

Propriétés et applications de l'information mutuelle entre extrinsèques

Florence ALBERGE

Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506)
Univ. Paris-Sud - CNRS - CentraleSupélec
3, rue Joliot-Curie, 91192 Gif-sur-Yvette
florence.alberge@u-psud.fr

Résumé – Nous présentons dans cette article une nouvelle métrique, l'information mutuelle entre extrinsèques, et nous considérons son application dans le cadre du décodage itératif. L'objectif est de montrer que, des techniques de type EXIT charts, qui permettent de choisir une combinaison efficace de codes peuvent aussi aider à calibrer correctement le récepteur. Nous montrons que l'information mutuelle entre extrinsèques peut être estimée avec précision au niveau du récepteur et être utilisée comme critère d'arrêt ou pour pondérer de manière optimale les quantités propagées dans un décodage itératif.

Abstract – In this manuscript, a new metric called mutual information between extrinsics is presented and applied to iterative decoding. The main objective is to prove that EXIT charts, a powerful method for designing an efficient combination of codes, can also be useful for a proper design of the receiver. It is proved that the mutual information between extrinsics can be estimated with good accuracy at the receiver side. As a consequence, this metric is a good candidate as stopping criterion and also to determine the optimal weighting factor on exchanged quantities.

1 Introduction

Le décodage itératif est une procédure qui consiste à décoder un code complexe à l'aide de décodeurs simples qui échangent de l'information (extrinsèques). L'apparition des turbo-codes en 1993 a permis de montrer que des performances proches de la capacité pouvaient être atteintes grâce à l'utilisation du décodage itératif. Le comportement d'un turbo-récepteur peut être prédit à l'aide d'une technique appelée Exit charts [1] qui permet alors de déterminer une association efficace de codes. Le principe est assez général et peut s'appliquer aux turbo-codes comme aux codes LDPC [2]. Un EXIT Chart est obtenu en traçant, pour chaque décodeur individuel, l'évolution de l'information mutuelle entre le message à décoder X et l'information extrinsèque L disponible en sortie du décodeur et exprimée sous la forme d'un logarithme du rapport de vraisemblance (LLR). Le récepteur n'a aucune connaissance a priori de X . Par conséquent, l'information mutuelle $I(L; X)$ n'est pas une métrique facilement utilisable en réception. Nous considérons ici l'information mutuelle $I(L_y, L_z)$ entre deux extrinsèques obtenues à la sortie de deux décodeurs distincts. Nous montrons le lien entre cette mesure et l'information mutuelle classiquement utilisée dans les EXIT charts. Puisque $I(L_y, L_z)$ ne fait intervenir que des quantités présentes au niveau du décodeur, elle peut être évaluée en réception et fournir des informations utiles pour un réglage adapté. Nous proposons ici un estimateur de $I(L_y, L_z)$ pour des séquences de longueur finie dont nous évaluons la précision. A titre d'application, nous montrons que $I(L_y, L_z)$ est une métrique adaptée pour stopper le processus itératif et pour déterminer le facteur de pondération optimal à appliquer aux LLR extrinsèques.

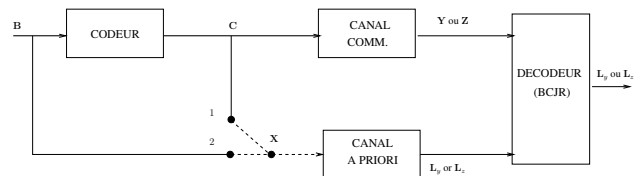


FIGURE 1 – Modèle générique de Codeur/Décodeur

2 Modèle et notations

Nous représentons les variables aléatoires (v.a.) à l'aide de lettres majuscules. Les réalisations particulières seront en minuscules. Les séquences de v.a. et les vecteurs apparaissent en gras. Le modèle considéré est schématisé [3] dans la figure 1. Le canal a priori modélise l'information présente à l'entrée du décodeur alors que le canal de communications modélise le milieu de transmission. Nous considérons ici un canal gaussien. Ce schéma correspond aussi bien à un turbo-code série qu'à un turbo-code parallèle. Dans le premier cas, le décodeur en sortie (outer) sera correctement modélisé avec l'interrupteur en position 1 et un canal de communication inactif. Dans le second cas, l'interrupteur sera en position 2 et le canal de communication sera actif. Ce modèle peut aussi être utilisé pour représenter la relation entrée/sortie au niveau des noeuds de parité ou de variable d'un décodeur LDPC. Dans tout ce qui suit, on notera $\mathbf{X} = (X_1, \dots, X_K)$ le message binaire à déterminer. Au niveau du récepteur, un décodeur donné reçoit en entrée une information a priori L_y sous la forme d'une séquence de longueur K , l'élément k est noté $L_{y,k}$ et représente une LLR extrinsèque sur le bit k . Ce même décodeur fournit en

sortie une information rafraîchie L_z qui servira d'information a priori au décodeur suivant. Dans une approche statistique, $L_{y,k}$ (resp. $L_{z,k}$) est vue comme une réalisation de la v.a. L_y (resp. L_z) et sera notée dans la suite ℓ_y (resp. ℓ_z) ou simplement ℓ quand une équation ou une propriété s'applique à ℓ_y et à ℓ_z . Nous supposons vérifiées les trois propriétés suivantes : *Symétrie* : $p_L(\ell|X) = p_L(-\ell - X)$, *Consistance généralisée* : $\frac{p_L(\ell|X=1)}{p_L(-\ell|X=1)} = e^{\alpha\ell}$ avec $\alpha \in \mathbb{R}^+$, *Support* : $\ell \in]-\infty; +\infty[$. Ce modèle est utilisé dans les Exit charts [1] avec $\alpha = 1$. On considère alors que L est une estimation bruitée de X telle que

$$L = \frac{\sigma^2}{2}X + \sigma n \quad (1)$$

avec $X \sim \mathcal{B}(1/2)$ et $n \sim \mathcal{N}(0, 1)$. La propriété de consistance généralisée devient $\frac{p_L(\ell|X=1)}{p_L(-\ell|X=1)} = e^\ell$. Il est montré dans [4] que la distribution des LLR peut être approchée par une loi gaussienne comme dans (1) mais que l'utilisation d'une seule variable pour représenter à la fois moyenne et variance est inappropriée. Nous considérerons ici le modèle suivant :

$$L = \alpha \frac{\sigma^2}{2}X + \sigma n \quad (2)$$

avec $n \sim \mathcal{N}(0, 1)$. Il est plus général que (1) et respecte les trois propriétés. Nous nous intéressons, dans la section suivante, au calcul de l'information mutuelle entre les LLR extrinsèques en sortie de deux décodeurs distincts.

3 Information mutuelle

L'information mutuelle $I(L_y, L_z)$ entre extrinsèques est définie comme suit :

$$I(L_y, L_z) = E_{p_{L_y, L_z}(\ell_y, \ell_z)} \left[\log_2 \left(\frac{p_{L_y, L_z}(\ell_y, \ell_z)}{p_{L_y}(\ell_y)p_{L_z}(\ell_z)} \right) \right] \quad (3)$$

Elle peut être reliée à l'information mutuelle tracée dans les Exit charts par l'intermédiaire de la relation :

$$I(L_y, L_z) = I(L_y, X) + I(L_z, X) - I(L_y + L_z, X) \quad (4)$$

avec $I(L, X) = 1 - \int_{-\infty}^{+\infty} p_L(\ell|X=1) \log_2(1 + e^{-\alpha\ell}) d\ell$. La démonstration utilise les propriétés listées dans la section 2 et l'indépendance de L_y et L_z conditionnellement à X [5]. L'évolution de $I(L_y, L_z)$ peut donc être simulée en utilisant la méthode de l'histogramme pour évaluer chacune des trois informations mutuelles de (4). L'expression de $I(L, X)$ pour L vérifiant le modèle (2) est donnée par

$$I(L, X) = 1 - \int_{-\infty}^{+\infty} \frac{e^{-\frac{(\ell - \alpha\mu)^2}{2\alpha^2\sigma^2}}}{\sqrt{2\pi\alpha\sigma}} \log_2(1 + e^{-\ell}) d\ell \quad (5)$$

ou de manière équivalente par

$$J(\alpha\sigma) = 1 - \int_{-\infty}^{+\infty} \frac{e^{-\frac{(\ell - \frac{\alpha^2\sigma^2}{2})^2}{2\alpha^2\sigma^2}}}{\sqrt{2\pi\alpha\sigma}} \log_2(1 + e^{-\ell}) d\ell \quad (6)$$

où J est une fonction monotone croissante [1]. On peut voir que même si L est définie à l'aide des deux variables α et σ , l'information mutuelle $I(L, X)$ est une fonction de la seule variable $\alpha\sigma$.

Pour établir, les propriétés ci-dessous, nous supposons que : (a) $I(L_y, L_z)$ s'écrit sous la forme (4) avec $I(L_y, X) = g(a_{L_y})$, $I(L_z, X) = g(a_{L_z})$, $I(L_y + L_z, X) = g(u(a_{L_y}, a_{L_z}))$ où g est une fonction croissante, (b) $u(a_{L_y}, a_{L_z}) \geq \max(a_{L_y}, a_{L_z})$. Ces hypothèses sont vérifiées avec les modèles (1) et (2) pour lesquels $u(a_{L_y}, a_{L_z}) = \sqrt{a_{L_y}^2 + a_{L_z}^2}$ et $g = J$.

Propriété 1. Si les hypothèses (a) et (b) sont vérifiées alors $I(L_y, L_z) \leq \min(I(L_y, X), I(L_z, X))$ et $I(L_y, L_z) \leq I(L_y + L_z, X)$.

Démonstration. $I(L_y + L_z, X) \geq g(\max(a_{L_y}, a_{L_z}))$. En utilisant la monotonie de la fonction g , on en déduit que $I(L_y + L_z, X) \geq \max(I(L_y, X), I(L_z, X))$ et en remplaçant dans (4) que $I(L_y, L_z) \leq \min(I(L_y, X), I(L_z, X))$ et que $I(L_y, L_z) \leq I(L_y + L_z, X)$. \square

L'information mutuelle entre extrinsèques est donc une borne inférieure de l'information mutuelle extrinsèque/message et également de l'information mutuelle entre la LLR a posteriori $L_y + L_z$ et le message qui représente la quantité de connaissance acquise par le récepteur sur le message X .

Corrolaire 1. Si $I(L_y, L_z) \geq 1 - \epsilon$ alors $I(L_y + L_z, X) \geq 1 - \epsilon$.

L'information mutuelle $I(L_y, L_z)$ a les bonnes propriétés pour être utilisée comme critère d'arrêt ainsi que comme un détecteur d'erreurs. Pour être utilisable au niveau du récepteur, nous devons trouver un estimateur de $I(L_y, L_z)$ pour des séquences de longueur finie. C'est l'objet de la section suivante.

4 Calcul efficace pour des séquences de longueur finie

Pour évaluer l'information mutuelle au niveau du récepteur, nous allons considérer la moyenne sur l'ensemble de la séquence. Dans un premier temps, nous considérons Q séquences toutes de longueur B et nous notons $K = QB$. Nous définissons l'information mutuelle moyenne entre extrinsèques :

$$I_M := \frac{1}{Q} \sum_{k=1}^Q I(L_{y_k}, L_{z_k}) \quad (7)$$

puis l'information mutuelle moyenne extrinsèque/message :

$$I_E := \frac{1}{Q} \sum_{k=1}^Q I(L_k, X_k) \quad (8)$$

Pour calculer ces deux quantités au niveau du récepteur, on utilise le lemme et le théorème ci dessous. Ces deux résultats sont issus de [3]. Ils sont rappelés ici pour le cas binaire.

Lemme 1. Soit $\ell_{y,k}$ la LLR extrinsèque à la sortie d'un décodeur donné, soit $\ell_{z,k}$ la LLR a priori à l'entrée du même décodeur et soit \mathbf{y} le vecteur d'observation. Soit $\ell_{z,[k]}$ la séquence ℓ_z de laquelle le terme d'indice k a été retiré. Alors pour tout $X_k \in \{-1; 1\}$, $p(X_k = x_k | L_{Y,k} = \ell_{y,k}) = p(X_k = x_k | \mathbf{y}, \ell_{z,[k]})$.

Ce lemme prouve que l'extrinsèque à la sortie du décodeur contient la même information sur X_k que le vecteur d'observation et la séquence d'a priori. Grâce à ce résultat, on obtient un estimateur de $I(L, X)$.

Théorème 1. Soit $p_z(x_k) = p(X_k = x_k | \mathbf{y}, \ell_{z,[k]})$ une séquence de v.a. ergodiques pour tout X_k dans $\{-1; 1\}$, on a

$$I_E = 1 + \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \sum_{x_k} p_z(x_k) \log_2(p_z(x_k)) \quad (9)$$

Cette quantité pourra être estimée au niveau du récepteur en moyennant sur des séquences de longueur $K = BQ$. On peut déduire du théorème 1 et de (4) une évaluation de I_M . On notera dans toute la suite $p_z(x_k) = p(X_k = x_k | \mathbf{y}, \ell_{z,[k]})$ et $p_y(x_k) = p(X_k = x_k | \mathbf{y}, \ell_{y,[k]})$.

Résultat 1. L'information mutuelle moyenne I_M s'écrit

$$I_M^{(a)} = 1 + \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \sum_{x_k} \left(p_z(x_k) \log_2(p_z(x_k)) + p_y(x_k) \log_2(p_y(x_k)) - \frac{p_y(x_k)p_z(x_k)}{s_k} \log_2 \left(\frac{p_y(x_k)p_z(x_k)}{s_k} \right) \right) \quad (10)$$

où p_y et p_z sont normalisées et où $s_k = \sum_{x_k} p_y(x_k)p_z(x_k)$.

La définition de l'information mutuelle donne une autre alternative pour exprimer I_M .

Résultat 2. L'information mutuelle I_M s'écrit

$$I_M^{(b)} = 1 + \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \sum_{x_k} \log_2(s_k) \quad (11)$$

avec $s_k = \sum_{x_k} p_y(x_k)p_z(x_k)$.

Démonstration. Par définition de l'information mutuelle,

$$I_M = \frac{1}{Q} \sum_{k=1}^Q E \left[\log_2 \left(\frac{p(\ell_{y,k}, \ell_{z,k})}{p(\ell_{y,k})p(\ell_{z,k})} \right) \right]$$

avec $p(\ell_{y,k}, \ell_{z,k}) = \sum_{x_k} \frac{p(x_k | \ell_{y,k})p(x_k | \ell_{z,k})}{p(x_k)} p(\ell_{y,k})p(\ell_{z,k})$. On obtient (11) en supposant ergodiques les v.a. considérées. \square

Nous venons de donner l'expression de deux estimateurs de $I(L_y, L_z)$ notés $I_M^{(a)}$ et $I_M^{(b)}$. Ils ont une complexité calculatoire linéaire avec K . Le nombre d'opérations est toutefois plus faible pour $I_M^{(b)}$ que pour $I_M^{(a)}$. Pour tester la précision de nos estimateurs, nous considérons un turbo-code série (SCTC) constitué d'un code $(5, 7)_8$ (outer code) et d'un second code

convolutif de polynôme générateur $\frac{1}{1+D}$. La taille K de la séquence évolue entre 500 et 8000 et on choisit $EbN_0 = 2dB$. La précision des estimateurs est évaluée par comparaison avec la valeur donnée par la méthode de l'histogramme (qui nécessite la connaissance du message X) et que nous noterons $I_E^{(hist)}$ et $I_M^{(hist)}$. Les résultats sont donnés sur la figure 2 sur laquelle est tracée l'erreur quadratique moyenne entre ces quantités. Nous pouvons constater la supériorité en terme de pré-

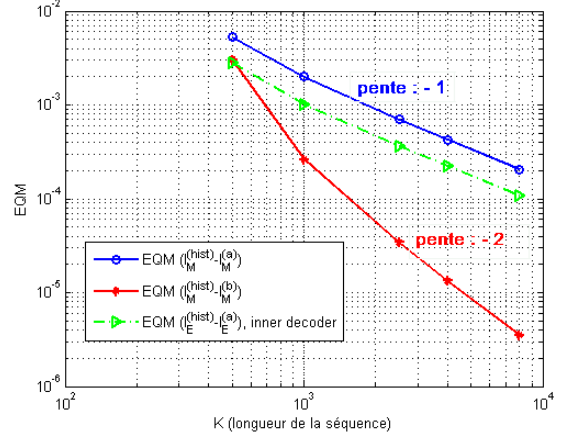


FIGURE 2 – Comparaison des estimateurs proposés.

sion de l'estimateur $I_M^{(b)}$ que nous retenons pour la suite. Dans un processus itératif, on propage en général des LLR. Notons $L_z = \log \frac{p(X_k=1 | \mathbf{y}, \ell_{z,[k]})}{p(X_k=-1 | \mathbf{y}, \ell_{z,[k]})}$ (notation similaire pour L_y).

Résultat 3. L'information mutuelle I_M est donnée par

$$I_M^{(b)} = 1 + \lim_{K \rightarrow \infty} \frac{1}{\log(2)K} \sum_{k=1}^K f(\ell_{y,k} + \ell_{z,k}) - f(\ell_{y,k}) - f(\ell_{z,k}) \quad (12)$$

avec $f(\ell) = \max(0, \ell) + \log(1 + \exp(-|\ell|))$.

Démonstration. On utilise la relation $\log(\exp(a) + \exp(b)) = \max(a, b) + \log(1 + \exp(-|a - b|))$. \square

L'expression (12) est équivalent asymptotiquement à (10). Mais, pour un bloc de taille finie, les estimateurs issus de (12) et (10) sont sensiblement différents (voir Figure 2). Dans cette section, nous avons établi l'expression d'un estimateur précis de $I(L_y, L_z)$. Nous montrons dans la section suivante son application potentielle au niveau d'un récepteur à décodage itératif.

5 Application au décodage itératif

5.1 $I(L_y, L_z)$ comme critère d'arrêt

La convergence d'un décodeur itératif vers la solution X du problème est caractérisée par des LLR de forte amplitude (en valeur absolue) et un accord entre les deux décodeurs individuels ($sign(L_{y,k}) = sign(L_{z,k})$). Nous appellerons ces deux

propriétés fiabilité et similitude. La plupart des critères d'arrêt utilisés dans la littérature ne sont sensibles qu'à une seule des deux conditions. Les critères basés sur la valeur absolue des LLR (comparées à un seuil) ne testent que la fiabilité des LLR alors que les critères de type cross-entropy ne testent que la similitude. Nous montrons dans le résultat ci-dessous que $I_M^{(b)}$ est sensible à la fois à la fiabilité et à la similitude.

Résultat 4. $I(L_y, L_z) \approx I_M^{(b)} = 1$ ssi $\text{sign}(L_{y,k}) = \text{sign}(L_{z,k})$ et $|L_{y,k}| = |L_{z,k}| = +\infty \forall k \in \{1, 2, \dots, K\}$.

Démonstration. $I_M^{(b)}$ peut s'écrire sous la forme $I_M^{(b)} = 1 + \frac{1}{K} \sum_{k=1}^K \log_2 \left(\frac{1+e^{L_{y,k}+L_{z,k}}}{(1+e^{L_{y,k}})(1+e^{L_{z,k}})} \right)$. $I_M^{(b)} = 1$ est équivalent à $\frac{1+e^{L_{y,k}+L_{z,k}}}{(1+e^{L_{y,k}})(1+e^{L_{z,k}})} = 1 \forall k \in \{1, 2, \dots, K\}$. Cette condition est satisfaite ssi $\text{sign}(L_{y,k}) = \text{sign}(L_{z,k})$ et $|L_{y,k}| = |L_{z,k}| = +\infty \forall k \in \{1, 2, \dots, K\}$. \square

Nous pouvons donc utiliser $I_M^{(b)}$ comme critère d'arrêt, pour tester la convergence vers la bonne décision ou pour stopper le processus itératif si $I_M^{(b)}$ n'a pas atteint un seuil fixé. Ce seuil, propre au système, pourra être facilement déterminé avec un Exit chart.

5.2 $I(L_y, L_z)$ pour estimer α

Il a été montré dans plusieurs publications [6, 7] qu'une correction linéaire des LLR est utile pour compenser les imperfections du récepteur. Nous proposons ici une méthode analytique permettant de calculer la valeur optimale du facteur de pondération. La méthode proposée est adaptative et ne nécessite pas l'estimation de la séquence X ce qui la distingue des contributions précédentes [6, 7]. Estimer la pondération optimale des LLR extrinsèques est équivalent à estimer α dans le modèle (2) [4]. Nous utilisons la méthode des histogrammes appliquée à (4) pour tracer l'évolution de α en fonction de celle de $I(L_y, L_z)$. Les résultats sont présentés sur la figure 3 pour le SCTC de la section 4. Nous notons α_y la valeur calculée à partir de L_y (sortie inner decoder) et α_z celle correspondant à la sortie de l'autre décodeur (outer). Il a déjà été observé dans la littérature que la valeur de α dépend du SNR et du numéro de l'itération rendant nécessaire une estimation adaptative. Nous voyons sur la figure 3 que α peut s'exprimer à l'aide d'une variable unique : $I(L_y, L_z)$; ce qui est un résultat nouveau. Nous avons appliqué le même procédé sur d'autres types de turbo-codes ainsi que sur des codes LDPC réguliers et irréguliers et nous avons toujours observé que $\alpha = F(I(L_y, L_z))$ où F dépend du système étudié. Nous en déduisons une méthode pour un choix optimal de α au niveau du récepteur : hors-ligne, tracer $\alpha = F(I(L_y, L_z))$ et en déduire l'expression de F par une méthode d'ajustement (polynomial); au récepteur, estimer $I(L_y, L_z)$ par $I_M^{(b)}$ et en déduire la valeur de α grâce à l'expression de F .

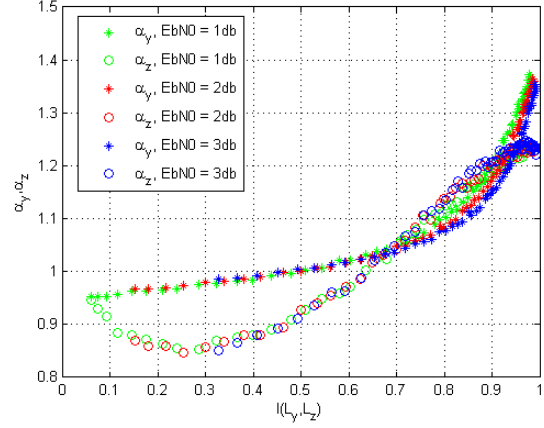


FIGURE 3 – Évolution de α vs $I(L_y, L_z)$

6 Conclusion

Nous avons présenté une nouvelle métrique, l'information mutuelle entre extrinsèques. Nous avons montré comment l'évaluer au niveau du récepteur, sur des séquences de longueur finie et sans connaissance du message. Nous avons également évalué la précision de l'estimateur. Nous avons enfin montré deux applications possibles de cette mesure. Au delà des exemples traités, l'information mutuelle entre extrinsèques est utilisable pour la sélection de relais dans un système coopératif ou encore comme indicateur de la connaissance accumulée au récepteur dans un système adaptatif à ré-émission des données.

Références

- [1] S. ten Brink. Convergence behavior of iteratively decoded parallel concatenated codes. *IEEE Trans. on Commun.*, 49 :1727–1737, Oct 2001.
- [2] E. Sharon, A. Ashikhmin, and S. Litsyn. Analysis of low-density parity-check codes based on EXIT functions. *IEEE Trans. on Commun.*, 54(8) :1407–1414, 2006.
- [3] J. Kliewer, S. Xin Ng, and L. Hanzo. Efficient computation of EXIT functions for nonbinary iterative decoding. *IEEE Trans. on Commun.*, 54(12) :2133–2136, December 2006.
- [4] M. Fu. Stochastic analysis of turbo decoding. *IEEE Trans on Inform. Theory*, 51(1) :81–100, 2005.
- [5] F. Alberge. On some properties of the mutual information between extrinsics with application to iterative decoding. Submitted to *IEEE Trans. on Commun.*, 2015.
- [6] A. Alvarado, V. Nunez, L. Szczecinski, and E. Agrell. Correcting suboptimal metrics in iterative decoders. In *ICC '09*, pages 1–6, June 2009.
- [7] J. Wu, M. El-Khamy, J. Lee, and I. Kang. BICM performance improvement via online LLR optimization. In *WCNC'13*, pages 3850–3855, April 2013.