WORST-CASE POWER-CONSTRAINED NOISE FOR BINARY-INPUT CHANNELS

Shlomo Shamai (Shitz)
Dept. Electrical Eng.
Technion
Haifa 32000, ISRAEL

Sergio Verdú
Dept. Electrical Eng.
Princeton University
Princeton, NJ 08544, USA

The error probability and the capacity of binary-input additive-noise channels are well-known if the noise is Gaussian. A basic problem in communication theory is to find the worst-case performance achievable by any noise distribution as a function of the signal-to-noise ratio. This paper gives a complete solution to this problem for the two major performance measures: error probability and capacity. Those results are obtained as an application of a general framework developed in [1] which applies to many other performance functionals of information-theoretic interest besides error probability and capacity, such as divergence, cutoff rate, random-coding error exponent, and Chernoff entropy. Those general results show that the worst-case performance functional is given by the convex hull of the functional obtained by minimizing only over power-constrained noise distributions which place all their mass on a lattice whose span is equal to the distance between the two inputs. This implies that the least-favorable distribution is, in general, the mixture of two lattice probability mass functions. This conclusion can actually be generalized to m-ary input constellations on finite-dimensional spaces, as long as the input constellation puts its mass on a lattice. The proof of the results presented in this paper can be found in its journal version [1].

Consider the binary equiprobable hypothesis testing problem:

$$H_1: \quad \mathbf{Y} = +1 + \mathbf{N}$$
$$H_0: \quad \mathbf{Y} = -1 + \mathbf{N}$$

where $\mathbf{N}$ is a real-valued random variable constrained to satisfy an average-power limitation $E[\mathbf{N}^2] \le \sigma^2$. For $k = 1, 2, \cdots$ let $\sigma_k^2 \triangleq (k^2 - 1)/3$. The worst-case probability of error is

$$P_e(\sigma^2) = \frac{1}{2} - \frac{1}{2\sqrt{1 + 3\sigma_k^2}} + \frac{3}{2}\,\frac{\sigma^2 - \sigma_k^2}{k(k+1)(2k+1)}$$

for $\sigma_k^2 \le \sigma^2 \le \sigma_{k+1}^2$.

A single span-2 lattice achieves the maximum probability of error only when the allowed noise power is equal to σ_k^2, $k = 1, 2, \cdots$ Those worst-case distributions are symmetric and distribute their mass evenly on k atoms. (Those atoms are located at $0, \pm 2, \pm 4, \cdots$ if k is odd and at $\pm 1, \pm 3, \cdots$ if k is even.) When the allowed noise power lies strictly between $\sigma_k^2 < \sigma^2 < \sigma_{k+1}^2$, then a single span-2 lattice is no longer least favorable. Instead, the worst-case distribution is the unique span-1 lattice which is a mixture of the span-2 lattices that are least-favorable for σ_k^2 and σ_{k+1}^2 with respective weights $(\sigma^2 - \sigma_{k+1}^2)/(\sigma_k^2 - \sigma_{k+1}^2)$ and $(\sigma_k^2 - \sigma^2)/(\sigma_k^2 - \sigma_{k+1}^2)$. In particular, if $SNR > 0$ dB ($\sigma^2 < 1$), then the worst-case noise is symmetric with nonzero atoms at -1, 0, $+1$, i.e., the channel becomes a symmetric erasure channel. Thus, the noise distribution that maximizes error probability puts all its mass on the integers $\{-M, \cdots M\}$, where M depends on the signal-to-noise ratio and the weight assigned to each of those integers depends (in addition to the signal-to-noise ratio) only on whether the integer is even or odd. Note that for low signal-to-noise ratios, the worst-case noise cdf does not become asymptotically Gaussian, as might have been surmised from capacity considerations.

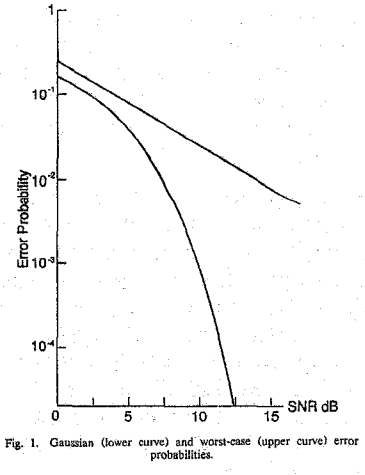

Fig. 1. Gaussian (lower curve) and worst-case (upper curve) error probabilities.

The worst-case capacity problem is

$$C(\sigma^2) = \min_{\substack{\mathbf{N} \\ E[\mathbf{N}^2] \le \sigma^2}} \max_{\mathbf{X}} I(\mathbf{X}; \mathbf{X} + \mathbf{N}) = \max_{\mathbf{X}} \min_{\substack{\mathbf{N} \\ E[\mathbf{N}^2] \le \sigma^2}} I(\mathbf{X}; \mathbf{X} + \mathbf{N})$$

where the maximum ranges over all distributions on $\{-1,1\}$ and the second equality follows from the concavity-convexity of mutual information in the respective arguments. It is verified in [1] that the least-favorable noise distribution puts its mass in the lattice $\{ \cdots -4, -2, 0, +2, +4 \cdots \}$ with a probability mass function that satisfies

$$\log(1 + \frac{p_{-1}}{p_0}) + \log(1 + \frac{p_1}{p_0}) - \log(1 + \frac{p_{k+1}}{p_k}) - \log(1 + \frac{p_{k-1}}{p_k}) + \lambda k^2 = 0$$

where p_k is the mass at $2k$. For low SNRs the least-favorable cdf approaches a Gaussian shape, whereas in the region $SNR > 0$ dB the least-favorable distribution is indistinguishable from a three-mass distribution with weights $(\sigma^2/8, 1 - \sigma^2/4, \sigma^2/8)$ at $(-2, 0, 2)$. The maximum difference between Gaussian capacity and worst-case capacity is 0.118 bit and occurs at 7.2 dB, whereas the maximum relative decrease is 12.5% occuring at 6.7 dB.

References

1. S. Shamai (Shitz) and S. Verdu, "Worst-Case Power Constrained Noise for Binary-Input Channels," *IEEE Transactions on Information Theory*, pp. 1494-1511, Sep. 1992.

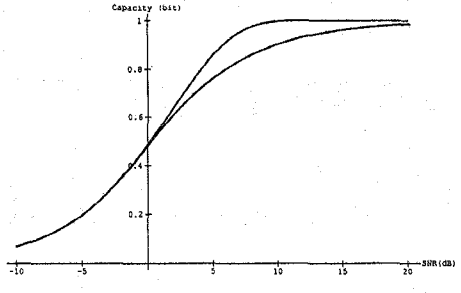

Fig. 2. Gaussian (upper curve) and worst-case (lower curve) channel capacities.

This work was supported in part by the Office of Naval Research under Grant N00014-90-J-1734.