Error Bounds for Parallel Distributed Detection under the Neyman–Pearson Criterion

Po-Ning Chen, *Member, IEEE*, and Adrian Papamarcou, *Member, IEEE*

***Abstract*—The Neyman–Pearson performance of a distributed detection system is considered wherein n independent and identically distributed observations are quantized locally into M-ary messages and transmitted to a fusion center. Under fairly general assumptions, it is shown that the type II error probability achieved by the best identical-quantizer system is at most a fixed (in n) multiple of that achieved by the absolutely optimal system.**

***Index Terms*— Distributed detection, quantization, error exponents, large deviations, asymptotic expansions.**

I. Introduction

Consider the distributed binary detection system $\mathcal{S}_n$ depicted in Fig. 1. The sensor observations $Y_1, \ldots, Y_n$ are assumed independent and identically distributed given each of the two hypotheses H_0 (signal absent) and H_1 (signal present). Each observation Y_i is locally quantized into an M-ary message $U_i = g_i(Y_i)$, which is then transmitted to the fusion center. The fusion center output is the global decision $\mathcal{D}(U_1, \ldots, U_n)$.

The Neyman–Pearson optimization problem for the above system is formulated as follows: choose $g_1, \ldots, g_n$ and $\mathcal{D}$ so as to minimize the type II error probability $\beta_n(\alpha)$ subject to an upper bound α on the type I error probability. This is a difficult computational task which can be simplified considerably by introducing the additional constraint that all local quantizers g_i be identical. Under fairly general conditions, it can be shown [8] that this identical-quantizer constraint does not affect the optimal *error exponent*. In symbols, if superscripts "$*$" and "$\diamond$" denote the *optimal* and *best identical-quantizer* solutions, respectively, then it is generally true that

$$\lim_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\alpha) = \lim_{n\to\infty} -\frac{1}{n}\log\beta_n^\diamond(\alpha)$$

provided $\alpha \in (0, 1)$ is fixed in n.

The aim of this correspondence is to establish a much stronger result, namely that under general conditions, the ratio $\beta_n^*(\alpha)/\beta_n^\diamond(\alpha)$ is lower bounded by a constant independent of n (it is, of course, upper bounded by unity). A nearly identical statement was proved for the Bayes optimization problem in a recent paper [4]. Despite the similarity between the statements of the Neyman–Pearson and Bayes results, the structure of the underlying proofs is altogether different.

This brief introduction should suffice for this correspondence; for a fuller background discussion, the reader is referred to [4, Sections I and II]. Otherwise, the material presented herein is entirely self-contained and independent of [4].

Manuscript received October 27, 1993; revised September 15, 1994. The material in this correspondence was presented in part at the 27th Annual Conference on Information Sciences and Systems, The Johns Hopkins University, Baltimore, MD, Mar. 24–27 1993. This research was supported by the Institute for Systems Research (a National Science Foundation Engineering Research Center) at the University of Maryland, College Park.

P.-N. Chen was with the Electrical Engineering Department, University of Maryland, College Park, MD 20742 USA. He is now with Wan Ta Technologies Inc., Taipei, Taiwan, ROC.

A. Papamarcou is with the Electrical Engineering Department, University of Maryland, College Park, MD 20742 USA.

IEEE Log Number 9408288.

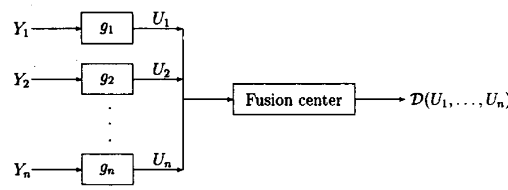

Fig. 1. Distributed detection in $\mathcal{S}_n$.

II. Preliminaries

A triangular array of indices is used for sensors, with the subscript ni denoting the ith sensor in $\mathcal{S}_n$.

Each sensor observation $Y = Y_{ni}$ takes values in the measurable space $(\mathcal{Y}, \mathcal{B})$. The distributions of Y under the null (H_0) and alternative (H_1) hypotheses are denoted by P and Q, respectively.

Assumption 1: P and Q are mutually absolutely continuous, i.e., $P \equiv Q$.

Under Assumption 1, the (pre-quantization) log-likelihood ratio

$$X(y) \triangleq \log\frac{dP}{dQ}(y)$$

is well-defined for $y \in \mathcal{Y}$ and is almost surely (with respect to both P and Q) finite. To exclude trivial cases, the following assumption is also made.

Assumption 2: Every measurable M-ary partition of $\mathcal{Y}$ contains an atom over which $X = \log(dP/dQ)$ is *not* almost everywhere constant.

The message U is generated from Y by an M-ary quantizer g which is either deterministic or randomized. A deterministic quantizer is a measurable mapping of $\mathcal{Y}$ into the message space $\{1, \cdots, M\}$. A randomized quantizer [9] is a random choice from a class of deterministic quantizers; this choice is made according to a fixed probability distribution, is independent of the observed Y, and is performed independently across sensors. The class of all M-ary quantizers (deterministic and randomized) will denoted by $\mathcal{G}_M$.

For the problem in hand, it is known that optimal solutions can be found in the subclass of deterministic and randomized likelihood ratio threshold quantizers [10]. This property will be of no particular value in what follows; this is because quantizers will be characterized almost entirely in terms of the induced *post-quantization distribution pair*. For a quantizer $g \in \mathcal{G}_M$, this is defined as the pair of vectors $\boldsymbol{p} \triangleq (p(1), \cdots, p(M))$ and $\boldsymbol{q} \triangleq (q(1), \cdots, q(M))$, where

$$p(u) = P\{U = u\} \qquad \text{and} \qquad q(u) = Q\{U = u\}$$

for $1 \leq u \leq M$. Write $\boldsymbol{f} = (\boldsymbol{p}, \boldsymbol{q})$, and define the *post-quantization log-likelihood ratio* of g by

$$X_{\boldsymbol{f}}(u) \triangleq \log\frac{p(u)}{q(u)}.$$

Finally, denote by $\mathcal{F}_M \subset [0, 1]^{2M}$ the class of all $\boldsymbol{f}$'s induced by quantizers in $\mathcal{G}_M$.

The moments of $X_{\boldsymbol{f}}$ under P will play an important role in the subsequent discussion. The following notation will be used:

$$\mu(\boldsymbol{f}) \triangleq E_P[X_{\boldsymbol{f}}]$$

$$\sigma^2(\boldsymbol{f}) \triangleq E_P[|X_{\boldsymbol{f}} - \mu(\boldsymbol{f})|^2]$$

and

$$\rho(\boldsymbol{f}, \delta) \triangleq E_P[|X_{\boldsymbol{f}} - \mu(\boldsymbol{f})|^{2+\delta}].$$

0018–9448/95$04.00

Note that $\mu(\boldsymbol{f})$ is just a Kullback–Leibler divergence, i.e.

$$\mu(\boldsymbol{f}) = D(\boldsymbol{p}\|\boldsymbol{q}) = \sum_{u=1}^{M} p(u)\log\frac{p(u)}{q(u)}.$$

By Stein's lemma [5], $D(\boldsymbol{p}\|\boldsymbol{q})$ is the asymptotic type II error exponent obtained when all sensors employ a common quantizer with post-quantization distribution pair $\boldsymbol{f}$. Consequently,

$$D_M \triangleq \sup_{\boldsymbol{f}\in\mathcal{F}_M} \mu(\boldsymbol{f})$$

equals the error exponent of the best identical-quantizer system and also—by the results in [9]—that of the optimal system. Under Assumption 2, $D_M > D_{M-1}$ for all $M \geq 2$.

The main result of this correspondence is obtained under the following assumption.

Assumption 3: There exists $\delta > 0$ for which

$$\sup_{\boldsymbol{f}\in\mathcal{F}_M} E_P\left[|X_{\boldsymbol{f}}|^{2+\delta}\right] < \infty. \quad (1)$$

In *every* occurrence, δ will be understood as positive and satisfying (1), and thus $\mu(\boldsymbol{f})$, $\sigma^2(\boldsymbol{f})$, and $\rho(\boldsymbol{f},\delta)$ will always be bounded above. Note that (1) is implied by the stronger (and easier to verify) condition that the pre-quantization log-likelihood X satisfies

$$E_P\left[|X|^{2+\delta}\right] < \infty.$$

Throughout this correspondence, $\boldsymbol{f}_{ni} = (\boldsymbol{p}_{ni}, \boldsymbol{q}_{ni})$ will denote the post-quantization distribution pair of the ith quantizer in $\mathcal{S}_n$. The corresponding log-likelihood ratio $X_{\boldsymbol{f}_{ni}}(U_{ni})$ will be abbreviated as X_{ni}. The quantities $\mu(\boldsymbol{f}_{ni})$, $\sigma^2(\boldsymbol{f}_{ni})$, and $\rho(\boldsymbol{f}_{ni},\delta)$ will be abbreviated as μ_{ni}, σ^2_{ni}, and $\rho_{ni}(\delta)$, respectively. Then

$$m_n \triangleq \sum_{i=1}^{n}\mu_{ni}, \qquad s_n^2 \triangleq \sum_{i=1}^{n}\sigma^2_{ni}, \quad \text{and} \quad r_n(\delta) \triangleq \sum_{i=1}^{n}\rho_{ni}(\delta).$$

The functions $\phi(\cdot)$ and $\Phi(\cdot)$ will denote the unit Gaussian pdf and cdf, respectively. It is assumed throughout that $M \geq 2$.

III. Properties of Post-Quantization Distributions

The purpose of this section is to highlight some properties of the functions $\mu(\boldsymbol{f})$, $\sigma^2(\boldsymbol{f})$, and $\rho(\boldsymbol{f},\delta)$ defined in Section II and to introduce an additional assumption which is needed for our main results.

It is well-known (see, e.g., [9]) that the set $\mathcal{F}_M$ of post-quantization distribution pairs is a closed convex subset of

$$\mathcal{X}_M \triangleq \left\{(\boldsymbol{p},\boldsymbol{q}) \in [0,1]^{2M} : \sum_{u=1}^{M} p(u) = \sum_{u=1}^{M} q(u) = 1\right\}.$$

Assumption 1 implies that for any $(\boldsymbol{p},\boldsymbol{q}) \in \mathcal{F}_M$, $p(u) = 0$ if and only if (iff) $q(u) = 0$; while Assumption 2 guarantees that $\mathcal{F}_M$ and $\mathcal{X}_M$ have the same dimension, namely $2M - 2$. From [9], all distribution pairs $(\boldsymbol{p},\boldsymbol{q})$ on the relative boundary of $\mathcal{F}_M$ can be generated by randomized likelihood-ratio threshold quantizers. The set $\mathcal{F}_2$ is depicted in Fig. 2.

The functions $\mu(\boldsymbol{f})$, $\sigma^2(\boldsymbol{f})$, and $\rho(\boldsymbol{f},\delta)$ are continuous at any point $\boldsymbol{f} \in \mathcal{F}_M$ with strictly positive coordinates. Under Assumption 3, $\mu(\boldsymbol{f})$ and $\sigma^2(\boldsymbol{f})$ are also continuous on the whole domain $\mathcal{F}_M$. Indeed, for any $\boldsymbol{f}' \in \mathcal{F}_M$

$$p'(u)\left|\log\frac{p'(u)}{q'(u)}\right|^2 \leq B^{\frac{2}{2+\delta}}(p'(u))^{\frac{\delta}{2+\delta}}. \quad (2)$$

where B is the value of the supremum in (1). Since the right-hand side tends to zero together with $p'(u)$, both $\mu(\boldsymbol{f})$ and $\sigma^2(\boldsymbol{f})$ are also continuous at points $\boldsymbol{f} = (\boldsymbol{p},\boldsymbol{q})$ such that $p(u) = q(u) = 0$ for some u.

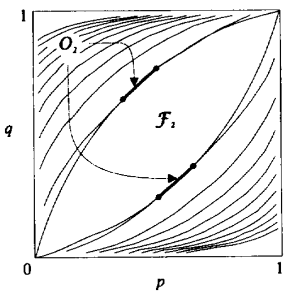

Fig. 2. $\mathcal{F}_2$ and $\mathcal{O}_2$ as subsets of the two-dimensional set $\mathcal{X}_2$. Shown also are contours of $\mu(\cdot)$.

Consider now the set

$$\mathcal{O}_M \triangleq \{\boldsymbol{f} \in \mathcal{F}_M : \mu(\boldsymbol{f}) = D_M\}$$

where D_M is the maximum of $\mu(\boldsymbol{f})$ over $\mathcal{F}_M$. Continuity of $\mu(\boldsymbol{f})$ on $\mathcal{F}_M$ implies that $\mathcal{O}_M$ is a nonempty closed subset of $\mathcal{F}_M$. Convexity of $\mu(\boldsymbol{f})$—which is strict on any convex subdomain that excludes points with $\boldsymbol{p} = \boldsymbol{q}$ [6, Thm. 2.7.2]—implies that $\mathcal{O}_M$ consists of extreme points of $\mathcal{F}_M$ (see Fig. 2). From [9], it is also known that the distribution pairs in $\mathcal{O}_M$ can be generated by deterministic likelihood-ratio threshold quantizers.

It is also useful to note that every point in $\mathcal{O}_M$ has strictly positive entries. If this were not true, i.e., if $p(u) = q(u) = 0$ for some $(\boldsymbol{p},\boldsymbol{q}) \in \mathcal{O}_M$ and some u, then $(\boldsymbol{p},\boldsymbol{q})$ would also correspond to an $(M-1)$-ary quantizer. Using Assumption 2 and standard properties of the Kullback–Leibler divergence, it would then be possible to refine this $(M-1)$-ary quantizer to an M-ary one yielding a higher value of $\mu(\cdot)$ and thus also a contradiction.

Let $\Delta(\boldsymbol{f},\mathcal{O}_M)$ denote the minimum Euclidean distance of $\boldsymbol{f}$ from the closed set $\mathcal{O}_M$, and define

$$\mathcal{N}_a(\mathcal{O}_M) \triangleq \{\boldsymbol{f} \in \mathcal{F}_M : \Delta(\boldsymbol{f},\mathcal{O}_M) \leq a\}.$$

By the foregoing discussion, $a > 0$ can be chosen such that

$$\mathcal{N}_a(\mathcal{O}_M) \subset [a, 1-a]^{2M}. \quad (3)$$

From (3) it follows that on the closed set $\mathcal{N}_a(\mathcal{O}_M)$, all partial derivatives (of any order) of $\sigma^2(\cdot)$ are continuous and hence bounded. It is then straightforward to show that exists a finite A_1 such that

$$(\forall \boldsymbol{f}_1 \in \mathcal{N}_a(\mathcal{O}_M),\ \boldsymbol{f}_2 \in \mathcal{F}_M)$$
$$|\sigma^2(\boldsymbol{f}_1) - \sigma^2(\boldsymbol{f}_2)| \leq A_1 \cdot \|\boldsymbol{f}_1 - \boldsymbol{f}_2\|. \quad (4)$$

Also from (3), it follows that both $\sigma^2(\boldsymbol{f})$ and $\rho(\boldsymbol{f},\delta)$ are bounded away from zero on $\mathcal{N}_a(\mathcal{O}_M)$.

The final assumption is as follows.

Assumption 4: There exists $A_2 > 0$ such that for all $\boldsymbol{f} \in \mathcal{F}_M$,

$$D_M - \mu(\boldsymbol{f}) \geq A_2 \cdot \Delta^2(\boldsymbol{f},\mathcal{O}_M).$$

A few comments on the above condition are in order here:

1) Although it is possible to construct examples of sets $\mathcal{F}_M$ that violate Assumption 4, these examples are not derived from parametric families of distributions.
2) Assumption 4 is satisfied in the case of a finite observation space $\mathcal{Y}$. This is because $\mathcal{F}_M$ is polygonal, hence $\mathcal{O}_M$ consists solely of vertices of $\mathcal{F}_M$. Strict convexity of $\mu(\boldsymbol{f})$ along the edges of $\mathcal{O}_M$ then leads to the required inequality.

3) In cases where the distribution of the pre-quantization log-likelihood ratio X is absolutely continuous, each distribution pair $\boldsymbol{f}$ on the relative boundary of $\mathcal{F}_M$ can be parametrized by the likelihood-ratio threshold quantizer that generates $\boldsymbol{f}$, i.e., $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{t})$, where $\boldsymbol{t}$ is a $(M-1)$-dimensional vector of (increasing) thresholds. It can then be shown that Assumption 4 is satisfied if $\mu(\boldsymbol{f}(\boldsymbol{t}))$ is twice continuously differentiable with respect to $\boldsymbol{t}$ and achieves its global maxima at points $\boldsymbol{t}$ where its Hessian is (strictly) negative-definite. For $M = 2$ and $M = 3$, this condition was not violated in any of parametric tests investigated with the aid of MATHEMATICA, including testing for a fixed signal in Gaussian noise.

IV. Error Probability Bounds

This section investigates upper and lower bounds on the error probabilities of hypothesis tests based on the messages $U_{n1}, \cdots, U_{nn}$.

Lemma 1 (Lower Bound on Type II Error Probability): Let $\lim_{n\to\infty} n^{-1} m_n = D_M$ and $\delta \le 1$. Then for n sufficiently large, the minimum type II error probability $\beta_n(\alpha)$ attainable subject to an upper bound α on the type I probability satisfies

$$\beta_n(\alpha) \ge \frac{\varphi(\Phi^{-1}(\alpha))}{2\, s_n\, e} \exp\left\{-\frac{36}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{1+\delta}}\right\} \cdot \exp\left\{-s_n\, \Phi^{-1}(\alpha) - m_n\right\}.$$

Proof: Since the messages $U_{n1}, \cdots, U_{nn}$ are independent, a sufficient statistic for the optimal global decision rule is the sum of the post-quantization log-likelihoods. Here, the normalized form

$$s_n^{-1} \sum_{i=1}^{n} (X_{ni} - \mu_{ni})$$

is used, which has zero mean and unit variance under P. The cdf of

$$s_n^{-1} \sum_{i=1}^{n} (X_{ni} - \mu_{ni})$$

under P is denoted by $H_n(\cdot)$.

From Assumption 3 and the condition

$$\lim_{n\to\infty} n^{-1} m_n = D_M$$

it is easily deduced (see the proof of Lemma 3) that

$$\lim_{n\to\infty} \frac{r_n(\delta)}{s_n^{2+\delta}} = 0, \tag{5}$$

$$\lim_{n\to\infty} s_n = \infty. \tag{6}$$

The first equality is just Lyapounov's condition, under which

$$s_n^{-1} \sum_{i=1}^{n} (X_{ni} - \mu_{ni})$$

converges in distribution to a unit Gaussian variable [1, Thm. 27.3]. The Berry–Esseen Theorem [7, sec. 2.5.1, Thm. 2.6] yields that

$$(\forall n) \quad \sup_{t\in\Re} |H_n(t) - \Phi(t)| \le 6\frac{r_n(\delta)}{s_n^{2+\delta}} \tag{7}$$

provided $\delta \le 1$.

Let

$$\varepsilon_n \triangleq \Phi^{-1}\left(\alpha + 6 r_n(\delta) s_n^{-2-\delta}\right).$$

Using a first-order Taylor series expansion of $\Phi(\cdot)$ and (6), is is easy to show that

$$0 \le \varepsilon_n - \Phi^{-1}(\alpha) \le \frac{12}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{2+\delta}} \to 0. \tag{8}$$

If the acceptance region for the alternative hypothesis is chosen as

$$\mathcal{U}_n \triangleq \left\{\frac{1}{s_n}\sum_{i=1}^{n}(X_{ni} - \mu_{ni}) \le \varepsilon_n\right\}$$

then by (7)

$$P(\mathcal{U}_n) = H_n(\varepsilon_n) \ge \Phi(\varepsilon_n) - 6\frac{r_n(\delta)}{s_n^{2+\delta}} = \alpha.$$

Although this choice violates the type I error constraint, it is nevertheless useful for lower bounding $\beta_n(\alpha)$. Indeed, $\mathcal{U}_n$ is defined via a likelihood ratio partition and hence by the Neyman–Pearson lemma, $P(\mathcal{U}_n) \ge \alpha$ implies $Q(\mathcal{U}_n^c) \le \beta_n(\alpha)$. Therefore, for $\eta_n > 0$

$$\begin{aligned}\beta_n(\alpha) &\ge Q(\mathcal{U}_n^c)\\ &\ge Q\left\{\varepsilon_n + \eta_n \ge \frac{1}{s_n}\sum_{i=1}^{n}(X_{ni} - \mu_{ni}) > \varepsilon_n\right\}\\ &\ge P\left\{\varepsilon_n + \eta_n \ge \frac{1}{s_n}\sum_{i=1}^{n}(X_{ni} - \mu_{ni}) > \varepsilon_n\right\}\\ &\quad \cdot \exp\{-s_n(\varepsilon_n + \eta_n) - m_n\}\\ &\ge h_n(\eta_n)\exp\{-s_n\,\eta_n\}\exp\{-s_n\,\varepsilon_n - m_n\}\end{aligned} \tag{9}$$

where

$$h_n(\eta_n) \triangleq \Phi(\varepsilon_n + \eta_n) - \Phi(\varepsilon_n) - 12 r_n(\delta) s_n^{-2-\delta}$$

and the last inequality follows from (7). If η_n is chosen so as to approach zero as n tends to infinity, then

$$\lim_{n\to\infty}\frac{\Phi(\varepsilon_n + \eta_n) - \Phi(\varepsilon_n)}{\eta_n} = \varphi(\lim_{n\to\infty}\varepsilon_n) = \varphi(\Phi^{-1}(\alpha)).$$

By (5) and (6), an appropriate choice for η_n is

$$\eta_n \triangleq \frac{24}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{2+\delta}} + \frac{1}{s_n}.$$

Hence, for all sufficiently large n

$$h_n(\eta_n) \ge \frac{1}{2}\varphi(\Phi^{-1}(\alpha))\,\eta_n - 12\frac{r_n(\delta)}{s_n^{2+\delta}} = \frac{\varphi(\Phi^{-1}(\alpha))}{2\, s_n}. \tag{10}$$

The combination of (9), (10), and (8) yields

$$\begin{aligned}\beta_n(\alpha) &\ge \frac{\varphi(\Phi^{-1}(\alpha))}{2\, s_n}\exp\left\{-\frac{24}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{1+\delta}} - 1\right\}\\ &\quad \cdot \exp\left\{-s_n\left(\Phi^{-1}(\alpha) + \frac{12}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{2+\delta}}\right) - m_n\right\}\\ &= \frac{\varphi(\Phi^{-1}(\alpha))}{2\, s_n\, e}\exp\left\{-\frac{36}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{1+\delta}}\right\}\\ &\quad \cdot \exp\left\{-s_n\,\Phi^{-1}(\alpha) - m_n\right\},\end{aligned}$$

thereby completing the proof. ■

Remark: Although the assumption $\lim_{n\to\infty} n^{-1} m_n = D_M$ in the hypothesis of Lemma 1 may appear somewhat contrived, it is, as will be seen in the following section, satisfied by both the optimal and best-identical quantizer designs.

The derivation of the upper bound on the type II error probability employs a basic large deviations technique in addition to the Berry-Esseen theorem.

Lemma 2 (Upper Bound on Type II Error Probability): Let $\lim_{n\to\infty} n^{-1} m_n = D_M$ and $\delta \le 1$. Then for n sufficiently large, the minimum type II error probability $\beta_n(\alpha)$ attainable subject to an upper bound α on the type I probability satisfies

$$\beta_n(\alpha) \le \frac{\varphi(\Phi^{-1}(\alpha))}{s_n}\left(2 + \frac{12}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{1+\delta}}\right) \cdot \exp\left\{\frac{12}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n(\delta)}{s_n^{1+\delta}}\right\}\exp\left\{-s_n\,\Phi^{-1}(\alpha) - m_n\right\}.$$

Proof: The large deviations notation used here is the same as in [2]. Let

$$Y_n \triangleq \sum_{i=1}^{n} (X_{ni} - \mu_{ni})$$

and denote the distribution of Y_n under Q by $F_n(\cdot)$. The *tilted* distribution with parameter θ corresponding to $F_n(\cdot)$ is defined by

$$dF_n^{(\theta)}(x) = \frac{\exp(\theta x) dF_n(x)}{\int \exp(\theta x') dF_n(x')} = \frac{\exp(\theta x) dF_n(x)}{M_n(\theta)} \tag{11}$$

where $M_n(\theta)$ is the moment generating function of the distribution $F_n(\cdot)$. Note that

$$M_n(1) = \exp\{-m_n\}$$

and

$$F_n^{(1)}(s_n x) = H_n(x)$$

where $H_n(\cdot)$ is the cdf of $s_n^{-1} Y_n$ under P (also used in the proof of Lemma 1).

As in the proof of Lemma 1, consider an acceptance region for the alternative hypothesis given by

$$\mathcal{V}_n \triangleq \left\{ \frac{1}{s_n} \sum_{i=1}^{n} (X_{ni} - \mu_{ni}) \le \zeta_n \right\}$$

where $\zeta_n \triangleq \Phi^{-1}\left(\alpha - 6r_n(\delta)s_n^{-2-\delta}\right)$. By analogy to (8), it is now true that

$$0 \le \Phi^{-1}(\alpha) - \zeta_n \le \frac{12}{\varphi(\Phi^{-1}(\alpha))} \frac{r_n(\delta)}{s_n^{2+\delta}} \to 0. \tag{12}$$

Also, by (7)

$$P(\mathcal{V}_n) = H_n(\zeta_n) \le \Phi(\zeta_n) + 6r_n(\delta)s_n^{-2-\delta} = \alpha.$$

Thus $\mathcal{V}_n$ satisfies the type I error constraint and

$$\begin{aligned}\beta_n(\alpha) &\le Q(\mathcal{V}_n^c) \\ &= \int_{s_n \zeta_n}^{\infty} dF_n(x) \\ &= M_n(1) \int_{s_n \zeta_n}^{\infty} \exp(-x)\, dF_n^{(1)}(x) \\ &= \exp\{-m_n\} \int_{\zeta_n}^{\infty} \exp(-s_n x)\, dH_n(x) \\ &= \exp\{-s_n \zeta_n - m_n\} \cdot I_n, \end{aligned} \tag{13}$$

where

$$I_n = \int_0^{\infty} \exp\{-s_n x\} dH_n(x + \zeta_n).$$

Note that $\lambda_n(dx) \triangleq s_n \exp\{-s_n x\}\, dx$ defines an exponential distribution on $[0, \infty)$. Integration by parts and (7) then yield

$$\begin{aligned} I_n &= \int_0^{\infty} [H_n(x+\zeta_n) - H_n(\zeta_n)] \lambda_n(dx) \\ &\le \int_0^{\infty} \left[\Phi(x+\zeta_n) - \Phi(\zeta_n) + 12 \frac{r_n(\delta)}{s_n^{2+\delta}} \right] \lambda_n(dx) \\ &= \frac{1}{s_n} \int_0^{\infty} \varphi(x+\zeta_n) \lambda_n(dx) + 12 \frac{r_n(\delta)}{s_n^{2+\delta}}. \end{aligned} \tag{14}$$

From (12), it follows that

$$\begin{aligned} \left| \varphi(x+\zeta_n) - \varphi(x + \Phi^{-1}(\alpha)) \right| &\le \sup_{x \in \Re} |\varphi'(x)| \cdot \left| \zeta_n - \Phi^{-1}(\alpha) \right| \\ &\le \frac{12}{\sqrt{2\pi e}\, \varphi(\Phi^{-1}(\alpha))} \frac{r_n(\delta)}{s_n^{2+\delta}} \end{aligned}$$

which in the light of (6) implies that $\varphi(x+\zeta_n)$ converges uniformly to $\varphi(x + \Phi^{-1}(\alpha))$. By (6), $\lambda_n(\cdot)$ converges to a degenerate distribution with unit mass at the origin. The first integral in (14) will therefore converge to $\varphi(\Phi^{-1}(\alpha))$ (see [2, Lemma 1.]). Thus if n is sufficiently large

$$I_n \le \frac{\varphi(\Phi^{-1}(\alpha))}{s_n} \left(2 + \frac{12}{\varphi(\Phi^{-1}(\alpha))} \frac{r_n(\delta)}{s_n^{1+\delta}} \right). \tag{15}$$

From (12), (13), and (15) it follows that

$$\begin{aligned} \beta_n(\alpha) \le{}& \frac{\varphi(\Phi^{-1}(\alpha))}{s_n} \left(2 + \frac{12}{\varphi(\Phi^{-1}(\alpha))} \frac{r_n(\delta)}{s_n^{1+\delta}} \right) \\ &\cdot \exp\left\{ \frac{12}{\varphi(\Phi^{-1}(\alpha))} \frac{r_n(\delta)}{s_n^{1+\delta}} \right\} \exp\{-s_n \Phi^{-1}(\alpha) - m_n\} \end{aligned}$$

as required. ■

Clearly, the validity of Lemmas 1 and 2 extends to *centralized* Neyman–Pearson testing with independent—but not necessarily identically distributed—observations under appropriate assumptions. In the special case of i.i.d. observations for which the third moment of the log-likelihood is finite, the type II error probability of the centralized system is given by

$$\beta_n(\alpha) = C_1(n) \frac{\varphi(\Phi^{-1}(\alpha))}{\sqrt{n}\sigma} \exp\{-\sqrt{n}\sigma\Phi^{-1}(\alpha) - n\mu\}$$

where $C_1(n)$ is a bounded constant satisfying

$$\begin{aligned} \exp\left\{-1 - \frac{36\rho(1)}{\varphi(\Phi^{-1}(\alpha))\sigma^2}\right\} \le C_1(n) \le{}& \left(2 + \frac{12\rho(1)}{\varphi(\Phi^{-1}(\alpha))\sigma^2}\right) \\ &\cdot \exp\left\{\frac{12\rho(1)}{\varphi(\Phi^{-1}(\alpha))\sigma^2}\right\} \end{aligned}$$

for all sufficiently large n.

V. Main Result

This section compares the error performance of the optimal and best identical quantizer systems $\mathcal{S}_n^*$ and $\mathcal{S}_n^\circ$. Basic properties of $\mathcal{S}_n^*$ are given in the following lemma.

Lemma 3: $\mathcal{S}_n^*$ satisfies

1) $\lim_{n\to\infty} n^{-1} m_n = D_M$.
2) $r_n(\delta) s_n^{-1-\delta} = O(n^{(1-\delta)/2})$.

Proof: As is shown in [3, Thm. 1], Assumption 3 implies

$$\beta_n(\alpha) \ge \exp\{-C\sqrt{n} - m_n\}$$

for some $C > 0$. Using this lower bound, it is straightforward to show that if statement 1) is not true, then the error exponent of $\beta_n(\alpha)$ is less than D_M, which is impossible for $\mathcal{S}_n^*$.

From 1) it follows that in $\mathcal{S}_n^*$, the fraction of quantizers with post-quantization distribution pairs that lie outside $\mathcal{N}_a(\mathcal{O}_M)$ (defined in Section III) is $o(1)$. This, in conjunction with the fact that both $\sigma^2(\boldsymbol{f})$ and $\rho(\boldsymbol{f}, \delta)$ are bounded away from zero and infinity on $\mathcal{N}_a(\mathcal{O}_M)$, readily yields statement 2). A similar argument proves (5) and (6). ■

In upper-bounding the type II error probability of $\mathcal{S}_n^\circ$, it is clear that any identical-quantizer system $\mathcal{S}_n^\circ$ with

$$\mu_{ni} = \mu_n^\circ \qquad \text{and} \qquad \sigma_{ni}^2 = (\sigma_n^\circ)^2$$

can be used. Statement 2) of Lemma 3 is trivially true for $\mathcal{S}_n^\circ$. Under the assumption

$$\lim_{n\to\infty} \mu_n^\circ = D_M \tag{16}$$

statement 1) will also be true. Both $\mathcal{S}_n^*$ and $\mathcal{S}_n^\circ$ will then satisfy the hypotheses of Lemmas 1 and 2, and

$$\begin{aligned} \frac{\beta_n^*(\alpha)}{\beta_n^\circ(\alpha)} &\ge \frac{\beta_n^*(\alpha)}{\beta_n^\circ(\alpha)} \\ &\ge C_2(n) \exp\{\Phi^{-1}(\alpha)\left(\sqrt{n}\sigma^\circ - s_n^*\right) + (n\mu_n^\circ - m_n^*)\} \end{aligned}$$

where

$$C_2(n) = \frac{1}{2e}\left(2 + \frac{12}{\varphi(\Phi^{-1}(\alpha))}\frac{r_n^\circ(\delta)}{(s_n^\circ)^{1+\delta}}\right)^{-1} \cdot \exp\left\{-\frac{12}{\varphi(\Phi^{-1}(\alpha))}\left(3\frac{r_n^*(\delta)}{(s_n^*)^{1+\delta}} + \frac{r_n^\circ(\delta)}{(s_n^\circ)^{1+\delta}}\right)\right\}.$$

Note that by Lemma 3(2), $C_2(n)$ is of the form

$$\exp\{-c(\delta,\alpha)n^{(1-\delta)/2}\}.$$

In particular, $C_2(n)$ is bounded below if $\delta = 1$.

The main result of this correspondence is as follows.

Theorem 1: Let $\delta \le 1$ satisfy (1). If $\alpha \le 1/2$, or if $\alpha > 1/2$ and Assumption 4 holds, then

$$\frac{\beta_n^*(\alpha)}{\beta_n^\circ(\alpha)} \ge \exp\{-c'(\delta,\alpha)n^{\frac{1-\delta}{2}}\}$$

for some constant $c' = c'(\delta,\alpha)$. In particular (under the same hypothesis), if the supremum of $E_P\left[|X_{\boldsymbol f}|^3\right]$ over $\mathcal{F}_M$ is finite, then the ratio $\beta_n^*(\alpha)/\beta_n^\circ(\alpha)$ is bounded from below.

Proof: By the foregoing discussion, it suffices to prove that there is a choice of μ_n° which satisfies (16) and yields

$$\liminf_{n\to\infty}\left[\Phi^{-1}(\alpha)\left(\sqrt{n}\sigma_n^\circ - s_n^*\right) + (n\mu_n^\circ - m_n^*)\right] > -\infty.$$

The following arguments will show that there exist such choices of μ_n° that yield the above bound for *arbitrary* parameters

$$m_n = \sum_{i=1}^n \mu_{ni}$$

and

$$s_n^2 = \sum_{i=1}^n \sigma_{ni}^2$$

replacing m_n^* and $(s_n^*)^2$, respectively. For simplicity, the superscript "°" will be dropped from μ_n° and σ_n°.

Case 1: $\alpha \le 1/2$. In this case, $\Phi^{-1}(\alpha) = -\gamma^2 \le 0$ and (see (17) at the bottom of this page) where the Cauchy–Schwarz inequality

$$n\cdot\sum_{i=1}^n z_i^2 \ge \left(\sum_{i=1}^n z_i\right)^2$$

was used.

The function $\mu(\boldsymbol f) - \gamma^2 n^{-1/2}\sigma(\boldsymbol f)$ is continuous on $\mathcal{F}_M$, hence it achieves a global maximum at (say) $\boldsymbol f = \boldsymbol f_n$. By the fact that $\sigma^2(\boldsymbol f)$ is bounded, this maximum value approaches D_M as n tends to infinity. Therefore, the choice $(\mu_n, \sigma_n^2) = (\mu(\boldsymbol f_n), \sigma^2(\boldsymbol f_n))$ satisfies (16) and yields a nonnegative value for the lower bound in (17). The required result follows.

Case 2: $\alpha > 1/2$. In this case $\Phi^{-1}(\alpha) = \gamma^2 > 0$. Let $\boldsymbol f = \boldsymbol f_0$ be the post-quantization distribution pair in $\mathcal{O}_M$ with the highest $\sigma^2(\boldsymbol f)$, and for all n take

$$(\mu_n, \sigma_n^2) = (\mu(\boldsymbol f_0), \sigma^2(\boldsymbol f_0)) = (D_M, \sigma^2).$$

Note that this choice trivially satisfies (16). Also, denote by J_n the set of indices i for which $\sigma_{ni}^2 \ge \sigma^2$. Then

$$\begin{aligned} s_n - \sqrt{n}\sigma &= \frac{s_n^2 - n\sigma^2}{s_n + \sqrt{n}\sigma} \\ &\le \frac{s_n^2 - n\sigma^2}{\sqrt{n}\sigma} \\ &\le \frac{1}{\sqrt{n}\sigma}\sum_{i\in J_n}\left[\sigma_{ni}^2 - \sigma^2\right]. \end{aligned}$$

Let $(\mu_{ni}, \sigma_{ni}^2)$ correspond to $\boldsymbol f_{ni}$ and consider for each i a point $\boldsymbol f'_{ni}$ in $\mathcal{O}_M$ such that $\|\boldsymbol f_{ni} - \boldsymbol f'_{ni}\| = \Delta(\boldsymbol f, \mathcal{O}_M)$. Since σ^2 is the maximum variance attainable on $\mathcal{O}_M$, it follows with the aid of (4) that

$$\begin{aligned} \sum_{i\in J_n}\left[\sigma_{ni}^2 - \sigma^2\right] &\le \sum_{i\in J_n}\sigma_{ni}^2 - (\sigma'_{ni})^2 \\ &\le A_1\cdot\sum_{i\in J_n}\Delta(\boldsymbol f_{ni}, \mathcal{O}_M) \\ &\le A_1\cdot\sum_{i=1}^n\Delta(\boldsymbol f_{ni}, \mathcal{O}_M) \end{aligned}$$

and, therefore, that

$$s_n - \sqrt{n}\sigma \le \frac{A_1}{\sqrt{n}\sigma}\sum_{i=1}^n\Delta(\boldsymbol f_{ni}, \mathcal{O}_M). \tag{18}$$

Assumption 4 and the Cauchy–Schwarz inequality yield

$$\begin{aligned} nD_M - m_n &= \sum_{i=1}^n\left[D_M - \mu_{ni}\right] \\ &\ge A_2\cdot\sum_{i=1}^n\Delta^2(\boldsymbol f_{ni}, \mathcal{O}_M) \\ &\ge \frac{A_2}{n}\left(\sum_{i=1}^n\Delta(\boldsymbol f_{ni}, \mathcal{O}_M)\right)^2. \end{aligned} \tag{19}$$

From (18) and (19) it follows that

$$\begin{aligned} &\Phi^{-1}(\alpha)\left(\sqrt{n}\sigma_n - s_n\right) + (n\mu_n - m_n) \\ &\quad = \gamma^2\left(\sqrt{n}\sigma - s_n\right) + (nD_M - m_n) \\ &\quad \ge -\frac{\gamma^2 A_1}{\sqrt{n}\sigma}\sum_{i=1}^n\Delta(\boldsymbol f_{ni}, \mathcal{O}_M) + \frac{A_2}{n}\left(\sum_{i=1}^n\Delta(\boldsymbol f_{ni}, \mathcal{O}_M)\right)^2. \end{aligned}$$

The required result holds because the quadratic $A_2 z^2 - \gamma^2\sigma^{-1}A_1 z$ has a finite minimum as z ranges over $\Re$. ■

$$\begin{aligned} \Phi^{-1}(\alpha)\left(\sqrt{n}\sigma_n - s_n\right) + (n\mu_n - m_n) &= \gamma^2\left(s_n - \sqrt{n}\sigma_n\right) + (n\mu_n - m_n) \\ &\ge \gamma^2\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n\sigma_{ni} - \sqrt{n}\sigma_n\right) + (n\mu_n - m_n) \\ &= \sum_{i=1}^n\left[\left(\mu_n - \frac{\gamma^2}{\sqrt{n}}\sigma_n\right) - \left(\mu_{ni} - \frac{\gamma^2}{\sqrt{n}}\sigma_{ni}\right)\right] \end{aligned} \tag{17}$$

VI. Conclusions

In designing a large parallel distributed system for Neyman–Pearson detection based on i.i.d. observations, the use of identical quantizers is marginally suboptimal. While previous results had shown that the best identical-quantizer system $\mathcal{S}_n^\circ$ achieves the same error exponent as the absolutely optimal system $\mathcal{S}_n^*$, a much stronger result was obtained in this correspondence, namely that the ratio between the actual type II error probabilities is in many cases bounded.

It is noteworthy that Theorem 1 differentiates between the cases $\alpha \leq 1/2$ and $\alpha > 1/2$, requiring an additional regularity condition (Assumption 4) for the latter case. If Assumption 4 is violated, then it is possible to have $\beta_n^*(\alpha)/\beta_n^\circ(\alpha) \to 0$ even when the third moment of the post-quantization log-likelihood is bounded, i.e., $\delta = 1$ in (1). Boundedness of third moments is not an unreasonable requirement, considering that boundedness of second moments ($\delta = 0$ in (1)) is needed in order to show equality of error exponents in $\mathcal{S}_n^\circ$ and $\mathcal{S}_n^*$.

Finally, it should be noted that the lower bound of Theorem 1 also holds for the limiting case $\delta = 0$, but that the proof requires a more elaborate central limit theorem argument.

References

[1] P. Billingsley, *Probability and Measure*, 2nd. ed. New York: Wiley, 1986.

[2] J. A. Bucklew and J. S. Sadowsky, "A contribution to the theory of Chernoff bounds," *IEEE Trans. Inform. Theory*, vol. IT-39, no. 1, pp. 249-254, 1993.

[3] P.-N. Chen and A. Papamarcou, "Asymptotic results on the symmetry of distributed detection," in *Proc. 26th Conf. on Information Sciences and Systems* (Princeton, NJ, Mar. 1992), pp. 180-184.

[4] ——, "New asymptotic results in parallel distributed detection," *IEEE Trans. Inform. Theory*, vol. IT-39, no. 6, pp. 1847–1863, 1993.

[5] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations," *Ann. Math. Stat.*, vol. 23, pp. 493–507, 1952.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[7] P. Hall, *Rates of Convergence in the Central Limit Theorem.* Boston, MA: Pitman, 1982.

[8] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Contr., Sig. Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[9] ——, "Extremal properties of likelihood-ratio quantizers," in *Proc. 29th Conf. on Decision and Control* (Honolulu, HI, 1990), pp. 2680–2685.

[10] D. Warren and P. Willett, "Optimal decentralized detection for conditionally independent sensors," in *Proc. American Control Conf.*, 1989, pp. 1326–1329.

Min–Max Signal Design for Detection Subject to Linear Distortion

Boaz Rippin and Ehud Weinstein, *Fellow, IEEE*

***Abstract*— We consider the problem of deciding between M signals that are observed through an unknown FIR system (channel). We suggest a generalized likelihood ratio test, and develop an upper bound on the detection performance under the worst possible channel. We then suggest a signal design strategy that is immune to the unknown channel effect, in which we construct the transmitted signals using an autoregressive mechanism. The signals in this class are distinguished only by the location of their poles. We show that in some cases these signal designs achieve the bound.**

***Index Terms*— Signal detection, composite hypotheses testing, signal design.**

I. Introduction

Consider the problem of deciding between the following: M hypotheses:

$$H_m: y_n = \sum_{k=0}^{K-1} s_{n-k}^{(m)} h_k + v_n, \quad \begin{matrix} n = 0, 1, 2, \cdots, N-1 \\ m = 1, \cdots, M \end{matrix} \tag{1}$$

where $\{y_n\}_{n=0}^{N-1}$ are the observed data samples, $\{s_n^{(m)}\}_{n=-(K-1)}^{N-1}$ are the samples of the mth signal, $\{h_k\}_{k=0}^{K-1}$ are the unknown impulse response coefficients of the system (channel), $K < N$, and v_n represents additive noise, modeling errors, etc.

Since the impulse response of the channel is unknown, we have a composite hypotheses testing problem. It is equivalent to the statistical problem of discriminating between nonnested linear regression models [2], [3], [9].

Expressing (1) for $n = 0, 1, \cdots, (N-1)$

$$H_m: \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} s_0^{(m)} & s_{-1}^{(m)} & \ddots & s_{-K+1}^{(m)} \\ s_1^{(m)} & s_0^{(m)} & \ddots & s_{-K+2}^{(m)} \\ s_2^{(m)} & s_1^{(m)} & \ddots & s_{-K+3}^{(m)} \\ \ddots & \ddots & \ddots & \ddots \\ s_{N-1}^{(m)} & s_{N-2}^{(m)} & \ddots & s_{N-K}^{(m)} \end{bmatrix} \times \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{K-1} \end{bmatrix} + \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \\ v_{N-1} \end{bmatrix}. \tag{2}$$

or in matrix notation

$$H_m: \underline{y} = S_m \underline{h} + \underline{v}. \tag{3}$$

To decide which hypothesis is true, we search for the combination of signal and channel response that best fit the data in the sense of

Manuscript received June 4, 1993; revised September 13, 1994.

B. Rippin is with the Department of Electrical Engineering-Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel.

E. Weinstein is with the Department of Electrical Engineering-Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel, and the Woods Hole Oceanographic Institute, Woods Hole, MA 02543 USA.

IEEE Log Number 9408365.