A Nash Equilibrium related to the Poisson Channel

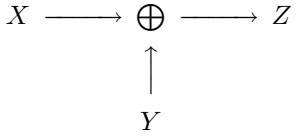
P. Harremoës and C. Vignat

Abstract

An information theoretic game associated to the Poisson distribution is studied and it is demonstrated that the game has a Nash equilibrium.

I. Introduction

We consider the following problem: an emitter transmits a message X on a channel; a jammer sends a message Y on the same additive channel,



so that the received message writes

$$Z = X + Y.$$

The objective of the emitter is to design his message so as to maximize the transmission rate i.e. the mutual information $I(X;Z)$. The objective of the jammer is to choose the distribution of Y independent of X so as to minimize the transmission rate.

The case when X and Y are continuous random variables with power constraints of the form $E\left(X^2\right) \leq P$ and $E\left(Y^2\right) \leq N$ has been studied by T. Cover [1, Ex. 1, page 263]. In this case the normal distributions with mean 0 and variances P and N form a Nash equilibrium. The entropy power inequality plays an essential role in the proof of the Nash equilibrium condition. There has been several attempts to provide discrete versions of the Entropy Power Inequality, and for Bernoulli sources with addition modulo 2 results have been obtained in a series of papers [2], [3], [4] and [5] which proved the existence of a Nash equilibrium. In this paper we shall consider a similar setup with discrete random variables and *classical* addition, and this should be seen as a step in the direction of a discrete Entropy Power Inequality. This paper follows ideas related to discrete Fisher information [6] [7].

P. Harremoës is with Dept. of Math., University of Copenhagen, Denmark, moes@math.ku.dk

C. Vignat is with Laboratoire Systèmes de Communications, Université de Marne la Vallée, France, vignat@univ-mlv.fr . This work was done during a visit at Dept. of Math., University of Copenhagen in March 2003.

The first author is supported by a post-doc fellowship from the Villum Kann Rasmussen Foundation and by INTAS (project 00-738) and by the Danish Natural Science Research Council.

The Poisson distribution with mean value λ will be denoted $Po(\lambda)$. The binomial distribution with number parameter n and success probability p will be denoted $Bi(n,p)$. When no confusion seems to occur $Po(\lambda)$ and $Bi(n,p)$ will be used to denote a random variable with this distribution. In what follows, we denote as $B_n(\lambda)$ the set of sums of n independent Bernoulli random variables, each having parameter $p_{i,1\leq i\leq n}$ with $\sum_{i=1}^{n} p_i = \lambda$. Let us define the set of *Bernoulli sums* as $B_*(\lambda) = \cup_{n=1}^{\infty} B_n(\lambda)$. Further $B_\infty(\lambda)$ shall denote the set of infinite sums $S = \sum_{i=1}^{\infty} X_i$ where X_i denotes independent Bernoulli random variables such that $\sum_{i=1}^{\infty} p_i = \lambda$. Finally the set of *generalized Bernoulli sums* is the closure $\overline{B_*(\lambda)} = \overline{B_\infty(\lambda)}$ and the elements in this set have the form

$$X = S + Po(\mu)$$

where $0 \leq \mu \leq \lambda$ and $S \in B_\infty(\lambda - \mu)$.

When X and Y are independent discrete random variables we have the formula

$$I(X;Z) = H(Z) - H(Z|X).$$

In order to address a realistic situation, both messages are subject to constraints

$$EX \leq \lambda_{\text{in}}, \;\; EY \leq \lambda_{\text{noise}}$$

where λ_{in} and λ_{noise} are positive constants. In this paper we shall consider the case when the messages X and Y are supposed to belong to the set $\overline{B_\infty(\lambda)}$ of generalized Bernoulli sums: $X \in \overline{B_\infty(\lambda_{in})}$ and $Y \in \overline{B_\infty(\lambda_{noise})}$. We see that we have a 2-persons zero sum game with the emitter and the jammer as players. This game will be called *the discrete transmission game*. The sets of strategies are not convex so Von Neumann's classical result on existence of a game theoretic equilibrium cannot be used. The main result in this paper is that the Poisson distributions forms a Nash equilibrium of the discrete transmission game. Thus the *Poisson channel*, i.e. an information channel with Poisson distributed noise, is in a natural way related to the discrete transmission game.

II. Main Result

We need the following extension of Hoeffding's theorem [8].

Lemma 1: If $S \in \overline{B_\infty(\lambda)}$ and f is a concave function such that $E|f(Po(\lambda))| < \infty$, then

$$E[f(S)]$$

is minimized for $S \sim Po(\lambda)$.

Proof: Since $S \in \overline{B_\infty}(\lambda)$, then, choosing $n \in \mathbb{N}$, S writes as

$$S = S_1^{(n)} + S_2$$

where $S_1^{(n)}$ is a sum of n independent Bernoulli random variables, each with success probability $p_i^{(n)}$. Now write

$$\begin{aligned} E[f(S)] &= E_{S_1^{(n)}} E_{S_2}\left[f\left(S_1^{(n)} + S_2\right)\right] \\ &= \sum_i \Pr\{S_2 = i\} E_{S_1^{(n)}}\left[f\left(S_1^{(n)} + i\right)\right] \\ &\geq \sum_i \Pr\{S_2 = i\} E\left[f\left(Bi\left(n, \frac{E\left(S_1^{(n)}\right)}{n}\right) + i\right)\right] \\ &= E\left[f\left(Bi\left(n, \frac{E(S_1)}{n}\right) + S_2\right)\right] \end{aligned}$$

where the inequality is a direct application of Hoeffding's theorem. This shows that $\inf_{S \in \overline{B_\infty}(\lambda)} E[f(S)]$ can be searched in the restricted set of $S \in \overline{B_\infty(\lambda)}$ such that $\max_{1 \leq i \leq n}\left\{p_i^{(n)}\right\} \leq \varepsilon$. Now consider a sequence $\{S_n\}$ such as

$$\lim_{n \to \infty} \max_{1 \leq i \leq n}\left\{p_i^{(n)}\right\} \to 0.$$

Then $S_n \to Po(\lambda)$. But since [9]

$$\Pr\{S_n = i\} \leq e^\lambda \Pr\{Po(\lambda) = i\},$$

the absolute expectation $E|f(S_n)|$ can be bounded as follows:

$$\begin{aligned} E|f(S_n)| &= \sum_i |f(i) \Pr\{S_n = i\}| \\ &\leq \sum_i |f(i)| e^\lambda \Pr\{Po(\lambda) = i\} \\ &= e^\lambda E|f(Po(\lambda))| < \infty, \end{aligned}$$

so that, using the dominated convergence criterion, we conclude that

$$\inf_{S \in \overline{B_\infty}(\lambda)} E[f(S)] = E[f(Po(\lambda))].$$

■

The main result of this paper is the following.

Theorem 2: The 2-persons zero sum game has the Poisson distributions as a Nash equilibrium. In other terms, if $Y \sim Po(\lambda_{\text{noise}})$ then

$$Po(\lambda_{\text{in}}) \sim \underset{X \in \overline{B_\infty}(\lambda),\, EX \leq \lambda_{\text{in}}}{\arg\max} I(X;Z) \ ,$$

and if $X \sim Po(\lambda_{\text{in}})$ then

$$Po(\lambda_{\text{noise}}) \sim \underset{Y \in \overline{B_\infty}(\lambda),\, EY \leq \lambda_{\text{noise}}}{\arg\min} I(X;Z) \ .$$

Proof: First rewrite the mutual information as

$$\begin{aligned} I(X;Z) &= H(Z) - H(Z|X) \\ &= H(X+Y) - H(X+Y|X) \, , \end{aligned}$$

and remark that the conditional entropy can be simplified as

$$H(X+Y|X) = H(Y) \ ,$$

so that the mutual information writes

$$I(X;Z) = H(X+Y) - H(Y) \, .$$

Suppose first that $Y \sim Po(\lambda_{\text{noise}})$: we look for the distribution of X that solves the following problem:

$$X = \arg \max_{X \in \overline{B_\infty}(\lambda), EX \leq K_X} H(X + Po(\lambda_{\text{noise}})) \, .$$

In [10, Thm. 8], it is shown that the solution of this problem is

$$X \sim Po(\lambda_{\text{in}}) \, .$$

Suppose reciprocally that $X \sim Po(\lambda_{\text{in}})$. We thus have to solve the following optimization problem

$$Y = \arg \min_{Y \in \overline{B_\infty}(\lambda), EY \leq \lambda_{\text{noise}}} I(X,Z) \ .$$

Assume that $EY < \lambda_{noise}$. Let Z denote $X + Y$ and let V be a Bernoulli sum with mean $\lambda_{noise} - EY$ and independent of X and Y. Put $W = V + Z$. Then

$$\begin{aligned} I(X;W) &\leq I(X;(Z,W)) \\ &= I(X;Z) + I(X;W \mid Z) \\ &= I(X;Z). \end{aligned}$$

Therefore $Y + V$ gives a smaller transmission rate than Y. It is seen that only the case $EY = \lambda_{\text{noise}}$ has to be taken into consideration.

For $\lambda \geq 0$ we denote by $\left\{P_k^\lambda\right\}$ the point probabilities of $Y + Po(\lambda)$, and remark that this distribution satisfies the following differential equation

$$\frac{\partial P_k^\lambda}{\partial \lambda} = -\left(P_k^\lambda - P_{k-1}^\lambda\right),$$

so that its entropy satisfies

$$\frac{\partial}{\partial \lambda} H\left(P_k^\lambda\right) = -\frac{\partial}{\partial \lambda} \sum_{k=0}^{\infty} P_k^\lambda \log P_k^\lambda = \sum_{k=0}^{\infty} P_k^\lambda \log \frac{P_k^\lambda}{P_{k+1}^\lambda}.$$

For $k = 0, 1, 2, ...$ put

$$Q_k^\lambda = \frac{(k+1)}{\lambda + EY} P_{k+1}^\lambda,$$

and remark that Q_k^λ is a probability distribution. Now we write

$$\log \frac{P_k^\lambda}{P_{k+1}^\lambda} = \log \frac{P_k^\lambda}{Q_k^\lambda} + \log \frac{k+1}{\lambda + EY},$$

so that

$$\begin{aligned} \frac{\partial}{\partial \lambda} H\left(P_k^\lambda\right) &= \sum_{k=0}^{\infty} P_k^\lambda \log \frac{P_k^\lambda}{Q_k^\lambda} + \sum_{k=0}^{\infty} P_k^\lambda \log \frac{k+1}{\lambda + EY} \\ &= D\left(P_k^\lambda || Q_k^\lambda\right) + \sum_{k=0}^{\infty} P_k^\lambda \log \frac{k+1}{\lambda + EY}, \end{aligned}$$

where $D\left(P_k^\lambda||Q_k^\lambda\right)$ denotes the (Kullback-Leibler) information divergence from P_k^λ to Q_k^λ. Thus, the mutual information to be minimized writes

$$\begin{aligned} H\left(Po\left(\lambda_{\text{in}}\right)+Y\right)-H\left(Y\right) &= \int_0^{\lambda_{\text{in}}} \frac{\partial}{\partial\lambda} H\left(P_k^\lambda\right) d\lambda \\ &= \int_0^{\lambda_{\text{in}}} \left(D\left(P_k^\lambda||Q_k^\lambda\right) + \sum_{k=0}^{\infty} P_k^\lambda \log\frac{k+1}{\lambda+EY} \right) d\lambda \,. \end{aligned}$$

This integral is clearly minimized if the integrand is minimized simultaneously for all λ. Moreover, the information divergence $D\left(P_k^\lambda||Q_k^\lambda\right)$ is positive, and equals zero when $Y \sim Po\left(EY\right)$. The other term can be written as

$$\sum_{k=0}^{\infty} P_k^\lambda \log\frac{k+1}{\lambda+EY} = E_{P^\lambda}\left[\log\left(Y+1\right)\right] - \log\left(\lambda+EY\right) \,.$$

Now Lemma 1 can be applied and it is seen that the term $E_{P^\lambda}\left[\log\left(Y+1\right)\right]$ is minimized when Y is Poisson distributed with mean λ_{noise}. ■

III. Acknowledgements

The authors want to thank Rudolf Ahlswede for useful discussions and drawing our attention to earlier work on the continuous and the discrete Entropy Power Inequalities.

References

[1] T. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[2] A. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications: Part I," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 769–772, Nov. 1973.

[3] A. Wyner, "A theorem on the entropy of certain binary sequences and applications: Part II," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 772–777, Nov. 1973.

[4] H. S. Witsenhausen, "Entropy inequalities for discrete channels," *IEEE Trans. Inform Theory*, vol. IT-20, pp. 610–616, Sept. 1974.

[5] S. Shamai and A. Wyner, "A binary analog to the entropy-power inequality," *IEEE Trans. Inform. Theory*, vol. IT-36, pp. 1428–1430, Nov. 1990.

[6] A. Kagan, "A discrete version of the Stam inequality and a characterization of the Poisson distribution," *J. Statist. Planning and Inference*, vol. 92, pp. 7–12, 2001.

[7] I. Kontoyiannis, P. Harremoës, and O. Johnson, "Entropy and the law of small numbers." Submitted for publication.

[8] W. Hoeffding, "On the distribution of the number of successes in independent trials," *Ann. Math. Statist.*, vol. 27, pp. 713–721, 1956.

[9] P. Harremoës, "Convergence to the Poisson distribution in information divergence," Tech. Rep. 2, Mathematical department, University of Copenhagen, Feb. 2003.

[10] F. Topsøe, "Maximum entropy versus minimum risk and applications to some classical discrete distributions," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2368–2376, Aug. 2002.