

Égalisation par Prédiction Basée sur des Réseaux de Neurons et des Fonctions Objectifs obtenues à partir de la Divergence de Kullback-Leibler et du Constant Modulus

Charles C. CAVALCANTE¹, João Cesar M. MOTA², Jugurta R. MONTALVÃO FILHO³, Bernadette DORIZZI⁴

¹DECOM, UNICAMP
BP 6101, 13083-970, Campinas-SP, Brésil

²DEE, UFC
BP 6001, 60455-760, Fortaleza-CE, Brésil

³CCFT, UNIT
Av. Murilo Dantas, 300, CEP: 49032-490, Aracaju-SE, Brésil

⁴EPH, INT
9 rue Charles Fourier, 91011, Évry Cedex, France
charles@decom.fee.unicamp.br, mota@dee.ufc.br
jugurta.montalvao@netdados.com.br, Bernadette.Dorizzi@int-evry.fr

Résumé – Une structure de filtrage non linéaire pour l'égalisation aveugle est présentée. Cette structure est basée sur un réseau de neurones, ce qui permet l'inclusion de non linéarités dans la structure du filtre. D'autre part, la stratégie d'apprentissage du réseau est séparée en deux parties : une supervisée et l'autre auto-organisée. La divergence de Kullback-Leibler est utilisée comme base pour une fonction de coût d'une règle d'apprentissage auto-organisée, tandis que le critère du « constant modulus » est utilisé dans la partie supervisée. Les résultats des simulations comparent la performance de cette stratégie par rapport aux stratégies classiques d'égalisation adaptative. Les résultats montrent que, pour certains canaux, la stratégie proposée est plus performante que l'égaliseur à retour des décisions (DFE) supervisé.

Abstract – A nonlinear structure of filtering for blind equalization is presented. The neural network-based structure is used in order to provide nonlinearity on the filter structure while the learning strategy is divided in two stages. The Kullback-Leibler divergence is used as the base for the cost function of a self-organized rule and constant modulus criterion for the supervised one. Simulation results illustrate the performance of the strategy compared to classical ones for adaptive equalization. The results show that the proposed strategy outperforms even trained DFE for some cases of channels.

1 Introduction

Dans le cadre de l'égalisation aveugle, il est bien connu que lorsque les symboles émis sont i.i.d., on peut utiliser la prédiction [1] dans le but d'optimiser les paramètres de l'égaliseur. De plus, en prenant le problème autrement, c'est-à-dire, en regardant l'égalisation comme une tâche de classification des observations à la sortie du canal, nous pouvons énumérer certains aspects intéressants [2, 3, 4]:

1. L'égalisation correspond à l'obtention d'une fonction séparatrice de classes ;
2. Les fonctions séparatrices les plus souhaitables sont, en général, non linéaires ;
3. Certains types de canaux ne peuvent pas être égalisés par des structures linéaires, comme le filtre transversal.

D'après les points 2 et 3, l'utilisation des structures non linéaires, au lieu des linéaires, semble être un choix pertinent pour des canaux « difficiles ». Ainsi, étant donné que l'approche d'égalisation par prédiction peut être vue comme un problème d'interpolation [4], nous nous sommes

intéressés à des structures capables d'engendrer les surfaces d'interpolation cherchées.

Par conséquent, nous avons proposé, dans [3], un égaliseur prédictif basé sur un réseau de neurones qui utilise une stratégie mixte d'apprentissage : une partie supervisée basée sur la minimisation de l'erreur de prédiction, et une autre basée sur la minimisation de la divergence de Kullback-Leibler. Cette stratégie nous a permis d'égaliser des canaux non minimum de phase (NMP) en utilisant seulement des statistiques d'ordre 2. Nous avons appelé cette structure « *Neural Predictive Structure* » (NPS). En plus, dû à l'utilisation de la minimisation de l'erreur de prédiction, cette stratégie est appelée « *NPS Minimizing Prediction Error* » (NPS-MPE).

Certains travaux proposent l'utilisation d'une classe de fonctions de coût basées sur la minimisation du critère du « *constant modulus* » (CMC) pour les stratégies prédictives - comme dans [6], par exemple. Les bons résultats obtenus nous ont poussé vers l'utilisation du CMC dans le NPS. Ce travail est donc, consacré à l'étude et à la comparaison de cette nouvelle stratégie, appelée désormais « *NPS using*

Constant Modulus Criterion » (NPS-CMC), avec d'autres stratégies d'égalisation, y compris la NPS-MPE.

Avant la présentation de ce travail, nous rappelons dans la section 2 la structure du NPS en mettant en évidence la stratégie mixte d'apprentissage. Dans la section 3 nous présentons la nouvelle stratégie. Les résultats des simulations sont présentés dans la section 4 pour l'évaluation de performance vis-à-vis des autres stratégies. Finalement, les conclusions et perspectives sont présentés dans la section 5.

2 Rappel sur le NPS

Soit $a(n)$ une variable aléatoire (v.a.), avec une loi de probabilité uniforme, dont les réalisations représentent les symboles émis. Soit aussi un canal dont la réponse impulsionnelle est donnée par

$$F(z) = \sum_{n=0}^{N-1} f_n z^{-n} \quad (1)$$

il est aussi possible de la représenter avec la notation vectorielle $\mathbf{f} = [f_0 \ \cdots \ f_{N-1}]^T$. Alors, on peut écrire les états non bruités du canal selon l'équation suivante

$$\bar{\mathbf{x}}(n) = \mathbf{F}^H \mathbf{a}(n) \quad (2)$$

où \mathbf{F} est la matrice de convolution du canal [2, 4] et $\mathbf{a}(n) = [a(n) \ a(n-1) \ \cdots \ a(n-N+1)]^T$. Donc, nous pouvons calculer les entrées de l'égaliseur comme:

$$\mathbf{x}(n) = \bar{\mathbf{x}}(n) + \mathbf{n}(n) \quad (3)$$

où $\mathbf{n}(n)$ est un vecteur de variables aléatoires normales (Gaussiennes) qui représentent le bruit.

Le modèle d'un égaliseur prédictif général est montré sur la figure 1.

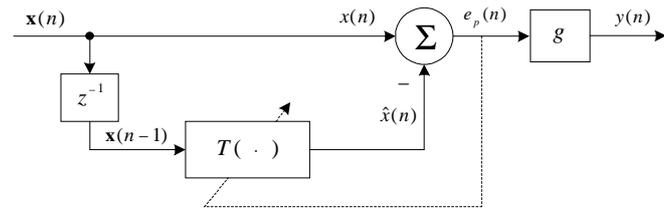


FIG. 1: Modèle d'un égaliseur prédictif.

Etant donné que nous sommes dans un cadre de prédiction, qui est lié à l'extraction de l'innovation, nous remarquons que l'utilisation d'une stratégie non linéaire permet l'usage d'une seule entrée pour le prédicteur, ceci veut dire que nous prédisons $x(n)$ par $T_{NL}(x(n-1))$, où $T_{NL}(\bullet)$ est une transformation non linéaire [3, 4], et g représente le gain responsable par récupérer la puissance des symboles émis.

Dans une approche de classification, cela correspond à trouver une fonction capable de séparer les différentes classes d'observations unidimensionnelles à la sortie du canal. Pour la communication numérique, ces classes correspondent aux symboles de l'alphabet de modulation [2]. Une illustration avec un alphabet binaire est montrée sur la figure 2.

La structure a été adoptée à cause de sa facilité d'implémentation d'une fonction comme celle représentée dans la figure 2 comme $\psi_{RN}(x(n-1))$.

Bien entendu, trouver automatiquement ce genre de fonction à partir seulement des observations (adaptation aveugle) n'est pas une tâche simple. Et pour y aboutir, nous avons développé un égaliseur prédictif basé sur des réseaux de neurones.

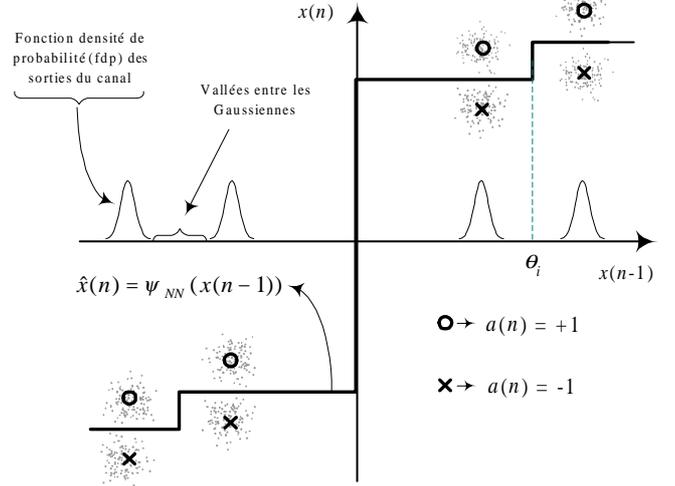


FIG. 2: Surface de séparation pour signaux binaires.

La figure 3 met en évidence la structure du NPS, où $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_M]^T$ est le vecteur de neurones de biais, $\varphi_{NPS}(\cdot)$ est la fonction seuil et $\boldsymbol{\beta} = [\beta_1 \ \cdots \ \beta_M]^T$ est le vecteur des coefficients linéaires des neurones.

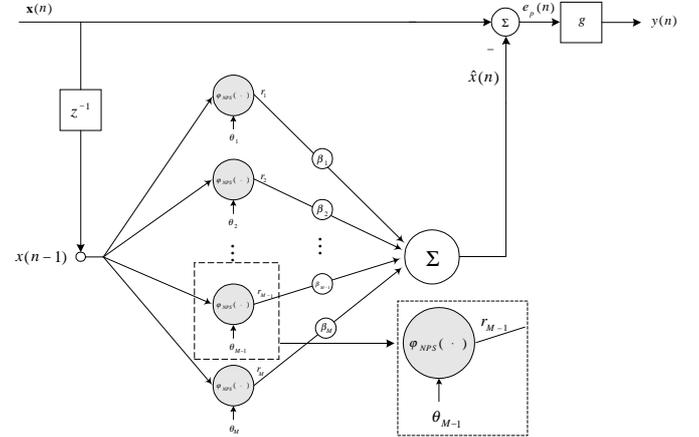


FIG. 3: Structure de la NPS (Neural Predictive Structure).

Cette structure étant choisie, il faut maintenant définir une stratégie d'apprentissage qui soit capable d'approximer proprement et rapidement la fonction cible $\psi(x(n-1))$ dans chaque cas. Nous avons choisi de couper ce problème en deux parties, à savoir:

1. L'apprentissage des paramètres $\boldsymbol{\theta}$;
2. L'apprentissage des paramètres $\boldsymbol{\beta}$,

Ceci sera décrit dans les prochaines sections.

2.1 L'algorithme auto-organisé

Il est connu que l'utilisation d'un modèle linéaire pour le canal et d'un modèle gaussien additif de bruit amène à une fonction de densité de probabilité (fdp) du signal reçu sous la forme d'un mélange de gaussiennes centrées sur les sorties non bruitées du canal (appelées désormais *états du canal*) [2, 4].

En plus, il est montré sur la figure 2 que les transitions rapides de la fonction $\psi(x(n-1))$ doivent être placées entre les gaussiennes de la fdp du signal reçu. Ainsi, dans la première partie de l'algorithme d'apprentissage du NPS, il faut d'abord trouver ces « vallées » entre les gaussiennes, ce qui est en rapport étroit avec les biais des neurones.

Pour ce faire, nous avons créé un algorithme adaptatif de détection de vallées du type LMS par la minimisation de la divergence de Kullback-Leibler [5] entre fonctions du type $|x - \theta_i|$ (fonctions en forme de 'V') et la fdp du signal reçu. Cet algorithme est auto-organisé parce qu'il n'utilise que les observations à la sortie du canal. En plus, il amène à une règle d'apprentissage du type Anti-Hebbienne [5]. Nous avons appelé cet algorithme SOFVA (*Self-Organized for Finding Valleys Algorithm*). Pour plus de détails voir [3, 4]. La fonction de coût est donnée par :

$$J_{\text{SOFVA}}(\boldsymbol{\theta}) = -\mathbb{E} \{ \ln(|x - \boldsymbol{\theta}| + \kappa) \} \quad (4)$$

où $\mathbb{E} \{ \cdot \}$ est l'opérateur espérance et où κ est une constante positive qui sert à assurer que la fonction est définie positive. Ainsi, le gradient de cette fonction par rapport au vecteur $\boldsymbol{\theta}$ est donné par :

$$\nabla J_{\text{SOFVA}}(\boldsymbol{\theta}) = -\mathbb{E} \left\{ \frac{\text{sgn}(x - \boldsymbol{\theta})}{|x - \boldsymbol{\theta}| + \kappa} \right\} \quad (5)$$

2.2 L'algorithme supervisé

Pour l'optimisation des paramètres β nous avons choisi une stratégie LMS classique pour la prédiction. C'est-à-dire, la minimisation de la puissance de l'erreur de prédiction. Cela correspond à un processus de blanchissement sur la sortie du prédicteur [7]. Il faut noter que cet algorithme est supervisé seulement dans le sens où le réseau des neurones dispose d'une sortie désirée, qui est la réalisation de $x(n)$. Cependant, l'égaliseur reste toujours autodidacte (ou aveugle) car il ne connaît pas le symbole émis.

La fonction de coût pour l'algorithme « supervisé » est

$$J_{\text{SUP-MPE}}(\boldsymbol{\beta}) = \mathbb{E} \left\{ |e_p(n)|^2 \right\} \quad (6)$$

et la version stochastique de l'algorithme d'adaptation est donnée par

$$\begin{aligned} e_p(n) &= x(n) - \psi(x(n-1), \boldsymbol{\theta}(n), \boldsymbol{\beta}(n-1)) \\ \nabla J_{\text{SUP-MPE}}(\boldsymbol{\beta}) &= e_p(n) \cdot \mathbf{r}(n) \\ \boldsymbol{\beta}(n+1) &= \boldsymbol{\beta}(n) + \mu \cdot \nabla J_{\text{SUP-MPE}}(\boldsymbol{\beta}) \end{aligned} \quad (7)$$

où $\psi(\cdot)$ est la fonction implémentée par la NPS, μ est le pas de convergence et $\mathbf{r}(n) = [r_1 \ \dots \ r_M]^T$ est le vecteur de sortie des neurones (voir la figure 3).

La fonction de coût pour le NPS-MPE est donc

$$\begin{aligned} J_{\text{NPS-MPE}}(\boldsymbol{\theta}, \boldsymbol{\beta}) &= J_{\text{SUP-MPE}}(\boldsymbol{\beta}) + J_{\text{SOFVA}}(\boldsymbol{\theta}) \\ J_{\text{NPS-MPE}}(\boldsymbol{\theta}, \boldsymbol{\beta}) &= \mathbb{E} \left\{ |e_p(n)|^2 \right\} + \mathbb{E} \{ \ln(|x - \boldsymbol{\theta}| + \kappa) \} \end{aligned} \quad (8)$$

3 La stratégie NPS-CMC

Inspirés par les bons résultats de [6], qui montrent une convergence plus rapide pour les égaliseurs prédictifs avec le CMC, nous avons aussi étudié l'apport d'un critère du type CM pour optimiser les coefficients β .

Pour la stratégie que nous avons appelée NPS-CMC, la fonction de coût de la partie supervisée a été remplacée par :

$$J_{\text{SUP-CMC}}(\boldsymbol{\beta}) = E \left\{ \left(|y(n)|^2 - R_2 \right)^2 \right\} \quad (9)$$

où $R_2 = \frac{\mathbb{E}\{|a(n)|^4\}}{\mathbb{E}\{|a(n)|^2\}}$ est le rayon d'égalisation.

Donc l'algorithme stochastique d'adaptation est :

$$\begin{aligned} e_p(n) &= x(n) - \psi(x(n-1), \boldsymbol{\theta}(n), \boldsymbol{\beta}(n-1)) \\ \nabla J_{\text{SUP-CMC}}(\boldsymbol{\beta}) &= g \cdot e_p(n) \cdot \mathbf{r}(n) \cdot \left(R_2 - |g \cdot e_p(n)|^2 \right) \\ \boldsymbol{\beta}(n+1) &= \boldsymbol{\beta}(n) + \mu \cdot \nabla J_{\text{SUP-CMC}}(\boldsymbol{\beta}) \end{aligned} \quad (10)$$

où g est le contrôle automatique de gain (CAG) adapté de l'algorithme présenté dans [1].

La stratégie NPS-CMC a donc les équations d'adaptation suivantes

$$\begin{aligned} J_{\text{NPS-CMC}}(\boldsymbol{\theta}, \boldsymbol{\beta}) &= J_{\text{SUP-CMC}}(\boldsymbol{\beta}) + J_{\text{SOFVA}}(\boldsymbol{\theta}) \\ J_{\text{NPS-CMC}}(\boldsymbol{\theta}, \boldsymbol{\beta}) &= \mathbb{E} \left\{ \left(|y(n)|^2 - R_2 \right)^2 \right\} \\ &\quad - \mathbb{E} \{ \ln(|x - \boldsymbol{\theta}| + \kappa) \} \end{aligned} \quad (11)$$

4 Simulations

Dans l'étude par simulations suivante, le but est la comparaison entre nos algorithmes et différents types de stratégies d'égalisation. Pour ce faire, nous avons fait 100 simulations de Monte-Carlo (avec la modulation BPSK) pendant lesquelles l'erreur de décision quadratique (EDQ) - $[\text{Dec}(y(n)) - y(n)]^2$, où $\text{Dec}(\cdot)$ est la décision - a été mesurée. Les résultats ont été lissés par un filtre passe-bas avec $\omega_c = 10^{-2}$.

Pour définir le « status » du canal, c'est-à-dire, si il est dans la condition de « l'œil ouvert » ou bien « l'œil fermé », nous avons utilisé le critère de distorsion maximale (MD) [2, 4].

Les autres algorithmes simulés pour comparaison sont : le CMA avec un filtre transversal, le prédicteur linéaire avec des statistiques d'ordre 2 et l'égaliseur à retour de décisions (DFE) supervisé. Comme référence pour les structures linéaires, la solution de Wiener [7] est aussi montrée.

En ce qui concerne l'initialisation du NPS-MPE et du NPS-CMC, pour que le SOFVA puisse aboutir à une bonne estimation des positions des « vallées » entre les gaussiennes, l'initialisation des biais des neurones joue un rôle très important. Par conséquent, nous avons choisi d'utiliser un assez grand nombre de neurones, dont les biais sont initialisés aléatoirement. Cela permet qu'au moins un certain nombre des transitions de la fonction $\psi(\cdot)$ soient bien placées avant même le démarrage de l'algorithme d'apprentissage auto-organisé.

Dans les simulations nous avons utilisé deux canaux minimum de phase: le premier sous la condition d'œil ouvert ($\mathbf{f}_1 = [1 \ 0.6 \ 0.2]^T$) voir figure 4, et l'autre sous la condition d'œil fermé ($\mathbf{f}_2 = [1 \ 0.8 \ 0.4]^T$) voir figure 5. Dans ces deux cas nous avons SNR = 40 dB.

Il est à noter, sur les figures 4 et 5, que le NPS-CMC est plus performant que les autres stratégies, y compris le DFE supervisé. La meilleure performance du NPS-CMC par rapport au NPS-MPE est due à l'utilisation d'une stratégie où l'erreur instantanée - utilisée pour l'adaptation stochastique des paramètres - converge vers le zéro (le critère CM), tandis que pour le MPE cette erreur converge vers une variable dont la variance est égale à celle du signal à l'entrée du canal, donc non nulle.

Enfin, d'autres simulations ont été faites avec un canal NMP ($\mathbf{f}_3 = [0.5 \ 1 \ -0.6]^T$), voir figure 6. Dans ce cas, nous n'avons pas eu une grande différence de performance entre le NPS-CMC et le NPS-MPE. Cela vient du fait que la stratégie est limitée par l'utilisation d'une seule entrée (prédicteur sans mémoire) [4].

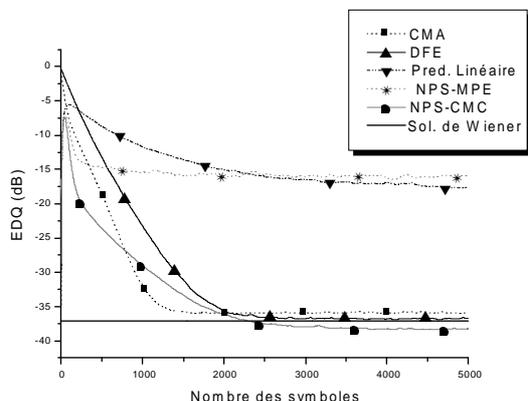


FIG. 4: Évolution de l'EDQ pour le canal \mathbf{f}_1 .

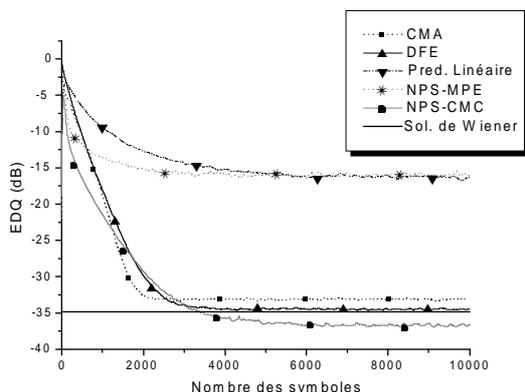


FIG. 5: Évolution de l'EDQ pour le canal \mathbf{f}_2 .

5 Conclusions

Dans ce travail nous avons présenté une structure prédictive basée sur un réseau des neurones avec une stratégie d'apprentissage mixte: une partie auto-organisée

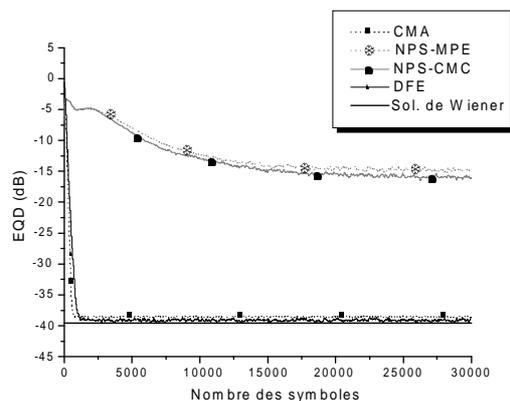


FIG. 6: Évolution de l'EDQ pour le canal \mathbf{f}_3 .

et l'autre supervisée.

Un des deux algorithmes présentés s'est montré plus performant que d'autres algorithmes d'égalisation classiques lorsque le canal est du type minimum de phase. En revanche, pour les canaux NMP, la performance n'est pas assez bonne, ce que nous fait envisager des améliorations qui peuvent être apportées au NPS-CMC, comme par exemple l'augmentation du nombre d'entrées (mémoire) du prédicteur.

D'autre part, l'algorithme SOFVA est encore un sujet d'étude assez vif. En effet, nous sommes maintenant en train d'étudier les performances de nouvelles stratégies pour accélérer sa convergence et aussi pour avoir une « régularisation » de la fonction de coût, ce qui doit réduire la complexité de calcul due au nombre de neurones.

Finalement, étant donné que l'algorithme SOFVA est également capable d'estimer des vallées de fonctions même quand les noyaux qui composent ces fonctions ne sont pas forcément gaussiens, il semble que le NPS-CMC est aussi une stratégie valable dans le cas de canaux non linéaires, ce qui sera étudié par la suite.

Références

- [1] Macchi O. *Adaptive Processing: The Least Squares with Applications in Transmission*. John Wiley & Sons, 1995.
- [2] Montalvão Filho, J.R., "Égalisation et identification de canaux de communication numérique: une approche par reconnaissance des formes et mélange de Gaussiennes", Thèse de doctorat, Université Paris XI, France, Novembre, 2000.
- [3] Cavalcante, C.C., J.R. Montalvão Filho, B. Dorizzi and J.C.M. Mota, "A Neural Predictor for Blind Equalization of Digital Communication Systems", in Proc. of AS-SPCC, Lake Louise, Canada, Octobre 2000.
- [4] Cavalcante, C.C., "Prédiction neural et estimation de la fonction de densité de probabilité utilisées à l'égalisation aveugle", *Master Thesis* (en Portugais), Université Fédérale du Ceara (UFC), Brésil, Février, 2001.
- [5] Haykin, S., *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice Hall, 1998.
- [6] Barbosa, L.J.M., J.C.M. Mota and F.R.P. Cavalcanti, "A Family of Predictive Constant Modulus Algorithms for Blind Equalization", in in Proc. of AS-SPCC, Lake Louise, Canada, Octobre 2000.
- [7] Haykin, S., *Adaptive Filter Theory*, 3rd edition, Prentice-Hall, 1996.