

On the Consistency of Certain Identification Methods for Linear Parameter Varying Systems ^{*}

Mark Butcher, Alireza Karimi and Roland Longchamp

*Automatic Control Laboratory, Ecole Polytechnique Fédérale de
Lausanne (EPFL), 1015 Lausanne, Switzerland.
(e-mail: alireza.karimi@epfl.ch)*

Abstract: The consistency of identification methods for input-output models of Linear Parameter Varying systems is considered. In order to perform a consistency analysis the applicability of ergodicity is required, which is not obvious with these types of nonstationary systems. It is therefore shown that, when the scheduling parameter satisfies certain conditions, an ergodicity-type result can be applied to the signals considered. An analysis is then carried out for two cases: that of noise-free measurements of the scheduling parameter, and then the more realistic case of noisy scheduling parameter measurements. The latter is shown to be an errors-in-variables type problem. Since the least squares technique does not give consistent estimates, the instrumental variables method is proposed to achieve this property. The analysis carried out is reinforced by simulation results.

1. INTRODUCTION

In many real world control applications a Linear Time Invariant (LTI) assumption is made on the system's dynamics in order to use the well developed identification and control techniques available for this class of systems. This assumption is valid when the system remains within a certain operating zone. However, when the operating point changes considerably the identified model is no longer valid and the controlled system's performance is degraded, and in the worst case it becomes unstable.

Therefore in order to achieve good performance throughout the operating zone but still use the linear system techniques the class of Linear Parameter Varying (LPV) systems has been defined (Shamma and Athans, 1991).

LPV systems can be thought of as a weighted combination of linear models. The weightings are a function of the operating point, which is, in turn, a function of certain scheduling parameters. These scheduling parameters can either be endogenous signals, such as the system's states or outputs, or exogenous signals, which cause the dynamics to change as a function of time according to the trajectories of these signals. The former case is sometimes referred to as Local Linear Models (Verdult, 2002). The latter means that the LPV system represents a family of Linear Time Varying (LTV) systems, each system corresponding to a particular set of trajectories.

In this paper the consistency of certain identification methods for LPV systems will be considered. A certain amount of work has already been done into identification of LPV systems. Two directions have been investigated principally: subspace methods and 'classical' identification techniques to find input-output models.

^{*} This work is supported by the Swiss National Science Foundation under Grant No. 200021-116156/1.

In Nemani et al. (1995) a method is proposed for the identification of LPV state-space models. It, however, requires full-state measurement, which is often not possible, and is only capable of handling one scheduling parameter. Multiple scheduling parameters can be dealt with using the method proposed in Lovera et al. (1998), but full-state measurement is still required. In Verdult and Verhaegen (2002) full-state knowledge is no longer required as a scheme for estimating the states is proposed. In order to obtain consistent estimates of the states, when the scheduling parameter is noise-free, the use of instrumental variables (IV) is proposed.

In Lee and Poolla (1999) the output-error identification problem for LPV state-space models is examined. The scheduling parameters are, again, considered noise-free. In this case, parameter estimates are consistent when the real system is in the model set and the global minimum of the proposed quadratic cost function is found.

A technique for the identification of LPV systems in input-output form is proposed in Bamieh and Giarré (2002). It is shown that with the proposed parameterisation the prediction error can be written in linear regression form. Using the quadratic loss function proposed the parameter estimation problem becomes one of linear least squares, therefore finding a global minimum. A white equation error is, however, assumed and no explicit analysis of the consistency of the method is made.

In this paper the ergodicity of the signals used in the identification of LPV systems is studied for different assumptions on the scheduling parameter. It is shown that the signals are ergodic if the scheduling parameter is noise-free or the measurement noise on the scheduling parameter is independent of that affecting the output. In these cases the use of IV leads to consistent estimation of the parameters. However, for the cases that the scheduling parameter

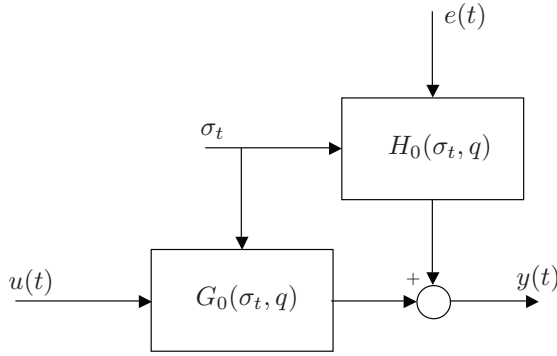


Fig. 1. True System

is correlated with the output signal or the LPV system does not have an affine dependency on the noisy scheduling parameter, the ergodicity of the signals cannot be proved.

The paper is organized as follows. Notation and preliminaries are given in Section 2. The non-noisy scheduling parameter case is considered in Section 3. Then the noisy scheduling parameter case is analysed in Section 4. Simulation results are presented in Section 5. Finally, some concluding remarks are made in Section 6.

2. PRELIMINARIES

2.1 True System Representation

Suppose that the true output of the SISO LPV system $G_0(\sigma_t, q)$ is given by (see Fig 1):

$$y(t) = G_0(\sigma_t, q)u(t) + H_0(\sigma_t, q)e(t) \quad (1)$$

where $\sigma_t \in \mathbb{R}^{n_\sigma}$ is the measurable scheduling parameter vector of dimension n_σ at time t , q is the forward-shift time operator and $e(t)$ is a sequence of independent random variables with zero mean values. $G_0(\sigma_t, q)$ and $H_0(\sigma_t, q)$ are the system and noise models, respectively, and are assumed to be uniformly stable for all σ_t in the operating zone. $G_0(\sigma_t, q)$ can be represented in transfer operator form as:

$$G_0(\sigma_t, q) = \frac{q^{-d}B_0(\sigma_t, q)}{A_0(\sigma_t, q)} \quad (2)$$

where:

$$B_0(\sigma_t, q) = b_{0,0}^0(\sigma_t) + b_{0,1}^0(\sigma_t)q^{-1} + \dots + b_{0,n_b}^0(\sigma_t)q^{-n_b} \quad (3)$$

and

$$A_0(\sigma_t, q) = 1 + a_{1,1}^0(\sigma_t)q^{-1} + \dots + a_{n_a, n_a}^0(\sigma_t)q^{-n_a}. \quad (4)$$

n_b and n_a are the numerator and denominator orders, respectively. The coefficients of $B_0(\sigma_t, q)$ are given by:

$$b_i^0(\sigma_t) = b_{i,0}^0\sigma_{t,0} + b_{i,1}^0\sigma_{t,1} + \dots + b_{i,n_\sigma}^0\sigma_{t,n_\sigma}. \quad (5)$$

Here $\sigma_{t,j}$ represents the j th element of σ_t . This parameterisation allows a wide range of dependence on the scheduling parameter to be described. For example each $\sigma_{t,j}$ could represent a function of a different scheduling parameter. Alternatively, as used in Bamieh and Giarré (2002), the numerator and denominator coefficients could be polynomially dependent on a single scheduling parameter i.e.

$$\sigma_{t,j} = \bar{\sigma}_t^j, \quad (6)$$

where $\bar{\sigma}_t$ is the single measured scheduling parameter. A similar representation to (5) exists for $a_i^0(\sigma_t)$.

Using (2), (1) can be rewritten as:

$$\begin{aligned} A_0(\sigma_t, q)y(t) &= q^{-d}B_0(\sigma_t, q)u(t) \\ &\quad + A_0(\sigma_t, q)H_0(\sigma_t, q)e(t) \\ &= q^{-d}B_0(\sigma_t, q)u(t) + e_0(t) \end{aligned} \quad (7)$$

This, in turn, can be written in linear regression form as:

$$y(t) = \phi^T(t)\theta_0 + e_0(t), \quad (8)$$

where the regressor vector is given by:

$$\phi(t) = \varphi(t) \otimes \sigma_t \quad (9)$$

where \otimes is the Kronecker product and

$$\begin{aligned} \varphi^T(t) &= [-y(t-1), -y(t-2), \dots, -y(t-n_a), \\ &\quad u(t-d), u(t-d-1), \dots, u(t-d-n_b)] \end{aligned} \quad (10)$$

and the true parameter vector is:

$$\begin{aligned} \theta_0^T &= [a_{1,0}^0, a_{1,1}^0, \dots, a_{1,n_\sigma}^0, \dots, a_{n_a,0}^0, a_{n_a,1}^0, \dots, a_{n_a,n_\sigma}^0, \\ &\quad b_{0,0}^0, b_{0,1}^0, \dots, b_{0,n_\sigma}^0, \dots, b_{n_b,0}^0, b_{n_b,1}^0, \dots, b_{n_b,n_\sigma}^0]. \end{aligned} \quad (11)$$

2.2 Model Representation

The model of the system is written as:

$$A(\sigma_t, q)y(t) = q^{-d}B(\sigma_t, q)u(t) \quad (12)$$

which allows the predictor:

$$\hat{y}(t, \theta) = \phi^T(t)\theta \quad (13)$$

to be defined, along with the prediction error:

$$\epsilon(t, \theta) = y(t) - \hat{y}(t, \theta) = y(t) - \phi^T(t)\theta. \quad (14)$$

As consistency is being considered in this paper, it will be assumed that the model order is always chosen equal to the true system order.

2.3 Single Realisation Behaviour and Ergodicity

It is often useful to be able to equate the time average properties of a signal over a single realisation with the ensemble average taken over many realisations. Signals with this property are called ergodic and Theorem 2B.1 in Ljung (1999) indicates when certain types of nonstationary signals can be ergodic in the correlation.

Theorem 2B.1 Let $\{P_\rho(q), \rho \in D_\rho\}$ be a uniformly stable family of filters, and assume that the family of deterministic signals $\{w_\rho(t)\}$, $\rho \in D_\rho$, $t = 1, 2, \dots$, is subject to

$$|w_\rho(t)| \leq C_w, \quad \forall \rho, \forall t. \quad (15)$$

Let the signal $s_\rho(t)$ be defined, for each $\rho \in D_\rho$, by

$$s_\rho(t) = P_\rho(q)v(t) + w_\rho(t) \quad (16)$$

where

$$v(t) = \sum_{k=0}^{\infty} l_k(t)e_l(t-k) = L(t, q)e_l(t) \quad (17)$$

and $e_l(t)$ is a sequence of independent random vectors with zero mean values, $E\{e_l(t)e_l^T(t)\} = \Lambda_t$ and bounded fourth moments, and $\{L(t, q), t = 1, 2, \dots\}$ is a uniformly stable family of filters. Then:

$$\sup_{\rho \in D_\rho} \left\| \frac{1}{N} \sum_{t=1}^N [s_\rho(t)s_\rho^T(t) - E\{s_\rho(t)s_\rho^T(t)\}] \right\| \rightarrow 0 \quad (18)$$

w.p. 1, as $N \rightarrow \infty$,

where $\| \cdot \|$ is the Frobenius norm and N is the number of data measured.

3. NON-NOISY SCHEDULING PARAMETERS

In this section the case of non-noisy scheduling parameters is considered. This assumption is normally made in the LPV identification literature, and is somewhat unrealistic unless the scheduling parameter is a function of the system input. It is considered here as it is a simpler case of the noisy scheduling parameters discussed in the next section. It also allows the polynomial scheduling parameter (6) to be analysed when the degree of the polynomial is greater than 1, which is not possible in the next section. Moreover, in the case of a very low noise-to-signal ratio it might be possible to ignore the noise on the scheduling parameter.

One restriction is that the scheduling parameter cannot be taken as the system output $y(t)$, because this is expressly considered noisy.

As in Bamieh and Giarré (2002), we propose minimisation of a quadratic criterion for the estimation of the system parameters. It is given by:

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^N \epsilon^2(t, \theta). \quad (19)$$

This criterion can be minimised using a number of well known algorithms including the Recursive Least Squares algorithm, as proposed in Bamieh and Giarré (2002), it is also possible to use the standard least squares algorithm for batch data, giving:

$$\hat{\theta}_N^{LS} = \left[\frac{1}{N} \sum_{t=1}^N \phi(t) \phi^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) y(t). \quad (20)$$

Replacing $y(t)$ in (20) with (8) gives:

$$\hat{\theta}_N^{LS} = \theta_0 + \left[\frac{1}{N} \sum_{t=1}^N \phi(t) \phi^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) e_0(t). \quad (21)$$

For the estimate $\hat{\theta}_N^{LS}$ to be consistent i.e. that $\hat{\theta}_N^{LS}$ converges in probability to θ_0 , it is well known that it is necessary that:

$$i) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi(t) \phi^T(t) \text{ be nonsingular.} \quad (22)$$

$$ii) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi(t) e_0(t) = 0. \quad (23)$$

Condition i) is a persistency of excitation condition, which in the LPV case requires more attention, as discussed in Bamieh and Giarré (2002) for the polynomial dependence case (6).

Condition ii) can be analysed using Theorem 2B.1, given in Subsection 2.3. We first note that $\phi(t)e_0(t)$ contains terms such as $\sigma_{t,j}u(t-p)e_0(t)$ and $\sigma_{t,j}y(t-p)e_0(t)$. We then define:

$$s_1(t) = \begin{bmatrix} \sigma_{t,j}u(t-p) \\ \sigma_{t,j}y(t-p) \\ e_0(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma_{t,j}H_0(\sigma_{t-p}, q)e(t-p) \\ A_0(\sigma_t, q)H_0(\sigma_t, q)e(t) \end{bmatrix} + \begin{bmatrix} \sigma_{t,j}u(t-p) \\ \sigma_{t,j}G_0(\sigma_{t-p}, q)u(t-p) \\ 0 \end{bmatrix} = \begin{bmatrix} v_1^1(t) \\ v_2^1(t) \\ v_3^1(t) \end{bmatrix} + \begin{bmatrix} w_1^1(t) \\ w_2^1(t) \\ w_3^1(t) \end{bmatrix} \quad (24)$$

where $w_1^1(t)$ and $w_2^1(t)$ satisfy (15), due to the assumed uniform stability of $G_0(\sigma_t, q)$ and the boundedness of σ_t^j and $u(t)$. Also $v_1^1(t)$ and $v_2^1(t)$ fit in with the desired form of (17) due to the assumed uniform stability of $H_0(\sigma_t, q)$.

The components of $s_1(t)s_1^T(t)$ give, amongst others, $\sigma_{t,j}u(t-p)e_0(t)$ and $\sigma_{t,j}y(t-p)e_0(t)$. Then (18) means that it is possible to write:

$$\left\| \frac{1}{N} \sum_{t=1}^N \phi(t) e_0(t) - \frac{1}{N} \sum_{t=1}^N E\{\phi(t) e_0(t)\} \right\| \rightarrow 0 \quad (25)$$

w.p. 1, as $N \rightarrow \infty$.

Analysing the components of $E\{\phi(t)e_0(t)\}$ we see that:

$$E\{\sigma_{t,j}u(t-p)e_0(t)\} = \sigma_{t,j}u(t-p)E\{e_0(t)\} = 0 \quad (26)$$

and

$$E\{\sigma_{t,j}y(t-p)e_0(t)\} = \sigma_{t,j}E\{[\phi^T(t-p)\theta_0 + e_0(t-p)]e_0(t)\} \quad (27)$$

which is only zero when either $e_0(t)$ is a zero-mean, white noise sequence, or $n_a = 0$. Neither of these cases occur often in practice, so the second sum in (25) will, in general, be non-zero, and thus also the first. The least squares method therefore typically does not give consistent parameter estimates.

In order to obtain consistent parameter estimates the Instrumental Variable (IV) technique can be used. The IV estimate is given by:

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N} \sum_{t=1}^N \zeta(t)^T \phi(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \zeta(t) y(t) \quad (28)$$

where $\zeta(t)$ is the Instrumental Variable vector, which must be correlated with the regressor $\phi(t)$ but not with the noise $e_0(t)$ in order that the algorithm be consistent. It should be mentioned that the IV vector for LPV system identification should be a function of the scheduling parameters, as $\phi(t)$ is. If the IV vector is only a function of $u(t)$ the variance of the estimates becomes too large. A good choice is to use an auxiliary LPV model (identified by the LS method) to generate the IV vector, or use a second experiment.

4. NOISY SCHEDULING PARAMETERS

In this section the more realistic case of the measured values of σ_t being contaminated by noise will be examined. To do this, we express the noisy, measured vector, σ_t^e , as the sum of a noise-free component and a noisy component i.e.

$$\sigma_t^e = \sigma_t + e_{\sigma_t}. \quad (29)$$

$\sigma_t^e \in \mathbb{R}^{n_\sigma}$ and $e_{\sigma_t} \in \mathbb{R}^{n_\sigma}$ are, therefore, stochastic vectors.

The measured regressor vector is now given by:

$$\phi_e(t) = \phi(t) + \phi_{e_\sigma}(t), \quad (30)$$

where $\phi(t)$ is as defined in (9) and

$$\phi_{e_\sigma}(t) = \varphi(t) \otimes e_{\sigma_t}. \quad (31)$$

If the least-squares algorithm is now applied in order to estimate the parameters we have:

$$\hat{\theta}_N^{LS} = \left[\frac{1}{N} \sum_{t=1}^N \phi_e(t) \phi_e^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \phi_e(t) y(t) \quad (32)$$

where it can be seen that now the measured version $\phi_e(t)$ is used. Substituting in (8) for $y(t)$ gives:

$$\begin{aligned} \hat{\theta}_N^{LS} &= \left[\frac{1}{N} \sum_{t=1}^N \phi_e(t) \phi_e^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \phi_e(t) (\phi^T(t) \theta_0 + e_0(t)) \\ &= \theta_0 - \\ &\quad \left[\frac{1}{N} \sum_{t=1}^N \phi_e(t) \phi_e^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \phi_e(t) (\phi_{e_\sigma}^T(t) \theta_0 - e_0(t)). \end{aligned} \quad (33)$$

It is clear that this expression has an extra term over (21), which comes from the fact that not only the output $y(t)$ is noisy but also the scheduling parameter signal. This can be recognised as an errors-in-variables type problem.

For the estimate to be consistent we require, in a similar fashion to before, that:

$$\text{i) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi_e(t) \phi_e^T(t) \text{ be nonsingular.} \quad (34)$$

$$\text{ii) } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \phi_e(t) (\phi_{e_\sigma}^T(t) \theta_0 - e_0(t)) = 0. \quad (35)$$

As before, Condition i) is a persistency of excitation condition.

Condition ii) can, again, be analysed using Theorem 2B.1. We first consider the matrix $\phi_e(t) \phi_{e_\sigma}^T(t)$. This contains the following terms:

$$(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_{\sigma_{t,i}} y(t-n), \quad (36)$$

$$(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_{\sigma_{t,i}} u(t-n) \text{ and} \quad (37)$$

$$(\sigma_{t,j} + e_{\sigma_{t,j}}) u(t-p) e_{\sigma_{t,i}} u(t-n). \quad (38)$$

Considering (36), it is clearly equal to:

$$\begin{aligned} &\sigma_{t,j} [G_0(\sigma_{t-p}, q) u(t-p) + H_0(\sigma_{t-p}, q) e(t-p)] \\ &e_{\sigma_{t,i}} [G_0(\sigma_{t-n}, q) u(t-n) + H_0(\sigma_{t-n}, q) e(t-n)] \\ &+ e_{\sigma_{t,j}} [G_0(\sigma_{t-p}, q) u(t-p) + H_0(\sigma_{t-p}, q) e(t-p)] \\ &e_{\sigma_{t,i}} [G_0(\sigma_{t-n}, q) u(t-n) + H_0(\sigma_{t-n}, q) e(t-n)]. \end{aligned} \quad (39)$$

So, referring to Theorem 2B.1, we can write:

$$\begin{aligned} s_2(t) &= \begin{bmatrix} \sigma_{t,j} H_0(\sigma_{t-p}, q) e(t-p) \\ e_{\sigma_{t,i}} G_0(\sigma_{t-p}, q) u(t-p) \\ e_{\sigma_{t,j}} H_0(\sigma_{t-p}, q) e(t-p) \\ e_{\sigma_{t,i}} G_0(\sigma_{t-n}, q) u(t-n) \\ e_{\sigma_{t,i}} H_0(\sigma_{t-n}, q) e(t-n) \\ 0 \end{bmatrix} + \\ &\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_{t,j} G_0(\sigma_{t-p}, q) u(t-p) \end{bmatrix} = \begin{bmatrix} v_1^2(t) \\ v_2^2(t) \\ v_3^2(t) \\ v_4^2(t) \\ v_5^2(t) \\ v_6^2(t) \end{bmatrix} + \begin{bmatrix} w_1^2(t) \\ w_2^2(t) \\ w_3^2(t) \\ w_4^2(t) \\ w_5^2(t) \\ w_6^2(t) \end{bmatrix}. \end{aligned} \quad (40)$$

In order to write $v_3^2(t) = e_{\sigma_{t,j}} H_0(\sigma_{t-p}, q) e(t-p)$ and $v_5^2(t) = e_{\sigma_{t,i}} H_0(\sigma_{t-n}, q) e(t-n)$ it is necessary that e_{σ_t} and $e(t)$ are uncorrelated. This condition is reasonable so long as the scheduling parameter is not the system output $y(t)$. If they are correlated the expected value of their product is non-zero and does not satisfy the theorem's assumptions on the stochastic component. The ergodicity of the signals

used in the identification method is, thus, not provable in this case.

Additionally it is not possible to establish the ergodicity of the signals when the scheduling parameter has the polynomial dependence discussed in (6) which is of a degree greater than 1. The reason is that we would have higher order moments of the noise term affecting σ_t , which are non-zero mean. This, in turn, would imply that $e_{\sigma_{t,j}}$ is non-zero mean and thus $v_2^2(t)$ and $v_4^2(t)$ are non-zero mean, violating the theorem's assumptions.

With these conditions in mind, we see that amongst the elements of $s_2(t) s_2^T(t)$ are all the cross-terms found in (39), and thus:

$$\begin{aligned} &\left\| \frac{1}{N} \sum_{t=1}^N [(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_{\sigma_{t,i}} y(t-n) \right. \\ &\quad \left. - E\{(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_{\sigma_{t,i}} y(t-n)\}] \right\| \rightarrow 0 \\ &\text{w.p. 1, as } N \rightarrow \infty. \end{aligned} \quad (41)$$

Similar results exist for (37) and (38), meaning that:

$$\begin{aligned} &\left\| \frac{1}{N} \sum_{t=1}^N \phi_e(t) \phi_{e_\sigma}^T(t) - \frac{1}{N} \sum_{t=1}^N E\{\phi_e(t) \phi_{e_\sigma}^T(t)\} \right\| \rightarrow 0 \\ &\text{w.p. 1, as } N \rightarrow \infty. \end{aligned} \quad (42)$$

Considering, for example, $E\{(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_{\sigma_{t,i}} y(t-n)\}$ from (41) it contains terms like

$E\{(e_{\sigma_{t,j}} G_0(\sigma_{t-p}, q) u(t-p)) (e_{\sigma_{t,i}} G_0(\sigma_{t-n}, q) u(t-n))\}$ which are not necessarily equal to zero. This, with other terms, will mean that the second sum in (42) will not, in general, be zero. This implies that the first sum will not be zero either.

Now examining the second term in (35) we have that the vector $\phi_e(t) e_0(t)$ contains the terms

$$(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_0(t) \text{ and } (\sigma_{t,j} + e_{\sigma_{t,j}}) u(t-p) e_0(t).$$

Again referring to Theorem 2B.1 we can write:

$$\begin{aligned} s_3(t) &= \begin{bmatrix} \sigma_{t,j} H_0(\sigma_{t-p}, q) e(t-p) \\ e_{\sigma_{t,j}} G_0(\sigma_{t-p}, q) u(t-p) \\ e_{\sigma_{t,j}} H_0(\sigma_{t-p}, q) e(t-p) \\ A_0(\sigma_t, q) H_0(\sigma_t, q) e(t) \\ 0 \end{bmatrix} + \\ &\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \sigma_{t,j} G_0(\sigma_{t-p}, q) u(t-p) \end{bmatrix} = \begin{bmatrix} v_1^3(t) \\ v_2^3(t) \\ v_3^3(t) \\ v_4^3(t) \\ v_5^3(t) \end{bmatrix} + \begin{bmatrix} w_1^3(t) \\ w_2^3(t) \\ w_3^3(t) \\ w_4^3(t) \\ w_5^3(t) \end{bmatrix}. \end{aligned} \quad (43)$$

The components of $s_3(t) s_3^T(t)$ include the terms $(\sigma_{t,j} + e_{\sigma_{t,j}}) y(t-p) e_0(t)$ and $(\sigma_{t,j} + e_{\sigma_{t,j}}) u(t-p) e_0(t)$. So it is possible to write:

$$\begin{aligned} &\left\| \frac{1}{N} \sum_{t=1}^N \phi_e(t) e_0(t) - \frac{1}{N} \sum_{t=1}^N E\{\phi_e(t) e_0(t)\} \right\| \rightarrow 0 \\ &\text{w.p. 1, as } N \rightarrow \infty. \end{aligned} \quad (44)$$

From (44) we have

$$E\{\phi_e(t) e_0(t)\} = E\{(\phi(t) + \phi_{e_\sigma}(t)) e_0(t)\} \quad (45)$$

which contains terms similar to those found in (27), so, as in Section 3, typically $E\{\phi(t) e_0(t)\} \neq 0$. Thus the second

sum, and so the first, in (44) will usually be non-zero. This, together with the result above for the first matrix in (35), means that Condition ii) is not normally satisfied and again the estimates obtained using the least squares algorithm are usually not consistent.

The instrumental variables method can be used to find a consistent estimate of the parameters in this case also. The estimate is given by:

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N} \sum_{t=1}^N \zeta(t)^T \phi_e(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \zeta(t) y(t). \quad (46)$$

This time, however, not only should $\zeta(t)$ be correlated with the regressor $\phi(t)$ and not with the noise $e_0(t)$, it should also be uncorrelated with $\phi_{e_\sigma}^T(t)$ in order that the algorithm be consistent.

5. SIMULATION RESULTS

Simulations are carried out to see how the identification techniques perform for a finite number of data.

5.1 Simulation 1

The case of non-noisy parameter scheduling variables is first tested in simulation. The true system is given by:

$$G_0(\sigma_t, q) = \frac{q^{-1}(b_0^0(\sigma_t) + b_1^0(\sigma_t)q^{-1})}{1 + a_1^0(\sigma_t)q^{-1} + a_2^0(\sigma_t)q^{-2}}. \quad (47)$$

The coefficient dependence on the scheduling parameter is chosen as polynomial in a single parameter, as in (6), giving:

$$\begin{aligned} a_1^0(\sigma_t) &= (1 - 0.5\bar{\sigma}_t + 0.2\bar{\sigma}_t^2) \\ a_2^0(\sigma_t) &= (0.5 - 0.7\bar{\sigma}_t - 0.1\bar{\sigma}_t^2) \\ b_0^0(\sigma_t) &= (0.5 - 0.4\bar{\sigma}_t + 0.01\bar{\sigma}_t^2) \\ b_1^0(\sigma_t) &= (0.2 - 0.3\bar{\sigma}_t - 0.02\bar{\sigma}_t^2). \end{aligned}$$

The true vector of parameters is, thus:

$$\theta_0^T = [1, -0.5, 0.2, 0.5, -0.7, -0.1, 0.5, -0.4, 0.01, 0.2, -0.3, -0.02]. \quad (48)$$

The signal $u(t)$ is chosen as a pseudo random binary signal (PRBS) with a shift register of length 12, giving $N = 4095$ points, and an amplitude of 1. $\bar{\sigma}_t$ is selected as a sinusoid with a period equal to N . It oscillates in the interval (0,1). In order for the algorithm to be consistent it is necessary that the persistency of excitation conditions are met. With the choice of $u(t)$ and $\bar{\sigma}_t$ used, this is easily the case. The noise on the output is taken such that $e(t)$ is a zero-mean, normally distributed white noise with a variance of 0.005, and $H_0(\sigma_t, q) = 1$. The least squares and instrumental variables method are tested. The instrumental variable vector is taken as the regression vector using $y(t)$ calculated from the model formed using $\hat{\theta}^{LS}$ i.e.

$$\zeta(t) = \varphi_{ls}(t) \otimes \sigma_t \quad (49)$$

where

$$\varphi_{ls}^T(t) = [-y_{ls}(t-1), -y_{ls}(t-2), \dots, -y_{ls}(t-n_a), u(t-d), u(t-d-1), \dots, u(t-d-n_b)] \quad (50)$$

It therefore satisfies the conditions in Section 3. The simulations are carried out 200 times to calculate the bias on the parameters found using the method, as well

Table 1. Results for Simulation 1 with noise-free σ_t

		LS	IV
Identification:	$\ \theta - E\{\hat{\theta}\}\ _2$	1.3505	0.0164
	$\ E\{(\hat{\theta} - E\{\hat{\theta}\})^2\}\ _2$	0.0112	0.0250
Validation:	$E\{\ y(t) - \hat{y}(t)\ _2\}$	0.0514	0.0099

as their variance. Table 1 shows the 2-norms of the bias and variance of the parameters found using the different methods. The expectation operators are calculated as averages over the simulations.

It can be seen from the table that the IV method can be considered to give unbiased parameter estimates compared to the LS method. The IV method does, however, have a larger variance, a known problem with these methods, which comes as a trade-off with their low computational complexity. This can often be improved using the Multi-step Algorithm in Ljung (1999).

A noise-free validation simulation is carried out for the θ estimated, using different signals for $u(t)$ and $\bar{\sigma}_t$ to those used in identification. They are 2 uniformly distributed random sequences varying in the intervals [-1,1] and [0,1] respectively, and of length 1024 points. The 2-norm of the prediction error is calculated for each simulation and the average values over the 200 simulations for each method are shown in Table 1. It can be seen that due to the fact that the parameters identified using the instrumental variables are asymptotically unbiased, the prediction error achieved in the validation experiment is much smaller using this method.

5.2 Simulation 2

Next a simulation is done to examine the case where the scheduling parameter is contaminated by noise. To be in accordance with the case analysed in Section 4 i.e. polynomial dependence no greater than affine, the system is given as:

$$G_0(\sigma_t, q) = \frac{q^{-1}(b_0^0(\sigma_t) + b_1^0(\sigma_t)q^{-1})}{1 + a_1^0(\sigma_t)q^{-1} + a_2^0(\sigma_t)q^{-2}} \quad (51)$$

where

$$\begin{aligned} a_1^0(\sigma_t) &= (1 - 0.5\bar{\sigma}_t), & a_2^0(\sigma_t) &= (0.5 - 0.7\bar{\sigma}_t), \\ b_0^0(\sigma_t) &= (0.5 - 0.4\bar{\sigma}_t), & b_1^0(\sigma_t) &= (0.2 - 0.3\bar{\sigma}_t). \end{aligned}$$

Thus, the true vector of parameters is this time given by:

$$\theta_0^T = [1, -0.5, 0.5, -0.7, 0.5, -0.4, 0.2, -0.3]. \quad (52)$$

The signals $u(t)$, $\bar{\sigma}_t$ and the noise on the output are chosen as before. The noise on $\bar{\sigma}_t$ is taken such that $e_{\sigma_{t,1}}$ is a zero-mean, normally distributed white noise with a variance of 0.0005. The least squares and instrumental variables method are tested. This time it is not possible to use the regression vector generated using the output of the model formed from $\hat{\theta}_{LS}$ as an instrumental variable vector. This is because it would still use σ_t^e from the first simulation. The instrumental variable vector is therefore formed by using the measured values of $y(t)$ and σ_t^e from a second simulation, which has different realisations of $e(t)$ and e_{σ_t} , and are therefore not correlated with those in the first simulation.

Table 2. Results for Simulation 2 with noisy σ_t

		LS	IV
Identification:	$\ \theta - E\{\hat{\theta}\}\ _2$	0.3161	0.0042
	$\ E\{(\hat{\theta} - E\{\hat{\theta}\})^2\}\ _2$	0.000505	0.0024
Validation:	$E\{\ y(t) - \hat{y}(t)\ _2\}$	0.0376	0.0051

The simulations are, again, carried out 200 times to calculate the bias on the parameters identified, as well as their variance. Table 2 shows the 2-norms of the bias and variance of the parameters found using the different methods.

Again the results show that the IV method can be considered to give unbiased parameter estimates compared to the LS method, though once more the estimate variance is larger.

A noise-free validation simulation is done again, using the same signals as in the validation in Simulation 1. The mean 2-norm values of the prediction error are shown in Table 2. Again we see that the prediction error achieved using the instrumental variables in the validation experiment is much smaller than that using the LS method.

6. CONCLUSIONS

In this paper the consistency of certain LPV identification methods is considered. It has been shown that ergodicity results, under certain conditions, can be applied to the signals in these LPV identification methods in order to carry out a consistency analysis.

When the scheduling parameter is noisy, it has been shown that an errors-in-variables type identification problem occurs and consistent estimates can be calculated using the IV method. In this case ergodicity was only demonstrable when the noise affecting the scheduling parameter and that affecting the output are uncorrelated. Furthermore, ergodicity is not provable when a polynomial dependence on the scheduling parameter of degree greater than 1 is present.

Further work would involve the analysis of the case where the scheduling parameter's noise is correlated with that of the output and also for the complete polynomial dependence case when noise is present on the scheduling parameter.

REFERENCES

- B. Bamieh and L. Giarré. Identification of linear parameter varying models. *International Journal of Robust and Nonlinear Control*, 12:841–853, 2002.
- L.H. Lee and K. Poolla. Identification of linear parameter-varying systems using nonlinear programming. *Journal of Dynamic Systems, Measurement and Control*, 121:71–78, 1999.
- L. Ljung. *System Identification - Theory for the User*. Prentice Hall, NJ, USA, second edition, 1999.
- M. Lovera, M. Verhaegen, and C.T. Chou. State space identification of MIMO linear parameter varying models. In *Proceedings of the International Conference on Mathematical Theory of Networks and Systems*, page 839842, Padova, Italy, 1998.

- M. Nemani, R. Ravikanth, and B.A. Bamieh. Identification of linear parametrically varying systems. In *34th IEEE Conference on Decision and Control*, pages 2990–2995, New Orleans, LA, USA, December 1995.
- W. J. Rugh and J. S. Shamma. Research on gain scheduling. *Automatica*, 36(10):1401–1425, October 2000.
- J. Shamma and M. Athans. Guaranteed properties of gain scheduled control for linear parameter-varying plants. *Automatica*, 27(3):559–564, 1991.
- V. Verdult. *Nonlinear system identification: A state-space approach*. PhD thesis, University of Twente, Enschede, The Netherlands, 2002.
- V. Verdult and M. Verhaegen. Subspace identification of multivariable linear parameter-varying systems. *Automatica*, 38:805–814, 2002.