Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-26 (2018) 74–80

A Rank Minimization Formulation for Identification of Linear Parameter Varying Models ⋆

Rajiv Singh * **Mario Sznaier** ** **Lennart Ljung** ***

* *The MathWorks, Inc., 1 Apple Hill Drive, Natick, MA 01760 USA (e-mail: rajiv.singh@mathworks.com)*
** *ECE Dept., Northeastern University, Boston, MA 02115 USA (e-mail: msznaier@coe.neu.edu)*
*** *Div. of Automatic Control, Linköping University, Sweden (e-mail: ljung@isy.liu.se)*

Abstract: We explore the problem of identification of LPV models when the scheduling variables are not known in advance and the model parameters exhibit a dynamic dependence on them. We consider an affine ARX model structure whose parameters vary with time. We solve for the model's parameters and scheduling variables in two steps. In the first step, we use the measured input-output data to realize a parameter trajectory by solving a regularized Hankel matrix rank minimization problem. The regularization penalty is guided by the prior knowledge regarding the nature of system's time variation. In the second step, the scheduling variables are estimated as parameters of a sparse ARX structure relating the model's parameters to the measured input-output variables. The effectiveness of the proposed approach is illustrated with two practical examples.

Keywords: LPV, subspace, convex optimization, rank minimization, system identification, bilinear, LTV

1. INTRODUCTION

A common description of dynamic phenomena is a nonlinear state-space model subject to inputs u(t), containing states x(t) and generating outputs y(t). The set of inputs and states define its operating conditions. When this model is linearized, the resulting linear models are dependent upon the operating points about which the linearization occurred. In a linear parameter varying (LPV) modeling approach, the dependence of the linear model on its operating point is projected into a lower dimensional space called the scheduling space. This projection is guided by physical intuition, or an analysis of the manner in which the inputs and states affect the output. This task is not always apparent and bad assumptions, such as choosing scheduling variables based on data acquisition convenience, may cause loss of fidelity in capturing the observed behavior. Hence it is imperative to consider if scheduling information can be extracted from the data along with the model coefficients, while judiciously using prior knowledge about the system's behavior.

Extraction of switching regimes or schedule has been considered in the context of piecewise affine model identification by Sznaier and Bemporad, see Ozay et al. (2012), and Breschi et al. (2016) among others. The scheduling variables are usually the entire regressor set that is used to define the separating hyperplanes. This form does not capture the case where local dynamics are inherently time-varying, for example, the flight dynamics of a rocket that is continuously losing mass in addition to experiencing discrete events where an entire propulsion stage is dropped. One can also consider a linear/affine model structure whose coefficients are continuous functions of certain scheduling variables. A popular class of such models are bilinear in structure and are typically identified using a subspace approach. These approaches rely on extension of linear system concepts of reachability and controllability in order to express the model's states as a linear combination of certain basis functions. Such approaches are usually computationally expensive since the block sizes of the data matrices grow exponentially with length of data (van Wingerden and Verhaegen (2008),Verdult and Verhaegen (2002)). They also often impose a restriction on the input (Favoreel et al. (1999)) or scheduling (Felici et al. (2007)) to make the problem tractable. For a comprehensive summary of LPV identification approaches, see Toth (2010) and dos Santos et al. (2012).

The goal of this paper is to provide a computationally efficient algorithm for identifying linear input-output models with time-varying parameters from experimental data and some minimal a-priori information about the model structure. As shown in the paper, this can be accomplished by following a two step procedure: In the first step, a suitable parameter trajectory is obtained by solving a regularized Hankel matrix rank minimization problem. In the second step, the scheduling variables are estimated as

⋆ This work was supported in part by NSF grants ECCS–1404163, CNS–1646121 and CMMI–1638234; AFOSR grant FA9550-12-1-0271; and the Alert DHS Center of Excellence under Award Number 2008-ST-061-ED0001.

10.1016/j.ifacol.2018.11.165

parameters of a sparse ARX structure relating the model's parameters to the measured input-output variables. The paper is organized as follows: Section 2 introduces the model structure under consideration and formally states the problem of interest. Section 3 presents the proposed algorithm and illustrates the ideas behind it by analyzing some simple cases. Section 4 illustrates the effectiveness of the proposed method with two practical examples. Finally, Section 5 present some conclusions and points out to some open questions.

2. MODEL STRUCTURE

In this paper we consider the input-output model form with time-varying parameters. The dependence of parameters on scheduling variables is not known in advance. The model structure chosen for analysis is an ARX structure:

$$A(t, q)y(t) = B(t, q)u(t) + e(t) \tag{1}$$

where $A(t, q)$ and $B(t, q)$ are time-varying polynomials in delay operator q^{-1}:

$$\begin{aligned} A(t, q) &= 1 + a_1(t)q^{-1} + a_2(t)q^{-2} + \ldots + a_{na}(t)q^{-na} \\ B(t, q) &= b_1(t)q^{-nk} + b_2(t)q^{-nk-1} + \ldots + b_{nb}(t)q^{-nb-nk+1} \end{aligned} \tag{2}$$

nk denotes input-to-output lag which can be zero. The model's parameter vector is:

$$\Theta(t) = [a_1(t), a_2(t), \ldots, a_{na}(t), b_1(t), \ldots, b_{nb}(t)]^T \tag{3}$$

The model structure can also be written as:

$$y(t) = \Theta(t)^T \Phi(t) + e(t) \tag{4}$$

where $\Phi(t)$ is the vector of model's regressors composed of lagged input-output variables. The length of $\Phi(t)$ is $n = na + nb$. The models parameters $\Theta(t)$ are assumed to evolve according to an affine auto-regressive process driven by the "inputs" $u(t)$ and $y(t)$:

$$F(q)(\Theta(t) - \bar{\Theta}) = G_1(q)u(t) + G_2(q)y(t) \tag{5}$$

where $F(q)$, $G_1(q)$, $G_2(q)$ are constant-coefficient polynomials of arbitrary orders and $\bar{\Theta}$ is the affine term. $G_2(q)$'s leading coefficient is zero so that there is at least one sample lag in contribution of $y(t)$. The free entries of $\bar{\Theta}$, F, G_1 and G_2 can be thought of as original model's hyper-parameters. Equation (5) allows a rational dependence of model's parameters on the system's states and inputs. Note that this form of parameter representation makes the model essentially a bilinear structure. Such forms are appealing candidates for modeling many nonlinear processes such as those arising in the areas of fMRI deconvolution and nonlinear tracking, see, for example, Bruni et al. (1974), Penny et al. (2005), and Priestley (1991). In this context, the problem of interest in this paper can be precisely stated as:

Problem 1. Given input/output data, and a-priori bounds on na, nb, nk and $\|e\|$, find a model of the form (4)-(5) that explains the observed data within the approximation error bounds.

Remark 1. Note that the problem above is ill posed, since typically there are multiple parameterizations that can generate the experimental data. Several regularizations that exploit additional information to remove this ambiguity will be discussed in Section 3.

The approach pursued in this paper to solve Problem 1 is to first realize the $\Theta(t)$ trajectory of Equation (4) under suitable constraints. Then use the estimated $\Theta(t)$ and the input-output data measurements to estimate the values of $\bar{\Theta}$, $F(q)$, $G_1(q)$ and $G_2(q)$ coefficients. This would deliver $\Theta(t)$ expressed as a function of model regressors in a rational form. If F, G_1 and G_2 are *sufficiently sparse*, we can treat the contributing regressors as scheduling variables.

3. RANK MINIMIZATION FORMULATION FOR ESTIMATING PARAMETER TRAJECTORY

Suppose $u(t)$ and $y(t)$ are uniformly sampled and N measurements for $t = 1, 2, \ldots, N$ are available. Consider a state-space realization of the $\Theta(t)$ dynamics in Equation (5):

$$\begin{aligned} X(t+1) &= A_\theta X(t) + B_\theta^1 u(t) + B_\theta^2 y(t) \\ \Theta(t) &= C_\theta X(t) + D_\theta^1 u(t) + \bar{\Theta} \end{aligned} \tag{6}$$

Let $U(t) = [u(t), y(t), 1(t)]^T$ be the augmented input vector of length $p = ny+nu+1$, where the step input $1(t)$ is added to account for the affine term $\bar{\Theta}$. Then as described in subspace identification literature (Overschee and Moor (1994)), the minimal order of parameter model (Equation (6)) is equal to the rank of the matrix $H_{n,m,N}(\Theta)H_U^\perp$ where:

$$H_{n,m,N}(\Theta) = \begin{bmatrix} \Theta(1) & \Theta(2) & \ldots & \Theta(N-m+1) \\ \Theta(2) & \Theta(3) & \ldots & \Theta(N-m+2) \\ \vdots & \vdots & \ddots & \vdots \\ \Theta(m) & \Theta(m+1) & \ldots & \Theta(N) \end{bmatrix} \tag{7}$$

$H_U^\perp \in \Re^{(N-m+1)\times q}$ is a matrix whose columns form an orthogonal basis for the null space (nullity q) of the Hankel matrix $H_{p,m,N}(U)$:

$$H_{p,m,N}(U) = \begin{bmatrix} U(1) & U(2) & \ldots & U(N-m+1) \\ U(2) & U(3) & \ldots & U(N-m+2) \\ \vdots & \vdots & \ddots & \vdots \\ U(m) & U(m+1) & \ldots & U(N) \end{bmatrix} \tag{8}$$

From the discussion above, it follows that the parameterization that explains the observed data with the lowest order model for the evolution of the parameter Θ can be found by minimizing the rank of $H_{n,m,N}(\Theta)H_U^\perp$ subject to the constraint that the final model should interpolate the observed data up to some model error.

$$\begin{aligned} \underset{\Theta}{\text{minimize}} \quad & \text{rank}(H_{n,m,N}(\Theta)H_U^\perp) \\ \text{subject to} \quad & \\ & \left\|y(t) - \Theta(t)^T\Phi(t)\right\| \le \delta_1 \\ & \textit{plus additional constraints} \end{aligned} \tag{9}$$

where the minimization is over the entire $\Theta(t)$ sequence of N samples. δ_1 is a measure of maximum output disturbance. Since rank minimization is computationally NP-hard, a convex relaxation of the problem above is obtained by using an iteratively re-weighted trace minimization heuristics (Mohan and Fazel (2010), Sznaier et al. (2014b)), summarized in Algorithm 1. The choice of additional constraints reflects our prior knowledge about the system behavior. For example, we can impose constraints that $\Theta(t)$ changes more slowly than the output of $y(t)$ of the model, the changes are "smooth", and/or limited to a 1-norm ball with unknown center. Finally, once a trajectory for $\Theta(t)$ is obtained, a linear (affine) model is fit to it using the standard subspace approach.

Algorithm 1 Reweighted $\|.\|_*$ based rank minimization

Initialize: $k = 0, \mathbf{W}_y(0) = \mathbf{I}, \mathbf{W}_z(0) = \mathbf{I}$, δ_o small
repeat
Solve

$$\min_{\mathbf{X}^{(k)},\mathbf{Y}^{(k)},\mathbf{Z}^{(k)}} \text{Trace} \begin{bmatrix} \mathbf{W}_y^{(k)}\mathbf{Y}^{(k)} & 0 \\ 0 & \mathbf{W}_z^{(k)}\mathbf{Z}^{(k)} \end{bmatrix}$$

$$\text{subject to: } \begin{bmatrix} \mathbf{Y}^{(k)} & \mathbf{\Theta}^{(k)} \\ \mathbf{\Theta^T}^{(k)} & \mathbf{Z}^{(k)} \end{bmatrix} \succeq 0$$

$$\mathbf{\Theta}^{(k)} \in \mathcal{S}$$

where $\mathcal{S}$ is the feasible set in (9).
Decompose $\mathbf{\Theta}^{(k)} = \mathbf{UDV}^T$.
Set $\delta \leftarrow \min[\text{diag}(\mathbf{D})] + \delta_0$.
Set $\mathbf{W}_y^{(k+1)} \leftarrow \left(\mathbf{Y}^{(k)} + \delta\mathbf{I}\right)^{-1}$
Set $\mathbf{W}_z^{(k+1)} \leftarrow \left(\mathbf{Z}^{(k)} + \delta\mathbf{I}\right)^{-1}$
Set $k \leftarrow k + 1$.
until a convergence criterion is reached. **return** $\mathbf{\Theta}^{(k)}$

Note in passing, that as expected, in the case of a linear time-invariant process (that is, A, B are constant coefficient polynomials in Equation (1)), the algorithm above yields a constant value for Θ. This is due to the fact that the minimal value of $\|H_{n,m,N}(\Theta)H_U^\perp\|_*$ is zero, and $D_\theta^1 \equiv 0$. To see this, note that $H_{p,m,N}(U)$ has m rows of all ones owing to the $\bar{\Theta}$ term. Since $H_{p,m,N}(U)H_U^\perp = 0$, the columns of $H_U^\perp$ must sum to zero. Since $H_{n,m,N}(\Theta) = [I, I, \ldots I]^T \times \bar{\Theta} \times [1, 1, \ldots 1]$ in the LTI case, we have $H_{n,m,N}(\Theta)H_U^\perp = 0$.

Similarly, for time-varying processes where the parameters $\Theta(t)$ show a static dependence on the inputs $u(t)$, the minimal value of $\|H_{n,m,N}(\Theta)H_U^\perp\|_*$ is zero but $D_\theta^1 \neq 0$. This is also easily seen by taking $\Theta(t) = KU(t)$, where K is an $n \times p$ matrix. Then $H_{n,m,N}(\Theta) = \text{block-diag}(K)H_{p,m,N}(U)$ and zero value of $\|H_{n,m,N}(\Theta)H_U^\perp\|_*$ follows from the orthogonality of $H_U^\perp$ to $H_{p,m,N}(U)$.

3.1 Additional Regularization Constraints

As noted in Section 2, Problem 1 is typically ill-posed, since it admits multiple solutions. These additional solutions can be ruled out by imposing "regularization constraints" that reflects our prior knowledge about the system behavior. Examples of additional information that can be captured by these constraints include:

(1) Only a few of the parameters in the Θ vector vary with time.
(2) The parameters vary either intermittently or slowly relative to the rate of change of model's states and outputs.

In the sequel we indicate how to incorporate these constraints into the proposed formulation and illustrate the use of the resulting framework for solving different types of parameter-varying identification problems.

3.2 LTI Systems

The approach is first applied to the well-known case of linear time-invariant (LTI) systems (Overschee and Moor (1994), Liu and Vandenberghe (2009), Fazel et al. (2013), Sznaier et al. (2014a)) as a simple proof of concept. The approach is also used to fit the dynamic model (Equation (6)) to the $\Theta(t)$ trajectory.

For a linear time-invariant system with an offset term, a large L1-penalty on the change in parameters is imposed, so that the objective is:

$$\begin{aligned} &\underset{\Theta}{\text{minimize}} \quad \left\|H_{n,m,N}(\Theta)H_U^\perp\right\|_* + \lambda \left\|\Delta\Theta\right\|_1 \\ &\text{subject to} \\ &\qquad \left\|y(t) - \Theta^T\Phi(t)\right\| \leq \delta_1 \end{aligned} \tag{10}$$

where $\Delta\Theta \in \Re^n$. $\Delta\Theta(i)$ is the maximum change in parameter $\Theta(i)$ between successive time samples, over the available data's time span. Another formulation is to treat the constraint on prediction error as a penalty in the objective, which yields an elastic-net type of structure:

$$\begin{aligned} \underset{\Theta}{\text{minimize}} \quad &\left\|H_{n,m,N}(\Theta)H_U^\perp\right\|_* + \lambda_1 \left\|\Delta\Theta\right\|_1 + \\ &\lambda_2 \left\|y(t) - \Theta^T\Phi(t)\right\| \end{aligned} \tag{11}$$

where λ_1 and λ_2 are regularization constants. λ_1 is typically large.

Sparsity Constraints for Parameter Dynamics Our objective is to describe $\Theta(t)$ using as few past values of $u(t)$ and $y(t)$ in linear (affine) equation (6). The minimal order of the dynamics are determined by solving Equation (9) which addresses the main sparsity consideration regarding use of as few regressors as possible. The only additional consideration is the choice of a minimal subset of input/output variables $(u(t), y(t))$, that is, making B_θ^1, B_θ^2 and D_θ^1 column-sparse.

A more direct way of inducing sparsity in the linear model for $\Theta(t)$ is to use an ARX parameterization obtained by a matrix fraction description (MFD) of the dynamics. Suppose nx denotes the estimated rank of $H_{n,m,N}(\Theta)H_U^\perp$. Then:

$$A(q)\Theta(t) = B_1(q)u(t) + B_2(q)y(t) + B_0 1(t) + e(t) \tag{12}$$

A(q) is a diagonal polynomial matrix of auto-regressive terms for each Θ variable, each diagonal of order $\leq nx$. B_0 is an n-by-1 constant vector to account for the affine term and $B_1(q)$, $B_2(q)$ are vectors of nx-order polynomials; the leading coefficients of $B_2(q)$ are zero. The polynomial coefficients can be determined by solving:

$$\underset{\beta}{\text{minimize}} \quad \left\|vec(\Theta(t)) - \beta^T\hat{\Phi}(t)\right\| + L\left\|\beta\right\|_1 \tag{13}$$

where $\beta = vec(A, B_0, B_1, B_2)$ is the vector of polynomial coefficients and $\hat{\Phi}(t)$ is the regressor matrix. $L\left\|\beta\right\|_1$ is the regularizing penalty. If the goal is to reduce the number of input/output variables from the set $u(t), y(t)$ participating in the model, then the penalty can be changed to a Group Lasso one:

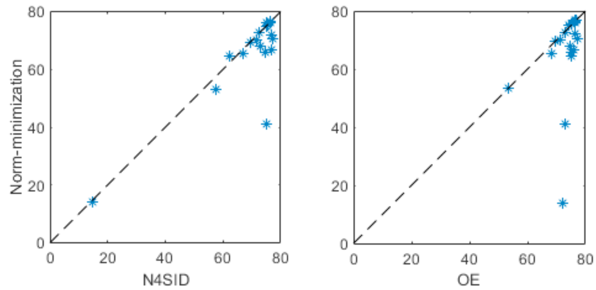

Fig. 1. Linear system results: Scatter plot of Norm-minimization vs. N4SID, OE validation fit values for linear Output-error model identification.

$$\underset{\beta}{\text{minimize}} \left\| vec(\Theta(t)) - \beta^T \hat{\Phi}(t) \right\| + L \left\| vec(\|\beta_A\|_\infty, \|\beta_{B_0}\|_\infty, \|\beta_{B_1}\|_\infty, \|\beta_{B_2}\|_\infty) \right\|_1 \quad (14)$$

where β_A is the vector of coefficients of the $A(q)$ polynomial etc.

As an example, consider 20 random realizations of a second order transfer function simulated using a PRBS signal. Gaussian white noise with SNR = 12dB is added to the simulated outputs. Each record contains N=100 samples. The results from the proposed rank-minimization approach are generated using Equations (10) and (13). For comparison, standard subspace (N4SID) approach and Output Error (OE) estimation in the MATLAB® System Identification Toolbox™ [Ljung (2017)] are also generated. All the models are validated using an independent dataset. The fit metrics are computed using a Normalized Root Mean Squared (NRMSE) goodness of fit metric, expressed as percentage. The results are shown in Figure 1. It is observed that the rank-minimization approach is able to meet the performance of the standard linear identification software results.

3.3 Bilinear Systems

The proposed representation of $\Theta(t)$ dynamics means that we essentially have a bilinear system with terms composed of lagged input-output variables. As an example, consider the system:

$$A(q)y(t) = B_0 u(t) F(u(t)) + B_1(q)u(t) + e(t) \quad (15)$$

where $A(q)$ and $B_1(q)$ are fixed coefficient polynomials, B_0 is a constant and $F(.)$ is a low-pass filter. This can be expressed in LPV form:

$$A(q)y(t) = B(q,t)u(t) + e(t) \quad (16)$$

where $B(q,t)$ is a time-varying polynomial. For example, if $A(q)$ and $B_1(q)$ are second-order, $B_1(q)$ has no feedthrough term and $F(.)$ is a third-order moving average filter, then $B(q,t) = [b_0(t), b_1 q^{-1}, b_2 q^{-2}]$, $b_0(t) = \sum_{i=0}^{2} \alpha_i u(t-i)$. We then have a second order ARX model with a time-varying input gain $b_0(t)$. The objective function is:

$$\begin{aligned} &\underset{B,\Theta}{\text{minimize}} \left\| H_{n,m,N}(\Theta) H_U^{\perp} \right\|_* \\ &\text{subject to} \\ &\left\| y(t) - \Theta^T \Phi(t) \right\| \leq \delta_1 \\ &\left\| \Delta\Theta_{max} \right\|_1 \leq \delta_2 \end{aligned} \quad (17)$$

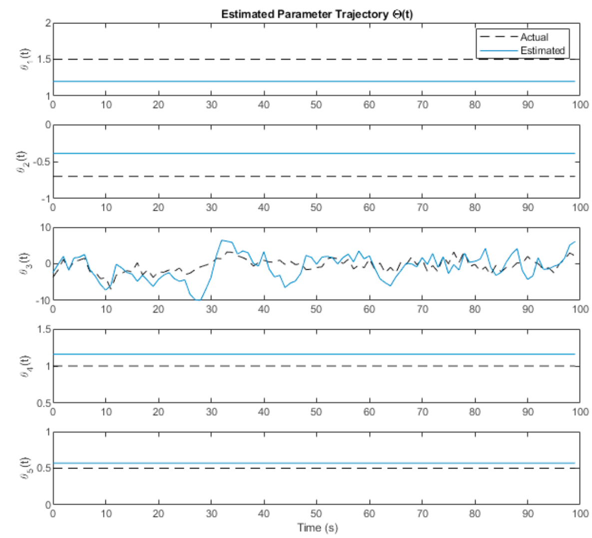

Fig. 2. Estimated $\Theta(t)$ trajectory compared against the true values.

The first constraint checks prediction error, the second imposes a limit on the rate of change of parameters $\Theta(t)$. Here $\Delta\Theta_{max}$ is a vector of maximum allowable parameter changes, such that for the i^{th} parameter, the value is $\max_t(|\theta_i(t+1) - \theta_i(t)|)$.

A simulation using low-pass filtered input sequence was performed of a second order polynomial model containing a bilinear term such that $F(.)$ is a fourth order FIR filter. Gaussian white noise of 14dB SNR was added to the simulated output. The resulting single-input, single-output data was split into three portions. The first portion was used as the main estimation data. The second portion was used in a cross-validation test to determine a good value for δ_1. The third portion was reserved for model validation. The value of δ_2 was chosen such that it was the minimum value for which the estimation problem was feasible (1e-4 here). A time-varying ARX model of order $na = 2$, $nb = 3$, $nk = 0$ was fit to the estimation data. The estimated parameter trajectory compared against their true values is shown in Figure 2. The singular values of $H(\Theta)H_U^{\perp}$ are shown in Figure 3.

The SVD plot shows 5 significant singular values. Hence a fifth order ARX model with L1-penalty was fit to the parameter trajectory. The resulting hyper-parameters were used to simulate the model response to the validation data input with zero initial conditions. If the nature of the time variation of the parameter $b_0(t)$ is known in advance, the model parameters can also be determined by a linear ARX estimation by treating each bilinear term as a known input. This can be treated as "Oracle" result for comparing the limit of performance of the proposed estimation algorithm. A typical fit to the validation data is shown in Figure 4.

50 similar experiments were performed for this system for SNRs around 10dB, 20 dB and 30dB. The NRMSE fit value ranges and the corresponding "Oracle" results are shown in the box plot of Figure 5.

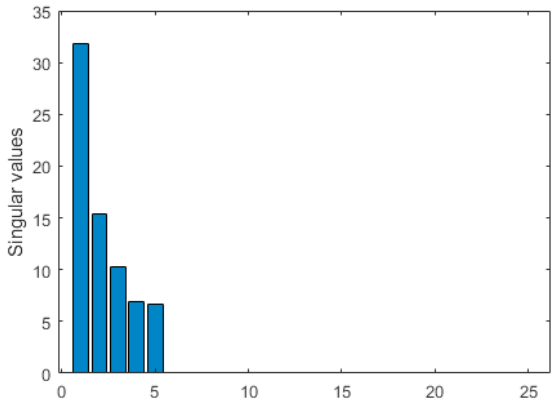

Fig. 3. Estimated singular values for the bilinear identification problem. First five singular values are found to be significant.

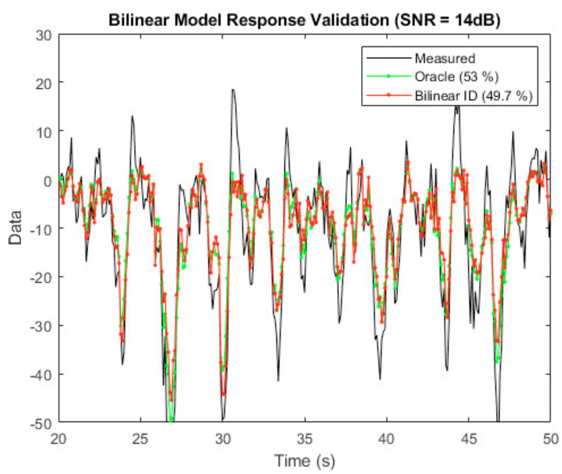

Fig. 4. Bilinear model response validation. Shown are the measured data, "Oracle" model response and Bilinear identification by the proposed rank minimization approach.

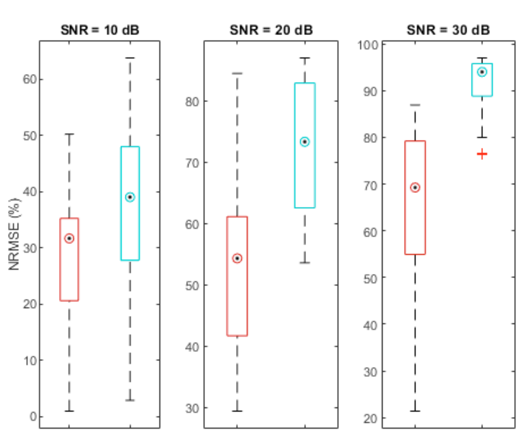

Fig. 5. NRMSE fit distribution for the bilinear identified model (red) and the linear "Oracle" model (cyan).

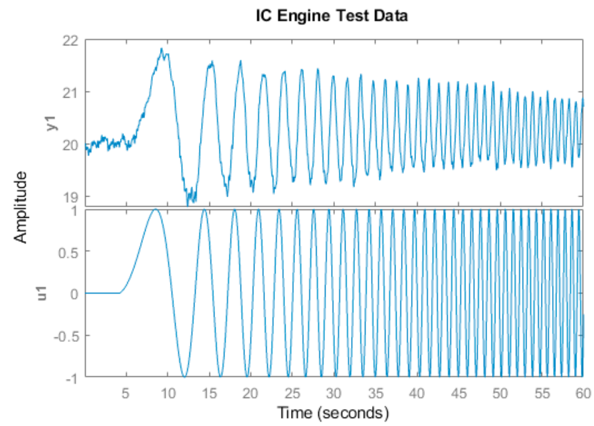

Fig. 6. IC Engine input-output data. y1 is the output (RPM/100) and u1 is the input (V).

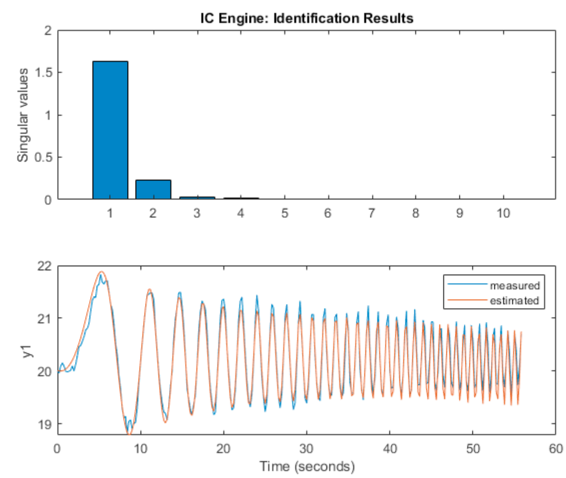

Fig. 7. Fit results for the engine RPM. Top plot shows the singular values while the bottom plot shows fit to the output.

4. EXAMPLES

In this section we illustrate the potential of the proposed approach with two practical examples:

4.1 IC Engine Dynamics

The dynamic relationship between the voltage controlling the Bypass Idle Air Valve (BPAV) and the engine speed is known to be nonlinear. 1500 samples of the control voltage V and the engine speed (RPM/100) were collected at a sampling rate of 0.04 seconds. The measured input-output data is shown in figure 6.

Orders $na = 4$, $nb = 2$, $nk = 2$ were used for estimation using the first half of the measured data downsampled by a factor of five. The prior knowledge is that the nonlinearity is usually of low-order polynomial in nature. Maximum rate of change and smoothness constraints were imposed. Regularization constants were tweaked until a low-order $\Theta(t)$) model that was able to provide good fit to the estimation data was obtained. The fit to the whole dataset is shown in figure 7.

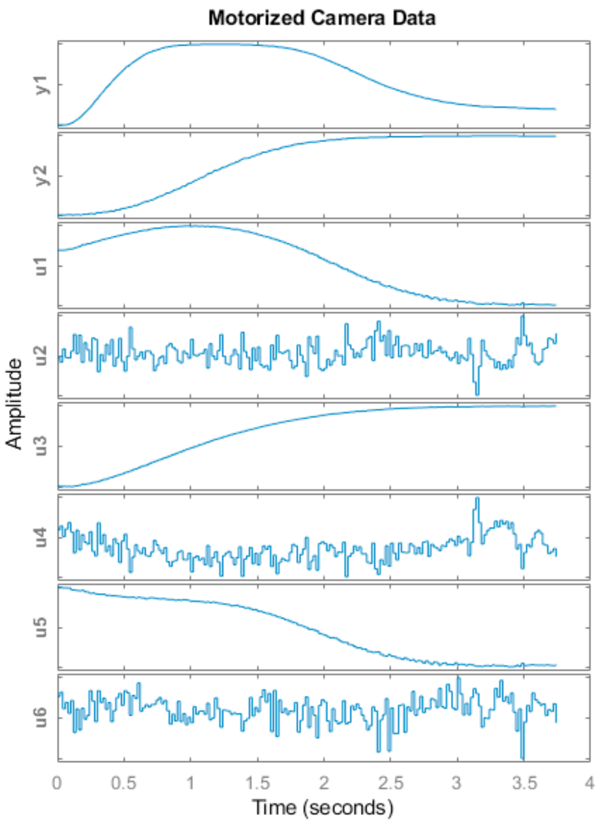

Fig. 8. Motorized camera input-output data. $y1, y2$ are the outputs, $u1, u2, \ldots, u6$ are the inputs.

4.2 Motorized Camera Dynamics

This example shows the identification of a multi-input, multi-output (MIMO) system describing a motorized camera. The input vector u(t) is composed of 6 variables: the 3 translation velocity components in the orthogonal X-Y-Z coordinate system fixed to the camera (m/s), and the 3 rotation velocity components around the X-Y-Z axis (rad/s). The output vector y(t) contains 2 variables: the position (in pixel) of a point which is the image taken by the camera of a fixed point in the 3D space. See figure 8.

The model orders were guided by results from using linear ARX identification, which even though suboptimal, provides a good estimate of the number of regressors involved. A constant input was added to account for offset. The orders chosen were: $na = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, $nb = \left(\begin{smallmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \end{smallmatrix}\right)$ and $nk = \left(\begin{smallmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{smallmatrix}\right)$.

The main regularization idea is to force as many parameters to constant values as possible while limiting the rate of change of the variation in the free parameters. The selection of free parameters was made by independent trials on combinations that forced all except a few parameters to be fixed. A first order model was fit separately to each of the free parameters, the choice of orders being guided by the number of non-zero singular values estimated by the rank minimization procedure. The results are shown in Fig. 9.

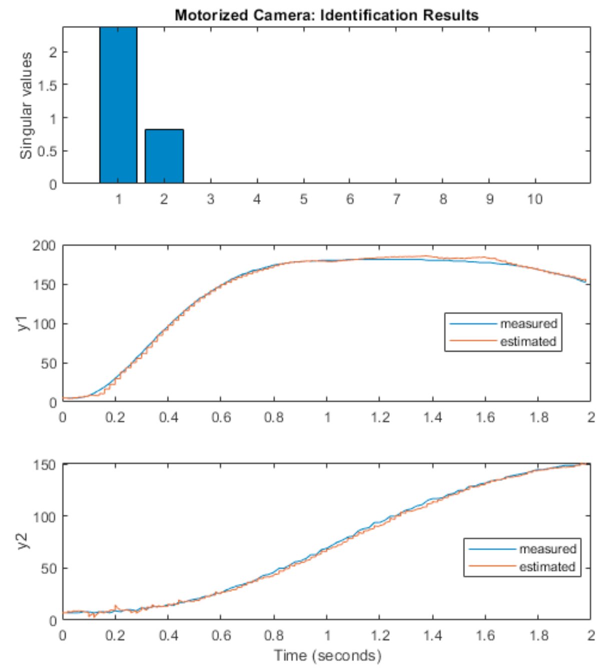

Fig. 9. Fit results for the 2 outputs. Top plot shows the singular values while the second and the third plots show fits to the two outputs.

5. CONCLUSIONS

The well-known rank-minimization approach that is the foundation of linear subspace identification methods is generalized for identification of time-varying linear ARX models. The approach is useful for estimation of LPV models where the choice of scheduling parameters, or the nature of dynamic dependence of the model parameters on them, is not known in advance. It is shown that such an approach offers a feasible framework for analysis of such systems provided the constraints are carefully derived using prior knowledge and applied to the minimization objective.

5.1 Ongoing and Future Work

Setting up the right constraints to ensure identifiability is one of the main difficulties. If we do not fully know apriori what regularization constants to use, we do not know what parameter trajectory $\Theta(t)$ is more reasonable than others. Use of cross-validation to tune regularization constants, exploiting physical knowledge regarding the nature of parameters (in grey box case), and imposing smoothness/sparsity constraints all seem to hold promise. Working with 3 datasets is often helpful - one for parameter estimation, one for tuning regularization constants by cross-validation, and one for independent validation. We are investigating how more information regarding the constraints can be extracted from data itself by means of appropriately designed experiments, or by using a Bayesian estimation framework to turn soft information (such as exponential stability) into concrete constraint information (see, for example, Chen et al. (2012)).

REFERENCES

Breschi, V., Bemporad, A., and Piga, D. (2016). Identification of hybrid and linear parameter varying models via recursive piecewise affine regression and discrimination. In *2016 European Control Conference (ECC)*, 2632–2637.

Bruni, C., DiPillo, G., and Koch, G. (1974). Bilinear systems: an appealing class of nearly linear systems in theory and applications. *IEEE Transactions on Automatic Control*, 19, 334–348.

Chen, T., Ohlsson, H., and Ljung, L. (2012). On the estimation of transfer functions, regularizations and gaussian processes - revisited. *Automatica*, 48, 1525–1535.

dos Santos, P.L., Perdicoulis, T.P.A., Novara, C., Ramos, J.A., and (Editor), D.E.R. (eds.) (2012). *Linear Parameter-Varying System Identification: New Developments and Trends*. World Scientific.

Favoreel, W., Moor, B.D., and Overschee, P.V. (1999). Subspace identification of bilinear systems subject to white inputs. *IEEE Transactions on Automatic Control*, 44, 1157–1165.

Fazel, M., Pong, T.K., Sun, D., and Tseng, P. (2013). Hankel matrix rank minimization with applications to system identification and realization. *SIAM J. Matrix Anal. & Appl.*, 34, 946–977.

Felici, F., van Wingerden, J.W., and Verhaegen, M. (2007). Subspace identification of MIMO LPV systems using a periodic scheduling sequence. *Automatica*, 43(10), 1684–1697.

Liu, Z. and Vandenberghe, L. (2009). Interior-point method for nuclear norm approximation with application to system identification. *SIAM J. Matrix Anal. & Appl.*, 31, 1235–1256.

Ljung, L. (2017). *System Identification Toolbox for use with MATLAB, Version 9.7*. Natick, MA, USA, 9th edition.

Mohan, K. and Fazel, M. (2010). Reweighted nuclear norm minimization with application to system identification. In *Proc. American Control Conference*.

Overschee, P.V. and Moor, B.D. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30, 75–93.

Ozay, N., Sznaier, M., Lagoa, C.M., and Camps, O.I. (2012). A sparsification approach to set membership identification of switched affine systems. *IEEE Transactions on Automatic Control*, 57(3), 634–648.

Penny, W., Ghahramani, Z., and Friston, K. (2005). Bilinear dynamical systems. *Philosophical Transactions of the Royal Society B*, 360, 983–993.

Priestley, M. (1991). *Non-linear and non-stationary time series analysis*. Academic Press.

Sznaier, M., Ayazoglu, M., and Inanc, T. (2014a). Fast structured nuclear norm minimization with applications to set membership systems identification. *IEEE Transactions on Automatic Control*, 59, 2837–2842.

Sznaier, M., Camps, O., Ozay, N., and Lagoa, C. (2014b). Surviving the upcoming data deluge: A systems and control perspective. *53rd IEEE Conference on Decision and Control*, 59, 1488–1498.

Toth, R. (2010). *Modeling and Identification of Linear Parameter-Varying Systems*. Springer-Verlag Berlin Heidelberg.

van Wingerden, J.W. and Verhaegen, M. (2008). Subspace identification of MIMO LPV systems: The PBSID approach. *47th IEEE Conference on Decision and Control*, 4516–4521.

Verdult, V. and Verhaegen, M. (2002). Subspace identification of multivariable linear parameter - varying systems. *Automatica*, 38(5), 805–814.