A NEW SUBSPACE IDENTIFICATION METHOD FOR OPEN AND CLOSED LOOP DATA

Magnus Jansson *,1

* *Royal Inst. of Technology (KTH), Stockholm, Sweden*

Abstract: Subspace methods have emerged as useful tools for the identification of linear time invariant discrete time systems. Most of the methods have been developed for the open loop case to avoid difficulties with data correlations due to the feedback. This paper extends some recent ideas for developing subspace methods that can perform well on data collected both in open and closed loop conditions. Here, a method that aims at minimizing the prediction errors in several approximate steps is proposed. The steps involve using constrained least squares estimation on models with different degrees of structure such as block-toeplitz, and reduced rank matrices. The statistical estimation performance of the method is shown to be competitive to existing subspace methods in a simulation example.

Keywords: multivariable, identification

1. INTRODUCTION

The subspace methods for system identification estimate linear state-space models directly from time-discrete observations. They have become part of the standard tools for the analysis of data from dynamical systems, especially for systems with multiple inputs and multiple outputs. Many of the ideas behind these methods come from classical state-space realization theory. The main observation used in the more recent algorithms (Van Overschee and De Moor, 1996; Verhaegen, 1994; Larimore, 1983; Peternell *et al.*, 1996) is that, under the assumption that there exists a true underlying finite order linear time invariant system, an estimate of the observability matrix or the state-trajectory can be obtained from the singular value decomposition of a certain data cross correlation matrix. Many of the subspace methods are also computationally attractive since only standard matrix operations are utilized to calculate the estimates and they do not use iterative optimization techniques as, in general, are needed for the maximum likelihood or the prediction error methods (PEM) (Ljung, 1987).

The "standard" subspace methods referred to above have problems when data are collected in closed loop. The estimates will then in general be biased since the feedback introduces a correlation between the input and the noise. It is well known that PEM can provide consistent estimates even on closed loop data (Ljung, 1987; Forssell and Ljung, 1999). (Ljung and McKelvey, 1996) used this fact and proposed to utilize a high order ARX model to circumvent the problem of subspace methods for closed loop data. They used the ARX model to build a bank of predictors from which the state sequence can be estimated similarly to the standard subspace methods. Many other subspace methods for closed loop data have also appeared in the literature (see, e.g., (Verhaegen, 1993; Van Overschee and De Moor, 1997; Chou and Verhaegen, 1999; Gustafsson, 2001; Qin and Ljung, 2003)). However, no method seems to perform satisfactorily in all cases or they utilize additional information about the feedback. Similar to PEM it would be desirable with a subspace method that works satisfactorily re-

[1] Corresponding author: Magnus Jansson, Dept. of Signals, Sensors and Systems, Signal Processing, Royal Inst. of Technology (KTH), SE-100 44 Stockholm, Sweden. Email: magnus.jansson@s3.kth.se, fax +46 8 7907260.

gardless of whether the data are collected in open or closed loop.

The method of this paper is influenced by the ideas in (Jansson, 2003; Ljung and McKelvey, 1996) and is related to the canonical correlation analysis (CCA) method (Peternell *et al.*, 1996; Larimore, 1983). A similar perspective is also given in (Chiuso, 2004; Chiuso and Picci, 2004) in which the methods of (Qin and Ljung, 2003) and (Jansson, 2003) are analyzed.

2. PROBLEM FORMULATION

Consider a time-discrete linear time-invariant dynamic system described by the state-space model in innovations form

$$\begin{cases} \mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}\mathbf{e}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{e}(t) \end{cases} \tag{1}$$

where $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ and $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ denote the observed input and output signals, respectively. $\mathbf{x}(t) \in \mathbb{R}^n$ is the state-vector and n is the system order. We assume that the system is minimal in the sense that the system cannot be described by a state-space model of order less than n. $\mathbf{e}(t) \in \mathbb{R}^{n_y}$ denotes the zero mean white innovation process.

In the model (1), the matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times n_u}$, $\mathbf{C} \in \mathbb{R}^{n_y\times n}$, $\mathbf{D} \in \mathbb{R}^{n_y\times n_u}$, $\mathbf{K} \in \mathbb{R}^{n\times n_y}$ and the covariance matrix of the innovations are unknown and need to be estimated from observations of $\mathbf{u}(t)$ and $\mathbf{y}(t)$ for $t = 1,2,\ldots,N$. Typically, also the model order n is unknown and needs to be estimated. However, this topic will not be covered herein (see, e.g., (Bauer, 2001) for such discussions).

If the observed data used for the estimation are collected in closed loop, it is assumed that the feedback loop contains a delay. In case the controller does not contain a delay we need to assume that the system has a delay and, hence, that $\mathbf{D} \equiv 0$.

3. ESTIMATION METHOD

In the following we will present the ideas behind the new method for estimating $\{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{K}\}$ of this paper.

Let us first introduce some preliminary notation. Define the vectors of stacked inputs, outputs and innovations as

$$\begin{aligned} \mathbf{y}_f(t) &= \begin{bmatrix} \mathbf{y}^T(t) & \mathbf{y}^T(t+1) & \ldots & \mathbf{y}^T(t+f-1) \end{bmatrix}^T \\ \mathbf{u}_f(t) &= \begin{bmatrix} \mathbf{u}^T(t) & \mathbf{u}^T(t+1) & \ldots & \mathbf{u}^T(t+f-1) \end{bmatrix}^T \\ \mathbf{e}_f(t) &= \begin{bmatrix} \mathbf{e}^T(t) & \mathbf{e}^T(t+1) & \ldots & \mathbf{e}^T(t+f-1) \end{bmatrix}^T \end{aligned}$$

where $f > n$ is an integer chosen by the user. In addition, define

$$\begin{aligned} \tilde{\mathbf{A}} &= \mathbf{A} - \mathbf{K}\mathbf{C} \\ \tilde{\mathbf{B}} &= \mathbf{B} - \mathbf{K}\mathbf{D}. \end{aligned}$$

It is well known that we can rewrite (1) as follows

$$\begin{cases} \mathbf{x}(t+1) = \tilde{\mathbf{A}}\mathbf{x}(t) + \tilde{\mathbf{B}}\mathbf{u}(t) + \mathbf{K}\mathbf{y}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{e}(t) \end{cases} \tag{2}$$

Let us now turn to the idea of the method of this paper. First we will use (2) to form the subspace data model:

$$\mathbf{y}_f(t) = \tilde{\Gamma}\mathbf{x}(t) + \tilde{\Phi}\mathbf{u}_f(t) + \tilde{\Psi}\mathbf{y}_f(t) + \mathbf{e}_f(t) \tag{3}$$

where

$$\tilde{\Gamma} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\tilde{\mathbf{A}} \\ \vdots \\ \mathbf{C}\tilde{\mathbf{A}}^{f-1} \end{bmatrix} \tag{4}$$

$$\tilde{\Phi} = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{C}\tilde{\mathbf{B}} & \mathbf{D} & & \\ \mathbf{C}\tilde{\mathbf{A}}\tilde{\mathbf{B}} & \mathbf{C}\tilde{\mathbf{B}} & \ddots & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{C}\tilde{\mathbf{A}}^{f-2}\tilde{\mathbf{B}} & & \mathbf{C}\tilde{\mathbf{B}} & \mathbf{D} \end{bmatrix} \tag{5}$$

$$\tilde{\Psi} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{C}\mathbf{K} & \mathbf{0} & & \\ \mathbf{C}\tilde{\mathbf{A}}\mathbf{K} & \mathbf{C}\mathbf{K} & \ddots & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{C}\tilde{\mathbf{A}}^{f-2}\mathbf{K} & & \mathbf{C}\mathbf{K} & \mathbf{0} \end{bmatrix}. \tag{6}$$

From (2) it is also clear that

$$\begin{aligned} \mathbf{x}(t) = &\sum_{k=0}^{p-1} \tilde{\mathbf{A}}^k [\mathbf{K}\mathbf{y}(t-k-1)\ \tilde{\mathbf{B}}\mathbf{u}(t-k-1)] \\ &+ \tilde{\mathbf{A}}^p \mathbf{x}(t-p) \end{aligned}$$

Assuming $\tilde{\mathbf{A}}$ is stable, this implies that the state can be written as a linear combination of past inputs and outputs. Furthermore, the state can be estimated arbitrarily well by a finite linear combination of past inputs and outputs by choosing p large enough. Let us replace the state in (3) by the estimate

$$\hat{\mathbf{x}}(t) = \mathscr{K}\mathbf{p}(t), \tag{7}$$

where $\mathscr{K}$ is a matrix of unknown coefficients and $\mathbf{p}(t)$ a vector containing delayed inputs and outputs p steps back:

$$\begin{aligned} \mathbf{p}(t) = \big[&\mathbf{y}^T(t-1)\ \mathbf{y}^T(t-2)\ \ldots\ \mathbf{y}^T(t-p) \\ &\mathbf{u}^T(t-1)\ \mathbf{u}^T(t-2)\ \ldots\ \mathbf{u}^T(t-p)\big]^T. \end{aligned}$$

Hence, consider the following approximate data model

$$\mathbf{y}_f(t) \approx \Gamma\mathscr{K}\mathbf{p}(t) + \tilde{\Phi}\mathbf{u}_f(t) + \tilde{\Psi}\mathbf{y}_f(t) + \mathbf{e}_f(t) \tag{8}$$

where the approximation is due to the fact that the true state has been replaced by the truncated state estimator (7). This is a non-linear regression problem. In the following, we will attempt to solve it by relaxation using linear and reduced-rank regression techniques.

The first step of the proposed method views (8) as a linear regression where $\Gamma\mathscr{K}$ is considered to be a full rank matrix of unknown parameters whereas $\tilde{\Phi}$ and

$\tilde{\Psi}$ are constrained to be lower triangular block toeplitz matrices of unknown parameters (cf. (5) and (6)). Note that $\tilde{\Psi}$ contains zeros on the block diagonal while the block diagonal of $\tilde{\Phi}$ may contain the **D** matrix if it is to be estimated.

The block toeplitz constraints are linear and can easily be imposed (cf. (Peternell *et al.*, 1996)). Let $\widehat{\Gamma\mathscr{K}}$, $\hat{\Phi}$ and $\hat{\Psi}$ denote the solutions to the above linear regression problem.

The next step is to utilize $\hat{\Phi}$ and $\hat{\Psi}$ in an attempt to remove the effects of the future inputs and "pre-whiten" the future outputs by forming

$$\mathbf{z}(t) \triangleq \mathbf{y}_f(t) - \hat{\Phi}\mathbf{u}_f - \hat{\Psi}\mathbf{y}_f(t) \approx \Gamma\mathscr{K}\mathbf{p}(t) + \mathbf{e}_f(t). \quad (9)$$

This equation can be viewed as a reduced rank linear regression problem in $\Gamma\mathscr{K}$. Estimates of Γ and $\mathscr{K}$ can be obtained by performing a canonical correlation analysis on $\mathbf{z}(t)$ and $\mathbf{p}(t)$ as follows: Let

$$\mathbf{M} = (\hat{\mathbf{R}}_{\mathbf{zz}})^{-1/2}(\hat{\mathbf{R}}_{\mathbf{zp}})(\hat{\mathbf{R}}_{\mathbf{pp}})^{-1/2}$$

where we have defined the sample correlation matrix between two signals $\mathbf{z}(t)$ and $\mathbf{p}(t)$ as

$$\hat{\mathbf{R}}_{\mathbf{zp}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{z}(t)\mathbf{p}^T(t).$$

Next, compute the singular value decomposition

$$\mathbf{USV}^T = \mathbf{M}$$

where **U** and **V** are orthonormal matrices of left and right singular vectors, respectively, and **S** is a diagonal matrix containing the singular values in nonincreasing order along the diagonal. Let $\mathbf{U}_n\mathbf{S}_n\mathbf{V}_n^T$ denote the partitioning of the SVD matrices corresponding to the n largest singular values. The estimates of Γ and $\mathscr{K}$ are then given by

$$\hat{\Gamma} = (\hat{\mathbf{R}}_{\mathbf{zz}})^{1/2}\mathbf{U}_n\mathbf{S}_n \quad (10)$$

$$\hat{\mathscr{K}} = \mathbf{V}_n^T(\hat{\mathbf{R}}_{\mathbf{pp}})^{-1/2} \quad (11)$$

This ends the second step of the proposed method. The idea of this step was to get initial estimates of Γ and $\mathscr{K}$. This is also a step where the system order can be estimated if it is not known a priori.

In the third step, we return to the regression in (8). Now (8) is considered to be a reduced-rank regression problem with additional block toeplitz terms. To solve this we will use the idea of the "one step correction" method of (Werner and Jansson, 2004). In this approach, one linearizes around the initial estimate of $\Gamma\mathscr{K}$ and minimizes with respect to the first order correction terms of $\hat{\Gamma}$ and $\hat{\mathscr{K}}$ along with the block toeplitz parameters. This can be shown to be identical to taking a newton step from the initial estimates of $\hat{\Gamma}$ and $\hat{\mathscr{K}}$ (Werner and Jansson, 2004).

From the solution of the third step we get an updated estimate of $\mathscr{K}$ which is denoted $\hat{\hat{\mathscr{K}}}$. The corresponding estimated state sequence is

$$\hat{\mathbf{x}}(t) = \hat{\hat{\mathscr{K}}}\mathbf{p}(t). \quad (12)$$

The system matrices can now be estimated by linear regression in the state space model equations (1) by replacing the true state with the estimate (12). Estimates of **C** and **D** are obtained by regressing $\mathbf{y}(t)$ on $\hat{\mathbf{x}}(t)$ and $\mathbf{u}(t)$. An estimate of the innovation sequence is then obtained as the residual of that regression. Finally, the matrices **A**,**B** and **K** are estimated by regressing $\hat{\mathbf{x}}(t+1)$ on $\hat{\mathbf{x}}(t)$, $\mathbf{u}(t)$, and the estimated innovation sequence (cf. the CCA method (Peternell *et al.*, 1996; Larimore, 1983)).

We finally note that in the multiple output case, it may be useful to include a weighting in the regression problems above to pre-whiten the prediction errors in $\mathbf{e}_f(t)$.

4. SIMULATION EXAMPLE

To study the performance of the suggested approach, a simulation is performed using a similar example as was used in (Ljung and McKelvey, 1996). This example considers the identification of the following system based on data collected both in open and closed loop:

$$y(t) = \frac{0.21q^{-1} + 0.07q^{-2}}{1 - 0.6q^{-1} + 0.8q^{-2}}u(t) + \frac{1}{1 - 0.98q^{-1}}e(t).$$

Here, $e(t)$ is zero mean white Gaussian noise with variance 4. In the open loop simulation, the input $u(t)$ is zero mean Gaussian white noise with unit variance. In the closed loop case, the input is given by

$$u(t) = r(t) - y(t)$$

where the external input $r(t)$ is white Gaussian noise with unit variance.

In both the open and closed loop simulations, 50 independent Monte Carlo experiments are used. In each simulation run, third order state space models are estimated by different identification methods based on $N = 3000$ samples of $u(t)$ and $y(t)$. In addition to the proposed method of this paper (marked as NEW in the plots), the studied identification methods include CCA (Larimore, 1983; Peternell *et al.*, 1996), Parsim_e (Qin and Ljung, 2003), SSARX (Jansson, 2003), SSNEW (Ljung and McKelvey, 1996), and the prediction error method (Ljung, 1987). The **D** matrix is constrained to be zero in all methods and the indices $f = p = 10$ for all subspace methods. The Figures 1 and 2 show the average and the root mean square errors (RMSEs) of the estimated magnitude of the transfer function from $u(t)$ to $y(t)$ for the open and closed loop simulation, respectively.

For the open loop simulation it can be seen in Figure 1 that CCA, SSARX and the new method perform equally well. Parsim estimates the resonance peak better than the mentioned methods but has a slightly higher estimation error variance in this example. Clearly, SSNEW performs the worst but it can be noted that its performance is comparable to that

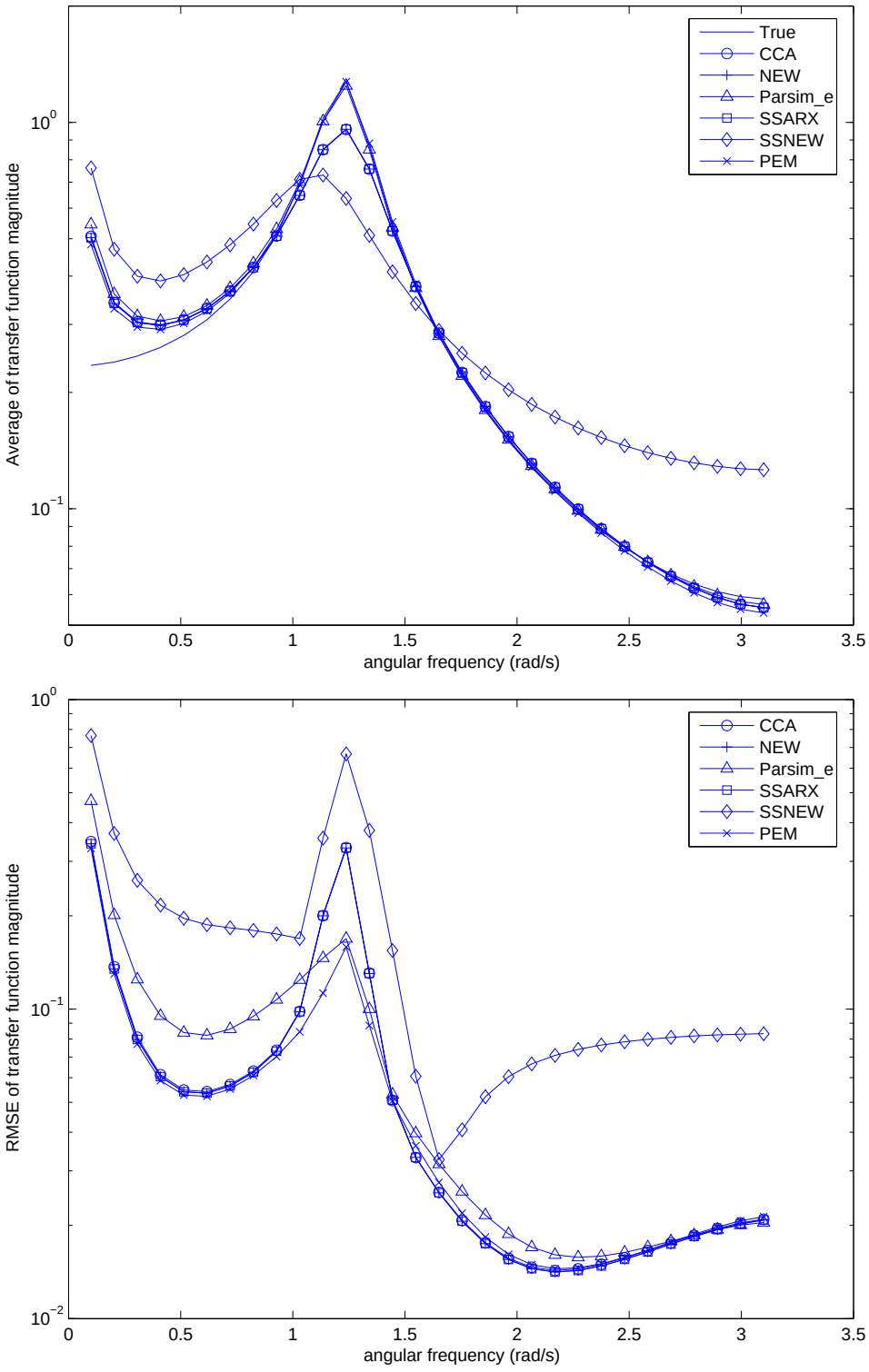

Fig. 1. Open loop simulation: The top graph shows the average of the 50 estimated magnitudes of the transfer function from $u(t)$ to $y(t)$. The bottom graph shows the corresponding RMS errors of the estimated magnitudes. The user defined indices are $f = p = 10$ for all subspace methods.

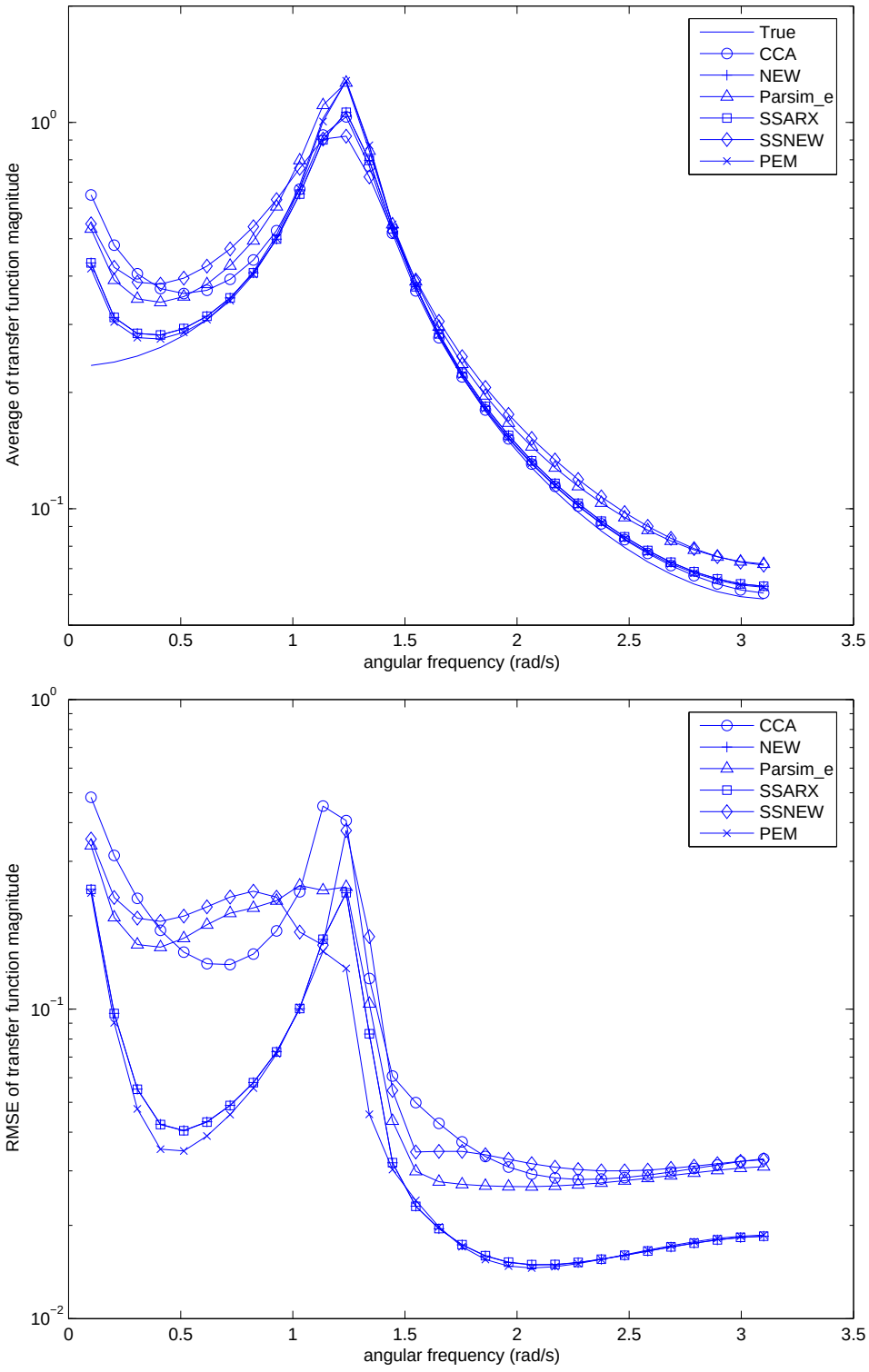

Fig. 2. Closed loop simulation: The top graph shows the average of the 50 estimated magnitudes of the transfer function from $u(t)$ to $y(t)$. The bottom graph shows the corresponding RMS errors of the estimated magnitudes. The user defined indices are $f = p = 10$ for all subspace methods.

of MOESP (Verhaegen, 1994) and N4SID (Van Overschee and De Moor, 1996) (the simulation results of MOESP and N4SID are not included in order not to clutter the figure).

In the closed loop simulation (see Figure 2), the performance of CCA degrades as expected (although it performs reasonably well in this example when $r(t)$ is white). SSARX and the new method still perform almost identical while Parsim and SSNEW have a larger bias and variance.

5. CONCLUSIONS

In this paper we have presented a new subspace identification method which is able to handle data collected both in open and closed loop. The method relies on a slightly reformulated subspace data model compared to previous similar models in the literature. The new model highlights more clearly the connection to prediction error minimization. Based on this formulation, the state space model parameters are estimated by the solution of a sequence of regression problems with linear and rank constraints.

REFERENCES

Bauer, D. (2001). Order estimation for subspace methods. *Automatica* **37**(10), 1561–73.

Chiuso, A. (2004). Asymptotic variance of a certain closed-loop subspace identification method. In: *Proc. of the 43rd IEEE Conf. on Decision and Control*.

Chiuso, A. and G. Picci (2004). Consistency analysis of certain closed-loop subspace identification methods. In: *Proc. of the 43rd IEEE Conf. on Decision and Control*.

Chou, C. T. and M. Verhaegen (1999). Closed-loop identification using canonical correlation analysis. In: *Proc. of the European Control Conference*. Karlsruhe, Germany.

Forssell, U. and L. Ljung (1999). Closed-loop identification revisited. *Automatica* **35**, 1215–1241.

Gustafsson, T. (2001). Subspace identification using instrumental variable techniques. *Automatica* **37**, 2005–2010.

Jansson, M. (2003). Subspace identification and ARX modeling. In: *Proc. 13th IFAC Symp. on System Identification*. Rotterdam, The Netherlands.

Larimore, W. E. (1983). System identification, reduced-order filtering and modeling via canonical variate analysis. In: *Proc. ACC*. pp. 445–451.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall. Englewood Cliffs, NJ. 2nd ed. 1999.

Ljung, L. and T. McKelvey (1996). Subspace identification from closed loop data. *Signal Processing* **52**(2), 209–216.

Peternell, K., W. Scherrer and M. Deistler (1996). Statistical analysis of novel subspace identification methods. *Signal Processing* **52**(2), 161–177.

Qin, S. J. and Lennart Ljung (2003). Closed-loop subsapce identification with innovation estimation. In: *Proc. 13th IFAC Symposium on System Identification*. Rotterdam, The Netherlands. pp. 887–892.

Van Overschee, P. and B. De Moor (1996). *Subspace Identification for Linear Systems: Theory–Implementation–Applications*. Kluwer Academic Publishers.

Van Overschee, P. and B. De Moor (1997). Closed loop subspace system identification. In: *Proc. Conference on Decision and Control*. New York, USA.

Verhaegen, M. (1993). Application of a subspace model identification technique to identify LTI systems operating in closed-loop. *Automatica* **29**(4), 1027–40.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica* **30**(1), 61–74.

Werner, K. and M. Jansson (2004). Weighted low rank approximation and reduced rank linear regression. In: *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*. Vol. 2. Montreal, Canada. pp. 501–4.