Technical report from Automatic Control at Linköpings universitet

# Order and Structural Dependence Selection of LPV-ARX Models Using a Nonnegative Garrote Approach

Roland Tóth, Christian Lyzell, Martin Enqvist, Peter S. C. Heuberger, Paul M. J. Van den Hof

Division of Automatic Control

E-mail: r.toth@tudelft.nl, lyzell@isy.liu.se, maren@isy.liu.se, p.s.c.heuberger@tudelft.nl, p.m.j.vandenhof@tudelft.nl

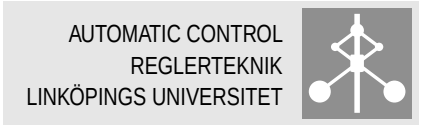
3rd March 2010

Report no.: LiTH-ISY-R-2937

Accepted for publication in 48th IEEE Conference on Decision and Control, Shanghai, China

Address:
Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se



## Abstract

In order to accurately identify Linear Parameter-Varying (LPV) systems, order selection of LPV linear regression models has prime importance. Existing identification approaches in this context suffer from the drawback that a set of functional dependencies needs to be chosen a priori for the parametrization of the model coefficients. However in a black-box setting, it has not been possible so far to decide which functions from a given set are required for the parametrization and which are not. To provide a practical solution, a nonnegative garrote approach is applied. It is shown that using only a measured data record of the plant, both the order selection and the selection of structural coefficient dependence can be solved by the proposed method.

**Keywords:** Linear Parameter-Varying, ARX, identification, order selection

Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009



# Order and Structural Dependence Selection of LPV-ARX Models Using a Nonnegative Garrote Approach

R. Tóth, C. Lyzell, M. Enqvist, P. S. C. Heuberger and P. M. J. Van den Hof

*Abstract*— **In order to accurately identify Linear Parameter-Varying (LPV) systems, order selection of LPV linear regression models has prime importance. Existing identification approaches in this context suffer from the drawback that a set of functional dependencies needs to be chosen a priori for the parametrization of the model coefficients. However in a black-box setting, it has not been possible so far to decide which functions from a given set are required for the parametrization and which are not. To provide a practical solution, a nonnegative garrote approach is applied. It is shown that using only a measured data record of the plant, both the order selection and the selection of structural coefficient dependence can be solved by the proposed method.**

*Index Terms*— **Linear Parameter-Varying, ARX, identification, order selection**

## I. Introduction

Since the introduction of *Linear Parameter-Varying* (LPV) systems in the 1990s, LPV control has rapidly grown into a well established framework with a wide range of applications. The practical use of LPV control design is stimulated by the fact that it extends the results of *Linear Time-Invariant* (LTI) control theory to nonlinear, time-varying plants via *gain scheduling* [1] or by LPV synthesis techniques like $\mu$-synthesis [2] or optimal control [3]. These approaches use LPV models where the signal relations are considered to be linear just as in the LTI case, but the model parameters are assumed to be functions of a measurable time-varying signal, the so-called *scheduling variable* $p : \mathbb{Z} \mapsto \mathbb{P}$. The compact set $\mathbb{P} \subseteq \mathbb{R}^{n_\mathbb{P}}$ denotes the *scheduling space*. Using scheduling variables as changing operating conditions or endogenous/free signals of the plant, the LPV system class can describe both nonlinear and time-varying phenomena.

LPV system identification and modeling have not been able to follow the rapidly advancing control field. Only recently several methods have been proposed, aiming at global identification of discrete-time LPV models from measured data using limited (grey-box) or no structural knowledge (black-box) about the data-generating system. These methods can be categorized mainly based on the used model structures: *State-Space* (SS) methods [4], [5], *Input/Output* (IO) representation based techniques [6], [7], [8] and truncated series expansions based approaches, e.g. *Orthonormal Basis Function* (OBF) techniques [9], [10], [11]. The LPV-IO approaches can be seen as an extension of LTI prediction-error methods. Using the similarities in terms of the model structures and the identification setting, the strength of these approaches lies in the possibility to extend the well-established results of the LTI case. Thus, they offer to solve the LPV identification problem in a simple manner even if transformation of LPV-IO models to LPV-SS descriptions is more complicated than in the LTI case (see [11], [12]). Investigations of the LPV prediction-error setting in terms of experiment design [13], consistency of model estimates [14], and persistency of excitation [15] have only recently appeared, indicating that many important questions still need to be explored. One of these issues concerns the order selection of the LPV-ARX model structure, proposed in [6].

Estimating adequate orders of ARX models is a widely studied topic in the LTI context, see e.g. [16], [17]. It can be seen as the basic step of the model structure selection phase of the identification cycle. Proper selection of the order assures accurate representation of the process dynamics with a limited number of parameters to be estimated. This means not only an adequate complexity of the obtained model, but also a decreased variance of the model estimate. For the same reasons, order selection of LPV-ARX models is also a question of main importance. Additionally, coefficients of these model structures, like $a_i(p(k))$, are often parametrized in the form

$$a_i(p(k)) = \theta_{i0} + \theta_{i1}\psi_{i1}(p(k)) + \ldots + \theta_{i1}\psi_{is_i}(p(k)), \quad (1)$$

where $\theta_{ij} \in \mathbb{R}$ are the unknown parameters and $\{\psi_{ij}\}_{j=1}^{s_i}$ with $\psi_{ij} : \mathbb{P} \mapsto \mathbb{R}$ is a set of a priori chosen functional dependencies. Thus, the set $\{\psi_{ij}\}$ represents an extra freedom in the model structure which needs to be selected a priori. This implies that determining which of these functions are required in the parametrization can be interpreted as an additional selection problem. It must also be noted that in [8], an alternative to the functional dependence estimation has recently been introduced through a "non-parametric" approach which uses dispersion functions instead of (1).

Additionally, it has been shown in [12] that for transformation between LPV-IO and LPV-SS models it is required that the model coefficients not only depend on the instantaneous value of the scheduling variable (*static dependence*) but also on its time shifted versions (*dynamic dependence*). Thus, estimating LPV-ARX models with dynamic dependence is often required for obtaining accurate models of the underly-

R. Tóth, P. S. C. Heuerger and P. M. J. Van den Hof are with the Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands, email: {r.toth,p.s.c.heuberger,p.m.j.vandenhof}@tudelft.nl.

C. Lyzell and M. Enqvist are with the Division of Automatic Control, Linköping Universitet, SE-581 83, Linköping, Sweden, email: {lyzell,maren}@isy.liu.se.

ing system. Selection of the required order of the time-shifts in the scheduling is a parametrization problem and can be interpreted as an adequate selection of $\{\psi_{ij}\}$.

In the LTI literature, recently an order selection method based on a statistical regularization approach, the *Nonnegative Garrote* (NNG) [18], has been proposed for LTI-ARX models [19]. In this method, a natural ordering of model complexity is inflicted to the parameters, which provides the possibility to estimate the order of poles and zeros independently. This yield insight into which parameters are the most important for obtaining a good adaption to the data. Due to the possibility of multidimensional ordering of model complexity, this technique can also be used in the LPV-ARX case to select the order of the input and output side polynomials simultaneously with the ordering of the required structural dependence of the coefficients. In this paper, we aim at this extension of the NNG approach to the LPV case, providing a practically useful tool for LPV-IO approaches.

The paper is organized as follows: In Section II, a short review of the LPV-ARX model structure and its linear-regression based identification method is given, defining the problem setting for order selection. Section III gives an introduction to the NNG approach and presents how its modified form can be used to solve the order selection problem. In Section III-C, an algorithm is proposed to solve the modified NNG problem and in Section IV this algorithm is validated on simulated data. Finally, in Section V, conclusions are drawn and perspectives on future work are given.

## II. LPV IDENTIFICATION VIA ARX MODELS

In this paper we focus on the LPV-ARX model structure, defined in the SISO case as

$$y(k)+\sum_{i=1}^{n_\mathrm{a}} a_i(p(k))y(k-i)=\sum_{j=0}^{n_\mathrm{b}} b_j(p(k))u(k-j)+e(k), \quad (2)$$

where $u$, $y$ and $e$ denote the input, the output, and the noise signals, respectively. Furthermore, the coefficient functions $a_i, b_j : \mathbb{P} \mapsto \mathbb{R}$ have static dependence on $p$. Introduce

$$\begin{bmatrix} \phi_1 & \dots & \phi_{n_\mathrm{g}} \end{bmatrix}^\top \triangleq \begin{bmatrix} a_1 & \dots & a_{n_\mathrm{a}} & b_0 & \dots & b_{n_\mathrm{b}} \end{bmatrix}^\top,$$

with $n_\mathrm{g} \triangleq n_\mathrm{a} + n_\mathrm{b} + 1$. Assume that each function $\phi_i$ is linearly parameterized as

$$\phi_i(\centerdot) = \theta_{i0} + \sum_{j=1}^{s_i} \theta_{ij}\psi_{ij}(\centerdot), \quad (3)$$

where $\{\theta_{ij}\}_{i=1,j=1}^{n_\mathrm{g},s_i}$ are unknown parameters and $\{\psi_{ij}\}_{i=1,j=1}^{n_\mathrm{g},s_i}$ are functions chosen by the user. In this case, (2) can be rewritten as

$$y(k) = \varphi^\top(k)\theta + e(k), \quad (4)$$

where

$$\begin{aligned}\theta &= \begin{bmatrix} \theta_{1,0} & \dots & \theta_{1,s_1} & \theta_{2,0} & \dots \theta_{n_\mathrm{g},s_{n_\mathrm{g}}} \end{bmatrix}^\top \\ \varphi(k) &= \begin{bmatrix} -y(k-1) & -\psi_{11}(p(k))y(k-1) & \dots \\ & -\psi_{1s_1}(p(k))y(k-1) & -y(k-2) & \dots \\ & -\psi_{n_\mathrm{a}s_{n_\mathrm{a}}}(p(k))y(k-n_\mathrm{a}) & u(k) & \dots \end{bmatrix}^\top.\end{aligned}$$

Note that a LPV-FIR model structure or other series expansion types of structures like OBF models can be seen as special cases of the LPV-ARX family with $n_\mathrm{a} = 0$. An additional difference in the OBF case is that instead of time-shifted version of $u$, $\varphi(k)$ is formed from the outputs of a preselected set of LTI-OBF filters applied on $u$ (see [10]).

Given a data set

$$Z^N \triangleq \big(u(k), p(k), y(k)\big)_{k=1}^N, \quad (5)$$

the *least-squares* (LS) parameter estimate for the linear regression model (4) is

$$\hat{\theta}_N \triangleq \underset{\theta\in\mathbb{R}^n}{\arg\min} V_N(\theta, Z^N), \quad (6)$$

where $n = \sum_{i=1}^{n_\mathrm{g}} 1 + s_i$ (according to (3)), and

$$V_N(\theta, Z^N) \triangleq \frac{1}{N}\sum_{k=1}^{N} \big(y(k) - \varphi^\top(k)\theta\big)^2. \quad (7)$$

To guarantee a unique solution of (6) it is assumed that $\{\psi_{ij}\}_{i=1,j=1}^{n_\mathrm{g},s_i}$ are chosen such that (2) is globally identifiable (there exist no $\theta$ and $\theta'$, such that the 1-step ahead predictor resulting from (2) is not distinguishable for $\theta$ and $\theta'$) and that $Z^N$ provides a *persistently exciting* regressor in (4) (see [20]). By organizing the data as

$$Y = \begin{bmatrix} y(1) & y(2) & \dots & y(N) \end{bmatrix}^\top, \quad (8a)$$

$$\Phi = \begin{bmatrix} \varphi(1) & \varphi(2) & \dots & \varphi(N) \end{bmatrix}^\top, \quad (8b)$$

the optimal solution to (6) can be written as

$$\hat{\theta}_N = \big(\Phi^\top\Phi\big)^{-1}\Phi^\top Y \triangleq \Phi^\dagger Y. \quad (9)$$

## III. ORDER SELECTION BY USING THE NNG

### A. The general NNG

The *Nonnegative Garrote* (NNG) method was first presented in [18] as a coefficient shrinkage method for linear regression models in statistics. As the celebrated Lasso method [21], it uses regularization to penalize the size of the parameter $\theta$. However, instead of affecting the parameters directly, the NNG method penalizes the least-squares solution by attaching weights to it, which in turn are regularized. Thus, given the least-squares estimate $\hat{\theta}_N$ of the parameters of a linear regression model like (4), the NNG problem can be written as

$$\min_{w} \sum_{k=1}^{N}\left(y(k) - \sum_{i=1}^{n_\mathrm{g}}\sum_{j=0}^{s_i} w_{ij}\varphi_{ij}(k)\hat{\theta}_{ij}\right)^2 + \lambda\sum_{i=1}^{n_\mathrm{g}}\sum_{j=0}^{s_i} w_{ij} \quad (10a)$$

$$\text{s.t. } w \succeq 0 \quad (10b)$$

where $\lambda$ is the model complexity parameter, $\varphi_{ij}(k)$ is the $(j + \sum_{\tau=1}^{i-1}(s_\tau + 1))$-th element of the vector $\varphi(k)$, $w \triangleq \begin{bmatrix} w_{10} & \dots & w_{n_\mathrm{g}s_{n_\mathrm{g}}} \end{bmatrix}^\top$ are the weights, and $\succeq$ denotes componentwise inequality. For a given $\lambda$, (10a-b) is a convex optimization problem in the decision variable $w$, and the NNG parameter estimate has the elements $w_{ij}\hat{\theta}_{ij}$, $1 \le i \le n_\mathrm{g}$, $0 \le j \le s_i$, where $w_{ij}$ is the optimal solution to (10a-b). As $\lambda$ increases, the weights of the less important regressors will shrink, and finally end up exactly zero. Thus, as $\lambda$ increases, the model becomes less complex.

### B. Modification for the LPV case

In system identification, one is typically interested in the estimation of dynamical models, in contrast to the static models commonly used in statistics. In dynamic linear regression models, the regressors are naturally ordered by their time lag. The higher the model order, the more data is needed. The original NNG method (10a-b) does not take such orderings into consideration. It just sets the weights of the less important regressors low, not considering their order. On the other hand, it is a particular feature of LPV linear regression models that besides the natural ordering of time lags, there is a lack of natural ordering of the functional terms $\psi_{ij}$ in the parametrization (3) of the $p$-dependent coefficients.

By taking into account the natural ordering of time lags it is possible to penalize a higher model order in the NNG estimate, leading to an approach to model order selection. To achieve this, without introducing ordering with respect to the parameters of the functional terms in each coefficients, one could modify (10a-b) by adding some constraints on the weights. For LPV-ARX models, these constraints could be

$$1 \geq \sum_{j=0}^{s_1} w_{1j} \geq \sum_{j=0}^{s_2} w_{2j} \geq \ldots \geq \sum_{j=0}^{s_{n_\mathrm{a}}} w_{n_\mathrm{a} j}, \tag{10c}$$

$$1 \geq \sum_{j=0}^{s_{n_\mathrm{a}+1}} w_{(n_\mathrm{a}+1)j} \geq \ldots \geq \sum_{j=0}^{s_{n_\mathrm{g}}} w_{n_\mathrm{g} j}. \tag{10d}$$

This is a natural[1] extension of the NNG method, for order selection of LPV-ARX models in system identification. In (10c-d), the ordering of the weights associated with $a_i$ and $b_j$ is independent. This yields automatic order selection, and a natural way to choose the importance between input lag and output lag, as their weightings remain independent. Moreover, (10c-d) does not unnecessarily constrain the choices of basis functions $\psi_{ij}$ within each group $\phi_i$ in (3). This provides a way to select the most adequate structural dependencies for the parametrization of the coefficients, independently from the model order. Note that this particular freedom of the NNG method represents an advantage over the use of classical regressor selection approaches of the LTI case, like AIC, BIC, etc. (see [17]) for LPV-ARX models. In these approaches, there is no possibility to provide both order and structural dependence selection.

The modified NNG problem (10a-d) can be written as a quadratic problem with linear inequality constraints, i.e.

$$\min \quad \frac{1}{2} w^\top Q w + f^\top w + \lambda E^\top w, \tag{11a}$$

$$\text{s.t.} \quad \Gamma w \preceq b, \tag{11b}$$

where $Q = 2\hat{\Theta}\Phi^\top\Phi\hat{\Theta}$, $f = -2\hat{\Theta}\Phi^\top Y$, $\hat{\Theta} \triangleq \operatorname{diag}(\hat{\theta})$, $E = [\ 1 \quad \ldots \quad 1\ ]$, and the inequality constraints (11b) are derived from (10a-d). Given the solution $w_\lambda$ to (11a-b), for a specific $\lambda$, the modified NNG parameter estimate is $\hat{\theta}_\lambda = \Theta w_\lambda$.

[1]Note that other choices for the ordering of the parameters, e.g. the maximum instead of the sum, are also possible. The effect of using different choices has not been evaluated.

### C. The algorithm

Basically, what we need to do is to solve (11a-b) for increasing values of $\lambda$, resulting in less and less complex model estimates, as long as the overall fit of the model estimate on validation data is still acceptable. An efficient way to implement this strategy is to use a path following parametric estimation. For this purpose a Lagrangian multipliers based method has been proposed in [19]. Starting form $\lambda = 0$, this method calculates a piece-wise affine solution path for $\lambda$. In this way it efficiently explores the change in the model fit as a function of $\lambda$. For more details see [19].

## IV. Simulations

In order to test the applicability of the proposed method, two examples are considered where the modified NNG is applied to simulated data.

### A. LPV-ARX model

In the first simulation example, the data-generating system is an LPV-ARX$(9, 3)$ model:

$$A(q,p)y = B(q,p)u + e, \tag{12}$$

where the noise $e$ is white with a Gaussian distribution $\mathcal{N}(0, 0.1)$, $p(k) \in \mathbb{P}$ with $\mathbb{P} = [-2\pi, 0]$ and

$$\begin{aligned} A(q,p) = 1 &+ (0.24 + 0.1p)q^{-1} - (0.1\sqrt{-p} - 0.6)q^{-2} \\ &+ 0.3\sin(p)q^{-3} + (0.17 + 0.1p)q^{-4} \\ &+ 0.3\cos(p)q^{-5} - 0.27q^{-6} + (0.01p)q^{-7} \\ &- 0.07q^{-8} + 0.01\cos(p)q^{-9}, \\ B(q,p) = 1 &+ (1.25 - p)q^{-1} - (0.2 + \sqrt{-p})q^{-2}, \end{aligned}$$

are polynomials in $q$ with a static coefficient dependence on $p$. In Figure 1a, the poles of (12) are plotted for all constant trajectories of $p$. As all frozen poles are in the unit disc, the LPV-ARX$(9, 3)$ model is stable for all constant trajectories of $p$ (*uniform frozen stability*). Figure 1a also indicates that the model has fast and slow modes which change rapidly with the variation of $p$.

The system (12) is simulated using a white noise $u$ with distribution $\mathcal{N}(0, 1)$ and a white noise $p$ with uniform distribution $\mathcal{U}(-2\pi, 0)$. With these signals, $2N$ data points are collected with $N = 5000$, and the obtained data record is divided equally in an estimation and a validation part. Under these conditions, the *Signal to Noise Ratio* (SNR) in the generated data set is 35 dB.

To evaluate the different model outcomes, the *Best Fit Rate* (BFR) is used [22]:

$$\mathrm{BFR} = 100\% \cdot \max\left(1 - \frac{\|y(k) - \hat{y}(k|\theta_\lambda)\|_2}{\|y(k) - \bar{y}\|_2}, 0\right), \tag{13}$$

where $\bar{y}$ is the mean of $y$. The BFR measures how much better the model describes the process compared to the mean of the output. In order to compare the results to other LPV identification approaches (like the subspace methods, see [4]), the *Variance Accounted For* (VAF) percentage is also computed:

$$\text{VAF} = 100\% \cdot \max\left(1 - \frac{\text{var}\left(y(k) - \hat{y}(k|\theta_\lambda)\right)}{\text{var}\left(y(k)\right)}, 0\right), \quad (14)$$

which is a measure of the percentage of the observed output variation that is explained by the model.

*1) Perfect model order:* As a first step, a LPV-ARX$(9, 3)$ model is estimated based on the collected data and using the coefficient parametrization (3) with

$$\psi_{i1}(p) = p, \qquad \psi_{i3}(p) = \sin(p),$$
$$\psi_{i2}(p) = \sqrt{-p}, \qquad \psi_{i4}(p) = \cos(p),$$

for all $i$, i.e. $s_1 = \ldots = s_{12} = 4$. This parametrization corresponds to $5 \cdot 12 = 60$ unknown $\theta_{ij}$'s to be estimated. The obtained LS estimate has been computed with a slightly modified version of the `arx` command in MATLAB. Note that (12) is in the model class and the model order is correct, but the coefficients are overparametrized, as only a subset of $\{\psi_{ij}\}$ is required for the estimation of each $\phi_i$.[2]

Plugging this estimate into the NNG problem (11a-b) and solving it with the proposed algorithm of [19] yield a piecewise affine solution path $w_\lambda$. For this solution path, the BFR and VAF of the associated model estimates are calculated for the validation data. As the performance path is similar for the BFR and VAF error measures, except that in the VAF case it is in the $100\%$ - $99\%$ region, only the BFR path is shown in Figure 1b. Note that the calculation time of this figure together with the solution of the NNG problem and the model estimate only takes a few seconds on a Pentium 4, 2.8 GHz PC running under Windows XP with SP2. The maximum of fit occurs for $\lambda_{42}$ with $\text{BFR} = 99.63\%$ and $\text{VAF} = 99.99\%$, for which the corresponding model is a LPV-ARX$(9, 3)$ with coefficients given in Table I. The coefficient dependencies clearly indicate a high similarity to the original $A(q, p)$ and $B(q, p)$. From the obtained graph it is also obvious that a model reduction is possible without too much loss in BFR. By choosing $\lambda = \lambda_{66}$, the corresponding model is a LPV-ARX$(8, 3)$ model with coefficients given in Table I. Note that all parameters of $a_7$ are set to zero by the method. This model approximates (12) with $\text{BFR} = 95.36\%$ and $\text{VAF} = 99.79\%$, which implies that the NNG method may also be used as a model reduction method for LPV-ARX models.

*2) Overfitting:* Now consider the situation of overfitting by estimating an LPV-ARX$(12, 6)$ model based on the collected data and using the coefficient parametrization (3) with

$$\psi_{i1}(p) = p, \qquad \psi_{i3}(p) = \sin(p), \qquad \psi_{i5}(p) = p^2,$$
$$\psi_{i2}(p) = \sqrt{-p}, \qquad \psi_{i4}(p) = \cos(p),$$

for all $i$, i.e. $s_1 = \ldots = s_{18} = 5$. Note that the true system (12) is again in the model class, but both the model order and the coefficient dependencies are overparametrized (108 parameters compared to 17). Plugging this estimate into the NNG problem (11a-b) and solving it with the proposed algorithm yield a piecewise affine solution path $w_\lambda$ for which the fit values are shown in Figure 1c. Again calculations only take a few seconds on the specified PC.

In Figure 1c, the fit values have an obvious maximum at $\lambda = \lambda_{103}$, which corresponds to a LPV-ARX$(9, 2)$ with coefficients given in Table I. Values for the parameters of $p^2$ are not reported as they are all zero. Note that this model has the correct model order and coefficient dependencies of the original data-generating plant. This underlines the value of the proposed method.

*Remark 1:* If the data-generating system is not in the model class due to undermodelling or inappropriate choice of the noise model, the proposed NNG approach still provides a reliable solution (see [19]). However, the investigation of the effect of structural modeling error in terms of the used $\{\psi_{ij}\}$ functions remains the objective of future research.

### *B. LPV-SS model*

As a next example, the identification of an LPV-SS model is demonstrated by using the LPV-ARX structure with dynamic dependence. In this setting, the NNG method is used to select the required dynamic dependence of the ARX model coefficients in order to deliver an adequate estimate of the data-generating system.

Consider the LPV-SS model

$$qx = \overbrace{\begin{bmatrix} 0 & p \\ 1 & p \end{bmatrix}}^{A(p)} x + \overbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}^{B(p)} u + \begin{bmatrix} 1 \\ 1 \end{bmatrix} e, \quad (15a)$$

$$y = \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{C(p)} x \quad (15b)$$

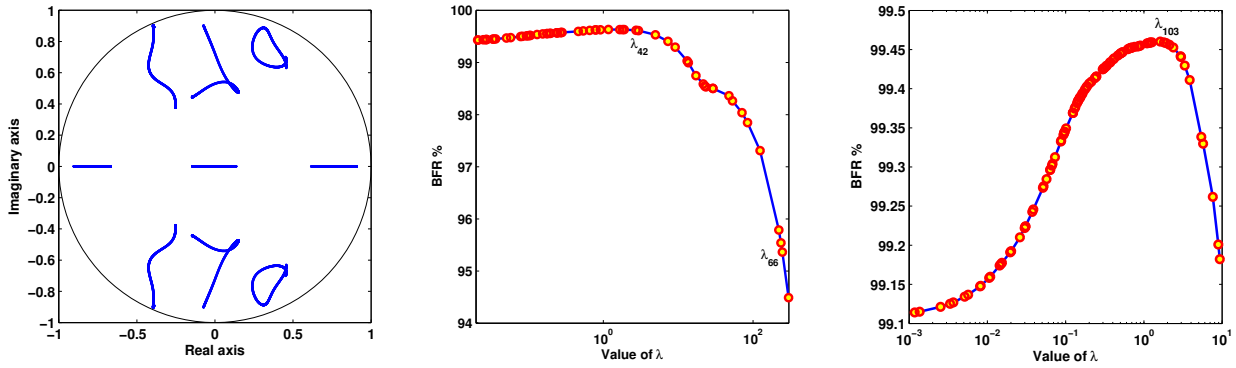
where $x$ denotes the state variable, $e$ is white noise with distribution $\mathcal{N}(0, 0.02)$ and $p \in \mathbb{P}$ where $\mathbb{P} = [-0.4, 0.4]$. Note that the matrices in (15a-b) depend only on the instantaneous value of $p$. Based on the frozen poles of (15a-b) given in Figure 2a, this system is uniformly frozen stable. Using the transformation theory presented in [11], [12], the equivalent LPV-IO realization of (15a-b) reads as

$$\left(1 - p(k-1)\left[q^{-1} + q^{-2}\right]\right) y(k) = u(k-1) + e(k-1). \quad (16)$$

Note in (16) that the coefficients of the output side polynomial depend on the time-shifted value of $p$, which is called dynamic dependence. Recently, it has been shown that in order to estimate adequate models of physical systems, possible dynamic dependence of the model coefficients must be taken into account as is obvious from (16). However, guessing the required order of time-shifts in the scheduling variable only from measured data is a non-trivial problem.

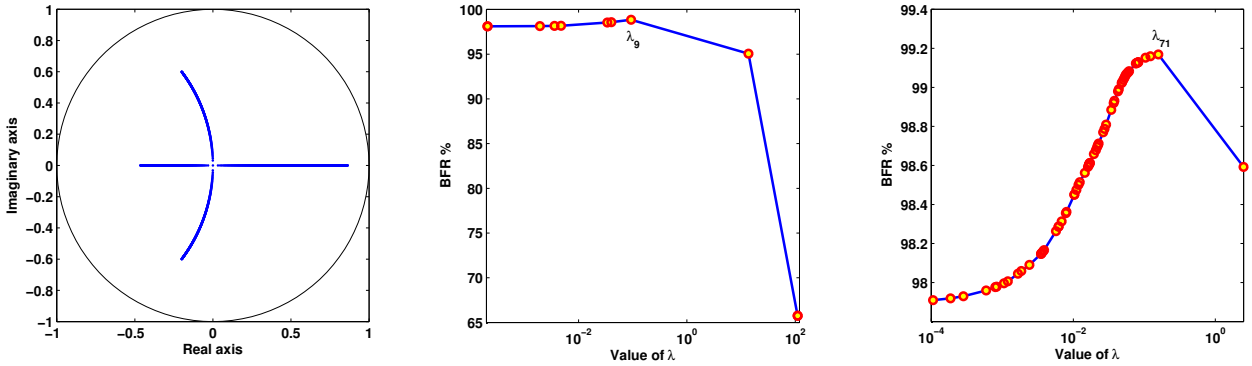
To test the NNG method in this setting, the system (15a-b) is simulated using a white noise $u$ with distribution $\mathcal{N}(0, 1)$ and a white noise $p$ with uniform distribution $\mathcal{U}(-0.4, 0.4)$. With these signals, just like in the previous case, $2N$ data points are collected with $N = 1000$, and the obtained data record is divided equally into an estimation and a validation data set. In the obtained data the SNR is 35 dB.

[2] In the noise-free case, the initial parameter estimate for the overparametrized model will lead to zero elements in $\hat{\theta}$. This, in turn, yields columns that are zero in the regressor matrix for the weights $w$ (10a) and the problem will be rank-deficient. Thus, special care is needed, where a rank-revealing decomposition may be used to transform the problem into a well conditioned one.




(a) (b) (c)

Fig. 1. (a) The pole locations of the LPV-ARX$(9,3)$ data-generating system (see (12)) for all constant trajectories of $p$ with $\mathbb{P} = [-2\pi, 0]$. (b) Using the model order $(9,3)$ in the NNG problem, the BFR for the break points of the piecewise affine solution path in terms of $\lambda$, calculated for the validation data. (c) Piecewise affine solution path of $\lambda$ for model order $(12,6)$.




(a) (b) (c)

Fig. 2. (a) The pole locations of the LPV-SS data-generating system (see (15a-b)) for all constant trajectories of $p$ with $\mathbb{P} = [-0.4, 0.4]$. (b) Using the model order $(2,2)$ in the NNG problem, the BFR for the break points of the piecewise affine solution path in terms of $\lambda$, calculated for the validation data. (c) Piecewise affine solution path of $\lambda$ for model order $(5,5)$.

*1) Perfect model order:* We demonstrate the capabilities of the NNG method to solve the dynamic dependence selection problem by estimating a LPV-ARX$(2,2)$ model based on the collected data and by using the coefficient parametrization (3) with

$$\psi_{i1}(p) = p, \qquad \psi_{i2}(p) = q^{-1}p,$$

for all $i$, i.e. $s_1 = \ldots = s_4 = 2$. Note that the true system (15a-b) is in the model class.

Again the initial LS model estimate has been computed in MATLAB and the NNG problem (11a-b) has been solved for all nonnegative $\lambda$. The resulting piecewise affine solution path $w_\lambda$ is depicted in Figure 2b in terms of the achieved BFR and VAF of the associated model estimates with respect to the validation data. By choosing $\lambda = \lambda_9$, which has the highest BFR $= 99.10\%$ and VAF $= 99.99\%$, the corresponding model is a LPV-ARX$(2,2)$ model with coefficients:

$$\begin{aligned} a_1(p)(k) &= -1.01p(k-1), & b_0(p)(k) &= -0.0083, \\ a_2(p)(k) &= -0.99p(k-1), & b_1(p)(k) &= 0.9971, \end{aligned}$$

which has only coefficient dependencies on $p(k-1)$. By considering that $b_0 \approx 0$, the obtained model is almost a perfect match with (16). This proves that the NNG method correctly selects the required dynamic dependence for the identification of LPV-ARX models.

*2) Overfitting:* Now we consider an overfitting scenario by estimating an LPV-ARX$(5,5)$ model based on the collected data and by using the coefficient parametrization (3) with

$$\begin{aligned} \psi_{i1}(p) &= p, & \psi_{i2}(p) &= q^{-1}p, \\ \psi_{i3}(p) &= q^{-2}p, & \psi_{i4}(p) &= q^{-3}p, \end{aligned}$$

for all $i$, i.e. $s_1 = \ldots = s_{10} = 4$. Note that the true system (11a-b) is again in the model class, but both the model order and the coefficient dependencies are overparametrized and except for $\psi_{i2}$ none of $\{\psi_{ij}\}$ show up in the true system as a part of any coefficient dependence (see (16)). Again, plugging this estimate into the NNG problem (11a-b) and solving it with the proposed algorithm yield a piecewise affine solution path $w_\lambda$ for which the fit values are shown in Figure 2c.

In Figure 2c, the fit values have an obvious maximum at $\lambda = \lambda_{71}$ with BFR $= 99.17\%$ and VAF $= 99.99\%$, which corresponds to a LPV-ARX$(2,2)$ model with coefficients:

$$\begin{aligned} a_1(p)(k) &= -1.022p(k-1), & b_0(p)(k) &= 0.0148p(k-3), \\ a_2(p)(k) &= -0.989p(k-1), & b_1(p)(k) &= 0.9963, \end{aligned}$$

TABLE I

PARAMETERS OF THE MODEL ESTIMATES IN EXAMPLE IV-A FOR DIFFERENT VALUES OF $\lambda$. (ONLY THE NON-ZERO ROWS ARE GIVEN)

| | | True | $\lambda_{42}$ | $\lambda_{66}$ | $\lambda_{103}$ |
|---|---|---|---|---|---|
| $a_1$ | $1$ | $0.24$ | $0.2401$ | $0.2425$ | $0.2434$ |
| | $p$ | $0.1$ | $0.1003$ | $0.1021$ | $0.1060$ |
| | $\sqrt{-p}$ | $0$ | $0$ | $0$ | $0.0068$ |
| $a_2$ | $1$ | $0.6$ | $0.6001$ | $0.5710$ | $0.5911$ |
| | $p$ | $0$ | $0$ | $0$ | $0.0022$ |
| | $\sqrt{-p}$ | $-0.1$ | $-0.1001$ | $-0.0856$ | $-0.0908$ |
| $a_3$ | $1$ | $0$ | $0.0001$ | $0$ | $-0.0075$ |
| | $p$ | $0$ | $-0.0002$ | $0$ | $0.0054$ |
| | $\sqrt{-p}$ | $0$ | $-0.0070$ | $0$ | $0.0139$ |
| | $\sin(p)$ | $0.3$ | $0.2993$ | $0.2991$ | $0.3008$ |
| | $\cos(p)$ | $0$ | $0$ | $0$ | $0.0028$ |
| $a_4$ | $1$ | $0.17$ | $0.1710$ | $0.1448$ | $0.1679$ |
| | $p$ | $0.1$ | $0.1002$ | $0.09039$ | $0.0995$ |
| $a_5$ | $\sin(p)$ | $0$ | $0.0014$ | $0$ | $0$ |
| | $\cos(p)$ | $0.3$ | $0.3006$ | $0.3018$ | $0.3007$ |
| $a_6$ | $1$ | $-0.27$ | $-0.2714$ | $-0.2692$ | $-0.2698$ |
| $a_7$ | $p$ | $0.01$ | $0.0097$ | $0$ | $-0.0120$ |
| | $\sin(p)$ | $0$ | $0$ | $0$ | $0.0018$ |
| $a_8$ | $1$ | $-0.07$ | $-0.0693$ | $-0.0645$ | $-0.0698$ |
| $a_9$ | $1$ | $0$ | $0.0004$ | $0$ | $0$ |
| | $\cos(p)$ | $0.01$ | $0.0091$ | $0$ | $0.0116$ |
| $b_0$ | $1$ | $1$ | $1.0069$ | $1.0226$ | $1.0328$ |
| | $p$ | $0$ | $-0.0003$ | $0$ | $-0.0704$ |
| | $\sqrt{-p}$ | $0$ | $-0.0070$ | $-0.0227$ | $-0.1141$ |
| | $\sin(p)$ | $0$ | $-0.0004$ | $0$ | $-0.0121$ |
| | $\cos(p)$ | $0$ | $-0.0005$ | $0$ | $0$ |
| $b_1$ | $1$ | $1.25$ | $1.2474$ | $1.1978$ | $1.2578$ |
| | $p$ | $-1$ | $-1.0017$ | $-1.0103$ | $-1.0006$ |
| $b_2$ | $1$ | $-0.2$ | $-0.1815$ | $0$ | $-0.1748$ |
| | $\sqrt{-p}$ | $-1$ | $-1.0117$ | $-1.1134$ | $-1.0122$ |

Note that this model has the correct model order and coefficient dependencies of the original data-generating plant if one considers $b_0$ to be approximately zero. This underlines the value of the proposed method, giving strong evidence that it is able to select appropriate order and structural dependence of LPV-ARX model structures.

*Remark 2:* Ordering of the dynamical coefficient dependence can also be introduced in (10c-d), by penalizing higher lags in $p$ (similar to input and output lags).

## V. CONCLUSIONS AND FUTURE WORK

In this paper a method for order and structural dependence selection of LPV-ARX models was introduced as the extension of the order selection approach presented in [19] for LTI-ARX models. The method is a modified variant of the NNG method [18], where constraints on the weights are added according to the natural ordering of the regressors in ARX models. At the same time, the weights of the prior given set of candidate scheduling dependencies are left unconstrained for each coefficient, to give equal chances for the selection of the most important candidates. This also provides the possibility to estimate the order of the required dynamic dependence in LPV-ARX models, giving a practical tool for the support of LPV-IO identification approaches. The proposed method is extendable to the multivariable case, providing an important objective for further research.

The order and structure selection problem of LPV-ARX models is a special case of the related NARX problem, which will be studied in the future together with the *Instrumental Variable Regression Shrinkage* version of the presented algorithm.

## REFERENCES


[1] W. Rugh and J. Shamma, "Research on gain scheduling," *Automatica*, vol. 36, no. 10, pp. 1401–1425, 2000.

[2] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice-Hall, 1998.

[3] C. W. Scherer, "Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control for time-varying and linear parametrically-varying systems," *Int. Journal of Robust and Nonlinear Control*, vol. 6, no. 9-10, pp. 929–952, 1996.

[4] J. W. van Wingerden and M. Verhaegen, "Subspace identification of bilinear and LPV systems for open- and closed-loop data," *Automatica*, vol. 45, pp. 372–381, 2009.

[5] V. Verdult and M. Verhaegen, "Subspace identification of multivariable linear parameter-varying systems," *Automatica*, vol. 38, no. 5, pp. 805–814, 2002.

[6] B. Bamieh and L. Giarré, "Identification of linear parameter varying models," *Int. Journal of Robust and Nonlinear Control*, vol. 12, pp. 841–853, 2002.

[7] X. Wei, "Adaptive LPV techniques for diesel engines," Ph.D. dissertation, Johannes Kepler University, Linz, 2006.

[8] K. Hsu, T. L. Vincent, and K. Poolla, "Nonparametric methods for the identification of linear parameter varying systems," in *Proc. of the Int. Symposium on Computer-Aided Control System Design*, San Antonio, Texas, USA, Sept. 2008, pp. 846–851.

[9] R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof, "Flexible model structures for LPV identification with static scheduling dependency," in *Proc. of the 47th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 4522–4527.

[10] ——, "Asymptotically optimal orthonormal basis functions for LPV system identification," *Automatica*, vol. 45, no. 6, pp. 1359–1370, 2009.

[11] R. Tóth, "Modeling and identification of linear parameter-varying systems, an orthonormal basis function approach," Ph.D. dissertation, Delft University of Technology, 2008.

[12] R. Tóth, F. Felici, P. S. C. Heuberger, and P. M. J. Van den Hof, "Discrete time LPV I/O and state space representations, differences of behavior and pitfalls of interpolation," in *Proc. of the European Control Conf.*, Kos, Greece, July 2007, pp. 5418–5425.

[13] A. A. Khalate, X. Bombois, R. Tóth, and R. Babuška, "Optimal experimental design for LPV identification using a local approach," in *Proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009, pp. 162–167.

[14] M. Butcher, A. Karimi, and R. Longchamp, "On the consistency of certain identification methods for linear parameter varying systems," in *Proc. of the 17th IFAC World Congress*, Seoul, Korea, July 2008, pp. 4018–4023.

[15] X. Wei and L. Del Re, "On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque," in *Proc. of the 14th IFAC Symposium on System Identification*, Newcastle, Australia, Mar. 2006, pp. 517–522.

[16] K. Åström and P. Eykhoff, "System identification - a survey," *Automatica*, vol. 7, pp. 123–162, 1971.

[17] L. Ljung, *System Identification, theory for the user*. Prentice Hall, 1999.

[18] L. Breiman, "Better subset regression using the nonnegative garotte," *Technometrics*, vol. 37, pp. 373–384, 1995.

[19] C. Lyzell, J. Roll, and L. Ljung, "The use of nonnegative garrote for order selection of ARX models," in *Proc. of the 45th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 1974–1979.

[20] M. Gevers, A. Bazanella, and L. Mišković, "Informative data: how to get just sufficiently rich?" in *Proc. of the 47th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 4522–4527.

[21] R. Tibshirani, "Regression shrinkage and selection with the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, pp. 267–288, 1996.

[22] L. Ljung, *System Identification Toolbox, for use with Matlab*. The Mathworks Inc., 2006.

**Avdelning, Institution**
Division, Department

Division of Automatic Control
Department of Electrical Engineering

**Datum**
Date

2010-03-03

**Språk**
Language

☐ Svenska/Swedish
☒ Engelska/English

☐ _____

**Rapporttyp**
Report category

☐ Licentiatavhandling
☐ Examensarbete
☐ C-uppsats
☐ D-uppsats
☒ Övrig rapport
☐ _____

**ISBN**
—

**ISRN**
—

**Serietitel och serienummer**
Title of series, numbering

**ISSN**
1400-3902

**URL för elektronisk version**

`http://www.control.isy.liu.se`

LiTH-ISY-R-2937

**Titel**
Title

Order and Structural Dependence Selection of LPV-ARX Models Using a Nonnegative Garrote Approach

**Författare**
Author

Roland Tóth, Christian Lyzell, Martin Enqvist, Peter S. C. Heuberger, Paul M. J. Van den Hof

**Sammanfattning**
Abstract

In order to accurately identify Linear Parameter-Varying (LPV) systems, order selection of LPV linear regression models has prime importance. Existing identification approaches in this context suffer from the drawback that a set of functional dependencies needs to be chosen a priori for the parametrization of the model coefficients. However in a black-box setting, it has not been possible so far to decide which functions from a given set are required for the parametrization and which are not. To provide a practical solution, a nonnegative garrote approach is applied. It is shown that using only a measured data record of the plant, both the order selection and the selection of structural coefficient dependence can be solved by the proposed method.

**Nyckelord**
Keywords

Linear Parameter-Varying, ARX, identification, order selection