[16] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inform. Theory,* vol. 37, no. 5, pp. 1412–1418, Sept. 1991.
[17] G. D. Forney, Jr., "Dimension/length profiles and trellis complexity of linear block codes," *IEEE Trans. Inform. Theory,* vol. 40, no. 6, pp. 1741–1752, Nov. 1994.
[18] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977.
[19] P. E. Allard, S. G. S. Shiva, and S. E. Tavares, "A note on the decomposition of cyclic codes into cyclic classes," *Inform. Contr.,* vol. 22, no. 1, pp. 100–106, Feb. 1973.
[20] E. R. Berlekamp, "The technology of error-correcting codes," *Proc. IEEE,* vol. 68, no. 5, pp. 564–593, May 1980.

General Formulas for the Neyman–Pearson Type-II Error Exponent Subject to Fixed and Exponential Type-I Error Bounds

Po-Ning Chen, *Member, IEEE*

Abstract—**The general formulas for the Neyman–Pearson type-II error exponent subject to two different type-I error constraints, as indicated in the title of the correspondence, are established. As revealed in the formulas, the type-II error exponents are fully determined by the ultimate statistical characteristic of the normalized log-likelihood ratio evaluated under the null hypothesis distribution. Applications of the general formulas to distributed Neyman–Pearson detection, and the channel reliability function are also demonstrated.**

Index Terms— **Neyman–Pearson hypothesis testing, error exponent, distributed detection, channel reliability function.**

I. Introduction

In this correspondence, we consider a Neyman–Pearson hypothesis testing problem for testing a null hypothesis $H_0 : P_{X^n}$ against an alternative hypothesis $H_1 : Q_{X^n}$ based on a sequence of random observations $X^n = (X_1, \cdots, X_n)$, which is supposed to exhibit a probability distribution of either P_{X^n} or Q_{X^n}. Upon receipt of the n observations, a final decision about the nature of the random observations is made so that the type-II error probability β_n, subject to an upper bound on the type-I error probability α_n, is minimized.[1]

Two different type-I error bounds will be considered: a constant bound ε and an exponential bound e^{-nr} for $r > 0$ fixed. These two situations cover a fairly general class of problems in simple hypothesis testing, including the Bayes testing which stipulates minimization of the weighted sum of α_n and β_n.

For most cases of interest, the minimum type-II error probabilities $\beta_n^*(\varepsilon)$ and $\beta_n^*(e^{-nr})$ (as functions of the type-I error bound) vanish at an exponential rate as n tends to infinity. It thus becomes legitimate to adopt a measure of asymptotic performance based on the *error exponents*. A simple example is that the data source is independent and identically distributed (i.i.d.) given each hypothesis, i.e.

$$P_{X^n} = P_{X_1} \times \cdots \times P_{X_n} \triangleq P_X^n$$

and

$$Q_{X^n} = Q_{X_1} \times \cdots \times Q_{X_n} \triangleq Q_X^n.$$

In such case,

$$(\forall \varepsilon \in (0,1)) \quad \lim_{n\to\infty} -\frac{1}{n} \log \beta_n^*(\varepsilon) = D(P_X \| Q_X) \triangleq \int \log \frac{P_X(x)}{Q_X(x)}\, dP_X(x) \quad (1)$$

and[2]

$$(\forall D(Q_X\|P_X) > r > 0) \quad \lim_{n\to\infty} -\frac{1}{n} \log \beta_n^*(e^{-nr}) = D(\Omega_X^{(\theta)} \| Q_X)$$

where

$$\Omega_X^{(\theta)}(x) \triangleq \frac{\exp\left\{\theta \log \dfrac{P_X(x)}{Q_X(x)}\right\} dQ_X(x)}{\int \exp\left\{\theta \log \dfrac{P_X(x')}{Q_X(x')}\right\} dQ_X(x')}$$

and θ is chosen such that $D(\Omega_X^{(\theta)} \| P_X) = r$ [2, pp. 312]. Equation (1) is known as Stein's lemma. The functional $D(\cdot\|\cdot)$ between two probability measures as defined in (1) is called the (Kullback–Leibler, informational) *divergence*, or *relative entropy*.

By properly modifying the proof of Stein's lemma, (1) can be generalized to

$$(\forall \varepsilon \in (0,1)) \quad \lim_{n\to\infty} -\frac{1}{n} \log \beta_n^*(\varepsilon) = \lim_{n\to\infty} \frac{1}{n} D(P_{X^n} \| Q_{X^n}) \quad (2)$$

if the normalized log-likelihood ratio converges to a limit in probability under the null hypothesis. However, for a more general case such as the random observations are nonstationary, the determination of the exponential decay of the minimum type-II error probabilities becomes intractable, and no known methodology seems to work in such situations, which, to some extent, limit the applications of hypothesis testing to the most general cases.

Due to the recent advanced work done by the information theorists, channels without statistical assumptions such as memorylessness, information stability, stationarity, causality, and ergodicity can be

Manuscript received September 7, 1994; revised June 29, 1995. Part of the correspondence has been included in a term paper for the course, EE728R, at the Univeristy of Maryland, College Park, 1994.

The author was with Wan Ta Technologies Inc., Taiwan. He is now with the Advanced Technology Center, Computer & Communication Research Laboratories, Industrial Technology Research Institute, E000, 195 11, Sec. 4, Chutung, Hsinchu, Taiwan 310, R.O.C.

Publisher Item Identifier S 0018-9448(96)00435-X.

[1] Mathematically

$$\beta_n = \inf_{\{\mathcal{A}\subset\mathcal{X}^n : P_{X^n}(\mathcal{A}) \le \alpha_n\}} Q_{X^n}(\mathcal{A}).$$

[2] For simplicity, we use the notation $\log[P_{X^n}(x^n)/Q_{X^n}(x^n)]$ to represent the log-likelihood ratio of P_{X^n} w.r.t. Q_{X^n}. When P_{X^n} is absolutely continuous w.r.t. Q_{X^n}, the log-likelihood ratio is understood as the log of the Radon–Nikodym derivative of P_{X^n} w.r.t. Q_{X^n}, which is conventionally denoted by $\log[dP_{X^n}/dQ_{X^n}](x^n)$. At the case that P_{X^n} is not absolutely continuous w.r.t. Q_{X^n}, since both P_{X^n} and Q_{X^n} are finite measures (and hence, σ-finite measures), by the Lebesque Decomposition Proposition [4, pp. 278], P_{X^n} can be uniquely written as sum of two measures, $P_{X^n}^{(0)}$ and $P_{X^n}^{(1)}$, where $P_{X^n}^{(0)}$ is absolutely continuous w.r.t. Q_{X^n} and $P_{X^n}^{(1)}$ is singular w.r.t. Q_{X^n}. Thus the log-likelihood ratio is understood as $\log[dP_{X^n}^{(0)}/dQ_{X^n}](x^n)$ if x^n lies in the support of $P_{X^n}^{(0)}$, and it is infinite if x^n lies in the support of $P_{X^n}^{(1)}$. Note that only those x^n that lie in the support of P_{X^n} will affect the minimum type-II error probability of the hypothesis testing system. In addition, P_{X^n} and Q_{X^n} must be defined over the same σ-field.

handled by employing the *liminf in probability* and *limsup in probability*[3] of the information spectrum. As a result, the channel capacity is shown to equal the supremum, over all input processes, of the input–output *inf-information rate* defined as the liminf in probability of the normalized information density [6]; while the channel resolvability, defined as the number of random bits required per channel use in order to generate an input that achieves arbitrarily accurate approximation of the output statistics for any given input process, is exactly the limsup in probability of the normalized information density [7]. It is thus natural to ask whether or not we can apply such methodology to the problem of Neyman–Pearson hypothesis testing, and yield a completely general formula for the Neyman–Pearson error exponent.

We found that the answer to the above inquiry is affirmative. As it turns out, the general formulas of the type-II error exponents have an analogous form to the general ε-capacity formula in [6]. In addition, similarly to the information spectrum in the channel capacity and channel resolvability, the *spectrum of the log-likelihood ratio*, defined as the distribution of the random variables $(1/n)\log[P_{X^n}(X^n)/Q_{X^n}(X^n)]$ evaluated under the null hypothesis distribution, can fully characterize the ultimate type-II error exponent of the hypothesis testing system.

The significance of the general type-II error exponent formula of constant test level becomes transparent when the spectrum of the log-likelihood ratio converges in probability under P to a random variable Z with invertible cumulative distribution function $F(\cdot)$ in which case the type-II error exponent can be explicitly written as

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

for $\varepsilon \in (0,1)$. A more extreme case is that Z is almost surely a constant which is

$$\lim_{n\to\infty} \frac{1}{n} D(P_{X^n}\|Q_{X^n}).$$

This result coincides with (2).

As for the general type-II error exponent formula subject to an exponential type-I error bound, the *large deviation spectrum* (specified in Section III) of the normalized log-likelihood ratio can be considered as a *system operating characteristic* from which the relationship of the test level exponent and the type-II error exponent, as well as the log-likelihood threshold that achieves these exponents, can be clearly illustrated.

This correspondence is arranged in the following fashion. Section II is devoted to the general formula of the type-II error exponent under a fixed test level. The general formula of the type-II error exponent subject to an exponential type-I error bound will be established in Section III. Two applications of the general type-II error exponent formulas are also demonstrated in this correspondence. Section IV covers the discussion of the distributed Neyman–Pearson detection for i.i.d. local observations, in which case the type-II error exponent could become a function of type-I error bound; while in Section V we explore the general formula of the partitioning upper bound for channel reliability function. Some concluding remarks appear in Section VI.

With no ambiguity, we will drop the subscript "X^n" of $P_{X^n}(x^n)$ and $Q_{X^n}(x^n)$ throughout the rest of the correspondence.

[3] If A_n is a sequence of random variables, its *liminf in probability* is the largest extended real number α such that for all $\xi > 0$, $\lim_{n\to\infty}\Pr[A_n \le \alpha - \xi] = 0$. Similarly, its *limsup in probability* is the smallest extended real numbers β such that for all $\xi > 0$, $\lim_{n\to\infty}\Pr[A_n \ge \beta + \xi] = 0$.

II. Generalized Neyman–Pearson Testing Subject to Fixed Type-I Error Constraint

In this section, we derive a general formula for the type-II error exponent under a sequence of arbitrary random observations subject to a constant upper bound ε on the type-I error probability. This is given in the following theorem.

Theorem 1: Given a sequence of random observations $X^n = (X_1,\cdots,X_n)$ which is assumed to have a probability distribution of either P_{X^n} or Q_{X^n}, the type-II error exponent satisfies

$$\sup\{D : \underline{F}(D) < \varepsilon\} \le \limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \le \sup\{D : \underline{F}(D) \le \varepsilon\} \quad (3)$$

$$\sup\{D : \bar{F}(D) < \varepsilon\} \le \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \le \sup\{D : \bar{F}(D) \le \varepsilon\} \quad (4)$$

where

$$\underline{F}(D) \triangleq \liminf_{n\to\infty} P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le D\right]$$

and

$$\bar{F}(D) \triangleq \limsup_{n\to\infty} P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le D\right]$$

and $\beta_n^*(\varepsilon)$ represents the minimum type-II error probability subject to a fixed type-I error bound $\varepsilon \in (0,1)$.

Proof:

Lower bound of (3): For any D satisfying $\underline{F}(D) < \varepsilon$, there exists $\delta > 0$ such that $\underline{F}(D) < \varepsilon - 2\delta$; and hence, by the definition of $\underline{F}(D)$, ($\exists$ a subsequence $\{n_j\}$ and N) such that for $j > N$

$$P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \le D\right] \le \varepsilon - \delta < \varepsilon.$$

$$\begin{aligned}\therefore\ \beta_{n_j}^*(\varepsilon) &\le Q\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} > D\right] \\ &\le P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} > D\right]\cdot \exp\{-n_jD\} \\ &\le \exp\{-n_jD\}.\end{aligned} \quad (5)$$

Therefore

$$\limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \ge \limsup_{j\to\infty} -\frac{1}{n_j}\log\beta_{n_j}^*(\varepsilon) \ge D$$

for any D with $\underline{F}(D) < \varepsilon$.

Upper bound of (3): Let $\mathcal{U}_n$ be the optimal acceptance region for alternative hypothesis under likelihood ratio partition, which is defined as follows.

$$\mathcal{U}_n \triangleq \left\{\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} < \tau_n\right\} + \eta_n\left\{\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} = \tau_n\right\} \quad (6)$$

for some log-likelihood threshold τ_n, and possible randomization factor $\eta_n \in [0,1)$. Then $P(\mathcal{U}_n) = \varepsilon$.

Let $\underline{D} = \sup\{D : \underline{F}(D) \le \varepsilon\}$. Then $\underline{F}(\underline{D}+\delta) > \varepsilon$ for any $\delta > 0$. Hence, $(\exists\gamma = \gamma(\delta) > 0)$, $\underline{F}(\underline{D}+\delta) > \varepsilon + \gamma$.

By the definition of $\underline{F}(\underline{D}+\delta)$, $(\exists N)(\forall n > N)$

$$P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \underline{D}+\delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Therefore

$$\begin{aligned}\beta_n^*(\varepsilon) &= Q\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} > \tau_n\right] \\ &\quad + (1-\eta_n)\cdot Q\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} = \tau_n\right] \\ &\geq Q\left[\underline{D}+\delta \geq \frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} > \tau_n\right] \\ &\quad + (1-\eta_n)\cdot Q\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} = \tau_n\right] \\ &\geq \Bigg(P\left[\underline{D}+\delta \geq \frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} > \tau_n\right] \\ &\quad +(1-\eta_n)\cdot P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} = \tau_n\right]\Bigg) \\ &\qquad \times\ \exp\{-n(\underline{D}+\delta)\} \\ &= \Bigg(P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq \underline{D}+\delta\right] - P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} < \tau_n\right] \\ &\quad -\eta_n P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} = \tau_n\right]\Bigg) \\ &\qquad \times\ \exp\{-n(\underline{D}+\delta)\} \\ &\geq \left(\varepsilon+\frac{\gamma}{2}-\varepsilon\right)\exp\{-n(\underline{D}+\delta)\}, \quad \text{for } n > N \\ &= \frac{\gamma}{2}\exp\{-n(\underline{D}+\delta)\}, \quad \text{for } n > N \end{aligned} \tag{7}$$

$$\therefore \quad \limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \leq \underline{D}+\delta.$$

Since δ can be made arbitrarily small

$$\limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \leq \underline{D}.$$

Lower bound of (4): For any D satisfying $\bar{F}(D) < \varepsilon$, $(\exists\delta > 0)$ such that $\bar{F}(D) < \varepsilon - 2\delta$; and hence, by the definition of $\bar{F}(D)$, $(\exists N)(\forall n > N)$

$$P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq D\right] \leq \varepsilon - \delta < \varepsilon.$$

By following the same procedure as in (5), we have for $n > N$,

$$\beta_n^*(\varepsilon) \leq \exp\{-nD\}.$$

Therefore

$$\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \geq D$$

for any D with $\bar{F}(D) < \varepsilon$.

Upper bound of (4): Let $\bar{D} = \sup\{D : \bar{F}(D) \leq \varepsilon\}$. Then $\bar{F}(\bar{D}+\delta) > \varepsilon$ for any $\delta > 0$. Hence, $(\exists\gamma = \gamma(\delta) > 0)$, $\bar{F}(\bar{D}+\delta) > \varepsilon + \gamma$.

By the definition of $\bar{F}(\bar{D}+\delta)$, ($\exists$ a subsequence $\{n_j\}$ and N) such that for $j > N$

$$P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \leq \bar{D}+\delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Therefore, by following the same procedure as (7), we have for $j > N$

$$\beta_{n_j}^*(\varepsilon) \geq \frac{\gamma}{2}\exp\{-n_j(\bar{D}+\delta)\}$$

$$\therefore \quad \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \leq \liminf_{j\to\infty} -\frac{1}{n_j}\log\beta_{n_j}^*(\varepsilon) \leq \bar{D}+\delta.$$

Since δ can be made arbitrarily small

$$\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) \leq \bar{D}.$$

□

Remarks:

- Both $\bar{F}(D)$ and $\underline{F}(D)$ are nondecreasing; hence, the number of discontinuous points for both functions is countable.
- When the normalized log-likelihood ratio converges in probability to a constant C under the null distribution, which is the case for most detection problems encountered, the type-II error exponent is that constant C, and is independent of the type-I error bound ε. For example, in the special case of the i.i.d. data source with $E_P[\log(P(X)/Q(X))] < \infty$, both functions degenerate to the form

$$\bar{F}(D) = \underline{F}(D) = 1, \quad \text{if } D > D_c \tag{8}$$

$$\bar{F}(D) = \underline{F}(D) = 0, \quad \text{if } D < D_c \tag{9}$$

where $D_c \triangleq E_P[\log(P(X)/Q(X))]$. As a result

$$\limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) = \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon) = D_c.$$

The general formula derived above, however, reveals that the type-II error exponent, in general, is a function of ε. Examples can be easily created from the formula derived in Theorem 1. In Section IV, a more interesting (nontrivial) example in which the Neyman–Pearson error exponent depends on the fixed type-I error bound will be demonstrated.

III. Generalized Neyman–Pearson Testing Subject to Exponential Type-I Error Constraint

In the previous section, we have shown that the type-II error exponent subject to a fixed bound on the type-I error probability is determined by the ultimate cumulative distribution function or the spectrum of the normalized log-likelihood ratio. As for an exponential type-I error bound, the type-II error exponent is characterized by a different statistical quantity of the normalized log-likelihood ratio, which is the *large deviation spectrum* of the normalized log-likelihood ratio. It is quantitatively defined as

$$\bar{f}(D) \triangleq \limsup_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq D\right]$$

and

$$\underline{f}(D) \triangleq \liminf_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq D\right].$$

Observation 1: Both $\bar{f}(D)$ and $\underline{f}(D)$ are nonincreasing functions; and hence they are continuous except at countably many points.

Proof: This observation immediately follows from the inequality

$$(\forall D < D') \quad P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq D\right] \leq P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \leq D'\right].$$

□

Observation 2: Both $\bar{f}(D)$ and $\underline{f}(D)$ are nonnegative functions.

Although the general formula proved later also holds when $\bar{f}(D)$ and $\underline{f}(D)$ are extended real-valued functions, we feel that the inclusion of such cases may detract from the clarity of the proof. Hence, in what follows, we assume:

Assumption 1: $\bar{f}(D)$ and $\underline{f}(D)$ are finite for $D \in \Re$.

Let $\beta_n^*(e^{-nr})$ represent the minimum type-II error probability subject to the type-I error bound e^{-nr}. Let

$$\bar{D}_r \triangleq \sup\{D : \bar{f}(D) \geq r\}$$

and

$$\underline{D}_r \triangleq \sup\{D : \underline{f}(D) \geq r\}.$$

Also let

$$\bar{d}_r \triangleq \sup\{D : \bar{f}(D) > r\}$$

and

$$\underline{d}_r \triangleq \sup\{D : \underline{f}(D) > r\}.$$

Define τ_n (the threshold of the likelihood ratio test) as the same as (6). Then we know that

$$P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \tau_n\right] \ge e^{-nr} \ge P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} < \tau_n\right].$$

Lemma 1: Under Assumption 1, if $r > 0$

a) $\bar{D}_r \ge \limsup_{n\to\infty} \tau_n \ge \bar{d}_r$

b) $\underline{D}_r \ge \liminf_{n\to\infty} \tau_n \ge \underline{d}_r$.

Proof: Note that under Assumption 1, $\bar{D}_r$, $\underline{D}_r$, $\bar{d}_r$, and $\underline{d}_r$ are all finite.

Lower bound of a): Suppose $\limsup_{n\to\infty} \tau_n < \bar{d}_r - \xi$ for some $\xi > 0$. Then $(\exists N)(\forall n > N)$ $\tau_n < \bar{d}_r - \xi/2$, which implies

$$e^{-nr} \le P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \tau_n\right] \le P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \bar{d}_r - \frac{\xi}{2}\right].$$

$$\Rightarrow \quad r \ge \limsup_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \bar{d}_r - \frac{\xi}{2}\right] = \bar{f}(\bar{d}_r - \xi/2)$$

$$\therefore \quad \sup\{D : \bar{f}(D) > r\} \le \bar{d}_r - \frac{\xi}{2}$$

a contradiction.

Upper bound of a): Suppose

$$\limsup_{n\to\infty} \tau_n > \bar{D}_r + \xi$$

for some $\xi > 0$. Then ($\exists N$ and a subsequence $\{n_j\}$)($\forall n_j > N$) $\tau_{n_j} > \bar{D}_r + \xi/2$, which implies

$$\begin{aligned} e^{-n_j r} &\ge P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} < \tau_{n_j}\right] \\ &\ge P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \le \bar{D}_r + \frac{\xi}{2}\right]. \end{aligned}$$

$$\Rightarrow \quad r \le \limsup_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \bar{D}_r + \frac{\xi}{2}\right] = \bar{f}(\bar{D}_r + \xi/2)$$

which is also a contradiction.

Lower bound of b): Suppose $(\exists \xi > 0)$ such that

$$\liminf_{n\to\infty} \tau_n < \underline{d}_r - \xi.$$

Then ($\exists N$ and a subsequence $\{n_j\}$)($\forall n_j > N$) $\tau_{n_j} < \underline{d}_r - \xi/2$.

$$\begin{aligned} \therefore\ e^{-n_j r} &\le P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \le \tau_{n_j}\right] \\ &\le P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \le \underline{d}_r - \xi/2\right]. \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad r &\ge \liminf_{j\to\infty} -\frac{1}{n_j}\log P\left[\frac{1}{n_j}\log\frac{P(X^{n_j})}{Q(X^{n_j})} \le \underline{d}_r - \frac{\xi}{2}\right] \\ &\ge \liminf_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \underline{d}_r - \frac{\xi}{2}\right] \\ &= \underline{f}\left(\underline{d}_r - \frac{\xi}{2}\right). \end{aligned}$$

Hence

$$\sup\{D : \underline{f}(D) > r\} \le \underline{d}_r - \frac{\xi}{2}$$

which is a contradiction.

Upper bound of b): Suppose $(\exists \xi > 0)$ such that

$$\liminf_{n\to\infty} \tau_n > \underline{D}_r + \xi.$$

So $(\exists N)(\forall n > N)$ $\tau_n > \underline{D}_r + \xi/2$.

$$\begin{aligned} \therefore\ e^{-nr} &\ge P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} < \tau_n\right] \\ &\ge P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \underline{D}_r + \frac{\xi}{2}\right]. \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad r &\le \liminf_{n\to\infty} -\frac{1}{n}\log P\left[\frac{1}{n}\log\frac{P(X^n)}{Q(X^n)} \le \underline{D}_r + \frac{\xi}{2}\right] \\ &= \underline{f}(\underline{D}_r + \xi/2) \end{aligned}$$

which contradicts to the definition of $\underline{D}_r$. □

We need the following theorem to prove the main result, for which the proof can be found in [3, Theorem 1, pp. 443].

Theorem 2: Let α^* and β^* be the type-I error and type-II error from the test procedure that H_0 is accepted if $aP(X) > bQ(X)$, H_1 is accepted if $aP(X) < bQ(X)$, and either H_0 or H_1 may be accepted if $aP(X) = bQ(X)$. Then for any admissible type-I error α and type-II error β,

$$a\alpha^* + b\beta^* \le a\alpha + b\beta.$$

Theorem 3 (Main result): Under Assumption 1, if $r > 0$, then

$$\text{a)} \quad r + \underline{D}_r \ge \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(e^{-nr}) \ge r + \underline{d}_r$$

$$\text{b)} \quad r + \bar{D}_r \ge \limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(e^{-nr}) \ge r + \bar{d}_r$$

for those r satisfying $r + \underline{d}_r > 0$ for (a), and $r + \bar{d}_r > 0$ for (b).

Proof: For each n, the log-likelihood threshold τ_n and the randomization factor η_n (defined the same as in (6)) are chosen such that type-I error $\alpha_n = e^{-nr}$. Hence from Theorem 2, the weighted sum

$$\gamma_n \triangleq e^{-n\tau_n}e^{-nr} + \beta_n^*(e^{-nr})$$

is minimized for each n. Therefore, if $e^{-n\tau_n}e^{-nr}$ and $\beta_n^*(e^{-nr})$ do not have the same exponential behavior, then for some subsequence $\{n_j\}_{j=1}^{\infty}$, one will dominate the other in the sense of the exponent for n_j large enough. We can then create a new partition (with some randomization if necessary) such that for n_j large enough

$$e^{-n_j\tau_{n_j}}\alpha_{n_j} + \beta_{n_j} < \gamma_{n_j}$$

which contradicts the definition of γ_n. Consequently

$$r + \limsup_{n\to\infty} \tau_n = \limsup_{n\to\infty} -\frac{1}{n}\log\beta_n^*(e^{-nr})$$

and

$$r + \liminf_{n\to\infty} \tau_n = \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(e^{-nr}).$$

Note that the above argument holds only when γ_n decays to zero exponentially. Hence a sufficient condition for the validity of the above argument is that $r + \limsup_{n\to\infty} \tau_n$ and $r + \liminf_{n\to\infty} \tau_n$ are positive. The theorem is completed by employing the results of Lemma 1. □

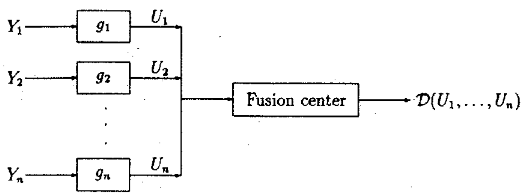

Fig. 1. Distributed detection in $\mathcal{S}_n$.

Example: Suppose X^n is i.i.d. with null distribution *Normal(0,1)* and alternative distribution *Normal(1,1)*. Then

$$f(D)=\bar{f}(D)=\underline{f}(D)=\begin{cases}\frac{1}{2}\left(\frac{1}{2}-D\right)^2, & \text{for } D\le\frac{1}{2}\\ 0, & \text{for } D>\frac{1}{2}.\end{cases}$$

So the type-II error exponent is

$$\lim_{n\to\infty}-\frac{1}{n}\log\beta_n^*(e^{-nr})=r+d_r=r+\frac{1}{2}-\sqrt{2r}$$

for $0<r<\frac{1}{2}$. □

From the above example, we have the following observation. As known in the literature, the Neyman–Pearson receiver operating characteristic (ROC) curve can clearly demonstrate the relationship between the test-level ε and the minimum type-II error β. In this example, the large deviation spectrum $f(D)$ can also be treated as a characteristic curve for testing two hypotheses via a sequence of random observations. When the system requires the exponent of test level to be r, by choosing the test log-likelihood threshold as $d_r=f^{-1}(r)$, the resulting type-II error exponent is $r+d_r$. Hence, the relationship of the test level exponent and the type-II error exponent, as well as the log-likelihood threshold that ultimately achieves these exponents, can be easily obtained from the function $f(D)$.

IV. Parallel Distributed Neyman–Pearson Detection

Consider the distributed binary detection system $\mathcal{S}_n$ depicted in Fig. 1. The sensor observations $Y_1,\cdots,Y_n$ are assumed independent and identically distributed given each of the two hypotheses. 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 $\mathcal{D}(U_1,\cdots,U_n)$.

The Neyman–Pearson optimization problem for the above system is formulated as follows: choose $g_1,\cdots,g_n$ and $\mathcal{D}$ so as to minimize the type-II error probability subject to a fixed upper bound ε on the type-I error probability. Since for any local quantizers $g_1,\cdots,g_n$ chosen, the global decision $\mathcal{D}$ is nothing but a likelihood ratio test of the joint post-quantization distribution of $(U_1,\cdots,U_n)$, which is uniquely determined by $g_1,\cdots,g_n$, the minimization of the type-II error probability suffices to be taken over all possible combinations of $g_1^n=(g_1,\cdots,g_n)$.

Let $\beta_n^*(\varepsilon;g_1^n)$ denote the minimum type-II error attainable subject to a type-I error bound ε under given quantizers $g_1,g_2,\cdots,g_n$. The set of all possible m-level quantizers will be denoted by $\mathcal{G}_m$. Define

$$\tilde{\beta}_n^*(\varepsilon)\triangleq\inf_{(g_1,\cdots,g_n)\in\mathcal{G}_m^n}\beta_n^*(\varepsilon;g_1^n).$$

(In notation, we use $\tilde{\beta}$ in order to distinguish the minimum type-II error of the distributed system from that of the centralized system.)

From Theorem 1, we can immediately obtain

$$\sup\{D:\underline{F}(D;\boldsymbol{g})<\varepsilon\}\le\limsup_{n\to\infty}-\frac{1}{n}\log\beta_n^*(\varepsilon;g_{n1},\cdots,g_{nn})\le\sup\{D:\underline{F}(D;\boldsymbol{g})\le\varepsilon\}$$

and

$$\sup\{D:\bar{F}(D;\boldsymbol{g})<\varepsilon\}\le\liminf_{n\to\infty}-\frac{1}{n}\log\beta_n^*(\varepsilon;g_{n1},\cdots,g_{nn})\le\sup\{D:\bar{F}(D;\boldsymbol{g})\le\varepsilon\}$$

where

$$\underline{F}(D;\boldsymbol{g})\triangleq\liminf_{n\to\infty}P\left[\frac{1}{n}\sum_{i=1}^{n}\log\frac{P(g_{ni}(Y_i))}{Q(g_{ni}(Y_i))}\le D\right]$$

$$\bar{F}(D;\boldsymbol{g})\triangleq\limsup_{n\to\infty}P\left[\frac{1}{n}\sum_{i=1}^{n}\log\frac{P(g_{ni}(Y_i))}{Q(g_{ni}(Y_i))}\le D\right]$$

and $\boldsymbol{g}\triangleq((g_{11}),(g_{21},g_{22}),(g_{31},g_{32},g_{33}),\cdots)$ denotes a choice of sequences (triangular array) of quantizer group. We can then optimize the ultimate type-II error exponent (bounds) over all possible quantizer sequences $\boldsymbol{g}$.

The evaluations of

$$\sup_{\boldsymbol{g}}\sup\{D:\bar{F}(D;\boldsymbol{g})<\varepsilon\}\quad\text{and}\quad\sup_{\boldsymbol{g}}\sup\{D:\bar{F}(D;\boldsymbol{g})\le\varepsilon\}$$

$$\sup_{\boldsymbol{g}}\sup\{D:\underline{F}(D;\boldsymbol{g})<\varepsilon\}\quad\text{and}\quad\sup_{\boldsymbol{g}}\sup\{D:\underline{F}(D;\boldsymbol{g})\le\varepsilon\}$$

involve finding the optimal quantizer sequences for ultimate functions, which in our opinion are difficult to compute. However, by slightly modifying the proof of Theorem 1, it can be shown that we can interchange the order of supremum and limitation, and yield more useful forms of these bounds.

Theorem 4:

$$\sup\{D:\underline{H}(D)<\varepsilon\}\le\limsup_{n\to\infty}-\frac{1}{n}\log\tilde{\beta}_n^*(\varepsilon)\le\sup\{D:\underline{H}(D)\le\varepsilon\}\tag{10}$$

$$\sup\{D:\bar{H}(D)<\varepsilon\}\le\liminf_{n\to\infty}-\frac{1}{n}\log\tilde{\beta}_n^*(\varepsilon)\le\sup\{D:\bar{H}(D)\le\varepsilon\}\tag{11}$$

where

$$\bar{H}(D)\triangleq\limsup_{n\to\infty}\inf_{(g_1,\cdots,g_n)\in\mathcal{G}_m^n}P\left[\frac{1}{n}\sum_{i=1}^{n}\log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))}\le D\right]$$

and

$$\underline{H}(D)\triangleq\liminf_{n\to\infty}\inf_{(g_1,\cdots,g_n)\in\mathcal{G}_m^n}P\left[\frac{1}{n}\sum_{i=1}^{n}\log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))}\le D\right].$$

Proof:

Lower bound of (10): For any D satisfying $\underline{H}(D)<\varepsilon$, $\underline{H}(D)<\varepsilon-2\delta$ for some $\delta>0$; hence, $(\exists N$ and a subsequence $\{n_j\})(\forall j>N)$

$$\inf_{(g_1,\cdots,g_{n_j})\in\mathcal{G}_m^{n_j}}P\left[\frac{1}{n_j}\sum_{i=1}^{n_j}\log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))}\le D\right]\le\varepsilon-\delta.$$

Therefore, $(\exists(\tilde{g}_1,\cdots,\tilde{g}_{n_j})\in\mathcal{G}_m^{n_j})$ such that

$$P\left[\frac{1}{n_j}\sum_{i=1}^{n_j}\log\frac{P(\tilde{g}_i(Y_i))}{Q(\tilde{g}_i(Y_i))}\le D\right]\le\varepsilon-\frac{\delta}{2}<\varepsilon.$$

Therefore, by following the same procedure as (5), we obtain for $j>N$

$$\tilde{\beta}_{n_j}^*(\varepsilon)\le\beta_{n_j}^*(\varepsilon;\tilde{g}_1^{n_j})\le\exp\{-n_jD\}.$$

Upper bound of (10): Given quantizers $g_1^n = (g_1, \cdots, g_n)$, let $\mathcal{V}_n = \mathcal{V}_n(g_1^n)$ be the optimal acceptance region for the alternative hypothesis under likelihood ratio partition, which is defined as

$$\mathcal{V}_n \triangleq \left\{ \frac{1}{n} \sum_{i=1}^n \log \frac{P(g_i(Y_i))}{Q(g_i(Y_i))} < \tau_n \right\} + \eta_n \left\{ \frac{1}{n} \sum_{i=1}^n \log \frac{P(g_i(Y_i))}{Q(g_i(Y_i))} = \tau_n \right\}$$

for some log-likelihood ratio threshold τ_n, and randomization factor $\eta_n \in [0, 1)$. Then $P(\mathcal{V}_n) = \varepsilon$.

Let

$$D' = \sup\{D : \underline{H}(D) \le \varepsilon\}.$$

Then $\underline{H}(D' + \delta) > \varepsilon$ for any $\delta > 0$. Hence, $(\exists \gamma = \gamma(\delta) > 0)$, $\underline{H}(D' + \delta) > \varepsilon + \gamma$. So $(\exists N)(\forall n > N)$

$$\inf_{(g_1,\cdots,g_n)\in\mathcal{G}_m^n} P\left[\frac{1}{n}\sum_{i=1}^n \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D' + \delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Therefore, for any $(g_1, \cdots, g_n) \in \mathcal{G}_m^n$

$$P\left[\frac{1}{n}\sum_{i=1}^n \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D' + \delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Following the same procedure as (7), we obtain

$$\beta_n^*(\varepsilon; g_1^n) \ge \frac{\gamma}{2} \exp\{-n(D' + \delta)\}, \quad \text{for any } (g_1, \cdots, g_n) \in \mathcal{G}_m^n.$$

Consequently

$$\limsup_{n\to\infty} -\frac{1}{n} \log \tilde{\beta}_n^*(\varepsilon) \le D' + \delta$$

for arbitrary small $\delta > 0$.

Lower bound of (11): For any D satisfying $\bar{H}(D) < \varepsilon$, $\bar{H}(D) < \varepsilon - 2\delta$ for some $\delta > 0$; hence, $(\exists N)(\forall n > N)$

$$\inf_{(g_1,\cdots,g_n)\in\mathcal{G}_m^n} P\left[\frac{1}{n}\sum_{i=1}^n \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D\right] \le \varepsilon - \delta.$$

Therefore, $(\exists(\hat{g}_1, \cdots, \hat{g}_n) \in \mathcal{G}_m^n)$ such that

$$P\left[\frac{1}{n}\sum_{i=1}^n \log\frac{P(\hat{g}_i(Y_i))}{Q(\hat{g}_i(Y_i))} \le D\right] \le \varepsilon - \frac{\delta}{2} < \varepsilon.$$

Therefore, by following the same procedure as (5), we obtain for $n > N$

$$\tilde{\beta}_n^*(\varepsilon) \le \beta_n^*(\varepsilon; \hat{g}_1^n) \le \exp\{-nD\}.$$

Upper bound of (11): Let $\mathcal{V}_n = \mathcal{V}_n(g_1^n)$ be defined the same as above.

Let

$$D'' = \sup\{D : \bar{H}(D) \le \varepsilon\}.$$

Then $\bar{H}(D'' + \delta) > \varepsilon$ for any $\delta > 0$. Hence, $(\exists \gamma = \gamma(\delta) > 0)$, $\bar{H}(D'' + \delta) > \varepsilon + \gamma$. So ($\exists N$ and a subsequence $\{n_j\}$)$(\forall j > N)$

$$\inf_{(g_1,\cdots,g_{n_j})\in\mathcal{G}_m^{n_j}} P\left[\frac{1}{n_j}\sum_{i=1}^{n_j} \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D'' + \delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Therefore, for any $(g_1, \cdots, g_{n_j}) \in \mathcal{G}_m^{n_j}$

$$P\left[\frac{1}{n_j}\sum_{i=1}^{n_j} \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D'' + \delta\right] > \varepsilon + \frac{\gamma}{2}.$$

Following the same procedure as (7)

$$\beta_{n_j}^*(\varepsilon; g_1^{n_j}) \ge \frac{\gamma}{2} \exp\{-n_j(D'' + \delta)\}, \quad \text{for any } (g_1, \cdots, g_{n_j}) \in \mathcal{G}_m^{n_j}.$$

Consequently

$$\liminf_{n\to\infty} -\frac{1}{n} \log \tilde{\beta}_n^*(\varepsilon) \le D'' + \delta$$

for arbitrary $\delta > 0$. □

In the proof of Theorem 4, we did not use the fact that the original observations are i.i.d.; hence, the same proof works for the most general cases.

It is easy to show that in a distributed system $\mathcal{S}_n$ (with i.i.d. local observations and m-ary quantization), when the second moment of the post-quantization log-likelihood ratio is uniformly bounded, i.e.

$$\sup_{g\in\mathcal{G}_m} E_P\left[(\log[P(g(Y))/Q(g(Y))])^2\right] < \infty$$

$\bar{H}(D)$ and $\underline{H}(D)$ degenerate to the same form as (8) and (9) for some constant

$$D_m \triangleq \sup_{g\in\mathcal{G}_m} E_P[\log[P(g(Y))/Q(g(Y))]].$$

Hence, the type-II error exponent is D_m, and is independent of the test level. However, in absence of the above boundedness condition, it is shown in [5, Theorem 4] that the type-II error exponent could become a function of type-I error bound. For the completeness of this discussion, we quote the result in [5] in the following.

Theorem 5: Define

$$\nu_t \triangleq P\left(\log\frac{P(Y)}{Q(Y)} \le t\right) \log \frac{P\left(\log\frac{P(Y)}{Q(Y)} \le t\right)}{Q\left(\log\frac{P(Y)}{Q(Y)} \le t\right)} + P\left(\log\frac{P(Y)}{Q(Y)} > t\right) \log \frac{P\left(\log\frac{P(Y)}{Q(Y)} > t\right)}{Q\left(\log\frac{P(Y)}{Q(Y)} > t\right)}.$$

Let

$$\nu_0 \triangleq \limsup_{t\to\infty} \nu_t$$

and

$$\bar{\nu} \triangleq \sup_{g\in\mathcal{G}_m} E_P\left[\left|\log\frac{P(g(Y))}{Q(g(Y))}\right|\right].$$

If $0 < \nu_0 < \infty$, there exists a subsequence $\{n_k\}$ such that

$$\liminf_{k\to\infty} -\frac{1}{n_k} \log \tilde{\beta}_{n_k}^*(\varepsilon) \ge \frac{\nu_0}{\log(1/\varepsilon)} \vee D_m \tag{12}$$

and

$$\limsup_{n\to\infty} -\frac{1}{n} \log \tilde{\beta}_n^*(\varepsilon) \le \frac{\bar{\nu}}{1-\varepsilon}.$$ □

The above theorem can be easily justified using the general formula derived in Theorem 4. Indeed, it suffices to show that

$$\underline{H}(D) \ge 1 - \frac{\bar{\nu}}{D} \tag{13}$$

$$\bar{H}(D) = 0, \quad \text{if } D < D_m \tag{14}$$

and

$$\limsup_{k\to\infty} \inf_{(g_1,\cdots,g_{n_k})\in\mathcal{G}_m^{n_k}} P\left[\frac{1}{n_k}\sum_{i=1}^{n_k} \log\frac{P(g_i(Y_i))}{Q(g_i(Y_i))} \le D\right] \le \exp\{-\nu_0/D\}, \quad \text{if } D > D_m. \tag{15}$$

The proof of (13) is an immediately result of the Chebeshev inequality. Equation (14) can be proved by properly choosing a quantizer g for which $E_P[\log[P(g(Y))/Q(g(Y))]]$ is arbitrarily close to D_m, and letting $g_1 = g_2 = \cdots = g_n = g$.

To prove (15), we first note that $(\forall\gamma > 0)(\exists$ a sequence $\{t_k\}_{k=1}^{\infty})$ such that $t_k \to \infty$ and (see the bottom of this page). Let $g_{(k)}$ be the corresponding binary quantizer with log-likelihood threshold t_k, and let

$$n_k \triangleq \left\lfloor \frac{1}{D+\gamma} \log \frac{P(\log[P(Y)/Q(Y)] > t_k)}{Q(\log[P(Y)/Q(Y)] > t_k)} \right\rfloor.$$

Then by choosing $g_1 = g_2 = \cdots = g_{n_k} = g_{(k)}$.

$$\begin{aligned} P&\left[\frac{1}{n_k}\sum_{i=1}^{n_k} \log \frac{P(g_{(k)}(Y_i))}{Q(g_{(k)}(Y_i))} > D\right] \\ &\geq 1 - \left(1 - P\left[\log \frac{P(g_{(k)}(Y_i))}{Q(g_{(k)}(Y_i))} > t_k\right]\right)^{n_k} \\ &\geq 1 - \left(1 - \frac{\nu_0 - \gamma}{(n_k+1)(D+\gamma)}\right)^{n_k} \\ &\stackrel{k\to\infty}{\longrightarrow} 1 - \exp\left\{-\frac{\nu_0-\gamma}{D+\gamma}\right\}. \end{aligned}$$

Since γ can be made arbitrarily small, (15) holds.

V. Generalized Partitioning Upper Bound of the Channel Reliability Function

Consider a noisy channel with input sequence X^n, output sequence Y^n, and channel transition probability $W_{Y^n|X^n}(y^n|x^n)$. The channel reliability function is the largest exponent (among all feasible codes) of the probability of average decoding error w.r.t. the codeword length n, subject to an exponential bound $\exp\{nR\}$ on the size of the code book. By the definition of the channel capacity C, the channel reliability function is zero for $R > C$.

In [6], Verdú and Han have shown that the general formula of the channel capacity for arbitrary channel is

$$C = \sup_{\boldsymbol{X}} \underline{I}(\boldsymbol{X};\boldsymbol{Y})$$

where $\underline{I}(\boldsymbol{X};\boldsymbol{Y})$ is the liminf in probability of the sequence of the random variables

$$\frac{1}{n} i_{X^nW^n}(x^n;y^n) \triangleq \frac{1}{n}\log \frac{W_{Y^n|X^n}(y^n|x^n)}{\sum_{z^n} P_{X^n}(z^n) W_{Y^n|X^n}(y^n|z^n)}.$$

The partitioning upper bound of the channel reliability function was derived in terms of transforming the problem into a hypothesis testing setup. Hence, with the general formula of the Neyman–Pearson type-II error exponent derived in Section III, we provide a general partitioning upper bound for channel reliability function.

It is worth mentioning that for discrete memoryless channels, the partitioning upper bound is tight for rates near capacity.

A. Transformation of Problem from Channel Reliability Function to Hypothesis Testing

Let $\mathcal{U}_m$ be the decoding set for codeword m for $1 \leq m \leq M = \exp\{nR\}$. Observe that for any probability distribution on the channel output $P_{Y^n}(\cdot)$, there exists a codeword m such that

$$\sum_{y^n \in \mathcal{U}_m} P_{Y^n}(y^n) \leq \frac{1}{M}.$$

Furthermore, the probability of decoding error by transmitting codeword m is

$$P_{e|m} = \sum_{y^n \in \mathcal{U}_m^c} W_{Y^n|X^n}(y^n|c_m)$$

where c_m represents the mth codeword. Hence, we can consider the problem as a hypothesis testing problem of $H_0 : W_{Y^n|X^n}(\cdot|c_m)$ versus $H_1 : P_{Y^n}(\cdot)$.

Let E denote the desired exponent for probability of decoding error, i.e., we dictate

$$P_{e|m} \leq \exp\{-nE\}.$$

Then the rate must satisfy

$$\beta_n^*(\exp\{-nE\}) \leq P_{Y^n}(\mathcal{U}_m) \leq \frac{1}{M} = \exp\{-nR\}.$$

Hence, by Theorem 3,

$$R \leq \liminf_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\exp\{-nE\}) \leq E + \sup\{D : \ell_m(D) \geq E\}$$

where

$$\begin{aligned} \ell_m(D) &\triangleq \liminf_{n\to\infty} -\frac{1}{n}\log \Pr\left\{\frac{1}{n}\log\frac{W_{Y^n|X^n}(Y^n|c_m)}{P_{Y^n}(Y^n)} \leq D\right\} \\ &= \liminf_{n\to\infty} -\frac{1}{n}\log \Pr\left\{\frac{1}{n} i_{X^nW^n}(c_m;Y^n) \leq D\right\}. \end{aligned}$$

Since the exponent of the minimum average decoding error probability is equal to that of the minimum $\max_{1\leq m\leq M} P_{e|m}$, by using the same argument as in [1, Theorem 10.2.1], we obtain the rate must be bounded above by

$$R \leq \bar{R}(E) \triangleq E + \sup_{\boldsymbol{X}} \sup\{D : L_{\boldsymbol{X}}(D) \geq E\}$$

where

$$L_{\boldsymbol{X}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log\Pr\left\{\frac{1}{n} i_{X^nW^n}(X^n;Y^n) \leq D\right\}.$$

We therefore obtain a dual function of the generalized partitioning upper bound of the channel reliability function $\bar{R}(E)$.

Observation 3:

$$\bar{R}(0^+) \triangleq \lim_{E\downarrow 0} \bar{R}(E) \leq C = \sup_{\boldsymbol{X}} \underline{I}(\boldsymbol{X};\boldsymbol{Y}).$$

Proof: Observe that for any sequence of input distributions chosen,

$$\limsup_{n\to\infty} \Pr\left\{\frac{1}{n} i_{X^nW^n}(X^n;Y^n) \leq \underline{I}(\boldsymbol{X};\boldsymbol{Y}) + \delta\right\} > 0$$

for any $\delta > 0$. Hence

$$L_{\boldsymbol{X}}(\underline{I}(\boldsymbol{X};\boldsymbol{Y}) + \delta) = 0. \tag{16}$$

Since (16) is true for all possible sequences of input distributions, $\sup_{\boldsymbol{X}} L_{\boldsymbol{X}}(C+\delta) = 0$. We therefore have

$$\bar{R}(E) \leq E + \sup_{\boldsymbol{X}} \sup\{D : L_{\boldsymbol{X}}(D) > 0\} \leq E + (C+\delta).$$

Since δ is arbitrary, $\bar{R}(0^+) \leq C$. □

$$0 \geq P(\log[P(Y)/Q(Y)] \leq t_k)\log\frac{P(\log[P(Y)/Q(Y)] \leq t_k)}{Q(\log[P(Y)/Q(Y)] \leq t_k)} \geq -\gamma$$

$$\nu_0 + \gamma \geq P(\log[P(Y)/Q(Y)] > t_k)\log\frac{P(\log[P(Y)/Q(Y)] > t_k)}{Q(\log[P(Y)/Q(Y)] > t_k)} \geq \nu_0 - \gamma.$$

Remark: It is known that for a discrete memoryless channel, $\bar{R}(0^+) = C$; however, from the general formula derived above, this is not true in general. An additional condition, i.e.

$$(\forall \delta > 0) \quad \sup_{\boldsymbol{X}} L_{\boldsymbol{X}}(C - \delta) > 0 \tag{17}$$

is necessary in order for $\bar{R}(0^+) = C$ (see the proof of Observation 4).

B. A Lower Bound of the Channel Reliability Function

By [6, Theorem 1], we can also derive a lower bound on the channel reliability function, which, together with the upper bound obtained above, results in the next corollary.

Corollary 1: Let $R(E)$ be the dual function of the channel reliability function (if it exists). Then for $E > 0$

$$\left(E + \sup_{\boldsymbol{X}} \inf\{D : L_{\boldsymbol{X}}(D) < E\}\right) \vee 0 \geq R(E) \geq \left(\sup_{\boldsymbol{X}} \inf\{D : L_{\boldsymbol{X}}(D + E) < E\}\right) \vee 0$$

where

$$L_{\boldsymbol{X}}(D) = \liminf_{n \to \infty} -\frac{1}{n} \log \Pr\left\{\frac{1}{n} i_{X^n W^n}(X^n; Y^n) \leq D\right\}.$$

Proof: Since $L_{\boldsymbol{X}}(D)$ is a nonincreasing function, the upper bound follows from

$$\sup\{D : L_{\boldsymbol{X}}(D) \geq E\} = \inf\{D : L_{\boldsymbol{X}}(D) < E\}.$$

From Theorem 1 in [6], there exists an

$$(n, \exp\{nR\}, \exp\{-nE\})$$

code satisfying

$$\exp\{-nE\} \leq \Pr\left[\frac{1}{n} i_{X^n W^n}(X^n; Y^n) \leq R + E + \gamma\right] + \exp\{-n(E + \gamma)\}$$

for any $\gamma > 0$ and any input distributions P_{X^n}. Therefore, it is possible to find a sequence of input distributions such that

$$E \geq \liminf_{n \to \infty} -\frac{1}{n} \log \Pr\left[\frac{1}{n} i_{X^n W^n}(X^n; Y^n) \leq R + E + \gamma\right] = L_{\boldsymbol{X}}(R + E + \gamma).$$

Hence

$$R(E) \geq \sup_{\boldsymbol{X}} \inf\{D : L_{\boldsymbol{X}}(D + E) < E\}. \qquad \square$$

Observation 4: $R(0^+) = C$ if, and only if, (17) holds.

Proof: Suppose (17) holds. Let

$$\gamma \triangleq \min\left\{\sup_{\boldsymbol{X}} L_{\boldsymbol{X}}(C - \delta), \delta\right\}.$$

Then $(\exists \tilde{\boldsymbol{X}})$ such that

$$L_{\tilde{\boldsymbol{X}}}(C - \delta) = L_{\tilde{\boldsymbol{X}}}(C - (\delta + \gamma/2) + \gamma/2) \geq \gamma/2.$$

Hence

$$\begin{aligned} R(\gamma/2) &\geq \sup_{\boldsymbol{X}} \inf\{D : L_{\boldsymbol{X}}(D + \gamma/2) < \gamma/2\} \\ &\geq \inf\{D : L_{\tilde{\boldsymbol{X}}}(D + \gamma/2) < \gamma/2\} \\ &\geq C - (\delta + \gamma/2). \end{aligned}$$

By letting $\delta \to 0$, we obtain $R(0^+) \geq C$, which, together with the result of Observation 3, implies $R(0^+) = C$.

Suppose $\exists\ \delta > 0$ such that

$$\sup_{\boldsymbol{X}} L_{\boldsymbol{X}}(C - \delta) = 0.$$

Then for all $\delta' \in (0, \delta)$

$$\sup_{\boldsymbol{X}} L_{\boldsymbol{X}}(C - \delta') = 0.$$

Therefore

$$R(0^+) \leq \bar{R}(0^+) \leq C - \delta < C. \qquad \square$$

VI. Concluding Remarks

Inspired by the work of Verdú and Han on channel capacity, a general expression for the Neyman–Pearson type-II error exponent based on arbitrary random observations is derived. It provides a new view of the simple hypothesis testing problem. For example, in light of the formula in Theorem 1, we know that the existence of the type-II error exponent of fixed test level is not limited to the case where the normalized log-likelihood ratio converges to a limit in probability under the null hypothesis. Futhermore, the type-II error exponent, even if it exists, is in general a function of the test level. Pertinent applications of the formulas also shows that there are situations where it is adveantageous to work with the general expressions.

Acknowledgment

Discussions with Prof. T.-Y. Lee who is with the Department of Mathematics at the University of Maryland, College Park, are gratefully acknowledged.

References

[1] R. E. Blahut, *Principles and Practice of Information Theory.* Reading, MA: Addison-Wesley, 1988.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.
[3] M. H. DeGroot, *Probability and Statistics*, 2nd ed. Reading, MA: Addison-Wesley, 1986.
[4] H. L. Royden, *Real Analysis*, 3rd ed. New York: Macmillan, 1988.
[5] P.-N. Chen and A. Papamarcou, "New asymptotic results in parallel distributed detection," *IEEE Trans. Inform. Theory*, vol. 39, no. 6, pp. 1847–1863, Nov. 1993.
[6] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1147–1157, July 1994.
[7] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.