Off-line State-dependent Parameter Models Identification using Simple Fixed Interval Smoothing

Elvis Omar Jara Alegria, Hugo Tanzarella Teixeira and Celso Pascoli Bottura
Semiconductors, Instruments, and Photonics Department, School of Electrical and Computer Engineering, State University of Campinas - UNICAMP, Av. Albert Einstein, N. 400 - LE31 - CEP 13081-970, Sao Paulo, Brazil

Keywords: Time Series Identification, State-dependent Parameter, Data Reordering.

Abstract: This paper shows a detailed study about the Young's algorithm for parameter estimation on ARX-SDP models and proposes some improvements. To reduce the high entropy of the unknown parameters, data reordering according to a state ascendant ordering is used on that algorithm. After the Young's temporal reordering process, the old data do not necessarily continue so. We propose to reconsider the forgetting factor, internally used in the exponential window past, as a fixed and small value. This proposal improves the estimation results, especially in the low data density regions, and improves the algorithm velocity as experimentally shown. Other interesting improvement of our proposal is characterized by the flexibility to the changes on the state-parameter dependency. This is important in a future On-Line version. Interesting features of the SDP estimation algorithm for the case of ARX-SDP models with unitary regressors and the case with correlated state-parameter are also studied. Finally a example shows our results using the INCA toolbox we developed for our proposal.

1 INTRODUCTION

Parameters of linear regression models can be satisfactorily estimated by using conventional estimation methods for Time Variable Parameters (TVP) based on least squares techniques, but its time parameter variations must be slow compared to the system state variations (Yaakov Bar-Shalom, 2001). But when the system presents State Dependent Parameters (SDP) (Priestley, 1988) the model response can be heavily nonlinear and even chaotic. The regression models with SDP, called SD-ARX (Priestley, 1988; Young et al., 2001) or quasi-ARX (Hu et al., 2001; F. Previdi, 2003), are always non-linear due to the product between the regressor function and the SDP. Conventional methods can't estimate SDP models because parameters vary very fast.

Young proposed a solution to SDP estimation (Young et al., 2001). It's based in a special temporal data reordering to smoothen the parameter variations. This proposed algorithm uses an optimal Fixed Interval Smoothing (FIS) algorithm as a first approximation and next a model with reordered data is recursively estimated. His justification for that data reordering is based on the fact that if this dependence between a parameter and a state exists, then both should react the same manner to this data reordering. Young shows that when the data and estimated parameters are returned to the normal temporal order, then the relationship among the SDP and the respective state is turned evident. When the time series are sorted in ascending order of magnitude, then the rapid natural variations are effectively eliminated from the data and replaced by much smoother and less rapid variations.

This paper studies the Young's algorithm in detail for SDP estimation and the forgetting factor α of the Exponentially-Weighted-Past (EWP) used on the FIS estimations with data reordering. We propose this factor should be reconsidered because: after the Young's reordering process, the old data not necessarily continues so. To obtain a more accurate and faster result, we propose using a fixed and low forgetting factor, here called as a filter factor, instead of using the current CAPTAIN tootlbox (Taylor et al., 2007) as it satisfies the continuous optimization of α for each iteration. The justification for this is probably that a small set of reordered samples corresponds with to the real samples. Then a small window past may be more useful in the data reordering case than a conventional optimized α. A practical advantage of this is the flexibility for structural changes modeling, it is

Omar Jara Alegria E., Tanzarella Teixeira H. and Pascoli Bottura C..
Off-line State-dependent Parameter Models Identification using Simple Fixed Interval Smoothing.
DOI: 10.5220/0005573903360341
In *Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2015), pages 336-341
ISBN: 978-989-758-122-9

important for a future On-Line version algorithm.

Two interesting cases of state-parameter dependence estimation of ARX-SDP models are shown also. The first one is when unitary regressors are considered and the other one is when the state and the parameter are correlated. Finally a example shows the use of our proposed Off-Line state-parameter dependence estimation algorithm using the INCA[1] (Alegria, 2015b).

2 SIMPLE FIXED INTERVAL SMOOTHING ALGORITHM

For the autoregressive with exogenous input (ARX) model:

$$y(k) = \mathbf{z}^T(k)\rho(k) + e(k); \qquad e(k) = N(0, \sigma^2) \qquad (1)$$

where,

$$\begin{aligned} \mathbf{z}^T(k) &= \begin{bmatrix} -y(k-1) & -y(k-2) & \cdots & -y(k-n) \\ u(k-\delta) & \cdots & u(k-\delta-m) & \end{bmatrix} \\ \rho(k) &= \begin{bmatrix} \rho_1(k) & \rho_2(k) & \cdots & \rho_{n+m+1}(k) \end{bmatrix}^T \end{aligned}$$

the Simple Fixed Interval Smoothing (SFIS) algorithm for TVP estimation could be used. The regression vector $\mathbf{z}(k)$ is composed of the scalar measure $y(k)$ and of the scalar exogenous input $u(k)$. The TVP are expressed as $\rho_i(k), i = 1, ..., n+m+1,$. Also, δ is a pure time delay on the input variable and $e(k)$ is a zero mean white noise.

The SFIS algorithm is a simple but useful recursive estimation method and can be obtained by a simple combination of the recursive estimation with forward data $\hat{\rho}_f$ and backward data $\hat{\rho}_b$, i.e. with data from sample k to N and from N to k respectively. To allow the TVP in each one of these cases, an Exponentially-Weighted-Past (EWP) with a fixed forgetting factor α is used. In the case of conventional FIS algorithm, an optimal value of α is obtained by the hyper-parameters optimization (Jazwinski, 2007; Young, 2011). For these reasons, in this paper we propose a SFIS instead of an optimal FIS algorithm. The cost function using the exponential weighting factor α is (Young, 2011):

$$\mathcal{J}_{EWP} = \sum_{i=1}^{k} \left[y(i) - \mathbf{z}^T(i)\hat{\rho} \right]^2 \alpha, \qquad 0 < \alpha < 1.0 \qquad (2)$$

[1]***I**dentificação **N**ão Linear para **C**ontrole **A**utomático* (INCA) toolbox we developed for our proposal for SDP estimation using SFIS algorithm

The SFIS algorithm for TVP estimation based on the *Recursive Least Squares* (RLS) method is:

Forward estimation $k = 1, 2, \cdots, N$:

$$\begin{aligned} \hat{\rho}_f(k) &= \hat{\rho}_f(k-1) + \mathbf{g}(k)\left\{ y(k) - \mathbf{z}^T(k)\hat{\rho}_f(k-1) \right\} \\ \mathbf{g}(k) &= \mathbf{P}(k-1)\mathbf{z}(k)\left[\alpha + \mathbf{z}^T(k)\mathbf{P}(k-1)\mathbf{z}(k) \right]^{-1} \\ \mathbf{P}(k) &= \frac{1}{\alpha}\left\{ \mathbf{P}(k-1) - \mathbf{g}(k)\mathbf{z}^T(k)\mathbf{P}(k-1) \right\} \end{aligned}$$

Backward estimation $k = N, N-1, \cdots, 1$:

$$\begin{aligned} \hat{\rho}_b(k) &= \hat{\rho}_b(k-1) + \mathbf{g}(k)\left\{ y(k) - \mathbf{z}^T(k)\hat{\rho}_b(k-1) \right\} \\ \mathbf{g}(k) &= \mathbf{P}(k-1)\mathbf{z}(k)\left[\alpha + \mathbf{z}^T(k)\mathbf{P}(k-1)\mathbf{z}(k) \right]^{-1} \\ \mathbf{P}(k) &= \frac{1}{\alpha}\left\{ \mathbf{P}(k-1) - \mathbf{g}(k)\mathbf{z}^T(k)\mathbf{P}(k-1) \right\} \end{aligned}$$

SFIS Estimation (Smoothing):

$$\hat{\rho}(k) = \frac{\hat{\rho}_f(k) + \hat{\rho}_b(k)}{2}$$

3 OFF-LINE IDENTIFICATION

An autoregressive with exogenous input and state dependent parameters (ARX-SDP) model, shown below in its simplest single input and singel output form, is characterized by the dependence among the parameter $\rho_i(k)$ and the state $x_i(k)$, $k = 1, 2, ..., N$, on the linear regression model:

$$y(k) = \mathbf{z}^T(k)\rho\{\mathbf{x}(k)\} + e(k); \qquad e(k) = N(0, \sigma^2) \qquad (3)$$

where,

$$\begin{aligned} \mathbf{z}^T(k) &= \begin{bmatrix} -y(k-1) & -y(k-2) & \cdots & -y(k-n) \\ u(k-\delta) & \cdots & u(k-\delta-m) & \end{bmatrix} \\ \rho\{\mathbf{x}(k)\} &= \begin{bmatrix} \rho_1\{x_1(k)\} & \rho_2\{x_2(k)\} & \cdots \\ \rho_{n+m+1}\{x_{n+m+1}(k)\} \end{bmatrix}^T. \end{aligned}$$

In this ARX-SDP model, the SDP ρ_i is assumed to be functions of only one state that could be one regression vector element or other variable that may affect the relationship between these two primary variables, $\rho_i(k)$ and $x_i(k)$, e.g. a regressors combination.

3.1 Recursive Off-line Identification of ARX-SDP Model

Here we the modify Young's SDP algorithm by using a SFIS algorithm. The reason of using a SFIS instead of an optimal FIS is because it allow the use of a fixed forgetting factor α, as justified in section 3.2. From equation (3), a first SFIS parameter estimation $\hat{\rho}_i^{first}$ is computed. This does not show any state-parameter dependence yet, but serves as a starting point to the recursive algorithm:

$$\hat{\rho}_i^k(k) = \hat{\rho}_i^{first}(k) \tag{4}$$

To analyze the parameters ρ_i of the ARX-SDP model one by one, first the Modified Dependent Variable y_{mdv_i} is calculated as follows:

$$y_{mdv_i}(k) = y(k) - \sum_{j \neq i} z_j(k)\,\hat{\rho}_j^k(k) \tag{5}$$

The signal y_{mdv_i} is obtained from equation (5); then the one-parameter model is:

$$y_{mdv_i}(k) = z_i(k)\,\rho_i\,\{x_i(k)\}\,. \tag{6}$$

Because signals y_{mdv_i}, z_i, x_i have high entropy, they are reordered based on the ascendant value of the state x_i. The justification for this reordering process is detailed in (Young et al., 2001) and an interesting observation is shown in the numerical example, see section 4. In this paper, the symbol $(^*)$ is used to represent the reordered data:

$$x_i \longrightarrow x_i^*, \qquad y_{mdv} \longrightarrow y_{mdv}^*, \qquad z_i \longrightarrow z_i^*.$$

then, the reordered equation (6) is:

$$y_{mdv_i}^*(k) = z_i^*(k)\,\rho_i^*\,\{x_i(k)\} \tag{7}$$

where, the parameter ρ_i^* is smoothed by the dependence with x_i; the signal y_{mdv_i} is lightly smoothed because it depends on ρ_i^* and also on the regressor z_i. The smoothing of the regressor z_i is dependent on the correlation among z_i and x_i. Experimentally, an uncorrelated selection of them is recommended.

From equation (7), a SFIS parameter estimation $\hat{\rho}_i^*$ is computed, it is reordered with respect to the normal ordering of the state x_i, i.e:

$$\hat{\rho}_i^*\,\{x_i(k)\} \longrightarrow \hat{\rho}_i\,\{x_i(k)\} \tag{8}$$

where, $\hat{\rho}_i$ is a more exact estimation than $\hat{\rho}_i^{first}$ and it is able to show dependency with the state x_i.

The steps shown from equations (4) to (8) describe the iterative procedure for the parameter ρ_i; this procedure should be used for the other parameters. Notice that only for the first iteration $\hat{\rho}_i^{first}$ is used.

3.2 Forgetting Factor Considerations

The Young's FIS estimation algorithm uses an *Exponential Windows Past* (EWP) with a *forgetting factor* α to consider parametric changes (Young, 2011).

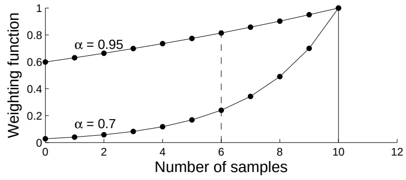

Figure 1: Weighting function for $\alpha = 0.95$ and $\alpha = 0.7$.

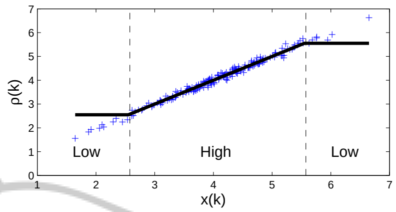

Figure 2: Example of a data density regions, and its SDP estimation using the CAPTAIN toolbox.

The FIS algorithm uses a variable factor $\alpha(k)$, initially smaller than α but approaches it asymptotically. In the CAPTAIN toolbox (Taylor et al., 2007), this value is optimized by maximum likelihood.

Here we propose an important consideration about this *forgetting factor* α because the data affected by it is reordered. This consideration is because the old data not necessarily remain old after the reordering process. Then, instead studying a method to choose a *forgetting factor* α suitable to the reordering process, this paper proposes to consider a small and constant value of α.

A small value of α is proposed because after the reordering transformation, only the nearby data probably corresponds to the unsorted data or at least to an *affine data*. For example, when $\alpha = 0.7$ the weight for the fourth older data is much lower than weight when $\alpha = 0.95$, see Fig. 1. In this sense, the forgetting factor here should be called a *filter factor* α.

We do not recommend a high value for the forgetting factor α because it could disregard sampled data with low density, e.g. the sampled data on the extremes of a Gaussian distribution, because after de reordering process, the low density data is averaged with the high state density. The Fig. 2 shows a example of data density distribution, divided by dashed lines and labeled as low and high, and the typical fault identification in the low data density region using Young's algorithm and the CAPTAIN toolbox.

3.3 Unit Regressors Model Considerations for SDP Estimation

This paper also considers the case of an ARX model without regressors, i.e. $z_i(k) = 1, k = 1, 2, \ldots, N$. In this case the equation model (3) with unitary regressors could be simply expressed as:

$$y(k) \quad = \quad \rho(k) + e(k) \qquad (9)$$

A practical advantage of considering unitary regressors is that the shape of the state-parameter dependency detection is better than when non-unitary regressors are considered. For example, let us analyze the equation (5) with unitary regressors:

$$y_{mdv_i}(k) \quad = \quad y(k) - \sum_{j \neq i} z_j(k)\, \hat{\rho}_j^k(k) \qquad (10)$$

$$y_{mdv_i}(k) \quad = \quad y(k) - \sum_{j \neq i} \hat{\rho}_j^k(k) \qquad (11)$$

We can see that, for unitary regressors, y_{mdv_i} is simplified to the measured signal subtracted by the effects of the other parameters ρ_j, $j \neq i$. Then y_{mdv_i} should be very similar to ρ_i.

Now, lets analyze the equation (6), see section 3.1; for unitary regressors $y_{mdv_i} = \rho_i\{x_i\}$ and the mean value $E[y_{mdv_i}] = E[\rho_i\{x_i\}]$. Then, the respective estimated parameter should have the same mean that y_{mdv_i} and it should be identical to the sum of the other parameters. Then it is obvious that an offset is generated for each parameter $\hat{\rho}_i(k) = \rho_i(k) + offset_i$.

Nevertheless, the estimated model is consistent with the input-output data, because when the estimated parameters are replaced on the regression model, the sum of all parameters offsets are compensated, i.e: $\sum_{i=1}^{n+m} offset_i \approx 0$

4 NUMERIC EXAMPLE

Initially we treat the estimation process to detect three state-parameter dependences in the model:

$$\begin{aligned} y(k) \quad = \quad & \rho_1\{x_1(k)\}\, u_1(k-1) + \rho_2\{x_2(k)\}\, u_2(k-1) \\ & + \rho_3\{x_3(k)\}\, u_3(k-1) + e(k) \qquad (12) \end{aligned}$$

The dependences to be estimate are:

$$\rho_1\{x_1(k)\} \quad = \quad sin\left(\frac{\pi}{2} x_1(k-1)\right) + 2 \qquad (13)$$

$$\rho_2\{x_2(k)\} \quad = \quad x_2^2(k-1) + 1 \qquad (14)$$

$$\rho_3\{x_3(k)\} \quad = \quad 0.8 x_3(k-1) + 3 \qquad (15)$$

where, the states $x_i; i = 1,2,3,$ are white noises; the inputs $u_i; i = 1,2,3,$ are white noises with displaced mean to $2, -1$ and 1 respectively and the measure noise $e(k); k = 1,2,\ldots,N; N = 1500,$ is white and zero mean. The signal-noise ratio (SNR) is 33% and all signals are uncorrelated.

Now, the detailed process from equations (5) to (8) is implemented and, in order to understand how the reordering process is useful to detect state-parameter dependence, the first iteration of the first parameter is detailed. The modified dependent variable y_{mdv_1}, corresponding to the first parameter ρ_1 is:

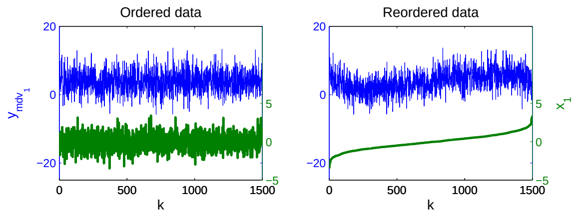

Figure 3: Temporally ordered data (left) an the effect of the reordering process for y_{mdv_1}.

$$y_{mdv_1}(k) = y(k) - \rho_2\{x_2(k)\}\, u_2(k-1) - \rho_3\{x_3(k)\}\, u_3(k-1) \qquad (16)$$

The signal y_{mdv_1} is the measure y without the effect provided by the parameters ρ_2 and ρ_3, each one multiplied by the regressors z_2 and z_3, respectively. The modified dependent variable in the reordered space corresponding to the first transformed parameter $y_{mdv_1} \rightarrow y^*_{mdv_1}$ is:

$$y^*_{mdv_1}(k) = \rho^*_1(k)\{x^*_1(k)\}\, u^*_1(k-1) \qquad (17)$$

We can say that the relationship $y^*_{mdv_1}$ has only the first parameter effect. Then the relation among x^*_1 and $y^*_{mdv_1}$ shows the shape of the dependence between x_1 and ρ_1 after of the reordering process. This effect is shown in Fig. 3: The original sinusoidal shape of the state-parameter dependence ρ_1 is evident after de reordering of y_{mdv_1} based on the ascendant value of the state x_1.

This process should be repeated for the other parameters ρ_2 and ρ_3 iteratively. Only for the first iteration $\rho_i^{first}, i = 1,2,3,$ is used. A stopping criteria for the algorithm should be a maximum iteration number or a minimal difference in two sequential estimations of ρ_i; in our tests 40 iterations were used for each case to contrast our results. This example was implemented in the INCA toolbox, for the cases of filter factor $\alpha = 0.9$ and $\alpha = 0.95$. The state-parameter dependence non-parametrically estimated is shown in Fig. 4.

In order to contrast graphically our dependence estimation, Fig. 5 shows the state-parameter dependences estimation using the function `sdp.m` of CAPTAIN. Note the differences of figures 4 and 5 especially on the low data density regions. The Table 1 shows results of the state-parameter dependency for the three unknown parameters using the INCA toolbox, i.e with low and fixed α, and the CAPTAIN toolbox, i.e. with optimized α.

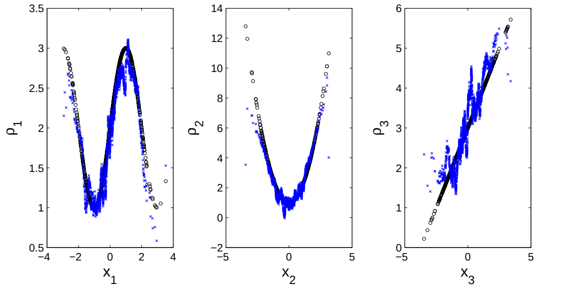

Figure 4: SDP estimation using the INCA for $\alpha = 0.95$ (blue) and reference (black).

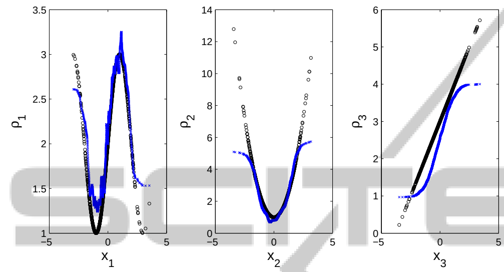

Figure 5: SDP estimation using the CAPTAIN (blue) and the reference (black).

Table 1: Estimation error of the state-parameter dependency using the INCA and CAPTAIN toolboxes.

SDP	Method	α	MAE$(\hat{\rho}_i - \rho_i)$	Time[s]
	INCA	0.90	0.4423	16
$\hat{\rho}_1$	INCA	0.95	**0.2917**	15
	CAPTAIN	-	0.7564	31
	INCA	0.90	0.4267	16
$\hat{\rho}_2$	INCA	0.95	0.3104	15
	CAPTAIN	-	**0.2956**	31
	INCA	0.90	0.2521	16
$\hat{\rho}_3$	INCA	0.95	**0.1665**	15
	CAPTAIN	-	0.2847	31

4.1 Estimation of ARX-SDP Model with Unitary Regressors

Now we consider the ARX-SDP model with unitary regressors:

$$\begin{aligned} y(k) &= \rho_1\{x_1(k)\} + \rho_2\{x_2(k)\} \\ &\quad + \rho_3\{x_3(k)\} + e(k) \end{aligned} \tag{18}$$

The dependences to be estimate are:

$$\begin{aligned} \rho_1\{x_1(k)\} &= u_1(k-1)\left(sin\left(\frac{\pi}{2}x_1(k-1)\right)+2\right) \\ \rho_2\{x_2(k)\} &= u_2(k-1)\left(x_2^2(k-1)+1\right) \\ \rho_3\{x_3(k)\} &= u_3(k-1)\left(0.8x_3(k-1)+3\right) \end{aligned}$$

The modified dependent variable y_{mdv} in this case is simply:

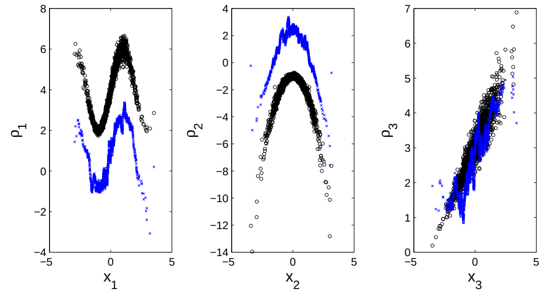

Figure 6: SDP estimation using the INCA toolbox, when model has unitary regressors (blue) and reference (black).

$$y_{mdv_1}(k) = y(k) - \hat{\rho}_2(k) + \hat{\rho}_3(k) \tag{19}$$

Notice that y_{mdv_1} is exactly the measurement signal without the other two parameters effect $\hat{\rho}_2$ and $\hat{\rho}_3$. Then the signal y_{mdv_1} should have shape similarity with the first parameter ρ_1. In the previous case, we can say that y_{mdv_1} also have ρ_1 information but perturbed by the regressor effects z_i, see equations (16) and (17). The dependence estimation result of this case is shown in Fig. 6.

We recommend to select a model with unitary regressors when the state-regressor dependence shape is more important than its exact value. e.g. in the case of parametric fault detection it could be more important to monitor the structure invariance of the model than its exact parameter value.

4.2 Estimation of ARX-SDP Model with Equal Regressors and States

Finally we consider ARX-SDP model with correlated regressors and states, exactly $z_i = x_i$:

$$\begin{aligned} y(k) &= \rho_1\{x_1(k)\}x_1(k-1) + \rho_2\{x_2(k)\}x_2(k-1) \\ &\quad + \rho_3\{x_3(k)\}x_3(k-1) + e(k) \end{aligned} \tag{20}$$

Notice that in equations (12) both signals are uncorrelated then the product of them keep, in some manner, the shape of the unknown parameter. In (20) the shape is completely affected because the parameter and the regressor are correlated. Thus, $\rho_i\{x_i(k)\}x_i(k-1)$ has a completely different shape from parameter $\rho_i\{x_i(k)\}$. In the models 12 and 18 the parameters and states were uncorrelated and their product shapes was conserved. For this case the modified dependent variable for the first parameter ρ_1 is:

$$\begin{aligned} y_{mdv_1}(k) &= \rho_1\{x_1(k)\}x_1(k) && (21) \\ &= x_1(k)\sin\left(\frac{\pi}{2}x_1(k)\right) + 2x_1(k) && (22) \end{aligned}$$

and its sinusoidal form isn't maintained, because the regressor $x_1(k)$ is correlated with the parameter

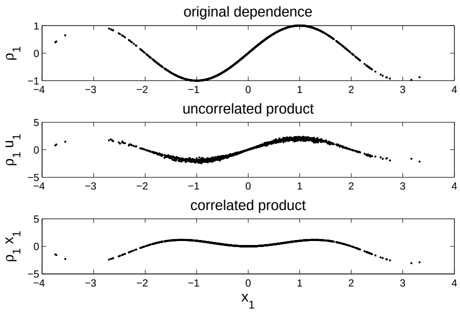

Figure 7: Original state-parameter dependence (top) and relation between state and regressor-parameter product, when the factors are uncorrelated (middle) and correlated (below).

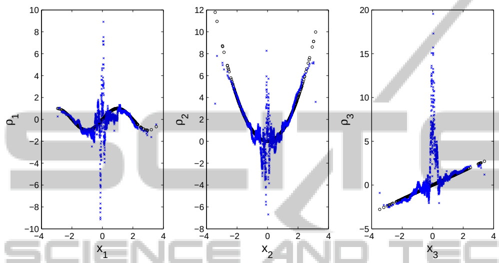

Figure 8: SDP estimation using the INCA toolbox, when state and parameter are correlated.

$\rho_1\{x_1(k)\}$, see Fig. 7. Finally Fig. 8 shows the SDP estimation for the three parameters when regressors and parameters are correlated.

5 CONCLUSIONS

Inspired on Young's algorithm, temporal data reordering strategy to reduce the data entropy, the state-parameter dependence estimation for ARX-SDP model is studied in this paper. Firstly, we proposed a fixed and relatively low value of the forgetting factor α instead of considering its optimal value. It showed a good estimation performance especially in the low density regions. It is because after de reordering process, the low density data is averaged with the high state density.

In spite of our proposal do not result in smoother state-parameter dependency as in CAPTAIN, this is not disadvantageous, because it is still necessary a parameterization stage, e.g. by using support vector regression (Alegria, 2015a). An important consequence of our proposal, low and fixed forgetting factor, is the flexibility to structural changes that our parameter-state dependency algorithm presents. This is very important for a future On-Line version of SDP estimation.

The three ARX-SDP estimation examples showed the usefulness of our proposal and implementation for the SFIS algorithm. For the first case the parameter and regressor are uncorrelated, and the SDP estimation results were good. For the second case with unitary regressors the SDP estimation results are also very good but with offsets. For the third case when the parameters and regressors are correlated the SDP estimation is poor. It is interesting to observe that, for the three parameters, the SDP estimation around the zero state has greater errors. This is due to the very low value of the regressor-parameter product that reduces the data richness.

ACKNOWLEDGEMENTS

We thank Brazilian agency CAPES for the financial support.

REFERENCES

Alegria, E. J. (2015a). State-dependent parameter models identification using data transformations and support vector regression. In *12th International Conference on Informatics in Control, Automation and Robotics ICINCO*.

Alegria, E. J. (2015b). State dependent parameters on-line estimation for nonlinear regression models. Master's thesis, Universidade estadual de Campinas - UNICAMP.

F. Previdi, M. L. (2003). Identification of a class of nonlinear parametrically varying models. *International Journal of Adaptive Control and Signal Processing*, 17:33–50.

Hu, J., Kumamura, K., and Hirasawa, K. (2001). A quasi-ARMAX approach to modeling of nonlinear systems. *International Journal of Control*, 74:1754–1766.

Jazwinski, A. H. (2007). *Stochastic Processes and Filtering Theory*. Dover publications.

Priestley, M. (1988). Non-linear and non-stationary time series analysis. *Academic Press, London*.

Taylor, C., Pedregal, D., Young, P., and Tych, W. (2007). Environmental time series analysis and forecasting with the captain toolbox. *Environmental Modelling & Software*, 22(6):797–814.

Yaakov Bar-Shalom, X. Rong Li, T. K. (2001). *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc. cop., New York (N.Y.), Chichester, Weinheim.

Young, P., McKenna, P., and Bruun, J. (2001). Identification of non-linear stochastic systems by state dependent parameter estimation. *International Journal of Control*, 74(18):1837–1857.

Young, P. C. (2011). *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, 2 edition.