Sensitivity of Quadratic Gaussian Matching to Interference

Lav R. Varshney, *Member, IEEE*, and Sanjoy K. Mitter, *Life Fellow, IEEE*

Abstract—It is well known that uncoded transmission of a memoryless Gaussian source over a memoryless additive white Gaussian noise channel results in optimal performance theoretically attainable. When there is additional interference in the channel, uncoded transmission is robust. It achieves the same sensitivity performance as optimal performance, measured using sensitivity results of Pinsker, Prelov, and Verdú.

Index Terms—Uncoded transmission, joint source-channel coding, interference, sensitivity.

I. INTRODUCTION

RESULTS on end-to-end information transmission systems have provided non-separation based, uncoded approaches to achieving optimal performance [1]. Optimal uncoded transmission requires that certain system components be matched in a certain way. As Shannon had described [2], "Solving this problem corresponds, in a sense, to finding a source that is just right for the channel and the desired cost. ... Solving this problem corresponds to finding a channel that is just right for the source and allowed distortion level."

As has long been recognized, the matching conditions are satisfied by memoryless Gaussian sources and memoryless additive white Gaussian noise (AWGN) channels [1], [3]–[5], [6, pp. 100–101], [7, Section II.E]. In this note, we investigate the degradation from optimal performance when the components are slightly mismatched. In particular, we focus on the scenario when there is weak additive interference, in addition to the modeled channel noise. We develop results for the case when the receiver has access to partial channel side information and when it does not. Comparisons to optimal performance show that to first-order approximation, the uncoded systems perform as well as optimal systems in the face of weak interference.

Unmodeled weak additive interference might arise in wireless networks, communication near ultrawideband systems, or cognitive radio [8], [9]. Robustness to such modeling errors is desirable in these communication systems. Prior work on compound channel capacity as well as on the capacity of arbitrarily varying channels (AVCs) similarly deals with additional interference. For AVCs, the interference is modeled as adversarial rather than as stochastic. Moreover, the coding schemes used for showing achievability are much
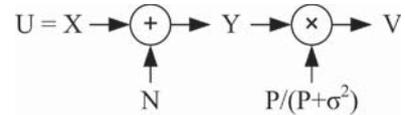
Fig. 1. Optimal quadratic Gaussian system.

more complicated than uncoded transmission, sometimes even making use of common randomness between the transmitter and receiver [10], [11]. Some work related to AVCs, but dealing directly with end-to-end transmission is [12], [13].

II. QUADRATIC GAUSSIAN MATCHING

The fact that a memoryless Gaussian source is matched to a memoryless AWGN channel with quadratic cost, under squared error distortion has long been known; here we make this explicit. Consider the system shown in Fig. 1. The source/channel input is distributed as $U = X \sim \mathcal{N}(0, P)$; the independent additive noise is distributed as $N \sim \mathcal{N}(0, \sigma^2)$; the channel input cost function is quadratic; and distortion is the squared difference between U and V. Computing the performance of this scheme, we find that the expected input power and expected end-to-end distortion pair (B, Δ) is given by

$$B = P, \quad \Delta = \frac{P\sigma^2}{P + \sigma^2}.$$

Optimality of this single-letter coding scheme can be seen by noting that the scheme simultaneously achieves the rate-distortion and the capacity-cost bounds. The rate-distortion function for a Gaussian source with squared error distortion is well known to be

$$R(\Delta) = \begin{cases} \frac{1}{2} \log_2 \left(\frac{P}{\Delta} \right), & \Delta < P \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

and the capacity-cost function for an AWGN channel with quadratic cost is well known to be

$$C(B) = \frac{1}{2} \log_2 \left(1 + \frac{B}{\sigma^2} \right). \quad (2)$$

Equating these two expressions yields the fact that the smallest achievable distortion at cost $\Gamma = P$ is

$$\Delta_{\min} = \frac{P\sigma^2}{P + \sigma^2},$$

and since the single-letter scheme achieves this, it is optimal. Next, we will look at what happens when there is mismatch induced by additive interference.
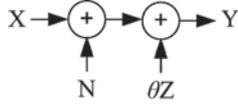
Fig. 2. Quadratic AWGN channel with additive interference.

III. SENSITIVITY OF CHANNEL CAPACITY

Before proceeding with the analysis, we review some results of Pinsker, Prelov, and Verdú on the sensitivity of channel capacity to weak additive interference [14].

Consider the channel shown in Fig. 2, with average power constraint B on the input X; nominal Gaussian noise $N \sim \mathcal{N}(0, \sigma^2)$; and contaminating noise normalized so that $E[Z] = 0$ and $E[Z^2] = 1$. The scalar θ determines the contaminating noise power, $\theta^2 = \zeta$. We denote the capacity of this channel as $C_B(\theta)$, and for any $\theta > 0$, this capacity is usually not known in closed form; for $\theta = 0$, the channel reduces to the usual AWGN channel which has known solution. Since we cannot find $C_B(\theta)$, we would like to find an approximation to it for small θ, based on the known value for $C_B(0)$. One can think of this as the technique from perturbation theory which uses the result of a known solvable problem to find an approximation to the solution of a difficult problem. Here, we aim to find the linearization of $C_B(\theta)$ around $\theta = 0$.

We define the sensitivity of channel capacity with respect to contaminating noise power as

$$S_B = \lim_{\theta \to 0} -\frac{C_B(\theta) - C_B(0)}{\theta^2}.$$

By the limit definition of the derivative, this is also

$$S_B = -\left. \frac{\partial C_B}{\partial \zeta} \right|_{\zeta=0}.$$

Then the linearization of the channel capacity is

$$C_B(\theta) \approx C_B(0) - S_B \theta^2.$$

When the contaminating noise is Gaussian, we can find the sensitivity directly from the definition:

$$S_B = \frac{B}{2\sigma^2 (B + \sigma^2) \log_e 2}.$$

This is clearly an upper bound on arbitrary white noise, since Gaussian noise results in worst case capacity degradation. A surprising result in [14], however is that the sensitivity is actually equal to the upper bound for any contamination noise Z that is drawn i.i.d. Thus we can approximate the channel capacity for a wide class of interference noise distributions as

$$C_B(\theta) \approx \frac{\log_2 \left(1 + \frac{B}{\sigma^2}\right)}{2} - \frac{B\theta^2}{2\sigma^2 (B + \sigma^2) \log_e 2}. \quad (3)$$

IV. SENSITIVITY OF DISTORTION TO INTERFERENCE

The previous section established the sensitivity of channel capacity to interference. Now we want to find the sensitivity of end-to-end distortion to interference. For a $\mathcal{N}(0, B)$ source, the distortion-rate function, the inverse of the rate-distortion function, is

$$\Delta(R) = B2^{-2R}. \quad (4)$$

Using a separation-style method, we want to use the distortion-rate function operating at capacity to find the sensitivity of distortion to the contamination noise power. Using the chain rule of differentiation, and assuming that (3) is exact, we find that

$$-\left. \frac{\partial \Delta}{\partial \zeta} \right|_{\zeta=0} = -\left. \left(\frac{\partial \Delta}{\partial C_B} \frac{\partial C_B}{\partial \zeta} \right) \right|_{\zeta=0} \quad (5)$$

$$= -\left. \left(\frac{\partial \Delta}{\partial C_B} \right) \right|_{C_B=C_B(0)} \left. \left(\frac{\partial C_B}{\partial \zeta} \right) \right|_{\zeta=0}$$

$$= \left. \left(\frac{\partial}{\partial C_B} B2^{-2C_B} \right) \right|_{C_B=C_B(0)} S_B$$

$$= \left(-\frac{2B \log_e 2}{1 + \frac{B}{\sigma^2}} \right) \left(\frac{B}{2\sigma^2 (B + \sigma^2) \log_e 2} \right)$$

$$= -\frac{B^2}{(B + \sigma^2)^2}.$$

The Gaussian rate-distortion function is differentiable, so this computation makes sense. We have found the sensitivity of distortion to changes in contamination noise power. By the limit definition of the derivative, this is also

$$S_{\Delta,B}^{\text{optimal}} \triangleq \lim_{\theta \to 0} \frac{\Delta_B(0) - \Delta_B(\theta)}{\theta^2} = -\left. \frac{\partial \Delta}{\partial \zeta} \right|_{\zeta=0} = -\frac{B^2}{(B + \sigma^2)^2},$$

where we have defined the new quantity, $S_{\Delta,B}$, the distortion sensitivity to contamination noise power for a fixed cost B. Thus the end-to-end distortion, for fixed cost B, as a function of contamination noise power may be approximated by

$$\Delta_B(\theta) \approx \Delta_B(0) - \theta^2 S_{\Delta,B}$$

$$\approx \frac{B\sigma^2}{B + \sigma^2} + \frac{B^2 \theta^2}{(B + \sigma^2)^2}.$$

This approximation involves two stages of linearization, first the capacity is linearized using the results of [14], and then the distortion is linearized using the chain rule of differentiation. Alternatively, we can get a better approximation of the distortion function if we use only one stage of linearization. This would involve substituting (3) into (4) to get

$$\Delta(C_B(\theta)) = \Delta_B(\theta) \approx \frac{B\sigma^2}{B + \sigma^2} 2^{S_B \theta^2}.$$

V. PERFORMANCE OF SINGLE LETTER CODES IN THE PRESENCE OF INTERFERENCE

Now that we know an approximation of the best that we can do, let us determine how well our original single letter scheme does. The new system is created by substituting the channel in Fig. 2 into the system in Fig. 1. Since everything is mean zero, the second moments and the variances are equal. Working through the calculation, making use of various independence relationships, we get that

$$\Delta_B(\theta) = \frac{B\sigma^4 + B^2 \theta^2 + B^2 \sigma^2}{(B + \sigma^2)^2}.$$

Taking the derivative of this with respect to θ^2 and evaluating the negative at $\theta = 0$ yields

$$S_{\Delta,B}^{\text{uncoded}} = -\frac{B^2}{(B + \sigma^2)^2}. \quad (6)$$

Now consider a decoder that has knowledge of the interference statistics and takes that into account. The decoder is now a scalar multiplication by $B/(B + \sigma^2 + \theta^2)$. Computing the distortion, we get

$$\Delta_B(\theta) = \frac{B\left(\sigma^2 + \theta^2\right)}{B + \sigma^2 + \theta^2}.$$

The associated sensitivity under partial information (PI) is

$$S_{\Delta,B}^{\mathrm{uncoded,PI}} = -\frac{B^2}{(B + \sigma^2)^2}. \tag{7}$$

Evidently, partial knowledge of the interference noise statistics does not reduce the sensitivity to the interference noise power.

Proposition 1: To a first-order approximation, the end-to-end distortion performance in the single-letter mismatched case, in the single-letter partially-matched case, and in the optimal case are the same.

Proof: Compare (6), (7), and (5). ∎

VI. DIRTY PAPER CODING AND AVCS

Since we are looking at channels with interference, for completeness, we should also comment on dirty paper [15] (or more precisely generalized writing on dirty paper [16]) results. In the previous section where we considered a partially informed single-letter decoder, the decoder had access to the variance of the contamination noise (dirt) distribution. In the dirty paper scenario, the encoder has access to the actual realization of the dirt, not just a statistical parameter. It has been shown that for any i.i.d. interference noise, the capacity is not reduced, so the distortion is not increased, and so the sensitivity of distortion to interference, $S_{\Delta,B}^{\mathrm{dirtypaper}}$, is zero.

One may also note that the basic result on Gaussian AVCs shows that when the jammer power constraint is known and the system has access to common randomness, the capacity is simply the AWGN capacity with interference power treated as additional Gaussian noise. This also holds true using deterministic codes when the jammer power constraint is smaller than the channel input power constraint [17]. Thus, sensitivity of distortion to interference is also the same as in Proposition 1.

VII. CONCLUSIONS

First, we have shown that distortion as a function of interference noise power is continuous for arbitrary zero-mean, memoryless interference. This follows from the continuity of the channel capacity as a function of interference noise power and the continuity of the distortion-rate function for a Gaussian source with respect to squared error. One can regard the optimal cost-distortion point as a saddlepoint, with perturbations causing smooth changes.

Second, we have found an approximation of the optimal sensitivity of distortion to interference noise power through a first-order approximation of the channel capacity as a function of interference noise power.

Third, we have found the distortion as a function of interference noise power for two single-letter coding schemes, one that uses the decoder for no interference, and the other that uses the decoder that takes the interference noise power into account. These distortion-interference noise power functions give associated sensitivity values.

We find that the sensitivity in the optimal case, the partially informed decoder case, and the uninformed decoder case are all equal. Thus to a first-order approximation for weak additive interference, the distortion-interference noise power functions are all equal. That is to say, in an approximate sense, single-letter codes perform as well as the optimality bound.

Additive interference reduces capacity. In a separation based approach with simple random codes, when the channel capacity reduces slightly below the rate, the probability of error goes to one [18]. To continue operating at optimal levels, one would require the use of new source and channel codes. In the single letter coding scheme, the coding scheme remains fixed, but the performance degradation is commensurate with the optimal performance degradation, so there is a great deal of robustness to uncertain weak channel interference. One may also study similar questions for discrete alphabet systems [19, Sec. 2.7].

REFERENCES

[1] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[2] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE Nat. Conv. Rec., Part 4*, Mar. 1959, pp. 142–163.

[3] T. J. Goblick, Jr., "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 558–567, Oct. 1965.

[4] T. Berger, "Living information theory," in *2002 IEEE Int. Symp. Inf. Theory*, July 2002 (Shannon Lecture).

[5] R. J. McEliece, *The Theory of Information and Coding: A Mathematical Framework for Communication*, ser. Encyclopedia of Mathematics and Its Applications. Addison-Wesley, 1977, vol. 3.

[6] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, 1971.

[7] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1998.

[8] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1745–1755, Apr. 2006.

[9] R. Menon, R. M. Buehrer, and J. H. Reed, "On the impact of dynamic spectrum sharing techniques on legacy radio systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4198–4207, Nov. 2008.

[10] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 2, pp. 267–284, Mar. 1987.

[11] A. D. Sarwate, "Robust and adaptive communication under uncertain interference," Ph.D. dissertation, University of California, Berkeley, 2008.

[12] T. Basar, "The Gaussian test channel with an intelligent jammer," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 1, pp. 152–157, Jan. 1983.

[13] T. Basar and Y.-W. Yu, "Solutions to a class of minimax decision problems arising in communication systems," *J. Optim. Theory Appl.*, vol. 51, no. 3, pp. 375–404, Dec. 1986.

[14] M. S. Pinsker, V. V. Prelov, and S. Verdú, "Sensitivity of channel capacity," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1877–1888, Nov. 1995.

[15] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[16] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, June 2002.

[17] I. Csiszár and P. Narayan, "Capacity of the Gaussian arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 18–26, Jan. 1991.

[18] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[19] L. R. Varshney, "Optimal information storage: nonsequential sources and neural channels," S.M. thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2006.