ON A SIMPLE OVERLAPPING STATE-SPACE PARAMETRIZATION FOR LINEAR TIME SERIES MODELS

Andrea Gombani [*] **Thomas Ribarits** [**]

[*] *LADSEB-CNR*
Corso Stati Uniti 4
35127 Padova, Italy
[**] *Institute for Mathematical Methods in Economics*
Research Unit Econometrics and System Theory
Vienna University of Technology
Argentinierstrasse 8, 1040 Vienna, Austria

`gombani@ladseb.pd.cnr.it`
`thomas.ribarits@tuwien.ac.at`

Abstract: We consider a new state-space parametrization for linear time series models: data driven coordinates (DDC), which provides an atlas for the manifold of (stable) $p \times m$ transfer functions of fixed McMillan degree n. Hence, DDC has similar desirable properties as more traditional overlapping parametrizations and better than classical canonical forms. Moreover, the choice of charts can be done in a data-driven manner in a very simple way. Althugh not yet as good numerically as the parametrization by data driven local coordinates (DDLC), this parametrization has the advantage of not being *local*. The application of DDC to maximum likelihood identification is exemplified.

Keywords: Parametrization, Linear multivariable systems, State-space models, Identifiability, System Identification.

1. INTRODUCTION

In this paper a new state-space parametrization for linear time series models is introduced: data driven coordinates (DDC). DDC provides an atlas of uncountably many overlapping charts for the manifold of (stable) $p \times m$ transfer functions of fixed McMillan degree n. Each coordinate neighborhood is a generic subset of this manifold. Hence, DDC has similar desirable properties as more traditional overlapping parametrizations and classical canonical forms. Moreover, the choice of charts out of an uncountable set of possible charts can be done in a data-driven manner in the course of the estimation procedure in a very simple way. In this respect, DDC resembles the parametrization by data driven local coordinates (DDLC); while (DDLC) still provides slightly better results, this parametrization has the advantage of not being *local*. The application of DDC to maximum likelihood identification is exemplified. The charts we use are by no means the only possible ones and we conjecture that there is a wide margin for improvement here.

Let us denote by $\mathcal{W}_n^{p,m}$ the set of stable $p \times m$ continuous time transfer functions of McMillan degree n with the topology induced by $\mathcal{H}_+^2$. Recall that stability amounts to analyticity in the right half plane $\mathbb{C}^+$ and $\mathcal{H}_+^2$ denotes the Hardy space of stable matrix valued functions with its usual inner product. Unless explicitly stated otherwise, all statements are to be understood in continuous time in the sequel. By

$\mathcal{S}_c \subseteq \mathbb{R}^{n^2+mn}$ we denote the set of stable and controllable real matrix pairs (e.g. (A_1,B_1)) and by $\mathcal{S}_m$ we denote the set of all stable and minimal state-space realizations (e.g. $(A_1,B_1,C_1,D_1) \subseteq \mathbb{R}^{n^2+n(m+p)+mp}$). It is easily seen that $\mathcal{S}_c$ is generic (i.e. open and dense) in the set of all stable matrix pairs and that $\mathcal{S}_m$ is generic in the set of all stable state-space realizations.

Consider the mapping π attaching transfer functions to system matrices:

$$\begin{array}{cccc} \pi: & \mathcal{S}_m & \to & \mathcal{W}_n^{p,m} \\ & (A_1,B_1,C_1,D_1) & \mapsto & \left(\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right) = C_1(sI-A_1)^{-1}B_1 + D_1 \end{array} \tag{1}$$

As is easily seen, π is continuous and it is an open mapping (where the relative topology corresponding to the Euclidean norm is considered in $\mathcal{S}_m$); see e.g. (Ribarits, 2002).

The paper is organized as follows: In Section 2 we introduce a continuous time parametrization. Section 3 then introduces two variants of DDC: First, the continuous time parametrization derived in Section 2 is translated to the discrete time case by means of the bilinear transform, leading to what is called DDC_{cont}. A second approach uses the same (state) feedback structure as in the continuous time case to derive a parametrization directly in discrete time (DDC_{dis}). In Section 4 simulation experiments are presented, where DDC is compared to more common parametrizations in a maximum likelihood identification setting. Finally, Section 5 contains the conclusions.

2. PARAMETRIZATION IN CONTINUOUS TIME

Although in most applications we deal with real coefficient matrices, we develop our results for the complex case. So, if A is a complex matrix, A^* denotes its transpose conjugate, while A' denotes its transpose. Also note that a rational matrix valued $m \times m$ transfer function $Q(s)$ is called inner if it is stable and satisifies $Q^*(s)Q(s) = I_m$, where $s = i\omega$, for (almost all) $\omega \in \mathbb{R}$; notice that we have $Q^*(s) = Q(-s)'$.

We will need a simple result in interpolation theory (see (Ball *et al.*, 1990) for the formulation of the problem). Let $\mathcal{A} \in \mathbb{C}^{n\times n}$ be stable, i.e. its eigenvalues are in the open left half plane $\mathbb{C}^-$, and let $U \in \mathbb{C}^{n\times m}$ be given. Moreover, let Γ denote any closed curve around the spectrum of $-\mathcal{A}^*$ which stays entirely in $\mathbb{C}^+$. We want to find the solutions to the generalized interpolation problem, i.e. we want to find an inner function $Q(s)$ of fixed McMillan degree n such that

$$\frac{1}{2\pi i}\int_\Gamma Q(s)U^*(sI+\mathcal{A}^*)^{-1}ds = V^* \quad Q(\infty) = I_m \tag{2}$$

Lemma 1. Let $(\mathcal{A},U,V)$ be given, where $\mathcal{A} \in \mathbb{C}^{n\times n}$ is stable and $U,V \in \mathbb{C}^{n\times m}$. Assume that the solution P to the Lyapunov equation $\mathcal{A}P + P\mathcal{A}^* + UU^* - VV^* = 0$ is positive definite. Then there exists a unique inner $Q(s)$ of McMillan degree n satisfying the interpolation conditions (2) and it is given by

$$Q(s) = \left(\begin{array}{c|c} -\mathcal{A}^* - BU^* & B \\ \hline -(U-V)^* & I \end{array}\right) \quad B = P^{-1}(U-V) \tag{3}$$

Moreover, P^{-1} is the controllability Gramian corresponding to the pair $(-\mathcal{A} - BU^*, B)$.

PROOF. See (Gombani and M.Olivi, 2000).

The following Sylvester equation will be of great importance below:

$$A_1\mathcal{Y} + \mathcal{Y}\mathcal{A}^* + B_1U^* = 0 \tag{4}$$

Theorem 2. Let the matrix pair $(\mathcal{A},U) \in \mathcal{S}_c$ be given. Define the set $\mathcal{V}_{(\mathcal{A},U)}$ of transfer functions and the function $\phi_{(\mathcal{A},U)}$ as follows:

$$\mathcal{V}_{(\mathcal{A},U)} := \{W(s) = \left(\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right) \in \mathcal{W}_n^{p,m} : \text{the solution } \mathcal{Y} \text{ in (4) is non singular}\} \tag{5}$$

$$\phi_{(\mathcal{A},U)} : \begin{cases} \mathcal{V}_{(\mathcal{A},U)} \to \mathbb{R}^{n(m+p)+pm} \\ W(s) \mapsto (\mathcal{Y}^{-1}B_1, C_1\mathcal{Y}, D_1) \end{cases}$$

Then the family $(\mathcal{V}_{(\mathcal{A},U)}, \phi_{(\mathcal{A},U)}, (\mathcal{A},U) \in \mathcal{S}_c)$ forms an atlas for the set $\mathcal{W}_n^{p,m}$ whose topology coincides with the one induced by the topology of $\mathcal{H}_+^2$ and each coordinate neighborhood $\mathcal{V}_{(\mathcal{A},U)}$ is generic in $\mathcal{W}_n^{p,m}$. The inverse map is given by

$$\phi^{-1}_{(\mathcal{A},U)}(B,C,D) = \left(\begin{array}{c|c} -\mathcal{A}^* - BU^* & B \\ \hline C & D \end{array}\right) \tag{6}$$

where the domain of definition is given by triples (B,C,D) such that (6) is stable and minimal.

PROOF. Notice first that the spectra of A_1 and $\sigma(-\mathcal{A}^*)$ are disjoint, and thus the solution to (4) is well known to be unique; it is also easily seen that the definition of $\phi_{(\mathcal{A},U)}(W)$ is independent of the realization and thus the function is well defined.

We claim that $\phi_{(\mathcal{A},U)}$ is a homeomorphism between $\mathcal{V}_{(\mathcal{A},U)}$ and an open subset of $\mathbb{R}^{n(m+p)+pm}$. To see this, we use the Douglas-Shapiro-Shields factorization $W = \overline{W}Q$ where $\overline{W}$ is antistable and Q is inner and the degree of Q is minimal. It's well known (and simple to check) that, for any given realization of Q, we can choose a realization for W so that the pair (A,B) is the same. Now, a realization for Q can be chosen as in (3), where $B = P^{-1}(U-V)$ and V is given by (2). Thus V depends continuously on Q. Thus so do P and B. Since the DSS factorization for rational functions is continuous, we get that B depends continuously on W. But then, so does A; as for C, defining $H(Q) = span\{\xi^*(sI-A)^{-1}B$, and denoting by $\mathcal{P}_{H(Q)}$ the projection onto $H(Q)$, and using the well known projection formula, we see that

$$\mathcal{P}_{H(Q)}W = \langle W, (sI-A)^{-1}B \rangle = (CP)P^{-1}(sI-A)^{-1}B$$

and so C represents the projection of W onto $H(Q)$ in the basis $(sI-A)^{-1}B$, and thus it is a continuous function of W. The function is clearly injective (two different functions cannot have the same realization). Surjectivity is clear by construction. The inverse map is thus well defined and it is given by (6). Continuity is also obvious.

The sets $\mathcal{V}_{(\mathcal{A},U)}$ are open. In fact, for given $(\mathcal{A},U)$, the set of state-space systems defined by the conditions (4) and $\det \mathcal{Y} = 0$, is closed and algebraic. Hence, the condition defining $\mathcal{V}_{(\mathcal{A},U)}$ yields an open and dense set of state-space systems unless the set is empty. To rule out the latter possibility, put $A_1 = \mathcal{A}$ and $B_1 = U$ (for given $(\mathcal{A},U)$), which clearly yields a non singular solution to (4). As the mapping π in (1) is an open and continuous mapping, this shows that the set $\mathcal{V}_{(\mathcal{A},U)}$ is generic in $\mathcal{W}_n^{p,m}$ for any given $(\mathcal{A},U) \in \mathcal{S}_c$.

If $(\mathcal{A}_1,U_1)$ and $(\mathcal{A}_2,U_2)$ define two charts such that $\mathcal{V}_{(\mathcal{A}_1,U_1)} \cap \mathcal{V}_{(\mathcal{A}_2,U_2)} \neq \emptyset$, the computation of the function $\phi_{(\mathcal{A}_1,U_1)} \circ \phi^{-1}_{(\mathcal{A}_2,U_2)}$ entails the solution of the linear system (4) and the inversion of this solution (which exists by assumption). Thus the function is clearly continuous and differentiable with its inverse.

Finally, we show that for any $\mathcal{W} \in \mathcal{W}_n^{p,m}$ with state-space realization (A_1,B_1,C_1,D_1) there exists a $\mathcal{V}_{(\mathcal{A},U)} \ni W$. Set P to be the solution to $A_1P + PA_1^* + B_1B_1^* = 0$. Then, setting $\mathcal{A} := P^{-1}A_1P$ and $U := P^{-1}B_1$, we obviously have $A_1 = -PA_1^*P^{-1} - B_1B_1^*P^{-1} = -\mathcal{A}^* - B_1U^*$ and thus $W \in \mathcal{V}_{(\mathcal{A},U)}$.

3. PARAMETRIZATION IN DISCRETE TIME

We now discuss two possibilities to apply Theorem 2 to the discrete time case. Let $\mathcal{Z}_n^{p,m}$ denote the set of discrete time stable $p \times m$ transfer functions of degree n with the topology induced by $\mathcal{H}(\partial\mathbb{D})^2_+$, the Hardy space of discrete time stable matrix valued functions with its usual inner product. Likewise, let $\mathcal{Z}_n^p$ denote the set of discrete time stable and strictly minimum phase $p \times p$ transfer functions of degree n induced with the relative topology.

3.1 DDC$_{cont}$

The following mapping is known under the term *bilinear transformation*:

$$\rho_1 : \mathbb{C} \to \mathbb{C} \quad z \mapsto \frac{1-z}{1+z} = s \tag{7}$$

Note that ρ_1 is a *bijection* on the compactified complex plane (with inverse $\rho_1^{-1}(s) = z = \frac{1-s}{1+s}$) mapping the complement of the closed unit disk onto the open left half plane. Therefore, the mapping

$$Z(z) = W(\rho_1(z))) \tag{8}$$

preserves the stability and minimum-phase property. In terms of state-space representations (A_c,B_c,C_c,D_c) for $W(s)$ and (A,B,C,D) for $Z(z)$ the transformation (8) can be chosen to be of the form given below:

Theorem 3. Let (A_c,B_c,C_c,D_c) be a (not necessarily minimal) state-space representation of some (not necessarily stable) continuous time transfer function $W(s)$ and let $\lambda_i(A_c) \neq 1$, $i = 1,\ldots,n$ hold true. Then $(A,B,C,D) = \rho(A_c,B_c,C_c,D_c)$ is a state-space representation of the discrete time transfer function $Z(z)$ given in (8) where

$$\rho(A_c,B_c,C_c,D_c) = (A,B,C,D) \tag{9}$$
$$A = (I+A_c)(I-A_c)^{-1} \quad B = \sqrt{2}(I-A_c)^{-1}B_c$$
$$C = \sqrt{2}C_c(I-A_c)^{-1} \quad D = D_c + C_c(I-A_c)^{-1}B_c$$

The mapping ρ has the following properties:

(i) ρ is a *homeomorphism* between the set of minimal and stable (and strictly minimum phase) continuous time systems and the set of minimal and stable (and strictly minimum phase) discrete time systems.

(ii) ρ preserves observational equivalence ($\sim$) for minimal systems: $(A_c,B_c,C_c,D_c) \sim (A_{c,1},B_{c,1},C_{c,1},D_{c,1})$ $\Leftrightarrow \rho(A_c,B_c,C_c,D_c) \sim \rho(A_{c,1},B_{c,1},C_{c,1},D_{c,1})$

(iii) The controllability and observability Gramians stay invariant under the transformation ρ.

Its inverse is given by

$$\rho^{-1}(A,B,C,D) = (A_c,B_c,C_c,D_c) \tag{10}$$
$$A_c = (I+A)^{-1}(A-I) \quad B_c = \sqrt{2}(I+A)^{-1}B$$
$$C_c = \sqrt{2}C(I+A)^{-1} \quad D_c = D - C(I+A)^{-1}B$$

PROOF. All results are well known. For a proof of this particular Theorem see e.g. the proof of Theorem A.4.1 in the appendix of (Ribarits, 2002).

Theorem 4. (DDC$_{cont}$). Let the controllable and (discrete time) stable matrix pair $(\mathcal{A}_d,U_d)$ be given and let $(\mathcal{A},U) = ((I+\mathcal{A}_d)^{-1}(\mathcal{A}_d - I), \sqrt{2}(I+\mathcal{A}_d)^{-1}B_d)$ be the corresponding controllable continuous time stable matrix pair. Furthermore, for any given (A_d,B_d,C_d,D_d), let $(A_1,B_1,C_1,D_1) = \rho^{-1}(A_d,B_d,C_d,D_d)$ denote the corresponding continuous time state-space matrices. Define the set $\mathcal{V}_{(\mathcal{A}_d,U_d)}$ of transfer functions and the function $\phi_{(\mathcal{A}_d,U_d)}$ as follows:

$$\mathcal{V}_{(\mathcal{A}_d,U_d)} := \{Z(z) = \left(\begin{array}{c|c} A_d & B_d \\ \hline C_d & D_d \end{array}\right) \in \mathcal{Z}_n^{p,m} : \text{the solution } \mathcal{Y} \text{ in (4) is unique and non singular}\}$$

$$\phi_{(\mathcal{A}_d,U_d)} : \begin{cases} \mathcal{V}_{(\mathcal{A}_d,U_d)} & \rightarrow \quad \mathbb{R}^{n(m+p)+pm} \\ Z(z) & \mapsto \quad (\mathcal{Y}^{-1}B_1, C_1\mathcal{Y}, D_1) \end{cases} \tag{11}$$

Then the family $(\mathcal{V}_{(\mathcal{A}_d,U_d)}, \phi_{(\mathcal{A}_d,U_d)}, (\mathcal{A}_d,U_d)$ stable and controllable$)$ forms an atlas for the set $\mathcal{Z}_n^{p,m}$ whose topology coincides with the one induced by the topology of $\mathcal{H}(\partial\mathbb{D})_+^2$ and each coordinate neighborhood $\mathcal{V}_{(\mathcal{A}_d,U_d)}$ is generic in $\mathcal{Z}_n^{p,m}$. The inverse map is given by

$$\phi^{-1}_{(\mathcal{A}_d,U_d)}(B,C,D) = \left(\begin{array}{c|c} A_d & B_d \\ \hline C_d & D_d \end{array}\right) \tag{12}$$

where $(A_d,B_d,C_d,D_d) = \rho(-\mathcal{A}^* - BU^*, B, C, D)$.

PROOF. From (i) and (ii) in Theorem 3 it follows that any overlapping form e.g. for stable continuous time state-space representations directly induces an overlapping form for stable discrete time state-space representations. We use this fact and apply (9) to the continuous time state-space matrices parametrized in (6) to get the parametrization by DDC_{cont} given in (12).

Remark 5. Note that an atlas for the set $\mathcal{Z}_n^p$ can be defined completely analogously as ρ in (9) also preserves the strict minimum phase property. Similarly, stability need not be imposed, i.e. an atlas for the set of $p \times m$ transfer functions of fixed McMillan degree can be defined. Clearly, the parameter spaces, i.e. the image of the corresponding sets of transfer functions under the mapping $\phi_{(\mathcal{A}_d,U_d)}$ change, but remain open subsets of a Euclidean space.

Remark 6. The derivatives of the continuous time state-space matrices with respect to the parameters are easily computed as the state-space matrices are an affine function of the parameters; see (6). In order to determine the derivatives of the discrete time state-space matrices, we use equation (13) below. Note that a 'dot' stands for the derivative with respect to some entry in the parameter vector:

$$\begin{pmatrix} vec\dot{A} \\ vec\dot{B} \\ vec\dot{C} \\ vec\dot{D} \end{pmatrix} = \begin{pmatrix} (I-A_c)^{-1'} \otimes (I+A) & 0 & 0 & 0 \\ B' \otimes (I-A_c)^{-1} & \sqrt{2}I_m \otimes (I-A_c)^{-1} & 0 & 0 \\ (I-A_c)^{-1'} \otimes C & 0 & (I-A_c)^{-1'} \otimes \sqrt{2}I_p & 0 \\ B' \otimes \frac{1}{2}C & \frac{1}{\sqrt{2}}I_m \otimes C & B' \otimes \frac{1}{\sqrt{2}}I_p & I_m \otimes I_p \end{pmatrix} \begin{pmatrix} vec\dot{A}_c \\ vec\dot{B}_c \\ vec\dot{C}_c \\ vec\dot{D}_c \end{pmatrix} \tag{13}$$

The derivation of formula (13) is straightforward; see Section 5.4 in (Ribarits, 2002).

A common problem in system identification is the optimization of a criterion function over the manifold $\mathcal{Z}_n^{p,m}$. In many cases this optimization has to be performed by means of an iterative search procedure. In the course of such a search procedure, DDC_{cont} offers a simple way to change charts by an appropriate choice of $(\mathcal{A},U) = ((I+\mathcal{A}_d)^{-1}(\mathcal{A}_d - I), \sqrt{2}(I+\mathcal{A}_d)^{-1}B_d)$. Starting from a state-space realization (A_d,B_d,C_d,D_d) of $Z(z)$, where $(A_1,B_1,C_1,D_1) = \rho^{-1}(A_d,B_d,C_d,D_d)$, one can change charts e.g. by setting

- ► $(\mathcal{A},U) = (P^{-1}A_1P, P^{-1}B_1) \Rightarrow \mathcal{Y} = I$, $(\mathcal{A}_d,U_d) = ((I+\mathcal{A})(I-\mathcal{A})^{-1}, \sqrt{2}(I-\mathcal{A})^{-1}U) \Rightarrow \phi_{(\mathcal{A}_d,U_d)}(Z) = (B_1,C_1,D_1)$ with new controllability Gramian P
- ► $(\mathcal{A},U) = (P^{-\frac{1}{2}}A_1P^{\frac{1}{2}}, P^{-\frac{1}{2}}B_1) \Rightarrow \mathcal{Y} = P^{\frac{1}{2}}$, $(\mathcal{A}_d,U_d) = ((I+\mathcal{A})(I-\mathcal{A})^{-1}, \sqrt{2}(I-\mathcal{A})^{-1}U) \Rightarrow \phi_{(\mathcal{A}_d,U_d)}(Z) = (P^{-\frac{1}{2}}B_1, C_1P^{\frac{1}{2}}, D_1)$ with the identity as new controllability Gramian

Here P denotes the controllability Gramian, i.e. the solution to $A_1P + PA_1^* + B_1B_1^* = 0$, which coincides with the controllability Gramian of the discrete time system (A_d,B_d,C_d,D_d). Similarly, a change of charts can be performed by a proper choice of $(\mathcal{A},U)$ in such a way that the new realization becomes Lypunov balanced; see Section 4.

3.2 DDC_{dis}

Another possibility is to consider equation (4) directly in discrete time. Hence, (A_1,B_1,C_1,D_1) and $(\mathcal{A},U)$ are to be understood in discrete time in the sequel.

Theorem 7. (DDC_{dis}). Let a controllable and (discrete time) stable matrix pair $(\mathcal{A},U)$ be given. Define the set $\mathcal{V}_{(\mathcal{A},U)}$ of transfer functions and the function $\phi_{(\mathcal{A},U)}$ as follows:

$$\mathcal{V}_{(\mathcal{A},U)} := \{Z(z) = \left(\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right) \in \mathcal{Z}_n^p : \text{the solution } \mathcal{Y} \text{ in (4) is unique and non singular}\}$$

$$\phi_{(\mathcal{A},U)} : \begin{cases} \mathcal{V}_{(\mathcal{A},U)} & \rightarrow \quad \mathbb{R}^{n(m+p)+pm} \\ Z(z) & \mapsto \quad (\mathcal{Y}^{-1}B_1, C_1\mathcal{Y}, D_1) \end{cases} \tag{14}$$

Then the family $(\mathcal{V}_{(\mathcal{A},U)}, \phi_{(\mathcal{A},U)}, (\mathcal{A},U)$ stable and controllable$)$ forms an atlas for the set $\mathcal{Z}_n^p$ whose topology coincides with the one induced by the topology of $\mathcal{H}(\partial\mathbb{D})_+^2$ and each coordinate neighborhood $\mathcal{V}_{(\mathcal{A},U)}$ is generic in $\mathcal{Z}_n^p$. The inverse map is given by

$$\phi^{-1}_{(\mathcal{A},U)}(B,C,D) = \left(\begin{array}{c|c} -\mathcal{A}^* - BU^* & B \\ \hline C & D \end{array}\right) \tag{15}$$

PROOF.

The proof proceeds along the same lines as the proof of Theorem 2 and is omitted.

Remark 8. Theorem 7 applies to the set $\mathcal{Z}_n^p$. However, it can be straightforwardly be used to derive an overlapping parametrization for the set $\mathcal{Z}_n^{p,(p+m)}$ of transfer functions of the form $(K(z), L(z))$ where $K(z)$ is a stable and strictly minimum phase transfer function (corresponding to the noise model) and $L(z)$ is a stable transfer function. In terms of a state-space realization, this corresponds to

$$\left.\begin{array}{l} x_{t+1} = Ax_t + Bu_t + K\varepsilon_t \\ y_t = Cx_t + Du_t + E\varepsilon_t \end{array}\right\} \quad (K(z), L(z)) = C(z^{-1}I - A)^{-1}(B,K) + (D,E) \tag{16}$$

In fact, neither the stability nor the strict minimum phase property need to be imposed (see Remark 5).

Note that $\mathtt{DDC}_{dis}$ also offers the possibility to change charts easily e.g. in the course of a search algorithm. Starting from a state-space realization (A_1, B_1, C_1, D_1) of $Z(z)$, one can change charts e.g. by setting

► $(\mathcal{A}, U) = (-(A_1 - B_1 D_1^{-1} C_1)^*, -(D_1^{-1} C_1)^*) \Rightarrow \mathcal{Y} = I \Rightarrow \phi_{(\mathcal{A},U)}(Z) = (B_1, C_1, D_1)$ with new controllability Gramian P

► $(\mathcal{A}, U) = -(P^{-\frac{1}{2}}(A_1 - B_1 D_1^{-1} C_1)^* P^{\frac{1}{2}}, P^{-\frac{1}{2}}(D_1^{-1} C_1)^*)$ $\Rightarrow \mathcal{Y} = P^{\frac{1}{2}}, \Rightarrow \phi_{(\mathcal{A},U)}(Z) = (P^{-\frac{1}{2}} B_1, C_1 P^{\frac{1}{2}}, D_1)$ with the identity as new controllability Gramian

Here P denotes the controllability Gramian, i.e. the solution to $P - A_1 P A_1^* - B_1 B_1^* = 0$. Similarly, a change of charts can be performed by a proper choice of $(\mathcal{A}, U)$ in such a way that the new realization becomes Lypunov balanced; see Section 4.

Note that by the above mentioned choices of $(\mathcal{A}, U)$ the *parameters B_1 and C_1 determine the zeros and poles of the transfer function independent of each other!* If, for instance, one wants to change the poles of (A_1, B_1, C_1, I) without changing its zeros, then a change of B_1 with an unchanged C_1 will do. The converse direction is obvious for any choice of $(\mathcal{A}, U)$: Changing the zeros of (A_1, B_1, C_1, I) without changing its poles has to be accomplished by a change of C_1 without changing B_1.

4. SIMULATION STUDIES

For the sake of a numerical comparison, 800 different minimal, stable and strictly minimum phase time series models (without exogenous inputs) of the form

$$x_{t+1} = Ax_t + B\varepsilon_t, \quad y_t = Cx_t + \varepsilon_t \tag{17}$$

are randomly generated. Here, (ε_t) is a white noise process, and the output y_t is two dimensional. The models are of order 2, 4, ..., 16. The model orders are given in the first row of the tables below. For each order, 100 models are generated.

Simulation data comprising $T = 500$ output observations are created, where the white noise sequence (ε_t) is chosen to be Gaussian distributed with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

All computations are carried out using the system identification toolbox of the software package `MATLAB`, version `6.5.1.199709 (R13)`. Initial estimates are computed by means of a subspace procedure (`n4sid`). It is ensured that the initial models are minimal, stable and minimum phase. The identification procedure itself is performed by using the built-in function `pem`. The option `SearchDirection` is set to `Gn` (a plain Gauss-Newton type algorithm is used for minimizing the criterion function).

In the sequel, we compare two variants of `DDC` ($\mathtt{DDC}_{cont}^{rebal}$ and $\mathtt{DDC}_{dis}$) with the echelon canonical form (`Can`) and the parametrization by data driven local coordinates ($\mathtt{DDLC}_{bal}$, starting from a Lyapunov balanced realization of the initial subspace estimate). Note that for $\mathtt{DDC}_{cont}^{rebal}$ a change of charts is performed such that the realizations remain Lyapunov balanced in the course of the optimization algorithm; see the paragraph below Remark 6. For $\mathtt{DDC}_{dis}$ charts are chosen according to the first choice described in the paragraph below Remark 8 (without rebalancing in the course of the optimization procedure).

It should be noted that the echelon canonical form is also kown under the term *observable canoncial form*; see e.g. Chapter 2 in (Hannan and Deistler, 1988) and Appendix 4A in (Ljung, 1999). `DDLC`, which has been introduced in (McKelvey *et al.*, 2004), does *not* provide an identifiable parametrization, and it does *not* describe a generic subset of $\mathcal{Z}_n^p$. In fact, it need not even describe an open subset of $\mathcal{Z}_n^p$; see (Ribarits *et al.*, 2004) for an analysis of `DDLC`. However, it serves as an important benchmark because it is the current standard parametrization used in the system identification toolbox of `MATLAB`.

Table 1. Percentage of failed runs.

	2	4	6	8	10	12	14	16
$DDLC^{bal}$	0	0	1	1	0	0	0	0
Can	0	0	4	12	15	17	23	28
DDC_{cont}^{rebal}	1	0	4	2	2	1	4	8
DDC_{dis}	1	13	11	22	24	18	28	36

For each order (each column corresponds to a fixed order) and each parametrization (each row corresponds to a fixed parametrization), Table 1 shows the percentage of failed runs out of the 100 estimation experiments. An identification experiment is considered to have failed if the algorithm yields a final parameter estimate where the value of the likelihood function is more than 20% worse than the value of the likelihood function at the true system. Concerning the *success rates*, $\mathtt{DDC}_{cont}^{rebal}$ clearly outperforms the classical echelon canonical form, but is a bit worse than $\mathtt{DDLC}_{bal}$. $\mathtt{DDC}_{dis}$ does not perform well.

Table 4.A shows the average number of iterations for successful experiments, and Table 4.B yields the same information for the experiments which failed. Concerning the *speed of 'convergence'* (i.e. the number of iterations until a termination criterion is met), $\mathtt{DDC}_{cont}^{rebal}$ on average needs the same number of iterations as $\mathtt{DDLC}_{bal}$, but considerably fewer iterations than the classical echelon canonical form. During the estimation process one can also record the Gauss-Newton approximation to the Hessian of the criterion function

Table 3. Average maximum condition number of the Gauss-Newton approximations to the Hessians for successful runs (A) and for failed runs (B). Test cases with no successful, respectively failed, runs are indicated by 0.

A	2	4	6	8	10	12	14	16
$DDLC^{bal}$	2.9e+7	2.4e+9	2.1e+9	2.5e+9	2.5e+9	5.2e+9	7.1e+8	8.5e+9
Can	1.3e+18	1.2e+17	1.8e+17	1.0e+19	2.8e+18	1.6e+16	1.5e+20	9.9e+19
DDC_{cont}^{rebal}	9.0e+7	2.6e+8	1.6e+10	3.8e+10	1.3e+11	2.6e+10	4.5e+10	9.6e+9
DDC_{dis}	3.0e+11	1.5e+13	2.4e+12	1.9e+16	5.4e+16	2.2e+14	7.6e+13	2.5e+15
B								
$DDLC^{bal}$	0.	0.	2.6e+7	2.6e+5	0.	0.	0.	0.
Can	0.	0.	4.9e+18	1.5e+18	1.6e+17	7.1e+21	5.1e+20	2.0e+19
DDC_{cont}^{rebal}	1.6e+16	0.	1.2e+10	6.6e+10	1.6e+9	4.4e+6	1.5e+10	4.6e+12
DDC_{dis}	1.1e+11	3.9e+11	2.6e+15	2.6e+11	1.2e+11	3.1e+12	2.7e+12	3.8e+17

Table 2. Average number of iterations ($\leq$ 50) for successful runs (A) and for failed runs (B). Test cases with no successful, respectively failed, runs are indicated by 0.

A	2	4	6	8	10	12	14	16
$DDLC^{bal}$	8	12	16	20	16	19	18	19
Can	9	18	24	33	35	35	37	37
DDC_{cont}^{rebal}	10	13	17	19	18	18	19	16
DDC_{dis}	20	28	30	26	26	24	29	26
B								
$DDLC^{bal}$	0	0	27	1	0	0	0	0
Can	0	0	21	26	18	18	14	20
DDC_{cont}^{rebal}	50	0	7	26	6	1	4	9
DDC_{dis}	50	18	34	18	14	17	14	12

at the current parameter estimate. This is done for all iterations, and then the maximum condition number of these matrices is stored. Table 4.A shows the average of these maximum condition numbers over all successful identification experiments. Table 4.B yields the same information for the experiments which failed. Concerning the magnitude of these *condition numbers* we see that the usage of $\mathtt{DDC}_{cont}^{rebal}$ leads to much lower condition numbers as compared to the echelon canonical form `Can`, and approximately the same, yet a bit worse, condition numbers as compared to $\mathtt{DDLC}_{bal}$.

It may also be interesting to examine the values of the likelihood function at the final parameter estimates. We have not included another table for reasons of space limitations, but it again turns out that *likelihood values* upon convergence are considerably lower (and therefore better) for $\mathtt{DDC}_{cont}^{rebal}$ than for `Can`. $\mathtt{DDLC}_{bal}$ again turns out to be a bit better than $\mathtt{DDC}_{cont}^{rebal}$.

5. CONCLUSIONS

A new state-space parametrization for linear time series models, `DDC`, has been introduced. `DDC` has similar desirable properties as more traditional overlapping parametrizations and classical canonical forms, additionally offering the possibility of a simple data-driven choice of charts out of an uncountable set of charts in the course of an estimation algorithm.

Simulation studies show that particular changes of charts may indeed be very beneficial for the estimation procedure. The usage of `DDC` has clear numerical advantages as compared to the more commonly used echelon canonical and overlapping forms. `DDC` also has much more favourable global properties than `DDLC`, the parametrization by data driven *local* coordinates which is the current standard parametrization in `MATLAB`. However, the price to be paid for this advantage is that simulation results for `DDC` are slightly worse than for `DDLC` if good initial (subspace) estimates are available. An important open problem is the question of choosing charts for `DDC` ‘optimally’.

6. ACKNOWLEDGEMENTS

Support by the Austrian ’Fonds zur Förderung der wissenschaftlichen Forschung’, Project P-17065 and grant SFB#010 (’Adaptive Information Systems and Modelling in Economics and Management Science’) is acknowledged. The work was also supported in part by the European union’s Human Potential Programme under contract HPRN-CT-2000-00100, DYNSTOCH.

REFERENCES

Ball, J. A., I. Gohberg and L. Rodman (1990). *Interpolation of rational matrix functions*. Birkhaeuser Verlag. Basel.

Gombani, A. and M.Olivi (2000). A new parametrization of rational inner functions of fixed degree: Schur parameters & realizations. *Mathematics of Control, Signals and Systems* **13**, 156–177.

Hannan, E. J. and M. Deistler (1988). *The Statistical Theory of Linear Systems*. John Wiley. New York.

Ljung, L. (1999). *System Identification: Theory for the User*. second ed.. Prentice Hall.

McKelvey, T., A. Helmersson and T. Ribarits (2004). Data driven local coordinates for multivariable linear systems and their application to system identification. *Automatica* **40**(9), 1629–1635.

Ribarits, T. (2002). The Role of Parametrizations in Identification of Linear Dynamic Systems. PhD thesis. TU Wien.

Ribarits, T., M. Deistler and T. McKelvey (2004). An analysis of the parametrization by data driven local coordinates for multivariable linear systems. *Automatica* **40**(5), 789–803.