

Quantification vectorielle haute résolution pour la détection de processus stationnaires

Joffrey VILLARD^{1*}, Pascal BIANCHI²

¹SUPELEC, Dpt. Télécommunications, Gif-sur-Yvette, France.

²LTCI Telecom Paristech, Paris, France.

joffrey.villard@supelec.fr, bianchi@telecom-paristech.fr

Résumé – Cet article traite de la quantification haute résolution pour la détection de processus stationnaires. Dans le contexte des réseaux de capteurs sans fil, il s’agit de détecter la présence d’un signal à partir des mesures d’un grand nombre de capteurs, préalablement rassemblées à un Centre de Fusion (CF). La présence de liens sans fil imparfaits requiert que les données soient quantifiées avant leur transmission au CF. Le but de ces travaux est d’une part de déterminer l’impact d’une telle quantification sur les performances du système (en termes de détection, mesurées via l’exposant d’erreur) et, d’autre part, de proposer des règles de quantification qui minimisent cet impact. Il est démontré que la probabilité de manque du test de Neyman-Pearson converge exponentiellement vers zéro, quand le nombre de capteurs tend vers l’infini et le taux de fausse alarme est fixé. Sous certaines hypothèses, valides pour une large classe de processus stationnaires à valeurs vectorielles, nous fournissons une expression compacte de la dégradation de l’exposant d’erreur due à la quantification dans le régime des hautes résolutions. À partir de cette expression, nous déterminons des stratégies de quantification appropriées qui permettent de réduire cette dégradation.

Abstract – This paper investigates the high-rate quantization for the detection of stationary processes. It is assumed that some sensors observe samples of a correlated stationary ergodic multivariate process. These measurements are transmitted to a remote Fusion Center (FC) which performs a binary hypothesis test. Due to imperfect links between the sensors and the FC, the measurements must be quantized prior to transmission. As a matter of fact, this quantization step may severely degrade the overall detection performance of the system. The aim of this work is to quantify the detection performance of a given quantizer and characterize quantization strategies which guarantee attractive performance at the FC. For any false alarm level, it is shown that the miss probability of the Neyman-Pearson test converges to zero exponentially as the number of sensors tends to infinity, assuming that the observed process satisfies certain mixing conditions. The main contribution of this paper is to provide a compact closed-form expression of the error exponent in the high-rate regime *i.e.*, when the number of quantization levels tends to infinity. As an application, relevant high-rate quantization strategies which lead to a large error exponent are determined.

1 Introduction

Considérons un Réseau de Capteurs Sans Fil (RCSF) dont le but est de détecter la présence d’un signal à partir des mesures d’un grand nombre de capteurs. Ce réseau comporte de plus un centre de fusion (CF) qui rassemble les données des capteurs et prend la décision finale. Dans le cas de tests d’hypothèse binaires, le CF doit décider parmi deux hypothèses possibles, H_0 et H_1 . La procédure de Neyman-Pearson (NP) fournit alors un test uniformément plus puissant [1]. Dans le contexte des réseaux de capteurs, un grand nombre de travaux a été consacré à l’étude des performances de ces tests NP dans le but de concevoir des RCSF présentant de bonnes performances en termes de détection. La majorité de ces travaux s’intéresse au cas où les observations des capteurs sont des variables aléatoires indépendantes. Cependant, la détection de processus spatialement corrélés est un problème crucial dans les applications RCSF. L’im-

portance de la densité de capteurs sur la détection lorsque les observations sont corrélées a été récemment étudié dans [2]. Dans le cas de la détection d’un signal gaussien markovien bruité, la probabilité de manque du test NP converge exponentiellement vers zéro lorsque le taux de fausse alarme est fixé [3]. Les résultats de [3] fournissent également une expression compacte de l’exposant d’erreur.

Ces travaux, dans leur majorité, supposent que le CF a une connaissance parfaite des mesures des capteurs. En pratique, la présence de liens sans fil imparfaits requiert que les données soient compressées/quantifiées avant leur transmission au CF. Au cours des dernières décennies, de nombreuses recherches ont été consacrées à la caractérisation de stratégies de quantification pour diverses applications. Bennett [4] a le premier étudié la quantification *haute résolution* (ou *haut débit*) pour la reconstruction de signaux scalaires. Ses résultats ont été par la suite étendus au cas d’observations vectorielles [5]. La quantification haute résolution dans le contexte des tests statistiques

*Les travaux de J. Villard sont soutenus par la DGA.

et de la détection décentralisée a cependant fait l'objet d'un nombre plus restreint de travaux. Gupta et Hero [6] ont déterminé la dégradation de performance en termes de détection due à la quantification, dans le cas où les observations sont indépendantes et identiquement distribuées (i.i.d.).

Dans cet article, nous considérons le cas d'observations corrélées et étudions les performances asymptotiques du test NP, lorsque chaque mesure est quantifiée sur $\log_2(N)$ bits. Il est démontré que la probabilité de manque du test NP converge exponentiellement vers zéro, quand le nombre de capteurs tend vers l'infini et le taux de fausse alarme est fixé. Nous généralisons l'idée initiale de [6] au cas non-i.i.d. et fournissons une expression compacte de la dégradation de l'exposant d'erreur due à la quantification dans le régime des hautes résolutions, c-à-d quand le nombre N de niveaux de quantification est grand. Ces résultats étendent ceux de [7] (valables pour des modèles de Markov cachés scalaires) à une large classe de processus stationnaires à valeurs vectorielles. Enfin, à partir de cette expression, nous déterminons des stratégies de quantification appropriées qui permettent de réduire cette dégradation.

2 Détection à partir des observations non-quantifiées

Considérons un processus $(Y_k)_{k \in \mathbb{Z}}$, stationnaire ergodique à valeurs dans un sous-ensemble convexe borné \mathcal{Y} de \mathbb{R}^d , de loi \mathbb{P}_0 (respectivement \mathbb{P}_1) sous l'hypothèse H0 (respectivement H1). On note \mathbb{E}_0 et \mathbb{E}_1 les espérances associées à \mathbb{P}_0 et \mathbb{P}_1 , respectivement. Pour tout $i \in \{0, 1\}$, et tout $n \geq 1$, on suppose que la mesure $\mathbb{P}_i[Y_{1:n} \in \cdot]$ admet une densité p_i par rapport à la mesure de Lebesgue, et que $p_i(y_{1:n}) > 0$ pour tout $y_{1:n} \in \mathcal{Y}^n$. On suppose de plus que pour tout $i \in \{0, 1\}$, $(\log p_i(Y_0|Y_{-m:-1}))_{m \geq 0}$ est une suite convergente de $L^1(\mathbb{P}_0)$.

Nous nous intéressons au problème de la détection de H1 contre H0 à partir d'un ensemble de n observations parfaites (non-quantifiées) $Y_{1:n} = (Y_1, \dots, Y_n)$. D'après le lemme de Neyman-Pearson, un test uniformément plus puissant est obtenu en rejetant l'hypothèse H0 quand la log-vraisemblance $L_n = \log \frac{p_1(Y_{1:n})}{p_0(Y_{1:n})}$ est supérieure à un certain seuil.

Pour tout $\alpha \in (0, 1)$, la probabilité de manque d'un test de niveau α , notée $\beta_n(\alpha)$, est une métrique clé pour en caractériser les performances. Malheureusement, elle ne possède généralement aucune expression simple. Dans la suite, nous étudions le comportement asymptotique de $\beta_n(\alpha)$ lorsque le nombre n d'observations tend vers l'infini. Dans ce régime, on peut montrer que $\beta_n(\alpha) \simeq \exp(-nK)$, où K est une constante appelée *exposant d'erreur* du test. Dans notre cas, nous pouvons démontrer le résultat suivant :

$$K = \lim_{m \rightarrow \infty} \mathbb{E}_0 \left[\log \frac{p_0}{p_1}(Y_0|Y_{-m:-1}) \right].$$

3 Quantification

Soit $N \geq 2$ un entier fixé. Un quantificateur à N niveaux est un triplet $(\mathcal{C}_N, \Xi_N, \xi_N)$ où $\mathcal{C}_N = \{C_{N,1}, \dots, C_{N,N}\}$ est un ensemble de N cellules (boréliens de volume non nul) qui forment une partition de \mathcal{Y} , où $\Xi_N = \{\xi_{N,1}, \dots, \xi_{N,N}\}$ est un ensemble arbitraire d'éléments distincts, et où $\xi_N : \mathcal{Y} \rightarrow \Xi_N$ est une fonction telle que $\xi_N(y) = \xi_{N,j}$ si $y \in C_{N,j}$.

Pour tous N, k , on note $Z_{N,k} = \xi_N(Y_k)$ la k -ème observation quantifiée sur $\log_2(N)$ bits. Dans notre modèle, toutes les observations sont quantifiées avec la même règle (les capteurs sont tous identiques). Bien sûr, le quantificateur $(\mathcal{C}_N, \Xi_N, \xi_N)$ est connu du CF. Le but est à présent de décider entre les hypothèses H0 et H1 à partir des observations quantifiées $Z_{N,1:n}$.

Pour tout $i \in \{0, 1\}$ et tout ensemble $(\xi_{N,j_1}, \dots, \xi_{N,j_n}) \in \Xi_N^n$, on définit

$$p_{i,N}(\xi_{N,j_1}, \dots, \xi_{N,j_n}) = \frac{1}{V_{N,j_1} \times \dots \times V_{N,j_n}} \mathbb{P}_i[Y_{1:n} \in C_{N,j_1} \times \dots \times C_{N,j_n}],$$

où $V_{N,j} = \int_{C_{N,j}} dy$ représente le volume de la cellule j . Cette quantité $p_{i,N}$ est la densité de probabilité des observations quantifiées $Z_{N,1:n}$ par rapport à la mesure de comptage des $(\xi_{N,j})_j$. De même que précédemment, si, pour tout $i \in \{0, 1\}$, $(\log p_{i,N}(Z_{N,0}|Z_{N,-m:-1}))_{m \geq 0}$ est une suite convergente de $L^1(\mathbb{P}_0)$, alors le test possède un exposant d'erreur K_N donné par

$$K_N = \lim_{m \rightarrow \infty} \mathbb{E}_0 \left[\log \frac{p_{0,N}}{p_{1,N}}(Z_{N,0}|Z_{N,-m:-1}) \right]. \quad (1)$$

Une question naturelle est alors la suivante : Comment le quantificateur choisi $(\mathcal{C}_N, \Xi_N, \xi_N)$ affecte-t-il l'exposant d'erreur K_N ? Malheureusement, l'équation (1) ne permet pas d'évaluer directement cet impact. En suivant l'approche de [5, 6], nous nous plaçons alors dans le régime de quantification *haute résolution* ($N \rightarrow \infty$).

4 Performance des quantificateurs haute résolution

On remarque que, pour tout N , l'exposant d'erreur K_N ne dépend pas du choix particulier de l'alphabet Ξ_N . Nous supposons donc sans perte de généralité que chaque $\xi_{N,j}$ coïncide avec le centre de la cellule $C_{N,j}$. La *densité spécifique* ζ_N et le *profil spécifique* M_N d'un quantificateur sont les fonctions définies sur \mathcal{Y} par :

$$\zeta_N(y) = \frac{1}{NV_{N,j}},$$

$$M_N(y) = \frac{1}{V_{N,j}^{1+2/d}} \int_{C_{N,j}} (y - \xi_{N,j})(y - \xi_{N,j})^T dy,$$

pour tout $y \in C_{N,j}$, ($j = 1, \dots, N$).

Considérons à présent une famille de quantificateurs $(\mathcal{C}_N, \Xi_N, \xi_N)_{N \geq 1}$. On suppose que, lorsque $N \rightarrow \infty$,

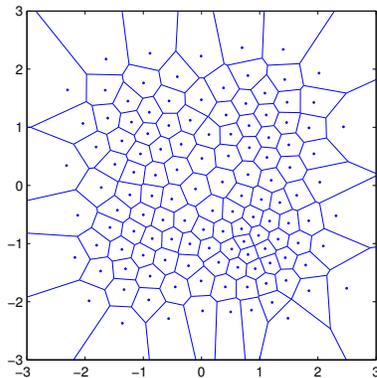


FIGURE 1 – Quantificateur MSE-optimal à 128 cellules ($M = 3, \sigma = 0.6$).

- ζ_N converge uniformément vers une fonction continue ζ (la *densité limite* de la famille $(\mathcal{C}_N, \Xi_N, \xi_N)_{N \geq 1}$) telle que $\inf_{y \in \mathcal{Y}} \zeta(y) > 0$;
- M_N converge uniformément vers une fonction continue (à valeurs matricielles) M (le *profil limite*) telle que $\sup_{y \in \mathcal{Y}} \|M(y)\| < \infty$.

Intuitivement, les quantificateurs haute résolution devraient être construits de manière à ce que $\zeta(y)$ soit grand aux points y pour lesquels une quantification fine est essentielle pour différencier les deux hypothèses. Le Théorème 1 ci-dessous donne une formulation rigoureuse de cette intuition. Ce théorème est valide sous certaines hypothèses (voir [8]) portant sur

- la douceur des densités de probabilité p_i des observations $(Y_k)_k$;
- la vitesse à laquelle les observations passées sont oubliées par les processus $(Y_k)_k$ et $(Z_{N,k})_k$, mesurée via des coefficients de « mélangeance conditionnelle ».

Par exemple, ces hypothèses sont vérifiées pour les processus à faible dépendance (tels que les processus à moyenne mobile), sous réserve que la densité des observations soit assez douce. De même pour une large classe de processus markoviens, sous une hypothèse de douceur sur la densité du noyau de transition. Plus généralement, elles sont valides pour une large classe de modèles de Markov cachés [8].

Théorème 1 ([8]) Lorsque N tend vers l’infini, $N^{2/d}(K - K_N)$ converge vers une constante D_e donnée par :

$$D_e = \frac{1}{2} \int \frac{p_0(y)F(y)}{\zeta(y)^{2/d}} dy,$$

où la fonction F est définie par

$$F(y) = \mathbb{E}_0 \left[\ell(Y_{\mathbb{Z}})^T M(Y_0) \ell(Y_{\mathbb{Z}}) \mid Y_0 = y \right],$$

et la variable aléatoire $\ell(Y_{\mathbb{Z}})$ est la limite dans $L^2(\mathbb{P}_0)$ de la suite $\left(\nabla_{y_0} \log \frac{p_0}{p_1}(Y_{-k:k}) \right)_{k \geq 0}$.

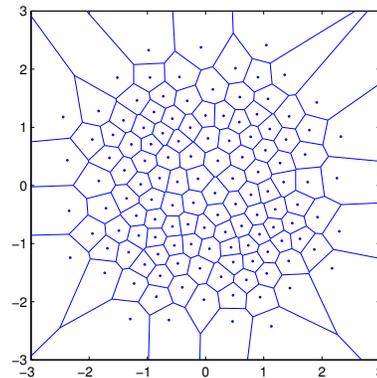


FIGURE 2 – Quantificateur proposé à 128 cellules ($M = 3, \sigma = 0.6$).

La perte asymptotique en exposant d’erreur D_e dépend du quantificateur à travers sa densité limite ζ et son profil limite M . Pour des observations scalaires ($d = 1$), on peut aisément calculer la règle de quantification optimale qui minimise la perte D_e . Dans le cas vectoriel ($d \geq 2$), l’optimisation conjointe de ζ et M n’est pas possible, car on ne connaît pas l’ensemble des fonctions M admissibles comme profils limite [9]. En suivant [6], concentrons-nous sur le cas où $M(y) = v I_d$, pour $v > 0$, où I_d représente la matrice identité. Ce profil limite correspond à des cellules congruentes à moment d’inertie minimal. Dans ce cas, la densité limite ζ qui minimise D_e peut facilement être calculée :

$$\zeta(y) = \frac{[p_0(y)\bar{F}(y)]^{\frac{d}{d+2}}}{\int [p_0(s)\bar{F}(s)]^{\frac{d}{d+2}} ds}, \quad (2)$$

où $\bar{F}(y) = \mathbb{E}_0 \left[\|\ell(Y_{\mathbb{Z}})\|^2 \mid Y_0 = y \right]$.

5 Illustrations

5.1 Détection d’une modulation QPSK/OQPSK

Dans cette section, nous présentons un exemple de modèle de Markov caché qui vérifie les hypothèses du Théorème 1. Il correspond au test entre deux modulations possibles (“quadrature phase-shift keying” QPSK et “offset quadrature phase-shift keying” OQPSK), dans le plan en-phase/quadrature [10, Chapter 3]). Nous considérons donc le modèle suivant pour des observations vectorielles de dimension $d = 2$:

$$Y_k = T(X_k) + W_k,$$

où $W_k \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, \sigma^2)$ représente un bruit gaussien circulaire de moyenne nulle et de variance σ^2 , $(X_k)_{k \in \mathbb{Z}}$ est un message sur 2 bits, à valeurs dans $\mathbf{X} = \{0, 1, 2, 3\}$, et $T(x)$ est la représentation 2-D de l’état x dans le plan en-phase/quadrature.¹

1. $T(0) = [-1; -1]$, $T(1) = [-1; 1]$, $T(2) = [1; 1]$, $T(3) = [1; -1]$.

Les observations sont de plus tronquées sur $Y = [-M; M]^2$ pour $M > 0$. Le processus $(X_k)_{k \in \mathbb{Z}}$ est i.i.d., uniformément distribué sous H_0 , et forme une chaîne de Markov sous H_1 . Plus précisément :

$$H_0 : X_k \stackrel{i.i.d.}{\sim} \mathcal{U}_{\{0,1,2,3\}}$$

$$H_1 : X_0 \sim \mathcal{U}_{\{0,1,2,3\}}, \mathbb{P}_1[X_{k+1} = x' | X_k = x] = q(x, x'),$$

$$\text{où } q = \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{bmatrix}.$$

La Figure 1 représente le quantificateur MSE-optimal (qui minimise l'erreur quadratique moyenne) à 128 cellules obtenu par l'algorithme LBG [11], avec $M = 3$ et $\sigma = 0.6$. La Figure 2 représente le quantificateur proposé correspondant, construit par l'algorithme LBG pour approcher la densité limite (2). Ces deux quantificateurs diffèrent sensiblement. En particulier, certains points de faible probabilité se révèlent significatifs pour le problème de détection considéré.

5.2 Détection d'un processus MA

Considérons à présent le modèle suivant, correspondant à la détection d'un processus à moyenne mobile (MA) :

$$H_0 : Y_k = W_k,$$

$$H_1 : Y_k = \sum_{\ell=0}^L h_\ell U_{k-\ell} + W_k.$$

où $W_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ est un bruit gaussien réel, U_k une source aléatoire transmise à travers un canal de propagation de coefficients (réels déterministes) h_0, \dots, h_L , et L un entier représentant la mémoire de ce canal. Dans la suite de cette section, nous prenons $L = 3$ et supposons que la source U_k est gaussienne : $U_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Nous étudions le cas où une quantification scalaire de chaque observation est réalisée avant le test d'hypothèse.

Notons que dans le modèle ci-dessus les densités de probabilité sont de support infini, de sorte que les hypothèses du Théorème 1 ne sont pas vérifiées. Cependant, de manière à obtenir quelque intuition à propos de la construction pratique de quantificateurs pour la détection, nous pouvons toujours utiliser la méthode décrite à la Section 4 et calculer la densité limite ζ qui minimise la quantité D_e .

Les performances dépendent évidemment de la variance du bruit σ^2 et des valeurs particulières des coefficients du canal. Nous supposons donc que ces coefficients h_0, \dots, h_L sont i.i.d., gaussiens de moyenne nulle et variance 1, et réalisons plusieurs simulations.

En considérant un système à $n = 80$ capteurs, des quantificateurs à $N = 4$ cellules, nous pouvons comparer les performances de différentes règles de quantification à travers leurs

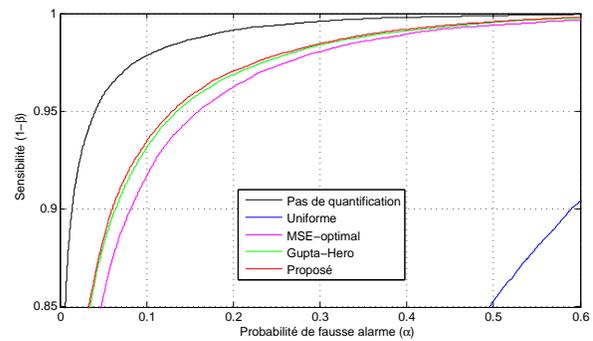


FIGURE 3 – Détection d'un processus MA – Courbes ROC ($h = [1.067, -0.593, 0.096]$, $\sigma = 1.5$, $n = 80$, $N = 4$).

courbes ROC respectives. La Figure 3 représente de telles courbes pour $h = [1.067, -0.593, 0.096]$, et $\sigma = 1.5$. Le quantificateur uniforme est défini sur $[-10\sigma, 10\sigma]$.

La règle de quantification proposée améliore les performances de détection. La courbe ROC est proche de celle obtenue avec le quantificateur décrit par Gupta et Hero [6], mais notons que celui-ci n'est pas défini dans d'autres contextes, où la distribution marginale de Y_k est identique sous H_0 et H_1 (p. ex. celui de la Section 5.1).

Références

- [1] E.L. Lehmann and J.P. Romano. *Testing Statistical Hypotheses (3rd Ed)*. Springer Texts in Statistics, 2005.
- [2] J.-F. Chamberland and V.V. Veeravalli. How dense should a sensor network be for detection with correlated observations? *IEEE Trans. Inf. Theory*, 52(11) :5099–5106, 2006.
- [3] Y. Sung, L. Tong, and H.V. Poor. Neyman–Pearson detection of Gauss–Markov signals in noise : closed-form error exponent and properties. *IEEE Trans. Inf. Theory*, 52(4) :1354–1365, 2006.
- [4] W.R. Bennett. Spectra of quantized signals. *Bell System Technical Journal*, 27 :446–472, 1948.
- [5] S. Na and D.L. Neuhoff. Bennett's integral for vector quantizers. *IEEE Trans. Inf. Theory*, 41(4) :886–900, 1995.
- [6] R. Gupta and A.O. Hero. High-rate vector quantization for detection. *IEEE Trans. Inf. Theory*, 49(8) :1951–1969, 2003.
- [7] J. Villard, P. Bianchi, E. Moulines, and P. Piantanida. High-rate quantization for the Neyman-Pearson detection of hidden Markov processes. In *Proc. ITW*, Cairo, Egypt, 2010.

- [8] J. Villard and P. Bianchi. High-rate vector quantization for the Neyman-Pearson detection of correlated processes. *IEEE Trans. Inf. Theory*, 57(8), 2011.
- [9] R.M. Gray and D.L. Neuhoff. Quantization. *IEEE Trans. Inf. Theory*, 44(6) :2325–2383, 1998.
- [10] J.G. Proakis and M. Salehi. *Digital communications (5th Ed)*. McGraw-Hill, 2007.
- [11] Y. Linde, A. Buzo, and R.M. Gray. An algorithm for vector quantizer design. *IEEE Trans. Commun.*, 28(1) :84–95, 1980.