

Identification de systèmes à commutation.

Abdelfettah Hocine, Didier Maquin, José Ragot

Centre de Recherche en Automatique de Nancy, CNRS UMR 7039

Institut National Polytechnique de Lorraine

2, Avenue de la forêt de Haye, 54516 Vandœuvre-les-Nancy Cedex, FRANCE

{Abdelfettah.Hocine, Didier.Maquin, José.Ragot}@ensem.inpl-nancy.fr

RESUME

L'objectif de cette étude est l'estimation de l'état d'un système à modes multiples connaissant ses entrées et ses sorties. La difficulté sous jacente est liée au fait que la matrice d'état du système n'est pas unique mais prend ses valeurs dans un ensemble fini à valeurs non connues. Le problème posé concerne donc la reconnaissance de la matrice d'état active à un instant donné, l'estimation des paramètres de cette matrice d'état et l'estimation de l'état du système. Se basant sur le principe utilisé dans l'approche EM (expectation maximization) mais sans utiliser de variables dites cachées pour affecter les données aux modes de fonctionnement, on propose une estimation simultanée des paramètres des modèles caractérisant les modes de fonctionnement et des instants de commutation d'un mode à un autre.

MOTS CLES : segmentation, commutation, changement de régime, mélange de modèles

ABSTRACT

Since several decades, researchers have been interested in various types of generalized regression models which admit changing parameters values at different time periods. The so-called regime switching models have given a lot of application in the fields of modelisation of complex systems, robust identification, detection of change of functioning and more generally in process diagnosis. Here we examine the case where the change in regime cannot be directly observed but may be estimated from observed variables (the input and the output of the process). For that purpose the well known EM approach may be applied ; to take into account the switches between the regimes, new variables (generally know as hidden or missing) are introduced in order to construct a complete data likelihood function. In this paper we show 1-how to directly formulate the estimation problem without introducing new variables, 2- a natural way to solve the obtained equations using hierarchical calculus. Some examples are given to illustrate how to use the proposed approach.

KEYWORDS : segmentation, data classification, switching, mixture of models

x	vecteur des observations
x_i	$i^{ème}$ composante de x
s_t	série temporelle
S_{t-1}	historique de la série $s_t : S_{t-1} = \{s_{t-i}\}_{i=1}^t$
θ	paramètres
σ^2	variance
α_i	paramètres de mélange
M	nombre de modèles locaux
N	nombre de mesures
$p(x/\theta)$	probabilité conditionnelle de x
ϕ	vecteur de regression
Φ	critère
a	paramètres du modèle
i	rang de l'observation courante
j	rang du modèle courant
$L(\theta/X)$	fonction de vraisemblance de X

Table 1: Liste des symboles

INTRODUCTION

La modélisation des systèmes physiques est l'étape préalable et indispensable à toute tentative d'amélioration de fonctionnement. La grande majorité des techniques de diagnostic, de surveillance, de régulation et d'optimisation se base sur la connaissance d'un modèle formalisé ou non du système à étudier. Le cas des systèmes à état discret ou à état continu a fait et fait toujours l'objet de très nombreux travaux. Par contre, pour les systèmes présentant simultanément des états continus et des états discrets, les études sont plus récentes. Parmi les travaux précurseurs, on peut mentionner ceux de Hudson [8], de Goldfeld [7] qui ont donné les premiers principes d'une telle représentation. Depuis, de très nombreux travaux ont précisé et développé ce formalisme dont les applications sont nombreuses dès qu'un système présente, de façon naturelle ou provoquée, des changements de régime [2], [18], [19].

De façon un peu arbitraire, on peut classer en deux catégories les types de modèles à commutation :

- les modèles dont le changement de régime est déterminé par une variable observée. Une situation typique est celle où le changement est réalisé à partir d'un seuil que

franchit ou non une variable d'état connue du système.

- les modèles dont le changement de régime n'est pas directement observé. Dans ce cas, le processus générateur du changement de régime doit être lui-même modélisé à partir de variables mesurées.

On peut aussi noter que les modèles hybrides recouvrent ces deux situations où, de plus, les changements de régime peuvent être décrits par un processus markovien [4], [12]. Avant de représenter le fonctionnement d'un système par un modèle à commutation, la question préliminaire à se poser est l'existence effective de plusieurs régimes de fonctionnement. La situation la plus difficile est celle où l'on ne dispose que de mesures des entrées et des sorties du système sans autre information sur l'existence physique de plusieurs régimes ; il faut donc découvrir ces derniers uniquement à partir des mesures disponibles.

Dans la suite de l'exposé, on s'intéresse à une classe particulière de système hybride connue sous le nom PWA (Piece-Wise Affine) [5] [25] et de leur extension aux systèmes dynamiques PWARX (Piece-Wise AutoRegressive eXogeneous). Ce type de modèle est construit en partitionnant l'espace des régresseurs en un nombre fini de polyèdres et en affectant un modèle local affine à chaque région ainsi obtenue. Le problème d'identification qui en résulte est généralement non linéaire [1] ; ceci est dû aux conditions logiques qui génèrent les passages d'une région à une autre région traduisant le point de vue physique du changement de régime ou de mode de fonctionnement. La difficulté essentielle résulte du couplage entre l'estimation des paramètres des modèles locaux et l'estimation du domaine de validité de ces modèles locaux. Il convient donc, et cela fait l'objet de notre présentation, d'exploiter simultanément plusieurs concepts : la segmentation des données, la classification des données, l'estimation des paramètres. Les approches conventionnelles utilisent généralement une procédure à deux niveaux [3] : le premier effectue un partitionnement des données, le deuxième procède à l'estimation des paramètres des modèles locaux dans l'espace ainsi partitionné.

Dans ce qui suit, on présente une approche globale estimant de façon conjointe les paramètres des modèles locaux et les coefficients d'appartenance des observations aux modes de fonctionnement.

REPRESENTATION DES MODELES A COMMUTATION

Un système à commutations peut être représenté par la famille de mélange de modèles :

$$p(x, \theta) = \sum_{j=1}^M \alpha_j p(x|\theta_j) \quad x \in \mathcal{R}^d \quad (1a)$$

$$\sum_{j=1}^M \alpha_j = 1 \quad \alpha_j \in \{0, 1\} \quad j = 1..M \quad (1b)$$

où M et d sont des entiers positifs a priori connus, α_j est positif, chaque $p(x|\theta_j)$ est une fonction de densité de probabilité paramétrisée par $\theta_j \in \mathcal{R}^p$. Les paramètres du modèle sont $\psi = (\theta_1 \dots \theta_M, \alpha_1 \dots \alpha_M)$, x est une variable aléatoire. Le paramètre de mélange α_j peut être interprété comme la probabilité d'activation du $j^{ième}$ modèle local particulier, tandis que θ_j caractérise la structure de ce modèle local. Les paramètres de mélange prennent uniquement les valeurs 0 ou 1 puisque les systèmes sont à commutation ; notons cependant qu'une représentation semblable vaut pour les multi-modèles [11], [6], [13], [15] à condition que les paramètres de mélanges prennent leurs valeurs dans l'intervalle $[0, 1]$; il y a sans doute de nombreux rapprochements à établir entre multi-modèles et modèles à commutation.

Le problème d'identification est le suivant : connaissant le nombre M de modèles locaux, leurs ordres respectifs et un ensemble de mesures des entrées et sorties du système, déterminer : 1) les ensembles de mesures associées à chaque modèle local, 2) les paramètres θ_j de chaque modèle local, 3) les paramètres de mélange α_j . Bien évidemment la difficulté sous-jacente est le couplage entre classification des données et identification de paramètres. En effet, pour estimer les paramètres d'un modèle local, il faudrait connaître les données qui caractérisent ce modèle local et il serait simple d'alouer les données à un modèle local si celui-ci était déjà caractérisé.

ESTIMATION DES PARAMETRES BASEE SUR LE MAXIMUM DE VRAISEMBLANCE

Formulation du problème

La variable x est observé et l'on dispose des données x_i supposées indépendantes et identiquement distribuées selon la loi de distribution p . La séquence des observations x est considérée comme une réalisation particulière d'une variable aléatoire X :

$$p(X|\theta) = \prod_{i=1}^N p(x_i|\theta) \quad (2)$$

A partir des données disponibles, et en utilisant le modèle de mélange (1a) pour chaque observation x_i , la fonction de vraisemblance s'écrit :

$$L(\psi|X) = \prod_{i=1}^N \sum_{j=1}^M \alpha_j p(x_i|\theta_j) \quad (3)$$

avec $\psi = (\alpha_1, \dots, \alpha_M, \theta_1, \dots, \theta_M)$. Les paramètres optimaux sont définis par

$$\hat{\psi} = \arg \max_{\psi} \prod_{i=1}^N \sum_{j=1}^M \alpha_j p(x_i|\theta_j) \quad (4)$$

Equations d'optimalité

Afin d'expliciter l'estimateur précédent, on définit la fonction de Lagrange :

$$\Phi = \sum_{i=1}^N \log \left(\sum_{j=1}^M \alpha_j p(x_i | \theta_j) \right) - \lambda \left(\sum_{j=1}^M \alpha_j - 1 \right) \quad (5)$$

Par la suite, on utilisera la notation :

$$p_i(\theta_j) = p(x_i | \theta_j) \quad (6)$$

Les équations d'optimalité de Φ par rapport aux paramètres du modèles s'explicitent :

$$\frac{\partial \Phi}{\partial \alpha_l} = \sum_{i=1}^N \frac{p_i(\theta_l)}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} - \lambda = 0, \quad l = 1..M \quad (7a)$$

$$\frac{\partial \Phi}{\partial \theta_l} = \sum_{i=1}^N \frac{\alpha_j \frac{\partial p_i(\theta_j)}{\partial \theta_j}}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} = 0, \quad l = 1..M \quad (7b)$$

$$\frac{\partial \Phi}{\partial \lambda} = \sum_{j=1}^M \alpha_j - 1 = 0 \quad (7c)$$

Mise à part la valeur de λ qui s'explicité $\lambda = N$, la résolution du système (7) est en général délicate du fait du couplage non linéaire entre les différents paramètres des modèles locaux, des paramètres des distributions et des paramètres de mélange. Dans la suite, l'usage de distribution gaussienne permet d'expliciter davantage les équations d'optimalité et de proposer une méthode de résolution hiérarchisée.

Cas d'un modèle linéaire dynamique sous hypothèse gaussienne

Un modèle local à temps discret sera choisi sous la forme ARMA :

$$x(t) + \sum_{i=1}^n \gamma_i x(t-i) = \sum_{i=0}^n \delta_i u(t-i) + e(t) \quad (8)$$

où u est une entrée exogène, $e(t)$ est un bruit blanc de moyenne nulle et d'écart-type σ et où n et m sont les ordres a priori connus du modèle. La prédiction à un pas s'écrit :

$$\hat{x}(t) = \phi(t)^T \theta \quad (9a)$$

$$\phi(t) = (-x(t-1) \dots -x(t-n) \quad u(t-1) \dots u(t-m)) \quad (9b)$$

$$\theta = (\gamma_1 \quad \dots \quad \gamma_n \quad \delta_1 \quad \dots \quad \delta_m) \quad (9c)$$

On considère maintenant un ensemble de M modèles locaux fournissant les prédictions :

$$\hat{x}_j(t) = \phi_j(t)^T \theta_j, \quad j = 1..M$$

auxquelles on associe les erreurs de prédiction :

$$\epsilon_j(t) = \hat{x}_j(t) - x(t)$$

Avec l'hypothèse d'une distribution gaussienne des erreurs de modèle ϵ_j :

$$p(\epsilon_j | \sigma_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left(-\frac{1}{2} \left(\frac{\epsilon_j}{\sigma_j} \right)^2 \right) \quad (10)$$

La distribution de x est donc :

$$p(x | \theta_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left(-\frac{1}{2} \left(\frac{x - \varphi^T a_j}{\sigma_j} \right)^2 \right) \quad (11)$$

où $\theta_j = (a_j, \sigma_j)$.

$$\Phi = \sum_{i=1}^N \log \sum_{j=1}^M \alpha_j \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left(-\frac{1}{2} \left(\frac{x_i - \varphi_i^T a_j}{\sigma_j} \right)^2 \right) - \lambda \left(\sum_{j=1}^M \alpha_j - 1 \right) \quad (12)$$

Les équations d'optimalité (7a,7b) par rapport à α_j , σ_j et a_j deviennent :

$$\sum_{i=1}^N \frac{p_i(\theta_l)}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} - N = 0, \quad l = 1..M \quad (13a)$$

$$\sum_{i=1}^N \frac{\alpha_j p_i(\theta_j)}{\sigma_j^3} \frac{(x_i - \varphi_i^T a_j)^2 - \sigma_j^2}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} = 0 \quad j = 1..M \quad (13b)$$

$$\sum_{i=1}^N \frac{\alpha_j p_i(\theta_j)}{\sigma_j^2} \frac{\varphi_i (x_i - \varphi_i^T a_j)}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} = 0 \quad j = 1..M \quad (13c)$$

Pour résoudre ce système, il peut être utile de faire les remarques suivantes relatives à sa structure. On peut tout d'abord noter que (13b) donne une forme un peu plus explicite de σ :

$$\sigma_j^2 = \frac{1}{N} \sum_{i=1}^N \frac{p_i(\theta_j)}{\sum_{j=1}^M \alpha_j p_i(\theta_j)} (x_i - \varphi_i^T a_j)^2 \quad j = 1..M \quad (14)$$

De plus l'équation (13a), dans le cas où $M = 2$ peut s'écrire :

$$\frac{1}{\alpha + \gamma_1} \dots + \frac{1}{\alpha + \gamma_N} = 0 \quad (15)$$

avec les définitions :

$$\alpha = \alpha_1 \quad \gamma_i = \frac{p_i(\theta_1)}{p_i(\theta_1) - p_i(\theta_2)}$$

La résolution de (15) par rapport à α , à partir d'algorithmes à itération, ne pose pas de problème particulier. Le structure exhibée dans le cas $M = 2$ s'étend au cas $M > 2$.

Enfin, avec

$$w_{ij} = \frac{\alpha_j p_i(\theta_j)}{\sigma_j^2 \sum_{j=1}^M \alpha_j p_i(\theta_j)}$$

l'équation (13c) s'écrit de façon plus condensée :

$$\sum_{i=1}^N w_{ij} \varphi_j (x_i - \varphi_i^T a_j) = 0$$

et dont la résolution par rapport aux coefficients des modèles locaux a_j est directe :

$$a_j = \frac{\sum_{i=1}^N w_{ij} \varphi_i x_i}{\sum_{i=1}^N w_{ij} \varphi_i \varphi_i^T} \quad (16)$$

Finalement, un algorithme à itération directe peut être utilisé pour la résolution de (10) :

Initialisation : choisir $(\alpha_j, \sigma_j, a_j), j = 1..M$,

Mettre à jour les écart-types σ_j (14)

Mettre à jour les probabilités $\theta_j, i = 1..N, j = 1..M$

Mettre à jour les paramètres α_j de mélange (15)

Mettre à jour les paramètres a_j des modèles locaux (16).

Tester la convergence des paramètres

RESULTATS

Un système à deux modèles locaux a été simulé. L'entrée et la sortie sont représentées à la figure 1. La commutation du mode 1 au mode 2 se fait avec une période de 1, la commutation du mode 2 au mode 1 avec une période de 2. La figure 2 indique l'évolution du système dans l'espace d'état ; aucune particularité de commutation n'apparaît visuellement sur cette figure.

$$y_{t+2} = a_j y_{t+1} + b_j y_t + c_j u_t \quad (17)$$

$$\begin{cases} j = 1 & a_1 = 1.2 & b_1 = -0.9 & c_1 = 0.2 \\ j = 2 & a_2 = .25 & b_2 = -.75 & c_2 = 0.1 \end{cases}$$

A partir des mesures de l'entrée et de la sortie, les deux modèles locaux ont été parfaitement identifiés (table 2), les valeurs initiales choisies pour les paramètres des modèles (et issues d'un tirage aléatoire) étant sensiblement différentes des valeurs vraies. La convergence est obtenue en une dizaine d'itération (figure 4). La figure 3 compare les commutations réelles et celles qui ont été identifiées.

Dans le cas de mesures bruitées, l'estimation des paramètres des modèles locaux reste correcte, celle des instants de commutation devient par contre erronée. Pour le même système que précédemment, mais avec des commutations à plus basse fréquence et un bruit d'écart-type 0.01 la figure 5 indique les commutations vraies et estimées montrant ainsi les défauts d'estimation. En dépit des défauts d'estimation des instants de commutation, et bien que les paramètres des modèles locaux aient été correctement identifiés, on peut penser qu'une contrainte additionnelle sur la répartition des fréquences de commutation puisse améliorer l'estimation. En effet, si l'on impose

une durée minimale entre deux commutations, c'est-à-dire un temps de séjour minimal dans un mode de fonctionnement, un filtrage des commutations semble réalisable.

par.	a_1	b_1	c_1	a_2	b_2	c_2
vrai	0.800	-0.100	0.250	-0.750	0.500	-0.400
init.	-0.208	1.208	-0.098	1.098	-0.466	1.466
est.	0.800	-0.100	0.250	-0.750	0.500	-0.400

Table 2: Paramètres

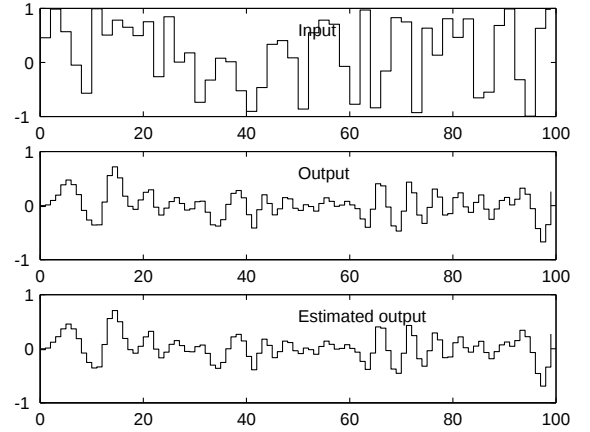


Figure 1: Entrée, sortie, sortie estimée

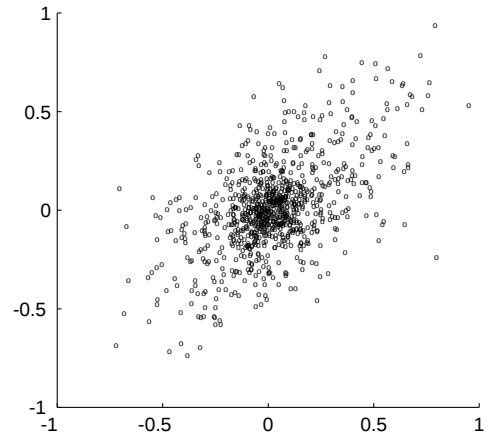


Figure 2: Espace d'état

DISCUSSIONS ET EXTENSIONS FUTURES

Bien qu'élaborée pour une classe particulière de systèmes dynamiques à commutation, la procédure proposée permet de déterminer les instants de commutation c'est-à-dire les instants de changement de régime si l'on s'intéresse au problème de surveillance. Il faut noter que le modèle du système n'est pas a priori connu et que la procédure estime de façon simultanée les paramètres

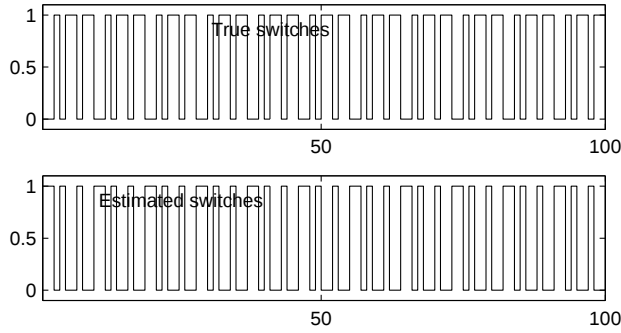


Figure 3: Signal de commutations réelles et estimées

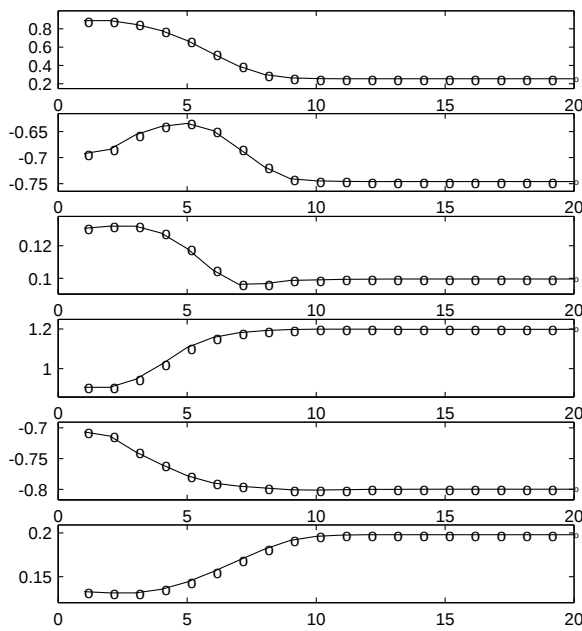


Figure 4: Vitesse de convergence des paramètres estimés

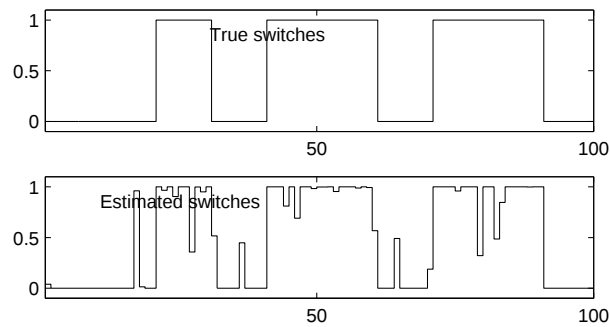


Figure 5: Signal de commutations réelles et estimées

des modèles et les instants de commutation. Ce résultat est obtenu en s'appuyant sur quelques hypothèses qu'il convient de discuter.

La première concerne le nombre a priori imposé de modes de fonctionnement, ce qui fixe ainsi le nombre de modèles locaux. C'est une hypothèse assez forte dans le cas de systèmes à comportement mal connus. Une solution envisageable est l'utilisation d'un nombre a priori grand de modèles locaux. Les premiers essais effectués dans ce sens montre qu'il est possible de déterminer ensuite les modèles voisins vis-à-vis d'un indice de similarité puis d'agréger ces modèles. La recherche du nombre de modèles locaux pourrait donc être déterminée de façon heuristique.

La deuxième hypothèse concerne l'ordre des modèles locaux. Elle paraît assez peu contraignante, car il reste possible d'utiliser des méthodes existantes de recherche d'ordre de modèle par exemple en se basant sur des critères de type Akaike.

La troisième hypothèse est celle de la distribution des erreurs de mesure. Tout d'abord, en ce qui concerne la loi utilisée ici (distribution normale), elle peut être remise en cause facilement, la procédure d'estimation développée ici pouvant s'adapter à d'autres distributions continues. Ce qui semble plus contraignant est sans doute le rapport signal à bruit. Les essais effectués montrent en effet une grande sensibilité du bruit vis-à-vis de l'estimation des instants de commutation et notamment à fréquence élevée de commutation ; ce constat est compréhensible car une erreur de mesure même de faible amplitude peut provoquer un franchissement de la surface de commutation, mais il conviendra sans doute d'améliorer la technique d'estimation peut être en remettant en cause l'utilisation du modèle prédictif à un pas au profit d'un modèle à erreur de sortie. De plus, une connaissance a priori de la distribution des fréquences de commutation pourrait être introduite.

Enfin, il semble envisageable d'identifier les paramètres de la loi de commutation lorsque celle-ci dépend de variables propres du système (entrée ou sortie). En effet, comme la procédure proposée identifie la séquence de commutation, un modèle explicatif vis-à-vis des mesures disponibles peut être recherché ; les premiers essais effectués dans ce sens confirment cette possibilité sous réserve de connaissance a priori de la classe des fonctions de commutation à utiliser ! Enfin, pour poursuivre ce travail, il conviendra de comparer l'approche proposée à celles existantes ; la récente publication de Juloski [10] peut constituer une base de comparaison.

References

- [1] A. Bemporad, J. Roll and L. Ljung. Identification of hybrid systems via mixed-integer programming. *40th IEEE Conference on Decision and Control*, pp. 786-792, 2001.
- [2] L. Breiman. Hinging hyperplanes for regression, classification and function approximation. *IEEE Transactions on Information Theory*, 39 (3), pp. 999-1013, 1993.
- [3] A.P. Dempster, N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm", *J. Royal Statistical Society, Series B* 39, pp. 1-38, 1977.
- [4] Z. Ding and L. Hong. An interactive multiple model algorithm with a switching markov chain. *Math. Comput. Modelling*, 25 (1), pp. 1-9, 1997.
- [5] G. Ferrari-Trecate, Muselli M., D. Liberati, M. Morari. Identification of piecewise affine and hybrid systems. *Proceedings of the american control conference*, pp. 3521-3526, 2001.
- [6] G. Gasso, G. Mourot and J. Ragot. Structure identification in multiple model representation: elimination and merging of local models. *IEEE Conference on Decision and Control*, Orlando, Floride, 2001.
- [7] S.M. Goldfeld, R.E. Quandt. A Markov model for switching regression. *Journal of econometrics*, 1, pp. 3-16, 1973.
- [8] D.J. Hudson. Finding segmented curves whose join points have to be estimated. *Journal of the american statistical association*, 61, pp. 1097-1129, 1966.
- [9] A. Juloski, W.P.M.H. Heemels, G. Ferrari-Trecate. Identification of an industrial hybrid system. *IFAC ADHS2003*, 16-18 June, Saint-Malo, France.
- [10] A. Juloski, W.P.M.H. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti, J.H.G.Niessen. Comparison of four procedures for the identification of hybrid systems. *Hybrid System Computation and Control*, Zurich, may 9-11, 2005.
- [11] McLachlan, G.J., Ng, S.K., and Peel, D. (2003). On clustering by mixture models. In *Studies in Classification, Data Analysis, and Knowledge Organization: Exploratory Data Analysis in Empirical Research*, O. Opitz and M. Schwaiger (Eds.). Berlin: Springer-Verlag, pp. 141-148.
- [12] Lee J.H. Non stationnary Markov switching models of exchange rates, the pound-dollar exchange rate. Dissertation, University of Pensilvania, 1991.
- [13] L. Mihaylova, V. Lampaert, H. Bruyninckx and J., Sweters. Hysteresis functions identification by multiple model approach. *International Conference MFI*, Baden-Baden, 2001.
- [14] E. Münz, T. Hodrus and V. Krebs. Top-down identification of hybrid characteristic maps. *Conference on analysis and design of hybrid systems*, pp. 34-39, Saint-Malo, France, June 16-18, 2003.
- [15] R. Murray-Smith and T.A. Johansen, *Multiple model approach in modelling and control*, Taylor and Francis, 1997.
- [16] D. Pearson, J. Ragot. Identifying parameters of local switching models: a geometrical approach. *11th International Workshop on Systems, Signals and Image Processing*, 2004
- [17] K. Plarre. A piecewise linear approximation of continuous functions and application to control. Final project ECE 417, 17 p., 2000.
- [18] P. Pucar and J. Sjöberg. On the hinge finding algorithm for higing hyperplanes. *IEEE Transactions of Information Theory*, 44 (3), pp. 1310-1319, 1998.
- [19] R.E. Quandt. The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the american statistical association*, pp. 873-880, 1958.
- [20] J. Ragot, A. Hocine, D. Maquin. Parameter estimation of switching systems. *International Conference on Computational Intelligence for Modelling Control and Automation*, 2004.
- [21] J. Ragot, D. Maquin and M. Pekpe. Signal segmentation and data classification. *10th International Workshop on Systems, Signals and Image Processing*, September 10-11, Prague, 2003.
- [22] J. Ragot, D. Maquin and E. Domlan. Switching time estimation of piecewise linear systems. Application to diagnosis. *Safeprocess*, 2003.
- [23] J. Roll. Robust verification and identification of piecewise affine systems. Thesis N 899, Linköping, 2001.
- [24] V. Verdult and M. Verhaegen. Identification of a weighted combination of multivariable local linear state-space sysems from input and output data. *40th IEEE Conference on Decision and Control*, pp. 4760-4765, 2001.
- [25] J.R. Yu, G.H. Tzeng and H.L. Li. General fuzzy piecewise regression analysis with automatic change point detection. *Fuzzy Sets and Systems*, 119, pp. 247-257, 2001.