

Identification of Polytopic Models for a Linear Parameter-varying System Performed on a Vehicle

Raluca Liacu, Dominique Beauvois and Emmanuel Godoy
Automatic Departement, Supelec Sciences of Systems (E3S), Paris, France

Keywords: LPV Model Identification, Polytopic Model, Automotive Identification.

Abstract: This paper deals with the parameter identification of continuous time polytopic models for a linear parameter-varying system (LPV). A continuous-time nonlinear identification approach is presented, a mix between a local approach and a global one is introduced in order to identify a LPV model for the lateral comportement of a vehicle. The proposed approach is based on the prediction error method for LTI systems, which is modified to take into account polytopic models and regularization terms. Using experimental data, different parameter-varying structures, explaining the lateral behavior of the vehicle, were identified by the proposed method considering the velocity as the scheduling parameter.

1 INTRODUCTION

In the last years, the field of Linear Parameter-Varying (LPV) systems has drawn a great attention in the industrial control with a growing number of successful applications.

Interest for such models is largely justified by the ability they offer to design gain scheduling control assuming guaranteed stability for nonlinear systems.

Classical LPV models are formed using polytopic structures constructed by a linear combination of multiple LTI models obtained at each operating point. The operating points are chosen from the range of the scheduling parameter. Such an experimental approach is not always feasible because of difficulties with experiments (safety, cost, duration...). For our application, identification of the lateral behavior of the vehicle, this leads us to an one-shot estimation of different LTI models based on a single record with a varying scheduling parameter.

Two situations were considered for the LTI models: the first situation assumes that we do not know how the model depends on the scheduling parameter. For this assumption we have created a fully parameterized structure which is suitable in any situation. In the second case, we consider that the dependence of the model on the scheduling parameter is known and the developed structure is based on the bicycle model.

In this paper the polytopic model is found using the prediction error method based on a quadratic criterion error in the time domain, leading to a global,

nonlinear optimization approach.

This paper presents adaptations of the standard prediction error method to the proposed polytopic structures.

The paper is organized as follows: in section 2 the polytopic model and the two LPV model structures are presented. Section 3 discusses the prediction error method for polytopic models. In section 4 the identification background is presented based on a real data experiment. In section 5 the results are analyzed and the two models for the lateral behavior of a vehicle are validated.

2 POLYTOPIC MODEL

A parameter varying system is considered. It is modeled in the state space and is described by the following equations:

$$\begin{cases} \frac{dx(t)}{dt} = A(p(t), \xi(p))(x(t)) + B(p(t), \xi(p))u(t) \\ \quad + K(p(t), \xi(p))e(t) \\ y(t) = C(p(t), \xi(p))(x(t)) + D(p(t), \xi(p))u(t) \\ \quad + e(t) \end{cases} \quad (1)$$

where $x(t)$, $y(t)$, $u(t)$, $e(t)$ represent the state variable, the output, the input and a stochastic white noise respectively. $p(t)$ is a measurable scheduling parameter with respect to time. $\xi(p)$ is the parameter vector which is varying with respect to p .

2.1 General Structure of a Polytopic Model

In this section the polytopic model for the considered system is presented. For each value p_i of p , the system is characterized by an LTI structure $(A_i(p_i, \xi_i), \dots, K_i(p_i, \xi_i))$, $\xi_i = \xi(p_i)$, named the i^{th} local LTI model. A polytopic structure is formed considering a state space representation given by:

$$\begin{aligned} A_c(p, \theta) &= \sum_{i=1}^r w_i(p) A_i(p_i, \xi_i) \\ B_c(p, \theta) &= \sum_{i=1}^r w_i(p) B_i(p_i, \xi_i) \\ C_c(p, \theta) &= \sum_{i=1}^r w_i(p) C_i(p_i, \xi_i) \\ D_c(p, \theta) &= \sum_{i=1}^r w_i(p) D_i(p_i, \xi_i) \\ K_c(p, \theta) &= \sum_{i=1}^r w_i(p) K_i(p_i, \xi_i) \end{aligned} \quad (2)$$

with $\theta = [\xi_1 \ \xi_2 \ \dots \ \xi_r]$ and $w_i(p)$ being the weighting functions satisfying the following relations :

$$w_i(p) \geq 0, \forall i, \quad \sum_{i=1}^r w_i(p) = 1 \quad (3)$$

where r is the number of operating points chosen to represent the model and p_i is the i -th operating point of the scheduled signal. Figure 1 shows the weighting functions chosen as triangular functions.

The objective of this paper is the following: using the input, output and the scheduled signal data (u, y, p) measured from the original system, estimate the vector θ in the polytopic form (2) so that the predicted output of the model denoted by $\hat{y}(t)$ is fitted to the output data $y(t)$ as closely as possible.

For simplicity, this approach is presented in the context of the determination of a second order model, which corresponds to a characterization of a lateral behavior of a vehicle. Two types of structures have been developed.

2.2 Fully Parameterized Model Structure

No knowledge about the dependence of the model matrices owing to the scheduling parameter is assumed; each local LTI model is chosen in a fully parameterized form. The weighting functions are those given previously. The LPV matrices are constructed as follows:

$$\begin{aligned} A_c(p, \theta) &= \sum_{i=1}^r w_i(p) A_i(\xi_i) \\ B_c(p, \theta) &= \sum_{i=1}^r w_i(p) B_i(\xi_i) \\ C_c(p, \theta) &= \sum_{i=1}^r w_i(p) C_i(\xi_i) \\ D &= 0 \end{aligned} \quad (4)$$

where

$$A_i = \begin{bmatrix} a_{11}^{(i)} & a_{21}^{(i)} \\ a_{12}^{(i)} & a_{22}^{(i)} \end{bmatrix}, \quad B_i = \begin{bmatrix} b_1^{(i)} & b_2^{(i)} \end{bmatrix}^T,$$

$$C_i = \begin{bmatrix} c_1^{(i)} & c_2^{(i)} \end{bmatrix}, \quad D_i = 0, \quad K_i = \begin{bmatrix} k_1^{(i)} & k_2^{(i)} \end{bmatrix}^T \quad (5)$$

and

$$\xi_i = \begin{bmatrix} A_i(\cdot)^T & B_i(\cdot)^T & C_i(\cdot)^T & K_i(\cdot)^T \end{bmatrix}, \quad (6)$$

where the notation (\cdot) is the Matlab operator which represents all the elements of a matrix, regarded as a single column.

$$\theta = \begin{bmatrix} \xi_1 & \xi_2 & \dots & \xi_r \end{bmatrix} \quad (7)$$

For a second order model the number of parameters to be identified is 10r.

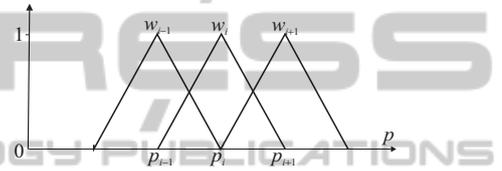


Figure 1: Triangular interpolative functions.

2.3 Analytical Bicycle Model

Usually, some analytical models can be established for a system, using physics of the phenomena. So, in our case, for small angles of deflection and drift, neglecting roll movement, the yaw angle of a vehicle is analytically described by an approximate model: the bicycle model (Mammar, 2006). This model (Figure 2) can be written as follows:

$$\begin{aligned} \begin{bmatrix} \dot{\beta} \\ \dot{\psi} \end{bmatrix} &= \begin{bmatrix} -\frac{c_f + c_r}{mv} & \frac{c_r l_r - c_f l_f}{mv^2} - 1 \\ \frac{c_r l_r - c_f l_f}{J_z} & -\frac{c_r l_r^2 + c_f l_f^2}{J_z v} \end{bmatrix} \begin{bmatrix} \beta \\ \psi \end{bmatrix} + \\ & \begin{bmatrix} \frac{c_f}{J_z} \\ \frac{c_f l_f}{J_z} \end{bmatrix} \delta; \quad \psi = [0 \ 1] \begin{bmatrix} \beta \\ \psi \end{bmatrix} \end{aligned} \quad (8)$$

where β is the slip angle, ψ the yaw movement, m the vehicle mass, v is the velocity, c_f and c_r the rigidities of the tires derivatives, l_f and l_r the distances between the front, rear wheels and the gravity center, J_z the yaw inertia and δ the steering angle of the front wheels. Defining $x(t)$, $y(t)$, $u(t)$ as

$$\text{aa} \quad x(t) = \begin{bmatrix} \beta \\ \psi \end{bmatrix}, \quad y(t) = \psi, \quad u(t) = \delta \quad (9)$$

a continuous time LPV model can be written as :

$$\begin{aligned} \frac{dx(t)}{dt} &= A_a(p, \theta^{(a)})x(t) + B_a(p, \theta^{(a)})u(t) \\ y(t) &= C_a x(t) + D_a u(t) \end{aligned} \quad (10)$$

with $p = \frac{1}{v}$ the scheduling parameter and $\theta^{(a)}$ defined as:

$$\theta^{(a)} = \begin{bmatrix} -\frac{c_f+c_r}{m} & \frac{c_r l_r - c_f l_f}{J_z} & \frac{c_r l_r - c_f l_f}{m} & -\frac{c_r l_r^2 + c_f l_f^2}{J_z} & \frac{c_f}{m} \\ \frac{c_f l_f}{J_z} \end{bmatrix} = \begin{bmatrix} \theta_1^{(a)} & \theta_2^{(a)} & \theta_3^{(a)} & \theta_4^{(a)} & \theta_5^{(a)} & \theta_6^{(a)} \end{bmatrix} \quad (11)$$

The model matrices have the following form:

$$A_a(p, \theta^{(a)}) = \begin{bmatrix} \theta_1^{(a)} p & \theta_3^{(a)} p^2 - 1 \\ \theta_2^{(a)} & \theta_4^{(a)} p \end{bmatrix}, \quad B_a(p, \theta^{(a)}) = \begin{bmatrix} \theta_5^{(a)} p \\ \theta_6^{(a)} \end{bmatrix}, \quad C_a = [0 \ 1], \quad D_a = 0. \quad (12)$$

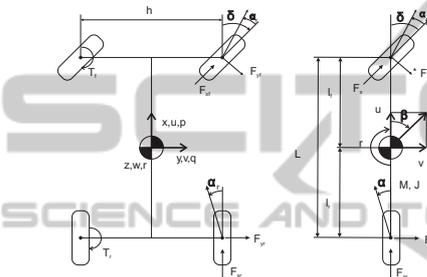


Figure 2: Bicycle model.

2.4 Structure based on the Bicycle Model

The approach is to use information available from an analytical model of the system. The structure corresponds to the case where dependence, owing to the scheduling parameter, is given to the deterministic part of each local LTI model of the polytopic structure in accordance with the structure of the analytical one. Using linear interpolation as shown in Figure 1 the LPV model matrices are constructed as follows:

$$\begin{aligned} A_c(p, \theta) &= \sum_{i=1}^r w_i(p) A_a(p_i, \theta^{(a)}) & C_c(p, \theta) &= C_a \\ B_c(p, \theta) &= \sum_{i=1}^r w_i(p) B_a(p_i, \xi_i) & D_c(p, \theta) &= D_a \end{aligned} \quad (13)$$

However, the stochastic part of each local LTI model is kept in a fully parameterized form:

$$K_c(p, \theta) = \sum_{i=1}^r w_i(p) K_i \quad (14)$$

So, each local model is function of the parameter

$$\xi_i = \begin{bmatrix} \theta^{(a)} & K_i(\cdot)^T \end{bmatrix}, \quad (15)$$

and the chosen structure of the LPV system is characterized by a parameter θ :

$$\theta = \begin{bmatrix} \theta^{(a)} & K_1(\cdot)^T & \dots & K_r(\cdot)^T \end{bmatrix} \quad (16)$$

The number of parameters to be identified is $6+2r$.

3 PREDICTION ERROR METHOD FOR POLYTOPIC MODELS

In this section, the prediction error method modified for polytopic models (Fujimori and Ljung, 2005) is presented. Compared to the case of LTI models, there are two specificities: the first is that the number of parameters to be estimated depends on the number of chosen operating points, as we saw in the previous section. The second one is an assumption on the discretization of the predictor and the gradient (details in section 3-1,2,3). An inconvenience treated here is that if we have a large number of operating points the calculation of the gradient numerically is time consuming, so we have adapted this procedure to our structure (details in section 3.3).

3.1 Predictor

In order to predict the output of the system, a Kalman predictor with the following structure is used:

$$\begin{cases} \frac{d\hat{x}(t, \theta)}{dt} = A_c(p, \theta)\hat{x}(t, \theta) + B_c(p, \theta)u(t) + \\ \quad + K_c(p, \theta)(y(t) - \hat{y}(t, \theta)) \\ \hat{y}(t, \theta) = C_c(p, \theta)\hat{x}(t, \theta) + D_c(p, \theta)u(t) \end{cases} \quad (17)$$

We can write the following equations in order to obtain the predictor:

$$\begin{aligned} \frac{d\hat{x}(t, \theta)}{dt} &= F_c(p, \theta)\hat{x}(t, \theta) + G_c(p, \theta)z(t) \\ \hat{y}(t, \theta) &= C_c(p, \theta)\hat{x}(t, \theta) + H_c(p, \theta)z(t) \end{aligned} \quad (18)$$

with the data set z defined as:

$$z(t) = \begin{bmatrix} y^T(t) & u^T(t) \end{bmatrix} \quad (19)$$

and

$$F_c(p, \theta) = \sum_{i=1}^r w_i(p) \{ A_i(p_i, \xi_i) - K_i(p_i, \xi_i) \sum_{j=1}^r w_j(p) C_j(p_j, \xi_j) \} \quad (20)$$

$$G_c(p, \theta) = \begin{bmatrix} G_{c1} & G_{c2} \end{bmatrix}$$

$$H_c(p, \theta) = \begin{bmatrix} 0 & \sum_{i=1}^r w_i(p) D_i(p_i, \xi_i) \end{bmatrix}$$

$$\begin{aligned} G_{c1} &= \sum_{i=1}^r w_i(p) K_i(p_i, \xi_i) \\ G_{c2} &= \sum_{i=1}^r w_i(p) \{ B_i(p_i, \xi_i) - \\ &\quad - K_i(p_i, \xi_i) \sum_{j=1}^r w_j(p) D_j(p_j, \xi_j) \} \end{aligned} \quad (21)$$

3.2 Discretization

For a sufficiently short sampling period of the data $z(t)$, $p(t)$ can be considered frozen during each sampling interval, and an Euler approximation (Heuberger, 2010) leads to the following discrete representation of (18)

$$p(t) = p^{(k)} = \text{const}; \quad kT \leq t < (k+1)T \quad (22)$$

$$\begin{aligned}\hat{x}((k+1)T, \theta) &= F_d(k, \theta)\hat{x}(kT, \theta) + G_d(k, \theta)z(kT) \\ \hat{y}(kT, \theta) &= C_d(k, \theta)\hat{x}(kT, \theta) + H_d(k, \theta)z(kT)\end{aligned}\quad (23)$$

$$\begin{aligned}F_d(k, \theta) &= I + TF_c(p^{(k)}, \theta); \quad G_d(k, \theta) = TG_c(p^{(k)}, \theta); \\ C_d(k, \theta) &= C_c(p^{(k)}, \theta); \quad H_d(k, \theta) = H_c(p^{(k)}, \theta)\end{aligned}\quad (24)$$

3.3 Prediction Error Method

Let N be the size of the sample data Z^N . The parameter, of the chosen structure, θ , is determined so as to minimize the sum of square of prediction errors $J_N(\theta, Z^N)$

$$J_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} e^T(t, \theta) e(t, \theta) \quad (25)$$

with $e(t)$ the prediction error vector defined as

$$e(t, \theta) = y(t) - \hat{y}(t, \theta). \quad (26)$$

The estimated parameter is obtained as:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} J_N(\theta, Z^N) \quad (27)$$

This optimization problem is solved using the Levenberg-Marquardt method (Roweis, 1996)(Osborne, 2007) which is supplied with the gradient of the prediction output, for sake of reduction of the calculation time (Ljung, 1999). This gradient

$$\Psi(t, \hat{\theta}) = \frac{\partial \hat{y}(t, \hat{\theta})}{\partial \hat{\theta}} \quad (28)$$

is obtained by the following discrete time state space representation:

$$\begin{aligned}\frac{\partial \hat{x}(k+1, \hat{\theta})}{\partial \theta_l} &= F_d(k, \hat{\theta}) \frac{\partial \hat{x}(k, \hat{\theta})}{\partial \theta_l} + \\ &+ \left[\frac{\partial F_d(k, \hat{\theta})}{\partial \theta_l} \quad \frac{\partial G_d(k, \hat{\theta})}{\partial \theta_l} \right] \begin{bmatrix} \hat{x}(k, \hat{\theta}) \\ z(k) \end{bmatrix} \\ \frac{\partial \hat{y}(k, \hat{\theta})}{\partial \theta_l} &= C_d(k, \hat{\theta}) \frac{\partial \hat{x}(k, \hat{\theta})}{\partial \theta_l} + \\ &+ \left[\frac{\partial C_d(k, \hat{\theta})}{\partial \theta_l} \quad \frac{\partial H_d(k, \hat{\theta})}{\partial \theta_l} \right] \begin{bmatrix} \hat{x}(k, \hat{\theta}) \\ z(k) \end{bmatrix}\end{aligned}\quad (29)$$

where θ_l represents the l^{th} component of the parameter vector. First of all the derivatives of

$$F_d(k, \theta), G_d(k, \theta), C_d(k, \theta), H_d(k, \theta) \quad (30)$$

were calculated by numerical approximations with the following expression:

$$\frac{\partial F_d(k, \theta)}{\partial \theta_l} \approx \frac{F_d(k, \theta + \delta_l) - F_d(k, \theta - \delta_l)}{2\delta_l} \quad (31)$$

where δ_l is a small value. We have noticed that this approach has a significant computational time, so we optimized the calculation time using an analytical determination for those derivatives.

Exploiting the form of the different matrices of the considered two polytopic structures, straightforward calculus leads to simple expressions for

$$\frac{\partial F_d(k, \theta)}{\partial \theta_l}, \frac{\partial G_d(k, \theta)}{\partial \theta_l}, \frac{\partial C_d(k, \theta)}{\partial \theta_l}, \frac{\partial H_d(k, \theta)}{\partial \theta_l} \quad (32)$$

these expressions let appear matrices with few non null elements depending on $T, w_i(p^{(k)}), w_{i+1}(p^{(k)})$ in case of the fully parameterized structure and on $T, w_i(p^{(k)}), w_{i+1}(p^{(k)}), p^{(k)}$ in case of bicycle model structure.

4 STUDY OF THE LATERAL BEHAVIOR OF A VEHICLE

4.1 Data for Parameter Estimation

An identification of LPV model corresponding to the lateral behavior of a vehicle is proposed. The vehicle used to provide the data is a standard Laguna II. The data were registered so as the experiment can be physically carried out; for sake of security, the test driver accepted to give "bang-bang" type movements to the steering wheel assuming some turns in the followed trajectory, while the vehicle speed is just decreasing from 110 Km/h to 30 km/h. In Figure 3 data recorded during experiments are given:

- first part gives the evolution of the control input of the model: the steering angle of front wheels.
- second part shows the scheduling parameter, the vehicle velocity decreasing from 110 Km/h to 30 Km/h. The velocity evolution was chosen such that the maneuver can be carried out.
- third part represents the measured output of the system: its yaw velocity.

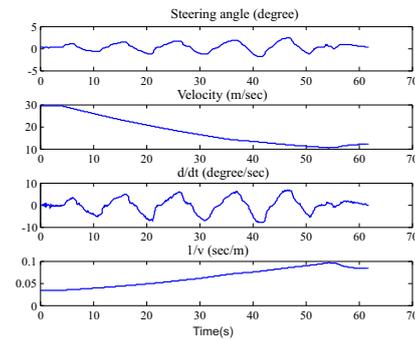


Figure 3: Experimental data.

As it has been mentioned previously the scheduling parameter is considered to be $1/v$, which is represented in the fourth part. The size of data is 50.000,

the sampling time is $T = 0.001$ sec. So, the assumption of a constant value of the scheduling parameter is clearly validated for this experiment.

4.2 Polytopic Structures Identifications

The two proposed structures are considered using the operating points corresponding to 30 Km/h, 50 Km/h, 110 Km/h. Optimization using Levenberg-Marquardt method needs an initialization of the parameter vector to identify; for both structures this initial vector was created from nominal values of $c_r, c_f, l_r, l_f, m, J_z$ given by the vehicle constructor.

Dealing with the fully parameterized polytopic structure, the problem consists in determining three parameter vectors ξ_i for the local LTI fully parameterized models, corresponding to the operating points, from data acquired during a global experience. As suggested in (McKelvey and Maureen, 1995), for this structure, it could be interesting to determine balanced realizations presenting a low sensitivity to parameters. A balanced realization is determined by balancing the observability and controllability gramians of the obtained state space model. Thus for this case, the prediction error method was slightly modified introducing a regularization term and the following algorithm has been used:

1. Obtain an initial state-space estimate θ_0
2. Convert the model to a balanced realization and let the corresponding parameter be θ_b^0
3. Let $I = 1$
4. Solve the minimization problem using Levenberg-Marquardt method:

$$\hat{\theta}_{N,\delta}^I = \underset{\theta}{\operatorname{argmin}} J_{N,\delta}(\theta)$$

$$J_{N,\delta}(\theta) = J_N(\theta, Z^N) + \frac{\delta}{2} \|\theta - \theta_b^{I-1}\|^2$$

5. Convert the obtained estimate to correspond to a balanced realization $\hat{\theta}_b^I$.
 if $J_{N,\delta}(\hat{\theta}_b^{I-1}) - J_{N,\delta}(\hat{\theta}_b^I) < e \rightarrow \text{fin}$;
 else $I = I + 1$ go to 4, where e is a *a priori* given constant.

5 RESULTS ANALYSIS

5.1 Identification Results

For each structure the identification algorithm has converged such that the estimated parameters lead to:

- insignificant prediction error, for the output ψ during the considered experience (with slow variation

of the velocity from 110 Km/h to 30 Km/h) as shown in Figure 4.

- reasonable errors simulating the experience with the obtained models using only the input data u_k, p_k .

Table I shows the Root Mean Square for the predicting and simulating errors. Better performances are obtained using the polytopic bicycle model structure, clearly because of introducing some a priori knowledge of the dependence on the scheduling parameter at the operating points.

Table 1: Root Mean Square.

Model	SMB	SFP
RMS(dgr/s) - predicted err.	0.0270	0.0277
RMS(dgr/s) - simulated err.	1.5851	3.4585

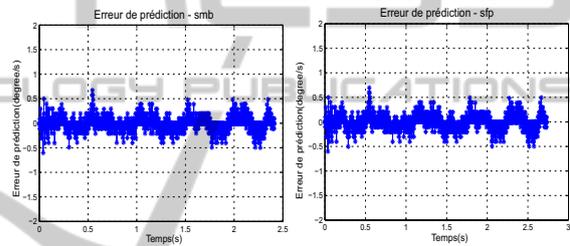


Figure 4: Comparison between the real vehicle output and the predicted outputs.

5.2 Validation

A new set of real data was used. We observe the results in time domain and analyze the model in the frequency domain.

5.2.1 Time Domain Analysis

The predictor was simulated with the new data set. As it can be seen in Figure 5 the outputs of the both structures reproduce well the system output. The notations Sfp correspond to the fully parameterized structure and Smb to bicycle model structure. In Figure 6, the model was simulated with the new data set. As we were looking for a predictor to minimize the prediction error, it is normal that the error between the predicted output of the model and the system output being more important. We can observe better performances in the case of the bicycle model structure, as we have more a priori information about the system. The validation also leads to insignificant prediction error, for any vehicle velocity and reasonable simulation error.

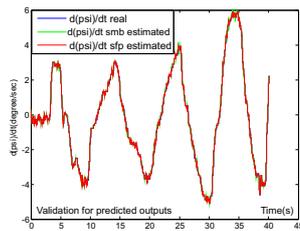


Figure 5: Comparison between the real vehicle output and the predicted output with a second set of data.

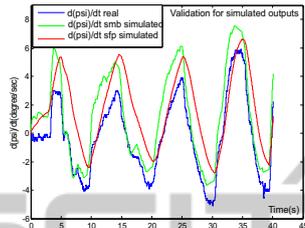


Figure 6: Comparison between the real vehicle output and the outputs simulated with a second set of data.

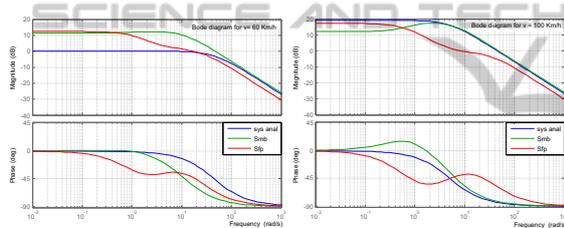


Figure 7: Bode diagrams for two arbitrary velocities.

5.2.2 Frequency Domain Analysis

Analysis of the analytical bicycle model, based on the physical parameters given by the vehicle constructor, can be done for the chosen operating points. The model is a second order with a zero that reduces the effect of the two poles, so it has the frequency response similar to that of a 1st order system. Bode diagrams (Figure 7) illustrate the influence of the scheduling parameter on the characteristics of the model, in the case of analytical model, high frequency behavior is the same whatever the velocity is (as expected from the fact that $C_a B_a$ is independent of the velocity), modifications of the velocity induce modifications of the static gain, and of the bandwidth in accordance. The frequency response of the two structures was compared with the Bode diagram of the analytical model, it can be observed that the identified polytopic bicycle model structure tracks well the analytical bicycle model dynamics at high frequencies, because of the coherence of the form of the CB products of the two structures; which is not the case for the fully parameterized polytopic model.

6 CONCLUSIONS

In this paper a LPV lateral behavior of a vehicle was represented by a polytopic model. It was considered that the system dependence on the scheduling parameter is known and a polytopic bicycle model structure was thus developed. The possibility of a model whose parameter dependence is not known was also taken into account, and a fully parameterized structure was created. Both models have been identified from real data set recorded during an experiment with a Laguna II vehicle. The pertinence of both identified structures was carried out with a validation data set. Better results were obtained with the polytopic bicycle model structure, knowing that *a priori* information of the system was available.

REFERENCES

- Fujimori, A. and Ljung, L. (2005). Parameter estimation of polytopic models for a linear parameter varying aircraft system. *submitted to Transactions of Japan Society for Aeronautical and Space Sciences (JSASS)*.
- Heuberger, P. S. C. (2010). On the discretization of lpv state-space representations. *Transformation*, 4(10):1–29.
- Ljung, L. (1999). *System Identification: Theory for the User (2nd Edition)*. Prentice Hall.
- Mammar, S. (2006). *Memoire presente en vue d'obtenir l'habilitation a diriger les recherches*. Universite d'Evry-Val d'Essone.
- McKelvey, T. and Maureen, T. (1995). *Identification of State-Space Models from Time and Frequency Data*.
- Osborne, M. R. (2007). Separable least squares, variable projection, and the Gauss-Newton algorithm. *Electronic Transactions on Numerical Analysis*, 28:1–15.
- Roweis, S. T. (1996). Levenberg-marquardt optimization.