

Algorithme Heuristique de Partitionnement Soustractif pour Reconnaissance Aveugle de Schéma de Modulation

Lionel DE GUENIN¹ Patrick ROSSON¹ Nicolas PETROCHILOS² Eric MOREAU³

¹CEA-Leti, 17 Av. des Martyrs, 38000 Grenoble, France

²CRESTIC Université de Reims, Moulin de la Housse, Chem. des Rouliers, 5100 Reims, France

³CNRS, LIS, Université de Toulon – Campus de La Garde – Bat X, CS 60584, 83041 Toulon CEDEX 9, France

Résumé – Ce papier présente un algorithme de reconnaissance aveugle d’ordre de modulation pour signaux à modulation d’amplitude en quadrature (QAM), fréquemment utilisés pour les communications sans-fil OFDM. La méthode proposée se base sur une méthode de reconnaissance de constellation, utilisant un algorithme de partitionnement soustractif. Une analyse en composante principale permet une réduction de bruit et une réduction de longueur du signal lorsqu’il est sur-échantillonné, ainsi qu’une estimation du bruit résiduel, cette information est utilisée pour effectuer un partitionnement soustractif. Enfin une classification en deux étapes exploite le nombre de groupes estimés ainsi que leur position dans le plan complexe pour inférer le schéma de modulation. Les résultats de simulation montrent que l’algorithme proposé permet de meilleures performances que l’état de l’art, en utilisant moins de paramètres libres.

Abstract – This paper presents a blind modulation recognition method for quadrature amplitude modulation (QAM) signals, which are broadly used in OFDM wireless communication systems. The proposed method takes the constellation based recognition approach, using a subtractive clustering algorithm. The principal component analysis allows noise-reduction and signal length reduction in over-sampled use cases, as it estimates leftover noise power, that information is used to perform said subtractive clustering. Subsequently a simple two-step classification exploits both the number of estimated cluster centers and their position in the complex plane to infer modulation scheme. Simulation results shows faster computing and better performance than similar methods, while using less free parameters.

1 Introduction

La reconnaissance de schéma de modulation est un domaine important de l’analyse des signaux télécoms. Elle est utilisée dans une variété de domaine, militaire et civil, tel que la surveillance de spectre, les contre-mesures électroniques, la radio logicielle, le radar passif etc. C’est une étape nécessaire pour la reconstruction aveugle de signal, dans le but de mettre en oeuvre un radar passif basé sur la forme d’onde LTE. La modulation d’amplitude en quadrature (QAM) est abondamment utilisée dans les réseaux cellulaires sous forme d’onde multi-porteuses, incluant le LTE mais aussi la 5G. D’où la nécessité de la reconnaissance aveugle d’ordre de modulation des signaux QAM.

Les algorithmes de partitionnement ont été démontrés comme particulièrement rapides et performants pour cette reconnaissance aveugle, puisqu’ils sont capables de distinguer un large nombre de schémas de modulation, généralement en s’appuyant sur une classification exploitant le nombre de centre de groupe estimé. Cette classification est pertinente tant que les schémas de modulation ont un nombre suffisamment différent de symboles possibles, ce qui est le cas pour les signaux QAM. Toutefois, des classifications plus complexes existent sans ces limitations [1].

Ce papier présente un algorithme de partitionnement pour la reconnaissance aveugle de signaux QAM encore plus rapide et performant. Il est organisé comme suit : un état de l’art est établi en Section 2, l’algorithme de reconnaissance aveugle basé sur une analyse en composante principale (PCA), le parti-

tionnement soustractif et une classification en deux étapes est discutés en Section 3, les résultats de simulation sont présentés en section 4, puis nous concluons en Section 5.

2 État de l’art

Des méthodes pertinentes de reconnaissance de schéma de modulation ont été présentées au cours des années passées, elles diffèrent par leur mode opératoire, la nature, le nombre de schéma de codage à distinguer et le nombre d’échantillons par symbole (ce qui sera référé comme le taux de sur-échantillonnage dans le cadre de cet article). Par exemple, dans [1] la correspondance d’images basée sur modèle est utilisée pour efficacement reconnaître les codages 16QAM, 32QAM et 64QAM, incluant leurs variantes de différentes formes. Pourtant, même avec un taux de sur-échantillonnage de 8, leur reconnaissance de la 64QAM ne fonctionne pas sous les 15 dB de Ratio Signal sur Bruit (SNR). La sensibilité des performances à l’ordre de modulation rend cette méthode inadaptée pour des ordres de modulation supérieurs.

Parmi les méthodes de reconnaissance aveugle, plusieurs sont basées sur des algorithmes de partitionnement pour inférer la constellation utilisée. Dans [2], on utilise le partitionnement hiérarchique en conjonction avec des algorithmes génétiques dans un cas d’usage canal plat avec Bruit Additif Gaussien (AWGN); ces derniers augmentent la complexité informatique en requérant plusieurs itérations pour converger. Dans [3] le partitionnement K-mean sert à grouper les points de la constellation, cette méthode nécessite plus d’itérations lorsque

l'ordre de modulation augmente, en conséquence elle gagne en complexité et perd en performance avec des ordres de QAM plus grands. Une autre solution populaire, [4], est d'appliquer un algorithme de partitionnement soustractif sur l'amplitude des signaux avec un taux de sur-échantillonnage de 8, sous hypothèse canal plat AWGN. A SNR faible, le taux de reconnaissance décline rapidement, de plus leur classification repose uniquement sur le nombre estimé de groupes.

Enfin, [5] propose le partitionnement soustractif dans le plan complexe et une classification basée sur le nombre de groupes pour reconnaître des QAM d'ordre important à faible SNR, et montrent que le partitionnement soustractif est plus rapide et obtient de meilleurs résultats que le partitionnement K-moyennes. Néanmoins, leur approche mène à un algorithme à 3 paramètres, cela rend des performances robustes difficiles à atteindre et à reproduire. Leur taux de sur-échantillonnage est de 20, avec à nouveau l'hypothèse canal plat AWGN.

3 Algorithme de reconnaissance de schéma de modulation

Notre méthode présente une amélioration à l'algorithme de [5] en considérant l'information supplémentaire donnée par le rayon des groupes estimés. Les groupes sont des amas de points dans l'espace I/Q, chaque point étant une réalisation de la mesure du signal reçu, sous hypothèse de synchronisation fréquentielle ils sont a priori situés à proximité de leur valeur attendue : un des symboles de la constellation utilisée. Un prétraitement de réduction de bruit permet d'estimer le bruit résiduel pour évaluer le rayon des groupes.

Premièrement, une PCA est appliquée au signal, la première composante est assignée comme symboles dé-bruités, tandis que les autres révèlent le niveau de bruit. Deuxièmement, on applique l'algorithme de partitionnement aux symboles dé-bruités. Troisièmement, une classification bidimensionnelle est effectuée à partir des estimations de la position et du nombre de groupes. Avant de détailler le processus de reconnaissance, le modèle, les notations et les hypothèses vont être introduits dans la sous-section suivante.

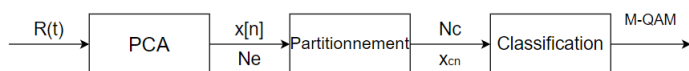


FIGURE 1 : Schéma bloc de l'algorithme

3.1 Modèle

Le signal reçu en bande de base d'un canal AWGN au récepteur s'exprime comme :

$$R(t) = S(t) + N(t) \quad (1)$$

Où S est le symbole M-QAM émis, et N le bruit additif. Bien que le catalogue de modulation étudié dans cet article se limite à : BSPK, QPSK, 4QAM, 16QAM, 32QAM, 64QAM, 128QAM et 256QAM, la méthode présentée peut s'étendre à d'autres ordres de modulation. L'hypothèse canal AWGN est la plus raisonnable dans notre étude, car elle s'applique à des

signaux OFDM égalisé (on peut aussi imaginer une application par sous-porteuse), de plus elle permet de se comparer aux méthodes similaires [5]. On définit notre SNR comme $\frac{std(S)}{std(N)}$. Si on considère le facteur de sur-échantillonnage q , et en assumant une synchronisation symbole, l'instant t est exprimé comme le $i^{\text{ème}}$ échantillon du $k^{\text{ème}}$ symbole :

$$t = (k + i) t \quad (2)$$

$$0 \leq i < q; 1 \leq k < K$$

Avec K le nombre total de symboles et t le temps d'échantillonnage.

3.2 Analyse en Composantes Principales

Les échantillons peuvent donc être réarrangés dans une matrice \mathbf{Y} de taille K par q . Cette matrice est centrée et normalisée, puis on lui applique l'algorithme PCA. Ce dernier minimise la perte d'information pour un bruit indépendant. De plus sa complexité est $O(\max(K, q) \min(K, q)^2)$ [6]. Il peut être interprété comme une variante de la décomposition en valeur singulière :

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (3)$$

Avec \mathbf{S} une matrice diagonale de taille K par q contenant les valeurs singulières de \mathbf{X} , \mathbf{U} et \mathbf{V} deux matrices carrées de vecteurs unitaires orthogonaux de taille respective $K \times K$ et $q \times q$. La matrice \mathbf{P} des composantes principales peut se calculer comme :

$$\mathbf{P} = \mathbf{U}\mathbf{S} \quad (4)$$

On note P_n la $n^{\text{ème}}$ composante principale de \mathbf{X} . La première composante P_1 de longueur K contient le signal dé-bruité sur lequel l'algorithme de partitionnement va travailler, voir Fig. 2. De meilleurs performances seront obtenues grâce à la réduction du niveau de bruit. La réduction de longueur du signal original permettra un traitement plus rapide.

Les propriétés statistiques des autres composantes donnent une estimation de la puissance de bruit rejetée et donc de la puissance de bruit résiduelle. Puisque c'est dans l'intérêt des bonnes performances de l'algorithme de partitionnement de légèrement surestimer le niveau de bruit N_e , ce dernier est exprimé comme un multiple de l'écart type de P_2 :

$$N_e = 1.5 \sqrt{\frac{1}{K} \sum_{i=1}^q (P_{2,i} - \overline{P_2})^2} \quad (5)$$

Avec $P_{2,i}$ le $i^{\text{ème}}$ élément de la seconde composante principale de \mathbf{X} et $\overline{P_2}$ sa moyenne. Ce bruit estimé est le rayon caractéristique de notre algorithme de partitionnement.

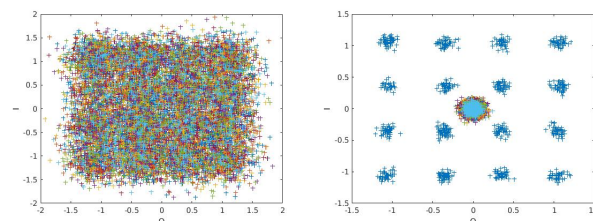


FIGURE 2 : 750 symboles de 16QAM sur-échantillonnés d'un facteur $q=20$ (gauche) à $SNR=3$ dB, et ses composantes principales (droite).

La Fig. 2 représente à gauche le signal reçu et à droite ses composantes principales. On observe la première composante qui met en évidence le partitionnement des points d'une 16QAM. Cette constellation ainsi obtenue sera l'entrée de l'algorithme de partitionnement. Les autres composantes, celles autour de l'origine ont un écart-type indicatif du rayon des groupes.

3.3 Algorithme de partitionnement soustractif

Les algorithmes de partitionnements [3][4][5][7][8] infèrent à partir des données la position et le nombre de symboles possible. L'algorithme estime la position centrale des groupes, points qui maximisent la densité spatiale dans le plan I/Q, sous hypothèse qu'un tel centre de groupe correspondrait à un symbole de la constellation. Ainsi, le nombre de centres de groupe indique intrinsèquement l'ordre de modulation. L'algorithme de partitionnement est itératif : chaque centre est estimé l'un après l'autre. Pour un jeu de données X de K points, la densité initiale D_0 à chaque point i est calculé de la manière suivante :

$$D_{0;i} = \prod_{j=1}^K \exp \left(-\frac{\|x_i - x_j\|^2}{r^2} \right) \quad (6)$$

Avec le paramètre r , rayon caractéristique du traitement. Les couples (x_i, x_j) tel que $\|x_i - x_j\| < r$ contribuent peu à la densité et inversement. Le rayon caractéristique r définit le rayon des groupes formés. Dans la mesure où la dispersion des points varie avec le bruit, il convient de prendre $r = N_e$.

Le point qui maximise la densité est alors défini comme premier centre de groupe c_1 et tous les points distant du centre de moins de r lui sont assignés. Une nouvelle densité réajustée D_n est ensuite calculée :

$$D_{n;i} = D_{n-1;i} \prod_{j=1}^K \exp \left(-\frac{\|x_i - x_{c_{n-1}}\|^2}{r^2} \right) \quad (7)$$

De manière similaire, un nouveau centre de groupe c_n est défini et de nouveaux points lui sont assignés. Le processus est réitéré jusqu'à assignation totale.

Dans le pire cas : autant de groupes que de points, la complexité calculatoire de l'algorithme de partitionnement est $O(K^2)$. Exécuter l'algorithme de partitionnement sur le signal original (comme dans [5]) a pour complexité : $O(K^2)$. Tandis qu'exécuter PCA puis l'algorithme de partitionnement sur la première composante a pour complexité : $(\max(K; K) \min(K; K)^2) + K^2$. Il s'avère donc que l'approche présentée est plus rapide, ce qui a été confirmé lors des simulations.

3.4 Classification

Cette dernière étape de classification a pour but de déduire le schéma de modulation correspondant au nombre de centres N_c de groupe mesurés et leur positions x_{c_n} donnée par l'étape précédente. Premièrement, une décision est faite à partir de N_c , elle est ensuite ajustée à l'aide d'une nouvelle métrique calculée à partir des positions $x_{c_i}; i \geq 1; n$.

Sans bruit, le nombre estimé de centres pour un signal M-QAM est M , en présence de bruit, ce nombre peut fluctuer

(voir Fig 3.). On construit empiriquement des seuils $Th_{c,n}$ sur les valeurs de N_c pour effectuer une première décision sur l'ordre de modulation (Voir Fig 5. et Table 1)

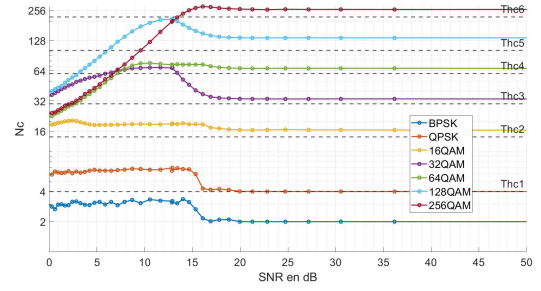


FIGURE 3 : Nombre de centres de groupe estimé pour différents schéma de modulation en fonction du SNR pour un taux de sur-échantillonnage de 20.

La seconde décision est prise à partir de la différence maximale ajustée au bruit entre la constellation estimée et l'ensemble complet de symbole possible de toutes les constellations des schémas de modulation que l'on cherche à distinguer. Cette métrique secondaire d_{max} est calculé comme suit :

Soit $S_i \in E_{M-QAM}$ avec E_{M-QAM} l'ensemble de tous les symboles de toutes les constellations considérées par l'algorithme. $E_{M-QAM} = \{BPSK, QPSK, 4QAM, \dots, 256QAM\}$. Pour chaque centre de groupe c_n , on définit :

$$d_{max;c_n} = \max_i (\|x_{c_n} - S_i\|) \quad (8)$$

Enfin :

$$d_{max} = \max_{c_n} (d_{max;c_n}) \cdot N_e \quad (9)$$

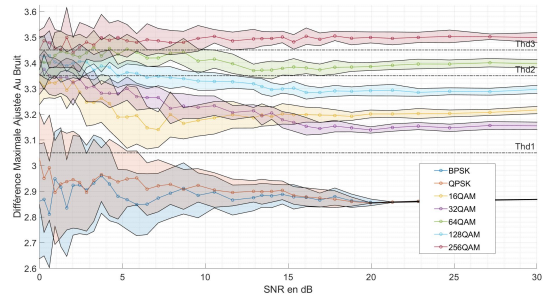


FIGURE 4 : Évolution de la moyenne et de l'écart-type de la métrique secondaire pour différents ordres de modulation, en fonction du SNR.

Comme le montre la Fig. 4, d_{max} augmente pour les constellations de formes non rectangulaire et pour les constellations de plus grande dispersion. Par conséquent, des seuils $Th_{d,n}$ peuvent être construit sur cette métrique pour ajuster la classification et enfin déterminer le schéma de modulation. L'arbre de décision complet est représenté en Fig. 5.

4 Simulations

Le signal composé de symboles sur-échantillonnés est généré en bande de base sur le catalogue de schéma de modula-

