Capacity Achieving Probability Measure of an Input-bounded Vector Gaussian Channel

Terence H. Chan[1], Steve Hranilovic, Frank R. Kschischang
Dept. of Electrical & Computer Engineering
University of Toronto
{terence, steve, frank}@comm.utoronto.ca

Abstract — A discrete-time memoryless additive vector Gaussian noise channel subject to average cost constraints and an input-bounded constraints is considered. The necessary and sufficient condition for an input distribution to be capacity achieving is derived, and the capacity achieving distribution is shown to be "discrete" in nature.

I. Introduction

In [1], Smith showed that the capacity achieving distribution of a discrete-time memoryless scalar Gaussian channel subject to both an average power constraint and a peak power constraint is a probability mass function. Using a similar approach as [1], Shamai *et al.* showed that the capacity achieving distributions of quadrature Gaussian channels are also discrete in nature [2]. In this paper, we generalize the results of Smith and Shamai by considering a class of more general channels, called vector Gaussian noise channels. Also, the input-bounded constraint is not restricted to be the peak power constraint as in previous work, and the average cost constraints we interested are the second-order constraints, of which the conventional average power constraint is a special case.

First, consider an additive vector Gaussian noise channel whose input $\mathbf{X} = [X_1, X_2, \cdots, X_{\mathrm{N}}]^\top$ and output $\mathbf{Y} = [Y_1, Y_2, \cdots, Y_{\mathrm{M}}]^\top$ is related by the equation $\mathbf{Y} = \mathbf{AX} + \mathbf{Z}$ where $\mathbf{A}$ is a fixed $\mathrm{M} \times \mathrm{N}$ matrix and $\mathbf{Z} = [Z_1, Z_2, \cdots, Z_{\mathrm{M}}]^\top$ is a zero-mean Gaussian noise vector. The communication systems are subject to two classes of constraints. The first class is second-order average cost constraints of the form $G_k(\mu) = \mathbf{E}_\mu\left[\mathbf{X}^\top \mathbf{Q}_k \mathbf{X} + \mathbf{L}_k \mathbf{X}\right] - \gamma_k \leq 0$ for $k = 1 \cdots \mathrm{K}$, where $\mathbf{Q}_k$ and $\mathbf{L}_k$ are $\mathrm{N}\times\mathrm{N}$ and $1\times\mathrm{N}$ matrices, and the expectations are taken with respect to the input distribution μ. The second class is the input-bounded constraint, which restricts the set of admissible channel inputs to be a closed and bounded subset $\mathcal{S}$ in $\mathbb{R}^{\mathrm{N}}$.

Example 1 *Consider a wireless optical channel in which the dominant noise source is the ambient light. Assuming that the symbol waveform is a linear combination of a basis of functions* $\{\phi_n(t) : n = 1, \cdots, \mathrm{N}\}$*, the channel is equivalent to a discrete time vector Gaussian noise channel* $[Y_1, \cdots, Y_{\mathrm{N}}]^\top = [X_1, \cdots, X_{\mathrm{N}}]^\top + [Z_1, Z_2, \cdots, Z_{\mathrm{N}}]^\top$*, such that the average power constraint is given by* $\mathbf{E}_\mu\left[\sum_{n=1}^{\mathrm{N}} x_n(\int_0^T \phi_n(t)\,dt)\right] - \gamma \leq 0$*. Moreover, the set of admissible channel inputs is the following closed and bounded subset [5]*

$$\mathcal{S} = \left\{(x_1, \cdots, x_{\mathrm{N}}) \in \mathbb{R}^{\mathrm{N}} : \forall 0 \leq t < T,\ 0 \leq \sum_{n=1}^{\mathrm{N}} x_n \phi_n(t) \leq s\right\}.$$

[1] This work is supported in part by The Croucher Foundation.

II. Results

Let $\mathbf{H}(\mu)$ and $\mathbf{I}(\mu)$ be the entropy of $\mathbf{Y}$ and the mutual information between $\mathbf{X}$ and $\mathbf{Y}$ given that the input distribution of $\mathbf{X}$ is μ respectively.

Theorem 1 (Existence) *There exists a probability distribution* μ_o *satisfying the average cost constraints and the input-bounded constraint, which maximizes* $\mathbf{I}(\mu)$*. Furthermore, if* $\mathbf{A}$ *has a left-inverse, then the capacity achieving distribution is unique.*

Let $h(\mathbf{x}; \mu_o) = -\int \mathrm{P}_{\mathbf{Z}}(\mathbf{y} - \mathbf{Ax}) \log \mathrm{P}_{\mathbf{Y}}(\mathbf{y}; \mu_o)\, d\mathbf{y}$ where $\mathrm{P}_{\mathbf{Z}}$ and $\mathrm{P}_{\mathbf{Y}}$ are the densities of $\mathbf{Z}$ and $\mathbf{Y}$ respectively. For any distribution μ_o, its set of points of increase E_o is defined by

$$E_o = \left\{\mathbf{x} \in \mathbb{R}^{\mathrm{N}} : \begin{array}{c} \mu_o(V) > 0 \text{ for any open} \\ \text{set } V \text{ containing } \mathbf{x} \end{array}\right\}.$$

Theorem 2 (Necessary and Sufficient Condition) *Let* μ_o *satisfy the average cost constraints and the input-bounded constraint, and* E_o *be its set of points of increase. Then* μ_o *is capacity achieving if and only if there exists* $\lambda_1 \cdots, \lambda_{\mathrm{K}} \geq 0$ *such that*

$$\begin{cases} h(\mathbf{x}; \mu_o) \leq \mathbf{H}(\mu_o) + \sum_{k=1}^{\mathrm{K}} \lambda_k (\mathbf{x}^\top \mathbf{Q}_k \mathbf{x} + \mathbf{L}_k \mathbf{x} - \gamma_k)\ \forall \mathbf{x} \in \mathcal{S} \\ h(\mathbf{x}; \mu_o) = \mathbf{H}(\mu_o) + \sum_{k=1}^{\mathrm{K}} \lambda_k (\mathbf{x}^\top \mathbf{Q}_k \mathbf{x} + \mathbf{L}_k \mathbf{x} - \gamma_k)\ \forall \mathbf{x} \in E_o. \end{cases}$$

A complex-valued function $f(\mathbf{w})$ defined on $\mathbb{C}^{\mathrm{N}}$ is called holomorphic if it is analytic in each individual variable. A subset F in $\mathbb{R}^{\mathrm{N}}$ is called *sparse* if there exists a nonzero holomorphic function $f(\mathbf{w})$, such that $f(\mathbf{w}) = 0$ for all $\mathbf{w} \in F$. For example, F is sparse if it is finite, whereas it is <u>not</u> sparse if it contains an open subset in $\mathbb{R}^{\mathrm{N}}$. In particular, if $\mathrm{N} = 1$ and F is bounded, then F is sparse if and only if it is finite.

Theorem 3 (Discreteness) *The capacity achieving distribution is "discrete" in the sense that its set of points of increase is sparse.*

References

[1] J.G. Smith, "The information capacity of amplitude and variance-constrained scalar Gaussian channels," *Inform. Contr.*, vol 18., pp. 203-219, 1971.

[2] S. Shamai(Shitz) and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inform. Theory*, vol 41, pp. 1060 -1071 July 1995.

[3] I.C. Abou-Faycal, M.D. Trott, S. Shamai(Shitz), "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290 -1301, May 2001.

[4] M. Katz, S. Shamai(Shitz), "On the capacity-achieving distribution of the discrete-time non-coherent additive white noise channel," *IEEE Int. Symp. Inform. Theory* 2002, pp. 165.

[5] S. Hranilovic, F.R. Kschischang, "Optical Intensity-Modulated Direct Detection Channels: Signal Space and Lattice Codes," to appear in *IEEE Trans. Inform. Theory*.