

**SUR LES SCHEMAS DE CODES CONCATENES
AVEC DECODAGES PONDERES**

Rémi SFEZ et Gérard BATTAIL

Ecole Nationale Supérieure des Télécommunications,
46, rue Barrault 75634 Paris CEDEX 13.

RÉSUMÉ

La concaténation de deux codes correcteurs est d'usage courant : un code intérieur, souvent décodé avec pondération à l'entrée, est concaténé à un code extérieur, décodé sans pondération. Mais si l'on pondère aussi la sortie du décodeur intérieur, le taux d'erreur résiduel est amélioré. Pour l'alphabet binaire et un canal gaussien sans mémoire, on propose une méthode de prévision des performances du schéma concaténé complet. Sont mis en évidence le rôle du codage intérieur et le gain imputable à la pondération du résultat du décodage intérieur.

SUMMARY

Concatenation of two error-correcting codes is current practice. An inner code, often decoded using a weighted input, is concatenated with an outer code decoded without such a weighting. If, however, the output of the inner decoder itself is weighted, the error rate is improved. When the binary alphabet is used over a memoryless Gaussian channel, we propose a means for predicting the performance of the whole concatenated scheme. The part of the inner coding and the benefit resulting from its weighted-output decoding are emphasized.

1. Introduction

Les domaines d'application du procédé de concaténation de codes correcteurs d'erreurs sont multiples, l'un des plus remarquables étant la communication avec les engins spatiaux : un code intérieur, en général décodé avec pondération à l'entrée, est concaténé à un code extérieur, décodé sans pondération. Selon Forney [1], le premier est destiné à éliminer la plupart des erreurs mais celles qu'il laisse subsister sont groupées en paquets (on utilise généralement pour cela un code en blocs court ou bien convolutif de faible longueur de contrainte) ; quant au second, il est spécialisé dans la correction des paquets d'erreurs ayant la taille attendue après le décodage intérieur (il s'agit le plus souvent d'un code de Reed-Solomon). Mais il est possible de pondérer la sortie du décodeur intérieur. La pondération renseigne alors sur la fiabilité des décisions qui y sont prises ; utilisée par le décodeur extérieur, elle permet de réduire le taux d'erreur résiduel par rapport au cas non pondéré.

Dans ce qui suit, nous envisagerons la concaténation de deux codes correcteurs, avec décodage intérieur pondéré en sortie. Nous considérerons un canal gaussien sans mémoire avec alphabet d'entrée binaire. La pondération

portera donc sur des symboles binaires.

2. Décodage pondéré en sortie, optimal pour les symboles

Un symbole binaire reçu et sa pondération peuvent être représentés conjointement par un nombre réel unique : u désignant le symbole émis, on déduit du démodulateur sa valeur relative *a priori* (avant décodage), notée a et définie ainsi :

$$a = \ln \frac{\Pr \{u = 0\}}{\Pr \{u = 1\}} \quad (1)$$

Par son signe, la valeur relative implique une décision ferme (+ pour 0, - pour 1) et, par son module, renseigne sur la fiabilité de cette décision. Pour une modulation binaire antipodale, en désignant par E/N_0 le rapport signal à bruit, il est facile de voir que les valeurs relatives *a priori* issues du démodulateur sont des variables aléatoires gaussiennes indépendantes, de moyenne et variance respectives :

$$E(a) = \pm 4 \frac{E}{N_0},$$



$$\text{Var}(a) = 8 \frac{E}{N_0}, \quad (2)$$

le signe "+" ou "-", dans l'expression de la moyenne $E(a)$ dépendant du symbole effectivement émis. Pour chaque symbole d'information, le décodage pondéré en sortie, symbole à symbole, consiste à en calculer la valeur relative *a posteriori*, définie elle aussi par la relation (1), mais où les probabilités sont calculées conditionnellement aux contraintes du codage et aux valeurs relatives *a priori*.

Nous envisageons maintenant un code C , défini par la liste exhaustive de ses mots (éventuellement de longueur infinie s'il est convolutif). Notant $C_{i,0}$ (resp. $C_{i,1}$) la classe des mots de C dont le symbole d'information u_i est égal à "0" (resp. à "1"), la règle de décodage pondéré en sortie de u_i , optimale, s'écrit [2] :

$$A_i = \ln \frac{\sum_{c_j \in C_{i,0}} \exp(\sum_{i \in I_j} a_i)}{\sum_{c_j \in C_{i,1}} \exp(\sum_{i \in I_j} a_i)}, \quad (3)$$

où I_j désigne l'ensemble des indices repérant les symboles égaux à "0" dans le mot c_j du code.

Dans le cas particulier d'un code linéaire en blocs, on peut récrire la règle (3) ; par exemple, si ce code est systématique, la relation (3) devient :

$$A_i = a_i + \ln \frac{\sum_{c_j \in C_{i,0}} \exp(\sum_{i \in I_j - \{i\}} a_i)}{\sum_{c_j \in C_{i,1}} \exp(\sum_{i \in I_j - \{i\}} a_i)}, \quad (4)$$

qui est l'une des formes de la règle de décodage par répliques [3].

3. Sur la loi des valeurs relatives *a posteriori* optimales

Pour un code linéaire en blocs on montre facilement [2] qu'avec un rapport signal à bruit "très grand", les valeurs relatives *a posteriori* sont gaussiennes. Mais rien ne permet d'extrapoler ce résultat au cas d'un rapport signal à bruit quelconque. En effet, selon la règle (4), chaque valeur relative *a posteriori* apparaît comme la somme de deux termes : le premier est gaussien par hypothèse (il s'agit de la valeur relative *a priori* a_i); quant au second, il se déduit de variables aléatoires gaussiennes par une expression non linéaire qui ne conserve pas, en principe, le caractère gaussien. Par conséquent, les valeurs relatives *a posteriori* n'ont aucune raison d'être gaussiennes dans le cas général.

Certes, d'après la règle (4), la loi d'une valeur relative *a posteriori* s'obtient par la convolution d'une loi, gaussienne par hypothèse, avec une autre loi qui ne l'est pas : il est intuitif que le résultat de cette convolution "se rapproche" du gaussien - en un sens qu'il faudrait préciser -, mais il est difficile de prévoir dans quelle mesure.

La figure 1 représente l'histogramme des valeurs relatives *a posteriori* optimales obtenues après le décodage simulé du code de Golay (23,12), avec un rapport signal à bruit par symbole d'information $E_b/N_0 = 3,8$ dB. La simu-

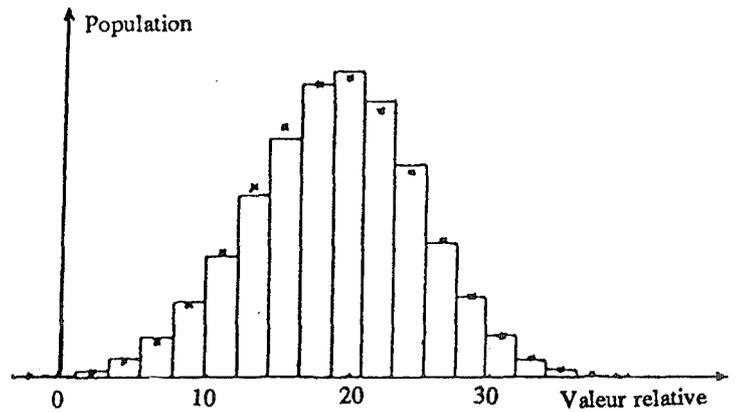


Fig. 1 - Histogramme des valeurs relatives après décodage

lation est effectuée en supposant le mot nul toujours émis (ce qui est légitime, pour un code régulier); 36 000 bits d'information ont ainsi été décodés.

Sur cet histogramme, les moyenne et variance estimées sont $m = 18,85$ et $\sigma^2 = 36,12$. Sont représentés par des croix quelques points de la courbe gaussienne de moyenne m et de variance σ^2 : il semble que le caractère gaussien soit assez bien conservé *a posteriori* ; en outre, on note que le rapport σ^2/m est sensiblement égal à 2, comme c'est le cas des valeurs relatives obtenues après démodulation (cf. relation (2), en supposant que le symbole émis est "0").

Un même histogramme a été tracé pour différents codes et diverses valeurs du rapport signal à bruit : pour des valeurs raisonnables du taux d'erreur résiduel, nous avons toujours constaté que les valeurs relatives étaient approximativement gaussiennes avec un rapport de leur variance à leur moyenne voisin de 2. Un test de gaussianité a confirmé ces observations [2].

L'hypothèse que les valeurs relatives *a posteriori* sont gaussiennes est nécessaire à la prévision des performances du décodage à sortie pondérée. Nous la ferons dans toute la suite ; la comparaison des résultats ainsi obtenus avec ceux de la simulation en confirmera la validité.

4. Performances du décodage pondéré en sortie, dans un schéma concaténé

4.1. Performances du décodage à vraisemblance maximale mot à mot

Nous commençons par rappeler quelques résultats classiques concernant la prévision des performances du décodage à vraisemblance maximale mot à mot. Celle-ci est rendue possible en utilisant des bornes "par réunion".

Pour un code en blocs de longueur n symboles binaires et de distance minimale d_{\min} , la probabilité d'erreur par bit après décodage à vraisemblance maximale mot à mot est majorée par la quantité

$$P_{e.b.} < \frac{1}{2n} \sum_{d=d_{\min}}^n d w_d \operatorname{erfc}(\sqrt{dE/N_0}), \quad (5)$$

où w_d est le nombre des mots du code de poids égal à d .

Pour un code convolutif de taux d'égal à b/n et de distance libre d_f , la même probabilité est majorée selon

$$P_{e.b.} < \frac{1}{2b} \sum_{i=1}^{+\infty} \sum_{d=d_f}^{+\infty} i a(d,i) \operatorname{erfc}(\sqrt{dE/N_0}), \quad (6)$$

où $a(d,i)$ est le nombre de chemins dans le treillis du code de poids égal à d , comportant i symboles d'information "1"; ce nombre se déduit de la fonction génératrice généralisée du code.

Mais les probabilités d'erreur par bit obtenues après décodage à vraisemblance maximale mot à mot d'une part, symbole à symbole d'autre part, sont très voisines quand elles sont petites (et donc quand le codage est utile), de sorte que (5) et (6) constituent aussi des prévisions valables pour le décodage optimal symbole à symbole d'un code en blocs et d'un code convolutif, respectivement.

4.2. Méthode de prévision

Sans perte de généralité, nous nous restreindrons à la concaténation de deux codes, avec entrelacement des symboles binaires émis par le codeur extérieur, comme l'indique la figure 2. Dans le cas (classique) où le décodage extérieur n'est pas pondéré, l'entrelacement, effectué par symboles du code extérieur et donc par blocs de symboles binaires, permet d'éviter un dépassement trop fréquent de la capacité de correction du code extérieur. Dans le cas pondéré, il apparaît qu'il est plus efficace quand on l'effectue symbole par symbole [2]. Son rôle est alors de décorrélérer les valeurs relatives issues du décodeur intérieur, de sorte que le décodeur extérieur puisse les traiter comme mutuellement indépendantes.

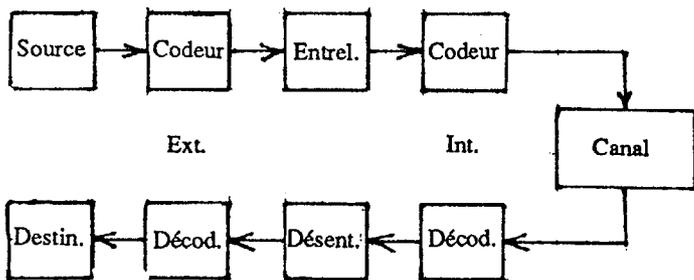


Fig. 2 - Concaténation de deux codes

En reprenant les observations mentionnées au paragraphe 3 et en supposant l'entrelacement parfait, la sortie du désentrelaceur est de même forme que celle d'un démodulateur optimal non quantifié (cf. paragraphe 2 : les valeurs relatives sont sensiblement gaussiennes, indépendantes, avec un rapport entre leur variance et leur moyenne voisin de 2). Le canal obtenu après désentrelacement peut alors être considéré comme gaussien sans mémoire et la prévision des résultats consiste à estimer son rapport signal à bruit.

Nous avons rappelé, au paragraphe précédent, la possibilité de prévoir la probabilité d'erreur par symbole (binaire) obtenue après décodage à maximum de vraisem-

blance mot à mot : nous avons souligné que cette prévision vaut également pour un décodage à vraisemblance maximale symbole à symbole (les deux critères impliquant des probabilités d'erreur par symbole très voisines).

Or, si le décodage est pondéré en sortie et si l'on suppose le symbole "0" systématiquement émis, la probabilité d'erreur par symbole est la probabilité d'obtenir une valeur relative négative (indiquant une décision en faveur du symbole "1"). Dans ces conditions, et en considérant les expressions des moyenne et variance des valeurs relatives après démodulation (cf. relation (2)), il vient, pour un canal gaussien sans mémoire :

$$\frac{E}{N_0} = \left[\operatorname{erfc}^{-1}(2P_{e.b.}) \right]^2. \quad (7)$$

Ainsi, $P_{e.b.}$ désignant ici la probabilité d'erreur par symbole après décodage intérieur (estimée par (5) ou (6), selon le code intérieur choisi), la relation (7) permet d'évaluer le rapport signal à bruit intermédiaire (après désentrelacement). Connaissant celui-ci, on peut calculer la probabilité d'erreur finale, en utilisant à nouveau l'une des relations (5) ou (6).

Nous avons dit que cette méthode repose conjointement sur l'hypothèse de gaussianité des valeurs relatives *a posteriori* optimales et sur la proximité des probabilités d'erreur par symbole après décodage à vraisemblance maximale mot à mot et optimale symbole à symbole. Dans ce qui suit, nous validons cette méthode dans son ensemble, en vérifiant la conformité de ses résultats avec ceux de la simulation.

5. Validation de la méthode de prévision

Nous étudions d'abord la concaténation de deux codes de Hamming, chacun de dimensions (7,4), afin de comparer les performances d'un tel schéma lorsque le décodage intérieur est à sortie pondérée optimale pour les symboles d'une part, à vraisemblance maximale mot à mot d'autre part. Le décodage extérieur est à vraisemblance maximale mot à mot, pondéré à l'entrée dans le premier cas mais sans pondération dans le second. Rien ne justifie en particulier l'emploi du code de Hamming (7,4), si ce n'est sa faible taille, qui le rend aisément décodable. Par ailleurs, nous utilisons un entrelacement parfait.

Nous représentons sur la figure 3 les taux d'erreur par symbole en fonction du rapport signal à bruit par symbole d'information, après un décodage complet du schéma concaténé, avec et sans pondération intermédiaire, obtenus par simulation.

On notera d'abord que le gain apporté par la pondération s'établit rapidement à environ 2 dB : cette valeur est à rapprocher des 2 dB perdus en quantifiant à deux niveaux la sortie d'un canal gaussien sans mémoire (pour une modulation binaire antipodale), par rapport au cas non quantifié : cela confirme l'hypothèse selon laquelle le canal obtenu après désentrelacement peut être considéré comme gaussien sans mémoire, ainsi que la méthode de

prévision que l'on en a déduite.

Par ailleurs, nous avons répété le décodage extérieur avec des valeurs relatives gaussiennes simulées, indépendantes, correspondant au rapport signal à bruit intermédiaire associé selon (7) au taux d'erreur mesuré en sortie du décodage intérieur. Les points représentés sur la figure 3 indiquent les taux d'erreur résiduels obtenus par cette méthode : ils sont assez proches de la courbe (a) pour confirmer la méthode de prévision proposée.

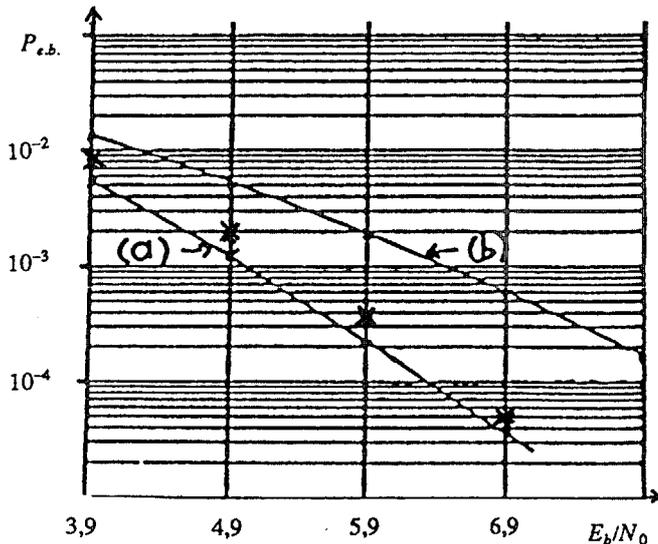


Fig. 3 - Résultats de simulation, concaténation du code de Hamming (7,4) avec lui-même.

Mieux encore, la prévision par (7) a été éprouvée sur le produit itéré de codes de parité [4]. La prévision de la probabilité d'erreur par symbole du décodage à vraisemblance maximale mot à mot de ces codes est possible par une méthode exacte, de sorte que (7) permet de définir un rapport signal à bruit équivalent après décodage. On peut calculer exactement la probabilité d'erreur correspondant à ce rapport signal à bruit puis à nouveau l'inverser par (7) : on obtient ainsi une probabilité d'erreur par bit et un rapport signal à bruit équivalent, pour chaque étape du décodage du produit itéré. Or, les résultats ainsi obtenus sont remarquablement voisins de ceux de la simulation, ce qui montre que l'emploi répété de (7) est robuste vis-à-vis de l'itération du décodage.

6. Conclusion

Pour un canal gaussien sans mémoire d'alphabet d'entrée binaire, nous avons présenté une méthode de prévision des performances d'un schéma concaténé complet où tous les décodages sont pondérés : un décodeur intérieur calcule donc des valeurs relatives qui sont utilisées par le décodeur extérieur qui lui fait suite. Cette pondération permet un gain sur le rapport signal à bruit voisin de 2 dB, pour un même taux d'erreur résiduel.

Il apparaît que les valeurs relatives obtenues par décodage optimal pondéré en sortie sont approximativement gaussiennes. Ce résultat n'est pas seulement établi par les histogrammes ou tests de gaussianité mais, aussi et surtout,

validé par ses conséquences, même quand il est fait un emploi répété de l'hypothèse gaussienne [2].

L'interprétation de Forney [1], qui conduisait à concaténer deux codes, est remplacée dans la perspective d'une pondération en sortie par la remarque que l'ensemble constitué du décodeur intérieur et du désentrelaceur se comporte comme un *amplificateur* du rapport signal à bruit, alimenté par la redondance. Le décodage devenu "transparent" à l'information, il n'y a plus d'inconvénient à l'effectuer successivement de nombreuses fois. Un code complexe peut donc être construit par la combinaison de nombreux codes simples décodés l'un après l'autre. Ainsi, le produit itéré de codes de parité combine de la façon la plus simple les codes les plus simples [4].

D'autre part, si l'on admet que le rapport signal à bruit du canal après décodage et désentrelacement est en relation biunivoque avec les performances du décodage à vraisemblance maximale mot à mot (selon (7)), la qualité d'un code peut être spécifiée par le même critère, qu'il soit décodé avec pondération en sortie ou seulement à l'entrée. Le choix de ce critère demeure un problème ouvert, très fondamental, car la distance de Hamming minimale s'y avère inadéquate [4].

Références bibliographiques

- [1] G.D. FORNEY, Concatenated Codes, M.I.T. Press, 1966
- [2] R. SFEZ, Sur les Systèmes de Codes Concaténés avec Décodages Intermédiaires Pondérés, Thèse de Doctorat, E.N.S.T., mars 1989
- [3] G. BATTAIL, M. DECOUVELAERE, Décodage par Répliques, Annales des Télécommunications, 31, n° 11-12, nov.-déc. 1976
- [4] G. BATTAIL, Construction Explicite de Bons Codes Longs, Annales des Télécommunications, à paraître