High-Rate Vector Quantization for the Neyman-Pearson Detection of Some Stationary Mixing Processes

Joffrey Villard* and Pascal Bianchi†

*SUPELEC, Telecom. Dpt., Gif-sur-Yvette, France

†LTCI Telecom Paristech, Paris, France

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

Abstract

This paper investigates the decentralized detection of spatially correlated processes using the Neyman-Pearson test. We consider a network formed by a large number of sensors, each of them observing a random data vector. Sensors' observations are non-independent, but form a stationary process verifying mixing conditions. Each vector-valued observation is quantized before being transmitted to a fusion center which makes the final decision. 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. A compact closed-form expression of the error exponent is provided in the high-rate regime *i.e.*, when fine quantization is applied. As an application, our results allow to determine relevant quantization strategies which lead to large error exponents.

I. Introduction

Consider a Wireless Sensor Network (WSN) whose aim is to detect the presence of a stochastic signal, based on a large number of sensors. We assume that a fusion center (FC) gathers information from the sensors and takes the final decision. Binary hypothesis tests are special cases where the FC has to decide between two possible hypotheses H0 and H1. In this case, the celebrated Neyman-Pearson (NP) procedure provides a uniformly most powerful test [1].

In the context of sensor networks, a large number of works have been devoted to the study of the performance of the NP test, with the aim to design WSN with attractive detection capabilities. Most of these works address the case where observations are independent random variables. However, the detection of a spatially correlated process is a crucial issue in WSN applications. In this case, fewer results are available in the literature. Chamberland and Veeravalli [2] analyse the impact of the density of sensors on

The work of J. Villard is supported by DGA (French Armement Procurement Agency).

the detection performance, when observations are correlated. In case of the detection of a Gauss-Markov signal in noise, Sung, Tong and Poor [3] prove that for a fixed false alarm level, the miss probability of the NP test converges exponentially to zero, and provide a closed form expression of the error exponent.

It is worth noting that in the above works, the FC is assumed to have a perfect knowledge of the sensors' measurements. In practice, the presence of imperfect wireless links between nodes requires to compress/quantize data before transmission to the FC. In the past decades, numerous papers were dedicated to the search for relevant quantization strategies for various target applications. Bennett [4] pioneered the study of *high-rate* (or *high-resolution*) quantization for the reconstruction of scalar signals. Extension of the works of Bennett to vector-valued observations was later achieved in [5]. On the otherhand, high-rate quantization in the framework of statistical tests and decentralized detection was subject to much fewer works. Gupta and Hero [6] determined the degradation of the detection performance due to quantization, in the particular case where observations are independent and identically distributed (i.i.d.). To our knowledge, a comprehensive analysis remains to be provided in the case of correlated observations.

In this paper, we investigate the case where sensors' measurements are non-independent. We study the asymptotic performance of the Neyman-Pearson detector when the number of sensors tends to infinity, assuming that each measurement is quantized on $\log_2(N)$ bits. For a fixed 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. Generalizing the initial idea of [6] to the non-i.i.d. case, we provide a compact expression of the error exponent in the case of high-rate quantization *i.e.*, when the number N of quantization levels is large. This paper extends our previous work [7] (which was specific to scalar hidden Markov models) to the case of an *arbitrary* process distribution and to *vector*-valued observations.

The paper is organized as follows. In Section II, we describe the observation model. We also recall some results on NP tests and evaluate the error exponent in the ideal case where the FC has perfect access to the measurements. Fixed-rate vector quantizers are defined in Section III. Next, we evaluate the error exponent in the case when the decision is made using quantized measurements. In Section IV, we evaluate the degradation of the error exponent due to quantization in the high-rate regime. We determine relevant quantization strategies allowing to reduce this degradation. Section V is devoted to numerical illustrations.

II. NP Detection with Perfect Observations

A. Observation Model

Consider two probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$ on a relevant probability space. Denote by $(Y_k)_{k\in\mathbb{Z}}$ a stationary ergodic process for both $\mathbb{P}_0$ and $\mathbb{P}_1$, taking its values in a bounded convex subset Y of $\mathbb{R}^d$. We associate an hypothesis (H0 and H1 respectively) to each of the two probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$ and investigate the problem of the detection of H1 *vs.* H0 based on a set of n observations $Y_{1:n} = (Y_1, \ldots, Y_n)$.

For convenience, we assume that $\mathbb{P}_0$ and $\mathbb{P}_1$ are probability distributions of process $(Y_k)_{k\in\mathbb{Z}}$ on the space $(\mathsf{Y}^{\mathbb{Z}}, B(\mathsf{Y}^{\mathbb{Z}}))$, where $B(\mathsf{Y}^{\mathbb{Z}})$ represents the Borel σ-field in $\mathsf{Y}^{\mathbb{Z}}$. Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be three sub σ-algebras of $B(\mathsf{Y}^{\mathbb{Z}})$. Define the following conditional ψ-mixing coefficient for each $i \in \{0,1\}$:

$$\psi_i(\mathcal{U},\mathcal{V}|\mathcal{W}) = \sup_{U\in\mathcal{U},\, V\in\mathcal{V}} \operatorname{ess\,sup} \left|1 - \frac{\mathbb{P}_i(U\cap V|\mathcal{W})}{\mathbb{P}_i(U|\mathcal{W})\mathbb{P}_i(V|\mathcal{W})}\right|$$

where the essential supremum is taken w.r.t. $\mathbb{P}_0$ and where we use the convention $0/0 = 1$. The above coefficient can be interpreted as a measure of independence between $\mathcal{U}$ and $\mathcal{V}$ conditionally to $\mathcal{W}$. In particular, it coincides with the traditional ψ-mixing coefficient $\psi(\mathcal{U},\mathcal{V})$ when $\mathcal{W}$ is taken to be the whole space $B(\mathsf{Y}^{\mathbb{Z}})$ [8]. For each $n \geq 1$, we define:

$$\psi_{i,n} = \sup_{m>0} \psi_i(\sigma(Y_{n+1}), \sigma(Y_{-m:0})|\sigma(Y_{1:n}))$$

and $\psi_{i,0} = \sup_{m>0} \psi_i(\sigma(Y_1), \sigma(Y_{-m:0}))$ when $n = 0$. In the sequel, we will generally assume that sequence $\psi_{i,n}$ converges to zero as $n \to \infty$. Loosely speaking, this means that for large values of n, the present observation Y_{n+1} is "nearly" independent of past observations $Y_{-m:0}$ conditionally to $Y_{1:n}$.

We denote by $\mathbb{E}_0$ and $\mathbb{E}_1$ the expectations associated with $\mathbb{P}_0$ and $\mathbb{P}_1$ respectively, and introduce the measure μ which coincides with the d-dimensional Lebesgue measure on Y.

Assumption 1. *For each $i \in \{0,1\}$, for each $n \geq 1$, measure $\mathbb{P}_i[Y_{1:n} \in \cdot\,]$ admits a density p_i w.r.t. $\mu^{\otimes n}$. Moreover, $p_i(y_{1:n}) > 0$ for each $y_{1:n} \in \mathsf{Y}^n$.*

In particular, Assumption 1 implies that the probability distribution of the observations $Y_{1:n}$ under hypotheses H0 and H1 respectively are absolutely continuous w.r.t. each other.

B. Likelihood Ratio Test

We now investigate the detection of H1 *vs.* H0 based on the **perfect** observation of n measurements $Y_{1:n}$. Due to the celebrated Neyman-Pearson's Lemma, a uniformly most powerful test is obtained by rejecting the null hypothesis when the following log-likelihood ratio (LLR)

$$L_n = \log \frac{p_1(Y_{1:n})}{p_0(Y_{1:n})}$$

is larger than a threshold, say γ. For each $\alpha \in (0,1)$, we define the miss probability of the Neyman-Pearson test of level α by:

$$\beta_n(\alpha) = \inf\{\mathbb{P}_1\left[L_n < \gamma\right] \; : \gamma \text{ s.t. } \mathbb{P}_0\left[L_n > \gamma\right] \leq \alpha\} \; .$$

Quantity $\beta_n(\alpha)$ is a key metric for characterizing the performance of the hypothesis test. Unfortunately, it usually does not admit a tractable closed form expression. In the sequel, we study the asymptotic behaviour of $\beta_n(\alpha)$ as the number of observations n tends to infinity. In this regime, it can be shown that $\beta_n(\alpha) \simeq \exp(-n\,K)$ for some constant K given below, which we shall refer to as the error exponent.

C. Error Exponent with Perfect Observations

The evaluation of the error exponent K fundamentaly relies on the following Lemma:

Lemma 1 ([9])**.** *Assume that* $(-1/n)L_n$ *converges in probability under* $\mathsf{H0}$ *to a deterministic constant* κ *s.t.* $0 < \kappa \leq \infty$*. Then, for any* $\alpha \in (0,1)$*,* $(-1/n)\log\beta_n(\alpha) \to \kappa$ *as* $n \to \infty$*.*

Therefore, the proof of existence of the error exponent and its evaluation reduce to the asymptotic study of the LLR. The following result is obtained by direct extension of the Shannon-McMillan-Breiman's Theorem [10].

Theorem 1. *Assume that* $(\psi_{1,n})_{n\geq 0}$ *converges to zero as* $n \to \infty$*. Assume that there exists an integer* $n_0 \geq 0$ *such that* $\mathbb{E}_0|\log p_1(Y_{n_0+1}|Y_{1:n_0})| < \infty$ *and* $\psi_{1,n_0} < 1$*. Then,*

$$\lim_{n\to\infty} \frac{1}{n} \log\beta_n(\alpha) = -K \ ,$$

where K *is the constant defined by*

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

Proof: We can prove the following inequality, for any fixed $k \geq 0$ and for each $m' \geq m$:

$$\left|\log p_1(Y_k|Y_{-m:k-1}) - \log p_1(Y_k|Y_{-m':k-1})\right| \leq -\log(1-\psi_{1,k+m}) \ . \tag{2}$$

Thus $(\log p_1(Y_k|Y_{-m:k-1}))_{m\geq -k}$ is a Cauchy sequence. Denote its limit by $\mathcal{L}_1(Y_{-\infty:k})$. Letting m' tend to infinity in (2) and using the triangular inequality, we obtain:

$$\left|\frac{1}{n}\sum_{k=1}^{n} \log p_1(Y_k|Y_{1:k-1}) - \frac{1}{n}\sum_{k=1}^{n} \mathcal{L}_1(Y_{-\infty:k})\right| \xrightarrow[n\to\infty]{\mathbb{P}_0\text{-a.s.}} 0 \ .$$

Under above assumptions, Equation (2) ensures that $\mathbb{E}_0|\mathcal{L}_1(Y_{-\infty:0})| < \infty$. As process $(Y_k)_{k\in\mathbb{Z}}$ is stationary ergodic under $\mathbb{P}_0$, the ergodic theorem yields:

$$\frac{1}{n}\sum_{k=1}^{n} \log p_1(Y_k|Y_{1:k-1}) \xrightarrow[n\to\infty]{\mathbb{P}_0\text{-a.s.}} \mathbb{E}_0\left[\mathcal{L}_1(Y_{-\infty:0})\right] \ .$$

On the otherhand, the Shannon-McMillan-Breiman's Theorem [10] implies that:

$$\frac{1}{n}\sum_{k=1}^{n} \log p_0(Y_k|Y_{1:k-1}) \xrightarrow[n\to\infty]{\mathbb{P}_0\text{-a.s.}} \lim_{m\to\infty} \mathbb{E}_0\left[\log p_0(Y_0|Y_{-m:-1})\right] \ .$$

This proves Theorem 1. ∎

III. Quantization

A. Definitions

Consider a fixed integer $N \geq 2$. An N-point quantizer is a triplet $(\mathcal{C}_N, \Xi_N, \xi_N)$ where $\mathcal{C}_N = \{C_{N,1}, \ldots, C_{N,N}\}$ is a set of N *cells* (Borel sets of Y with non-zero volume) which form a partition of Y, where $\Xi_N = \{\xi_{N,1}, \ldots, \xi_{N,N}\}$ is an arbitrary set of distinct elements and where $\xi_N : \mathsf{Y} \to \Xi_N$ is a function s.t. $\xi_N(y) = \xi_{N,j}$ whenever $y \in C_{N,j}$.

For each N, k, denote by $Z_{N,k} = \xi_N(Y_k)$ the quantized measurement on $\log_2(N)$ bits. Note that in our model, all measurements are quantized using the same quantization rule. Moreover, we assume that the quantizer $(\mathcal{C}_N, \Xi_N, \xi_N)$ is known at the decision device. The aim is to decide between hypotheses H0 and H1 based on the observation of $Z_{N,1:n}$.

B. Error Exponent

We define the LLR based on n quantized measurements by:

$$L_{n,N} = \log \frac{p_{1,N}\left(Z_{N,1:n}\right)}{p_{0,N}\left(Z_{N,1:n}\right)} \ ,$$

where for each $i \in \{0,1\}$ and for any set of quantization points $(\xi_{N,j_1}, \ldots, \xi_{N,j_n}) \in \Xi_N^n$,

$$p_{i,N}(\xi_{N,j_1}, \ldots, \xi_{N,j_n}) = \frac{\mathbb{P}_i[Y_{1:n} \in C_{N,j_1} \times \ldots \times C_{N,j_n}]}{V_{N,j_1} \times \ldots \times V_{N,j_n}} \ , \tag{3}$$

where $V_{N,j} = \int_{C_{N,j}} dy$ represents the volume of cell j. Function $p_{i,N}$ is the pdf of the observations $Z_{N,1:n}$ w.r.t. the (n-product) weighted counting measure of the points $\xi_{N,j}$'s.

For each $\alpha \in (0,1)$, we denote by $\beta_{n,N}(\alpha)$ the miss probability of the NP test of level α when quantization is applied *i.e.*, the infimum of $\mathbb{P}_1\left[L_{n,N} < \gamma\right]$ w.r.t. all γ s.t. $\mathbb{P}_0\left[L_{n,N} > \gamma\right] \leq \alpha$. We define for each $i \in \{0,1\}$, $n \geq 1$:

$$\bar{\psi}_{i,n}(\mathcal{C}_N) = \sup_{m>0} \ \psi_i(\sigma(Z_{N,n+1}), \sigma(Z_{N,-m:0}) | \sigma(Z_{N,1:n})) \ .$$

Corollary 1. *Assume that* $\bar{\psi}_{i,n}(\mathcal{C}_N) \to 0$ *as* $n \to \infty$. *Assume that there exists* $n_0 \geq 0$ *s.t.* $\mathbb{E}_0 |\log p_1(Z_{N,n_0+1}|Z_{N,1:n_0})| < \infty$ *and* $\bar{\psi}_{1,n_0}(\mathcal{C}_N) < 1$. *As* $n \to \infty$, $(-1/n)\log\beta_{n,N}(\alpha)$ *converges to the error exponent* K_N *given by:*

$$K_N = \lim_{m\to\infty} \mathbb{E}_0 \left[\log \frac{p_{0,N}}{p_{1,N}} (Z_{N,0}|Z_{N,-m:-1}) \right] \ . \tag{4}$$

A natural question is: How does the quantizer affect the error exponent? Unfortunately, Equation (4) is not informative enough to directly evaluate the impact of the quantizer. Following [5], [6], we thus study the *high-rate* quantization regime *i.e.*, the case where N tends to infinity.

IV. Performance of High-Rate Vector Quantizers

A. Notations and Assumptions

For each N, the error exponent K_N does not depend on the particular choice of the quantization alphabet Ξ_N. We then assume with no loss of generality that each $\xi_{N,j}$ coincides with the center of cell $C_{N,j}$. We introduce the *specific point density* ζ_N and the *specific covariation profile* M_N as the piecewise constant functions on Y respectively given by

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

whenever $y \in C_{N,j}$, $(j = 1, \ldots, N)$. Now consider a family of quantizers $(\mathcal{C}_N, \Xi_N, \xi_N)_{N\geq 1}$.

Assumption 2. *As* $N \to \infty$, ζ_N *converges uniformly to a continuous function* ζ *such that* $\inf_{y\in\mathsf{Y}} \zeta(y) > 0$. M_N *converges uniformly to a continuous (matrix-valued) function* M *such that* $\sup_{y\in\mathsf{Y}} \|M(y)\| < \infty$.

We will refer to ζ as the *model point density* of the family $(\mathcal{C}_N, \Xi_N, \xi_N)_{N\geq 1}$. It represents the fraction of cells in the neighborhood of a given point y. Function M will be refered to as the *model covariation profile*. It provides information about the shape of the cells. Intuitively, high-rate quantizers should be constructed in such a way that $\zeta(y)$ is large at those points y for which a fine quantization is essential to discriminate the two hypotheses. Theorem 2 below provides a more rigorous formulation of this intuition. We need further assumptions. For each $j \in \{1, \ldots, N\}$, denote by $d_{N,j} = \sup_{u,v\in C_{N,j}} \|u - v\|$ the diameter of cell j.

Assumption 3. *The following properties hold true.*

1) *For all N, $\sup_j d_{N,j} \leq C_d/(N^{1/d})$ for some $C_d > 0$.*
2) *For any $n \geq 1$, $y_{1:n} \mapsto p_i(y_{1:n})$ is of class C_3 on Y^n.*
3) $$\sup_{n\geq 1,\ 1\leq k,l,r\leq n,\ 1\leq h,i,j\leq d} \left\| \frac{\partial^3 \log p_i}{\partial y_k^{(h)}\, \partial y_\ell^{(i)}\, \partial y_r^{(j)}} \right\|_\infty < \infty \ .$$
4) *There exist constants C_e, $\epsilon > 0$ and an integer $n_0 \geq 0$ s.t. for each $i \in \{0, 1\}$, each $N \geq 2$ and each $n \geq n_0$,*
$$\max\left(\psi_{i,n},\ \bar{\psi}_{i,n}(\mathcal{C}_N)\right) \leq C_e/(n^{6+\epsilon}) \ .$$
5) *For each $i \in \{0, 1\}$ and each $-m' \leq -m \leq 0 \leq k$:*
$$\|\nabla_{y_0} \log p_i(Y_{0:k}|Y_{-m:-1}) - \nabla_{y_0} \log p_i(Y_{0:k}|Y_{-m':-1})\| \leq \varphi_m$$
$$\|\nabla_{y_0} \log p_i(Y_k|Y_{-m:k-1})\| \leq \rho_k \ ,$$
where $\sum_k \varphi_k$ and $\sum_k \rho_k$ are convergent series.

Assumption 3-4) can be interpreted as a condition on the speed at which past observations are forgotten. Assumption 3-5) can be interpreted similarly as a forgetting property which involves the derivative of the log-density of the observations.

For instance, Assumption 3 is simple to verify in case of short-dependency processes. A similar remark holds for a wide class of Markov chains. More generally, we prove in an extended version of this paper [11] that Assumption 3 holds for a wide class of hidden Markov models.

B. Error Exponent in the High-Rate Regime

Theorem 2 below states that when the order of the quantizer tends to infinity, the error exponent K_N associated with the NP test in the presence of quantization (4) converges at speed $N^{-2/d}$ to the error exponent K that one would have obtained in the absence of quantization (1). Loosely speaking, the approximation
$$\beta_{n,N}(\alpha) \simeq e^{-n\left(K - \frac{D_e}{N^{2/d}}\right)}$$

is valid when both the number n of sensors and the order N of quantization are large, but $n \gg N$. Quantity D_e represents the (normalized) loss in error exponent between the quantized and the unquantized cases, in the high-rate quantization regime.

Theorem 2. *Under Assumptions 1–3:*

1) *The following limit exists with probability one under $\mathbb{P}_0$:*

$$\ell(Y_{\mathbb{Z}}) := \lim_{k\to\infty} \lim_{m\to\infty} \nabla_{y_0} \log \frac{p_0}{p_1}(Y_{-m:k}) \ .$$

Moreover, $\|\ell(Y_{\mathbb{Z}})\| < C$ for some constant C.

2) *As N tends to infinity, $N^{2/d}(K - K_N)$ converges to a constant D_e given by*

$$D_e = \frac{1}{2} \int \frac{p_0(y)F(y)}{\zeta(y)^{2/d}}\, dy \ , \tag{5}$$

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

The particular situation where measurements $(Y_k)_{k\geq 0}$ are i.i.d. under both hypotheses was studied by Gupta and Hero [6]. Expression (5) of D_e is clearly consistent with the one of [6] in the i.i.d. case.

C. Short Sketch of the Proof

Due to the lack of space, we only provide some of the basic ideas underlying the proof of Theorem 2. A rigorous proof will be provided in an extended version of this paper [11].

It is straightforward to prove a counterpart of Equation (2) for density $p_{i,N}$ involving coefficients $\bar{\psi}_{i,n}$. Consequently, under Assumption 3-4), $(\log p_{i,N}(Z_{N,0}|Z_{N,-m,-1}))_{m\geq 0}$ is a Cauchy sequence. We denote its limit by $\mathcal{L}_{i,N}(Z_{N,-\infty:0})$. The error exponent associated with the NP test on quantized observations is then given by $K_N = K_{0,N} - K_{1,N}$ where for each $i \in \{0, 1\}$, $K_{i,N} = \mathbb{E}_0 [\mathcal{L}_{i,N}(Z_{N,-\infty:0})]$. First focus on $K_{1,N}$ and choose a sequence $m = m(N)$ of integers such that $\frac{N^{2/d}}{m^{6+\epsilon}} \to 0$ and $\frac{m^3}{N^{1/d}} \to 0$ as $N \to \infty$. We write

$$\begin{aligned} K_{1,N} - K_1 &= \mathbb{E}_0 [\mathcal{L}_{i,N}(Z_{N,-\infty:0}) - \mathcal{L}_1(Y_{-\infty:0})] \\ &= T_N + U_N + \delta_N \ , \end{aligned}$$

where

$$\begin{aligned} T_N &= \mathbb{E}_0 [\log p_{1,N}(Z_{N,0}|Z_{N,-m:-1}) - \log p_1(Z_{N,0}|Z_{N,-m:-1})] \\ U_N &= \mathbb{E}_0 [\log p_1(Z_{N,0}|Z_{N,-m:-1}) - \log p_1(Y_0|Y_{-m:-1})] \ . \end{aligned}$$

Under Assumption 3, from properties of the sequence $m(N)$, the remainder δ_N is a little-o of $N^{-2/d}$. The study of T_N is based on the following expansion:

$$T_N = \mathbb{E}_0 \left[\log \frac{p_{1,N}(Z_{N,-m:0})}{p_1(Z_{N,-m:0})} \right] - \mathbb{E}_0 \left[\log \frac{p_{1,N}(Z_{N,-m:-1})}{p_1(Z_{N,-m:-1})} \right] \ . \tag{6}$$

Plugging the Taylor-Lagrange expansion of function $y_{-m:u} \mapsto p_1(y_{-m:u})$ at point $\xi_N(y_{-m:u})$ in (3) leads to the following approximate:

$$\frac{p_{1,N}(\xi_{N,j_{-m:u}})}{p_1(\xi_{N,j_{-m:u}})} \approx 1 + \frac{1}{2N^{2/d}} \sum_{k=-m}^{u} \mathrm{Tr}\left(\frac{\nabla^2_{y_k} p_1(\xi_{N,j_{-m:u}})^{\mathrm{T}}}{p_1(\xi_{N,j_{-m:u}})} \frac{M_{N,j_k}}{\zeta^{2/d}_{N,j_k}} \right) .$$

We now plug the above expansion into (6) and approximate T_N by a series involving the second term of the rhs of the above equation. The study of U_N is also based on the expansion of function $y_{-m:u} \mapsto p_1(y_{-m:u})$ at point $\xi_N(y_{-m:u})$ and leads to a similar approximation. Bounding properly the remainders and using the fundamental mixing conditions of Assumption 3-5), it can be shown after tedious derivations that $N^{2/d}T_N$ and $N^{2/d}U_N$ respectively converge to some constants c_T and c_U as $N \to \infty$ which can be determined. Thus, $N^{2/d}(K_{1,N} - K_1)$ converges to the sum $c_T + c_U$. Proceeding in the same way for the study of $N^{2/d}(K_{0,N} - K_0)$, we prove Theorem 2.

D. Application: Determination of Relevant Quantizers

The asymptotic loss in error exponent D_e depends on the quantizer through its model point density ζ and its model covariation profile M. For scalar measurements, we can derive the optimal quantization rule, which minimizes the loss D_e. This derivation is more difficult in the vector case.

1) Vector case ($d \geq 2$): We first address the case where measurements $(Y_k)_{k\geq 0}$ are vector-valued. The determination of optimal high-rate quantization rules implies the joint minimization of expression (5) w.r.t. both functions ζ and M. Unfortunately, as remarked in [12], it is not known what functions M are allowable as covariation profiles. The determination of the set of admissible couples (ζ, M) is an open problem, which is beyond the scope of this paper.

However, when M is fixed, the point density ζ which minimizes D_e can be easily expressed as a function of M. Using Hölder's inequality on (5), it is straightforward to prove that $D_e \geq \frac{1}{2} \left(\int [p_0(y)F(y)]^{\frac{d}{d+2}} \, dy \right)^{\frac{d+2}{d}}$, where equality is achieved when the point density coincides with:

$$\zeta(y) = \frac{[p_0(y)F(y)]^{\frac{d}{d+2}}}{\int [p_0(s)F(s)]^{\frac{d}{d+2}} \, ds} . \tag{7}$$

In other words, one can easily provide the optimal high-rate quantization rule for a given limiting covariation profile.

Following the approach of [6], let us now focus on the case where $M(y) = \upsilon \, I_d$, for some $\upsilon > 0$, where I_d represents the $d \times d$ identity matrix. The reason for investigating this case is essentially practical. Indeed, practical construction of quantizers is usually achieved by means of well-established algorithms, the most popular of them being the Linde-Buzo-Gray (LBG) algorithm [13] which computes a (nearly) MSE-optimal N-point quantizer for a given pdf. Gersho [14] made the now widely accepted conjecture that when N is large, most cells of a d-dimensional MSE-optimal quantizer are approximately congruent to some tessellating polytope H_d^*. In such a case, $M(y) = \upsilon_d^* \, I$ for all $y \in \mathsf{Y}$, where υ_d^* is the inertia

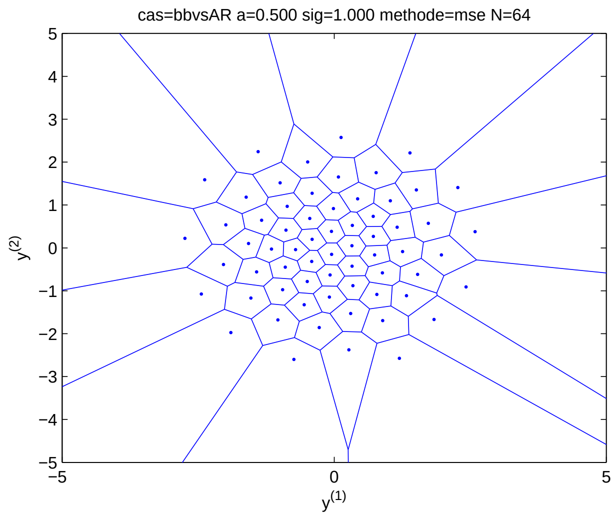

Fig. 1. MSE-optimal 64-cell quantizer ($\sigma = 1$)

of H_d^*. For such quantizers, the optimal point density (7) becomes:

$$\zeta(y) = \frac{\left[p_0(y)\bar{F}(y)\right]^{\frac{d}{d+2}}}{\int \left[p_0(s)\bar{F}(s)\right]^{\frac{d}{d+2}} ds} \ , \tag{8}$$

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

2) Scalar case ($d = 1$): We now study the case of real-valued observations. Assume without much loss of generality that each cell is connected (cells are intervals). In this case, a straightforward derivation leads to $M_N(y) = 1/12$ for each y and each N. Therefore, function F simplifies to:

$$F(y) = \frac{1}{12}\,\mathbb{E}_0\left[\ell(Y_{\mathbb{Z}})^2 \,\middle|\, Y_0 = y\right]$$

and Equation (7) (with $d = 1$) now provides the optimal high-rate quantization rule. Note that expression (7) is quite similar to Bennett's formula [4] which gives the optimal model point density in an MSE perspective.

V. Illustration

Consider the following model for each k:

$$Y_k = X_k + W_k \ , \tag{9}$$

where $W_k \overset{i.i.d}{\sim} \mathcal{CN}(0, \sigma^2)$ represents a zero mean complex circular Gaussian thermal noise with variance σ^2, and where $(X_k)_{k\in\mathbb{Z}}$ is a Gaussian process which is white under H0 and correlated (AR-1) under H1. More precisely,

$$\begin{aligned} \mathsf{H0}:&\ \ X_k \overset{i.i.d}{\sim} \mathcal{CN}(0,1) \\ \mathsf{H1}:&\ \ X_k = aX_{k-1} + \sqrt{1-a^2}\,U_k \ , \end{aligned} \tag{10}$$

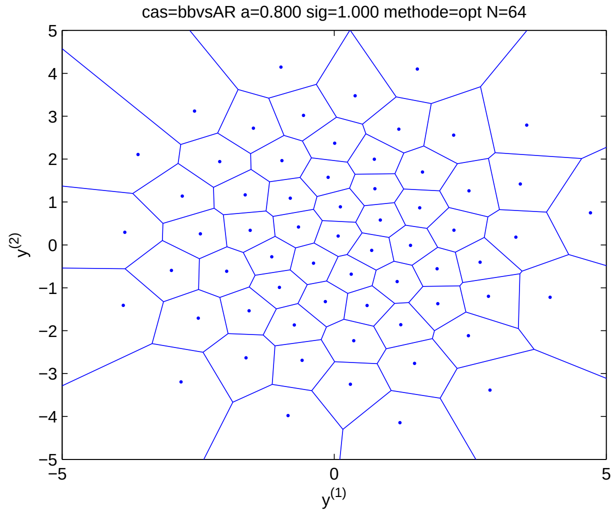

Fig. 2. Optimal 64-cell quantizer ($a = 0.8$, $\sigma = 1$)

where $a \in (0, 1)$ is the correlation coefficient and $U_k \stackrel{i.i.d}{\sim} \mathcal{CN}(0, 1)$ is the innovation process. In particular, $(Y_k)_{k\in\mathbb{Z}}$ is a white Gaussian process under H0 and is a hidden Markov process under H1, with the particular property that marginal distribution of single observations are indentical under both hypotheses. We mention that in the above model, densities have infinite support so that, strictly speaking, the assumptions made in this paper are not satisfied. Nevertheless, the above model can be slightly modified to be consistent with our assumptions. For instance, in order that the observations lie in a bounded subset of $\mathbb{R}^2$, it is sufficient to replace the distribution $\mathcal{CN}(0, 1)$ of U_k with the corresponding truncated distribution on an arbitrarily large support. In order to simplify the presentation, we do not go into details and keep model (9)-(10) with slight abuse.

As the marginal pdf does not depend on the hypothesis, it turns out that Gupta and Hero's quantizer [6], which minimizes the error exponent loss in case of i.i.d. observations, is not defined. We compare our quantization rule to the traditional MSE-optimal quantizer. Figure 1 represents the 64-cells quantizer computed by the LBG algorithm for pdf $p = p_0 = p_1$ with parameter $\sigma = 1$. Figure 2 represents the optimal (8) 64-cells quantizer for $a = 0.8$ and $\sigma = 1$. This optimal quantizer significantly differs from the MSE-optimal one. Low probability points turn out to be significant for the considered detection problem.

More examples will be provided in an extended version of this paper [11].

REFERENCES

[1] E. Lehmann and J. Romano, *Testing Statistical Hypotheses (3rd Ed).* Springer Texts in Statistics, 2005.

[2] J.-F. Chamberland and V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5099–5106, 2006.

[3] Y. Sung, L. Tong, and H. Poor, "Neyman–Pearson detection of Gauss–Markov signals in noise : closed-form error exponent and properties," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1354–1365, 2006.

[4] W. Bennett, "Spectra of quantized signals," *Bell System Technical Journal*, vol. 27, pp. 446–472, 1948.

[5] S. Na and D. Neuhoff, "Bennett's integral for vector quantizers," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 886–900, 1995.

[6] R. Gupta and A. Hero, "High-rate vector quantization for detection," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 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] R. Bradley, "Basic properties of strong mixing conditions. a survey and some open questions," *Probability Surveys*, vol. 2, pp. 107–144, 2005.

[9] P.-N. Chen, "General formulas for the Neyman–Pearson type-II error exponent subject to fixed and exponential type-I error bounds," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 316–323, 1996.

[10] A. Barron, "The strong ergodic theorem for densities: Generalized Shannon–McMillan–Breiman theorem," *The Annals of Probability*, vol. 13, no. 4, pp. 1292–1303, 1985.

[11] J. Villard and P. Bianchi, "High-rate vector quantization for the Neyman-Pearson detection of correlated processes," *Submitted to IEEE Trans. Inf. Theory*, 2010.

[12] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[13] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, 1980.

[14] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inf. Theory*, vol. 25, no. 4, pp. 373–380, 1979.