COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 3, No. 3, pp. 183-190, March 2004
© 2004 International Press
003

A NASH EQUILIBRIUM RELATED TO THE POISSON CHANNEL*

P. HARREMOËS† AND C. VIGNAT‡

Abstract. An information theoretical game is considered where both signal and noise are generalized Bernoulli sums with upper bounds on their mean values. It is shown that a pair of Poisson distributions is a Nash equilibrium pair.

1. Introduction. We consider the following setup: an emitter transmits a signal X through a channel; a jammer sends noise Y on the same additive channel, and the received signal is $Z = X + Y$. The signals X and Y are assumed to be independent. The emitter wants to maximize the transmission rate i.e. the mutual information $I(X;Z)$ by choosing an appropriate distribution of X. Conversely, the objective of the jammer is to choose the distribution of Y such that the transmission rate is minimized.

$$\begin{array}{ccccc} X & \longrightarrow & \bigoplus & \longrightarrow & Z \\ & & \uparrow & & \\ & & Y & & \end{array}$$

For continuous random variables X and Y with power constraints of the form $E\left(X^2\right) \leq P$ and $E\left(Y^2\right) \leq N$, this problem has been studied by T. Cover and S. Diggavi in [1, Exercise 1, p. 263] and in more detail in [2]. In this case the normal distributions with mean 0 and variances P and N respectively form a Nash equilibrium pair, in the sense that none of the players has any benefit of changing his strategy if the other player does not change his strategy either. The Entropy Power Inequality plays an essential role in the proof of the Nash equilibrium condition. For Bernoulli sources with addition modulo 2, the existence of a Nash equilibrium has been proved in a series of paper ([3], [4], [5], [6] and [7]). In this paper we shall consider a similar setup with discrete random variables and *usual* addition. This paper follows ideas related to discrete Fisher information (see [8], [9] and [10]) and the results should be considered as a step in the direction of a discrete Entropy Power Inequality.

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

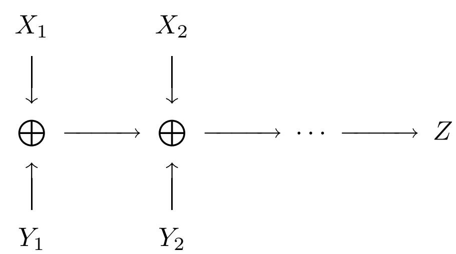
*Received on March 31, 2003; accepted for publication on January 22, 2004. The first author is supported by a post-doc fellowship from the Villum Kann Rasmussen Foundation and by INTAS (project 00-738) and by the Danish Natural Science Research Council.

†P. Harremoës is with Department of Math., University of Copenhagen, Denmark, E-mail: moes@math.ku.dk

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

$Bi(n,p)$. When no confusion seems likely $Po(\lambda)$ and $Bi(n,p)$ will also be used to denote random variables with these distributions. We denote by $B_n(\lambda)$ the set of sums of n independent Bernoulli random variables, each having success probability $p_i; 1 \leq i \leq n$ with $\sum_{i=1}^{n} p_i = \lambda$. Define the set of *Bernoulli sums* $B_*(\lambda)$ as the union $\cup_{n=1}^{\infty} B_n(\lambda)$. 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* $B(\lambda)$ is the total variation closure $cl(B_*(\lambda))$.

In this paper we shall consider the situation where the input signal X is the superposition of a number of independent signals X_i $(i = 1, ..., m)$ with values in $\{0, 1\}$. Thus $X = \sum_{i=1}^{m} X_i$. Similarly the noise Y is superposition of a number of independent noise signals Y_j $(j = 1, ..., n)$ with values in $\{0, 1\}$, and $Y = \sum_{j=1}^{n} Y_j$. Thus the output signal is $Z = \sum_{i=1}^{m} X_i + \sum_{j=1}^{n} Y_j$ and the mapping $(X_1, X_2, ..., X_m, Y_1, Y_2, ..., Y_n) \to Z$ can be considered as a multiple access channel with $m + n$ senders, where the first m senders want to send information to the receiver and the last n senders want to prevent the communication. We observe that we have a two-persons zero sum game with the emitter and the jammer as players.



If X is fixed and Y is binomial distributed $Bi(n, 1/2)$ then $I(X;Z) \to 0$ for $n \to \infty$. Similarly, if Y is fixed and X is binomially distributed $Bi(m, 1/2)$ then $I(X;Z) \to \infty$. Therefore, if there are no restrictions on the Bernoulli sums sent by the players, then the two-persons zero sum game will not reach an equilibrium. In this paper the strategies of both players are subject to constraints

$$E(X) \leq \lambda_{\text{in}}, \quad E(Y) \leq \lambda_{\text{noise}}$$

where λ_{in} and λ_{noise} are positive constants. Thus we shall consider the case when the messages X and Y are supposed to be generalized Bernoulli sums: $X \in B(\lambda), \lambda \leq \lambda_{\text{in}}$ and $Y \in B(\mu), \mu \leq \lambda_{\text{noise}}$. The reason to allow distributions from $B(\lambda)$ and not only from $B_*(\lambda)$ is for mathematical convenience. 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 form a Nash equilibrium pair 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.

2. Results.

LEMMA 1. *Let S be a generalized Bernoulli sum with $E(S)=\lambda$. Then, for any $k\geq 0$,*

$$P(S=k)\leq \exp(\lambda)\cdot Po(\lambda,k).$$

Proof. Assume that S is the Bernoulli sum $\sum_{i=1}^{n} X_i$ where X_i are independent Bernoulli random variables with success probabilities $p_i=P(X_i=1)$. For the Bernoulli random variable X_i we have $P(X_i=k)\leq \exp(p_i)\cdot Po(p_i,k)$. Then

$$\begin{aligned}P\left(\sum_{i=1}^{n} X_i=k\right) &= \sum_{(k_i)}\prod_{i=1}^{n} P(X_i=k_i)\\ &\leq \sum_{(k_i)}\prod_{i=1}^{n}\exp(p_i)\,Po(p_i,k_i)\\ &= \sum_{(k_i)}\exp\left(\sum_{i=1}^{n} p_i\right)\prod_{k=1}^{n} Po(p_i,k_i)\\ &= \exp\left(\sum_{i=1}^{n} p_i\right)\cdot Po\left(\sum_{i=1}^{n} p_i,k\right),\end{aligned}$$

where the summation is over all generalized indices $(k_i)\in\{0,1\}^n$ which satisfy $\sum_{i=1}^{n} k_i=k$. The inequality follows because $\sum_{i=1}^{n} p_i=\lambda$.

A generalized Bernoulli sum satisfies the same inequality because a generalized Bernoulli sum can be approximated by a sequence of Bernoulli sums. □

PROPOSITION 2. *For any generalized Bernoulli sum $T\in B(\lambda)$, there exists a real number $\mu\in[0;\lambda]$ and a random variable $S\in B_\infty(\lambda)$ such that $T=S+Po(\mu)$, where S is independent of the Poisson random variable.*

Proof. Let G be the set of random variables of the form $S+Po(\mu)$ where $S\in B_\infty(\lambda-\mu)$. First let us prove that G is closed. There is a correspondence between distribution of random variables of the form $S+Po(\mu)$ and the set of decreasing sequences $(p_i)_{i\in\mathbb{N}}$ such that $\sum_{i=1}^{\infty} p_i\leq\lambda$. The set of such sequences is compact when it is equipped with the topology of pointwize convergence. The mapping from sequences to distributions of the corresponding Bernoulli sums is continuous. Therefore G is compact and therefore also closed.

Now we have to demonstrate that for any element $S+Po(\mu)\in G$ there exists a sequence of Bernoulli sums $(S_n)_{n\in\mathbb{N}}$ with $S_n\in B_n(\lambda)$ such that the distribution of S_n converges to the distribution of $S+Po(\mu)$. The Poisson distribution $Po(\mu)$ can be approximated by a sequence of binomial distribution $Bi(n,\mu/n)$, and therefore it is sufficient to prove that $S\in B_\infty(\lambda-\mu)$ can be approximated by Bernoulli sums. Assume that $S=\sum_{i=1}^{\infty} X_i$ where X_i are independent Bernoulli random variables with success probabilities $p_i=P(X_i=1)$. Then $\sum_{i=n+1}^{\infty} p_i$ converges to zero. Thus, for n

sufficiently large, $\sum_{i=n+1}^{\infty} p_i \leq 1$, and there exists a Bernoulli random variable Y_n with $P(Y_n = 1) = \sum_{i=n+1}^{\infty} p_i$. Then the distribution of $\sum_{i=1}^{n} X_i + Y_n$ will approximate S. To prove this it is sufficient to check that

$$\begin{aligned}\left\| S - \left(\sum_{i=1}^{n} X_i + Y_n \right) \right\| &\leq \left\| \sum_{i=n+1}^{\infty} X_i - Y_n \right\| \\ &= \sum_{k=0}^{\infty} \left| P\left(\sum_{i=n+1}^{\infty} X_i = k \right) - P(Y_n = k) \right| \\ &\leq \sum_{k=1}^{\infty} 2 \exp\left(\sum_{i=n+1}^{\infty} p_i \right) Po\left(\sum_{i=n+1}^{\infty} p_i, k \right) \\ &\to 0 \text{ for } n \to \infty.\end{aligned}$$

□

We need the following extension of Hoeffding's theorem [11, Theorem 3].

Lemma 3. *If $S \in B(\lambda)$ and f is a strictly concave function, then $E[f(S)]$ is minimized for $S \sim Po(\lambda)$.*

Proof. Assume that $E|f(Po(\lambda))| = \infty$. The function f is concave and therefore it is upper bounded by a linear function. Thus the positive part of f is upper bounded on $\mathbb{R}_+$ by a linear function and therefore the mean value of the positive part is bounded, and the mean value of the negative part must be unbounded. This implies that $E[f(Po(\lambda))] = -\infty$, and this bound is obviously minimal.

Assume that $E|f(Po(\lambda))| < \infty$. Let P_n be a sequence of distributions of Bernoulli sums converging to Q. Then

$$|f(k) P_n(k)| \leq |f(k)| e^{\lambda} Po(\lambda, k)$$

and by the dominated convergence theorem

$$\sum_{k=0}^{\infty} f(k) P_n(k) \to \sum_{k=0}^{\infty} f(k) Q(k) \text{ for } n \to \infty.$$

Therefore $S \to E[f(S)]$ is a continuous function with a compact domain and thus has a minimum.

Assume that the minimum is reached at a generalized Bernoulli sum S of the form $T + U$ where T is a Bernoulli random variable with success probability $P(T = 1) = p$ and U is an independent generalized Bernoulli sum. Then

$$E(f(T + U)) = \sum_{k=0}^{\infty} \Pr(U = k) E[f(T + k)].$$

Now, replace T by a random variable $V \sim Bi(2, p/2)$. This random variable has the same mean value as T and

$$\begin{aligned}E\left[f\left(V+k\right)\right] &= \left(1-\frac{p}{2}\right)^2 f\left(k\right)+2\frac{p}{2}\left(1-\frac{p}{2}\right) f\left(k+1\right)+\left(\frac{p}{2}\right)^2 f\left(k+2\right)\\ &= \left(1-p\right) f\left(k\right)+pf\left(k+1\right)+\left(\frac{p^2}{4}f\left(k\right)-\frac{p^2}{2}f\left(k+1\right)+\frac{p^2}{4}f\left(k+2\right)\right)\\ &= E\left[f\left(T+k\right)\right]+\frac{p^2}{2}\left(\frac{f\left(k\right)+f\left(k+2\right)}{2}-f\left(k+1\right)\right)\\ &\geq E\left[f\left(T+k\right)\right].\end{aligned}$$

Equality holds if and only if $p=0$. Thus, at the minimum all the Bernoulli random variables in the generalized Bernoulli sum are deterministic and $S\sim Po\left(\lambda\right)$. □

The following theorem is a slight extension of theorems in [12] and [13].

THEOREM 4. *In the discrete transmission game the Poisson distribution is the optimal input distribution for any noise distributed according to a generalized Bernoulli sum, i.e.*

$$Po\left(\lambda_{in}\right)=\underset{X\in B(\lambda),\,\lambda\leq\lambda_{in}}{\arg\max}\; I\left(X;Z\right).$$

Proof. Rewrite the mutual information as

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

We look for the distribution of X which maximizes $H\left(X+Y\right)$. We have $H\left(X+Y\right)\leq H\left(X+Po\left(\lambda_{\text{in}}-E\left(X\right)\right)+Y\right)$, and therefore we may assume that $E\left(X\right)=\lambda_{\text{in}}$. Put $f\left(k\right)=\log\left(\Pr\left(Po\left(\lambda_{\text{in}}\right)+Y=k\right)\right)$. All Bernoulli sums are log-concave [12] and therefore the same property holds for generalized Bernoulli sums. According to [14, Theorem 6.1] it is sufficient to prove that $E\left(f\left(X\right)\right)$ is minimal for the Poisson distribution, but this follows from Lemma 3. □

Knowing that the Poisson distribution is the optimal input distribution, we choose a Poisson input, and look for the optimal noise from the jammer's point of view.

THEOREM 5. *If $X\sim Po\left(\lambda_{in}\right)$ in the discrete transmission game then the Poisson distribution is the optimal distribution for the jammer, i.e.*

$$Po\left(\lambda_{noise}\right)=\underset{Y\in B(\lambda),\,\lambda\leq\lambda_{noise}}{\arg\min}\; I\left(X;Z\right)\;.$$

Proof. Assume that $X\sim Po\left(\lambda_{\text{in}}\right)$. We thus have to solve the following optimization problem

$$Y=\underset{Y\in B(\lambda),\lambda\leq\lambda_{\text{noise}}}{\arg\min}\; I\left(X;Z\right)\;.$$

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

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

Therefore the 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 $k = 0, 1, 2, ...$ define

$$Q_k^\lambda = \frac{k+1}{\lambda + \lambda_{\text{noise}}} P_{k+1}^\lambda \ ,$$

and note that Q_k^λ is a sub-probability measure ($\sum_{k=1}^{\infty} Q_k^\lambda \leq 1$). Now we write

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

so that

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

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

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

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 if and only if $Y \sim Po\left(\lambda_{\text{noise}}\right)$. The second term can be written as

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

Now, according to Lemma 3 the term $E_{P^\lambda}\left[\log\left(Y+1\right)\right]$ is minimized when Y is Poisson distributed with mean λ_{noise}. □

The main result of this paper is now proved:

COROLLARY 6. *A pair of Poisson distributions is a unique Nash equilibrium pair in the discrete transmission game.*

3. 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. We have also had interesting discussions with Ioannis Kontoyiannis related to the possibility of an entropy power inequality related to the Poisson distribution. Finally we want to thank Flemming Topsøe who has given us several useful comments which have increased the quality of this paper.

REFERENCES

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

[2] S. N. Diggavi and T. M. Cover, *The worst additive noise under a covariance constraint*, IEEE Trans. Inform. Theory, IT-47(2001), pp. 3072–3081.

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

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

[5] H. S. Witsenhausen, *Entropy inequalities for discrete channels*, IEEE Trans. Inform Theory, IT-20(1974), pp. 610–616.

[6] R. Ahlswede and J. Körner, *On the connection between the entropies of input and output distributions of discrete memoryless channels*, In: Proceedings of the Fifth Conference on Probability Theory, 1974, Brasov, Romania (B. Bereanu, M. Iosifescu, and G. Popescu, eds.), (Bucuresti), pp. 13–22, The Centra of Mathematical Statistics of the Ministry of Education, Editura Academiei Republicii Socialiste Romania, 1977.

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

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

[9] I. Kontoyiannis and P. Harremoës, *Entropy and the law of small numbers*, In: Proceedings 2003 IEEE International Symposium on Information Theory, pp. 26, June 2003.

[10] I. Kontoyiannis, P. Harremoës, and O. Johnson, *Entropy and the law of small numbers*, Accepted for publication in IEEE Trans. Inform. Theory.

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

[12] P. Harremoës, *Binomial and Poisson distributions as maximum entropy distributions*, IEEE Trans. Inform. Theory, IT-47(2001), pp. 2039–2041.

[13] F. Topsøe, *Maximum entropy versus minimum risk and applications to some classical discrete distributions*, IEEE Trans. Inform. Theory, IT-48(2002), pp. 2368–2376.

[14] P. Harremoës and F. Topsøe, *Maximum entropy fundamentals*, Entropy, 3(2001), pp. 191–226. http://www.unibas.ch/mdpi/entropy/ [ONLINE].