A Generalized Poor-Verdú Error Bound for Multihypothesis Testing

Po-Ning Chen, *Senior Member, IEEE*, and Fady Alajaji, *Senior Member, IEEE*

Abstract—**A lower bound on the minimum error probability for multihypothesis testing is established. The bound, which is expressed in terms of the cumulative distribution function of the *tilted* posterior hypothesis distribution given the observation with tilting parameter $\theta \geq 1$, generalizes an earlier bound due the Poor and Verdú (1995). A sufficient condition is established under which the new bound (minus a multiplicative factor) provides the exact error probability asymptotically in θ. Examples illustrating the new bound are also provided.**

Index Terms—**Channel reliability function, converse channel coding theorems, hypothesis testing, maximum-a-posteriori estimation, probability of error.**

I. INTRODUCTION

IN [3], Poor and Verdú establish a lower bound to the minimum error probability of multihypothesis testing. Specifically, given two random variables X and Y with joint distribution $P_{X,Y}$, X taking values in a finite or countably-infinite alphabet $\mathcal{X}$ and Y taking values in an arbitrary alphabet $\mathcal{Y}$, they show that the optimal maximum-a-posteriori (MAP) estimation of X given Y results in the following lower bound on the probability of estimation error P_e:

$$P_e \geq (1-\alpha)P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}(x|y) \leq \alpha\right\} \quad (1)$$

for each $\alpha \in [0,1]$, where $P_{X|Y}$ denotes the posterior distribution of X given Y and the prior distribution P_X is arbitrary (not necessarily uniform). This bound has pertinent information-theoretic applications such as in the proof of the converse part of the channel coding theorem that yield formulas for both capacity and ε-capacity for general channels with memory (not necessarily information stable, stationary, etc) [5], [3]. It also improves upon previous lower bounds due to Shannon [4], [3, Eq. (7)] and to Verdú and Han [5], [3, Eq. (9)].

In this work, we generalize the above Poor-Verdú lower bound in (1) for the minimum error probability of multihypothesis testing. The new bound is expressed in terms of the cdf of the tilted posterior distribution of X given Y with tilting parameter $\theta \geq 1$, and it reduces to (1) when $\theta = 1$; see Theorem 1. We also provide a sufficient condition under which

our generalized Poor-Verdú bound, without the multiplicative factor $(1 - \alpha)$, is exact in the limit of θ going to infinity. Specifically, the sufficient condition requires having a unique MAP estimate of X from Y almost surely in P_Y, where P_Y is the distribution of Y; see Theorem 2. We present a few examples to illustrate the results of Theorems 1 and 2. Note that we will use the natural logarithm throughout.

II. MAIN RESULTS

Consider two (correlated) random variables X and Y, where X has a discrete (i.e., finite or countably infinite) alphabet $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$ and Y takes on values in an arbitrary alphabet $\mathcal{Y}$. The minimum probability of error P_e in estimating X from Y is given by

$$P_e \triangleq \Pr[X \neq e(Y)] \quad (2)$$

where $e(Y)$ is the MAP estimate defined as

$$e(Y) = \arg\max_{x \in \mathcal{X}} P_{X|Y}(x|Y). \quad (3)$$

Theorem 1: The above minimum probability of error P_e in estimating X from Y satisfies the following inequality

$$P_e \geq (1-\alpha)P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \leq \alpha\right\} \quad (4)$$

for each $\alpha \in [0,1]$ and $\theta \geq 1$, where for each $y \in \mathcal{Y}$,

$$P_{X|Y}^{(\theta)}(x|y) \triangleq \frac{P_{X|Y}^{\theta}(x|y)}{\sum_{x' \in \mathcal{X}} P_{X|Y}^{\theta}(x'|y)}, \quad x \in \mathcal{X}, \quad (5)$$

is the tilted distribution of $P_{X|Y}(\cdot|y)$ with parameter θ.

Note: When $\theta = 1$, the above bound in (4) reduces to the Poor-Verdú bound in (1).

Proof: Fix $\theta \geq 1$. We only provide the proof for $\alpha < 1$ since the lower bound trivially holds when $\alpha = 1$.

From (2) and (3), the minimum error probability P_e incurred in testing among the values of X satisfies

$$1 - P_e = \int_{\mathcal{Y}} \left(\max_{x \in \mathcal{X}} P_{X|Y}(x|y)\right) dP_Y(y) = E\left[\max_{x \in \mathcal{X}} f_x(Y)\right],$$

where $f_x(y) \triangleq P_{X|Y}(x|y)$. For a fixed $y \in \mathcal{Y}$, let $h_j(y)$ be the j-th element in the set

$$\{f_{x_1}(y), f_{x_2}(y), f_{x_3}(y), \ldots\}$$

such that its elements are listed in non-increasing order; i.e.,

$$h_1(y) \geq h_2(y) \geq h_3(y) \geq \cdots$$

and

$$\{h_1(y), h_2(y), h_3(y), \ldots\} = \{f_{x_1}(y), f_{x_2}(y), f_{x_3}(y), \ldots\}.$$

Then

$$1 - P_e = E[h_1(Y)]. \tag{6}$$

Furthermore, for each $h_j(y)$ above, define $h_j^{(\theta)}(y)$ such that $h_j^{(\theta)}(y)$ be the respective element for $h_j(y)$ satisfying

$$h_j(y) = f_{x_j}(y) = P_{X|Y}(x_j|y) \Leftrightarrow h_j^{(\theta)}(y) = P_{X|Y}^{(\theta)}(x_j|y).$$

Since $h_1(y)$ is the largest among $\{h_j(y)\}_{j \geq 1}$,

$$h_1^{(\theta)}(y) = \frac{h_1^{\theta}(y)}{\sum_{j \geq 1} h_j^{\theta}(y)} = \frac{1}{1 + \sum_{j \geq 2}[h_j(y)/h_1(y)]^{\theta}}$$

is non-decreasing in θ for each y; this implies that

$$h_1^{(\theta)}(y) \geq h_1(y) \quad \text{for } \theta \geq 1 \text{ and } y \in \mathcal{Y}. \tag{7}$$

For any $\alpha \in [0, 1)$, we can write

$$
\begin{aligned}
& P_{X,Y}\left\{(x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) > \alpha\right\} \\
& = \int_{\mathcal{Y}} P_{X|Y}\left\{x \in \mathcal{X} : P_{X|Y}^{(\theta)}(x|y) > \alpha\right\} dP_Y(y) \\
& = \int_{\mathcal{Y}} \left(\sum_{j=1}^{\infty} h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right)\right) dP_Y(y) \\
& \geq \int_{\mathcal{Y}} h_1(y) \cdot \mathbf{1}(h_1^{(\theta)}(y) > \alpha) dP_Y(y) \\
& \geq \int_{\mathcal{Y}} h_1(y) \cdot \mathbf{1}(h_1(y) > \alpha) dP_Y(y) \\
& = E[h_1(Y) \cdot \mathbf{1}(h_1(Y) > \alpha)], \tag{8}
\end{aligned}
$$

where $\mathbf{1}(\cdot)$ is the indicator function and the second inequality follows from (7). To complete the proof, we next relate $E[h_1(Y) \cdot \mathbf{1}(h_1(Y) > \alpha)]$ with $E[h_1(Y)]$, which is exactly $1 - P_e$. Invoking [3, eq. (19)], we have that for any $\alpha \in [0, 1]$ and any random variable U with $\Pr\{0 \leq U \leq 1\} = 1$, the following inequality holds with probability one

$$U \leq \alpha + (1 - \alpha) \cdot U \cdot \mathbf{1}(U > \alpha).$$

Thus

$$E[U] \leq \alpha + (1 - \alpha)E[U \cdot \mathbf{1}(U > \alpha)].$$

Applying the above inequality to (8) by setting $U = h_1(Y)$, we obtain

$$
\begin{aligned}
(1 - \alpha)P_{X,Y}\Big\{(x, y) & \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) > \alpha\Big\} \\
& \geq E[h_1(Y)] - \alpha \\
& = (1 - P_e) - \alpha \\
& = (1 - \alpha) - P_e,
\end{aligned}
$$

where the first equality follows from (6). ∎

We next show that if the MAP estimate $e(Y)$ of X from Y is almost surely unique in (3), then the bound of Theorem 1, without the $(1 - \alpha)$ factor, is tight in the limit of θ going to infinity.

Theorem 2: Consider two random variables X and Y, where X has a finite or countably infinite alphabet $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$ and Y has an arbitrary alphabet $\mathcal{Y}$. Assume that

$$P_{X|Y}(e(y)|y) > \max_{x \in \mathcal{X} : x \neq e(y)} P_{X|Y}(x|y) \tag{9}$$

holds almost surely in P_Y, where $e(y)$ is the MAP estimate from y as defined in (3); in other words, the MAP estimate is almost surely unique in P_Y. Then, the error probability in the MAP estimation of X from Y satisfies

$$P_e = \lim_{\theta \to \infty} P_{X,Y}\left\{(x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \leq \alpha\right\} \tag{10}$$

for each $\alpha \in (0, 1)$, where the tilted distribution $P_{X|Y}^{(\theta)}(\cdot|y)$ is given in (5) for $y \in \mathcal{Y}$.

Proof: It can be easily verified from the definitions of $h_j(\cdot)$ and $h_j^{(\theta)}(\cdot)$ that the following two limits hold for each $y \in \mathcal{Y}$:

$$\lim_{\theta \to \infty} h_1^{(\theta)}(y) = \frac{1}{\ell(y)},$$

where

$$\ell(y) \triangleq \max\{j \in \mathbb{N} : h_j(y) = h_1(y)\} \tag{11}$$

and $\mathbb{N} \triangleq \{1, 2, 3, \ldots\}$ is the set of positive integers, and

$$
\begin{aligned}
& \lim_{\theta \to \infty} h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right) \\
& = \begin{cases} h_j(y) \cdot \mathbf{1}\left(\frac{1}{\ell(y)} > \alpha\right) & \text{for } j = 1, \ldots, \ell(y); \\ 0 & \text{for } j > \ell(y). \end{cases} \tag{12}
\end{aligned}
$$

As a result, we obtain that for any $\alpha \in [0, 1)$,

$$
\begin{aligned}
& \lim_{\theta \to \infty} P_{X,Y}\left\{(x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) > \alpha\right\} \\
& = \lim_{\theta \to \infty} \int_{\mathcal{Y}} \left(\sum_{j=1}^{\infty} h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right)\right) dP_Y(y) \\
& = \int_{\mathcal{Y}} \lim_{\theta \to \infty} \left(\sum_{j=1}^{\infty} h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right)\right) dP_Y(y) \tag{13} \\
& = \int_{\mathcal{Y}} \left(\sum_{j=1}^{\ell(y)} h_j(y) \cdot \mathbf{1}\left(\frac{1}{\ell(y)} > \alpha\right)\right) dP_Y(y), \tag{14}
\end{aligned}
$$

where (13) holds by the dominated convergence theorem since

$$\left|\sum_{j=1}^{\infty} h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right)\right| \leq \sum_{j=1}^{\infty} h_j(y) = 1.$$

Furthermore, (14) holds since the limit (in θ) of

$$a_{\theta,j} \triangleq h_j(y) \cdot \mathbf{1}\left(h_j^{(\theta)}(y) > \alpha\right)$$

exists for every $j = 1, 2, \ldots$ by (12), hence implying that

$$\lim_{\theta \to \infty} \sum_{j=1}^{\infty} a_{\theta,j} = \sum_{j=1}^{\infty} \lim_{\theta \to \infty} a_{\theta,j}.$$

Now condition (9) is equivalent to

$$\Pr[\ell(Y) = 1] \triangleq P_Y\{y \in \mathcal{Y} : \ell(y) = 1\} = 1; \qquad (15)$$

thus,

$$\lim_{\theta \to \infty} P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) > \alpha\right\}$$
$$= \int_{\mathcal{Y}} h_1(y) \cdot \mathbf{1}(1 > \alpha)\, dP_Y(y) = E[h_1(Y)]$$
$$= 1 - P_e, \qquad (16)$$

where (16) follows from (6). This immediately yields that for $0 < \alpha < 1$,

$$P_e = 1 - \lim_{\theta \to \infty} P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) > \alpha\right\}$$
$$= \lim_{\theta \to \infty} P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \le \alpha\right\}.$$

∎

Observation 1: We first note that since the bound in (4) holds for every $\theta \ge 1$, it also holds in the limit of θ going to infinity (the limit exists as shown in the above proof):

$$P_e \ge (1-\alpha) \lim_{\theta \to \infty} P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \le \alpha\right\} \qquad (17)$$

for any $0 \le \alpha \le 1$.

Furthermore, if condition (9) does not hold (or equivalently from (15), if $\Pr[\ell(Y) = 1] < 1$), but there exists an integer $L > 1$ such that $\Pr[\ell(Y) \le L] = 1$, then using (14), we can write (17) as

$$P_e \ge (1-\alpha)\left[1 - \int_{\mathcal{Y}}\left(\sum_{j=1}^{\ell(y)} h_j(y) \cdot \mathbf{1}\left(\frac{1}{\ell(y)} > \alpha\right)\right) dP_Y(y)\right]$$

$$= (1-\alpha)\left[\int_{\mathcal{Y}}\left(\sum_{j=1}^{\infty} h_j(y)\right) dP_Y(y)\right.$$

$$\left. - \int_{\mathcal{Y}}\left(\sum_{j=1}^{\ell(y)} h_j(y) \cdot \mathbf{1}\left(\frac{1}{\ell(y)} > \alpha\right)\right) dP_Y(y)\right]$$

$$\overset{(a)}{=} (1-\alpha)\int_{\mathcal{Y}}\left(\sum_{j=1}^{\ell(y)} h_j(y) \cdot \mathbf{1}\left(\frac{1}{\ell(y)} \le \alpha\right)\right.$$

$$\left. + \sum_{j=\ell(y)+1}^{\infty} h_j(y)\right) dP_Y(y)$$

$$\overset{(b)}{=} (1-\alpha)\left[\int_{y:\ell(y)=1}\left(\sum_{j=1}^{1} h_j(y) \cdot \mathbf{1}(1 \le \alpha)\right.\right.$$

$$\left. + \sum_{j=2}^{\infty} h_j(y)\right) dP_Y(y)$$

$$+ \int_{y:\ell(y)=2}\left(\sum_{j=1}^{2} h_j(y) \cdot \mathbf{1}\left(\frac{1}{2} \le \alpha\right) + \sum_{j=3}^{\infty} h_j(y)\right) dP_Y(y)$$

$$+ \cdots + \int_{y:\ell(y)=L}\left(\sum_{j=1}^{L} h_j(y) \cdot \mathbf{1}\left(\frac{1}{L} \le \alpha\right)\right.$$

$$\left.\left. + \sum_{j=L+1}^{\infty} h_j(y)\right) dP_Y(y)\right].$$

To render this lower bound as large as possible, its formula above indicates that although the multiplicative constant $(1-\alpha)$ favors a small α, the integration term in (a) actually has its smallest value when α is less than $1/L$ (see (b)). Therefore, a compromise in the choice of α has to be made in order to maximize the bound.

III. EXAMPLES

A. Ternary Hypothesis Testing

We revisit the ternary hypothesis testing example examined in [3, Figs. 1 and 2], where random variables X and Y have identical alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$, X is uniformly distributed (i.e., $P_X(x) = 1/3 \,\forall x \in \mathcal{X}$) and Y is related to X via

$$P_{Y|X}(y|x) = \begin{cases} 1 - v_1 - v_2 & \text{if } y = x \\ v_1 & \text{if } x = 1 \text{ and } y = 0 \\ v_2 & \text{if } x = 2 \text{ and } y = 0 \\ v_1 & \text{if } y \ne x \text{ and } y = 1 \\ v_2 & \text{if } y \ne x \text{ and } y = 2 \end{cases}$$

where we assume that $1 - v_1 - v_2 > v_2 > v_1 > 0$. In [3], $v_1 = 0.27$ and $v_2 = 0.33$ are used.

A direct calculation reveals that the MAP estimation function (3) for guessing X from Y is given by $e(y) = y$ for every $y \in \mathcal{Y}$, resulting in a probability of error of $P_e = v_1 + v_2 = 0.6$ when $v_1 = 0.27$ and $v_2 = 0.33$. Furthermore, we obtain that P_e is exactly determined via

$$\lim_{\theta \to \infty} P_{X,Y}\left\{(x,y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \le \alpha\right\} = v_1 + v_2$$
$$= P_e$$

as predicted by Theorem 2, since condition (9) holds (since $\ell(Y) = 1$ almost surely in P_Y, where $\ell(\cdot)$ is defined in (11)).

We next compute the new bound in (4) for $v_1 = 0.27$, $v_2 = 0.33$ and for different values of $\theta \ge 1$ and plot it in Fig. 1, along with Fano's original bound (referred to as "Fano" in the figure) given by

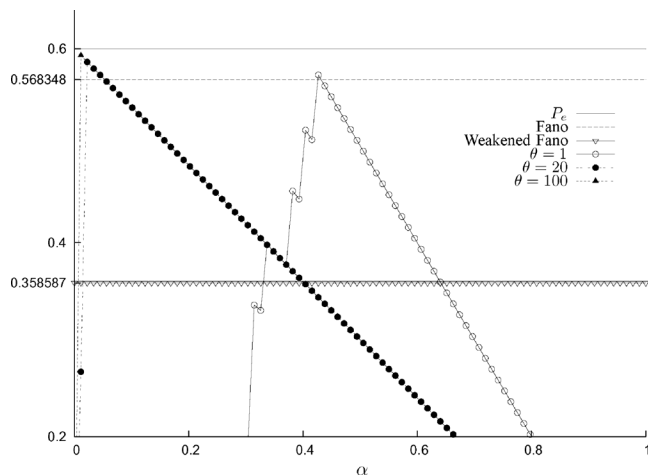$$P_e \ge \frac{\log 3 - I(X;Y) - \log 2}{\log 2} = 0.568348,$$

Fig. 1. Lower bounds on the minimum probability of error for Example III-A: bound (4) versus α for $\theta = 1, 20, 100$ and Fano's original and weakened bounds.
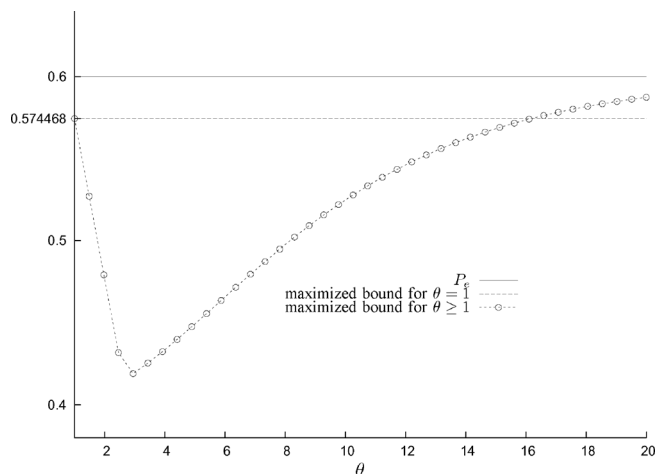

Fig. 2. Lower bounds on the minimum probability of error for Example III-A: bounds (1) and (4) versus θ optimized over α.

and Fano's weaker (but commonly used) bound

$$P_e \geq 1 - \frac{I(X;Y) + \log 2}{\log 3} = 0.358587$$

shown in [3, Fig. 2] (referred to as "Weakened Fano" in the figure). The case of $\theta = 1$ corresponds to the original Poor-Verdú bound in (1). As can be seen from the figure, bound (4) for $\theta = 20$ and 100 improves upon (1) and both Fano bounds and approaches the exact probability of error as θ is increased without bound (e.g., for $\theta = 100$ and $\alpha \downarrow 0$, the bound is quite close to P_e). In Fig. 2, bounds (4) and (1), maximized over $\alpha \in [0, 1]$, are plotted versus θ. It is observed that when $\theta \geq 16$, bound (4) improves upon (1).

B. Binary Erasure Channel

Suppose that X and Y are respectively the channel input and output of a BEC with erasure probability ε, where $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, \mathtt{E}\}$. Let $\Pr[X = 0] = 1 - p$ and $\Pr[X = 1] = p$

with $0 < p < 1/2$. Then, the MAP estimate of X from Y is given by

$$e(y) = \begin{cases} y & \text{if } y \in \{0, 1\} \\ 0 & \text{if } y = \mathtt{E} \end{cases}$$

and the resulting error probability is $P_e = \varepsilon p$.

Calculating bound (4) of Theorem 1 yields

$$(1 - \alpha)P_{X,Y}\left\{(x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \leq \alpha\right\}$$

$$= \begin{cases} 0 & \text{if } 0 \leq \alpha < \frac{p^\theta}{p^\theta + (1-p)^\theta} \\ \varepsilon p(1 - \alpha) & \text{if } \frac{p^\theta}{p^\theta + (1-p)^\theta} \leq \alpha < \frac{(1-p)^\theta}{p^\theta + (1-p)^\theta} \\ \varepsilon(1 - \alpha) & \text{if } \frac{(1-p)^\theta}{p^\theta + (1-p)^\theta} \leq \alpha < 1. \end{cases} \quad (18)$$

Thus, taking $\theta \uparrow \infty$ and then $\alpha \downarrow 0$ in (18) results in the exact error probability εp. Note that in this example, the original Poor-Verdú bound (i.e., with $\theta = 1$) also achieves the exact error probability εp by choosing $\alpha = 1 - p$; however this maximizing choice of $\alpha = 1 - p$ for the original bound is a function of system's statistics (here, the input distribution p) which is undesirable. On the other hand, the generalized bound (4) can herein achieve its peak by systematically taking $\theta \uparrow \infty$ and then letting $\alpha \downarrow 0$.

Furthermore, since in this example, $\ell(y) = 1$ for every $y \in \{0, 1, \mathtt{E}\}$, we have that (9) holds; hence, by Theorem 2, (10) yields

$$P_e = \lim_{\theta \to \infty} P_{X,Y}\left\{(x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \leq \alpha\right\}$$
$$= \varepsilon p \text{ for } 0 \leq \alpha < 1,$$

where the last equality follows directly from (18) without the $(1 - \alpha)$ factor.

C. Binary Input Observed in Gaussian Noise

We herein consider an example with a continuous observation alphabet $\mathcal{Y} = \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers. Specifically, let the observation be given by $Y = X + N$, where X is uniformly distributed over $\mathcal{X} = \{-1, +1\}$ and N is a zero-mean Gaussian random variable with variance σ^2. Assuming that X and N are independent from each other, then for $x \in \{-1, +1\}$ and $y \in \mathbb{R}$

$$P_{X|Y}(x|y) = \frac{1}{1 + \exp\{-\frac{2xy}{\sigma^2}\}}, \quad (19)$$

which directly yields a MAP estimate of X from Y given by $e(y) = +1$ if $y > 0$ and $e(y) = -1$ if $y \leq 0$. with a resulting error probability of $P_e = \Phi(-1/\sigma)$, where $\Phi(z) \triangleq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp-\frac{t^2}{2}dt$ is the cdf of the standard (zero-mean unit-variance) Gaussian distribution. Furthermore, since $x \in \{-1, +1\}$, we have

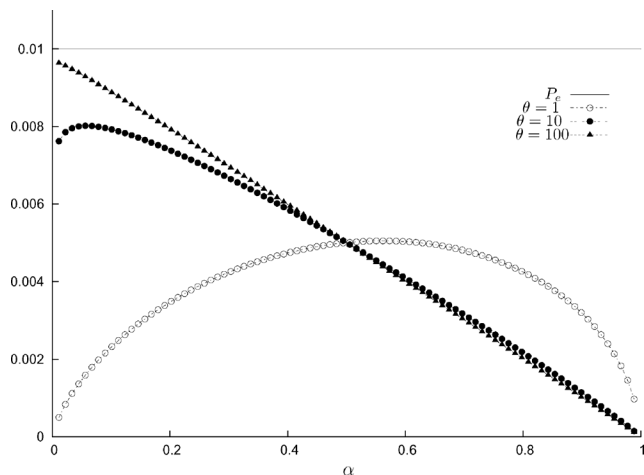$$P_{X|Y}^{(\theta)}(x|y) = \frac{1}{1 + \exp\{-\frac{2xy}{\sigma^2/\theta}\}},$$

Fig. 3. Example III-C: bound (4) versus α for $\theta = 1, 10, 100$; $\sigma = 0.429858$ and $P_e = 0.01$.

and the generalized Poor-Verdú bound (4) yields

$$
P_e \geq (1 - \alpha) P_{X,Y} \left\{ (x, y) \in \mathcal{X} \times \mathcal{Y} : P_{X|Y}^{(\theta)}(x|y) \leq \alpha \right\}
$$
$$
= (1 - \alpha) \int_{-\infty}^{-\frac{\sigma^2}{2\theta} \log\left(\frac{1}{\alpha} - 1\right) - 1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{t^2}{2\sigma^2} \right\} dt
$$
$$
= (1 - \alpha) \Phi \left(-\frac{\sigma}{2\theta} \log\left(\frac{1}{\alpha} - 1 \right) - \frac{1}{\sigma} \right). \quad (20)
$$

Now taking the limits $\theta \uparrow \infty$ followed by $\alpha \downarrow 0$ for the right-hand side term in (20) yields exactly $\Phi\left(-\frac{1}{\sigma}\right) = P_e$; hence the generalized Poor-Verdú bound (4) is asymptotically tight. The bound is illustrated in Fig. 3 for $\sigma = 0.429858$ which gives $P_e = 0.01$. It can be seen that for $\theta = 100$ and $\alpha \downarrow 0$, bound (4) is quite close to P_e. Finally note that (19) directly ascertains that condition (9) of Theorem 2 holds; thus P_e is given by (10).

IV. Concluding Remarks

We generalized the Poor-Verdú lower bound of (1) for the multihypothesis testing error probability. The new bound, given in (4), involves the tilted posterior distribution of the hypothesis given the observation with tilting parameter θ and reduces to the original Poor-Verdú bound when $\theta = 1$. We established a sufficient condition under which the bound (without its multiplicative factor) provides the exact error probability when $\theta \to \infty$. We also provided some examples to illustrate the tightness of the bound in terms of θ.

In [3], Poor and Verdú used (1) to show an information-spectrum based upper bound to the reliability function $E^*(R)$ – i.e., the optimal channel coding error exponent – for general channels [3, Eq. (14)] and they conjectured that the bound is tight. In [1], this bound was shown not to be tight at low rates for memoryless BECs.

We can similarly apply the new bound (4) to obtain a generalized upper bound on $E^*(R)$. Specifically, for a general channel [5], [2] $\mathbf{W} \triangleq \{ W^n = P_{Y^n|X^n} : \mathcal{X}^n \to \mathcal{Y}^n \}_{n=1}^\infty$ with input

$\mathbf{X} \triangleq \{ X^n = (X_1^{(n)}, \ldots, X_n^{(n)}) \}_{n=1}^\infty$ and corresponding output $\mathbf{Y} \triangleq \{ Y^n = (Y_1^{(n)}, \ldots, Y_n^{(n)}) \}_{n=1}^\infty$, we have

$$
E^*(R) \leq \sup_{\mathbf{X} \in \mathcal{Q}(R)} \liminf_{n \to \infty} E_{\mathrm{PV}}^{(\theta,n)}(R, \mathbf{X}) \quad (21)
$$

for every rate $R > 0$ and $\theta \geq 1$ where

$$
E_{\mathrm{PV}}^{(\theta,n)}(R, \mathbf{X}) \triangleq -\frac{1}{n} \log P_{X^n W^n} \{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n :
$$
$$
\frac{1}{n} j_{X^n W^n}^{(\theta)}(x^n; y^n) \leq R \},
$$
$$
\mathcal{Q}(R) \triangleq \Big\{ \mathbf{X} : \text{Each } X^n \text{ in } \mathbf{X} \text{ is uniformly distributed}
$$

over its support $\mathcal{S}(X^n)$, and $R < \liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{S}(X^n)| \Big\}$.

and

$$
j_{X^n W^n}^{(\theta)}(x^n; y^n) \triangleq \log \frac{P_{Y^n|X^n}^\theta(y^n|x^n)}{\sum_{\hat{x}^n \in \mathcal{X}^n} P_{X^n}(\hat{x}^n) P_{Y^n|X^n}^\theta(y^n|\hat{x}^n)}.
$$

Also, in light of (17), the following upper bound holds

$$
E^*(R) \leq \sup_{\mathbf{X} \in \mathcal{Q}(R)} \liminf_{n \to \infty} \lim_{\theta \to \infty} E_{\mathrm{PV}}^{(\theta,n)}(R, \mathbf{X}) \quad (22)
$$

for any $R > 0$. The bound in (22) is exact for a class of channels that satisfy Theorem 2 (which includes the finite-input memoryless Gaussian channel). However, the determination in closed form of this information-spectral multi-letter characterization for $E^*(R)$ remains an open challenging optimization problem.

References

[1] F. Alajaji, P.-N. Chen, and Z. Rached, "A note on the Poor-Verdú upper bound for the channel reliability function," *IEEE Trans. Inf. Theory*, vol. 48, no. 1, pp. 309–313, Jan. 2002.
[2] T. S. Han, *Information-Spectrum Methods in Information Theory*. New York: Springer, 2003.
[3] H. V. Poor and S. Verdú, "A lower bound on the probability of error in multi-hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1992–1994, Nov. 1995.
[4] C. E. Shannon, "Certain results in coding theory for noisy channels," *Inf. Control*, vol. 1, pp. 6–25, Sep. 1957.
[5] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

Po-Ning Chen (S'93–M'95–SM'01) was born in Taipei, Taiwan, in 1963. He received the B.Sc. and M.S. degrees in electrical engineering from the National Tsing-Hua University, Taiwan, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 1994. From 1985 to 1987, he was with Image Processing Laboratory in National Tsing-Hua University, where he worked on the recognition of Chinese characters. During 1989, he was with Star Tech. Inc., where he focused on the development of finger-print recognition systems. After the reception of Ph.D. degree in 1994, he joined Wan Ta Technology Inc. as a vice general manager, conducting several projects on Point-of-Sale systems. In 1995, he became a research staff in Advanced Technology Center, Computer and Communication Laboratory, Industrial Technology Research Institute in Taiwan, where he led a project on Java-based Network Managements.

Since 1996, he has been an Associate Professor in the Department of Communications Engineering at the National Chiao-Tung University, Taiwan, and was promoted to a full professor in 2001. He was elected to be the Chair of the IEEE Communications Society Taipei Chapter in 2006 and 2007, during which the IEEE ComSoc Taipei Chapter won the 2007 IEEE ComSoc Chapter Achievement Awards (CAA) and 2007 IEEE ComSoc Chapter of the Year (CoY). He has served as the chairman of the Department of Communications Engineering, National Chiao-Tung University, during 2007–2009. Dr. Chen received the annual Research Awards from the National Science Council, Taiwan, five years in a row since 1996. He then received the 2000 Young Scholar Paper Award from Academia Sinica, Taiwan. His Experimental Handouts for the course of Communication Networks Laboratory have been awarded as the Annual Best Teaching Materials for Communications Education by the Ministry of Education, Taiwan, in 1998. He has been selected as the Outstanding Tutor Teacher of the National Chiao-Tung University in 2002. He was also the recipient of the Distinguished Teaching Award from the College of Electrical and Computer Engineering, National Chiao-Tung University, Taiwan, in 2003. His research interests generally lie in information and coding theory, large deviation theory, distributed detection and sensor networks.

Fady Alajaji (S'90–M'94–SM'00) received the B.E. degree with distinction from the American University of Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990 and 1994, respectively. He held a postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland.

In 1995, he joined the Department of Mathematics and Statistics at Queen's University, Kingston, Ontario, where he is currently a Professor of Mathematics and Engineering. Since 1997, he has also been cross-appointed in the Department of Electrical and Computer Engineering at the same university. From 2003 to 2008, he served as chair of the Queen's Mathematics and Engineering program. His research interests include information theory, digital communications, error control coding, joint source-channel coding and data compression.

Dr. Alajaji currently serves as Area Editor and Editor for Source and Source-Channel Coding for the IEEE TRANSACTIONS ON COMMUNICATIONS. He served as co-chair of the 1999 Canadian Workshop on Information Theory, as co-chair of the Technical Program Committee (TPC) of the 2004 Biennial Symposium on Communications and as a TPC member of several international conferences and workshops. He received the Premier's Research Excellence Award from the Province of Ontario.