

Worst Case Additive Noise for Binary-Input Channels and Zero-Threshold Detection under Constraints of Power and Divergence

Andrew L. McKellips, *Student Member, IEEE*,
and Sergio Verdú, *Fellow, IEEE*

Abstract— Additive-noise channels with binary inputs and zero-threshold detection are considered. We study worst case noise under the criterion of maximum error probability with constraints on both power and divergence with respect to a given symmetric nominal noise distribution. Particular attention is focused on the cases of a) Gaussian nominal distributions and b) asymptotic increase in worst case error probability when the divergence tolerance tends to zero.

Index Terms—Detection, Gaussian error probability, hypothesis testing, Kullback–Leibler divergence, least favorable noise.

I. INTRODUCTION

Consider a binary-input channel with additive noise N and the associated hypothesis test

$$\begin{aligned} H_0 : Y &= +1 + N \\ H_1 : Y &= -1 + N \end{aligned} \quad (1)$$

where N is a random variable with probability density function (pdf) f_N . The probability of error achieved by a zero-threshold detector is given by

$$P = \frac{1}{2} \left\{ \int_{-\infty}^{-1} f_N(x) dx + \int_1^{\infty} f_N(x) dx \right\}. \quad (2)$$

In many situations of interest, the distribution of N is not known exactly, and we wish to consider worst case performance for a certain class of noise distributions.

The case where the uncertainty class is the set of all noise distributions bounded in second moment (power) is particularly interesting as it quantifies the worst case probability of error as a function of the channel signal-to-noise ratio (SNR). This case was considered in [1] for “very noisy” channels. A full solution for the maximum-likelihood detection problem was obtained in [2] where it was shown that the worst case error probability is given by

$$\begin{aligned} P(\sigma^2) &= \frac{1}{2} - \frac{1}{2\sqrt{1+3\sigma_k^2}} + \frac{3(\sigma^2 - \sigma_k^2)}{2k(k+1)(2k+1)}, \\ \sigma_k^2 &\leq \sigma^2 \leq \sigma_{k+1}^2, \quad \sigma_k^2 = \frac{1}{3}(k^2 - 1), \quad k = 1, 2, \dots \end{aligned} \quad (3)$$

where σ^2 is the noise power. It was also shown in [2] that the least favorable noise distribution achieving (3) is a mixture of two equiprobable distributions, each taking values on a span-2 lattice. An extension to the zero-sum error probability game between communicator and noise where the communicator is allowed to transmit an antipodal signal with pseudorandom amplitude pattern known only to the receiver is considered in [3].

Manuscript received November 3, 1995; revised December 23, 1996. This work was supported in part by the U.S. Army Research Office under Grant DAAH04-96-1-0379. The material in this correspondence was presented in part at the 33rd Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, October 1995.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA.

Publisher Item Identifier S 0018-9448(97)03873-X.

In many situations, there is some available information about the unknown noise distribution in addition to the power, leading to the consideration of a smaller uncertainty class centered around a nominal distribution. In particular, we can consider the uncertainty class defined by the intersection of the set of distributions with power not exceeding that of the nominal and the set of distributions differing from the nominal by no more than a specified amount, where the discrepancy between noise distributions is quantified in a specific manner. In this correspondence we use the (Kullback–Leibler) divergence (see [4], for instance) as a measure of distance between the worst case and nominal noise distributions. We will study the worst case error probability of the zero-threshold detector achievable by any noise distribution constrained in both power and divergence with respect to a given symmetric nominal distribution. It should be noted that, although the zero-threshold detector is optimum for the important special class of symmetric unimodal distributions, it does not follow that the pairing we obtain of threshold detector and least favorable noise forms a saddle point for the game where the detector is unconstrained; indeed, it is not a saddle point for nominals with symmetric continuous pdf's, since, as shown in Section II-A, the zero-threshold detector is not a maximum-likelihood detector for the least favorable distribution in this case.

In addition to its intrinsic interest, our solution has applications to problems in data communications subject to intersymbol interference, crosstalk, multiuser interference, or jamming, where the decision statistic is the sum of a signal component, a noise component which is easy to characterize, and an interference component whose distribution is hard to characterize, but such that the divergence between noise plus interference and noise can be studied analytically. For instance, a bound on the divergence-from-Gaussian of the total interference in a multiuser system with a linear minimum-mean-square-error (MMSE) transformation followed by a zero-threshold detector is obtained in [5].

Another direct application is the scenario where a jammer wishes to avoid being detected while degrading error performance in an additive-noise channel wherein the receiver makes zero-threshold decisions. Using training data sequences, the decision as to whether or not a jammer is active is based on the hypothesis test

$$\begin{aligned} H_0 : f_N \\ H_1 : f_{\bar{N}} \end{aligned}$$

where \bar{N} and N represent the channel noise with and without a jammer, respectively. The discernibility of this test can be quantified through the divergence between \bar{N} and N , making our result the best a jammer can do for a fixed acceptable detection level.

The general form of worst case noise is presented in Section II. A related problem of interest which is also investigated in that section arises when the noise uncertainty class is enlarged by dropping the power constraint. Section III studies the worst case error probability when the divergence constraint tends to zero; this reflects the common situation where the noise is known to be “very close” to some specified nominal. Throughout the correspondence, special emphasis is placed on the important case of Gaussian nominal pdf's.

II. WORST CASE NOISE

A. Power- and Divergence-Constrained pdf's

Given a choice of nominal pdf f_N and divergence tolerance δ , the optimization problem we face is that of finding the least favorable

noise \hat{N} with pdf $f_{\hat{N}}$ satisfying

$$\begin{aligned} \hat{P} &= \frac{1}{2} \left\{ \int_{-\infty}^{-1} f_{\hat{N}}(x) dx + \int_1^{\infty} f_{\hat{N}}(x) dx \right\} \\ &= \max_{f_{\mathcal{N}}} \frac{1}{2} \left\{ \int_{-\infty}^{-1} f_{\mathcal{N}}(x) dx + \int_1^{\infty} f_{\mathcal{N}}(x) dx \right\} \end{aligned} \quad (4)$$

subject to the constraints

$$\int_{-\infty}^{\infty} f_{\mathcal{N}}(x) dx = 1 \quad (5)$$

$$\int_{-\infty}^{\infty} x^2 f_{\mathcal{N}}(x) dx \leq \sigma^2 \quad (6)$$

$$D(f_{\mathcal{N}} \| f_N) = \int_{-\infty}^{\infty} f_{\mathcal{N}}(x) \log \frac{f_{\mathcal{N}}(x)}{f_N(x)} dx \leq \delta \quad (7)$$

where \hat{P} denotes the associated worst case probability of error.

Taking note from (2) that a zero-threshold error probability is restricted to the interval $[0, 1/2]$, we demonstrate in Appendix I that the constraints (6) and (7) are active whenever $0 < \hat{P} < 1/2$. We also demonstrate that the extreme case $\hat{P} = 0$ is trivially uninteresting and that the extreme case $\hat{P} = 1/2$ cannot occur in a channel with SNR > 0 dB, and hence restrict attention to the case $0 < \hat{P} < 1/2$. Our first observation concerns symmetry of $f_{\hat{N}}$ for symmetric nominal pdf's.

Proposition 2.1: Given a symmetric nominal pdf f_N , the worst case pdf $f_{\hat{N}}$ in (4) is also symmetric.

Proof: Assume by contradiction that $f_{\hat{N}}$ is not symmetric. Note that

$$D(f_{\hat{N}}(x) \| f_N(x)) = D(f_{\hat{N}}(-x) \| f_N(x)) = \delta$$

by the assumed symmetry of f_N and demonstrated activity of the divergence constraint. Define the pdf

$$f'_{\hat{N}}(x) = (f_{\hat{N}}(x) + f_{\hat{N}}(-x))/2.$$

Note that the second moment of $f'_{\hat{N}}$ is equal to that of $f_{\hat{N}}$ and that the achieved zero-threshold error probability matches that of $f_{\hat{N}}$ through (2). By strict convexity of the divergence measure $D(\cdot \| \cdot)$ in the first entry (see [4], for instance) we have that

$$\begin{aligned} D(f'_{\hat{N}}(x) \| f_N(x)) &< \frac{1}{2} \{ D(f_{\hat{N}}(x) \| f_N(x)) \\ &+ D(f_{\hat{N}}(-x) \| f_N(x)) \} = \delta \end{aligned}$$

which leads to the desired contradiction in view of the demonstrated activity of the divergence constraint. \square

We assume throughout that the nominal pdf f_N is symmetric, and restrict attention to symmetric feasible pdf's as a result of Proposition 2.1. Note that the nominal error probability is then given by

$$P = \int_1^{\infty} f_N(x) dx$$

and the worst case error probability by

$$\hat{P} = \int_1^{\infty} f_{\hat{N}}(x) dx.$$

Consider the Lagrangian

$$\begin{aligned} L(f_{\mathcal{N}}, \lambda_1, \lambda_2, \lambda_3) &= \int_{-\infty}^{\infty} f_{\mathcal{N}}(x) \left\{ \frac{1}{2} \mathbf{1}_{\{|x| \geq 1\}} + \lambda_1 + \lambda_2 x^2 + \lambda_3 \log \frac{f_{\mathcal{N}}(x)}{f_N(x)} \right\} dx \end{aligned} \quad (8)$$

corresponding to (4), where $\mathbf{1}_{\{\cdot\}}$ represents the standard indicator function. The Lagrangian (8) is maximized by

$$\begin{aligned} f_{\hat{N}}(x) &= f_N(x) \exp \{ -\lambda_2 x^2 / \lambda_3 \} \\ &\times \exp \{ -(2\lambda_1 + \mathbf{1}_{\{|x| \geq 1\}} + 2\lambda_3) / 2\lambda_3 \} \end{aligned} \quad (9)$$

which is justified since λ_3 is nonzero in light of the demonstrated activity of the divergence constraint [6, ch. 9]. An immediate consequence of (9) is that the optimal form obtained by ignoring the pointwise constraint $f_{\hat{N}}(x) \geq 0$ always leads to nonnegative-valued realizations of $f_{\hat{N}}$, justifying the relaxing of this constraint. Further analysis also reveals that worst case power-constrained noise, resulting from the dropping of the divergence constraint (7), can always be realized by a probability mass function with mass restricted to the set of three points $\{-1, 0, 1\}$, and is hence easy to study.

We can rewrite (9) more intuitively and with full generality as

$$f_{\hat{N}}(x) = \begin{cases} (1 - \bar{\eta}_1) f_{\hat{N}}(x), & |x| < 1 \\ (1 + \bar{\eta}) f_{\hat{N}}(x), & |x| \geq 1 \end{cases} \quad (10)$$

where $f_{\hat{N}}$ is a proper pdf, expressible as

$$f_{\hat{N}}(x) = C f_N(x) \exp \{ -c x^2 \} \quad (11)$$

$$C = \left(2 \int_0^{\infty} f_N(x) \exp \{ -c x^2 \} dx \right)^{-1} \quad (12)$$

and $c, \bar{\eta}$, and $\bar{\eta}_1$ are positive constants, which depend on the nominal f_N and choice of divergence tolerance δ . If we denote by \bar{P} the zero-threshold probability of error associated with the additive noise \bar{N} , then the constraint (5) yields

$$\bar{\eta}_1 = \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \quad (13)$$

$$\bar{P} = \int_1^{\infty} C f_N(x) \exp \{ -c x^2 \} dx. \quad (14)$$

The divergence constraint (7) can now be rewritten under activity in the equivalent form

$$\begin{aligned} (1 - 2\bar{P}) \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \log \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \\ + 2\bar{P}(1 + \bar{\eta}) \log(1 + \bar{\eta}) \\ + \log C - 2c \int_0^{\infty} x^2 f_{\hat{N}}(x) dx = \delta \end{aligned} \quad (15)$$

after some straightforward manipulation.

An important observation is that the term

$$2c \int_0^{\infty} x^2 f_{\hat{N}}(x) dx$$

in (15) is governed by the independent active power constraint (6), so that we can rewrite (15) as

$$\begin{aligned} (1 - 2\bar{P}) \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \log \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \\ + 2\bar{P}(1 + \bar{\eta}) \log(1 + \bar{\eta}) + \log C - c\sigma^2 = \delta. \end{aligned} \quad (16)$$

So, given a prescribed nominal distribution f_N and divergence tolerance δ , we can determine the worst case noise distribution $f_{\hat{N}}$ through the use of (13) and (16). Given an initial value of c , $\bar{\eta}_1$ is related to $\bar{\eta}$ through (13) using the definite integrals (12) and (14). Then $\bar{\eta}$ is related to δ through (16). The proper value of c is determined through (6) under activity by the relation

$$2 \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) B_0 + 2(1 + \bar{\eta}) B_1 = \sigma^2 \quad (17)$$

where

$$B_0 = C \int_0^1 x^2 f_N(x) \exp \{ -c x^2 \} dx \quad (18)$$

$$B_1 = C \int_1^{\infty} x^2 f_N(x) \exp \{ -c x^2 \} dx. \quad (19)$$

Each choice of c involves the computation of the four definite integrals (12), (14), (18), and (19). Upon solving the above system, the worst case error probability \tilde{P} is given by

$$\tilde{P} = \int_1^\infty f_{\tilde{N}}(x) dx = (\bar{\eta} + 1)\bar{P}$$

as evidenced by (10). A sample of numerical results is to be found in Section II-D.

The form of the worst case noise pdf $f_{\tilde{N}}$ leads to the following interesting observation.

Proposition 2.2: For any symmetric-continuous pdf f_N and any divergence tolerance $\delta > 0$, the zero-threshold detector falls outside the class of maximum-likelihood detectors for the worst case noise pdf $f_{\tilde{N}}$.

Proof: From (11) it is clear that $f_{\tilde{N}}$ is continuous whenever f_N is continuous. It follows directly from (10) that there exists $\varepsilon > 0$ such that $f_{\tilde{N}}(x) < f_{\tilde{N}}(y)$ for all $1 - \varepsilon \leq |x| < 1$ and $1 < |y| \leq 1 + \varepsilon$. With reference to the hypothesis test (1), any maximum-likelihood detector must, therefore, decide H_0 if $-\varepsilon \leq Y < 0$ and H_1 if $0 < Y \leq \varepsilon$, precluding maximum-likelihood performance for the zero-threshold detector. \square

A direct implication of Proposition 2.2 is that the pairing of a zero-threshold detector and worst case noise does not form a saddle point for the game of noise versus unconstrained detector for symmetric-continuous nominal pdf's.

B. Divergence-Constrained pdf

We turn attention to the determination of worst case divergence-constrained noise, where we drop the power constraint (6). That is, we wish to find $f_{\tilde{N}}$ such that

$$\begin{aligned} \tilde{P} &= \frac{1}{2} \left\{ \int_{-\infty}^{-1} f_{\tilde{N}}(x) dx + \int_1^{\infty} f_{\tilde{N}}(x) dx \right\} \\ &= \max_{f_{\tilde{N}}} \frac{1}{2} \left\{ \int_{-\infty}^{-1} f_{\tilde{N}}(x) dx + \int_1^{\infty} f_{\tilde{N}}(x) dx \right\} \end{aligned} \quad (20)$$

subject to

$$\int_{-\infty}^{\infty} f_{\tilde{N}}(x) dx = 1 \quad (21)$$

$$\int_{-\infty}^{\infty} f_{\tilde{N}}(x) \log \frac{f_{\tilde{N}}(x)}{f_N(x)} dx \leq \delta. \quad (22)$$

A useful observation is that the solution \tilde{P} provides an upper bound for \tilde{P} in the original problem (4), a result of the enlarged feasible class. We will see in this section that solving for \tilde{P} is a simpler task, and provides some idea of the accuracy of \tilde{P} when used as an approximation for \tilde{P} with a sample of numerical results in Section II-D.

As in the analysis of Section II-A, it is easy to show that the divergence constraint is active for $0 < \tilde{P} < 1/2$, and that the worst case pdf $f_{\tilde{N}}$ is symmetric for any symmetric nominal pdf f_N . Subsequent Lagrange-multiplier analysis yields the optimal form

$$f_{\tilde{N}}(x) = \begin{cases} K_1 f_N(x), & |x| < 1 \\ K_2 f_N(x), & |x| \geq 1 \end{cases} \quad (23)$$

where K_1 and K_2 are constants, which allows us to derive general results for \tilde{P} dependent on f_N only through the nominal error probability P .

Setting $K_2 = 1 + \eta$, and solving for K_1 using the constraint (21), we rewrite (23) as

$$f_{\tilde{N}}(x) = \begin{cases} (1 - \eta \frac{2P}{1-2P}) f_N(x), & |x| < 1 \\ (1 + \eta) f_N(x), & |x| \geq 1 \end{cases} \quad (24)$$

which, when subjected to the active divergence constraint, yields the single-variable two-parameter equation

$$(1 - 2P) \left(1 - \eta \frac{2P}{1-2P} \right) \log \left(1 - \eta \frac{2P}{1-2P} \right) + 2P(1 + \eta) \log(1 + \eta) = \delta \quad (25)$$

to solve for η .

As before, the special case $\tilde{P} = 0$ holds if and only if the nominal pdf f_N exhibits $P = 0$, and is hence trivially uninteresting. The case $\tilde{P} = 1/2$ requires more care than in the power-constrained analysis, since the divergence constraint may be inactive even for channels with an SNR > 0 dB. As before, there will exist for any nominal pdf with $P > 0$ some value δ_0 for which the worst case error probability is $\tilde{P} = 1/2$ under an active divergence constraint; for any value $\delta > \delta_0$, there will exist a class of worst case pdf's corresponding to the divergence tolerance δ_0 . A quick inspection of (25) shows that $\delta_0 = \log(1/2P)$, allowing us to rewrite the solution with complete generality as (24) with

$$(1 - 2P) \left(1 - \eta \frac{2P}{1-2P} \right) \log \left(1 - \eta \frac{2P}{1-2P} \right) + 2P(1 + \eta) \times \log(1 + \eta) = \delta_0, \quad \delta_0 = \min \left\{ \delta, \log \frac{1}{2P} \right\} \quad (26)$$

for all nominals f_N exhibiting $P > 0$.

Having solved (26) for a given P and δ , we can write

$$\tilde{P} = (\eta + 1)P \quad (27)$$

as the solution to (20) and as an upper bound on \tilde{P} . In terms of computational efficiency, it is important to note that (26) depends on f_N only through P . An interesting consequence is that all nominal random variables with the same zero-threshold error probability P exhibit the same worst case divergence-constrained probability of error \tilde{P} .

Our hope in utilizing the divergence-constrained solution \tilde{P} as an upper bound approximation to \tilde{P} in (4) is that the power constraint (6) has a significantly lower relative cost than the divergence constraint (7), rendering the upper bound relatively tight. Indeed, we will see in a later section that the power constraint is actually extraneous for nominal pdf's satisfying certain decay restrictions in the asymptotic case $\delta \rightarrow 0$.

C. Gaussian Nominals

The most interesting and natural choice of nominal noise is Gaussian, in which case we can significantly reduce the computation involved with determining worst case noise as outlined in Section II-A. Let N_σ be a zero-mean Gaussian random variable with variance σ^2 . A significant savings in computation stems from the fact that the pdf

$$f_{\tilde{N}}(x) = C f_{N_\sigma}(x) \exp\{-cx^2\}$$

as described by (11) is also Gaussian with variance $\bar{\sigma}^2 = \sigma^2(1 + 2c\sigma^2)^{-1}$. In other words, the worst case noise $f_{\tilde{N}}$ derives from first transforming the nominal Gaussian pdf with variance σ^2 into a Gaussian pdf with variance $\bar{\sigma}^2 < \sigma^2$, then performing the perturbation described in (10), in so doing increasing the second moment back to σ^2 . This observation allows us to avoid the computation of C by the integral equation (12), where we can now say

$$C = \sqrt{1 + 2c\sigma^2}$$

and allows us to express (14) as

$$\tilde{P} = Q \left(\frac{\sqrt{1 + 2c\sigma^2}}{\sigma} \right) \quad (28)$$

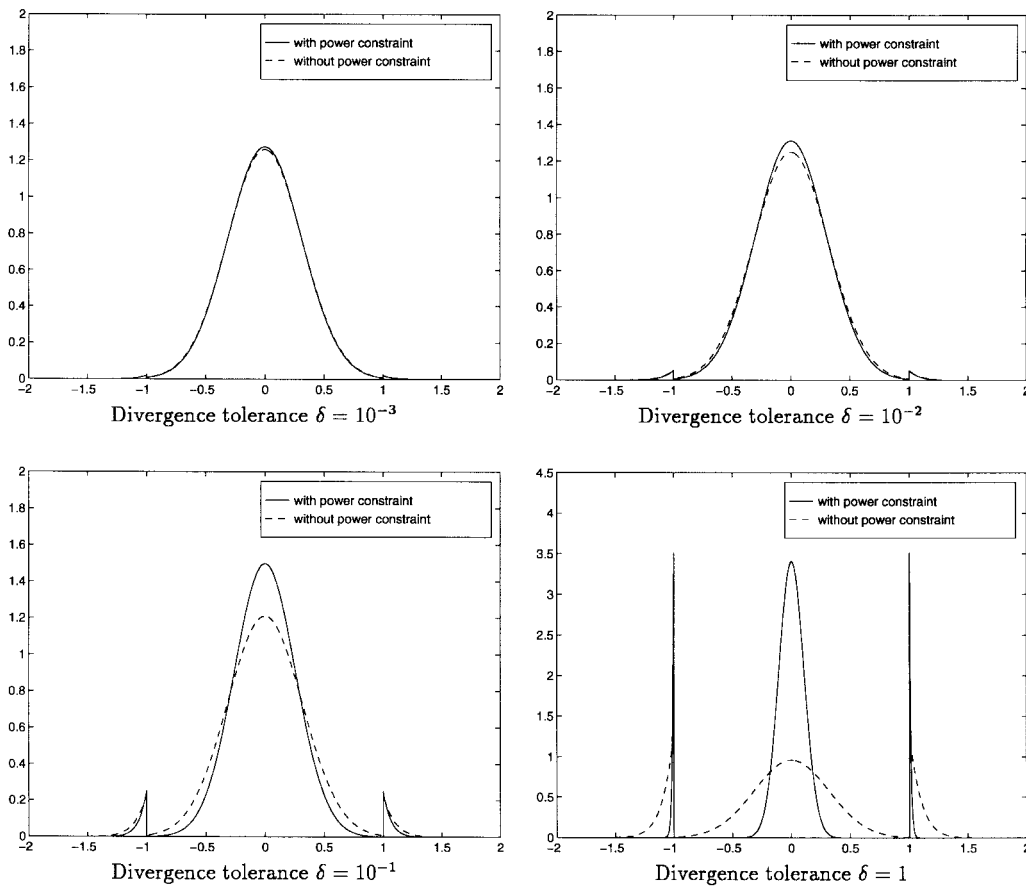


Fig. 1. Worst case zero-threshold pdf's for a Gaussian nominal, SNR = 10 dB.

where $Q(\cdot)$ is the well-studied tail of the standard Gaussian distribution. Furthermore, the quantities B_0 and B_1 as defined, respectively, in (18) and (19) satisfy the relation

$$\begin{aligned} 2(B_0 + B_1) &= 2 \int_0^\infty x^2 f_{N_{\bar{\sigma}}}(x) dx \\ &= \bar{\sigma}^2 \\ &= \frac{\sigma^2}{1 + 2c\sigma^2} \end{aligned}$$

reducing the computation of B_0 and B_1 to a single well-studied definite integral.

These observations reduce to two, the number of definite integrals that need be computed for each choice of c , both of which avail themselves to well-studied numerical approximations. To find $f_{N_{\bar{\sigma}}}$, we choose $\bar{\sigma}^2 = \sigma^2(1 + 2c\sigma^2)^{-1}$ and solve for $\bar{\eta}$ using the equation

$$\begin{aligned} (1 - 2\bar{P}) \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \log \left(1 - \bar{\eta} \frac{2\bar{P}}{1 - 2\bar{P}} \right) \\ + 2\bar{P}(1 + \bar{\eta}) \log(1 + \bar{\eta}) + \frac{1}{2} \log \frac{\sigma^2}{\bar{\sigma}^2} - c\sigma^2 = \delta \end{aligned} \quad (29)$$

as specified by (16), where $\bar{P} = Q(1/\bar{\sigma})$. The value of $\bar{\sigma}^2$ associated with the solution is that which meets the power constraint

$$2\bar{\eta} \left[B_1 - \frac{\bar{P}}{1 - 2\bar{P}} (\bar{\sigma}^2 - 2B_1) \right] = \sigma^2 - \bar{\sigma}^2 \quad (30)$$

where

$$B_1 = \int_1^\infty x^2 f_{N_{\bar{\sigma}}}(x) dx.$$

As above, the worst case error probability is given by

$$\hat{P} = (\bar{\eta} + 1)\bar{P}.$$

If the power constraint (6) is dropped, the solution for worst case divergence-constrained noise is simply given by (24) and (26) with $P = Q(1/\sigma)$.

D. Numerical Results

Worst case pdf's are presented in Fig. 1 for a Gaussian nominal with SNR = 10 dB and a variety of divergence tolerance values using the results of Sections II-A and II-B. Note that the power constraint precludes a full transportation of mass out of the interval $[-1, 1]$ in the power-constrained case. As the divergence tolerance grows unbounded, the worst case power- and divergence-constrained pdf will approach a three-point probability mass function with weight 0.05 at $x = \pm 1$ and weight 0.9 at $x = 0$, while the divergence-constrained pdf will, for large enough δ , take the form of weighted Gaussian tails void of mass in the interval $[-1, 1]$.

Similar curves are presented in Fig. 2 for a Gaussian nominal with SNR = 0 dB. In this case, full transportation of mass out of the interval $[-1, 1]$ is asymptotically achievable in the power-constrained case because of the increased nominal power; note that for $\delta = 1$ there is positive mass in the interval $[-1, 1]$ although it is not discernible at the resolution of the graph. As the divergence tolerance grows unbounded, the worst case power- and divergence-constrained pdf will approach a two-point probability mass function with equal weights at the points $x = \pm 1$. The divergence tolerance $\delta = 1$ is sufficiently large for the worst case divergence-constrained noise pdf to achieve the upper bound on zero-threshold error probability of $1/2$.

In order to develop a feel for the relative costs of the power and divergence constraints, Figs. 3–8 depict worst case error probabilities as computed via the analysis of this section for a variety of nominal pdf's and values of SNR and δ .

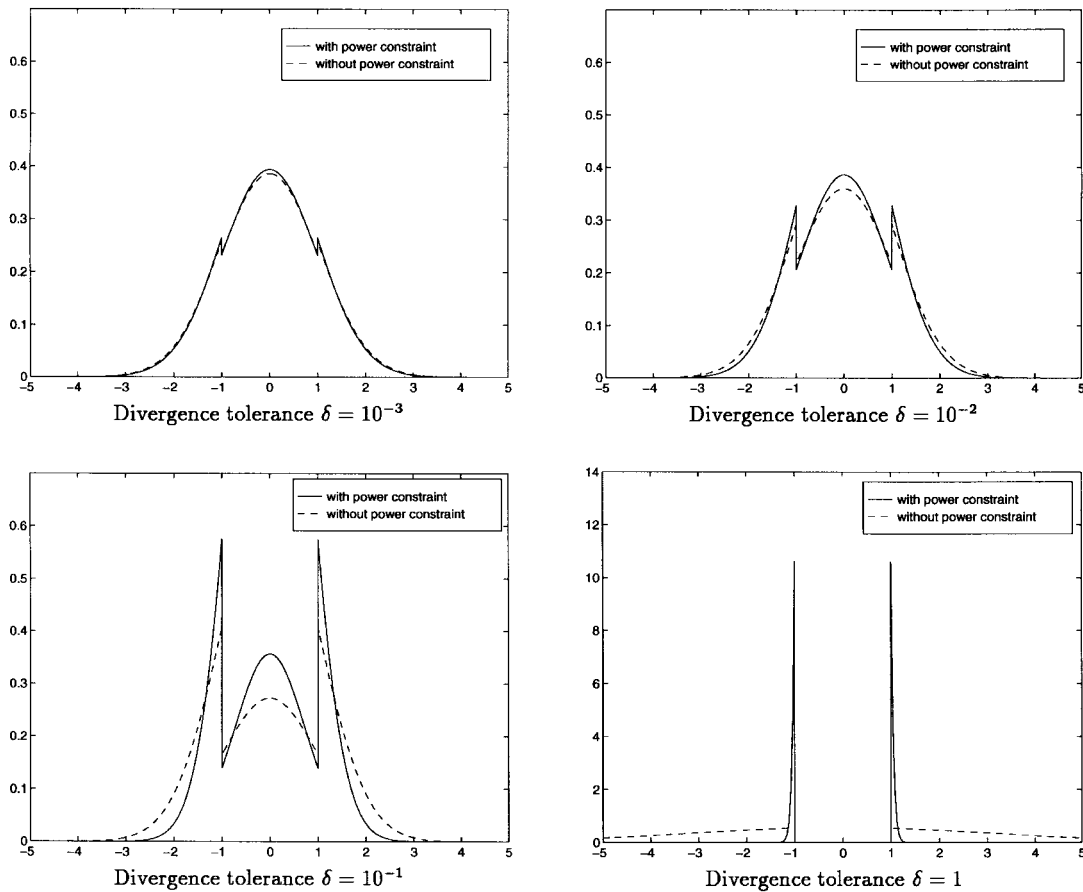


Fig. 2. Worst case zero-threshold PDF's for a Gaussian nominal, SNR = 0 dB.

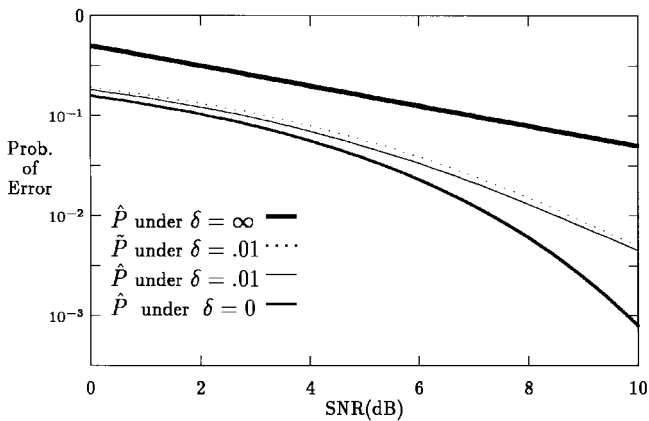


Fig. 3. Worst case error probability for Gaussian nominals, $\delta = 0.01$.

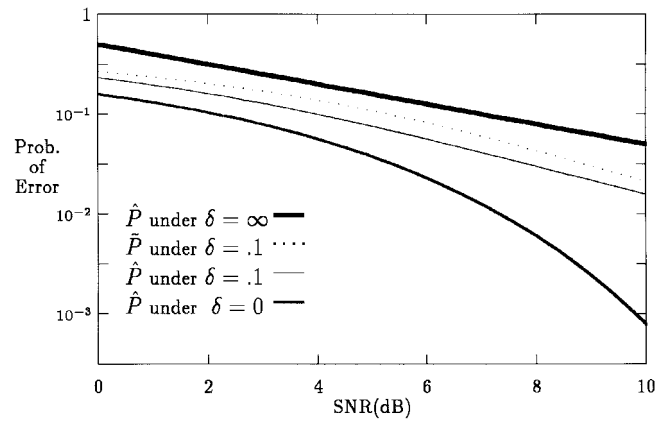


Fig. 4. Worst case error probability for Gaussian nominals, $\delta = 0.1$.

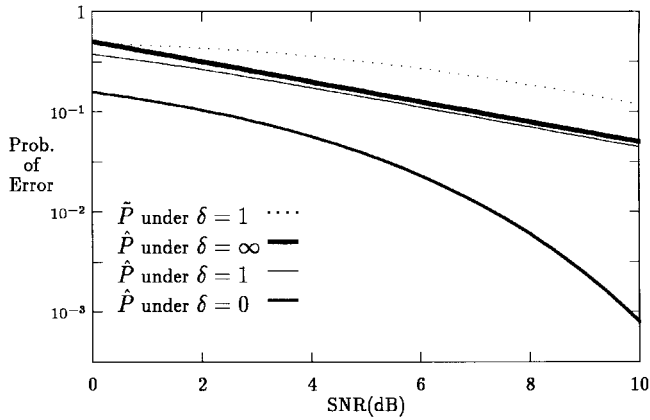
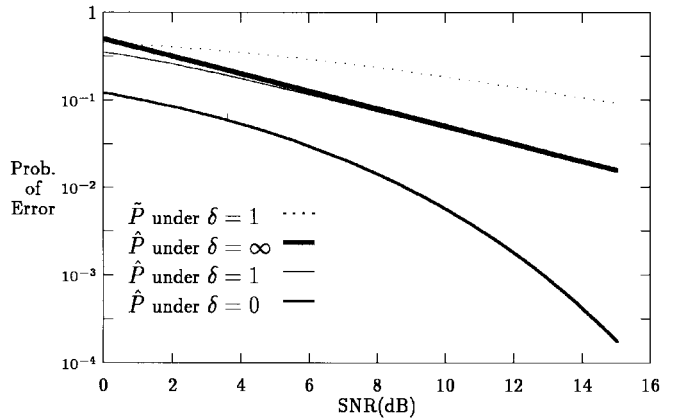
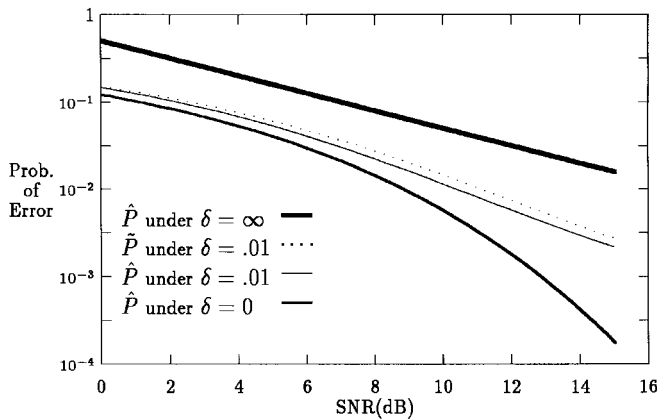
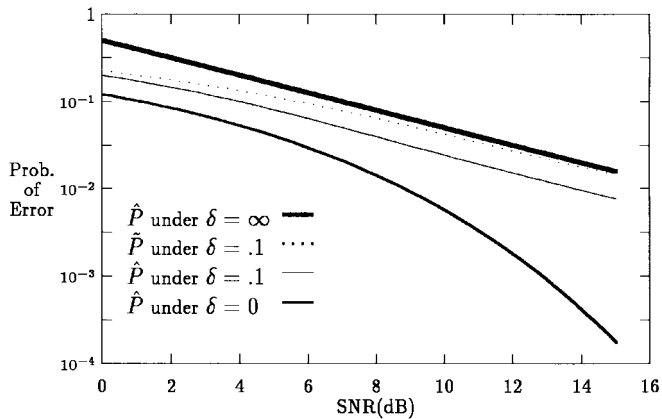
Figs. 3–5 depict worst case probability of error for Gaussian nominals over a range of SNR values. Each graph depicts curves corresponding to (4) under $\delta = 0$ and $\delta = +\infty$ for reference. Note that $\hat{P} = P$ for $\delta = 0$ and recall that worst case noise takes the form of a three-point probability mass function under $\delta = +\infty$ as discussed in Section II-A.

The graphs show that the effect of the imposed power constraint grows with increasing δ , which is expected since a larger transportation of mass will lead to a greater increase in second moment under transformations (10) and (24). In Fig. 5, it is interesting to note that the power constraint is actually more restrictive than the divergence constraint when $\delta = 1$, as evidenced by the relative positions of the curves corresponding to \hat{P} and the reference \hat{P} for $\delta = +\infty$.

Figs. 6–8 depict similar curves for the case of a Laplacian nominal with pdf

$$f_{N_\sigma}(x) = \frac{1}{\sigma^2} \exp\{-2|x|/\sigma^2\}.$$

Note that the effect of the imposed power constraint is more pronounced than in the Gaussian case. This can be explained intuitively by the fact that the tail of a Laplacian pdf contributes more to its second moment than that of a Gaussian; since the increase in second moment under either transformation (10) or (24) takes place in the tail, such effects are more pronounced in the Laplacian case.


 Fig. 5. Worst case error probability for Gaussian nominals, $\delta = 1$.

 Fig. 8. Worst case error probability for Laplacian nominals, $\delta = 1$.

 Fig. 6. Worst case error probability for Laplacian nominals, $\delta = 0.01$.

 Fig. 7. Worst case error probability for Laplacian nominals, $\delta = 0.1$.

III. ASYMPTOTIC BEHAVIOR

We turn attention to the asymptotic behavior of the worst case error probability achieved by the zero-threshold detector as the divergence tolerance δ tends to zero in (4). This will convey some idea of the expected worst case degradation in channel performance when the noise is “very close” to a given nominal, for instance the typical scenario of near-Gaussianness.

A. Divergence-Constrained

As before, it is interesting to consider worst case noise subject to a lone divergence constraint, without the power constraint (6). The goal here is to quantify the asymptotic behavior of the worst case

divergence-constrained error degradation as the divergence tolerance tends to zero.

Recall the expression for worst case divergence-constrained error probability developed in Section II-B, namely,

$$\tilde{P} = (\eta + 1)P$$

where P represents the nominal error probability and η satisfies (25) in full generality since we are taking $\delta \rightarrow 0$ and can hence assume $\delta < \delta_0$ in (26). By inspection of (25), we note that η tends to zero with δ , and for $\eta < 1$ we can rewrite (25) as

$$\begin{aligned} (1 - 2P) \left(1 - \eta \frac{2P}{1 - 2P} \right) \\ \times \left(-\eta \frac{2P}{1 - 2P} - \frac{\eta^2}{2} \left(\frac{2P}{1 - 2P} \right)^2 - \frac{\eta^3}{3} \left(\frac{2P}{1 - 2P} \right)^3 - o(\eta^4) \right) \\ + 2P(1 + \eta) \left(\eta - \frac{\eta^2}{2} + \frac{\eta^3}{3} - o(\eta^4) \right) = \delta \end{aligned}$$

which leads directly to the relation

$$\lim_{\delta \rightarrow 0} \frac{\eta}{\sqrt{\delta}} = \sqrt{\frac{1 - 2P}{P}}. \quad (31)$$

If we denote by ΔP_U the worst case divergence-constrained error degradation $\Delta P_U = \tilde{P} - P$, (31) yields the relation

$$\lim_{\delta \rightarrow 0} \frac{\Delta P_U}{\sqrt{\delta}} = P \lim_{\delta \rightarrow 0} \frac{\eta}{\sqrt{\delta}} = \sqrt{P(1 - 2P)}. \quad (32)$$

Note that (32) depends on the nominal pdf f_N only through the nominal error probability P , so that all nominal pdf's exhibiting the same nominal error probability P also exhibit the same asymptotic increase in worst case divergence-constrained error probability.

B. Power- and Divergence-Constrained pdf's: Upper Bound

An upper bound on worst case asymptotic error degradation subject to both power and divergence constraints follows directly from the fact that \tilde{P} (no power constraint) provides an upper bound for \hat{P} (imposed power constraint). We denote by ΔP the worst case degradation $\hat{P} - P$ and form from (32) the upper bound

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} \leq \lim_{\delta \rightarrow 0} \frac{\Delta P_U}{\sqrt{\delta}} = \sqrt{P(1 - 2P)}. \quad (33)$$

We now turn attention to a class of nominal pdf's for which we can demonstrate equality of the upper bound (33), providing an exact description of the asymptotic behavior of the worst case error

degradation ΔP . We restrict our consideration to nominal random variables N for which the following conditions hold:

$$\text{There exists } x^* \text{ such that } f_N(x) > 0 \text{ for all } x \geq x^* \quad (34)$$

and

$$\lim_{\bar{x} \rightarrow \infty} \bar{x}^2 \int_{\bar{x}}^{\infty} x^2 f_N(x) dx = \infty. \quad (35)$$

Condition (35) limits the rate of decay of the tail of the pdf $f_N(x)$, and is only satisfied by pdf's exhibiting decay less rapid than $o(1/x^5)$; it hence does not hold for the important class of Gaussian nominals, which we deal with in the following section. Condition (35) enables us to satisfy the power constraint (6) by transporting mass from the tail in a manner which negligibly affects the divergence, providing a construction which achieves worst case asymptotic error degradation equal to the upper bound (33).

Theorem 3.1: Given a nominal f_N satisfying conditions (34) and (35)

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} = \sqrt{P(1-2P)}. \quad (36)$$

Proof: A proof appears in Appendix II.

A direct consequence of Theorem 3.1 is that the power constraint (6) becomes extraneous as $\delta \rightarrow 0$ for nominals satisfying (34) and (35).

To gain further insight into the need for condition (35), we present in Appendix II an analysis of worst case noise based on a specific choice of discrete nominal distribution, where condition (35) is shown to be nonsuperfluous in order for (36) to hold. In this example

$$\lim_{\delta \rightarrow 0} \Delta P / \sqrt{\delta} = \sqrt{P(1-2P)}$$

if (35) holds, while

$$\lim_{\delta \rightarrow 0} \Delta P / \sqrt{\delta} = 0$$

if this condition is not satisfied. We will see later that in the case of Gaussian nominals, the quantity $\lim_{\delta \rightarrow 0} \Delta P / \sqrt{\delta}$ takes on strictly positive values, although these values may fall short of the quantity $\sqrt{P(1-2P)}$.

C. Power- and Divergence-Constrained pdf's:

Lower Bound for Gaussian Nominals

We now focus attention on constructing a lower bound for the important case of Gaussian nominals. This bound can be paired with the upper bound (33) to provide a relatively narrow envelope of uncertainty for the worst case error degradation in this special case. As in Section III-B, we would like to limit local dependence on the nominal f_N in order to obtain general results and to reduce required computation; however, we take note that a constructive lower bound must satisfy the power constraint (6). Nonetheless, we provide a specific construction which yields a simple and tight lower bound.

Given a Gaussian nominal pdf f_{N_σ} with variance σ^2 , we found in Section II-C that the solution for worst case error probability takes the form

$$\hat{P} = (\bar{\eta} + 1)\bar{P}$$

where $\bar{\eta}$ and \bar{P} solve (29) and where \bar{P} is the error probability associated with the Gaussian pdf $f_{N_{\bar{\sigma}}}$ once $\bar{\sigma}^2$ is chosen so that the second moment after subjecting $f_{N_{\bar{\sigma}}}$ to transformation (10) is σ^2 . We develop a lower bound P_L for \bar{P} by rewriting (29) as

$$(1-2\bar{P}) \left(1 - \bar{\eta} \frac{2\bar{P}}{1-2\bar{P}}\right) \log \left(1 - \bar{\eta} \frac{2\bar{P}}{1-2\bar{P}}\right) + 2\bar{P}(1+\bar{\eta}) \log(1+\bar{\eta}) = \delta + c\sigma^2 - \frac{1}{2} \log \frac{\sigma^2}{\bar{\sigma}^2}$$

and by making the key observation that for all $c \geq 0$

$$c\sigma^2 - \frac{1}{2} \log \frac{\sigma^2}{\bar{\sigma}^2} = c\sigma^2 - \frac{1}{2} \log(1+2c\sigma^2) \geq 0. \quad (37)$$

Hence, a valid constructive lower bound P_L results from choosing $\bar{\sigma}^2$ so that the second moment after subjecting $f_{N_{\bar{\sigma}}}$ to transformation (10) is no greater than σ^2 and where $\bar{\eta}$ and \bar{P} satisfy

$$(1-2\bar{P}) \left(1 - \bar{\eta} \frac{2\bar{P}}{1-2\bar{P}}\right) \log \left(1 - \bar{\eta} \frac{2\bar{P}}{1-2\bar{P}}\right) + 2\bar{P}(1+\bar{\eta}) \log(1+\bar{\eta}) = \delta \quad (38)$$

and then taking

$$P_L = (\bar{\eta} + 1)\bar{P}. \quad (39)$$

Note the strong match between (38) governing the lower bound P_L and (25) governing the upper bound \bar{P} , where the only difference is that P is the unadjusted nominal error probability while \bar{P} corresponds to a lower variance Gaussian chosen so that the solution meets the power constraint. This match enables us to easily analyze the discrepancy between these two bounding values.

The increase in second moment resulting from the application of (10) is given by

$$2\bar{\eta} \left(\bar{B} - \frac{(\bar{\sigma}^2 - 2\bar{B})\bar{P}}{1-2\bar{P}} \right) \leq 2\bar{\eta}\bar{B}$$

where

$$\bar{B} = \int_1^{\infty} \frac{x^2}{\sqrt{2\pi\bar{\sigma}}} \exp\{-x^2/2\bar{\sigma}^2\} dx$$

and hence we can satisfy the power constraint by taking $\bar{\sigma}^2 = \sigma^2 - 2\bar{\eta}\bar{B}$. If we denote by ΔP_L the new constructive lower bound on worst case error degradation, given by

$$\Delta P_L = (\bar{\eta} + 1)\bar{P} - P$$

then we have the asymptotic result

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{\Delta P_L}{\Delta P_U} &= \lim_{\delta \rightarrow 0} \frac{(\bar{\eta} + 1)\bar{P} - P}{\eta P} \\ &= 1 - \lim_{\delta \rightarrow 0} \frac{Q\left(\frac{1}{\sigma}\right) - Q\left(\frac{1}{\sqrt{\sigma^2 - 2\bar{\eta}\bar{B}}}\right)}{\eta P} \\ &= 1 - \frac{B \exp\{-1/2\sigma^2\}}{\sqrt{2\pi}P\sigma^3} \end{aligned}$$

where

$$B = \int_1^{\infty} \frac{x^2}{\sqrt{2\pi}\sigma} \exp\{-x^2/2\sigma^2\} dx.$$

This leads directly to the following lower bound.

Theorem 3.2: For Gaussian nominals

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} \geq \sqrt{P(1-2P)} \left(1 - \frac{B \exp\{-1/2\sigma^2\}}{\sqrt{2\pi}P\sigma^3}\right). \quad (40)$$

Fig. 9 depicts a graph of the bounds of Theorem 3.2 and result (33), providing an envelope of uncertainty for the asymptotic behavior of worst case error probability for Gaussian nominals over a range of SNR values. Note the increasing tightness of the bounds with increasing SNR, and in particular the narrow band of uncertainty in the typical range of 8 dB and above.

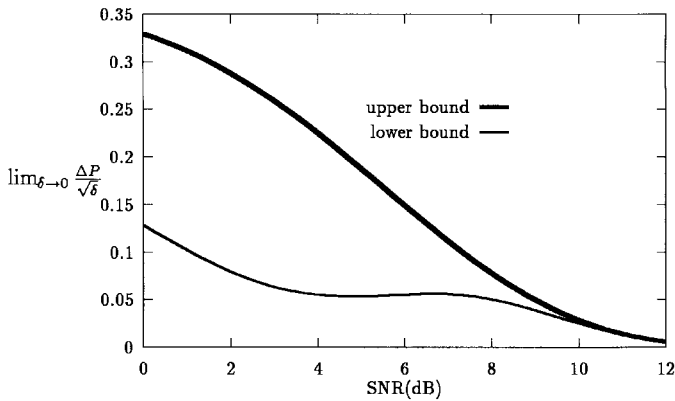


Fig. 9. Asymptotic worst case error probability for Gaussian nominals.

IV. CONCLUSION

We have seen that the worst case noise pdf $f_{\hat{N}}$ under constraints of power and divergence with respect to a prescribed nominal pdf f_N takes the form

$$f_{\hat{N}}(x) = \begin{cases} k_1 f_N(x) \exp\{-cx^2\}, & |x| < 1 \\ k_2 f_N(x) \exp\{-cx^2\}, & |x| \geq 1 \end{cases}$$

while dropping the power constraint leads to the worst case divergence-constrained noise pdf $f_{\tilde{N}}$ of the form

$$f_{\tilde{N}}(x) = \begin{cases} K_1 f_N(x), & |x| < 1 \\ K_2 f_N(x), & |x| \geq 1. \end{cases}$$

In Section II we demonstrated an efficient procedure for computing the parameters associated with $f_{\tilde{N}}$ and $f_{\hat{N}}$, and took advantage of the form of $f_{\tilde{N}}$ to reduce computation significantly in the case of Gaussian nominals.

Section III provided a study of the asymptotic behavior of the worst case error degradation as the divergence tolerance δ tends to zero. Without the power constraint, the worst case degradation ΔP_U was shown to satisfy the relation

$$\lim_{\delta \rightarrow 0} \frac{\Delta P_U}{\sqrt{\delta}} = \sqrt{P(1-2P)}$$

where P is the nominal probability of error. Upon imposing the power constraint, the upper bound on worst case degradation ΔP provided by ΔP_U was shown to be exact for nominal pdf's satisfying certain decay conditions, namely, (34) and (35). For Gaussian nominals we provided a lower bound for the quantity $\lim_{\delta \rightarrow 0} \Delta P/\sqrt{\delta}$ that performed very well for high values of SNR.

APPENDIX I

DEMONSTRATION OF CONSTRAINT ACTIVITY

We first demonstrate activity of the power constraint (6) and divergence constraint (7) when the worst case probability of error satisfies $0 < \hat{P} < 1/2$. We then study the extreme cases $\hat{P} = 0$ and $\hat{P} = 1/2$ in some detail, demonstrating that they are not typically of interest.

Assume for now that $0 < \hat{P} < 1/2$, and assume by contradiction that there exists a strict inequality in the divergence constraint. Pick any three ordered disjoint intervals I_1 , I_2 , and I_3 where $I_1 \in [0, 1)$, $I_2 \in [0, 1)$, $I_3 \in [1, \infty)$, and where each contains mass corresponding to the pdf $f_{\tilde{N}}$. It will always be possible (perhaps after interchanging I_1 and I_2) to displace mass from I_2 to both I_1 and I_3 in such a manner that the second moment remains unchanged and the divergence constraint remains satisfied, resulting in a larger error probability and thus rendering $f_{\tilde{N}}$ suboptimal. Similarly, if we

assume by contradiction that $f_{\hat{N}}$ exhibits a strict inequality in the power constraint, there must consequently be two ordered positive disjoint intervals J_1 and J_2 such that $f_{\hat{N}}(x) > f_N(x)$ throughout J_1 and $f_{\hat{N}}(x) < f_N(x)$ throughout J_2 . By displacing mass from J_1 to J_2 in such a manner that the power constraint remains satisfied, we can reduce the divergence and thus demonstrate nonoptimality with reference to the preceding argument. It should be noted that the last argument does not hold if the limit on noise power is general (and not necessarily set equal to the nominal power), making that problem more difficult to analyze.

Having demonstrated activity of both constraints for the case $0 < \hat{P} < 1/2$, we now consider the respective extreme cases. The case $\hat{P} = 0$ holds if and only if the nominal f_N exhibits zero error probability as a direct consequence of the divergence constraint (7), since $f_{\tilde{N}}$ must remain absolutely continuous with respect to f_N in order to achieve finite divergence, and is therefore trivially uninteresting. The case $\hat{P} = 1/2$ holds if and only if the worst case pdf is void of mass in the interval $[-1, 1]$, which in turn implies that the second moment of $f_{\tilde{N}}$ is greater than or equal to 1. According to the constraint (6), then, we can immediately assume activity of the constraints for any channel with SNR > 0 dB.

APPENDIX II

TIGHTNESS OF UPPER BOUND ON ASYMPTOTIC PERFORMANCE

Take any nominal pdf f_N for which there exists x^* such that $f_N(x) > 0$ for all $x^* > x$, and for which

$$\lim_{\bar{x} \rightarrow \infty} \bar{x}^2 \int_{\bar{x}}^{\infty} x^2 f_N(x) dx = \infty.$$

Fix $\bar{x} > 1$ and consider the random variable \tilde{N} with pdf

$$f_{\tilde{N}}(x) = \begin{cases} (1 - \eta_1) f_N(x), & |x| < 1 \\ (1 + \eta) f_N(x), & 1 \leq |x| \leq \bar{x} \\ 0, & |x| > \bar{x}. \end{cases} \quad (41)$$

The following quantities will be useful in our analysis

$$\begin{aligned} P_{\bar{x}} &= \int_1^{\bar{x}} f_N(x) dx & P^{\bar{x}} &= \int_{\bar{x}}^{\infty} f_N(x) dx \\ B_{\bar{x}} &= \int_1^{\bar{x}} x^2 f_N(x) dx & B^{\bar{x}} &= \int_{\bar{x}}^{\infty} x^2 f_N(x) dx. \end{aligned}$$

We also define $B = B_{\bar{x}} + B^{\bar{x}}$ and note that $P = P_{\bar{x}} + P^{\bar{x}}$. To make $f_{\tilde{N}}$ a proper pdf, we set

$$\eta_1 = \frac{2\eta P_{\bar{x}} - 2P^{\bar{x}}}{1 - 2P}.$$

In order to satisfy the power constraint (6), we use the degree of freedom provided by the choice of \bar{x} in (41), choosing \bar{x} as that for which

$$B^{\bar{x}} = \eta B \quad (42)$$

which is a well-defined specification of \bar{x} for any $0 < \eta < B^{\bar{x}}/B$ in light of (34).

With this choice of \bar{x} , we have that

$$\begin{aligned} \int_{-\infty}^{\infty} x^2 f_{\tilde{N}}(x) dx &= \sigma^2 - \eta_1(\sigma^2 - 2B) - 2B^{\bar{x}} + 2\eta B_{\bar{x}} \\ &\leq \sigma^2 - 2B^{\bar{x}} + 2\eta B \\ &= \sigma^2 \end{aligned}$$

where the last step follows directly from (42). Furthermore, we can write

$$P^{\bar{x}} = \int_{\bar{x}}^{\infty} f_N(x) dx$$

$$\begin{aligned} &\leq \bar{x}^{-2} \int_{\bar{x}}^{\infty} x^2 f_N(x) dx \\ &= \frac{(B^{\bar{x}})^2}{\bar{x}^2 B^{\bar{x}}} \\ &= \eta^2 \frac{B^2}{\bar{x}^2 B^{\bar{x}}} \end{aligned} \tag{43}$$

where the last step follows again from (42).

The divergence constraint (7) is satisfied as long as

$$\begin{aligned} &(1 - 2P)(1 - \eta_1) \log(1 - \eta_1) + 2P_{\bar{x}}(1 + \eta) \log(1 + \eta) \leq \delta \\ \Leftrightarrow &(1 - 2P) \left(-\eta_1 + \frac{\eta_1^2}{2} + \frac{\eta_1^3}{6} + o(\eta_1^4) \right) \\ &+ 2P_{\bar{x}} \left(\eta + \frac{\eta^2}{2} - \frac{\eta^3}{6} + o(\eta^4) \right) \leq \delta \\ \Leftrightarrow &-2\eta P_{\bar{x}} + 2P_{\bar{x}} + 2\eta P_{\bar{x}} + \frac{2(\eta P_{\bar{x}} - P_{\bar{x}})^2}{1 - 2P} \\ &+ \eta^2 P_{\bar{x}} + o(\eta^3) \leq \delta \\ \Leftrightarrow &2P_{\bar{x}} \left(1 - \frac{2\eta P_{\bar{x}}}{1 - 2P} \right) + 2 \frac{(P_{\bar{x}})^2}{1 - 2P} \\ &+ \eta^2 \frac{2(P_{\bar{x}})^2}{1 - 2P} + \eta^2 P_{\bar{x}} + o(\eta^3) \leq \delta \end{aligned}$$

which is satisfied by

$$\begin{aligned} &2\eta^2 \frac{B^2}{\bar{x}^2 B^{\bar{x}}} \left(1 - \frac{2\eta P_{\bar{x}}}{1 - 2P} \right) + 2\eta^4 \frac{B^4}{(1 - 2P)\bar{x}^4 (B^{\bar{x}})^2} \\ &+ \eta^2 \left(\frac{2(P_{\bar{x}})^2}{1 - 2P} + P_{\bar{x}} \right) + o(\eta^3) \leq \delta \end{aligned}$$

using (43), allowing us to constructively achieve

$$\lim_{\delta \rightarrow 0} \frac{\eta^2}{\delta} \left(\frac{2(P_{\bar{x}})^2}{1 - 2P} + P_{\bar{x}} + \frac{2B^2}{\bar{x}^2 B^{\bar{x}}} \right) = 1$$

implying that

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} \geq \sqrt{\frac{P^2(1 - 2P)}{P_{\bar{x}} + 2P_{\bar{x}}(P_{\bar{x}} - P)} + \frac{2B^2}{\bar{x}^2 B^{\bar{x}}}}$$

Since η tends to zero with δ , we have $\lim_{\delta \rightarrow 0} \bar{x} = +\infty$ by (34) and (42), so that

$$\lim_{\delta \rightarrow 0} P_{\bar{x}} = P$$

and

$$\lim_{\delta \rightarrow 0} 2B^2 / (\bar{x}^2 B^{\bar{x}}) = 0$$

by (35), allowing us to write

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} \geq \sqrt{P(1 - 2P)}.$$

Combining with (33) yields the limiting expression

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} = \sqrt{P(1 - 2P)}.$$

We now turn attention to the analysis of a relatively simple nominal probability mass function in order to demonstrate that (35) is nonsuperfluous for Theorem 3.1. Although a discrete nominal distribution is chosen for ease of analysis, the example is intuitively generalizable, for instance through approximation by a mixture of low-variance Gaussians.

Consider the problem (4) and take as nominal p_N the five-point probability mass function

$$p_N(x) = \begin{cases} 1 - 2P, & x = 0 \\ P - \varepsilon, & |x| = 1 \\ \varepsilon, & |x| = \bar{x} \end{cases}$$

where $\bar{x} > 1$, $0 < P < 1/2$, and $0 < \varepsilon < P$. We can immediately write

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} \leq \sqrt{P(1 - 2P)}$$

from (33). The exact behavior is also easy to determine in this case due to the limited number of degrees of freedom. Taking into account activity of the power constraint, the worst case probability mass function $p_{\bar{N}}$ must take the form

$$p_{\bar{N}}(x) = \begin{cases} 1 - 2P - 2\Delta P, & x = 0 \\ P - \varepsilon + \Delta P \frac{\bar{x}^2}{\bar{x}^2 - 1}, & |x| = 1 \\ \varepsilon - \Delta P \frac{1}{\bar{x}^2 - 1}, & |x| = \bar{x} \end{cases}$$

where ΔP must be chosen to satisfy the active divergence constraint, yielding

$$\lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} = \left(\frac{2}{1 - 2P} + \frac{\left(\frac{\bar{x}^2}{\bar{x}^2 - 1}\right)^2}{P - \varepsilon} + \frac{\frac{1}{(\bar{x}^2 - 1)^2}}{\varepsilon} \right)^{-\frac{1}{2}}.$$

By letting $\bar{x} \rightarrow \infty$ and $\varepsilon \rightarrow 0$, we obtain

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{\Delta P}{\sqrt{\delta}} &= \left(\frac{2}{1 - 2P} + \frac{1}{P} \right)^{-\frac{1}{2}} \\ &= \sqrt{P(1 - 2P)} \end{aligned}$$

as long as

$$\lim_{\bar{x} \rightarrow \infty, \varepsilon \rightarrow 0} 1 / (\varepsilon(\bar{x}^2 - 1)^2) = 0.$$

Since

$$\bar{x}^2 \int_{\bar{x}}^{\infty} x^2 f_N(x) dx = \varepsilon \bar{x}^4$$

this last condition matches exactly with (35), ensuring that it is not a superfluous requirement in Theorem 3.1.

REFERENCES

- [1] C. R. Cahn, "Worst interference for coherent binary channel," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 209–210, Mar. 1971.
- [2] S. Shamai (Shitz) and S. Verdú, "Worst case power-constrained noise for binary-input channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1494–1511, Sept. 1992.
- [3] M. A. Klimesh and W. E. Stark, "Worst case power-constrained noise for binary-input channels with varying amplitude signals," in *Proc. IEEE Int. Symp. on Information Theory* (Trondheim, Norway, 1994), p. 381.
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [5] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 858–871, May 1997.
- [6] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.