



HAL
open science

Test d'hypothèses distribué pour des modèles de sources générales non-iid, non-stationnaires, et non-ergodiques

Ismaila Salihou Adamou, Elsa Dupraz, Tadashi Matsumoto

► **To cite this version:**

Ismaila Salihou Adamou, Elsa Dupraz, Tadashi Matsumoto. Test d'hypothèses distribué pour des modèles de sources générales non-iid, non-stationnaires, et non-ergodiques. GRETSI 2023 : XXIXème Colloque Francophone de Traitement du Signal et des Images, Aug 2023, Grenoble, France. pp.1061-1064. hal-04184078

HAL Id: hal-04184078

<https://imt-atlantique.hal.science/hal-04184078v1>

Submitted on 21 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Test d'hypothèses distribué pour des modèles de sources générales non-iid, non-stationnaires, et non-ergodiques

Ismaila SALIHOU ADAMOU¹ Elsa DUPRAZ¹ Tad MATSUMOTO^{1,2}

¹IMT Atlantique, CNRS UMR 6285, Lab-STICC, 29280 Brest, France

²JAIST and University of Oulu (Emeritus)

Résumé – Cet article traite du problème de test d'hypothèses distribué, où une source \mathbf{X} est encodée sachant qu'une information adjacente \mathbf{Y} n'est disponible qu'au décodeur. Le décodeur doit décider entre les hypothèses H_0 et H_1 relatives à la distribution jointe de \mathbf{X} et \mathbf{Y} en se basant sur une version codée de \mathbf{X} . Contrairement à la plupart des travaux existants qui considèrent des modèles de sources i.i.d., cet article étudie des modèles de sources plus génériques qui ne sont pas nécessairement i.i.d., ni stationnaires ou ergodiques. L'article développe une analyse basée sur l'"information spectrum" pour fournir des formules générales sur l'exposant d'erreur de Type-II atteignable sous une contrainte sur l'erreur de Type-I. Le schéma de codage atteignable est basé sur une approche "quantize-and-binning". L'article montre que pour ce schéma de codage, l'exposant d'erreur se résume à un problème de compromis entre une erreur de binning et une erreur de décision, comme cela a déjà été observé pour les sources i.i.d. Enfin, La dernière partie de l'article traite le modèle particulier des sources gaussiennes, stationnaires et ergodiques.

Abstract – This paper investigates Distributed Hypothesis testing (DHT), in which a source \mathbf{X} is encoded given that side information \mathbf{Y} is available at the decoder only. Based on the received coded data, the receiver aims to decide on the two hypotheses H_0 or H_1 related to the joint distribution of \mathbf{X} and \mathbf{Y} . While most existing contributions in the literature on DHT consider i.i.d. assumptions, this paper assumes more generic, non-i.i.d., non-stationary, and non-ergodic sources models. It relies on information-spectrum to provide general formulas on the achievable Type-II error exponent under a constraint on the Type-I error. The achievability proof is based on a quantize-and-binning scheme. It is shown that with the quantize-and-binning approach, the error exponent boils down to a problem of making a trade-off between a binning error and a decision error, as already observed for the i.i.d. sources. The last part of the paper provides error exponents for particular source models, *e.g.*, gaussian, stationary, and ergodic models.

1 Introduction

Dans les réseaux de communication distribués, les données sont collectées à partir de différents nœuds distants puis envoyées à un serveur pour traitement. Souvent, l'objectif principal du serveur n'est pas de reconstruire les données, mais plutôt de prendre une décision à partir des données collectées. Ce type de configuration est appelé test d'hypothèses distribué, et a été étudié pour la première fois du point de vue de la théorie de l'information dans [3].

Dans ce cadre, une source \mathbf{X} est codée sachant qu'une information adjacente \mathbf{Y} est disponible uniquement au décodeur, comme illustré en Figure 1. Le décodeur vise à prendre une décision entre deux hypothèses H_0 ou H_1 , sur la distribution de probabilité jointe de (\mathbf{X}, \mathbf{Y}) . L'analyse du test d'hypothèses du point de vue de la théorie de l'information vise à déterminer l'exposant d'erreur réalisable pour l'erreur de Type-II, tout en imposant une contrainte fixe sur l'erreur de Type-I, et en considérant que la transmission de \mathbf{X} se fait sur un canal limité en débit [3]. La littérature sur ce sujet considère différents modèles de source pour \mathbf{X} et \mathbf{Y} .

La plupart des travaux existants supposent que les sources \mathbf{X} et \mathbf{Y} génèrent conjointement des suites de symboles (X_t, Y_t) indépendantes et identiquement distribuées (i.i.d.) [7, 9]. Certains modèles de source plus complexes ont également été étudiés, tels que ceux de [1, 8], qui supposent que les sources \mathbf{X} et \mathbf{Y} génèrent conjointement des suites de vecteurs gaussiens $(\mathbf{X}_t, \mathbf{Y}_t)$ ($t \in \llbracket 1, T \rrbracket$) avec des auto-corrélations inté-

grées dans chaque vecteur \mathbf{X}_t et \mathbf{Y}_t , ainsi qu'une corrélation croisée entre eux. Cependant, les modèles de [1, 8] sont de type bloc-i.i.d. dans le sens où les paires successives $(\mathbf{X}_t, \mathbf{Y}_t)$ sont supposées être i.i.d. avec t .

Cependant, les modèles i.i.d. et bloc-i.i.d. sont souvent inadaptés pour capturer les statistiques de signaux tels que les séries temporelles ou les vidéos, qui ne peuvent pas être décomposés en blocs indépendants de longueur fixe et sont fréquemment non-stationnaires et/ou non-ergodiques. Par conséquent, l'objectif de cet article est de considérer un modèle de sources non i.i.d. plus général, et qui pourra prendre en compte de signaux non-stationnaires et non-ergodiques, tout en englobant les modèles précédents en tant que cas particuliers. Pour étudier le problème de test d'hypothèses distribué dans ces conditions, nous proposons d'utiliser des outils d'"information-spectrum" [5], qui fournissent des résultats de théorie de l'information applicables à une large gamme de modèles de source. Il convient de noter que l'"Information-spectrum" a été précédemment utilisé pour le test d'hypothèses dans [4], mais uniquement pour l'encodage d'une source \mathbf{X} seule, sans l'information adjacente \mathbf{Y} .

Dans cet article, pour des modèles de sources générales \mathbf{X} et \mathbf{Y} (voir Section 2), nous proposons une expression générale pour l'exposant d'erreur de Type-II (Section 3), qui s'appuie sur un schéma d'atteignabilité de type "quantize-and-binning" (Section 4). Comme dans [7] (qui considère un modèle i.i.d.), notre exposant d'erreur se compose de deux termes : un pour l'erreur de binning et l'autre pour l'erreur de décision. Nous

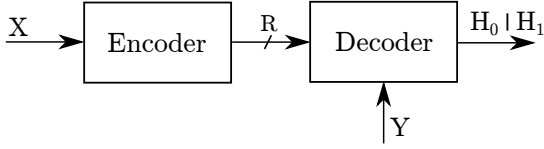


FIGURE 1 : Schéma de codage de test d'hypothèses distribué

spécifions ensuite notre exposant d'erreur pour des sources gaussiennes non i.i.d., stationnaires et ergodiques (Section 5).

2 Énoncé du Problème

Dans le problème de test d'hypothèses distribué illustré en Figure 1, le décodeur reçoit une description compressée de \mathbf{X} et observe l'information adjacente \mathbf{Y} , où \mathbf{X} et \mathbf{Y} sont corrélées.

2.1 Modèle de sources générales

Nous adoptons la définition de [4] pour le modèle de sources générales. Les sources \mathbf{X} et \mathbf{Y} génèrent deux suites infinies :

$$\begin{cases} \mathbf{X}^n = (X_1, X_2, \dots, X_n)_{n=1}^\infty, \\ \mathbf{Y}^n = (Y_1, Y_2, \dots, Y_n)_{n=1}^\infty \end{cases} \quad (1)$$

de variables aléatoires $\mathbf{X}^n, \mathbf{Y}^n$, chacune de dimension n . De plus, les symboles X_i, Y_i prennent leurs valeurs dans des alphabets \mathcal{X}, \mathcal{Y} , respectivement. Le modèle de [4] suppose ensuite que les distributions de probabilités jointes $P_{\mathbf{X}^n \mathbf{Y}^n}$ sont connues mais quelconques (pas d'hypothèse de stationnarité ou d'ergodicité).

On peut distinguer deux cas particuliers du modèle (1) : le modèle scalaire i.i.d. et le modèle bloc i.i.d. Dans le premier cas, qui a été étudié dans [7, 9] pour le test d'hypothèses distribué, les séquences \mathbf{X}^n et \mathbf{Y}^n produisent conjointement des suites de symboles (X_n, Y_n) i.i.d., selon une distribution P_{XY} fixée. Dans le second cas, qui a été étudié dans [1, 8], les séquences de sources \mathbf{X}^n et \mathbf{Y}^n sont définies comme :

$$\mathbf{X}^n = \{\mathbf{X}_t^M\}_{t=1}^n, \quad \mathbf{Y}^n = \{\mathbf{Y}_t^M\}_{t=1}^n, \quad (2)$$

où $\{\mathbf{X}_t^M\}_{t=1}^n$ et $\{\mathbf{Y}_t^M\}_{t=1}^n$ sont des suites de vecteurs aléatoires i.i.d. de dimension M . En d'autres termes, les couples successifs $(\mathbf{X}_t^M, \mathbf{Y}_t^M)$ sont i.i.d. selon la même distribution jointe $P_{\mathbf{X}^M \mathbf{Y}^M}$. Le fait que ces deux modèles soient i.i.d. simplifie leur analyse.

2.2 Test d'hypothèses distribué

Nous définissons maintenant le problème du test d'hypothèses distribué pour le modèle de sources générales que nous avons présenté précédemment. Nous supposons que la loi jointe du couple $(\mathbf{X}^n, \mathbf{Y}^n)$ dépend des hypothèses sous-jacentes H_0 et H_1 , définies comme suit :

$$H_0 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{\mathbf{X}^n \mathbf{Y}^n}, \quad (3)$$

$$H_1 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{\bar{\mathbf{X}}^n \bar{\mathbf{Y}}^n}. \quad (4)$$

Nous supposons de plus que les distributions de probabilités marginales $P_{\mathbf{X}^n}$ et $P_{\mathbf{Y}^n}$ sont les mêmes sous H_0 et H_1 afin

d'éviter que les observateurs prennent des décisions non triviales en se basant uniquement sur leurs propres observations. Nous adoptons le schéma de codage pour le test d'hypothèses distribué habituellement considéré dans la littérature [2, 3, 7] :

Définition 1 Les fonction d'encodage $f^{(n)}$ et de décodage $g^{(n)}$ sont définies comme suit :

$$f^{(n)} : \mathcal{X}^n \longrightarrow \mathcal{M}_n = \llbracket 1, M_2 \rrbracket, \quad (5)$$

$$g^{(n)} : \mathcal{M}_n \times \mathcal{Y}^n \longrightarrow \mathcal{H} = \{H_0, H_1\}, \quad (6)$$

telles que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_2 \leq R, \quad (7)$$

avec M_2 le cardinal de l'ensemble \mathcal{M}_n .

Définition 2 Les probabilités d'erreur de Type-I et de Type-II, α_n et β_n , sont définies comme suit :

$$\alpha_n = \mathbb{P} \left[g^{(n)} \left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n \right) = H_1 \mid H_0 \right] \quad (8)$$

$$\beta_n = \mathbb{P} \left[g^{(n)} \left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n \right) = H_0 \mid H_1 \right]. \quad (9)$$

Définition 3 Un exposant d'erreur de Type-II noté θ est dit atteignable si, pour des tailles n de séquences assez grandes, il existe des fonctions d'encodage et de décodage $(f^{(n)}, g^{(n)})$ telles que les probabilités d'erreur de Type-I et de Type-II α_n et β_n vérifient

$$\alpha_n \leq \epsilon, \quad (10)$$

et

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq \theta \quad (11)$$

pour tout $\epsilon > 0$.

3 Exposant d'erreur

Dans cette partie, nous présentons notre résultat principal sur l'exposant d'erreur de Type-II réalisable θ pour les sources générales définies dans la Section 2.1. Nous rappelons d'abord les définitions des outils de l'"information spectrum" de [5] qui seront utiles pour notre analyse.

3.1 Définitions

Tout d'abord, nous définissons respectivement la lim sup et la lim inf en probabilité d'une suite de variables aléatoires $\{Z_n\}_{n=1}^\infty$ comme suit [5] :

$$p - \limsup_{n \rightarrow \infty} Z_n = \inf \left\{ \alpha \mid \lim_{n \rightarrow +\infty} \mathbb{P}(Z_n > \alpha) = 0 \right\}$$

$$p - \liminf_{n \rightarrow \infty} Z_n = \sup \left\{ \alpha \mid \lim_{n \rightarrow +\infty} \mathbb{P}(Z_n < \alpha) = 0 \right\}$$

Ensuite, l'information mutuelle spectrale supérieure $\bar{I}(\mathbf{X}; \mathbf{U})$, l'information mutuelle spectrale inférieure $\underline{I}(\mathbf{U}; \mathbf{Y})$, la divergence spectrale inférieure $\underline{D}(P_{\mathbf{U}\mathbf{Y}} \| P_{\bar{\mathbf{U}}\bar{\mathbf{Y}}})$, et la divergence spectrale supérieure $\bar{D}(P_{\mathbf{U}\mathbf{Y}} \| P_{\bar{\mathbf{U}}\bar{\mathbf{Y}}})$ sont respectivement définies comme [5]

$$\bar{I}(\mathbf{X}; \mathbf{U}) = p - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \frac{P_{\mathbf{U}^n | \mathbf{X}^n}(\mathbf{U}^n | \mathbf{X}^n)}{P_{\mathbf{U}^n}(\mathbf{U}^n)}, \quad (12)$$

$$\underline{I}(\mathbf{U}; \mathbf{Y}) = p - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{P_{\mathbf{U}^n | \mathbf{Y}^n}(\mathbf{U}^n | \mathbf{Y}^n)}{P_{\mathbf{U}^n}(\mathbf{U}^n)}, \quad (13)$$

$$\underline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}}) = \mathbb{p} - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{P_{\mathbf{U}^n \mathbf{Y}^n}(\mathbf{U}^n, \mathbf{Y}^n)}{P_{\overline{\mathbf{U}^n \mathbf{Y}^n}}(\mathbf{U}^n, \mathbf{Y}^n)}, \quad (14)$$

$$\overline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}}) = \mathbb{p} - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \frac{P_{\mathbf{U}^n \mathbf{Y}^n}(\mathbf{U}^n, \mathbf{Y}^n)}{P_{\overline{\mathbf{U}^n \mathbf{Y}^n}}(\mathbf{U}^n, \mathbf{Y}^n)}. \quad (15)$$

Dans le cas i.i.d., on retrouve les définitions classiques de l'information mutuelle et de la divergence.

3.2 Exposant d'erreur atteignable pour des sources générales

Theorem 1 *Pour le schéma de codage de la Définition 1, l'exposant d'erreur θ suivant est atteignable pour des sources générales définies par (1) :*

$$\theta \leq \min \left\{ r - (\overline{I}(\mathbf{X}; \mathbf{U}) - \underline{I}(\mathbf{U}; \mathbf{Y})), \right. \\ \left. \underline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}}) + (\underline{I}(\mathbf{X}; \mathbf{U}) - \overline{I}(\mathbf{X}; \mathbf{U})) \right\}, \quad (16)$$

avec \mathbf{U} , variable aléatoire auxiliaire telle que la chaîne de Markov $\mathbf{U} \rightarrow \mathbf{X} \rightarrow \mathbf{Y}$ est satisfaite à la fois sous H_0 et sous H_1 . $P_{\mathbf{U}\mathbf{Y}}$ et $P_{\overline{\mathbf{U}\mathbf{Y}}}$ sont les distributions jointes de $(\mathbf{U}^n, \mathbf{Y}^n)$ sous H_0 et H_1 , respectivement, et $r \leq R$.

On remarque que lorsque les sources \mathbf{X}, \mathbf{Y} et \mathbf{U} sont i.i.d., notre exposant d'erreur se résume à celui trouvé dans [7]. Ceci montre la cohérence de notre analyse.

4 Démonstration du Théorème 1

En raison des contraintes d'espace, seules les grandes lignes de la démonstration du Théorème 1 sont présentées dans cet article.

4.1 Stratégie de codage et de décodage

Génération de codebook aléatoire et encodeur : On génère d'abord $M_1 = e^{nr_0}$ séquences \mathbf{u}^n indépendamment selon une distribution fixe $P_{\mathbf{U}^n|\mathbf{X}^n}$ et on les assigne de manière aléatoire à l'un des $M_2 = e^{nr}$ bins selon une distribution uniforme. On note $\mathbf{B}(\mathbf{u}^n) \in \llbracket 1, M_2 \rrbracket$, l'indice du bin auquel \mathbf{u}^n appartient. Ensuite, pour une séquence \mathbf{x}^n donnée, l'encodeur cherche une séquence \mathbf{u}^n telle que $(\mathbf{x}^n, \mathbf{u}^n) \in T_n^{(1)}$, où $T_n^{(1)}$ est défini sous H_0 par

$$T_n^{(1)} = \left\{ (\mathbf{x}^n, \mathbf{u}^n) \text{ t.q. } \frac{1}{n} \log \frac{P_{\mathbf{U}^n|\mathbf{X}^n}(\mathbf{u}^n | \mathbf{x}^n)}{P_{\mathbf{U}^n}(\mathbf{u}^n)} < r_0 - \varepsilon \right\}. \quad (17)$$

Si une telle séquence est trouvée, l'encodeur envoie l'indice $\mathbf{B}(\mathbf{u}^n)$ du bin auquel \mathbf{u}^n appartient. Sinon, il envoie un message d'erreur.

Décodeur : Le décodeur opère en deux étapes. D'abord, pour l'index de bin reçu et l'information adjacente \mathbf{y}^n , le décodeur teste les séquences \mathbf{u}^n une par une dans le bin en vérifiant si $(\mathbf{y}^n, \mathbf{u}^n) \in T_n^{(2)}$, où $T_n^{(2)}$ est défini sous H_0 par

$$T_n^{(2)} = \left\{ (\mathbf{y}^n, \mathbf{u}^n) \text{ t.q. } \frac{1}{n} \log \frac{P_{\mathbf{U}^n|\mathbf{Y}^n}(\mathbf{u}^n | \mathbf{y}^n)}{P_{\mathbf{U}^n}(\mathbf{u}^n)} > r' - \varepsilon \right\}. \quad (18)$$

Si une telle séquence n'est pas trouvée ou si le décodeur reçoit un message d'erreur de l'encodeur, le décodeur rejette l'hypothèse H_0 . Dans la deuxième étape, si une séquence est extraite du bin, le décodeur compare la séquence décodée avec une région d'acceptation \mathcal{A}_n définie sous H_0 par

$$\mathcal{A}_n = \left\{ (\mathbf{y}^n, \mathbf{u}^n) \text{ t.q. } S < \frac{1}{n} \log \frac{P_{\mathbf{U}^n \mathbf{Y}^n}(\mathbf{u}^n, \mathbf{y}^n)}{P_{\overline{\mathbf{U}^n \mathbf{Y}^n}}(\mathbf{u}^n, \mathbf{y}^n)} \right. \\ \left. \leq \overline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}}) \right\}, \quad (19)$$

où S est le seuil de décision. Si la séquence décodée appartient à cette région, le décodeur accepte l'hypothèse H_0 . Sinon, il rejette l'hypothèse H_0 . Nous notons que les paramètres r_0 , r' et S sont choisis tel que la contrainte sur l'erreur de Type-I soit satisfaite.

4.2 Analyse des probabilités d'erreur

Erreur de Type-I : Les événements d'erreur qui conduisent le décodeur à commettre une erreur de Type-I sont les suivants :

$$E_{11} = \left\{ \nexists \mathbf{u}^n \text{ tel que } (\mathbf{X}^n, \mathbf{u}^n) \in T_n^{(1)}, (\mathbf{Y}^n, \mathbf{u}^n) \in T_n^{(2)}, \right. \\ \left. (\mathbf{Y}^n, \mathbf{u}^n) \in \mathcal{A}_n \right\}, \quad (20)$$

$$E_{12} = \left\{ \exists \mathbf{u}'^n \neq \mathbf{u}^n \text{ tel que } \mathbf{B}(\mathbf{u}'^n) = \mathbf{B}(\mathbf{u}^n), (\mathbf{Y}^n, \mathbf{u}'^n) \in T_n^{(2)}, \right. \\ \left. \text{mais } (\mathbf{u}'^n, \mathbf{Y}^n) \notin \mathcal{A}_n \right\}. \quad (21)$$

L'évènement E_{11} correspond à une erreur pendant l'encodage, pendant le debinning, ou lors du test final. L'évènement E_{12} correspond au cas où une séquence incorrecte est extraite du bin. Par la borne de l'union, la probabilité d'erreur de Type-I est bornée par $\alpha_n \leq \mathbb{P}(E_{11}) + \mathbb{P}(E_{12})$. En appliquant [6, Lemma 1], on montre que si $r_0 \geq \overline{I}(\mathbf{X}; \mathbf{U})$, $r' \leq \underline{I}(\mathbf{Y}; \mathbf{U})$ et $S = \underline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}})$, alors $\mathbb{P}(E_{11}) \rightarrow 0$ lorsque n tend vers l'infini. De plus, en appliquant les mêmes étapes que dans la preuve de [6], on montre que $\mathbb{P}(E_{12}) \rightarrow 0$ lorsque n tend vers l'infini. Cela implique que pour tout $\varepsilon > 0$ et pour n assez grand, la condition $\alpha_n \leq \varepsilon$ est satisfaite.

Erreur de Type-II : Les événements d'erreurs qui peuvent mener à une erreur de Type-II sont

$$E_{21} = \left\{ \exists \tilde{\mathbf{u}}^n \neq \mathbf{u}^n : \mathbf{B}(\tilde{\mathbf{u}}^n) = \mathbf{B}(\mathbf{u}^n), (\overline{\mathbf{Y}}^n, \tilde{\mathbf{u}}^n) \in T_n^{(2)}, \right. \\ \left. \text{mais } (\overline{\mathbf{Y}}^n, \tilde{\mathbf{u}}^n) \in \mathcal{A}_n \right\},$$

$$E_{22} = \left\{ (\mathbf{u}^n, \overline{\mathbf{Y}}^n) \in T_n^{(2)}, (\mathbf{u}^n, \overline{\mathbf{Y}}^n) \in \mathcal{A}_n \right\}. \quad (22)$$

En appliquant la borne de l'union, on obtient $\beta_n \leq \mathbb{P}(E_{21}) + \mathbb{P}(E_{22})$. On borne la probabilité de l'évènement E_{21} en appliquant les mêmes étapes que pour l'étude de la convergence de E_{12} . Pour borner la probabilité de E_{22} , on utilise la définition de l'ensemble \mathcal{A}_n dans (19). On obtient que β_n peut être borné par

$$\beta_n \leq e^{-n(r - (\overline{I}(\mathbf{X}; \mathbf{U}) - \underline{I}(\mathbf{Y}; \mathbf{U})) - \varepsilon)} \\ + e^{-n(\underline{I}(\mathbf{X}; \mathbf{U}) - \overline{I}(\mathbf{X}; \mathbf{U}) + \underline{D}(P_{\mathbf{U}\mathbf{Y}}\|P_{\overline{\mathbf{U}\mathbf{Y}}}) - 2\varepsilon)}. \quad (23)$$

Enfin, en utilisant la définition de l'exposant d'erreur θ donnée par (11), nous montrons que (16) est atteignable, ce qui prouve le théorème 1.

5 Sources Gaussiennes

Nous appliquons maintenant le Théorème 1 à des modèles de sources Gaussiennes non i.i.d. mais stationnaires et ergodiques, telles que $\mathbf{X} \sim \mathcal{N}(0, \mathbf{K}_X)$ et $\mathbf{Y} \sim \mathcal{N}(0, \mathbf{K}_Y)$, où \mathbf{K}_X et \mathbf{K}_Y sont les matrices de covariance de \mathbf{X} et \mathbf{Y} , respectivement. Nous définissons les deux hypothèses suivantes :

$$H_0 : \begin{pmatrix} \mathbf{X}^n \\ \mathbf{Y}^n \end{pmatrix} \sim \mathcal{N}(0, \mathbf{K}), \quad (24)$$

$$H_1 : \begin{pmatrix} \mathbf{X}^n \\ \mathbf{Y}^n \end{pmatrix} \sim \mathcal{N}(0, \bar{\mathbf{K}}), \quad (25)$$

avec $\mathbf{K} = \begin{bmatrix} \mathbf{K}_X & \mathbf{K}_{XY} \\ \mathbf{K}_{YX} & \mathbf{K}_Y \end{bmatrix}$ et $\bar{\mathbf{K}} = \begin{bmatrix} \mathbf{K}_X & \bar{\mathbf{K}}_{XY} \\ \bar{\mathbf{K}}_{YX} & \mathbf{K}_Y \end{bmatrix}$,

Nous supposons que toutes les matrices \mathbf{K}_X , \mathbf{K}_Y , $\bar{\mathbf{K}}_{XY}$, \mathbf{K}_{XY} , and $\bar{\mathbf{K}}_{YX}$ sont définies positives. Nous définissons également

$$\mathbf{K}_{X|Y} = \mathbf{K}_X - \mathbf{K}_{XY}\mathbf{K}_Y^{-1}\mathbf{K}_{YX}. \quad (26)$$

Les valeurs propres de $\mathbf{K}_{X|Y}$ seront notées $\lambda_i^{(X|Y)}$.

Proposition 1 Si les sources \mathbf{X} et \mathbf{Y} sont stationnaires et ergodiques sous H_0 et H_1 , l'exposant d'erreur (16) devient :

$$\theta \leq \min \left\{ r - \lim_{n \rightarrow \infty} \left[\frac{1}{n} h(\mathbf{U}^n | \mathbf{Y}^n) - \frac{1}{n} h(\mathbf{U}^n | \mathbf{X}^n) \right], \lim_{n \rightarrow \infty} \frac{1}{n} D(P_{\mathbf{U}^n \mathbf{Y}^n} \| P_{\bar{\mathbf{U}}^n \bar{\mathbf{Y}}^n}) \right\}. \quad (27)$$

Cette proposition découle de la propriété "strong converse property" [5, Page 48-49]. En plus, si \mathbf{X} et \mathbf{Y} sont gaussiennes, stationnaires et ergodiques, les termes dans (27) se réduisent à :

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} h(\mathbf{U}^n | \mathbf{Y}^n) - \lim_{n \rightarrow \infty} \frac{1}{n} h(\mathbf{U}^n | \mathbf{X}^n) \\ &= \lim_{n \rightarrow \infty} \frac{1}{2n} \sum_{i=1}^n \log \frac{\lambda_i^{(X|Y)} + \kappa}{\kappa}, \end{aligned} \quad (28)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} D(P_{\mathbf{U}^n \mathbf{Y}^n} \| P_{\bar{\mathbf{U}}^n \bar{\mathbf{Y}}^n}) &= \lim_{n \rightarrow \infty} \frac{1}{2n} \left[\log \frac{|\bar{\Sigma}|}{|\Sigma|} \right. \\ & \left. - 2n + \text{tr} \left\{ \bar{\Sigma}^{-1} \Sigma \right\} \right]. \end{aligned} \quad (29)$$

où Σ et $\bar{\Sigma}$ sont les matrices de covariance conjointes de \mathbf{U} et \mathbf{Y} sous H_0 et H_1 , respectivement, $|\cdot|$ représente un déterminant, et $\text{tr}(\cdot)$ représente la trace. Les termes donnés par (28) and (29) sont obtenus en considérant que la source \mathbf{U} est gaussienne de sorte que $\mathbf{U} = \mathbf{X} + \mathbf{Z}$. Avec $\mathbf{Z} \sim \mathcal{N}(0, \kappa \mathbf{I}_n)$, indépendant de \mathbf{X} et \mathbf{I}_n est la matrice identité de dimension $n \times n$. Les matrices de covariance Σ et $\bar{\Sigma}$ sont alors définies comme $\Sigma = \begin{bmatrix} \mathbf{K}_U & \mathbf{K}_{UY} \\ \mathbf{K}_{YU} & \mathbf{K}_Y \end{bmatrix}$ et $\bar{\Sigma} = \begin{bmatrix} \mathbf{K}_U & \bar{\mathbf{K}}_{UY} \\ \bar{\mathbf{K}}_{YU} & \mathbf{K}_Y \end{bmatrix}$. On note que les matrices Σ et $\bar{\Sigma}$ sont de dimension $(2n) \times (2n)$, avec $n \rightarrow \infty$ dans les équations précédentes. Quand on considère un modèle de sources Gaussiennes particulières, il est donc nécessaire d'étudier la convergence des déterminants et trace impliqués dans les équations précédentes.

6 Conclusion

L'étude menée dans cet article utilise une approche "information-spectrum" pour obtenir une formule générale de l'exposant d'erreur de type-II pour les sources non-i.i.d., non-stationnaires et non-ergodiques. Le schéma de codage atteignable est basé sur le principe du "quantize-and-binning", et l'exposant d'erreur résultant représente un compromis entre une erreur de binning et une erreur de décision, comme pour les sources i.i.d. Les travaux futurs incluront la prise en compte d'un modèle de corrélation de Markov caché entre \mathbf{X} et \mathbf{Y} , et la conception de schémas de codage pratiques¹.

Références

- [1] Pierre ESCAMILLA, Abdellatif ZAIDI et Michele WIGGER : Some Results on the Vector Gaussian Hypothesis Testing Problem. *IEEE Int. Symp. Inf. Theory - Proc.*, 2020-June(May):2421–2425, 2020.
- [2] Eli HAIM et Yuval KOCHMAN : On Binary Distributed Hypothesis Testing. pages 1–37, 2017.
- [3] Te Sun HAN : Hypothesis Testing with Multiterminal Data Compression. *IEEE Trans. Inf. Theory*, 33(6):759–772, 1987.
- [4] Te Sun HAN : Hypothesis testing with the general source. *IEEE Trans. Inf. Theory*, 46(7):2415–2427, 2000.
- [5] Te Sun HAN : *Information-Spectrum Methods in Information Theory*. Baifukan, Tokyo, 1998.
- [6] Ken-ichi IWATA et Jun MURAMATSU : An information-spectrum approach to rate-distortion function with side information. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 85(6):1387–1395, 2002.
- [7] Gil KATZ, Pablo PIANTANIDA, Romain COUILLET et Merouane DEBBAH : On the necessity of binning for the distributed hypothesis testing problem. *IEEE Int. Symp. Inf. Theory - Proc.*, 2015-June:2797–2801, 2015.
- [8] Md Saifur RAHMAN et Aaron B. WAGNER : Vector gaussian hypothesis testing and lossy one-helper problem. *IEEE Int. Symp. Inf. Theory - Proc.*, pages 968–972, 2009.
- [9] Md Saifur RAHMAN et Aaron B. WAGNER : On the optimality of binning for distributed hypothesis testing. *IEEE Trans. Inf. Theory*, 58(10):6282–6303, 2012.

¹Cette recherche est financée par le laboratoire d'excellence CominLabs ANR-10-LABX-07-01