: ++31 317 482 014, Fax: ++31 317 484 957
†Dept. of Applied Mathematics, Faculty of EEMS, University of Twente
The Netherlands

D. Vries*, K.J. Keesman*, H. Zwart†

Abstract. In input-output relations of (compartmental) diffusive systems, physical parameters appear non-linearly, resulting in the use of (constrained) non-linear parameter estimation techniques with its short-comings regarding global optimality and computational effort. Given a LTI system in state space form, we propose an approach to get a linear regressive model structure and output predictor, both in algebraic form. We deduce the linear regressive model from a particular LTI state space system without the need of direct matrix inversion. As an example, two cases are discussed, each one a diffusion process which is approximated by a state space discrete time model with n compartments in the spatial plane. After a sequence of steps the system output can then be explicitly predicted by $\hat{y}_k = \hat{\theta}^T \phi_{k-n} - \check{\gamma}_{k-n}$ as a function of n, sensor and actuator position, the parameter vector θ, and input-output data. This method is attractive for estimation insight in experimental design issues, when physical knowledge in the model structure is to be preserved.

1 Introduction

Identification is about matching selected models to observations and is typically restricted to input-output data in finite time. For LTI systems, it is possible to estimate a (parametric) transfer function model with appropriate model order without using prior system's knowledge. However, it is not immediately clear how the structure of the transfer function model is linked to the underlying physical and/or (bio)chemical processes. Contrary to this so-called black box modelling, grey or white box modelling provides a model structure that is more suited for physical interpretation of *e.g.* (optimal) control solutions and experimental design issues like sensor and/or actuator placement. Therefore we would like to be able to estimate model parameters and states, *while preserving physical meaning*.

However, physical relevant parameters often appear non-linearly in an input-output model structure. Consequently, in general, non-linear optimization solvers with (costly) iterative procedures are used, where it frequently occurs that the parameter search gets stuck in local minima. The problem gets particularly more complex when dealing with infinite dimensional systems. In that case, it is a common approach to determine a minimal basis in order to solve the estimation problem in a finite-dimensional space. See *e.g*: Galerkin approximation schemes [1, 2], rational approximations [8], collocation methods [9], minimal finite element approximations [3] and subspace identification techniques [7]. Interestingly, the latter technique has the advantage over classical prediction error techniques by the absence of non-linear parametric optimizations.

Our approach is to handle the parameter estimation and prediction problem, initially for an LTI infinite-dimensional system, via discretization and a linear regressive parametric realization of the system in order to obtain unique estimates. However, unlike a 'data based' approach such as subspace identification, we will conserve the physical compartmental model structure due to the reasons mentioned before.

In particular, we will consider a set of finite LTI state space systems which can be regarded as compartmental diffusive systems. By the properties of such a system (denoted as Σ^d), we are able to find another realization of Σ^d which is suited for linear regressive estimation and prediction. After some integral transform of Σ^d, we obtain a set of linear equations of the form $\varphi^T M \psi = b$. It will be shown, that the inverse of M is the resolvent of the system matrix A in Σ^d. In the specific case that A is a symmetric tridiagonal matrix, explicit solutions for the inverse of M are known, see *e.g.* [5, 6]. The key here is to find M^{-1}, such that we may rewrite this as a linear regressive set of equations: $\theta^T \phi = \gamma$, with $\theta = \xi(\vartheta)$ a known reparametrization function of the physical parameter ϑ. From hereon, it is rather straightforward to arrive at an estimate $\hat{\theta}$ using existing estimation techniques. By a simple rearrangement of terms we get an explicit expression for the output at time instant k, *i.e.* $\hat{y}(k|\hat{\theta}; Z^-)$ with Z^- the current available input-output data set.

Summarizing, the key objectives of the paper are as follows:

(i) to show the procedure to obtain a linear regressive *mechanistic* (or also referred to as physical) model representation, say $\Sigma^{\text{LR}}(\phi, \gamma)$ from a linear state space model $\Sigma(A, B, C, D)$

(ii) to show the application of approach (i) to

Case A a (finite) compartmental approximation of a boundary control system with Dirichlet boundary conditions, and,

Case B a (finite) compartmental approximation of a boundary control system with one Neumann boundary condition

(iii) to show some estimation results of one case, that is, Case A.

The paper is organized as outlined above, that is, we will show how to fulfill objective (i) in section 2, objective (ii) in section 3 and some simulation results (iii) of Case A in section 4.

2 From state space to linear regressive form

In particular, we consider discrete time SISO LTI state space systems,

$$\Sigma^d(A,B,C,D) : \begin{cases} x(k+1) &= A(\vartheta)x(k) + B(\vartheta)u(k) \\ y(k) &= C(\vartheta)x(k) + D(\vartheta)u(k) \end{cases} \tag{1}$$

with $x \in \mathbb{R}^n = \mathcal{X}$ the state in state space, $y \in \mathbb{R}^m = Y$ the output in observation space, $u \in \mathbb{R}^p = U_{\text{adm}}$ the input variable in the admissible input set and A, B, C and D the matrices mapping inputs and states into state space. We assume some properties on A so that Σ^d can be regarded as a compartmental system with unknown (lumped) physical parameters ϑ. Furthermore, we assume there is no feed-through term, hence $D = 0$. Nevertheless, it is easy to extend the results presented here for the case $D \neq 0$. Note that in the sequel we will have slight abuse of notation by preserving D for another variable.

The key here is to find the linear regressive equivalent form of Σ^d by means of defining the parameter (vector) function $\xi : \vartheta \longmapsto \theta$, with $\theta \in \Theta_{\text{adm}}$ newly defined parameters in the admissible parameter space. The parameters θ can then be estimated by using existing linear estimation techniques. Further, by rearranging terms we also get an explicit expression for the output predictor at time instant k.

The case that the linear regressive form $\Sigma^{\text{LR}_d}(\phi,\gamma)$ is found from $\Sigma^d(A,B,C)$ has been proved under the condition that A and B are linear in ϑ. To arrive at this, the following sequence of steps is needed.

(1) Denote $M = qI - A$. Splitting the rational transfer function G of $\Sigma^d(A,B,C)$ in a numerator N and denominator part D, gives,

$$N(\vartheta,q)u_k = D(\vartheta,q)y_k \tag{2}$$

with N and D polynomials in ϑ and q. For instance, Pintelon and many others [8], split N and D in functions of the polynomial variable q (or the Laplace variable s) and the parameter vector θ, so that $N(\vartheta)n(q)u_k = B(\vartheta)d(q)y_k$ is considered. It is common to treat the entries of the vectors $N(\vartheta)$ and $D(\vartheta)$ as black-box parameters for further estimation and prediction. As a consequence, the link to the underlying 'white' parameters ϑ will soon be lost as the polynomial order of N and/or D increases. However, we will try to prevent this loss by decomposing N and D not only in a shift operator dependent, but also in a (physical) parameter dependent part so to obtain polynomial coefficient matrices R_N and R_D, as,

$$N = \varphi_N^T R_N \psi(q), \quad D = \varphi_D^T R_D \psi(q) \tag{3}$$

with $\varphi_{\boldsymbol{\cdot}}$ a vector containing (possibly lumped combinations of) the original parameters ϑ.The transfer function can then be written as a rational function of polynomials in ϑ and q,

$$G(\vartheta,q) = CM^{-1}B = \frac{N(\vartheta,q)}{D(\vartheta,q)} = \frac{\varphi_N^T(\vartheta)R_N\psi(q)}{\varphi_D^T(\vartheta)R_D\psi(q)} \tag{4}$$

where R_N and R_D become coefficient matrices which link the polynomial variables q and ϑ. This is in contrast to *e.g.* [8] and work by others, where the product $\varphi^T R_{(\boldsymbol{\cdot})}$ is considered as the (black box) parameter vector. In the sequel, q will be stacked by polynomial degree order in the vector ψ.

(2) The final step in the linear regressive reparametrization is to rewrite (2) to the form $\theta^T \phi_k = \gamma_k$. We write the polynomial coefficient matrices as,

$$R_N = \begin{pmatrix} R_{N_1} \\ R_{N_2} \end{pmatrix} \qquad R_D = \begin{pmatrix} R_{D_1} \\ R_{D_2} \end{pmatrix} \tag{5}$$

Furthermore, define

$$Z_k = \begin{pmatrix} U_k & Y_k \end{pmatrix}^T = \begin{pmatrix} \bar{U} & u_{k+n} & | & \bar{Y} & y_{k+n} \end{pmatrix}^T \tag{6}$$

where,

$$\bar{U}_k = \begin{pmatrix} u_k & u_{k+1} \cdots u_{k+n-1} \end{pmatrix}^T \qquad \bar{Y}_k = \begin{pmatrix} y_k & y_{k+1} \cdots y_{k+n-1} \end{pmatrix}^T$$

Rewrite (2) to,

$$\varphi^T \begin{pmatrix} R_N & -R_D \end{pmatrix} Z_k = \varphi^T \begin{pmatrix} R_{N_1} & -R_{D_1} \\ R_{N_2} & -R_{D_2} \end{pmatrix} Z_k = 0 \tag{7}$$

Now define,

$$\phi_k = \begin{pmatrix} R_{N_1} & | -R_{D_1} \end{pmatrix} Z_k \tag{8}$$

$$\gamma_k = \begin{pmatrix} -R_{N_2} & | \, R_{D_2} \end{pmatrix} Z_k \tag{9}$$

Observe that we can rewrite (7) to a linear regressive prediction, using the definitions (8) and (9) if we can write the parameter vector as a vector with unknown (to be estimated) part θ and known constant part, *i.e.* $\varphi = (\theta^T \;\; \text{constant})^T$. We get

$$y_{k+n} = \theta^T \phi_k - \check{\gamma}_k \tag{10}$$

where,

$$\check{\gamma}_k = c_D^{-1} \begin{pmatrix} -R_{N_2} & b_D \end{pmatrix} \check{Z}_k, \tag{11}$$

with $\check{Z}_k = \begin{pmatrix} \bar{U} & u_{k+n} \mid \bar{Y} \end{pmatrix}^T$ and $R_{D_2} = (b_D \mid c_D)$. After multiplication of (7) by q^{-n} (*i.e.* a backward time shift is applied), we may write (8)–(11) as our wanted linear regressive system:

$$\tilde{\Sigma}_d^{\text{LR}} : \begin{cases} \theta^T \phi_{k-n} & = \gamma_{k-n} \\ y_k & = \theta^T \phi_{k-n} - \check{\gamma}_{k-n} \end{cases} \tag{12}$$

Notice that we have obtained a linear function in θ and $\check{Z}_{k-n}$, see (8) and (11). This leads to the following proposition.

Proposition 2.1. *Given system Σ^d as in (1). Then,*

(i) exact explicit expressions of $N(\vartheta, q)$ and $D(\vartheta, q)$ as a function of n exist.

(ii) Σ^d (1) can be written in the form of $\tilde{\Sigma}_d^{LR}$ as in (12) with $\theta_i = \xi_i(\vartheta)$, a polynomial function.

Proof.

(i) Let $M = qI - A$ and the determinant of M be denoted by $\mathbf{M} = \det(M)$. Let $M_{i'j'}$ and $A_{i'j'}$ denote the submatrix of M and A respectively, both resulting from the deletion of row i and j. By Laplace expansion, the determinant of M is given by

$$\mathbf{M} = \sum_{j=1}^{n} (-1)^{i+j} m_{ij} \mathbf{M_{i'j'}}$$

$$= \sum_{i=1}^{n} (-1)^{i+j} m_{ij} \mathbf{M_{i'j'}}$$

$\forall i \le n,\ j \le n$ with $\mathbf{M_{i'j'}} = \det(M_{i'j'})$ assumed to be known. For any choice of row or column either expansion yields the determinant. The classical adjoint of M is defined by the transposed matrix of co-factors denoted by $\operatorname{adj} M = (p_{ij})$, with entries $p_{ij} = \sum_{i=1}^{n} (-1)^{i+j} \mathbf{M_{i'j'}}$. The inverse of M is $M^{-1} = \operatorname{adj} M / \mathbf{M}$, iff M non-singular. Because A is linear in ϑ, $M_{i'j'} = qI_{i'j'} - \bar{A}_{i'j'} - \tilde{A}_{i'j'}\vartheta$. From induction it follows that the determinant of M will be a polynomial in q and ϑ with maximal order n and the adjoint of M a matrix filled with polynomials in q and ϑ with maximal order $n-1$. Since $G = C(qI - A)^{-1}B = CM^{-1}B$ and B linear in ϑ, G becomes a rational function of polynomials in q and ϑ. It directly follows that $N(\vartheta, q)$ and $D(\vartheta, q)$ can be decomposed as in (3), with polynomial coefficient matrices $R_N(\vartheta, q)$ and $R_D(\vartheta, q)$ being functions on n.

(ii) Given $G(\vartheta, q) = \frac{N(\vartheta,q)}{D(\vartheta,q)}$, the transfer function of Σ_d. Then, $D(\vartheta, q)y_k = N(\vartheta, q)u_k$ and via direct algebra and the proof of part (i) one readily obtains $\theta^T \phi_k = \gamma_k$ with $\theta_i = \xi_i(\vartheta)$, a polynomial function. Notice that γ_k contains y_{k+n} and we can write the equivalent form $\tilde{\Sigma}_d$ (12) after multiplication with q^{-n}.

□

Remark 2.1. *From the proof of Proposition 2.1(i) it follows directly that,*

(a) The polynomial degree of $N(\vartheta, q)$ in q is determined by C and the adjoint of M. Let the entry c_r be non-zero, with $1 \leq r \leq n$. Then the polynomial in q is of maximum degree $n-r$.

(b) The polynomial degree of $N(\vartheta, q)$ in ϑ is determined by $B(\vartheta)$, C and the adjoint of M. Recall that $B(\vartheta)$ is linear in ϑ. Again, let the entry c_r be non-zero, with $1 \leq r \leq n$. Then the maximum degree of polynomial in ϑ is $n-r+1$.

(c) The last column of R_{N_1} is zero and $R_{N_2} = 0$, due to the maximum degree $n - 1$ of the polynomial $\operatorname{adj} M$ in ϑ and q. More generally, if the maximum degree of polynomials is $n - r$, $1 \leq r \leq n$, then the last r rows of R_{N_1} are 0 and the lower r rows of R_N is filled with zeros. This result is caused by causality of Σ^d.

Some further remarks regarding identifiability and predictability are given in [11].

In the following we will inspect model structures which exhibit a tridiagonal Toeplitz banded system matrix A. Such systems are interesting by their system matrix which is symmetric on the diagonal. Hence, the resulting matrix $M = qI - A$ is also tridiagonal and persymmetric. When equipped with explicit expressions for M^{-1}, it will be straightforward to get a linear regressive realization as in (12) via the mentioned steps. This is shown in the next section.

3 Application to diffusive compartmental models

Let us first examine two particular three-banded Toeplitz matrices. These matrices will be of special interest for our example cases A and B. Recall that the resolvent matrix $M = qI - A$ may originate from a discrete time LTI state space model Σ^d and the inverse of M is valuable for a linear regressive representation of Σ^d.

3.1 Tridiagonal matrices

Consider the nonsingular tridiagonal matrices $M_{\text{A}} \in \mathbb{R}^{n\times n}$ and $M_{\text{B}} \in \mathbb{R}^{n\times n}$. Note that we slightly abuse our notation by the subscripts A and B which do *not* correspond to the system matrix A and input mapping matrix B, but (may) correspond to our example cases A and B. We define,

$$M_{\text{A}} = \begin{pmatrix} a & b & 0 & \cdot & 0 \\ b & a & b & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & b & a & b \\ 0 & \cdot & 0 & b & a \end{pmatrix} \qquad M_{\text{B}} = \begin{pmatrix} c & b & 0 & \cdot & 0 \\ b & a & b & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & b & a & b \\ 0 & \cdot & 0 & b & a \end{pmatrix}, \tag{13}$$

where a, b and c are real scalars and assumed constant and $c \neq a \neq 0 \neq b$. We will study the inverse of M_{A} and M_{B}. First, write the inverse of both tridiagonal matrices as a rational function:

$$M_{\text{A}}^{-1} = \frac{V}{W} \qquad M_{\text{B}}^{-1} = \frac{\tilde{V}}{\tilde{W}}$$

Denote furthermore the dimensions in the superscript and the determinant in boldface, *i.e.* $\mathbf{M}^n = \det M^{n\times n}$.

By Laplace expansion of the determinants and (classical) adjoints of $M_{(\cdot)}$ we can easily see that the inverse of M_{A} and M_{B} are characterized by:

$$M_{\text{A}}^{-1} = \frac{1}{\mathbf{M}_{\text{A}}} \begin{pmatrix} v_{11} & \cdots & v_{1j} & \cdots & v_{1n} \\ \vdots & \cdot & \cdot & \cdot & \vdots \\ v_{i1} & \cdot & v_{ij} & \cdot & v_{in} \\ \vdots & \cdot & \cdot & \cdot & \vdots \\ v_{n1} & \cdots & v_{nj} & \cdots & v_{nn} \end{pmatrix}, \quad M_{\text{B}}^{-1} = \frac{1}{\mathbf{M}_{\text{A}}} \begin{pmatrix} v_{11} & v_{12} & v_{1j} & \cdots & v_{1n} \\ v_{21} & cv_{22} & cv_{2j} & \cdots & cv_{2n} \\ v_{i1} & cv_{i2} & * & * & * \\ \vdots & \vdots & * & * & * \\ v_{n1} & cv_{n2} & * & * & * \end{pmatrix} \tag{14}$$

where the entries with $*$ denote some polynomial in a, b and c. When writing the determinant of M as $\mathbf{M}^n := \det M^{n\times n}$, we get,

$$\mathbf{M}^{\mathbf{n}}_{\mathrm{B}} = c\mathbf{M}^{\mathbf{n-1}}_{\mathrm{A}} - b^2\mathbf{M}^{\mathbf{n-2}}_{\mathrm{A}}$$

Hence the denominator $\tilde{W}$ of the inverse of M_{B}, has been characterized in terms of W. By close inspection, it follows that we may decompose the denominator and nominator of M_{A}^{-1} and M_{B}^{-1} in a parameter and time shifted part. That is, (we omit the arguments of φ and ψ for eligiblity),

$$M_{\mathrm{A}}^{-1} = \frac{\varphi^T_{N_{\mathrm{A}}} R_{N_{\mathrm{A}}}\psi}{\varphi^T_{D_{\mathrm{A}}} R_{D_{\mathrm{A}}}\psi} = \frac{\varphi^T_{\mathrm{A}} R_{N_{\mathrm{A}}}\psi}{\varphi^T_{\mathrm{A}} R_{D_{\mathrm{A}}}\psi} \tag{15}$$

$$M_{\mathrm{B}}^{-1} = \frac{\varphi^T_{N_{\mathrm{B}}} R_{N_{\mathrm{B}}}(i^*,j^*)\psi}{\varphi^T_{D_{\mathrm{B}}} R_{D_{\mathrm{B}}}\psi} = \frac{\varphi^T_{\mathrm{B}} R_{N_{\mathrm{B}}}(i^*,j^*)\psi}{\varphi^T_{\mathrm{B}} R_{D_{\mathrm{B}}}\psi} \tag{16}$$

Notice that for the inverse of M, a compartmental model with as in (2) is described by $2n$ parameters, whereas it suffices to use n parameters, since it will follow that

$$\begin{aligned}\varphi_{\mathrm{A}} &= \varphi_{N_{\mathrm{A}}} = \varphi_{D_{\mathrm{A}}} \\ &= \begin{pmatrix} \theta_{\mathrm{A}} \\ 1 \end{pmatrix},\end{aligned} \tag{17}$$

with θ_{A} consisting of lumped combinations of the original parameters a and b.

Let a fixed entry at some row $i = i^*$ and some column $j = j^*$ of $R_{N_{(\cdot)}}$ be *a priori* defined by the input mapping matrix B and observation matrix C respectively, where $b_{i^*} = 1$ and $b_i = 0$ for $i \neq i^*$ and $c_{j^*} = 1$ and $c_j = 0$ for $j \neq j^*$, with $i, j \in \{1 \dots n\}$ and $i^*, j^* \in \{1 \dots n\}$.

For M_{B}, we will get fixed coefficient matrices $R_{D_{\mathrm{B}}}$ and $R_{N_{\mathrm{B}}}(i^*, j^*)$ for a fixed i^* and j^* when we set,

$$\begin{aligned}\varphi_{\mathrm{B}} &= \varphi_{N_{\mathrm{B}}} = \varphi_{D_{\mathrm{B}}} \\ &= \begin{pmatrix} \theta_{\mathrm{B}} \\ 1 \end{pmatrix},\end{aligned} \tag{18}$$

with,

$$\theta_{\mathrm{B}} = \begin{pmatrix} \frac{c}{b}\varphi_{\mathrm{A}} \\ \hline \theta_{\mathrm{A}} \end{pmatrix} = \begin{pmatrix} \frac{c}{b}\theta_{\mathrm{A}} \\ \frac{c}{b} \\ \hline \theta_{\mathrm{A}} \end{pmatrix} \tag{19}$$

where $\theta_{(\cdot)} \in \mathbb{R}^n$.

Notice that here a similar result for the parameter containing vector of M_{B}^{-1}, $\varphi_{\mathrm{B}} = (\theta_{\mathrm{B}} \;\; 1)^T$ is found.

3.2 Coefficient matrices for parametric estimations

Suppose we are interested in writing the numerator and denominator of $M_{(\cdot)}$ as a coefficient-vectorproduct, e.g. $z := \varphi^T R\psi$, where φ and ψ are column vectors containing combinations of the parameters of a and b respectively and R a coefficient matrix containing integer values. We arrive at the following proposition.

Proposition 3.1. *Given the symmetric 3-banded Toeplitz matrix $M_A(\vartheta, q)$ as in (15) and (17), then the entries for the coefficient matrices R_{D_A} and R_{N_A} of M_A are given by*

$$r^D_{ij} = \begin{cases} (-1)^{i+j-2}\binom{n+i}{n-i+1}\binom{i-1}{j-1} & \textit{if} \quad i \geq j \\ 0 & \textit{elsewhere} \end{cases}$$

$$r^N_{ij} = \begin{cases} \displaystyle\sum_{s=0}^{\min(i,j)-1} (-1)^{i+j-2}\binom{n-l_s+i}{n-l_s-i+1}\binom{i-1}{j-1} & \textit{if} \quad i \geq j \;\wedge\; i \leq n - l_s + 1 \\ 0 & \textit{elsewhere} \end{cases}$$

where,

$$\imath = \begin{cases} i^* & \text{if} \quad i^* \leq \frac{n}{2} \\ n - i^* + 1 & \text{elsewhere} \end{cases}$$

$$\jmath = \begin{cases} j^* & \text{if} \quad j^* \leq \frac{n}{2} \\ n - j^* + 1 & \text{elsewhere} \end{cases}$$

$$l_s = l^* + 2s + 1 \quad \text{with} \qquad l^* = \begin{cases} |i^* - j^*| & \text{if} \quad i^* < \frac{n}{2} \\ |j^* - i^*| & \text{elsewhere} \end{cases}$$

We have omitted the indication A in the superscript of r for reasons of eligibility. Denote again a shorthand notation to indicate the dimensions of the square coefficient matrix $R_{(.)}$ by using the superscript, *i.e.* $R^{(n+1)} = (n+1) \times (n+1)$. Then, we can easily write the numerator and denominator coefficient matrices in proposition as a summation of co-factors, or, since M is symmetric on the tridiagonals, in a summation of determinants of M with lower dimension. Indeed, we get,

Proposition 3.2. *Given the symmetric 3-banded Toeplitz matrix $M_A(\vartheta, q)$ as in (15) and (17), then the coefficient matrices of M_A are given by lower triangular matrices R_{D_A} and R_{N_A},*

$$R_{D_A}^{n+1} = \begin{pmatrix} \binom{n+1}{n} & & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ (-1)^{i-1}\binom{n+i}{n-i+1}\binom{i-1}{0} & & (-1)^{i+j-2}\binom{n+i}{n-i+1}\binom{i-1}{j-1} & & 0 \\ \vdots & & & \ddots & \\ (-1)^{n-1}\binom{2n}{1}\binom{n-1}{0} & \cdots & (-1)^{n+j-2}\binom{2n}{1}\binom{n-1}{j-1} & \cdots & (-1)^{2n-2}\binom{2n}{1}\binom{n-1}{n-1} \end{pmatrix}$$

$$R_{N_A}^{n+1}(i^*, j^*) = \sum_{s=0}^{\min(\imath,\jmath)-1} \bar{R}_{D_A}^{n-l_s}$$

where,

$$\bar{R}_{D_A}^{k_1} = \begin{pmatrix} \mathbf{0}^{(1\times k_2)} & \mathbf{0}^{(1\times(k_1-k_2))} \\ R_{D_A}^{k_2} & \mathbf{0}^{(k_2\times(k_1-k_2))} \\ \mathbf{0}^{((k_1-k_2)\times k_2)} & \mathbf{0}^{((k_1-k_2)\times(k_1-k_2))} \end{pmatrix} \quad \text{with } k_1 = (n - l_s) \in \{1, \cdots, n+1\}, \quad 1 \leq k_2 < k_1$$

$$\imath = \begin{cases} i^* & \text{if} \quad i^* \leq \frac{n}{2} \\ n - i^* + 1 & \text{elsewhere} \end{cases}$$

$$\jmath = \begin{cases} j^* & \text{if} \quad j^* \leq \frac{n}{2} \\ n - j^* + 1 & \text{elsewhere} \end{cases}$$

$$l_s = l^* + 2s + 1 \quad \text{with} \qquad l^* = \begin{cases} |i^* - j^*| & \text{if} \quad i^* < \frac{n}{2} \\ |j^* - i^*| & \text{elsewhere} \end{cases}$$

For the inverse of M_B, we propose the following:

Proposition 3.3. *Given the symmetric 3-banded Toeplitz matrix $M_B(\vartheta, q)$ as in (16) and (18)–(19), then the entries for the coefficient matrices R_{D_B} and R_{N_B} are given by,*

$$R_{D_B}^{n+1} = \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times(n-1)} \\ -R_{D_A}^{n} & \mathbf{0}^{n\times 1} \\ \hline \multicolumn{2}{c}{\begin{pmatrix} 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & R_{D_A}^{n} \end{pmatrix} - \begin{pmatrix} R_{D_A}^{n-1} & \mathbf{0}^{(n-1)\times 2} \\ \mathbf{0}^{2\times(n-1)} & \mathbf{0}^{2\times 2} \end{pmatrix}} \end{array} \right)$$

$$R^{n+1}_{N_B}(i^*,j^*) = \begin{cases} \left(\begin{array}{c} 0 \\ \hline -R^{n+1}_{N_A}(1,1) \end{array}\right) & \textit{if } i^*=1 \vee j^*=1 \\ \left(\begin{array}{cc} 0 & \mathbf{0}^{1\times n} \\ R^{n}_{N_A}(1,j^*{-}1) & \mathbf{0}^{n\times 1} \\ \hline 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & -R^{n}_{N_A}(i^*{-}1,j^*-1) \end{array}\right) & \textit{if } i^*=2 \wedge j^*\geq 2 \\ \left(\begin{array}{c} \begin{array}{cc} 0 & \mathbf{0}^{1\times n} \\ R^{n}_{N_A}(i^*{-}1,j^*-1) & \mathbf{0}^{n\times 1} \end{array} \\ \hline \begin{pmatrix} 0 & \mathbf{0}^{1\times n} \\ \mathbf{0}^{n\times 1} & -R^{n}_{N_A}(i^*-1,j^*-1) \end{pmatrix} + \begin{pmatrix} \mathbf{0}^{2\times(n-1)} & \mathbf{0}^{2\times 2} \\ R^{n-1}_{N_A}(i^*{-}2,j^*{-}2) & \mathbf{0}^{2\times(n-1)} \end{pmatrix} \end{array}\right) & \textit{if } i^*>2 \vee j^*\geq 2 \end{cases}$$

where the coefficient matrices R_{X_B} are divided in blocks which correspond to the blocks in φ_B, see (18).

4 Compartmental modelling for diffusion processes

4.1 Case A: Boundary control system with Dirichlet conditions

Consider an infinite dimensional system Σ^e_{A} of parabolic type on $[0,\infty)\times[0,\infty)$, see [4].

$$\Sigma^e_{\text{A}} : \begin{cases} \frac{\partial w}{\partial t}(x,t) & = \alpha^2 \frac{\partial^2 w}{\partial x^2}(x,t), \quad w(x,0) = w_0(x), \\ w(0,t) & = u(t) \\ y(t) & = w(x^*,t) + \beta \end{cases} \tag{20}$$

where $w_0(x) \in L_2(0,\infty)$, $x^* \in [0,\infty)$ and $U = \mathbb{R}$.

The solution to this problem when applying a step input $u(t) = 1_{[0,\infty)}(t)$ is well-known and is given by the output $w(x^*,t) = \text{erfc}\left(x^*/(2\alpha\sqrt{t})\right) 1_{[0,\infty)}(t)$, where α^2 can be interpreted as the diffusion constant and β a constant feed-through term. With this system some estimations of α is worked out for different feed-through terms and an added Gaussian white noise sequence $e(t_k)$.

By finite differences approximation, we obtain a discrete LTI state space system, with the resolvent of A as M^{-1}: $a = (q-1)/\vartheta + 2$, $b = 1$ and φ as given in (17). Further, we let $B = (\vartheta \;\; 0 \;\; \cdots \;\; 0)^T$. The observation $y(t) = w(t,x^*)$ is approximated by Cw^d_k, with C mapping a 'point' observation at the j-th compartment (*i.e* $c_{j^*}=1$, $j^* \in j$). Notice that a grid with n points directly leads to n states, because we have started with one state variable and one spatial direction in the PDE model Σ^e.

The reader is referred to an explicit solution of the linear regressive predictor in [11]. Because we are dealing with a single parameter, the polynomial functions ξ_i become $\xi_i(\vartheta)=\vartheta^{n-i+1}$, $i = 1,2,\cdots,n$, so that $\theta^T = \left(\vartheta^n \quad \vartheta^{n-1} \quad \cdots \quad \vartheta\right)$.

4.2 Case B: Boundary control system with a Neumann condition

We can approximate the following distributed parameter system,

$$\Sigma^e_{\text{B}} : \begin{cases} \frac{\partial w}{\partial t}(x,t) & = \alpha^2 \frac{\partial^2 w}{\partial x^2}(x,t), \quad w(x,0) = w_0(x), \\ \frac{\partial}{\partial x} w(0,t) & = -\frac{1}{\kappa}\left(w(0,t) - u(t)\right) \\ w(\infty,t) & = 0 \\ y(t) & = w(x^*,t) \end{cases} \tag{21}$$

where $w_0(x) \in L_2(0,\infty)$, $x^* \in [0,\infty)$ and $U = \mathbb{R}$, by finite differences to obtain the resolvent of A, $qI - A = M_{\text{B}}$. In M_{B} we encounter: $a = (q-1)/\vartheta + 2$, with the lumped parameter $\vartheta = \Delta\alpha^2$ consisting of the discretization parameter Δ and a diffusion coefficient α^2. Further, $b = 1$, $c = 1 + (q-1)/\vartheta$ and the input mapping matrix B similar as in [10] and properly defined. Further, φ is properly defined, *i.e.* it is similar as in (18) and scaled with $1/\kappa$ due to the definition of the boundary control mapping B.

In addition to case A, we have a second parameter κ, which corresponds to a specific heat capacity divided by a thermal conductivity. This case is merely to show that the inverse of M_{B} corresponds to a physical example. Case B is subject of further (simulation) study.

5 Simulation results

The above system is approximated with finite differences, yielding a compartmental model and where β is assumed unknown and not modelled. Keeping the sensor and input position as degrees of freedom, we have found explicit expressions for the coefficient matrices R_N and R_D of (4). Consequently, also explicit functions for $y(k|\theta;Z^-)$ as a function of the number of compartments n, sensor position j^* and the input compartment, i^*, are found.

In the simulation study, we fix $i^* = j^* = 1$ for our B and C matrix. Furthermore, we estimate our original parameter α with (i) non-linear least squares estimation from the disturbed classical solution (erfc-model) and (ii) with (linear) least squares modelled by Σ^{LR_A} with added disturbances. In the latter case, we approximate α as $\hat{\alpha}_{\mathrm{LR}} \approx |\frac{1}{n}\sum_{i=1}^{n}\hat{\vartheta}_i^{\frac{1}{2(n-i)}}|$. The system is approximated with 4 compartments, $\alpha = 0.2$, and bias β is added and has a range from -20% to 20% of the applied input step ($u_{\min} = 0$, $u_{\max} = 1$). It is shown in table 1, that the estimation of α in the linear regressive compartmental model is very insensitive to bias, but very sensitive to added Gaussian white noise, when the system is driven by a step input. Interestingly, table 2 illustrates that a pseudo random binary (PRBS) input sequence improves our linear regression results for the noiseless case. Since we do not know the classical solution with such an input sequence, we do not show non-linear least square estimation results.

Table 1: Least squares (LS) estimations of α, with different bias (column-wise) and measurement noise variance, σ_e^2, when a step input is applied.

	Non-linear LS estimation $\hat{\alpha}$, erfc-model					Linear regressive LS estimation $\hat{\alpha}_{\mathrm{LR}}$, Σ_A^{LR}				
	β=-0.2	β=-0.1	β=0	β=0.1	β=0.2	β=-0.2	β=-0.1	β=0	β=0.1	β=0.2
$\sigma_e = 0$	0.13	0.16	0.20	0.26	0.34	0.15	0.14	0.20	0.14	0.14
$\sigma_e = 0.01$	0.13	0.16	0.20	0.26	0.34	0.62	0.63	0.64	0.68	0.64

Table 2: Estimations of α with linear regression from Σ_A^{LR}, with different bias (column-wise) and measurement noise variance, σ_e^2, when a a PRBS is applied, with probability $p = 0.5$.

	β=-0.2	β=-0.1	β=0	β=0.1	β=0.2
$\sigma_e = 0$	0.20	0.20	0.20	0.20	0.20
$\sigma_e = 0.01$	0.27	0.24	0.27	0.27	0.27

We observe the opposite when non-linear least squares is used for estimating α in a erfc-model, that is, it is sensitive to bias and merely sensitive to added white noise. Iterative pre-filtering (see *e.g.* [10]) or more advanced least squares techniques may compensate the colored disturbances in the linear regressive model structure and applying these techniques is subject of further study.

6 Conclusions

Summarizing, the sketched methodology is attractive for linear regressive estimation and prediction. Furthermore, since we keep track of the physical parameters, physical knowledge in the model structure is preserved throughout analysis and estimation. For two compartmental diffusive models we have derived linear regressive structures. For one case, we illustrate the strength of linear regressive estimation by a simulation study.

Acknowledgments

This work was supported by the Technology foundation STW under the project number WWI.6345.

References

[1] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhäuser, Boston, 1989.

[2] J. Baumeister, W. Scondo, M.A. Demetriou, and I.G. Rosen. On line parameter estimation for infinite-dimensional dynamical systems. *SIAM J. Control Optim.*, 35:2:678–713, 1997.

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

[4] R.F. Curtain and H.J. Zwart. *An Introduction to Infinite Dimensional Linear Systems Theory*. Springer-Verlag, New York, 1995.

[5] G.Y. Hu and R.F. O'Connell. Analytical inversion of symmetric tridiagonal matrices. *J. Phys. A: Math. Gen.*, 29:1511–1513, 1996.

[6] Y. Huang and W.F. McColl. Analytical inversion of general tridiagonal matrices. *J. Phys. A: Math. Gen.*, 30:7919–7933, 1997.

[7] P. Van Overschee. Choice of state-space basis in combined deterministic-stochastic subspace identification. *Automatica*, 31:12:1877–1883, 1995.

[8] R. Pintelon, P. Guillaume, G. Vandersteen, and Y. Rolain. Analysis, development and applications of tls algorithms in frequency domain system identification. *SIAM J. Matrix Anal. Appl.*, 19:4:983–1004, 1998.

[9] T. Söderström and B. Bhikkaji. Reduced order models for diffusion systems via collacation methods. In *IFAC 12th Symposium on System Identification*. (CD-Rom). Santa Barbara, California, June 21-23 2000.

[10] T. Söderström and Susanne Remle. Parameter estimation and model approximation for diffusion models. In *IFAC 12th Symposium on System Identification*. (CD-Rom). Santa Barbara, California, June 21-23 2000.

[11] D. Vries, K.J. Keesman, and H. Zwart. Explicit linear regressive model structures for estimation, prediction and experimental design in compartmental diffusive systems. In *IFAC 14th Symposium on System Identification*. (CD-Rom). Newcastle, Australia, Accepted for presentation 2006.