

Egalisation de canaux non linéaires par filtrage de Kalman parallèle

Rim AMARA¹, Sylvie MARCOS¹

¹Laboratoire des signaux et systèmes
CNRS-Supelec, plateau de Moulon, 91192, Gif/Yvette, France
amara@lss.supelec.fr,marcos@lss.supelec.fr

Résumé – Dans ce papier, l'égalisation de canaux non linéaires est considérée dans un contexte d'estimation, où l'état inconnu à estimer est une séquence finie de symboles transmis. En linéarisant la fonction canal autour des états prédits, le Réseau de Filtres de Kalman (RFK) est étendu aux cas de canaux non linéaires (RFKE) [8]. Etant donné que la validité de cette linéarisation n'est pas toujours assurée, une variante du filtrage de Kalman proposée par Julier et al. [9] et basée sur un Monte-carlo déterministe pour le calcul des espérances conditionnelles mises en jeu dans le filtre de Kalman, a pu être adaptée dans ce même but.

Abstract – In this paper, nonlinear channel equalization is addressed in an estimation framework where the unknown state to estimate is a finite sequence of transmitted symbols. By linearizing the nonlinear channel function, a Network of Extended Kalman Filters (NEKF) based equalizer is developed [8]. Since the validity of the linearization is one of the major limitations of the EKF, a new Kalman filtering approach, the Unscented Kalman Filter (UKF) suggested by Julier et al. [9] and based on deterministic Monte-carlo simulation for conditionnal means determination, has been adapted to our equalization context.

1 Introduction

Le difficile problème de l'égalisation de canaux non linéaires a fait l'objet de recherches importantes depuis pas mal d'années sans encore donner entière satisfaction. En effet, dans certains domaines des communications numériques, comme les liaisons par satellite, l'enregistrement sur bande magnétique ...etc, le canal peut être modélisé par une fonction non linéaire, $h(\cdot)$, agissant sur une séquence finie de symboles transmis comme suit

$$y(k) = h(\mathbf{D}(k)) + b(k), \quad d(k) \in \gamma = \{d_i\}_{1 \leq i \leq q},$$

$\mathbf{D}(k) = [d(k), \dots, d(k-M+1)]^T \in \Gamma = \{\mathbf{D}_i\}_{1 \leq i \leq q^M}$, l'ensemble des états symboles possibles et $b(k)$ est un bruit blanc gaussien centré et de variance σ_b^2 . La fonction $h(\cdot)$ peut être une fonction polynômiale en les symboles, un filtre de volterra dont la sortie est une somme de produits croisés des symboles ou encore du type Linéarité + Nonlinéarité sans mémoire + Linéarité (LNL) [11]. Classiquement, les égaliseurs envisagés ont une structure paramétrée par un ensemble de paramètres W et de sortie $z(k) = g(y(k), \dots, y(k-m+1), W)$. W est optimisé pour produire, suivant un critère donné et après décision, la sortie désirée $d(k-D)$, où D est le retard d'estimation. En effet, à chaque valeur possible de $d(k-D)$ correspond une classe de vecteurs d'observations non bruités notée

$$\mathcal{C}_i = \{\hat{\mathbf{y}}(k) = [\hat{y}(k), \dots, \hat{y}(k-m+1)]^T | d(k-D) = d_i\}.$$

Optimiser, par exemple, les paramètres d'un filtre transverse de mémoire m correspond à chercher, dans l'espace \mathbf{R}^m , une droite séparatrice des classes $\{\mathcal{C}_i\}_{i=1, \dots, q}$. Ceci n'est pas toujours possible à partir du moment où suivant le canal (qu'il soit linéaire ou pas) et le retard D , les classes peuvent être linéairement séparables ou non.

Avec le DFE, la chance d'assurer une séparabilité linéaire dans l'espace des observations augmente grâce au retour sélectionnant plutôt des sous-classes à séparer [7], mais il reste encore sensible à la sévérité de la non linéarité. C'est pourquoi, les méthodes d'égalisation se sont orientées ces dernières années, vers les méthodes de classification et d'estimation Bayésienne pour améliorer les performances en terme de Taux d'Erreurs Binaire (TEB) [5][6][12][4]. Les égaliseurs Bayésiens sont basés sur la détermination récursive de la ddp *a posteriori* des symboles étant donné un horizon fini ou infini d'observations. Cette même ddp est utilisée par la suite comme fonction indicatrice de la classe à laquelle appartient le symbole qu'on désire estimer. Une autre approche utilisant la ddp comme information sur le signal est l'estimation d'état à partir d'une modélisation d'un système de communication numérique. Le filtre de Kalman constitue dans ce dernier contexte l'estimateur à minimum de variance de l'état dans le cas où le système est linéaire et gaussien. Pour l'égalisation, puisque l'état formé des symboles n'est pas gaussien, il a été suggéré dans [3] d'approximer $p(\mathbf{D}(k)|y^k)$ par une Somme Pondérée de Gaussiennes (SPG), dont les paramètres sont remis à jour à travers un Réseau de Filtres de Kalman (RFK). Pour un canal non linéaire, une linéarisation de $h(\cdot)$ autour des états prédits conduit à un struture d'égaliseur sous forme d'un Réseau de Filtres de Kalman Etendus (RFKE) [8] qui sera présenté dans la section 2. La validité de la linéarisation locale au premier ordre est en fait une limitation du Kalman étendu, ainsi une variante appelée "Unscented Kalman" [9] similaire aux Filtres par simulations de Monte-Carlo [13] est présentée dans la section 3 et son adaptation au cas de bruit d'état non gaussien est décrite dans le paragraphe 4. Enfin, nous donnons

notre conclusion.

2 L'égaliseur RFKE

L'égaliseur RFKE est basé sur une détermination récursive de la ddp *a posteriori* de l'état des symboles $p(\mathbf{D}(k)|y^k)$ où $y^k = \{y(0), \dots, y(k)\}$, approximée par SPG comme suit

$$p(\mathbf{D}(k)|y^k) = \sum_{i=1}^{\xi(k)} \alpha_i(k) \mathcal{N}(\mathbf{D}(k) - \hat{\mathbf{D}}_i(k|k), \hat{P}_i(k|k)) \quad (1)$$

L'état inobservable évolue dans un système de communication numérique suivant

$$\mathbf{D}(k+1) = F\mathbf{D}(k) + \mathbf{G}d(k+1), \quad (2)$$

$$y(k) = h(\mathbf{D}(k)) + b(k), \quad (3)$$

F est la matrice de transition à un pas et $\mathbf{G} = [1, 0, \dots, 0]^T$. Vu que le bruit d'état, affecté par le symbole futur, $d(k+1)$, prend des valeurs discrètes avec la même probabilité (les symboles sont i.i.d), la formulation d'état ci-dessus est celle d'un système non gaussien et non linéaire. Sorenson et Alspach ont proposé [1] d'approximer des densités non gaussiennes par des SPG et de linéariser les non linéarités (ici, celle de l'équation d'observation) autour de points de fonctionnement, correspondant à des états prédits comme c'est le cas pour le filtre de Kalman étendu [2]. Ainsi, il est montré dans [8, 3] que l'approximation de la ddp de l'état par une SPG se propage au cours du temps et que les paramètres de chaque SPG (sa moyenne et sa covariance) sont remis à jour à travers un filtre de Kalman étendu conditionné par une valeur d_i possible de $d(k+1)$. Les équations de l'algorithme du RFKE produisant l'estimée EQMM de l'état $\hat{\mathbf{D}}(k) = E\{\mathbf{D}(k)|y^k\}$ et la covariance de l'erreur d'estimation correspondante $\hat{P}(k)$ sont les suivantes

Etape de prédiction

$$\hat{\mathbf{D}}_i(k|k-1) = F\hat{\mathbf{D}}(k-1) + \mathbf{G}d_i \quad (4)$$

$$\hat{P}_i(k|k-1) = F\hat{P}(k-1)F^T + Q_i \quad (5)$$

$$e_i(k|k-1) = y(k) - h(\hat{\mathbf{D}}_i(k|k-1)) \quad (6)$$

$$\hat{\sigma}_i^2(k|k-1) = \mathbf{H}_i(k)^T \hat{P}_i(k|k-1) \mathbf{H}_i(k) + \sigma_b^2 \quad (7)$$

Etape de filtrage

$$\mathbf{K}_i(k) = \hat{P}_i(k|k-1) \mathbf{H}_i(k) / \hat{\sigma}_i^2(k|k-1) \quad (8)$$

$$\hat{P}_i(k|k) = (I_M - \mathbf{K}_i(k) \mathbf{H}_i(k)^T) \hat{P}_i(k|k-1) \quad (9)$$

$$\hat{\mathbf{D}}_i(k|k) = \hat{\mathbf{D}}_i(k|k-1) + \mathbf{K}_i(k) e_i(k|k-1) \quad (10)$$

Estimation EQMM de l'état

$$\beta_i(k) = \mathcal{N}(e_i(k|k-1), \hat{\sigma}_i^2(k|k-1)) \quad (11)$$

$$\alpha_i(k) = \frac{\beta_i(k)}{\sum_{j=1}^q \beta_j(k)} \quad (12)$$

$$\hat{\mathbf{D}}(k) = \sum_{i=1}^q \alpha_i(k) \hat{\mathbf{D}}_i(k|k) \quad (13)$$

$$\hat{P}(k) = \sum_{i=1}^q \alpha_i(k) \{ \hat{P}_i(k|k) \quad (14)$$

$$+ [\hat{\mathbf{D}}(k) - \hat{\mathbf{D}}_i(k|k)][\hat{\mathbf{D}}(k) - \hat{\mathbf{D}}_i(k|k)]^T \} \quad (15)$$

où $\mathbf{H}_i(k) = \frac{\partial h}{\partial \mathbf{D}(k)}(\hat{\mathbf{D}}_i(k|k-1))$ est la dérivée partielle de h évaluée en le *i^{ème}* état prédit $\hat{\mathbf{D}}_i(k|k-1)$ et Q_i intervient dans la SPG approximant la ddp du bruit d'état, donnée par

$$p(\mathbf{G}d(k)) = \sum_{i=1}^q \frac{1}{q} \mathcal{N}(\mathbf{G}d(k) - \mathbf{G}d_i, Q_i).$$

L'état estimé selon l'EQMM (13) est une combinaison convexe de q sorties de filtres de Kalman étendus.

Simulations : La figure (1) montre les bonnes performances du RFKE par rapport à l'égaliseur MAP optimal [4] et à l'égaliseur non linéaire développé par Lainiotis et al. dans [12]. Le canal LNL utilisé est supposé connu et donné par

$$\begin{aligned} y_1(k) &= \mathbf{C}^T \mathbf{D}(k), \quad \mathbf{C}^T = [1 \ 0.5] \\ y_2(k) &= y_1(k) + 0.1y_1^2(k) + 0.05y_1^3(k) \\ y(k) &= \gamma y_2(k) + b(k), \quad \gamma = 1 \end{aligned}$$

L'égaliseur de [12] est structuré en L filtres de Kalman étendus parallèles découplés, chaque filtre fournissant une estimation de l'état des symboles associée à une trajectoire nominale de l'état. Pour des symboles binaires, on remarque bien qu'avec seulement deux filtres de Kalman étendus, le RFKE fait mieux que l'égaliseur proposé dans [12] à 4 filtres parallèles.

Discussion : Notons que, l'état de symboles $\mathbf{D}(k)$ est un processus aléatoire qui prend des valeurs discrètes. La fonction $h(\cdot)$ n'est donc pas continue sur l'ensemble Γ et $h(\mathbf{D})|_{\mathbf{D} \in \Gamma}$ serait une distribution de points plutôt qu'une fonction en une variable aléatoire continue dans un région de \mathbf{R}^M . Par conséquent, linéariser la fonction canal $h(\cdot)$ autour des états prédits, qui sont des vecteurs dispersés dans \mathbf{R}^M , n'assure pas la validité de la linéarisation. Ainsi, les bonnes performances du RFKE seront dictées par une bonne étape de prédiction; cette dernière devant aboutir à un état prédit, $\hat{\mathbf{D}}_i(k|k-1)$, très proche du vrai état des symboles $\mathbf{D}(k)$. Dans ce cas, l'erreur de prédiction correspondant à cet état prédit sera négligeable et sa contribution dans l'estimée EQMM, à travers le poids $\alpha_i(k)$, sera maximale. Cette remarque nous a incité à envisager d'autres méthodes d'estimation d'état issu de systèmes non linéaires et plus particulièrement l'étude du "Unscented Kalman" développé par Julier et al. [9]. Cette variante est basée sur les récursions du filtre de Kalman en tant qu'estimateur linéaire à minimum de variance et sur du Monte-carlo déterministe pour le calcul des espérances conditionnelles mises en jeu.

3 "Unscented Kalman filter"

La logique du "Unscented Kalman Filter" (UKF) est qu'un nombre fini de points échantillons dans l'espace d'état peut être utilisé pour déterminer les espérances conditionnelles de l'état et de ses covariances; ainsi, il n'y a pas besoin de linéariser les non linéarités du système pour donner une forme explicite à ces moyennes et covariances. Pour une équation de transition d'état non linéaire

comme suit

$$\mathbf{D}(k+1) = f(\mathbf{D}(k), \mathbf{v}(k)), \quad \mathbf{v}(k) \sim \mathcal{N}(\mathbf{0}_q, Q_v)$$

Le modèle du processus est réécrit en l'état augmenté $\mathbf{D}^a(k) = [\mathbf{D}^T(k) \quad \mathbf{v}^T(k)]^T$ de dimension $n_a = M + q$,

$$\mathbf{D}(k+1) = f^a(\mathbf{D}^a(k)).$$

Les statistiques de l'estimateur à minimum de variance sont calculées à l'aide de $2n_a + 1$ points répartis autour de $\hat{\mathbf{D}}^a(k) = [\hat{\mathbf{D}}^T(k) \quad \mathbf{0}_q^T]^T$ (la moyenne conditionnelle de

$$\mathbf{D}_a(k)), \text{ avec une incertitude de } \hat{P}^a(k) = \begin{bmatrix} \hat{P}(k) & \hat{P}_{d,v}(k) \\ \hat{P}_{d,v}(k) & Q_v \end{bmatrix}$$

comme suit

- $\mathbf{D}_0^a(k|k) = \hat{\mathbf{D}}^a(k)$, $W_0 = \mathcal{K}/(n_a + \mathcal{K})$

- pour $i = 1, \dots, n_a$

$$\mathbf{D}_i^a(k|k) = \hat{\mathbf{D}}^a(k) + (\sqrt{(n_a + \mathcal{K})\hat{P}^a})_i, \quad W_i = 1/2(n_a + \mathcal{K})$$

$$\mathbf{D}_{i+n_a}^a(k|k) = \hat{\mathbf{D}}^a(k) - (\sqrt{(n_a + \mathcal{K})\hat{P}^a})_i, \quad W_i = 1/2(n_a + \mathcal{K})$$

$(\sqrt{(n_a + \mathcal{K})\hat{P}^a})_i$ est la i^{eme} colonne de la matrice racine carrée de $(n_a + \mathcal{K})\hat{P}^a$ et W_i est le poids correspondant à ce point (voir [9] pour le réglage de \mathcal{K}). L'algorithme se déroule comme suit

Etape de prédiction

$$\mathbf{D}_i(k+1|k) = f^a(\mathbf{D}_i^a(k|k)) \quad (16)$$

$$\hat{\mathbf{D}}(k+1|k) = \sum_{i=0}^{2n_a} W_i \mathbf{D}_i(k+1|k) \quad (17)$$

$$\mathbf{E}_i(k+1|k) = \mathbf{D}_i(k+1|k) - \hat{\mathbf{D}}(k+1|k) \quad (18)$$

$$\hat{P}(k+1|k) = \sum_{i=0}^{2n_a} W_i \mathbf{E}_i(k+1|k) \mathbf{E}_i^T(k+1|k) \quad (19)$$

$$\hat{y}_i(k|k-1) = h(\mathbf{D}_i(k+1|k)) \quad (20)$$

$$\hat{y}(k|k-1) = \sum_{i=0}^{2n_a} W_i \hat{y}_i(k|k-1) \quad (21)$$

$$e_i(k+1|k) = \hat{y}_i(k+1|k) - \hat{y}(k+1|k) \quad (22)$$

$$P_{ee}(k+1|k) = \sum_{i=0}^{2n_a} W_i e_i^2(k+1|k) + \sigma_b^2 \quad (23)$$

$$P_{ed}(k+1|k) = \sum_{i=0}^{2n_a} W_i e_i(k+1|k) \mathbf{E}_i(k+1|k) \quad (24)$$

Etape de filtrage

$$e(k+1|k) = y(k+1) - \hat{y}(k+1|k) \quad (25)$$

$$\hat{\mathbf{D}}(k+1|k+1) = \hat{\mathbf{D}}(k+1|k) + \mathbf{K}(k+1)e(k+1|k)$$

$$\hat{P}(k+1|k+1) = \hat{P}(k+1|k) - \mathbf{K}(k+1)P_{ee}(k+1|k)\mathbf{K}^T(k+1)$$

$$\mathbf{K}(k+1) = P_{ed}(k+1)/P_{ee}(k+1|k)$$

4 Algorithme de l'UKF appliqué à l'égalisation

L'utilisation de l'UKF pour contourner la linéarisation au niveau du RFKE a nécessité son adaptation au cas de

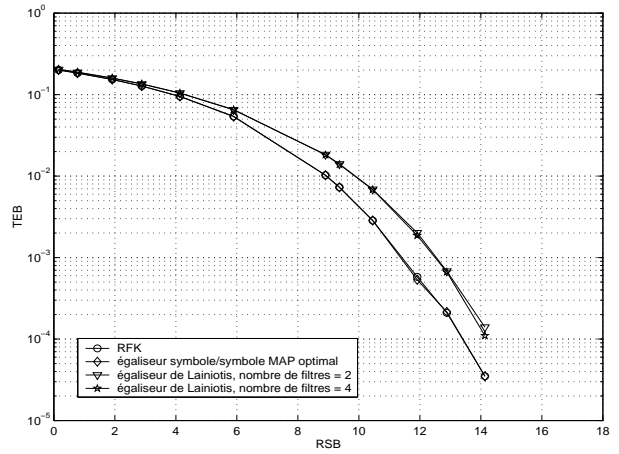


FIG. 1: $M = 2$ et $D = 0$

bruit d'état non gaussien. En effet, le bruit d'état dans (2) peut prendre q valeurs $\mathbf{G}d_i$ à chaque instant. Les équations (4)-(10) sont en fait celles d'un filtre de Kalman étendu pour un bruit d'état gaussien centré autour de $\mathbf{G}d_i$. Il est alors logique de suggérer l'utilisation des équations de l'UKF au lieu de l'EKF pour produire les états filtrés $\hat{\mathbf{D}}_i(k|k)$ (10). Ainsi, et de façon similaire au RFKE, un Réseau de q filtres UKF parallèles (RUKF) est construit; chacun faisant appel à un nuage de points dispersés autour d'un certain $\hat{\mathbf{D}}_i^a(k) = [\hat{\mathbf{D}}^T(k) \quad \mathbf{G}^T d_i]^T$ avec $Q_v = Q_i$. Chaque sortie d'un filtre UKF est ensuite pondérée par

$$\alpha_i(k+1) = \frac{\beta_i(k)}{\sum_{j=1}^q \beta_j(k)}, \quad \beta_i(k+1) = \mathcal{N}(e_i, P_{e_i e_i})$$

e_i est l'erreur de sortie prédite (25) conditionnée par $\mathbf{G}d_i$. L'estimée EQMM et sa covariance sont aussi données par les expressions (13) et (14).

Simulations: La figure (2) montre les courbes des TEB correspondant aux RFKE et le RUKF pour l'égalisation du canal polynomial

$$y(k) = \frac{1}{2}(\mathbf{C}^T \mathbf{D}(k))^3, \quad \mathbf{C} = [0.4084; 0.8164; 0.4084]^T$$

Dans le cas binaire $d(k) \in \{-1, +1\}$, les puissances paires des symboles sont égales à 1 et les puissances impaires au symbole même. Ainsi, la non linéarité polynomiale n'est en fait qu'un canal de Volterra d'ordre 3.

Un gain important est réalisé en terme de TEB pour une plage de (5-25 dB). Pour de forts RSB, les courbes de TEB se rejoignent, puisqu'à de faibles niveaux de bruits, le fonctionnement de l'algorithme du RFKE est piloté par l'étape de prédiction, qui est consistante dans le sens où l'un des états prédits est très proche de la valeur actuelle des symboles, produisant ainsi une erreur de sortie prédite négligeable.

Remarques :

- Suivant la formulation d'état (2)-(3), seule l'équation d'observation est non linéaire. Ainsi, nous pouvons garder les équations (4) et (5) pour le calcul des états prédits et

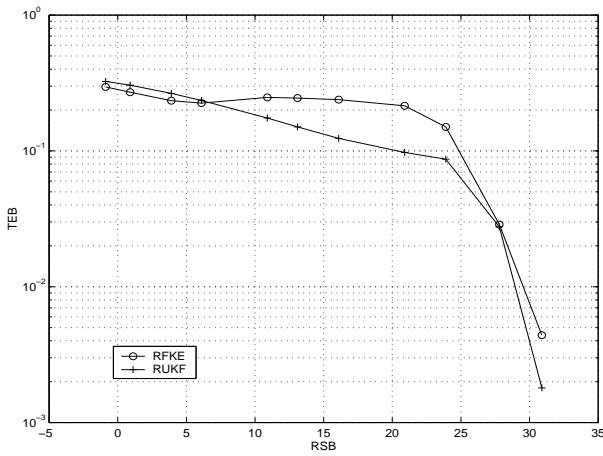


FIG. 2: $M = 3$ et $D = 2$

des matrices de covariances d'erreur de prédiction.

- La validité de la linéarisation menée au niveau du RFKE nécessite, principalement, une modélisation assez réaliste du canal (bien choisir la structure du canal) et son identification. En communications numériques, les effets du canal peuvent être décrits par sa réponse au nombre fini de séquences qui se présentent à lui, autrement, on peut se contenter de déterminer les sorties du canal non bruitées $\hat{y}_i = h(\mathbf{D}_i)$, $1 \leq i \leq q^M$. En s'intéressant à la seule identification des états de canal non bruités, l'effort de modélisation et d'identification est épargné d'une part et le risque d'une mauvaise linéarisation est écartée d'autre part. Avec le RUKF, le canal intervient au niveau du calcul des sorties prédites (20), un algorithme de clustering peut donc être utilisé pour identifier ces sorties [6].

- Pour des canaux modélisés par des filtres de Volterra, la sortie du canal peut être écrite comme suit

$$y(k) = \mathbf{f}^T \mathbf{X}(k) + b(k)$$

\mathbf{f} est le vecteur dont les composantes sont les noyaux du filtre de volterra d'ordre p et $\mathbf{X}(k)$ est le vecteur composé de tous les produits $d(k-i_1) \dots d(k-i_m)$, $1 \leq m \leq p$. En exploitant la linéarité du modèle par rapport à \mathbf{f} , l'identification des paramètres peut se faire par un simple LMS comme suit

$$\hat{\mathbf{f}}(k+1) = \hat{\mathbf{f}}(k) + \mu (y(k) - \hat{\mathbf{f}}(k)^T \mathbf{X}(k)) \mathbf{X}(k) \quad (26)$$

5 Conclusion

Le RFKE a été suggéré pour l'égalisation de canaux non linéaires à travers une approche par estimation d'une séquence de symboles issu d'un système non gaussien et non linéaire. Pour contourner la linéarisation au niveau du filtre de Kalman étendu, le filtre UKF peut être utilisé pour produire la sortie d'un filtre de Kalman du réseau. Le RUKF ainsi obtenu réalise de bonnes performances en terme de TEB et assure une estimation fiable de l'état. L'estimation par UKF étant basée sur le calcul des moyennes conditionnelles mises en jeu dans un algorithme de Kalman à l'aide d'un Monte-carlo déterministe; réduire le

nombre de points nécessaires pour le calcul des statistiques est possible d'après [10].

Références

- [1] D.L Alspach, H.W Sorenson. *Nonlinear Bayesian Estimation Using Gaussian sum Approximations*. IEEE Transactions on Automatic control, Vol. 17, Aout 1972.
- [2] A.H Jaswinski. *Stochastic processes and filtering theory*. Academic press 1970.
- [3] S. Marcos. *A network of adaptive Kalman filters for data channel equalization*. IEEE transactions on signal processing, September 2000.
- [4] R. Amara , S. Marcos. *Parallel Kalman filtering for optimal equalization*. submitted to IEEE Trans. on Signal Processing, Februray 2001.
- [5] B. Mulgrew. *Nonlinear signal processing for adaptive equalization and multiuser detection*. EUSIPCO 98, Rhodes, Greece, pp. 537-544, set. 1998.
- [6] S. Chen, B. Mulgrew and P.M Grant. *A clustering technique for digital communications channel equalization using Radial Basis Function Networks*. IEEE Trans. Neural Networks, vol. 4, no. 4, pp. 570-579, July 1993.
- [7] S. Chen, B. Mulgrew, L. Hanzo. *Asymptotic Bayesian DFE using a set of hyperplanes*. IEEE Trans. on Signal processing, vol.48, No.12, Dec. 2000.
- [8] P. Grohan, S. Marcos. *Nonlinear channel equalizer using gaussian sum approximations*. ICASSP, Munich, Avril 1997.
- [9] S.J. Julier, J.K. Uhlmann. *A New extension of the Kalman filter to nonlinear systems*. In Proc. of Aero.Sense, the 11th Int. Symp. on Aerospace.defense sensing, simulation and controls, 1997.
- [10] S.J. Julier. *Reduced sigma points filters for the propagation of means and covariances through nonlinear transformations*. Preprint submitted to Elsevier, Dec. 1998.
- [11] V.J. Mathews. *Adaptive polynomial filters*. IEEE Signal processing magazine 1991.
- [12] D.G. Lainiotis, P. Papaparaskeva. *A partitioned adaptive approach to nonlinear channel equalization*. IEEE Trans. commun., vol.40, no 10, October 1998.
- [13] H. Tanizaki. *Nonlinear Filters: Estimation and Applications, Lecture Notes in Economics and Mathematical Systems*. No.400, Springer-Verlag, 1993.