PERFORMANCES DES TURBO-CODES EN BLOCS Q-AIRES "REED-SOLOMON"

Omar AITSAB et Ramesh PYNDIAH

TELECOM BRETAGNE
Technopôle de Brest Iroise BP-832
29285 BREST

RÉSUMÉ

Grâce aux récents progrès dans le domaine du décodage itératif des codes concaténés, plusieurs voies d'exploration ont été ouvertes. Nous présentons dans cet article, une première approche du décodage itératif des codes produits Reed-Solomon (RS) "turbo-codes en blocs RS". Deux méthodes de construction de ces codes sont illustrées, ainsi que le principe du décodage itératif à décodage pondéré et décision pondérée. Les performances de ces codes ont été simulées sur un canal de Gauss à l'aide de simulation de type Monté Carlo. Les gains de codage obtenus sont de l'ordre de 5.5dB pour un TEB(Taux d'erreur Binaire) de 10^{-5}. Ce nouvel algorithme de décodage est très intéressant surtout dans le domaine du stockage des données où les codes produits RS sont souvent utilisés.

ABSTRACT

Thanks to recent progress in the iterative decoding of concatenated codes, several new fields of investigation have appeared. In this paper, we present a first approach of the iterative decoding of Reed-Solomon (RS) product codes: "Turbo-codes RS". Two methods to construct RS product codes are given. The iterative decoding of the RS product codes is based on the soft decoding and the soft decision of the component codes. The performance of RS turbo-codes have been evaluated on the gaussian channel using Monte Carlo simulation. Coding gains up to 5.5dB for a BER (Bit error rate) of 10^{-5} have been obtained. This new decoding algorithm is very interesting for storing data where the RS product codes are often used.

1. INTRODUCTION

L'utilisation des Codes Correcteurs d'Erreurs (CCE) est actuellement en pleine explosion, notamment dans les domaines de la télédiffusion et des radiocommunications mobiles. Cet intérêt pour les CCE résulte principalement de la réduction du débit au niveau du codage source qui rend les données trés sensibles aux erreurs. De plus, l'augmentation de la densité d'intégration des circuits VLSI permet d'implémenter des algorithmes de plus en plus complexes. D'où l'utilisation de code de plus en plus performant en termes de rendement et de gain de codage.

Les codes produits sont des codes puissants qui présentent de très grandes distances minimales pour un rendement donné. Malgré ces bonnes caractéristiques, il faut être prudent lors du décodage afin de bénéficier de toute la puissance de ces codes.

En 1993, C. Berrou [1] a obtenu des performances exceptionnelles avec les "turbo-codes" convolutifs. Ces turbo-codes sont construits à partir de deux codes convolutifs récursifs concaténés. C. Berrou a utilisé un algorithme de décodage itératif à décodage pondéré et à décision pondérée [2]. Suite aux excellents résultats des turbo-codes convolutifs, R. Pyndiah [3] a montré qu'il était possible d'obtenir d'aussi bons résultats avec des turbo-codes en blocs. Ces turbo-codes en blocs sont construits suivant le principe des codes produits [4] à partir de codes binaires du type BCH. Les turbo-codes en blocs sont décodés par un procédé itératif à décodage pondéré et à décision pondérée [3] également. De plus, ces travaux ont montré que les turbo-codes en blocs avaient de meilleurs performances pour les applications nécessitant des rendements de codage élevés ($R \geq 0.8$).

Les codes de Reed-Solomon sont des codes en blocs à éléments non binaires (q-aires). Ils offrent le meilleur rendement pour une distance de Hamming minimale donnée. Il était donc intéressant d'évaluer les performances d'un turbo-code en blocs à base de codes RS.

Dans la deuxième partie de cet article, nous présenterons deux méthodes utilisées pour la construction des codes produits RS. La troisième partie est consacrée au décodage pondéré des codes RS. La dernière partie concerne la pondération des décisions et le décodage itératif des turbo-code en blocs RS.

2. CONSTRUCTION DU CODE PRODUIT

2.1. Codes de Reed-Solomon

Les codes RS sont des codes BCH à éléments non binaires appartenant au corps de galois $GF\left(q = 2^m\right)$. Chaque symbole q-aires du corps peut être représenté par m éléments binaires. Les principaux paramètres d'un code RS sont (n, k, δ) où n est la longueur des mots de codes, k la longueur des messages d'informations et δ sa distance minimale de Hamming.

2.2. Codes produits

Les codes produits sont construits à partir de deux ou plusieurs codes en blocs élémentaires, généralement linéaires. Considérons deux codes en blocs élémentaires C^1 et C^2 ayant respectivement comme paramètres (n_1, k_1, δ_1) et (n_2, k_2, δ_2).

Le code produit se présente sous forme d'une matrice $C = C^1 \otimes C^2$ à n_1 lignes et n_2 colonnes où :

- les symboles d'information sont rangés dans une sous matrice M à k_1 lignes et k_2 colonnes,
- chacune des k_1 lignes est codée par le code C^2,
- chacune des n_2 colonnes est ensuite codée par le code C^1.

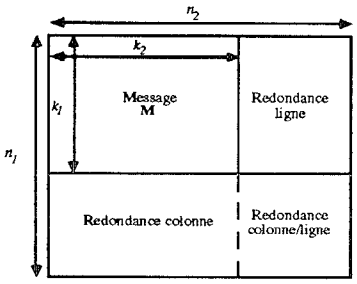

fig 1: Principe de codage d'un code produit

On démontre [4] que toutes les lignes sont des mots de code de C^1 et toutes les colonnes sont des mots de codes de C^2. Les paramètres (n, k, δ) du code produit C s'expriment en fonction de ceux des codes C^1 et C^2 par :

$$n = n_1 \times n_2,\ k = k_1 \times k_2 \text{ et } \delta = \delta_1 \times \delta_2.$$

Le rendement R du code produit est alors égal à : $R = R_1 \times R_2$ où R_i est le rendement du code C^i.

Cette méthode classique a été utilisée pour la construction des codes produits RS où la sous-matrice M contient $k_1 \times k_2$ symboles q-aires d'informations.

Une seconde méthode a été étudiée pour construire un code produit RS à partir d'une matrice d'éléments binaires. La matrice M contient $k_1 * m$ lignes et $k_2 * m$ colonnes d'éléments binaires. Chacune des $k_1 * m$ lignes est convertie en symboles q-aires puis codée par le code C^2. Les mots de codes obtenus sont reconvertis en binaires et chacune des $n_2 * m$ colonnes est convertie en symboles q-aires puis codée par le code C^1. Avec cette méthode chaque bit est codé deux fois mais pas avec les mêmes bits voisins.

Les $(n_1 - k_1) * m$ dernières lignes sont des mots de codes de C^2 de même que les $(n_2 - k_2) * m$ dernières colonnes sont des mots de codes de C^1.

Nous présentons dans cet article uniquement les résultats des codes produits RS construits par la méthode classique qui ont donné les meilleurs résultats.

3. DÉCODAGE PONDÉRÉ DES CODES RS

Dans le cas d'une transmission par modulation de phase (MDP-2, MDP-4) sur un canal à bruit blanc gaussien additif, l'entrée du décodeur est égale à :

$$\mathbf{R} = \mathbf{C} + \mathbf{B} \qquad (1)$$

où $\mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{1j} & \cdots & r_{1n} \\ \vdots & \vdots & r_{ij} & \vdots & \vdots \\ r_{mn} & \cdots & r_{mj} & \cdots & r_{mn} \end{pmatrix}$ est le mot reçu,

$\mathbf{C} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n} \\ \vdots & \vdots & c_{ij} & \vdots & \vdots \\ c_{mn} & \cdots & c_{mj} & \cdots & c_{mn} \end{pmatrix}$ $(c_{ij} = \pm 1)$ le mot émis

et $\mathbf{B} = \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & \vdots & b_{ij} & \vdots & \vdots \\ b_{mn} & \cdots & b_{mj} & \cdots & b_{mn} \end{pmatrix}$ la matrice de bruit blanc dont les composantes b_{ij} sont non corrélés, de moyenne nulle et de même variance σ^2.

Le décodage du mot $\mathbf{C}$ selon le critère du Maximum de Vraisemblance a Posteriori (MVP) conduit à prendre le décision :

$$\mathbf{D} = \mathbf{C}^i \text{ si } \Pr\left(\mathbf{C}^i / \mathbf{R}\right) > \Pr\left(\mathbf{C}^l / \mathbf{R}\right) \quad \forall l \neq i \qquad (2)$$

qui se traduit sur un canal à bruit blanc gaussien additif par :

$$\mathbf{D} = \mathbf{C}^i \text{ si } \left|\mathbf{R} - \mathbf{C}^i\right|^2 < \left|\mathbf{R} - \mathbf{C}^l\right|^2 \quad \forall l \neq i \qquad (3)$$

avec : $$\left|\mathbf{R} - \mathbf{C}^i\right|^2 = \sum_{j=1}^{n} \sum_{f=1}^{m} \left(r_{jf} - c_{jf}^i\right)^2 \qquad (4)$$

$\mathbf{C}^i$ est le $i$ème mot du code et $\mathbf{D} = \begin{pmatrix} d_{11} & \cdots & d_{1j} & \cdots & d_{1n} \\ \vdots & \vdots & d_{ij} & \vdots & \vdots \\ d_{mn} & \cdots & d_{mj} & \cdots & d_{mn} \end{pmatrix}$ la décision suivant le critère du MVP.

Pour les codes RS, le nombre q^k de mots de code est très grand et le décodage suivant le MVP est impossible à implémenter. En 1972, Chase [6] a proposé un algorithme sous optimal pour le décodage pondéré des codes en blocs, qui minimise le nombre de mots testés tout en gardant de bonnes performances. Il suffit de tester uniquement le sous-ensemble des mots de code les plus probables, en se basant sur les informations reçues $\mathbf{R}$, pour appliquer la règle de décision (3). Pour cela on prend une décision ferme sur $\mathbf{R}$ qui donne un mot $\mathbf{Y^0}$ à élément binaire (± 1). En modifiant le signe des éléments de $\mathbf{Y^0}$ qui correspondent aux p composantes les moins fiable de $\mathbf{R}$, on obtient $2^p - 1$ nouveaux vecteurs Y^l, $l = 1 \ldots \left(2^p - 1\right)$.

Le décodage algébrique des 2^p vecteurs Y^l donne le sous-ensemble de mots $\mathbf{C}^l$ à utiliser pour le décodage pondérée (3).

Résultats de simulation:

La borne supérieure du gain asymptotique (G_a) pour un décodage pondéré dépend de la distance minimale du code : $G_a \leq 10 \log(R.\delta)$. Nous avons opté pour l'utilisation des codes étendus puisqu'ils augmentent la distance minimale de 1 pour une légère diminution du rendement.

Nous avons simulé le code RS(16,11,6) sur un canal de Gauss décodé suivant l'algorithme de Chase. La *fig 2.* montre les courbes des TEB obtenus pour différentes séquences test :

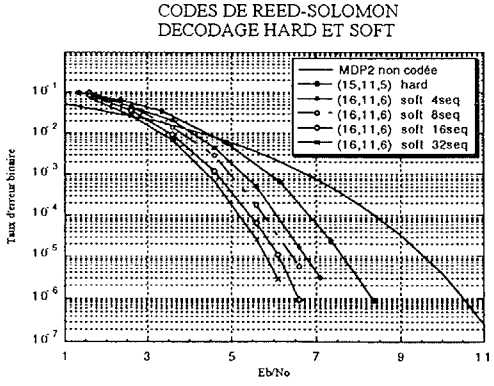

fig 2 : Décodage pondéré du code RS (16,11,6) avec différentes séquences test sur un canal de Gauss.

On remarque que plus on augmente le nombre de séquences test, plus le sous-ensemble de mots décodés à de chance de contenir le bon mot de code émis et plus le gain de codage est élevé. Le gain par rapport au décodage ferme atteint pour un (TEB) de 10^{-5} 1.85dB. Pour le code RS(16,13,4) ce gain est de l'ordre de 2.1dB avec 16 séquences test.

Sur la *fig 3.*, nous comparons les performances du RS(255,223,33)à décodage hard avec le code RS(32,27,6) à décodage soft.

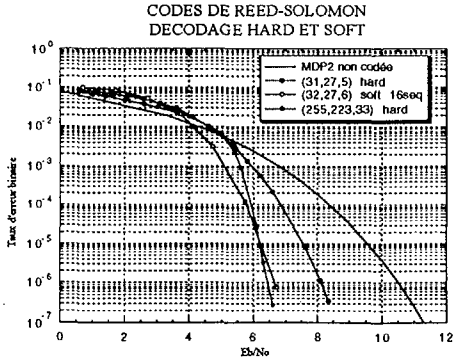

fig 3. : Comparaison du code RS(255,223,33) à décodage ferme avec le code RS(32,27,6) à décodage pondéré.

Les deux codes ont sensiblement le même rendement ($R \approx 0.85$) et donnent des résultats assez proches aux alentours d'un TEB de 10^{-5}. Notons que le pouvoir de correction du code RS(255,223,33) est t=16 tandis que celui du code RS(32,27,6) est t=2. Sachant que le nombre d'opérations effectuées par un décodeur algébrique est une puissance de t, le second code sera plus simple à implémenter, même si on utilise 16 séquences test pour le décodage pondéré. Donc cette comparaison est intéressante puisqu'elle montre qu'un petit code RS avec un décodage pondéré donne d'aussi bons résultats qu'un code, plus puissant avec un décodage ferme.

4. ALGORITHME DE DÉCODAGE DES CODES PRODUITS RS

4.1. Pondération des décisions

Nous avons utilisé l'algorithme "turbo" [3] pour la pondération des décisions. En sortie du décodeur, la pondération de la décision d_{jf} associée à chaque élément binaire c_{jf} du mot décodé $(1 < j < n \text{ et } 1 < f < m)$, est donnée par le Logarithme du Rapport de Vraisemblance (LRV) :

$$LRV_{jf} = \ln \frac{\Pr\{c_{jf} = 1 / R\}}{\Pr\{c_{jf} = -1 / R\}} \qquad (5)$$

On démontre [6], dans le cas d'un canal à bruit blanc additif gaussien, que le LRV_{jf} peut s'écrire sous la forme :

$$LRV_{jf} \approx \frac{1}{2\sigma^2}\left\{\left|R - C^{\min(-1)}\right|^2 - \left|R - C^{\min(+1)}\right|^2\right\} \qquad (6)$$

où $C^{\min(+1)}$ et $C^{\min(-1)}$ sont deux mots de code à distance minimale du l'observation $\mathbf{R}$ et ayant respectivement (+1) et (-1) comme élément binaire à la position (j,f).

En développant (6) à l'aide de (4) le LRV_{jf} devient :

$$LRV_{jf} \approx \frac{2}{\sigma^2}\left(r_{jf} + \sum_{x=1}^{n}\sum_{z=1}^{m} r_{xz} c_{xz}^{\min(+1)} \rho_{xz}\right) \qquad (7)$$

avec $((x,z) \neq (j,f))$, $\rho_{xz} = \begin{cases} 0 \text{ si } c_{xz}^{\min(+1)} = c_{xz}^{\min(-1)} \\ 1 \text{ si } c_{xz}^{\min(+1)} \neq c_{xz}^{\min(-1)} \end{cases}$.

En normalisant le LRV par rapport à $2/\sigma^2$, on peut écrire:

$$r'_{jf} = \frac{\sigma^2}{2} LRV_{jf} = r_{jf} + w_{jf} \qquad (8)$$

avec : $w_{jf} = \sum_{x=1}^{n}\sum_{z=1}^{m} r_{xz} c_{xz}^{\min(+1)} \rho_{xz} \quad ((x,z) \neq (j,f))$ (9)

Le LRV_{jf} normalisé (8) est égal à la somme de l'échantillon r_{jf} présent à l'entrée du décodeur et d'une quantité w_{jf}, indépendante de r_{jf} appelée information extrinsèque.

Pour déterminer l'expression simplifiée du LRV_{jf} d'un élément binaire en sortie du décodeur, il faut identifier les deux mots du code à la distance minimale de $\mathbf{R}$ et ayant des éléments de signe opposé en position (j,f). On utilise pour cela le sous-ensemble de mots de code obtenus par l'algorithme de Chase pour trouver les deux mots recherchés.

Soient $C^{\min(i)}$ et $C^{\min(-i)}$ ces deux mots avec $C^{\min(i)}$ plus proche de $\mathbf{R}$ que $C^{\min(-i)}$ alors d'après (6) le LRV_{jf} normalisé peut s'écrire comme :

$$r'_{jf} = \left(\frac{\left(M^{\min(-i)} - M^{\min(i)}\right)}{4}\right) c_{jf}^{\min(i)} \qquad (10)$$

où $M^{\min(-i)}$ et $M^{\min(i)}$ représentent respectivement le carré de la distance euclidienne entre $\mathbf{R}$ et $C^{\min(-i)}$ et $\mathbf{R}$ et $C^{\min(i)}$.

Il est très fréquent que l'on ne puisse pas trouver dans le sous-ensemble de mots, déterminé par l'algorithme de Chase, deux mots à distance minimale de $\mathbf{R}$ ayant des éléments de signe différent en position (j,f). Dans ce cas le LRV_{jf} normalisé est simplement pris égal à : $r'_{jf} = \beta . c_{jf}^{\min(i)}$ (11)

où β est une constante positive dont la valeur est fixée à priori.

4.2. Décodage itératif des codes produits RS

Pour un code produit RS, la matrice des éléments reçues $[\mathbf{R}]$ possède $n_1 \times m$ lignes et n_2 colonnes. Cette matrice sera décomposée en sous-matrices représentant chacune un mot de code RS. Le premier décodage des lignes permet de déterminer une matrice $[\mathbf{R'}]$ en utilisant l'algorithme décrit ci-dessus. L'information extrinsèque $[W_0] = [\mathbf{R'}] - [\mathbf{R}]$ est extraite et le décodage des colonnes s'effectue avec la nouvelle matrice :

$$[\mathbf{R_1}] = [\mathbf{R}] + \alpha_1 [W_0] \qquad (12)$$

où α_1 est une constante utilisée pour réduire la contribution de l'information extrinsèque dont la fiabilité est généralement faible durant les premières itérations. Après le décodage des colonnes, on obtient une nouvelle information extrinsèque que l'on reinjectera lors du prochain décodage ligne.

En généralisant cette méthode à l'itération p, on effectue le décodage suivant les lignes (ou les colonnes) à partir d'une matrice $[\mathbf{R}_p]$ de la forme :

$$\left[\mathbb{R}_p\right] = [\mathbb{R}] + \alpha_p \left[W_{p-1}\right] \text{ avec } \left[\mathbb{R}_0\right] = [\mathbb{R}] \qquad (13)$$

où $\left[W_{p-1}\right]$ est l'information extrinsèque calculée à l'itération $(p\text{-}1)$.

Lorsqu'il n'est pas possible de trouver les deux mots $C^{\min(-i)}$ et $C^{\min(i)}$, les éléments correspondants de la matrice $\left[W_p\right]$ sont déterminés à partir de la relation (11), où β est indicée β_p en fonction de l'itération p considérée. Vue que la fiabilité du décodage croit au fil des itérations les coefficient α_p et β_p augmentent avec p.

4.3. Résultats des Simulations

Nous avons utilisé des codes RS expurgés comme codes élémentaires car ils augmentent la distance minimale et nécessitent moins de séquences test. Les performances des turbo-codes RS ont été évaluées sur un canal de Gauss par des simulations de type Monté Carlo. Dans la *fig 4.*, nous présentons les TEB du turbo-code RS (15,12,4)*(15,12,4) aprés chaque itération:

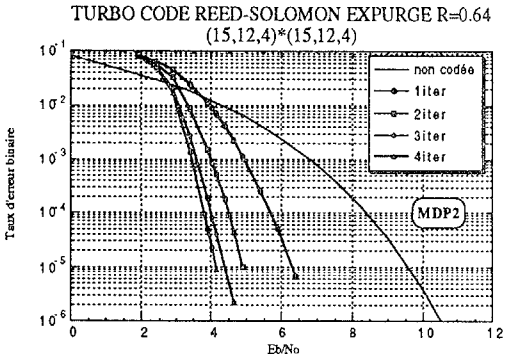

fig 4. : Décodage itératif du code produit RS (15,12,4)*(15,12,4) par l'algorithme "turbo".

On constate que le gain de codage s'améliore au fil des itérations. Pour un TEB de 10^{-5}, on gagne respectivement 1.4dB, 0.5dB et 0.25dB à chaque itération supplémentaire. Le gain global est d'environ 5.5dB par rapport à une transmission non codée, pour un TEB de 10^{-5}. Le phénomène turbo est bien mis en évidence. Cependant le fait d'augmenter les itérations au delà de 4 n'apporte plus de gain significatif.

La *fig 5.* représente le TEB du code RS(31,28,4)*(31,28,4) :

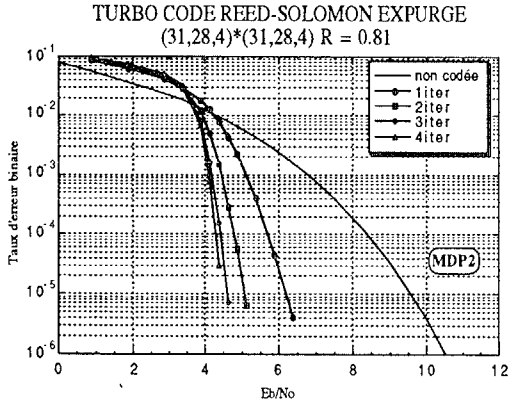

fig 5. : Décodage itératif du code produit RS (31,28,4)*(31,28,4) par l'algorithme "turbo".

Pour ce code, le gain de codage à la 4ème itération est de 5dB (TEB = 10^{-5}), et le rapport signal à bruit pour obtenir un TEB de 10^{-5} est à 3.5dB de la limite de Shannon pour cette transmission codée.

5. CONCLUSION

Dans cet article nous avons présenté les premiers résultats obtenus avec le nouvel algorithme [3] de décodage itératif appliqué aux codes produits RS. Les simulations effectuées ont montré de très bons résultats, meilleurs que ceux obtenus jusqu'à présent avec d'autres algorithmes. Nous avons obtenu des gains de codage allant jusqu'à 5.5dB après 4 itérations et ceci grâce à un décodage itératif basé sur :

1. un décodage *pondéré* de chaque ligne (colonne),
2. pour chaque élément binaire décodé, l'information extrinsèque est extraite à partir du LRV estimé,
3. cette information extrinsèque est pondérée par une constante α lors des premières itérations où le TEB est élevé.

Par contre, l'algorithme de décodage itératif semble moins efficace dans le cas des codes RS que dans le cas des codes BCH[6]. Il faudrait probablement tenir compte dans l'algorithme, du fait que le code RS est un code q-aires pour le rendre plus efficace. Néanmoins les codes de RS sont beaucoup plus utilisés que les codes BCH, surtout dans le domaine des transmissions par satellite et du stockage des données, ce qui offrent à ces codes de très bonnes perspectives dans l'avenir.

REMERCIMENTS

Ce travail a été financé par le Centre National d'Etudes des Télécommunications (CNET) et le Centre Commun d'Etudes de Télédiffusion et Télécommunications (CCETT). Les auteurs tiennent à remercier P. Combelles pour son aide précieuse.

REFERENCES

[1] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and décoding : Turbo-codes (1)", IEEE Int. Conf. on Comm. ICC'93, vol 2/3, May 93, pp. 1064-1071.

[2] C.Berrou, P. Adde, E. Angui and S. Faudeil, " A low complexity Soft-output Viterbi decoder", IEEE Int. Conf. on Comm. ICC'93, vol 2/3, May 93, pp.737-740.

[3] R. Pyndiah, A. Glavieux, A. Picart and S. Jacq,"Near optimum decoding of product codes", IEEE Globecom'94, vol 1/3, pp 339-343.

[4] F.J. Macwilliams and N.J.A. Slone, "The theory of error correcting codes", North-Holland publishing company, 1978, pp. 567-580.

[5] D. Chase, "A class of algorithms for decoding block codes with measurement information", IEEE Trans. Inform. Theory, vol IT-18, Jan. 72, pp. 170-182.

[6] S. Jacq, R. Pyndiah, A. Picart "Algorithme Turbo : un nouveau procédé de décodage pour les codes produits"sera publié au GRETSI 95