New Asymptotic Results in Parallel Distributed Detection

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

Abstract—The performance of a parallel distributed detection system is investigated as the number of sensors tends to infinity. It is assumed that the i.i.d. sensor data are quantized locally into m-ary messages and transmitted to the fusion center for binary hypothesis testing. The boundedness of the second moment of the postquantization log-likelihood ratio is examined in relation to the asymptotic error exponent. It is found that when that second moment is unbounded, the Neyman–Pearson error exponent can become a function of the test level, whereas the Bayes error exponent remains, as previously conjectured by Tsiatis, unaffected. Large deviations techniques are also employed to show that in Bayes testing, the equivalence of absolutely optimal and best identical quantizer systems is not limited to error exponents, but extends to the actual Bayes error probabilities up to a multiplicative constant.

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

I. INTRODUCTION

PERHAPS the most common architecture in distributed (or decentralized) detection is the parallel feedforward system \( \mathcal{P} \) depicted in Fig. 1. It consists of \( n \) geographically dispersed sensors, noiseless one-way communication links, and a fusion center. Each sensor makes an observation (denoted by \( Y \)) of a random source, quantizes \( Y \) into an \( m \)-ary message \( U = g(Y) \), and then transmits \( U \) to the fusion center. Upon receipt of \( (U_1, \ldots, U_n) \), the fusion center makes a global decision \( \mathcal{D}(U_1, \ldots, U_n) \) about the nature of the random source.

The optimal design of \( \mathcal{P} \) entails choosing quantizers \( g_1, \ldots, g_n \) and a global decision rule \( \mathcal{D} \) so as to optimize a given performance index. In this paper, we consider binary hypothesis testing under the classical (Neyman–Pearson) and Bayesian formulations. The first formulation dictates minimization of the type II error probability subject to an upper bound on the type I error probability, while the second stipulates minimization of the Bayes error probability, computed according to the prior probabilities of the two hypotheses.

![Diagram of distributed detection system](image)

Fig. 1. Distributed detection in \( \mathcal{P} \).

The joint optimization of entities \( g_1, \ldots, g_n \) and \( \mathcal{D} \) in \( \mathcal{P} \) is a hard computational task [22], except in trivial cases (such as when the observations \( Y \) lie in a set of size no greater than \( m \)). The complexity of the problem can only be reduced by introducing additional statistical structure in the observations. For example, it has been shown that whenever \( Y_1, \ldots, Y_n \) are independent given each hypothesis, an optimal solution can be found in which \( g_1, \ldots, g_n \) are threshold-type functions of the local likelihood ratio (possibly with some randomization for Neyman–Pearson testing). These results, which were developed—among others—[18], [23], [21], [25], are directly relevant to the discussion in this paper. Still, we should note that optimization of \( g_1, \ldots, g_n \) over the class of threshold-type likelihood-ratio quantizers is prohibitively complex when \( n \) is large.

Of equal importance are situations where the statistical model exhibits spatial symmetry in the form of permutation invariance with respect to the sensors. A natural question to ask in such cases is whether a symmetric optimal solution exists in which the quantizers \( g_i \) are identical; if so, then the optimal system design is considerably simplified. The answer is clearly negative for cases where sensor observations are highly dependent; as an extreme example, take \( Y_1 = \cdots = Y_n = Y \) with probability \( 1 \) under each hypothesis, and note that any two identical quantizers lead to a redundancy. If one heuristically interprets dependence as the existence of a common “core” in all observations, then the previous example would suggest that this core is not best handled by identical quantizers. Yet, as the core shrinks and the data become independent (and identically distributed), it also becomes conceivable that a symmetric solution consisting of identical quantizers may exist. Surprisingly, this is not always true; counterexamples to this symmetry conjecture were first given in [20], and were also explored in [24] and [7]. We now know that this asymmetry is by no means a

Manuscript received August 3, 1992. This work was supported in part by the Institute for Systems Research (a National Science Foundation Engineering Research Center) at the University of Maryland, College Park. This paper was presented in part at the 1992 Conference on Information Sciences and Systems, Princeton, NJ, Mar. 1992 and at the 1993 International Symposium on Information Theory, San Antonio, TX, Jan. 1993.

The authors are with the Department of Electrical Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742.

IEEE Log Number 9213030.
pathology, and arises quite frequently in numerical simula-
tions.

The initial goal of our inquiry was the characterization of
those i.i.d. source models for which the optimal system
$S_n$ consists of identical quantizers. Early results—both
analytical and numerical—showed that the presence of
nonidentical quantizers in the optimal $S_n$ is neither spo-
radic nor predictable. As a result, we found the determi-
nation of necessary and sufficient conditions for the exist-
ence of symmetric solutions quite intractable (notwith-
standing the contributions of [24], where certain necessary
conditions were derived). We subsequently focused on the
study of asymptotically (large $n$) optimal systems, thereby
continuing the inquiry initiated by Tsitsiklis in [20].

The general problem is as follows. System $S_n$ is used
for testing $H_0: P$ versus $H_1: Q$, where $P$ and $Q$ are one-di-

gensional marginals of the i.i.d. data $Y_1, \ldots, Y_n$. As $n$ tends
to infinity, both the minimum type II error probability $\beta_n^*(\alpha)$ (as a function of the type I error probability bound $\alpha$) and the Bayes error probability $\gamma_n^*(\pi)$ (as a function of the prior probability $\pi$ of $H_0$) vanish at an exponential rate. It thus becomes legitimate to adopt a measure of
asymptotic performance based on the error exponents

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

$$e_{B}^*(\pi) = \lim_{n \to \infty} \frac{-1}{n} \log \gamma_n^*(\pi).$$

It was shown in [20] that, under certain assumptions on the
hypothesis $P$ and $Q$, it is possible to achieve the same
error exponents using identical quantizers. Thus, if $\beta_n^*(\alpha)$,
$\gamma_n^*(\pi)$, $e_{SP}^*(\alpha)$, and $e_B^*(\pi)$ are the counterparts of $\beta_n^*(\alpha)$,
$\gamma_n^*(\pi)$, $e_{SP}^*(\alpha)$, and $e_B^*(\pi)$ under the constraint that the
quantizers $g_1, \ldots, g_n$ are identical, then

$$(\forall \alpha \in (0,1)) \quad e_{SP}^*(\alpha) = e_{SP}^*(\alpha)$$

and

$$(\forall \pi \in (0,1)) \quad e_B^*(\pi) = e_B^*(\pi).$$

[Of course, for all $n$, $\beta_n^*(\alpha) \geq \beta_n^*(\alpha)$ and $\gamma_n^*(\pi) \geq 
\gamma_n^*(\pi).$] This result provides some justification for restricting
attention to identical quantizers when designing a
system consisting of a large number of sensors.

Our work revolves around two issues that remained
open in [20]. The first issue is the exact asymptotics of
the minimum error probabilities achieved by the absolutely
optimal and best identical-quantizer systems. Note
that equality in the error exponents of $\gamma_n^*(\pi)$ and $\gamma_n^*(\pi)$
does not in itself guarantee that for any given $n$, the values
of $\gamma_n^*(\pi)$ and $\gamma_n^*(\pi)$ are in any sense close. In particular, the
ratio $\gamma_n^*(\pi)/\gamma_n^*(\pi)$ may vanish at a subexponential rate,
and thus the best identical-quantizer system may be vastly
inferior to the absolutely optimum system. [The same argu-
ment can be made for $\beta_n^*(\alpha)/\beta_n^*(\alpha).$]

In examining this problem, we first performed numeri-
cal simulations of Bayes testing in $S_n$. These showed that
the ratio $\gamma_n^*(\pi)/\gamma_n^*(\pi)$ is (apparently) bounded from
below by a positive constant which is, in many cases,
reasonably close to unity. In this paper, we substantiate
these findings by using large deviations techniques to
prove that $\gamma_n^*(\pi)/\gamma_n^*(\pi)$ is, indeed, always bounded from
below. For Neyman–Pearson testing, we state a weaker
lower bound on the ratio $\beta_n^*(\alpha)/\beta_n^*(\alpha)$. In either case,
we conclude that the optimal system essentially consists of
“almost identical” quantizers, and is thus only marginally
different from the best identical-quantizer system.

The second issue stemming from [20] concerns the
assumption under which identical-quantizer systems were
shown to attain the optimal error exponent. This assump-
tion stipulates that the variance of the log-likelihood ratio
at the quantizer output be bounded over all quantizer
mappings $g$, where this variance is computed w.r.t. $P$ in
the case of Neyman–Pearson testing, and w.r.t. both $P$
and $Q$ in the case of Bayes testing.

It was conjectured in [19] that the aforementioned
boundedness assumption may be superfluous in the case
of Bayes testing. We are able to give proof to this con-
jecture. On the other hand, we have found that the bound-
edness assumption is significant in Neyman–Pearson
testing. In its absence, the finite-sample exponent $(-1/n)$
$\log \beta_n^*(\alpha)$ exhibits three distinct modes, including one in
which the lim inf and lim sup are dependent on the type I
error probability $\alpha$. This is somewhat surprising, consid-
ering that the asymptotic Neyman–Pearson error exponent
is independent of $\alpha$ in most detection problems of inter-
est, including the present one when the boundedness
assumption holds.

The paper is organized as follows. Some general facts
on distributional distance, error exponents, and quantiza-
tion are covered in Section II. In Section III, a brief
analysis of Neyman–Pearson testing under the bounded-
ess assumption is given, followed by an examination of
models that violate this assumption. Section IV is a prepa-
ration for the main result on Bayes testing (i.e., that the
ratio $\gamma_n^*(\pi)/\gamma_n^*(\pi)$ is bounded from below, which is
proved in Section V. Some final comments appear in
Section VI. Appendix A contains proofs of two lemmas
from Section IV, while in Appendix B it is demonstrated
that the ratio $\gamma_n^*(\pi)/\gamma_n^*(\pi)$ need not tend to unity.

Sections IV and V on Bayes testing are quite indepen-
dent of Section III on Neyman–Pearson testing.

II. GENERAL BACKGROUND

A. Observational Model

Each sensor observation $Y = Y_t$ takes values in the
measurable space $(\mathcal{Y}, \mathcal{B})$. The distribution of $Y_t$ under the
null ($H_0$) and alternative ($H_1$) hypotheses is denoted
by $P$ and $Q$, respectively. For simplicity, the same
notation is used for the $n$-fold product of these measures
defined on $(\mathcal{Y}^n, \mathcal{B}^n)$.

Assumption 1: $P$ and $Q$ are mutually absolutely con-
tinuous, i.e., $P = Q$.

Although our analysis can be tailored to situations
where $P$ and $Q$ are not mutually absolutely continuous,
we feel that such problems are of somewhat secondary interest and that their inclusion in this paper would detract from its clarity and cohesiveness.

Under Assumption 1, the (prequantization) log-likelihood ratio

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

is well defined for $y \in \mathcal{Y}$ and is a.s. finite. (Since $P \equiv Q$, "almost surely" and "almost everywhere" are understood as under both $P$ and $Q$.) In $\mathcal{S}$, the variable $X(Y)$ will also be denoted as $X_{c}$.

It is well known that in a centralized system where the sensor observations $Y_{1},\ldots,Y_{n}$ are directly available to the fusion center, the optimal hypothesis test (Neyman–Pearson or Bayes) involves $(Y_{1},\ldots,Y_{n})$ only through the sufficient statistic $X_{c} = X_{1} + \cdots + X_{n}$. If each $X_{i}$ happens to be an m-ary random variable (so that it can be obtained from $Y_{i}$ via an m-ary quantizer), then the optimal centralized system can be realized by $\mathcal{S}$. This trivial case is excluded by the following assumption.

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

**B. Distance Functionals and Error Exponents**

The error exponents in centralized detection are given by the functionals $D(P\|Q)$ and $\rho(P, Q)$ defined below. Both entities are derived from the class of $f$-divergences [10] or Ali–Silvey distances [1].

The (Kullback–Leibler, informational) divergence, or relative entropy, of $P$ relative to $Q$ is defined by

$$D(P\|Q) \triangleq E_{P}[X] = \int \log \frac{dP}{dQ}(y) dP(y).$$

On the convex domain $\mathcal{M}(\mathcal{Y})$ consisting of distribution pairs $(P, Q)$ with the property $P = Q$, the functional $D(P\|Q)$ is nonnegative and convex. It is also equal [8] to the optimal Neyman–Pearson error exponent in testing $P$ versus $Q$ at any level $\alpha \in (0, 1)$ based on the i.i.d. observations $Y_{1},\ldots,Y_{n}$. The Chernoff exponent $\rho(P, Q)$ is derived from the moment generating function $\Phi(s)$ of $X$ under $Q$:

$$\Phi(s) \triangleq E_{Q}[\exp \{sX\}] = \int \left( \frac{dP}{dQ}(y) \right)^{s} dQ(y).$$

For fixed $s \in [0, 1]$, both $\Phi(s)$ and $\log \Phi(s)$ (the cumulant generating function) are finite-valued concave functionals of the pair $(P, Q)$ in $\mathcal{M}(\mathcal{Y})$, while for fixed $(P, Q) \in \mathcal{M}(\mathcal{Y})$, both $\Phi(s)$ and $\log \Phi(s)$ are finite and convex in $s \in [0, 1]$. This last property, together with the fact that $\Phi(0) = \Phi(1) = 1$, guarantees that $\Phi(s)$ has a minimum value which is less than or equal to unity, achieved by some $s^{*} \in (0, 1)$. We define

$$\rho(P, Q) \triangleq -\log \Phi(s^{*}) = -\log \left[ \min_{s \in (0, 1)} \Phi(s) \right].$$

and note [8] that the Chernoff exponent coincides with the Bayes error exponent in centralized hypothesis testing.

The functionals $D(\|\cdot\|)$ and $\rho(\cdot, \cdot)$ are also associated with error exponents in $\mathcal{S}$. As was shown in [20] under a boundedness assumption, the error exponents $\epsilon_{\rho}(n)$ and $\epsilon_{\Phi}(\pi)$ of the absolutely optimal system coincide with those achieved by the best identical-quantizer system, i.e., $\epsilon_{\rho}(n)$ and $\epsilon_{\Phi}(\pi)$, respectively. Now, in an $n$-sensor identical-quantizer system, the quantizer outputs $U_{i} = g(Y_{i})$ are clearly i.i.d. Thus, by the discussion of the previous paragraphs, a fixed (as $n$ varies) quantizer mapping $g$ yields error exponents $D(P_{g}\|Q_{g})$ for Neyman–Pearson testing and $\rho(P_{g}, Q_{g})$ for Bayes testing, where $P_{g}$ and $Q_{g}$ are the distributions of $U_{i}$ under the null and alternative hypotheses, respectively. The optimal error exponent is then obtained by choosing the mapping $g$ so as to maximize the appropriate functional.

**C. Deterministic and Randomized Quantization; LRQ’s**

We now consider quantization for $\mathcal{S}$ in somewhat greater detail.

A deterministic m-ary quantizer is a measurable mapping $g$ from the observation space $(\mathcal{Y}, \mathcal{B})$ to the message space $\mathcal{Y}_{m} \triangleq \{1,\ldots,m\}$.

For completeness, it also necessary to introduce randomized quantizers in $\mathcal{S}$. Each such quantizer $g$ (w.l.o.g. we use the same symbol) is specified by a finite vector $(g_{1},\ldots,g_{m})$ of deterministic quantizers together with a pmf vector $(\lambda_{1},\ldots,\lambda_{m})$. Before making an observation, each randomized quantizer in $\mathcal{S}$ independently selects a deterministic mapping according to its individual $(\lambda_{i},g_{i})$, and then proceeds to encode its observation using the chosen mapping. Independent (across sensors) randomization is needed in order to preserve independence of the messages $U_{i}$ (the possibility of cooperative, or synchronized, randomization has also been mentioned in [21] and [25]).

The distributions of the message $U \in \mathcal{Y}_{m}$ produced by $g$ are denoted by $P_{g}$ and $Q_{g}$, and are readily obtainable from $P$, $Q$, and $g$. The postquantization log-likelihood ratio is defined on $\mathcal{Y}_{m}$ by

$$X_{g}(U) \triangleq \log \frac{P_{g}(U)}{Q_{g}(U)}.$$ 

Clearly, if $g$ is deterministic, then both the output message $U$ and the log-likelihood ratio $X_{g}(U)$ are measurable functions of the observation $Y$. In that case, we obtain the smoothing property

$$\exp \{X_{g}(U)\} = \frac{1}{Q(g(Y) = U)} \int_{g(Y) = U} \frac{dP}{dQ}(y) dQ(y) \quad = E_{Q}[\exp \{X\}g(Y) = U].$$

Throughout the paper, $\mathcal{S}$ will employ $n$ distinct quantizers $g_{1},\ldots,g_{n}$. We will suppress the common “$n$” from most subscripts, and we will also abbreviate $X_{g}(U_{i})$ as $X_{g}$.
The joint distributions of the independent messages $U_1, \ldots, U_n$ will be denoted by $P_{g}$ and $Q_{g}$, where $P_{g} \triangleq P_{g_1} \times \cdots \times P_{g_n}$ and $Q_{g} \triangleq P_{g_1} \times \cdots \times P_{g_n}$.

Of special importance to our analysis are the so-called (log-) likelihood ratio quantizers (LRQ's). The definitions and properties given below are adapted from [23] and [21].

A deterministic m-ary LQ is specified by an m-ary partition

$$\tau = (I_1, \ldots, I_m)$$

of the real line, where the $I_i$'s are consecutive intervals (with $I_1$ being the leftmost). We call such $\tau$ a likelihood ratio partition (LRP), and emphasize that $I_g$ can be a singleton or even empty. The output of the quantizer is given by the rule

$$U(Y) = u_i \text{ iff } X(Y) \in I_u.$$

A randomized m-ary LQ is obtained by randomly selecting, according to a pmf $(p_i)$, one of finitely many deterministic m-ary LRP's $\tau$, where the $\tau_i$'s may differ only in the endpoints of the constituent intervals. For example, if $[a, b]$ is the $i$th interval in $\tau$, then the corresponding interval in $\tau_i$ must be one of $[a, b], [a, b)$, $(a, b]$, or $(a, b)$. As a consequence of this definition, if the observed log-likelihood ratio $X$ falls in the interior of one such interval, then the output of the randomized LQ is, with probability one, the index of that interval. Thus randomized LQ's are only marginally different from deterministic ones, and they become indistinguishable— in terms of the resulting output distributions $(P_{g}, Q_{g})$— from the latter when the distribution of $X$ (under either hypothesis) has no point masses.

The class of randomized m-ary quantizers will be denoted by $\mathcal{Q}_m$, and the deterministic m-ary LRP's will be denoted by $\mathcal{F}_m$. A deterministic LQ $g$ will commonly be designated by its corresponding LRP $\tau$, and $P_{g}$, $Q_{g}$, and $X_{g}$ will also appear as $P_{\tau}$, $Q_{\tau}$, and $X_{\tau}$, respectively. Note that the random variable $X_{\tau}$ is a measurable function of both $Y$ and $X(Y)$.

D. Optimality of LQ's

The optimality of LQ's in $\mathcal{F}_n$ was established in a series of results outlined below.

1) For Neyman–Pearson testing, it was shown in [23] that if the observations $Y_1, \ldots, Y_n$ are independent (but not necessarily identically distributed) under both hypotheses, then an optimal solution $(g_1, \ldots, g_n, \mathcal{Q})$ exists in which all $g_i$'s are randomized LQ's. In general, the optimal Neyman–Pearson fusion rule $\mathcal{Q}_n$ operating on the discrete messages $U_1, \ldots, U_n$ will also be randomized. However, for the special case where the distribution (under either hypothesis) of each quantization log-likelihood $X$ has no point masses, it can be shown [25] that the fusion rule $\mathcal{Q}_n$ must be deterministic. (We already know from the discussion of the previous subsection that the optimal $g_i$'s are also deterministic in this special case.)

2) For Bayesian testing with independent (but, again, not necessarily identically distributed) observations, it was shown in [18] that an optimal solution for $\mathcal{F}_n$ exists in which all $g_i$'s are deterministic LRQ's, and $\mathcal{Q}$ is also deterministic.

3) The error exponents achievable by the optimal $\mathcal{F}_n$ (assuming i.i.d. observations and the boundedness condition) are given by [20] by

$$\left(\forall \alpha \in (0, 1)\right) \quad e_{\alpha}^n(\alpha) = \sup_{g \in \mathcal{Q}_n} D(P_{g} \| Q_{g})$$

and

$$\left(\forall \pi \in (0, 1)\right) \quad e_{\pi}^n(\pi) = \sup_{g \in \mathcal{Q}_n} \rho(P_{g}, Q_{g})$$

As $g$ varies over $\mathcal{Q}_n$, the resulting range of output distribution pairs $(P_{g}, Q_{g})$ is a closed convex subset $\mathcal{L}(P, Q)$ of $[0, 1]^m$. Using the convexity properties of the functionals $D(P_{g} \| Q_{g})$ and $\Phi(\pi)$ defined earlier, one can show that both suprema will be achieved by distribution pairs that are extremal points of $\mathcal{L}(P, Q)$. These extremal points can, in turn, be generated by deterministic LRQ's, and it thus follows that

$$\sup_{g \in \mathcal{Q}_n} D(P_{g} \| Q_{g}) = \max_{\tau \in \mathcal{F}_n} D(P_{\tau} \| Q_{\tau}) = D_m \tag{2.2}$$

$$\sup_{g \in \mathcal{Q}_n} \rho(P_{g}, Q_{g}) = \max_{\tau \in \mathcal{F}_n} \rho(P_{\tau}, Q_{\tau}) = \rho_n. \tag{2.3}$$

This argument is developed in detail in [21], and similar results have appeared in [23].

In light of the above discussion, all problems studied in this paper have solutions which are randomized LRQ's, and the generic quantizer $g$ can be restricted to this class.

Notation: Throughout the paper, $a \wedge b \triangleq \min\{a, b\}$ and $a \vee b \triangleq \max\{a, b\}$.

III. THE BOUNDEDNESS ASSUMPTION IN NEYMAN–PEARSON TESTING

The equality (and finiteness) of $e_{\alpha}^n(\alpha)$ and $e_{\pi}^n(\alpha)$ was established in [20, Theorem 2] under the assumption that $\mathcal{E}_p(X) < \infty$. Actually, the proof only utilized the following weaker condition on the postquantization log-likelihood ratio.

Assumption 3: $\sup_{f \in \mathcal{F}_n} \mathcal{E}_f(X) < \infty$.

Let us briefly examine the above assumption. For an arbitrary randomized quantizer $g$, let $p_u = P_{g}(u)$ and $q_u = Q_{g}(u)$, then,

$$E_p[|X^2] = \sum_{u=1}^{m} p_u \log^2 \frac{p_u}{q_u}. \tag{2.4}$$

Our first observation is that the negative part $X_{-}$ of $X$ has bounded second moment under $P$, and thus Assumption 3, is equivalent to $\sup_{g \in \mathcal{F}_n} E_p[|X^2] < \infty$. Indeed, we have

$$E_p[|X^2] = \sum_{u: p_u < q_u} p_u \log \frac{p_u}{q_u} = \sum_{u: p_u < q_u} p_u \log^2 \frac{q_u}{p_u}.$$
and using the inequality \( \log |x| \leq \sqrt{x} - 1 \) for \( x \geq 1 \), we obtain

\[
E_p \left( X^2 \right) \leq 4 \sum_{u: p_u < q_u} p_u \left( \sqrt{\frac{q_u}{p_u}} - 1 \right)^2 \\
\leq 4 \sum_{u=1}^m \left( p_u + q_u \right) \leq 8.
\]

Our second observation is as follows.

**Theorem 1:** Assumption 3 is equivalent to the condition

\[
\sup_{\tau \in \mathcal{J}_s} E_p \left[ X_{\tau}^2 \right] < \infty. \quad (3.1)
\]

**Proof:** Assumption 3 clearly implies (3.1). To prove the converse, let \( \tau_\tau \) be the LRP in \( \mathcal{J}_s \) defined by

\[
\tau_\tau = ((-\infty, t], (t, \infty)), \quad (3.2)
\]

and let \( p(t) = P(X > t), q(t) = Q(X > t) \). By (3.1), there exists \( b < \infty \) such that for all \( t \in \mathcal{R} \),

\[
E_p \left[ X_{\tau}^2 \right] = (1 - p(t)) \log \left( \frac{1 - p(t)}{1 - q(t)} \right) + p(t) \log \left( \frac{p(t)}{q(t)} \right) \leq b. \quad (3.3)
\]

Consider now an arbitrary deterministic \( m \)-ary quantizer \( g \), with output pmfs \( (p_1, \ldots, p_m) \) and \( (q_1, \ldots, q_m) \), let the maximum of \( p_u \) \( \log \left( \frac{p_u}{q_u} \right) \) subject to \( p_u \geq q_u \) be achieved at \( u = u_\ast \). Then, from the first observation in this section, it follows that

\[
E_p \left[ X^2 \right] \leq mp_* \log \frac{p_*}{q_*} + 8.
\]

To see that \( p_* \log \left( \frac{p_*}{q_*} \right) \) is bounded from above if (3.3) holds, note that for a given \( p_* = P(Y \in C^*) \), the value of \( q_* = Q(Y \in C^*) \) can be bounded from below using the Neyman–Pearson lemma. In particular, there exist \( t \in \mathcal{R} \) and \( \mu \in [0, 1] \) such that

\[
p_* = \mu P(X = t) + p(t) \\
q_* \geq \mu Q(X = t) + q(t).
\]

If \( P(X = t) = 0 \), then

\[
p_* \log \frac{p_*}{q_*} \leq p(t) \log \frac{p(t)}{q(t)},
\]

where by virtue of (3.3), the RHS is upper-bounded by \( b \). Otherwise, the RHS is of the form

\[
(p(t) + \mu P(X = t)) \log \frac{p(t) + \mu P(X = t)}{q(t) + \mu Q(X = t)}
\]

for some \( \mu \in (0, 1) \). For \( \mu = 1 \), this can again be bounded using (3.1); take \( \tau' \) consisting of intervals \((-\infty, t] \) and \([t, \infty) \), or use \( \tau_\tau \) and a simple continuity argument. Then, the log-sum inequality [9, Theorem 2.7.1] can be applied together with the concavity of \( f(t) = \sqrt{t} \) to show that the same bound \( b \) is valid for \( \mu \in (0, 1) \) (details are omitted).

If \( g \) is a randomized quantizer, then the probabilities \( p_u \) and \( q_u \) defined previously will be expressible as \( \sum \lambda, p^{(l)} \) and \( \sum \lambda_q q^{(l)} \). Here, \( k \) ranges over a finite index set, and each pair \((p^{(l)}, q^{(l)})\) is derived from a deterministic quantizer. Again, using the log-sum inequality and the concavity of \( f(t) = \sqrt{t} \), one can obtain \( p_* \log \left( \frac{p_*}{q_*} \right) \leq b \).

**Remark:** It is easy to show that (3.1) is, in fact, equivalent to

\[
\limsup_{t \to \infty} E_p \left[ X_{\tau}^2 \right] < \infty \quad (3.4)
\]

where \( \tau_\tau \) is defined in (3.2).

Our next result is a refinement on the asymptotic equivalence of the best identical-quantizer system \( (\mu) \) and the absolutely optimal system \( (\mu) \). The proof is a variant of a standard argument for Stein’s lemma (see, e.g., [9, Theorem 12.8.1]) and also parallels the condensed proof of Theorem 2 in [20]; it can be found in [6, Theorem 1].

**Theorem 2:** If Assumption 3 holds, then for all \( \alpha \in (0, 1) \),

\[
\frac{1}{n} \log \beta^*(\alpha) = D_m + O(n^{-1/2})
\]

and

\[
\frac{1}{n} \log \beta^r(\alpha) = D_m + O(n^{-1/2}). \quad \square
\]

As an immediate corollary, we have \( e_{Q, p}^*(\alpha) = e_{Q, p}^r(\alpha) = D_m \), which is Theorem 2 in [20]. The stated result sharpens this equality by demonstrating that the finite-sample error exponents of the absolutely optimal and best identical-quantizer systems converge at a rate \( O(n^{-1/2}) \). It also motivates the following observations.

The first observation concerns the accuracy, or tightness, of the \( O(n^{-1/2}) \) convergence factor. Although the upper and lower bounds on \( \beta^*(\alpha) \) in Theorem 2 are based on a suboptimal null acceptance region, it seems that, in general, the \( O(n^{-1/2}) \) rate cannot be improved on. (We have found examples to support this conjecture in the context of centralized testing.) At the same time, it is rather unlikely that the ratio \( \frac{\beta^*(\alpha)}{\beta^r(\alpha)} \) decays as fast as \( \exp \left( -c' \sqrt{n} \right) \), there probably exists a lower bound which is tighter than what is implied by Theorem 2. Preliminary results based on large deviations techniques different from those used in Sections IV and V indicate that, under conditions stronger than Assumption 3, the ratio \( \frac{\beta^*(\alpha)}{\beta^r(\alpha)} \) may indeed be bounded below.

We will report on these findings in a future paper.

The second observation is about the composition of an optimal quantizer \( (g_1, \ldots, g_m) \) for \( \mathcal{J}_s \). The proof of Theorem 2 also yields an upper bound on the number of quantizers that are at least \( \epsilon \)-distant from the deterministic LRO that achieves \( e_{Q, p}^*(\alpha) \). Specifically, if \( K_\epsilon(\epsilon) \) is the number of indexes \( i \) for which

\[
D(P_{\epsilon, i}(Q_{\mathcal{J}})) < D_m - \epsilon
\]
(where $\epsilon > 0$), then
\[
K_n(\epsilon) = O(n^{-1/2}).
\]
Thus, in an optimal system, most quantizers will be "essentially identical" to the one that achieves $\epsilon_{NP}(\alpha)$. As we shall see in Section V, this conclusion can be significantly strengthened in the case of Bayesian testing, and an interesting counterexample can be constructed.

In the remainder of this section, we discuss the asymptotics of Neyman–Pearson testing in situations where Assumption 3 does not hold. By the remark following the proof of Theorem 1, this condition is violated if and only if
\[
\lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}^2] = \infty \tag{3.5}
\]
where $\tau_t$ is the binary LRP defined by
\[
\tau_t = (\langle -\infty, t \rangle, \langle t, \infty \rangle).
\]

We distinguish between three cases, the first two of which are completely subsumed under (3.5):

Case A:
\[
\lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}] = \infty.
\]

Case B:
\[
0 < \lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}] < \infty.
\]

Case C:
\[
\lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}] = 0 \quad \text{and} \quad \lim_{t \to \infty} E_p[X_{\tau_t}^2] = \infty.
\]

Example: Let the observation space be the unit interval $(0, 1)$ with its Borel field. For $a > 0$, define the distributions $P$ and $Q$ by
\[
P(Y \leq y) = y, \quad Q(Y \leq y) = \exp \left\{ a + \frac{a}{1 - y^a} \right\}.
\]
The pdf of $Q$ is strictly increasing in $y$, and thus the likelihood ratio $(dP/dQ)(Y)$ is strictly decreasing in $y$. Hence, the event $\{X > t\}$ can also be written as $\{Y < y_t\}$, where $y_t \to 0$ as $t \to \infty$. Using this equivalence, we can examine the limiting behavior of $E_p[X_{\tau_t}]$ and $E_p[X_{\tau_t}^2]$ to obtain

a) $a > 1$: $\lim_{t \to \infty} E_p[X_{\tau_t}] = \lim_{t \to \infty} E_p[X_{\tau_t}^2] = \infty$ (Case A).

b) $a = 1$: $\lim_{t \to \infty} E_p[X_{\tau_t}] = 2$, $\lim_{t \to \infty} E_p[X_{\tau_t}^2] = \infty$ (Case B).

c) $1/2 < a < 1$: $\lim_{t \to \infty} E_p[X_{\tau_t}] = 0$, $\lim_{t \to \infty} E_p[X_{\tau_t}^2] = \infty$ (Case C).

d) $a \leq 1/2$: $\lim_{t \to \infty} E_p[X_{\tau_t}^2] < \infty$ (Assumption 3 is satisfied).

In Case A, the error exponents $\epsilon_{NP}^a(\alpha)$ and $\epsilon_{NP}^a(\alpha)$ are both infinite. This result is neither difficult to prove nor surprising, considering that $E_p[X_{\tau_t}] = D(P_t||Q)$ can be made arbitrarily large by choice of the quantizer $g$.

Theorem 3: If $\lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}] = \infty$, then for all $m \geq 2$ and $\alpha \in (0, 1)$,
\[
\epsilon_{NP}^a(\alpha) = \epsilon_{NP}(\alpha) = \infty.
\]

We now turn to Case B, which is more interesting. Using the notation $p(t) = p_X(t) = q(t) = Q(X > t)$ introduced earlier, we have
\[
E_p[X_{\tau_t}] = (1 - p(t)) \log \frac{1 - p(t)}{1 - q(t)} + p(t) \log \frac{p(t)}{q(t)}.
\]
The first summand on the RHS clearly tends to zero (as $t \to \infty$, which is understood throughout); hence, the limit sup of the second summand $p(t) \log [p(t)/q(t)]$ is greater than zero. Since $p(t)$ tends to zero, both $\log [p(t)/q(t)]$ and $p(t) \log [p(t)/q(t)]$ have limit sup equal to infinity. Thus, in particular, (3.5) always holds in Case B.

A separate argument (which we omit) reveals that the centralized error exponent $D(P||Q) = E_p[X_{\tau_t}]$ is also infinite in Case B. Yet, unlike Case A, the decentralized error exponent $\epsilon_{NP}^d(\alpha)$ obtained here is not infinite. Quite surprisingly, if this exponent exists, then it must depend on the test level $\alpha$. This is stated in the following theorem.

Theorem 4: Consider hypothesis testing with $m$-ary quantization, where $m \geq 2$. If
\[
o < \lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}] < \infty, \tag{3.6}
\]
then there exist:

i) an increasing sequence of integers $\{n_k, k \in N\}$ and a function $L:(0, 1) \to (0, \infty)$ which is monotonically increasing to infinity, such that
\[
\lim_{k \to \infty} \frac{1}{n_k} \log \beta_{n_k}(\alpha) \geq L(\alpha) \vee D_m;
\]

ii) a function $M:(0, 1) \to (0, \infty)$ which is monotonically increasing to infinity and is such that
\[
\lim_{n \to \infty} \frac{1}{n} \log \beta_n(\alpha) \leq M(\alpha).
\]

Proof: i) Lower Bound: As was argued in the proof of Theorem 2, an error exponent equal to $D_m$ can be achieved using identical quantizers; hence, one part of the bound follows immediately. In what follows, we construct a sequence of identical-quantizer detection schemes with finite sample-error exponent almost exceeding $L(\alpha)$, where $L(\alpha)$ increases to infinity as $\alpha$ tends to unity.

Let $\nu = \lim_{t \to \infty} \sup_{\tau_t} E_p[X_{\tau_t}]$, so that $0 < \nu < \infty$ by (3.6). By subsequence selection, we obtain a sequence of LRP's $\tau_{n_k}$ with the property that $\nu_k \overset{\Delta}{=} E_p[X_{\tau_{n_k}}]$ converges to $\nu$ as $k \to \infty$. (We eliminate $t$ from all subscripts to simplify the notation.) Letting $p_k = p(t_k), q_k = q(t_k), \epsilon_k = \log [1 - p_k/(1 - q_k)]$ and $\zeta_k = \log (p_k/q_k)$, we can write
\[
\nu_k = (1 - p_k) + p_k \zeta_k
\]
where, by the discussion preceding this theorem, $p_k$ and $\epsilon_k$ tend to zero and $\zeta_k$ increases to infinity.
Fix \( \omega > 0 \) and assume w.l.o.g. that \( \xi_k > \xi_{k-1} + (1/\omega) \).
Consider a system consisting of \( n_k \) sensors, where \( n_k = \lfloor \omega \xi_k \rfloor \) and let each sensor employ the same binary LRO with LRP \( \tau_s \). This choice is clearly suboptimal (since \( m \geq 2 \)), but it suffices for our purposes.

Define the set \( \Delta_k \subset [1,2]^{n_k} \) by
\[
\Delta_k = \{(u_1, \ldots, u_n) : \text{at least one } u_i \text{ equals } 2 \}.
\]
Recalling that \( U_i = 2 \) iff the observed log-likelihood ratio is larger than \( \tau_s \), we have
\[
P_k(\Delta_k) = 1 - \left(1 - p_k\right)^{n_k},
\]
and thus
\[
\lim_{k \to \infty} (1 - P_k(\Delta_k)) = \lim_{k \to \infty} \left(1 - p_k\right)^{n_k}
= \lim_{k \to \infty} \left(1 - p_k\right)^{\omega \xi_k}
= \left(\lim_{k \to \infty} \left(1 - p_k\right)^{\xi_k}\right)^\omega = \exp\{-\nu \omega\}
\]
where the last equality follows from the fact that \( \xi_k \to \infty \) and \( p_k \xi_k \to \nu \) as \( k \to \infty \).

Thus, given any \( \delta > 0 \), for all sufficiently large \( k \), the set \( \Delta_k \) is admissible (albeit not necessarily optimal) as a null acceptance region for testing at level \( \alpha = \exp\{-\nu \omega + \delta\} \). For this value of \( \alpha \), we have
\[
\beta^*\alpha_k(\alpha) \leq Q_k(\Delta_k)
= \left[1 - \left(1 - q_k\right)^{n_k}\right]^{\omega}
= \left[1 - \left(1 - p_k\exp\{-\xi_k\}\right)^{n_k}\right]^{\omega}
= n_k p_k \exp\{-\xi_k\} - \frac{n_k(n_k - 1)}{2!} p_k^2 \exp\{-2\xi_k\}
+ \cdots + (-1)^{n_k-1} p_k^{n_k} \exp\{-n_k \xi_k\}
\leq n_k p_k \exp\{-\xi_k\} + n_k^2 p_k^2 \exp\{-2\xi_k\} + \cdots.
\]
Summing the geometric series, we obtain
\[
\beta^\omega\alpha_k(\alpha) \leq \frac{n_k p_k \exp\{-\xi_k\}}{1 - n_k p_k \exp\{-\xi_k\}}.
\]
The RHS denominator tends to unity because \( \xi_k \to \infty \) and \( n_k p_k \to \nu \omega \) as \( k \to \infty \). Since \( \xi_k / n_k \to 1/\omega \), we conclude that
\[
\lim_{k \to \infty} -\frac{1}{n_k} \log \beta^\omega\alpha_k(\alpha) \geq \frac{\nu}{\omega} = \log(1/\alpha) + \delta.
\]
As \( \omega > 0 \) and \( \delta > 0 \) were chosen arbitrarily, it follows that
\[
\lim_{k \to \infty} -\frac{1}{n_k} \log \beta^\omega\alpha_k(\alpha) \geq \frac{\nu}{\log(1/\alpha)}
\]
for all \( \alpha \in (0,1) \). The lower bound in statement i) of the theorem is obtained by taking \( L(\alpha) = \nu / \log(1/\alpha) \).

ii) Upper Bound: Consider an optimal detection scheme for \( \mathcal{A}_n \), with the same setup as in the proof of Theorem 2. Recall, in particular, that the fusion center employs a randomized test with log-likelihood threshold \( \eta_n \) and randomization constant \( \mu_n \).

For \( \theta \) to be an upper bound on the error exponent of \( \beta^\omega\alpha(\alpha) \), it suffices that \( n \theta \) be greater than \( \eta_n \) and such that the events \( \{\|X_{[n]} - X_{[n]}\| \leq n \theta\} \) and \( \{\|X_{[n]} - X_{[n]}\| \geq n \eta\} \) have significant overlap under \( P_\omega \). Indeed, if \( \theta > \eta_n / n \) is such that for all sufficiently large \( n \),
\[
\mu_n P_{\omega}\left\{\sum_{i=1}^{n} X_{[i]} < \eta_n\right\} + P_{\omega}\left\{n \theta \geq \sum_{i=1}^{n} X_{[i]} \geq \eta_n\right\} \geq \epsilon > 0,
\]
then \( \beta^\omega\alpha(\alpha) > \epsilon \exp\{-n \theta\} \), as required.

The threshold \( \eta_n \) is rather difficult to determine, so we use an indirect method for finding \( \theta \). We have
\[
P_{\omega}\left\{\sum_{i=1}^{n} X_{[i]} > n \theta\right\} \leq P_{\omega}\left\{\|X_{[n]} - X_{[n]}\| > n \theta\right\}
\leq \frac{1}{\theta} \sup_{x \in \mathbb{R}} P_{\omega}[|X_{[n]}|] = \frac{1}{\theta} E_{\omega}[|X_{[n]}|],
\]
where the last bound follows from the Markov inequality. We claim that the supremum in this relationship is finite. This is because the negative part of \( X_{[n]} \) has bounded expectation under \( P \) (see the discussion following Assumption 3), and the proof of Theorem 1 can be easily modified to show that \( \nu' = \sup_{x \in \mathbb{R}} E_{\omega}[|X_{[n]}|] \) is finite iff \( \nu' = \limsup_{n \to \infty} E_{\omega}[|X_{[n]}|] \) is (which is our current hypothesis). Thus,
\[
P_{\omega}\left\{\sum_{i=1}^{n} X_{[i]} > n \theta\right\} \leq \frac{\nu'}{\theta}.
\]

Now, let \( \epsilon > 0 \) and \( \theta = \nu' / (1 - \alpha - \epsilon) \), so that \( P_{\omega}\{\|X_{[n]} - X_{[n]}\| > n \theta\} = 1 - \alpha - \epsilon \). The Neyman-Pearson lemma immediately yields \( n \theta > \eta_n \). Also, using
\[
\mu_n P_{\omega}\left\{\sum_{i=1}^{n} X_{[i]} < \eta_n\right\} + P_{\omega}\left\{\sum_{i=1}^{n} X_{[i]} > \eta_n\right\} = 1 - \alpha
\]
and a simple contradiction, we obtain (3.7). Thus, the chosen value of \( \theta \) is an upper bound on \( \limsup_{n \to \infty} \frac{1}{n} \log \beta^\omega\alpha(\alpha) \). Since \( \epsilon > 0 \) can be made arbitrarily small, we have that
\[
M(\alpha) = \frac{\nu'}{1 - \alpha}
\]
is also an upper bound.

From Theorem 1, we conclude that in Case B, the error exponent \( \epsilon^B_n(\alpha) \) must lie between the bounds \( L(\alpha) \) and \( M(\alpha) \) whenever it exists as a limit. (Since \( \nu' \geq \Delta_n \geq \nu \) and \( 1 - \alpha \leq \log(1/\alpha) \), the inequality \( M(\alpha) \geq L(\alpha) \) is indeed true.) These bounds are shown in Fig. 2. We note that an earlier result [6, Theorem 3] employed Poisson's theorem to yield a weaker bound \( L(\alpha) \).

Finally, we briefly argue that the error exponents in Case C are identical to those obtained under Assumption
3 (boundedness of $E_{P}[X_{k}^{2}]$). Indeed, let $g_{1}, \ldots, g_{m}$ be optimal randomized LRO’s for $\mathscr{S}_{n}$. It is easy to show, by techniques similar to those employed in the proof of Theorem 1, that the condition $\lim_{n \to \infty} E_{P}[X_{k}^2] = 0$ implies uniform integrability in the following sense: given $\epsilon \in (0, 1)$, there exists $b = b(\epsilon, m)$ such that $E_{P}[|X_{k}|1_{|X_{k}| > b}] < \epsilon^2$ for all $m$-ary LRO’s $g$. Letting $X_{k}^\ast = X_{k}1_{|X_{k}| \leq b}$ and $X_{k}^\ast = X_{k}1_{|X_{k}| > b}$, we have

$$P_{P} \left( \sum_{i=1}^{n} X_{k}^\ast > n(D_{m} + 2\epsilon) \right) \leq P_{P} \left( \sum_{i=1}^{n} X_{k}^\ast > n(D_{m} + \epsilon) \right) + e^{-\epsilon^2}.$$

Now, $|X_{k}^\ast| \leq b + X_{k}^\ast$, hence, by the observation preceding Theorem 1, $\text{Var}_{P}[X_{k}^\ast]$ is bounded. Since $E_{P}[X_{k}^\ast] \leq D_{m} + \epsilon^2 < D_{m} + \epsilon$, we can apply the Chebyshev inequality to conclude that first summand on the RHS of (3.8) tends to zero as $n$ tends to infinity. As for the second summand, it is less than $\epsilon$ by the Markov inequality and the bound $E_{P}[X_{k}^\ast] < \epsilon^2$. Retracing the proof of the converse (upper bound) part of Theorem 4, we obtain $\lim_{n \to \infty} \left(-1/n\right) \log B_{n}(\alpha) \leq D_{m} + 2\epsilon$ for all $\epsilon \in (0, \alpha)$, and thus finally, $e^{\epsilon_{n}P}(\alpha) = e^{\epsilon_{P}(\alpha)} = D_{m}$.

IV. BAYES TESTING: PRELIMINARIES

We now turn to the asymptotic study of optimal Bayes detection in $\mathscr{S}_{n}$. The prior probabilities of $H_{0}$ and $H_{1}$ are denoted by $\pi$ and $1 - \pi$, respectively, the probability of error of the absolutely optimal system is denoted by $\gamma_{B}^{a}(\pi)$, and the probability of error of the best identical-quantizer system is denoted by $\gamma_{B}^{i}(\pi)$.

In our analysis, we will always assume that $\mathscr{S}_{n}$ employs deterministic $m$-ary LRO’s represented by LRP’s $\tau_{1}, \ldots, \tau_{m}$. This is clearly sufficient by the discussion in Section II-D.

Upon receiving the messages $U_{1}, \ldots, U_{n}$ produced by the quantizers, the fusion center uses the MAP rule to decide in favor of $H_{0}$ iff

$$\frac{P(U_{1}, \ldots, U_{n})}{Q(U_{1}, \ldots, U_{n})} > 1 - \frac{\pi}{\pi},$$

or, equivalently, iff

$$X_{\tau_{1}} + \cdots + X_{\tau_{m}} > \log \frac{1 - \pi}{\pi}.$$

The resulting probability of error is

$$\gamma_{B}^{a}(\pi) = \pi P(X_{\tau_{1}} + \cdots + X_{\tau_{m}} \leq \eta) + (1 - \pi) Q(X_{\tau_{1}} + \cdots + X_{\tau_{m}} > \eta) \quad (4.1)$$

where $\eta = \log(1 - \pi)/\pi$. When the LRP’s are optimal, then $\gamma_{B}^{a}(\pi) = \gamma_{B}^{i}(\pi)$. [Note that the $X_{\tau}$’s are measurable functions of the respective observations $Y_{i}$, and thus it is legitimate to use the product measure $P$ instead of $P_{i}$ in expressions such as (4.1).]

Our aim is to show eventually that the ratio $\gamma_{B}^{a}(\pi)/\gamma_{B}^{i}(\pi)$ is bounded away from zero as long as $P = Q$. The groundwork for this result is developed in this section, and is divided into five units for convenience.

A. Basic Large Deviations Results

The values of $\gamma_{B}^{a}(\pi)$ and $\gamma_{B}^{i}(\pi)$ can be approximated using a large deviations technique which originated in [4] and [2]. Our exposition here is fairly complete; for a general reference, see [5, ch. VII, A] or [15, sect. 5.4 and Appendix 5A].

For fixed $n$, let

$$\Phi_{s}(\cdot) = E_{Q}[\exp(sX_{\tau_{s}})],$$

and

$$M_{s}(\cdot) = E_{Q}[\exp\left\{s(X_{\tau_{1}} + \cdots + X_{\tau_{m}})\right\}].$$

Then, by independence of the $X_{\tau_{s}}$’s, we have $M_{n}(s) = \prod_{s=1}^{m} \Phi_{s}(s)$.

We recall from Section II-B that for each $i$, the function $\Phi_{s}(s) = \Sigma_{s=1}^{m} \log \Phi_{s}(s)$ is convex in $s$. This implies that $\log M_{n}(s) = \Sigma_{s=1}^{m} \log \Phi_{s}(s)$ is convex, and it is also strictly so unless all $X_{\tau}$’s are trivial (which can be safely excluded). Since $\log M_{n}(0) = \log M_{n}(1) = 0$, both $\log M_{n}(s)$ and $M_{n}(s)$ have a unique minimum achieved by $s_{n} \in (0, 1)$. We have

$$M_{n}(s_{n}) \geq \prod_{s=1}^{m} \min_{s(0, 1)} \Phi_{s}(s) = \prod_{s} \exp\left\{-p(\tau_{s}, Q_{s})\right\} \geq \exp\{-n p_{0}\} \quad (4.2)$$

where $p_{0}$ is defined in (2.3). Equality holds throughout iff every $\tau_{s}$ is an $m$-ary LRP that achieves $p_{0}$.

Let $\mathscr{X}_{s}$ be the ($m$-point) range of $X_{\tau_{s}}$, and denote the distributions of $X_{\tau_{s}}$ under $H_{0}$ and $H_{1}$ by $\rho_{0}$ and $\rho_{1}$, respectively. The tilted distribution of $X_{\tau_{s}}$ is the measure $\tilde{\rho}_{s}$ on $\mathscr{X}_{s}$ given by

$$\tilde{\rho}_{s}(x) = \frac{\exp\{s_{n} x\} \Phi_{s}(x)}{\Phi_{s}(s_{n})} \quad (4.3)$$
where \( s_r \) is as defined in the previous paragraph, i.e., it satisfies \( M_\ell(s_r) = 0 \).

By extension, the tilted distribution of the vector \((X_{s_1}, \ldots, X_{s_r})\) is defined as the product measure \( \mathcal{E}_1 \times \cdots \times \mathcal{E}_r \), which clearly preserves the independence of the \( X_i \)'s. Denoting expectation w.r.t. the tilted distribution by \( \bar{E}_r \), we have the identity
\[
E_r[X_{s_1} + \cdots + X_{s_r}] = 0. \tag{4.4}
\]

The Bayes error probability is given by (4.1). This can be expressed in terms of the tilted distribution by noting that
\[
Q(X_{s_1} + \cdots + X_{s_r} > \eta)
= \sum_{x_1 + \cdots + x_r > \eta} \mathcal{E}(x_1) \cdots \mathcal{E}(x_r)
= M_r(s_r) \sum_{x_1 + \cdots + x_r > \eta} \exp \{-s_r(x_1 + \cdots + x_r)\}
- \bar{E}_r(x_1) \cdots \bar{E}_r(x_r); \tag{4.5}
\]
and also since \( \mathcal{P}(x) = \exp \{x \mathcal{E}(x)\} \),
\[
P(X_{s_1} + \cdots + X_{s_r} \leq \eta)
- M_r(s_r) \sum_{x_1 + \cdots + x_r \leq \eta} \exp \{(1 - s_r)(x_1 + \cdots + x_r)\}
- \bar{E}_r(x_1) \cdots \bar{E}_r(x_r). \tag{4.6}
\]
Expressing the sums in (4.5) and (4.6) in terms of the tilted cdf \( \bar{F}_r(x) \) of \( X_{s_1} + \cdots + X_{s_r} \), we can rewrite (4.1) as
\[
\gamma_\ell(\pi) = M_r(s_r) \left[ \pi \int_{x \leq \eta} \exp \{(1 - s_r)x\} \bar{d}\bar{F}_r(x) \right.
+ (1 - \pi) \int_{x > \eta} \exp \{-s_r x\} \bar{d}\bar{F}_r(x) \right]. \tag{4.7}
\]

The Chernoff (upper) bound on \( \gamma_\ell(\pi) \) can be easily derived from (4.7). More useful for our purposes is the lower bound
\[
\gamma_\ell(\pi) \geq [\pi \wedge (1 - \pi)] M_r(s_r)
\int_R \exp \{(1 - s_r)x\} \bar{d}\bar{F}_r(x)
\geq [\pi \wedge (1 - \pi)] M_r(s_r) \int_R \exp \{-|x|\} \bar{d}\bar{F}_r(x), \tag{4.8}
\]
which is independent of \( \eta \) and holds for all \( s_r \in (0,1) \).

We note that (4.8) also yields the lower bound appearing in [17, relationship (3.42)], which—together with the Chernoff bound—was used in [20] to prove the result \( e'_n(\pi) = e'_n(\pi) \) under the boundedness assumption \( \sup_{x \in \mathcal{X}_r} \Phi(x) < \infty \) or its equivalent form (see also [6, Theorem 5])
\[
\sup_{x \in \mathcal{X}_r} \sup_{r \in \mathcal{D}} \bar{E}_r(x) < \infty. \tag{4.9}
\]

**B. CLT Approximations**

Tighter bounds on \( \gamma_\ell(\pi) \) can be derived from (4.7) by a central limit theorem approximation. Since the independent sum \( X_{s_1} + \cdots + X_{s_r} / \sigma_r \) has zero mean and unit variance under the tilted distribution, it may—under certain conditions—converge weakly to a Gaussian \( \mathcal{N}(0,1) \) variable. If so, then \( \bar{F}_r(x) \) can be approximated by a Gaussian cdf. In the (i.i.d.) case of identical LRPs \( \tau_{s_1} \), this technique yields (see, e.g., [2], [12], [13])
\[
\lim_{n \to \infty} \inf \frac{\gamma_n(\pi)}{\sqrt{n}} \exp \{n \rho(P, Q, \pi)\} = c_{\min} > 0 \tag{4.10}
\]
and
\[
\lim_{n \to \infty} \sup \frac{\gamma_n(\pi)}{\sqrt{n}} \exp \{n \rho(P, Q, \pi)\} = c_{\max} < \infty \tag{4.11}
\]
where \( c_{\min} \) and \( c_{\max} \) depend on \( P, Q, \pi, \) and \( \eta \). Furthermore, \( c_{\min} = c_{\max} \), except in certain cases where the variable \( X_n \) has a lattice form.

From (4.10), we immediately obtain
\[
\lim_{n \to \infty} \sup \frac{\gamma_n(\pi)}{\sqrt{n}} \exp \{n \rho_m\} < \infty. \tag{4.12}
\]
In Section V, we will show implicitly that
\[
\lim_{n \to \infty} \inf \frac{\gamma_n(\pi)}{\sqrt{n}} \exp \{n \rho_{m1}\} > 0 \tag{4.13}
\]
and that therefore \( \lim_{n \to \infty} \frac{\gamma_n(\pi)}{\sqrt{n}} > 0 \). The following lemma and its corollary will be essential in proving that result.

**Lemma 1:** If there are uniform (in \( n \) and \( i \leq n \)) bounds \( \alpha > b > 0 \) such that \( C_1 \mathcal{X}_{s_1} \leq b \) a.s., and \( C_2 \) \( \mathbb{V} \mathcal{A}_{s_1} \mathcal{X}_{s_1} \geq b^2 \), then
\[
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \int_R \exp \{-|x|\} \bar{d}\bar{F}_r(x) > 0. \tag{4.14}
\]

**Proof:** Let \( \bar{H}_r(x) \) be the tilted cdf of the normalized sum \( X_{s_1} + \cdots + X_{s_r} / \sigma_r \), and \( G(x) \) be the cdf of an \( \mathcal{N}(0,1) \) distribution. A CLT approximation due to Esseen [11, sect. XVI.3, Theorem 2] gives, for all \( x \) and \( n \),
\[
|\bar{H}_r(x) - G(x)| \leq \frac{6}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}_{\mathcal{Q}}[|X_{s_1} - E[X_{s_1}]|^3].
\]
By an elementary argument, (4.15) can be shown to imply both \( \mathbb{V} \mathcal{A}_{s_1} \mathcal{X}_{s_1} \leq b^2 \) and \( \mathbb{E}_{\mathcal{Q}}[|X_{s_1} - E[X_{s_1}]|^3] \leq b^4 \). As \( \alpha_n \geq \alpha(n) \), we obtain
\[
|\bar{H}_r(x) - G(x)| \leq 6(b/a)^3 n^{-1/2}. \tag{4.15}
\]
We have
\[
\int_R \exp \{-|x|\} \bar{d}\bar{F}_r(x) = \int_R \exp \{-|\alpha_n x|\} d\bar{H}_r(x).
\]
Restricting the range of the right-hand integral to 
\((-m^{-1/2}, m^{-1/2})\), where \(t > 0\) will be specified later, we obtain
\[
\int_R \exp \left\{ -|x| \right\} d\tilde{F}_n(x) \\
\geq \left[ \hat{H}_n(-m^{-1/2}) - \hat{H}_n(m^{-1/2}) \right] \exp \left\{ -t \sigma_n n^{-1/2} \right\} \\
\geq \left[ \hat{H}_n(-m^{-1/2}) - \hat{H}_n(m^{-1/2}) \right] \exp (-tb).
\]

It remains to show that for all sufficiently large \(n\), the difference \(\hat{H}_n(-m^{-1/2}) - \hat{H}_n(m^{-1/2})\) can be made greater than \(cn^{-1/2}\) for \(c > 0\). From the CLT approximation (4.15), we have
\[
\hat{H}_n(t^{-1/2}) - \hat{H}_n(-t^{-1/2}) \\
\geq \frac{G(t^{-1/2}) - G(-t^{-1/2})}{2 - 12(b/a)^3 n^{-1/2}} \\
\geq \frac{2/n^{-1/2}}{12(b/a)^3 n^{-1/2}} \\
= \frac{n^{-1/2}}{6(b/a)^3}.
\]

Since for fixed \(t\), \(G(t^{-1/2}) \to G(0) = (2\pi)^{-1/2}\) as \(n \to \infty\), any choice \(t > 6\sqrt{2\pi}(b/a)^3\) will suffice.

\begin{corollary}
Under conditions C1) and C2) in the hypothesis of Lemma 1,
\[
\lim_{n \to \infty} \gamma_n(\pi) \sqrt{n} \exp \left\{ -n \rho_n \right\} > 0.
\]

Proof: This follows immediately from (4.2), (4.8), and (4.14).
\end{corollary}

\section{C. Inequalities for Chernoff Exponents}

The following fact is well known (see, e.g., [16] or [14]), and is established using Jensen’s inequality together with (2.1).

\begin{lemma}
Let \(g\) be a deterministic \(m\)-ary quantizer. Then for all \(s \in (0, 1),
\[
E_g[\exp \{ sX_s \}] \geq E_g[\exp \{ sY \}]
\]
or, equivalently, \(\Phi(s) \geq \Phi(s)\), with equality if and only if \(X_s = X\) a.s.
\end{lemma}

Using Assumption 2 of Section II and a simple argument, we obtain the following corollary to Lemma 2.

\begin{lemma}
If \(l' < l \leq m\), then
\[
\rho_{l'} < \rho_{l} < \rho(P, Q).
\]
\end{lemma}

\section{D. A Classification of LRPs}

We denote by \(\mathcal{F}_n(\delta)\) the class of all \(\tau\)'s in \(\mathcal{F}_m\) with the property that
\[
(V u \in \mathcal{F}_m) \rightarrow P_{\tau}(u)^{\delta} Q_{\tau}(u) > \delta.
\]
Clearly, for \(\delta > \delta' > 0,\)
\[
\mathcal{F}_m(\delta) \subset \mathcal{F}_m(\delta') \subset \mathcal{F}_m(0) = \mathcal{F}_m.
\]

The complement \(\overline{\mathcal{F}_m(\delta)}\) w.r.t. \(\mathcal{F}\) is denoted by \(\mathcal{B}_m(\delta).\)

The next two lemmas together imply that the class \(\mathcal{F}_m(\delta)\) satisfies conditions C1) and C2) in the hypothesis of Lemma 1. The first result follows immediately from the definition of \(\mathcal{F}_m(\delta)\), while the second lemma is proved in Appendix A.

\begin{lemma}
Let \(\delta > 0\). If \(\mathcal{F}_m(\delta)\) is nonempty, then
\[
\sup_{r \in \mathcal{F}_m(\delta)} \sup_{u \in \mathcal{F}_m} |X_r(u)| < \log(1/\delta).
\]
\end{lemma}

\begin{lemma}
Let \(m \geq 2\) and \(\delta > 0\). If \(\mathcal{F}_m(\delta)\) is nonempty, then
\[
\inf_{r \in \mathcal{F}_m(\delta)} \inf_{u \in \mathcal{F}_m} \mathbb{E}_u[X_r(u)] > 0,
\]
where \(\mathbb{E}\) varies over all tilted distributions of \(X_r\).
\end{lemma}

Next, we consider the class \(\mathcal{B}_m(\delta) \supset \mathcal{F}_m(\delta)\) for \(\delta > 0\). It is reasonable to expect that as \(\delta\) decreases to zero, the performance of any LRP \(\tau\) in \(\mathcal{B}_m(\delta)\) (measured in terms of \(\rho(P, Q)\)) approaches that of an LRP in \(\mathcal{F}_m\). This is effectively stated in the following lemma, which is proved in Appendix A.

\begin{lemma}
For every \(2 \leq l \leq m,\)
\[
\lim_{\delta \to 0} \sup_{r \in \mathcal{F}_m(\delta)} \rho(P, Q_r) = \rho_{l-1}.
\]
\end{lemma}

\section{E. An Inequality for Bayes Testing}

The following lemma completes the preparation for the results of Section V.

\begin{lemma}
Consider two independent observations \(Y_A\) and \(Y_B\) such that
\[
H_0: Y_A \sim P_A \quad \text{and} \quad Y_B \sim P_B; \\
H_1: Y_A \sim Q_A \quad \text{and} \quad Y_B \sim Q_B.
\]
Let \(\gamma_A(\pi), \gamma_B(\pi), \) and \(\gamma_{AB}(\pi)\) be the the minimum probability of error attainable in testing \(H_0\) versus \(H_1\) on the basis of \(Y_A, Y_B, (Y_A, Y_B)\), respectively. Then, for all \(\pi \in [0, 1],\)
\[
\gamma_{AB}(\pi) \geq 2\gamma_A(\pi)\gamma_B(\pi)(1/2).
\]

Proof: Without loss of generality, assume that \(Y_A\) and \(Y_B\) are discrete observations with alphabets \(\mathcal{Y}_A\) and \(\mathcal{Y}_B\), respectively. Then
\[
\gamma_{AB}(\pi) = \sum_{y \in \mathcal{Y}_A} \sum_{y' \in \mathcal{Y}_B} \pi P_A(y) P_B(y') \left(1 - \pi\right) Q_A(y) Q_B(y') \\
\geq \sum_{y \in \mathcal{Y}_A} \sum_{y' \in \mathcal{Y}_B} \left(\pi P_A(y) Q_A(y') \left(1 - \pi\right) Q_B(y')\right) \\
\geq \sum_{y \in \mathcal{Y}_A} \sum_{y' \in \mathcal{Y}_B} \left[P_A(y) \left(1 - \pi\right) Q_A(y')\right] \\
\geq 2\gamma_A(\pi)\gamma_B(1/2),
\]
which completes the proof.
\end{lemma}

\section{V. BAYES TESTING: MAIN RESULT}

We now use the tools developed in Section IV to prove the following theorem.
Theorem 5: In Bayes testing with m-ary quantization,
\[ \lim_{n \to \infty} \gamma_n^\ast(\pi) > 0 \]  
(5.1)
for all \( \pi \in (0,1) \).

Proof: An upper bound on \( \gamma_n^\ast(\pi) \) is immediately obtained from (4.12). It remains to find a good lower bound on \( \gamma_n^\ast(\pi) \).

Step 1. Two Subsystems: Fix \( n \), and consider an absolutely optimal system employing LRP’s \( \tau_1, \ldots, \tau_n \). Also fix \( \delta > 0 \) (to be specified later), and assume w.l.o.g. that
\[
(\forall i \leq n_\delta) \quad \tau_i \in \mathcal{S}_\delta(\delta) \quad \text{and} \quad (\forall i > n_\delta) \quad \tau_i \in \mathcal{R}_\delta(\delta).
\]

Let \( n_B = n - n_\delta \), and define two systems \( \mathcal{S}_\delta \) and \( \mathcal{R}_\delta \) as follows. System \( \mathcal{S}_\delta \) observes \( Y_{i_1}, \ldots, Y_{i_n} \), compresses these observations using the LRP’s \( \tau_{i_1}, \ldots, \tau_{i_n} \), and then performs an optimal test based on \( U_{i_1}, \ldots, U_{i_n} \). While \( \mathcal{R}_\delta \) observes \( Y_{i_{n_\delta + 1}}, \ldots, Y_n \), quantizes using \( \tau_{i_{n_\delta + 1}}, \ldots, \tau_n \), and decides based on \( U_{i_{n_\delta + 1}}, \ldots, U_n \). The Bayes error probabilities for these two systems are denoted by \( \gamma_n(\pi) \) and \( \gamma_n^\ast(\pi) \). Since the vectors \( (U_{i_1}, \ldots, U_{i_n}) \) and \( (U_{i_{n_\delta + 1}}, \ldots, U_n) \) are independent, we can apply Lemma 7 to obtain
\[ \gamma_n^\ast(\pi) \geq 2 \gamma_n(\pi) \gamma_{n_\delta}(1/2). \]  
(5.2)

We examine \( \gamma_n(\pi) \) and \( \gamma_n^\ast(\pi) \) separately. By Lemmas 4 and 5, every LRP in \( \mathcal{S}_\delta(\delta) \) satisfies conditions C1 and C2 in the hypothesis of Lemma 1; hence, by Corollary 1,
\[ \lim_{n \to \infty} \gamma_n(\pi) n_\delta \exp(n_\delta \rho_m) > 0. \]  
(5.3)

Step 2. Conditioning in \( \mathcal{S}_\delta \): To lower-bound \( \gamma_n(\pi) \), we need consider only LRP’s in \( \mathcal{S}_\delta(\delta) \). In this case, Lemma 1 does not necessarily apply, nor is the weaker boundedness assumption (4.9) (which was used in the lower bound in [20]) valid in general. To remove these obstacles, we condition on the event \( \Delta \) as specified below. For each \( i \geq n_\delta + 1 \), we let
\[ J_i = \{ u \in \mathcal{U}_m; P_{\tau_i}(u) \land Q_{\tau_i}(u) \geq \delta \}. \]

Furthermore, we let \( \Delta_i \) be the event that \( X_{\tau_i} \) lies in \( J_i \), and define
\[ \Delta = \bigcap_{i = n_\delta + 1}^{n} \Delta_i. \]

From (4.1), we obtain the lower bound
\[ \gamma_n(\pi) \geq \pi P \left( \sum_{i = n_\delta + 1}^{n} X_{\tau_i} \leq \eta \Bigg| \Delta \right) \]
\[ + (1 - \pi) Q \left( \sum_{i = n_\delta + 1}^{n} X_{\tau_i} > \eta \Bigg| \Delta \right) \left( P(\Delta) \land Q(\Delta) \right) \]
where \( P \) and \( Q \) represent product measures by the convention of Section II. To estimate \( P(\Delta) \land Q(\Delta) \), we recall that since the measures \( P \) and \( Q \) on the marginal space \( (\mathcal{Y}, \mathcal{B}) \) are mutually absolutely continuous, the inequality \( P(C) \land Q(C) < \delta \) also implies \( P(C) \lor Q(C) < \xi(\delta) \), where \( \xi(\delta) \to 0 \) as \( \delta \to 0 \). Therefore,
\[ P(\Delta_i) \land Q(\Delta_i) \geq 1 - m \xi(\delta), \]
and by independence of the \( \Delta_i \)'s,
\[ P(\Delta) \land Q(\Delta) \geq (1 - m \xi(\delta))^{n_\delta}. \]

We conclude that
\[ \gamma_n(\pi) \geq \pi P \left( \sum_{i = n_\delta + 1}^{n} X_{\tau_i} \leq \eta \Bigg| \Delta \right) \]
\[ + (1 - \pi) Q \left( \sum_{i = n_\delta + 1}^{n} X_{\tau_i} > \eta \Bigg| \Delta \right) \exp(-n_\delta \epsilon(\delta)) \]  
(5.4)

where \( \epsilon(\delta) \to 0 \) as \( \delta \to 0 \). Our aim is to show that the first factor on the RHS of (5.4) decreases at a rate which bounded away from the optimal \( \rho_m \).

Step 3. Restoration of \( \mathcal{S}_\delta \): Consider next a system \( \mathcal{S}_\delta \) with independent observations \( Y_{i_1}, \ldots, Y_{i_n} \) each \( Y_{i} \) takes values in the restriction \( \Delta_i \) of \( (\mathcal{Y}, \mathcal{B}) \) on \( \Delta_i \), with distributions given, for \( C \in \mathcal{B}_m \), by
\[ P_{\Delta}(C) = \frac{P(C)}{P(\Delta)} \quad \text{and} \quad Q_{\Delta}(C) = \frac{Q(C)}{Q(\Delta)}. \]

For all \( y \in \Delta_i \), (a version of) the likelihood ratio \( dp_{\Delta} / dq_{\Delta} \) is given by
\[ dp_{\Delta}(y) = \frac{Q(\Delta_i)}{P(\Delta_i)} \quad \text{and} \quad dq_{\Delta}(y), \]
and thus the log-likelihood ratio \( X_{\Delta} \) satisfies
\[ X_{\Delta}(y) = X(y) + \mu \]
where \( \mu \equiv \log \left[ Q(\Delta) / P(\Delta) \right] \).

The last relationship implies that if \( g: \mathcal{Y} \to \mathcal{U}_m \) is a deterministic LRQ, then so is its restriction \( g_{\Delta} \) to \( \Delta_i \), and that is
\[ \tau = (1, \ldots, 1) \]
is an LRP corresponding to \( g \), then
\[ \tau_{\Delta} = (1, \mu, \ldots, \mu, \mu) \]
is an LRP corresponding to \( g_{\Delta} \). In the special case \( \tau = (1, \mu, \ldots, \mu, \mu) \) (the only case of interest), the definition of \( \Delta_i \) implies that the output of the quantizer \( g_{\Delta} \) will be a.s. lie in the set \( J_i \) where, by the assumption \( \tau_i \in \mathcal{R}_\delta(\delta) \), the size of \( J_i \) is at most \( m - 1 \). Thus, w.l.o.g., the partition \( \tau_{\Delta} \) can be coarsened by absorbing all intervals with indexes not in \( J_i \) into intervals with indexes in \( J_i \), i.e., \( \tau_{\Delta} \) can be taken as a \( |J_i| \)-ary LRP. Writing \( P_{\Delta} \) and \( Q_{\Delta} \) for the distributions induced on \( J_i \) by this trimmed version of \( \tau_{\Delta} \), we have
\[ (\forall u \in J_i) \quad P_{\Delta}(u) = \frac{P(\Delta_i)}{P(\Delta)} \quad \text{and} \quad Q_{\Delta}(u) = \frac{Q(\Delta_i)}{Q(\Delta)} \]
and thus for all \( u \in I_i, \)
\[
P_{\tau_i}^a(u) \wedge Q_{\tau_i}^a(u) \geq \delta. \tag{5.5}
\]

(One could say that \( \tau_i^a \) lies in \( S_{\tau_i}^a \), where the latter class is defined w.r.t. the observation space \( (\Delta_i, \mathcal{B}_i) \) and the base measures \( P_{\tau_i}^a \) and \( Q_{\tau_i}^a \).) Furthermore, if \( X_{\tau_i}^a \) is the postquantization log-likelihood ratio corresponding to \( \tau_i^a \), then, for all \( x \in \mathbb{R}, \)
\[
P_{\tau_i}^a \left\{ X_{\tau_i}^a = x \right\} = P \left\{ X_{\tau_i} = x - \mu_i, |\Delta_i| \right\} \tag{5.6}
\]
and
\[
Q_{\tau_i}^a \left\{ X_{\tau_i}^a = x \right\} = Q \left\{ X_{\tau_i} = x - \mu_i, |\Delta_i| \right\}. \tag{5.7}
\]

**Step 4. Relevance of \( S_{\tau_i}^a \):** In \( S_{\tau_i}^a \), the events \( \{ X_{\tau_i} = x \} \cap \Delta_i \) are independent, and thus
\[
P \left( \bigcap_{i = n_\tau + 1}^n \{ X_{\tau_i} = x_i \} \right) = \prod_{i = n_\tau + 1}^n P \left( X_{\tau_i} = x_i | \Delta_i \right).
\]
The RHS can be rewritten in terms of (5.6); hence,
\[
P \left( \bigcap_{i = n_\tau + 1}^n \{ X_{\tau_i} = x_i \} \right) = \prod_{i = n_\tau + 1}^n P_{\tau_i}^a \left\{ X_{\tau_i} = x_i + \mu_i \right\},
\]
and the same holds by (5.7) with \( Q \) replacing \( P \). Thus, letting \( \eta_\tau = \eta + \sum_{i = n_\tau + 1}^n \mu_i \) and defining product measures \( P = P_{\tau_1}^a \times \cdots \times P_{\tau_n}^a \) and \( Q = Q_{\tau_1}^a \times \cdots \times Q_{\tau_n}^a \), we obtain
\[
\pi P \left( \sum_{i = n_\tau + 1}^n X_{\tau_i} \leq \eta \Delta \right) + (1 - \pi)Q \left( \sum_{i = n_\tau + 1}^n X_{\tau_i} > \eta \Delta \right)
= \pi P^{\Delta} \left( \sum_{i = n_\tau + 1}^n X_{\tau_i} \leq \eta \Delta \right)
+ (1 - \pi)Q^{\Delta} \left( \sum_{i = n_\tau + 1}^n X_{\tau_i} > \eta \Delta \right).
\]
The RHS of this equation is the overall probability of error in \( S_{\tau_i}^a \) when the fusion center performs a log-likelihood ratio test threshold with test threshold \( \eta_\tau \). This is clearly less than the Bayes error probability \( \gamma_{\eta_\tau}^a (\tau) \), which is obtained by resetting \( \eta \) to \( \eta = \log \left( \frac{1 - \pi}{\pi} \right) \). Thus, in conjunction with (5.4), we conclude that
\[
\gamma_{\eta_\tau}^a (\tau) \geq \exp \left\{ - n_\tau \varepsilon_{\eta_\tau} (\delta) \right\} \gamma_{\eta_\tau}^a (\pi). \tag{5.8}
\]

**Step 5. Error Performance of \( S_{\tau_i}^a \):** We will use the results of Section IV to obtain a lower bound for \( \gamma_{\eta_\tau}^a (\tau) \). In what follows, the entities \( \Phi_{\tau_i}^a (\cdot), M_{\tau_i}^a (\cdot), s_{\tau_i}^a, \) and \( F_{\tau_i}^a \) will be defined—with obvious modifications—as in Section IV-A.

The idea is to adapt Lemma 1 and its corollary to \( S_{\tau_i}^a \).

By virtue of (5.5), each \( X_{\tau_i}^a \) is absolutely bounded by \( \log (1/\delta) \), and thus condition C1 in the hypothesis of Lemma 1 is satisfied, with \( X_{\tau_i}^a \) replacing \( X_{\tau_i} \). Condition C2 can be checked as follows.

Let \( i \) be such that \( |J_i| \geq 2 \), and take \( r > 0 \) as in the proof of Lemma 5:
\[
Q \left( X \leq x \right) \geq r \quad \text{and} \quad Q \left( X \geq x' \right) \geq r.
\]

Without loss of generality, we can take \( \delta \) sufficiently small so that \( m \xi (\delta) \leq r/2 \). Then,
\[
Q_i^a \left( X_{\tau_i}^a \leq x + \mu_i \right) = \frac{Q \left( \{ X_{\tau_i} \leq x \} \cap \Delta_i \right)}{Q (\Delta_i)} \geq \frac{r}{2Q (\Delta_i)} \geq \frac{r}{2}
\]
uniformly in \( i \), and similarly, \( Q_i^a \left( X_{\tau_i}^a \geq x' \right) \geq r/2 \). Retracing the proof of the lemma with \( r/2 \) replacing \( r \), we obtain
\[
\left| X_{\tau_i}^a (J_i) - X_{\tau_i}^a (1) \right| \geq b (x + \mu_i, x' + \mu_i, r/2) > 0.
\]
Continuing the argument using (5.5), we conclude that
\[
\text{Var}_{\delta} \left( X_{\tau_i}^a \right) \geq b^2 (x + \mu_i, x', \mu_i, r/2) \frac{\delta}{2} > 0.
\]

Thus, condition C2 is satisfied by \( X_{\tau_i}^a \) provided \( i \) is such that \( |J_i| \geq 2 \). Let \( n_{\tau_i} \) be the number of such indexes \( i \), and observe that the remaining \( n_{\tau_i} = n_{\tau_i}^a \) indexes are of no interest because \( X_{\tau_i}^a = 0 \) a.s. when \( |J_i| = 1 \). We can thus apply Lemma 1 to conclude that
\[
\liminf_{n \to \infty} {\sqrt{n}} \int_{-\infty}^{\infty} \exp \left\{ - |x| \right\} d\tilde{F}_{\tau_i}^a (x)
\]
\[
\geq \liminf_{n \to \infty} {\sqrt{n}} \int_{-\infty}^{\infty} \exp \left\{ - |x| \right\} d\tilde{F}_{\tau_i}^a (x) > 0 \tag{5.9}
\]
where \( \tilde{F}_{\tau_i}^a (\cdot) = \tilde{F}_{\tau_i}^a (\cdot) \) by the earlier remark.

In order to combine (5.9) with (4.8), we need a lower bound on \( M_{\tau_i}^a (s_{\tau_i}) \). We obtain this by writing, as in (4.2),
\[
M_{\tau_i}^a (s_{\tau_i}) \geq \prod_{i = n_\tau + 1}^{n_{\tau_i}} \min \Phi_{\tau_i}^a (s_i), \tag{5.10}
\]
and then by examining each \( \Phi_{\tau_i}^a (s_i) \) separately. We have, for \( s \in (0, 1), \)
\[
\Phi_{\tau_i}^a (s) = \sum_{u \in J_i} \left[ P_{\tau_i}^a (u) \right] \left[ Q_{\tau_i}^a (u) \right]^{-1} -
\]
\[
\geq \sum_{u \in J_i} \left[ P_{\tau_i}^a (u) \right] \left[ Q_{\tau_i}^a (u) \right]^{-1} -
\]
\[
\geq \Phi_{\tau_i} (s) - \sum_{u \in J_i} \left[ P_{\tau_i}^a (u) \right] \left[ Q_{\tau_i}^a (u) \right]^{-1} -
\]
\[
\geq \Phi_{\tau_i} (s) - \sum_{u \in J_i} \left[ P_{\tau_i} (u) \right] \left[ Q_{\tau_i} (u) \right] -
\]
\[
\text{where} \ v \cdot \delta \to 0 \text{ as} \ \delta \to 0. \text{ Since} \ \tau_i \in \mathcal{R}_\delta (\delta), \text{we can involve Lemma 6 to obtain}
\]
\[
\Phi_{\tau_i}^a (s) \geq \exp \left\{ - \rho_{m - 1} - \varepsilon_\delta (\delta) \right\}
\]

where, again, \( \epsilon_{n}(\delta) \to 0 \) as \( \delta \to 0 \). This holds for all \( s \in (0,1) \), and consequently, (5.8) yields

\[
M_{n_{A}}(x_{n_{A}}) \geq \exp \{-n_{B}(\rho_{m_{-1}} + \epsilon_{n}(\delta))\}.
\]

Using this lower bound on \( M_{n_{A}}(x_{n_{A}}) \) in conjunction with (5.9) and (4.8), we obtain

\[
\lim_{n \to \infty} \gamma_{n_{A}}^{+}(\pi) \sqrt{n_{B}} \exp \{n_{B}(\rho_{m_{-1}} + \epsilon_{n}(\delta))\} > 0. \tag{5.11}
\]

Step 6. Conclusion: In light of (5.8) and (5.11), we have

\[
\lim_{n \to \infty} \gamma_{n_{A}}^{+}(\pi) \sqrt{n_{B}} \exp \{n_{B}(\rho_{m_{-1}} + \epsilon_{n}(\delta))\} > 0 \tag{5.12}
\]

with \( \epsilon_{n}(\delta) \to 0 \) as \( \delta \to 0 \). Since the error exponent obtained by \( \mathcal{F}_{n_{A}} \) is at best only marginally greater than \( \rho_{m_{-1}} \), system \( \mathcal{F}_{n_{A}} \) is markedly inferior to the best identical quantizer of the same size. Hence, it is natural to expect that the size \( n_{B} \) of \( \mathcal{F}_{n_{A}} \) is bounded in \( n \). As we will see, this is indeed true, and implies that \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \) is bounded from below.

Combining (5.2) with (5.3) and (5.12), we obtain

\[
\lim_{n \to \infty} \frac{\gamma_{n_{A}}^{+}(\pi) \sqrt{n_{B}}}{\gamma_{n_{A}}^{-}(\pi)} \exp \{n_{B}(\rho_{m_{-1}} + \epsilon_{n}(\delta))\} > 0
\]

where we assume that \( n_{B} \geq 1 \). By (4.12), the quantity \( \gamma_{n_{A}}^{+}(\pi) \sqrt{n_{B}} \exp \{n_{B}(\rho_{m_{-1}} + \epsilon_{n}(\delta))\} \) is bounded from above; hence, there exists \( c > 0 \) such that

\[
\lim_{n \to \infty} \frac{\gamma_{n_{A}}^{+}(\pi)}{\gamma_{n_{A}}^{-}(\pi)} \geq \frac{1}{n_{A}} \sqrt{\frac{1}{n_{B}} + \frac{1}{n_{B}}}
\]

\[
\cdot \exp \{n_{B}(\rho_{m_{-1}} - \epsilon_{n}(\delta))\}.
\]

By Lemma 3, the difference \( \rho_{m_{-1}} - \epsilon_{n}(\delta) \) can be made larger than \( \rho_{m_{-1}} - \epsilon_{n}(\delta)/2 \) by choosing \( \delta \) suitably small, and thus,

\[
\lim_{n \to \infty} \frac{\gamma_{n_{A}}^{+}(\pi)}{\gamma_{n_{A}}^{-}(\pi)} \geq \frac{c}{\sqrt{n_{B}}} \exp \{n_{B}(\rho_{m_{-1}} - \epsilon_{n}(\delta))/2\} \tag{5.13}
\]

where the RHS is clearly bounded away from zero for all \( n_{B} \geq 1 \). This proves the statement of the theorem. Observe further that, although the RHS of (5.13) increases (in \( n_{B} \)) to infinity, we always have \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \leq 1 \); thus, \( n_{B} \) must also be bounded from above.

The case \( n_{B} = 0 \) is straightforward. Indeed, (5.1) follows directly from (5.3) and (4.12).

Theorem 5 refines the equality of error exponents \( \epsilon_{n}(\pi) \) and \( \epsilon_{n}(\pi) \) by showing that the actual ratio \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \) is bounded away from zero provided \( \rho = Q \). No further conditions, such as the boundedness assumption (4.9) used in [20], are needed. The latter fact was indeed conjectured in [19], together with the possibility that the optimal LRO's in \( \mathcal{F}_{n_{A}} \) employ thresholds that are confined in a fixed (as \( n \to \infty \)) finite interval. We do not have proof of the second—and quite stronger—conjecture; our proof only establishes that the number of LRO's with one or more thresholds outside a certain fixed interval \([-T(\delta), T(\delta)]\) must be bounded in \( n \).

Also related to the composition of an optimal set \( \tau_{1}, \ldots, \tau_{n} \) of LRQ's is the following observation.

Corollary 2. For \( \epsilon > 0 \), the number \( K_{n}(\epsilon) \) of optimal LRQ's \( \tau_{i} \) such that

\[
\rho(P_{\tau_{i}}, Q_{\tau_{i}}) < \rho_{m} - \epsilon \tag{5.14}
\]

is bounded.

Proof: Suppose that, in the proof of Theorem 5, \( n_{B} \) of the first \( n_{A} \) quantizers satisfy (5.14). Then, (5.3) can be strengthened into

\[
\lim_{n \to \infty} \gamma_{n_{A}}^{+}(\pi) \sqrt{n_{B}} \exp \{n_{B}(\rho_{m} - n_{B}\epsilon)\} > 0,
\]

in which case (5.13) becomes

\[
\lim_{n \to \infty} \frac{\gamma_{n_{A}}^{+}(\pi)}{\gamma_{n_{A}}^{-}(\pi)} \geq \frac{c}{\sqrt{n_{B}}} \exp \{n_{B}[(\rho_{m} - \rho_{m-1})/2] + n_{B}\epsilon\}.
\]

From \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \leq 1 \), it follows again that \( n_{B}, \rho_{m} \), and \( K_{n}(\epsilon) \) (which is at most \( n_{B} \)) are bounded. The argument is easily adapted to the case \( n_{B} = 0 \).

The significance of this result becomes transparent when the sample space \( \mathcal{S} \) is finite. In that case, there are at most \( \binom{m}{m - 1} \) deterministic LRQ's, and thus only finitely many pairs \( (P_{\tau_{i}}, Q_{\tau_{i}}) \) of output distributions. Clearly, we can choose \( \epsilon \) small enough so that \( n - K_{n}(\epsilon) \) is the exact number of quantizers in the system that achieve \( \rho_{m} \). Then by Corollary 2, the number of remaining quantizers (the ones that do not achieve \( \rho_{m} \)) must be bounded. If, in particular, there is a unique output distribution pair that achieves \( \rho_{m} \), then an optimal system exists in which the same quantizer is used by all but a bounded number of sensors.

As we pointed out in Section I, our numerical results have often indicated that the ratio \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \) is very close to unity. It is thus possible that under certain conditions as yet unknown to us, the ratio \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \) tends to unity as \( n \) approaches infinity. Nevertheless, we have a counterexample (appearing in Appendix B) which shows that this cannot be true in general. In that example, the ratio \( \gamma_{n_{A}}^{+}(\pi)/\gamma_{n_{A}}^{-}(\pi) \) is smaller than a constant \( r < 1 \) infinitely often in \( n \).

VI. CONCLUDING REMARKS

Our investigation of error exponents in the absence of boundedness conditions [such as Assumption 3 and (4.9)] has yielded interesting and illuminating results: notably Theorem 4 on the dependence of the Neyman–Pearson error exponent on the test level \( \alpha \), and Theorem 5, which affirms the previously conjectured redundancy of the boundedness assumption in Bayes testing.

The main conclusion of Theorem 5 is that the performance ratio between the absolutely optimal and best identical quantizer systems is bounded. This result, in conjunction with Corollary 2, shows that the degree of asymptotic equivalence of these two systems is far greater.
than what is implied by the equality of error exponents \( e_g^*(\sigma) \) and \( e_g^*(\tau) \). A stronger version of Theorem 5 could include a lower bound on \( \gamma^*_g(\sigma)/\gamma^*_g(\tau) \); although this was not attempted in this paper (the proof of Theorem 5 is already too technical), this refinement would be quite welcome.

Following a reviewer’s suggestion, we have ascertained that Theorem 5 also holds for \( L \)-ary hypothesis testing, where \( L > 2 \). In that case, it is known from [20, Theorem 1] that there exists an asymptotically exponentially optimal system \( \mathcal{X}^* \) employing— in fixed proportions as \( n \) varies—at most \( L(L-1)/2 \) fixed quantizers. We have obtained a variant of Lemma 1 based on a multidimensional Berry–Esseen theorem [3, Corollary 17.2], and proceeded to modify the proof of Theorem 5 so as to consider the \( L(L-1)/2 \) binary likelihood-ratio comparisons involved in \( L \)-ary testing (details omitted).

The apparatus developed in Section IV is not suited to the detailed asymptotic analysis of Neyman–Pearson testing. We will report on this problem in a forthcoming paper.

**Appendix A**

**Proof of Lemma 5**

Assumption 2 of Section II implies that the prequantization log-likelihood ratio \( X \) is not almost surely constant. Thus, there exist \( x \in \mathbb{R} \), \( x' > 0 \), and \( r > 0 \) such that

\[
Q(X \leq x) \geq r \quad \text{and} \quad Q(X \geq x') \geq r.
\]  

(1.1)

Now, let \( m > 2, \delta > 0 \), and consider an arbitrary LRP \( \tau = (I_1, \ldots, I_m) \) in \( \mathcal{F}_\delta \). The values of \( P_X \) and \( Q_X \) here are all greater than zero, so the equation \( X(u) = \log[P_X(u)/Q_X(u)] \) is meaningful. We always have

\[
X(1) \geq 0, \quad X(m) \geq 0.
\]  

(1.2)

In what follows, we will find a nontrivial lower bound to \( X(m) - X(1) \) in terms of \( x, x', \) and \( r \). For that purpose, we decompose \( I_1 \) into \( A = I_1 \cap (-\infty, x) \) and \( B = I_1 \cap (-\infty, x) \), and denote the \( (P, Q) \)-probabilities of these intervals by \( (p_A, q_A) \) and \( (p_B, q_B) \), respectively. Also, we fix a point \( t \) inside the interior of \( I_1 \) \( \cup \) \( I_m \).

Clearly, \( p_A \leq p_A \exp x \) and \( p_B \leq q_A \exp t \), and thus

\[
\exp \{X(1)\} = \frac{p_A + p_B}{q_A + q_B} \leq \frac{q_A \exp x + q_B \exp t}{q_A + q_B}.
\]

(1.3)

\[
= \frac{q_A}{q_A + q_B} \exp x + \frac{q_B}{q_A + q_B} \exp t.
\]

If \( x \not\in I_1 \), then the RHS equals \( e^x \), and by virtue of (1.2), we have

\[
X(m) - X(1) \geq -X(1) = -x > 0.
\]  

(1.4)

Otherwise, if \( x \in I_1 \), the bound is a proper mixture of \( \exp x \) and \( \exp t \) (greater than \( \exp x \)), whose value increases in \( q_B/q_A \). Noting that \( q_B/q_A < (1 - r)/r \) by (1.1), we obtain

\[
X(1) \leq \log \{ re^t + (1 - r)e^t \}.
\]  

(1.5)

An identical argument of \( I_m \) yields, for \( x' \in I_m \),

\[
X(m) - X(1) \geq X(m) - x' > 0;
\]  

(1.6)

and for \( x' \in I_m \),

\[
X(m) \geq \log \{ re^t + (1 - r)e^t \}.
\]  

(1.7)

It remains to find a lower bound to \( X(m) - X(1) \) when both \( x \in I_1 \) and \( x' \in I_m \). Using (1.4) and (1.6), we obtain for this case

\[
X(m) - X(1) \geq \log \frac{re^t + (1 - r)e^t}{re^t + (1 - r)e^t}.
\]  

(1.8)

Since \( t \in [x, x'] \) and the RHS is decreasing in \( t \), we can set \( t = x' \) to obtain

\[
X(m) - X(1) \geq -\log \frac{re^t}{re^t + (1 - r)e^t} > 0.
\]  

(1.9)

From (1.3), (1.5), and (1.8), we conclude that \( |X(m) - X(1)| \geq b(x, x', r) > 0 \), and that the bound is independent of \( \delta \) (we do, however, need \( \delta > 0 \)). The remainder is straightforward. We write for simplicity \( p_A = p_A(u) \) and \( q_A = Q_A(u) \), and note that the tilted distribution with parameter \( s \in (0, 1) \) satisfies

\[
\tilde{\Phi}_s(X(1)) = \frac{p_A^s q_A^{1-s}}{\sum_{i=1}^m p_i^s q_i^{1-s}}
\]

where the denominator of the rightmost term is just \( \Phi_s(x) \). Since for \( s \in (0, 1) \), \( \Phi_s(x) \leq 1 \) and \( p_i^{s} q_i^{1-s} \geq \delta \), we conclude that both \( \tilde{\Phi}_s(X(1)) \geq \delta \) and \( \tilde{\Phi}_s(X(m)) \geq \delta \). Therefore, \( \text{Var}_s[X_1] \geq \delta b^2(x, x', r)/2 \geq 0 \).

**Proof of Lemma 6**

The quantity

\[
r(\delta) = \sup_{\tau \in \mathcal{F}_\delta} \rho(P_\tau, Q_\tau)
\]

decreases together with \( \delta \), and thus the limit

\[
r^\vee = \lim_{\delta \to 0} r(\delta)
\]

is well defined.

First, we observe that \( r(\delta) \geq r_{\delta, 1} \) for every \( \delta > 0 \). Indeed, any LRP \( \tau_{\delta, 1} \) in \( \mathcal{F}_\delta \) that achieves \( r_{\delta, 1} \) can be refined to an LRP \( \tau_{\delta, t} \) in \( \mathcal{F}_\delta \) by splitting the rightmost interval of \( \tau_{\delta, 1} \) at a point \( t \) such that \( P(X \geq t) < \delta \). It follows easily from Lemma 2 that \( \rho(P_{\tau_{\delta, t}}, Q_{\tau_{\delta, t}}) \geq \rho(P, Q_{\tau_{\delta, t}}) \), and thus also \( r(\delta) \geq r_{\delta, 1} \).

It remains to show that \( r_{\delta, 1} \geq r \). For every \( \delta > 0 \), we choose a \( \tau_\delta = \tau_\delta(\delta) \) in \( \mathcal{F}_\delta \) such that

\[
\rho(P_{\tau_\delta}, Q_{\tau_\delta}) = -\log \Phi_{\tau_\delta} \geq r(\delta) - \epsilon
\]

(1.9)

where \( \epsilon > 0 \). Without loss of generality, we can assume that \( \tau_\delta = \frac{1}{\delta} \delta \) lies in a fixed (as \( \delta \) varies) closed interval \([a, b] \), where \( 0 < a < b < 1 \). This is because by Lemma 2 and the convexity of \( \Phi(s) \), we can find \( a = a(s) > 0 \) and \( b = b(s) < 1 \) such that for all \( s \in [0, a(e)] \cup [b(e), 1] \),

\[
1 \geq \Phi(s) \geq \Phi(s) \geq \exp \{-\epsilon s\}.
\]
Therefore, if \( s_i \) lies in \([0, a] \cup [b, 1]\), it must satisfy
\[
\exp \varepsilon \leq \frac{\Phi_{s_i}(s_i)}{\Phi_{s_i}(a)} \geq \exp(-\varepsilon),
\]
and thus also by (1.9),
\[
-\log \Phi_{s_i}(s_i) \geq r(\delta) - 2\varepsilon. \tag{1.10}
\]
We can thus replace \( s_i \) by \( a \) if \( s_i \notin [a, b] \).

In the usual notation, let \( P_i = (p_{i1}, \cdots, p_{iL}) \) and \( Q_i = (q_{i1}, \cdots, q_{iL}) \) be the output probability distributions generated by \( r_i = (I_i, \cdots, I_i) \). Then, there exists at least one index \( u \) for which \( p_{iu} \wedge q_{iu} < \delta \). Since \( P \) and \( Q \) are mutually absolutely continuous, \( P(A) \wedge Q(A) \leq \delta \) implies \( P(A) \vee Q(A) \leq \xi(\delta) \), where \( \xi(\delta) \to 0 \) as \( \delta \to 0 \). Thus, we also have \( p_{iu} \vee q_{iu} < \xi(\delta) \).

Taking w.l.o.g. the said index as \( u = 1 \), we can merge \( I_1 \) with \( I_i \) to obtain an LRP \( r_{i-1} \in \mathcal{F}_{l-1} \). This new \((l-1)\)-ary LRP is equivalent to the \( l \)-ary LRP
\[
r'_1 = (I_1, I_2, \cdots, I_{l-1}, I_{l}),
\]
and thus \( \Phi_{s_i}(s_i) \) for all \( s_i \). Also, observe that the pairs \((P_i, Q_i)\) and \((P_{i-1}, Q_{i-1})\) differ by no more than \( \xi(\delta) \) in any coordinate.

Now, when \( \tau \in \mathcal{F}_l \), the function \( \Phi_{s_i}(s_i) \) is obtained from
\[
\Psi(s; \lambda, \mu, \lambda) = \sum_{s} \lambda^m \mu^{n-m}
\]
by setting \( \lambda = P_i \) and \( \mu = Q_i \). It is straightforward to show that \( \Psi(s; \lambda, \mu) \) is continuous on \([a, b] \times [0, 1]^L \) provided \( 0 < a < b < 1 \). Recalling that \( \Phi_{s_i}(s_i) \) is bounded from below by \exp(-\rho(P, Q)) > 0, we conclude that
\[
\log \Phi_{s_i}(s_i) \leq (1/2) \sum_{u \in I} \log P_i(u)-1 \log Q_i(u)
\]
can be smaller than \( \varepsilon \) uniformly in \( s \in [a, b] \) by taking \( \delta \) small.

By virtue of (1.10), we now have \( p_{i-1} > r(\delta) - 3\varepsilon \geq r_i \), and since \( \varepsilon \) was arbitrary, we conclude that \( p_{i-1} \geq r_i \).

**APPENDIX B**

A Counterexample to \( \gamma^a(\pi) / \gamma^c(\pi) \to 1 \)

Consider a ternary observation space \( \mathcal{Y} = \{a_1, a_2, a_3\} \) with binary quantization. The two hypotheses are assumed equally likely, with
\[
\begin{align*}
    y & a_1 & a_2 & a_3 \\
    P(y) & 1/12 & 1/4 & 2/3 \\
    Q(y) & 1/3 & 1/3 & 1/3 \\
    (dP/dQ)(y) & 1/4 & 3/4 & 2.
\end{align*}
\]

There are only two nontrivial deterministic LRQ's: \( \mathcal{G} \), which partitions \( \mathcal{Y} \) into \([a_1] \) and \([a_2, a_3] \), and \( \mathcal{G} \), which partitions \( \mathcal{Y} \) into \([a_1, a_2] \) and \([a_3] \). The corresponding output distributions and log-likelihood ratios are given by
\[
\begin{align*}
    u & 1 & 2 \\
    P(u) & 1/12 & 11/12 \\
    Q(u) & 1/3 & 2/3 \\
    X(u) & -\log 4 & \log (11/8) \\
    u & 1 & 2 \\
    P(u) & 1/3 & 2/3 \\
    Q(u) & 2/3 & 1/3 \\
    X(u) & -\log 2 & \log 2.
\end{align*}
\]

Direct computation shows that
\[
\rho(P, Q) = \frac{9}{8} = 0.5098 > 0.5354 = \rho(P, Q),
\]
and thus for \( n \) sufficiently large, the best identical-quantizer system \( \mathcal{S}_n^a \) employs \( \mathcal{G} \) on all \( n \) sensors [the other choice yields a suboptimal error exponent, and thus eventually a higher value of \( \gamma(\pi) \)].

We will now show by contradiction that if \( \mathcal{S}_n^a \) is an absolutely optimal system consisting of deterministic LRO's, then for all even values of \( n \), at least one of the quantizers in \( \mathcal{S}_n^a \) must be \( \mathcal{G} \).

Assume the contrary, i.e., \( \mathcal{S}_n^a \) is such that for all \( i \leq n \), \( g_i = \mathcal{G} \). Now, consider the problem of optimizing the quantizer \( g_n \) in \( \mathcal{S}_n^a \) subject to the constraint that each of the remaining quantizers \( g_{i+1}, \cdots, g_{n-1} \) equals \( \mathcal{G} \). From the discussion in Section II-D, we know that each of the remaining quantizers \( g_{i+1}, \cdots, g_{n-1} \) equals \( \mathcal{G} \). From the discussion in Section II-D, we know that each of the remaining quantizers \( g_{i+1}, \cdots, g_{n-1} \) equals \( \mathcal{G} \). From the discussion in Section II-D, we know that each of the remaining quantizers \( g_{i+1}, \cdots, g_{n-1} \) equals \( \mathcal{G} \).

To see why this cannot be so if \( n \) is even, consider the Bayes error probability for this problem. Writing \( u_n^l \) for \((u_1, \cdots, u_n)\), we have
\[
\gamma_n = \frac{1}{2} \sum_{u \in \{1, 2\}^n} [P_n(u_n-1)Q_n(u_n{-1})] = 1 \sum_{u \in \{1, 2\}^n} [P_n(u_n-1)Q_n(u_n{-1})] = \sum_{u \in \{1, 2\}^n} \left[ \frac{P_n(u_n{-1}) + Q_n(u_n{-1})}{2} \right] \gamma_n = \left( \frac{P_n(u_n{-1}) + Q_n(u_n{-1})}{2} \right) \gamma_n \tag{2.2}
\]
where \( \gamma(\cdot) \) represents the Bayes error probability function of the \( n \)th sensor/quantizer pair [note that in this equation, the argument of \( \gamma(\cdot) \) is just the posterior probability of \( H_n \) given \( u_n{-1} \)].

We note from (2.1) that the log-likelihood ratio
\[
\frac{P_n(u_n{-1})}{Q_n(u_n{-1})} = X_n(u_n{1}) + \cdots + X_n(u_n{-1})
\]
can also be expressed as \((2l_{n-1}(u) - n + 1)(\log 2)\), where \( l_{n-1}(u) \) is the number of 2's in \( u_n{-1} \). Now, \( l_{n-1}(U) \) is a
binomial variable under either hypothesis, and we can rewrite (2.2) as
\[
\gamma(\frac{1}{2}) = \sum_{l=0}^{n-1} \binom{n-1}{l} \left( \frac{2^{l} + 2^{n-1-l}}{2 \cdot 3^{n-1}} \right) \gamma\left( \frac{2^{2l-n+1} + 1}{2^{2l-n+1} + 1} \right)
\]
(2.3)

The two candidates for \( \gamma \) and \( \hat{\gamma} \) and \( \tilde{\gamma} \), given by
\[
\hat{\gamma}(\pi) = \left[ \frac{1}{12} \pi + \frac{1}{3}(1 - \pi) \right] + \left[ \frac{11}{12} \pi + \frac{2}{3}(1 - \pi) \right]
\]
\[
\tilde{\gamma}(\pi) = \left[ \frac{1}{3} \pi + \frac{2}{3}(1 - \pi) \right] + \left[ \frac{2}{3} \pi + \frac{1}{3}(1 - \pi) \right]
\]
and shown in Fig. 3. Note that \( \hat{\gamma}(\pi) = \tilde{\gamma}(\pi) \) for \( \pi \leq 1/3 \), \( \pi = 4/7 \), and \( \pi \geq 4/5 \). Thus, the critical values of \( l \) in (2.3) are those for which \( 2^{2l-n+1}/(2^{2l-n+1} + 1) \) lies in the union of \( (1/3, 4/7) \) and \( (4/7, 4/5) \).

If \( n \) is odd, then the range of \( 2l-n+1 \) in (2.3) comprises the even integers between \(-n+1\) and \( n-1 \) inclusive. The only critical value of \( l \) is \( (n-1)/2 \), for which the posterior probability of \( H_0 \) is 1/2. Since \( \tilde{\gamma}(1/2) = 3/8 > 1/3 = \tilde{\gamma}(1/2) \), the optimal choice is \( \tilde{\gamma} \).

If \( n \) is even, then \( 2l-n+1 \) ranges over all odd integers between \( -n+1 \) and \( n-1 \) inclusive. Here, the only critical value of \( l \) is \( n/2 \), which makes the posterior probability of \( H_0 \) equal to 2/3. Since \( \tilde{\gamma}(2/3) = 5/18 < 1/3 = \tilde{\gamma}(2/3) \), \( \tilde{\gamma} \) is optimal.

We thus obtain the required contradiction, together with the inequality
\[
\gamma^{*}_{\frac{k}{2}} \left( \frac{1}{2} \right) - \gamma^{*}_{\frac{k}{2}} \left( \frac{1}{2} \right) \geq \left( \frac{2k-1}{k} \right) \left( \frac{2^{k} + 2^{k-1}}{2 \cdot 3^{k-1}} \right) \left[ \tilde{\gamma}\left( \frac{2}{3} \right) - \hat{\gamma}\left( \frac{2}{3} \right) \right]
\]
\[
= \frac{1}{8} \left( \frac{2k-1}{k} \right) \left( \frac{2^{k}}{9} \right).
\]

Stirling's formula gives \( (4\pi k)^{-1/2} \exp\{-k \log 4\}(1 + o(1)) \) for the binomial coefficient, and thus
\[
\lim_{k \to \infty} \gamma^{*}_{\frac{k}{2}} \left( \frac{1}{2} \right) - \gamma^{*}_{\frac{k}{2}} \left( \frac{1}{2} \right) \geq \frac{1}{16}.
\]
(2.4)

Since \( (9/8)^{k} = \exp\{2k p(P_{1}, Q_{2})\} = \exp\{2k \rho_{2}\} \), we immediately deduce from (2.4) and (4.12) that
\[
\lim_{k \to \infty} \sup \gamma^{*}_{\frac{k}{2}}(1/2) \leq \frac{23}{24}.
\]

From (2.4), we then obtain
\[
\lim_{k \to \infty} \sup \gamma^{*}_{\frac{k}{2}}(1/2) \leq \frac{23}{24}
\]

REFERENCES


