Cooperative Detection over Multiple Parallel Channels: a Principle Inspired by Nature

Rudolf Mathar
Inst. for Theoretical Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
mathar@ti.rwth-aachen.de

Anke Schmeink
UMIC Research Center
RWTH Aachen University
D-52056 Aachen, Germany
schmeink@umic.rwth-aachen.de

***Abstract*— We consider an information theoretic model that describes parallel reception of a common signal by many receivers. The signal is distorted by noise and independently decoded by each receiver. Receivers cooperate in the sense that the outcome is conveyed to a decision center, which makes the final decision on the received signal by a majority vote. A potential scenario which could be described by this model is a cluster of base stations jointly receiving signals from a mobile station and conjoining individual decodings. Similar principles apply for cognitive radio setups where a secondary user employs many weak subcarriers, subject to sudden drop out by occupation from the primary users. The whole system is modeled as a cascade of channels, and quality of information exchange is measured by mutual information. This allows for modeling and optimizing quantization and detection in a unifying approach. Numerical evaluations demonstrate that the increase of mutual information by using additional channels is only logarithmic. 4-QAM is investigated as a concrete example and system performance is numerically investigated. The model in this paper is motivated by one used for describing information exchange in biological neural networks, revealing so called stochastic resonance by adding noise to signals.**

I. Introduction and Motivation

Analytical models are needed in order to quantify the effect of transmission over parallel channels and cooperatively combining information. On one hand, such models should describe all relevant effects, on the other hand, they should be simple enough to allow for analytical calculus and conclusion. In this paper, we suggest a biophysically inspired model, which explains and quantifies at least partial aspects of cooperation when using wireless channels in parallel, making individual assessments and thereafter merging information into a single decision.

Present wireless communication systems are highly optimized by using coding and multiaccess schemes that achieve reliable communication at highest possible rates. Modern digital communication systems come very close to the Shannon bound, which cannot be exceeded by whatever practical implementation. Biological communication and information systems in contrast have been optimized by evolution subject to different design criteria. Biological information channels use a rate which ensures survival and proliferation of the species. Signaling and communciation between cells, within the brain or between entities of a species are developed by evolution to a degree such that creatures are able to cope with environmental challenges. Speed and capacity are not the primary concerns, if both are sufficient they will no more be an objective of evolutionary optimization.

For biological systems communication and information exchange has to be extremely reliable in a wide range of situations. If some information sources or channels are not operational, their role should be taken over by others, still functional. Massive parallelism seems to be the solution to this problem in biological systems. The retina in the human eye, inner hair cells in the auditory cochlea, and the semicircular canals of the human ear, e.g., process information in parallel and convey quantized signals to the brain. Nerve tracts serve as channels and information is mainly processed in the brain, often after local quantization and compression. Low energy consumption paired with simplicity, efficiency and adaptability are further important objectives for information exchange in living organisms.

In this work, we study bio-inspired communication over parallel channels by the following class of models. A common signal is observed by many receivers, each afflicted by noise. At each receiver, signals are decoded and the result is reported to a central decision making unit. The measurement of each node is then combined into a single decision, which is expected to represent the original input signal. Practical applications that are covered by this model may be cellular networks, where signals of a mobile station are received by a cluster of cooperating base stations. Furthermore, in OFDM cognitive radio idle channels may be occupied by secondary users. They must be vacated as soon as primary users raise demand. Hence, a system of parallel channels is cooperatively used, some of which may easily become unavailable. The goal of this paper is to quantify such effects.

The model in this paper is adapted from a subclass of so called *stochastic pooling networks*, a denotation first coined by [1] for a binary detection problem. This class of networks has emerged as a useful model for many applications and has ignited a whole series of publications. An overview of potential applications is given in [2], [3]. Artificial sensor networks, digitized beamforming, stochastic resonance, biological neurons, cochlear implants and also complex social networks are prominent examples and widely investigated.

The effect that the presence of noise can enhance the detection of signals is called *stochastic resonance* (SR). This phenomenon is often observed in nature and a variety of man-made systems. Natural systems such as the animal and human brain or the visual, auditory and olfactory systems utilize stochastic resonance to enhance stimulus detection. The effect of noise onto detecting below threshold, time variant signals in biology and technological applications is investigated in [4]. A comprehensive review of models and applications in biology is given in [5].

However, not only the detection of subthreshold signals may gain from SR, also *suprathreshold* SR can increase the total amount of information for parallel channels with multi-threshold devices, as is demonstrated in [6]. In this work, a common Gaussian input signal X is observed by n sensors, each subject to independent Gaussian noise, 0-1-quantization applies with equal threshold values $\vartheta = \mathrm{E}(X)$. The number of 1's then forms the output signal. Fig. 1 with binary quantization and $U = u(Y_1, \dots, Y_n) = \sum_{i=1}^n Y_i$ specifies this model. Simulation and numerical computations in [6] demonstrate in concert that mutual information between input and output is enlarged by the presence of noise, particularly as the number of threshold devices increases.

The final step of combining quantized information into a single decision corresponds to information fusion, which is a well investigated subject in the literature. However, most papers investigate error probabilities in a Bayesian setup, see [7], while in this work emphasis is laid on maximizing mutual information from a source to a decision center.

In this paper, we approach the problem from an information theoretic point of view. Mutual information, $I(X;U)$, between a certain stochastic signal X and output U is used to describe the amount of information a channel is able to convey. We investigate mutual information stagewise between (i) the input and noisy observations hereof, (ii) observations and decoded signal, and finally (iii) between individual decodings and the final output by the central decision making unit. The final decoding is based on a majority vote over all individual observations. Since separately for each stage the outcome is a deterministic function of a stochastic input, mutual information coincides with the entropy of the output, as is demonstrated in section II. The majority vote is described by the maximum number of decisions for some signaling point. This leads to the maximum component of multinomially distributed random variables, whose distribution is difficult to obtain. However, an efficent numerical algorithm is available and used to determine the entropy of the final decision step. In a series of numerical evaluations we investigate the resilience of such parallel systems by varying the number of available channels.

II. Mathematical Prerequisites

Mutual information between two random variables X and Y is denoted by $I(X;Y)$ and generally defined as

$$I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X).$$

Let q be a function such that $Y = q(X)$ is well defined. Since $H\big(q(X) \mid X\big) = 0$ it follows from the above that

$$I\big(X; q(X)\big) = H\big(q(X)\big). \tag{1}$$

This identity will be used a number of times in the present work. Of particular interest is the case that function q represents quantization.

We call each Lebesgue measurable function q from $\mathcal{V} \subseteq \mathbb{R}^N$, the N-dimensional Euclidean space, into the set of integers $\{1, \dots, M\}$ an M-quantizer, $q : \mathcal{V} \to \{1, \dots, M\}$.

Let $\mathcal{V}_i = q^{-1}\big(\{i\}\big)$ denote the preimage of the singleton set $\{i\}$. Obviously, $\mathcal{V}_1, \dots, \mathcal{V}_M$ is a partitioning of $\mathcal{V}$. Vice versa, any partitioning $\mathcal{V}_1, \dots, \mathcal{V}_M$ defines an M-quantizer q by setting $q(v) = i$ whenever $v \in \mathcal{V}_i$, $i = 1, \dots, M$.

Let V be a random variable with density f. Optimum quantizers q that maximize mutual information $I(V; q(V))$ between V and $q(V)$ are characterized in the following proposition.

Proposition 1: Some M-quantizer q^* is a solution of

$$\max\ I(V; q(V)) \text{ over all } M\text{-quantizers } q$$

if and only if $\int_{\mathcal{V}_i^*} f(v)dv = \frac{1}{M}$ for all $i = 1, \dots, M$, where $\mathcal{V}_i^* = {q^*}^{-1}\big(\{i\}\big)$ denotes the preimage of the singleton $\{i\}$.

Proof. From (1) we conclude that

$$I\big(V; q(V)\big) = H(g_1, \dots, g_M)$$

with $g_i = P(V \in \mathcal{V}_i)$, $i = 1, \dots, M$. $H(g_1, \dots, g_M)$ is maximized for the uniform distribution $g_i = \frac{1}{M}$ for all $i = 1, \dots, M$, which can always be achieved by an appropriate choice of the sets $\mathcal{V}_i$. ∎

Optimum quantization hence means to split the image of random variable V into M equiprobable disjoint subsets $\mathcal{V}_1^*, \dots, \mathcal{V}_M^*$, each representing a quantization stage. In [8], the same concept has been applied to the quantization of log-likelihood ratios in coded modulation systems.

III. The System Model

Assume that there are n nodes, e.g., base stations which listen to a transmission of a mobile, say. We consider transmission symbol-wise and assume that the mobile uses symbol constellation $\mathcal{S} = \{s_1, \dots, s_M\} \subset \mathbb{C}$. The incoming bit stream is groupwise mapped onto the symbols, which generates random variable X with support $\mathcal{S}$, governed by distribution $(p_1, \dots, p_M)$.

At each node i the transmitted symbol X is received, however, subject to additive random noise modeled by the random variables W_i, $i = 1, \dots, M$, each having density f. We assume that $W_1, \dots, W_n$ are stochastically independent, also independent of X. The noisy observation at each receiver is denoted by

$$V_i = X + W_i,\ i = 1, \dots, n.$$

Each receiver decodes the signal by optimum quantization forming subsets $\mathcal{V}_1^*, \dots, \mathcal{V}_M^*$, each related to one of the symbols s_i. This process generates discrete random variables

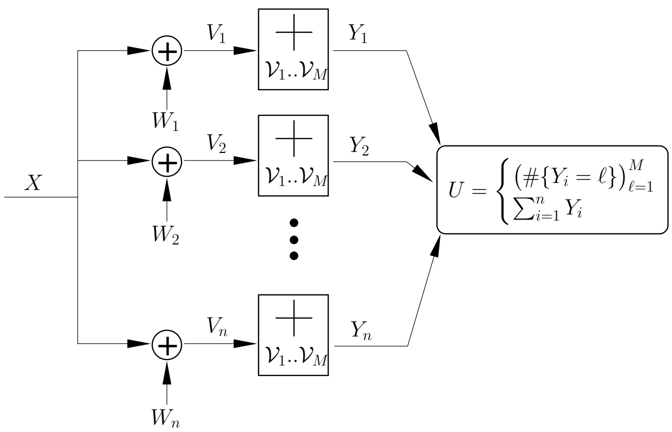

Fig. 1. A parallel channel model for joint detection.

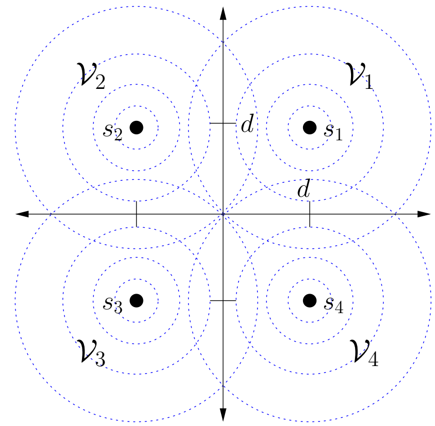

Fig. 2. 4-QAM, each point selected with equal probability $1/4$. Contourlines of the noise distribution are depicted by dotted circles.

$Y_1, \dots, Y_n$ each with support $\{1, \dots, M\}$. The final joint decision about the incoming signal X is made by a majority vote. Symbol k is chosen if it is decoded by the majority of receivers. The system model is depicted in Fig. 1.

IV. Optimum Quantization

We first consider optimum quantization separately for each subchannel in Fig. 1 and therefor momentarily fix index i. Optimality is meant in an information theoretic sense such that

$$\max I\big(V_i; q(V_i)\big) \text{ over all } M\text{-quantizers } q$$

must be solved.

The distribution of V_i itself is rather complicated. If W_i possesses some density f_W and the distribution of X with support $\{s_1, \dots, s_M\}$ is given by $(p_1, \dots, p_M)$, then the distribution of V_i is the same for each i, given by the mixture density

$$f_V(v) = \sum_{k=1}^{M} p_k f(v - s_k).$$

Its entropy is hard to compute in general. One of the few papers which deal with this problem in the case of Gaussian f is reference [9]. Numerical and asymptotic results are achieved for a symmetric one-dimensional mixture of two Gaussians.

Whatever this distribution in concrete is, according to Proposition 1 the optimum quantizer q^* will achieve a discrete uniform distribution over $\{1, \dots, M\}$ such that

$$P(Y_i = k) = \frac{1}{M}, \quad k = 1, \dots, M.$$

From the above, after optimum quantization $\boldsymbol{Y} = (Y_1, \dots, Y_n)$ is a discrete random vector with support $\{1, \dots, M\}^n$. All marginals are uniform, i.e., $P(Y_i = k) = \frac{1}{M}$ for all $k = 1, \dots, M$. Moreover, $Y_1, \dots, Y_n$ are conditionally stochastically independent, given X. The conditional distributions $P(Y_i = k \mid X = \ell)$, $k = 1, \dots, M$, given $\{X = \ell\}$, $\ell = 1, \dots, M$, are the same for all $i = 1, \dots, n$. We denote the conditional probabilities by

$$p_{\ell k} = P(Y_i = k \mid X = \ell), \;\; k, \ell = 1, \dots, M.$$

The joint distribution of $(Y_1, \dots, Y_n)$ may then be written as

$$P(Y_1 = k_1, \dots, Y_n = k_n) = \sum_{\ell=1}^{M} p_\ell \cdot p_{\ell,k_1} \cdots p_{\ell,k_n}.$$

Although the marginals are all known to be uniform when applying optimum quantization, the joint distribution of $(Y_1, \dots, Y_n)$ and the conditional probabilities $p_{k\ell}$ are difficult to determine in general. However, more can be done if the signaling constellation and the error distribution are known and not too complicated, as is the case for 4-QAM.

V. 4-QAM

We investigate the example of 4-QAM with four signaling points

$$s_1 = (d, d), \; s_2 = (-d, d), \; s_3 = (-d, -d), \; s_4 = (d, -d),$$

as depicted in Fig. 2, each occurring with probability $p_k = 1/4$. Let $f(w)$ denote the density of Gaussian noise with expectation zero and covariance matrix $\sigma^2 \boldsymbol{I}_2$.

The set $\mathcal{V}_\ell$ corresponding to decoding of s_ℓ is assumed to be the ℓ-th quadrant as shown in Fig. 2. It follows that

$$p_{\ell k} = P(Y_i = k \mid X = \ell) = \int_{\mathcal{V}_k} f(w - s_\ell) dw.$$

Because of symmetry only three different values occur and the matrix of $p_{\ell k}$ may be arranged as

$$(p_{\ell k}) = \begin{pmatrix} a & b & c & b \\ b & a & b & c \\ c & b & a & b \\ b & c & b & a \end{pmatrix}. \tag{2}$$

Let $\Phi_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-w^2/(2\sigma^2)} dw$ denote the cumulative distribution function of the Gaussian with variance σ^2. The values of a, b, c are given by

$$a = \big(1 - \Phi_\sigma(-d)\big)^2,$$
$$b = \big(1 - \Phi_\sigma(-d)\big)\, \Phi_\sigma(-d),$$
$$c = \Phi_\sigma^2(-d).$$

If X is uniformly distributed, then from the special form of (2) it follows that $P(Y_i = k) = 1/4$, which demonstrates optimality of the partitioning $\mathcal{V}_1, \dots, \mathcal{V}_4$ with $I(V_i; Y_i) = H(Y_i)$. Additionally using that $I(X; Y_i) = H(Y_i) - H(Y_i \mid X)$ we obtain with $M = 4$

$$I(V_i; Y_i) = \log M \quad \text{and} \quad I(X; Y_i) = \log M - H(a, b, b, c)$$

where $H(a, b, b, c)$ denotes the entropy of the stochastic vector (a, b, b, c).

VI. Detection Fusion

We investigate two different ways to combine information from the individual decodings Y_i, $i = 1, \dots, n$, into a single decision. First, the number of decision in favor of symbol ℓ is considered. The outcome will be a vector of integers $(N_1, \dots, N_M)$, component ℓ counting the number of decodings of symbol ℓ. Secondly, to further compress information into a single real value, $\sum_{i=1}^{N} Y_i$ is explored. This is a generalization of the sum criterion for binary quantization, see [2], [3], [6], and sensitive to identifying which symbol has been emitted by random varibale X.

To formalize the first approach we introduce random variables N_ℓ which count the number of detections of symbol ℓ,

$$N_\ell = \#\{1 \le i \le n \mid Y_i = \ell\}, \ \ell = 1 \dots, M.$$

We aim at determining the mutual information $I\big((Y_1, \dots, Y_n); (N_1, \dots, N_M)\big)$. By (1), mutual information coincides with the entropy $H(N_1, \dots, N_M)$. To compute this entropy the distribution of $(N_1, \dots, N_M)$ is needed.

For this purpose, we first condition on the event $\{X = \ell\}$. The conditional distribution of $(N_1, \dots, N_M)$ given $\{X = \ell\}$ is the multinomial with parameters $p_{\ell,1}, \dots, p_{\ell,M}$ and n,

$$\begin{aligned} P(N_1 = n_1, \dots, N_M = n_M \mid X = \ell) \\ = \frac{n!}{n_1! \cdots n_M!}\, p_{\ell 1}^{n_1} \cdots p_{\ell M}^{n_M} \end{aligned}$$

whenever $n_k \ge 0$, $\sum_{k=1}^{M} n_k = n$, and 0 otherwise. The joint distribution of $(N_1, \dots, N_M)$ then reads as

$$\begin{aligned} P(N_1 = n_1, \dots, N_M = n_M) \\ = \frac{n!}{n_1! \cdots n_M!} \sum_{\ell=1}^{M} p_\ell\, p_{\ell 1}^{n_1} \cdots p_{\ell M}^{n_M} \end{aligned} \qquad (3)$$

In a second approach, the information from $(Y_1, \dots, Y_n)$ is further compressed into a single real value by random variable

$$U = \sum_{i=1}^{n} Y_i = \sum_{k=1}^{M} k N_k.$$

The distribution of U is obtained as

$$P(U = s) = \sum_{\ell=1}^{M} p_\ell \sum_{i_1 + \dots + i_n = s} p_{\ell i_1} \cdots p_{\ell i_n} \qquad (4)$$

The entropy of $(N_1, \dots, N_M)$ and U can be easily computed from (3) and (4), respectively. An explicit form of the distributions seems to be hard to achieve. Numerical values for 4-QAM are determined in the next section.

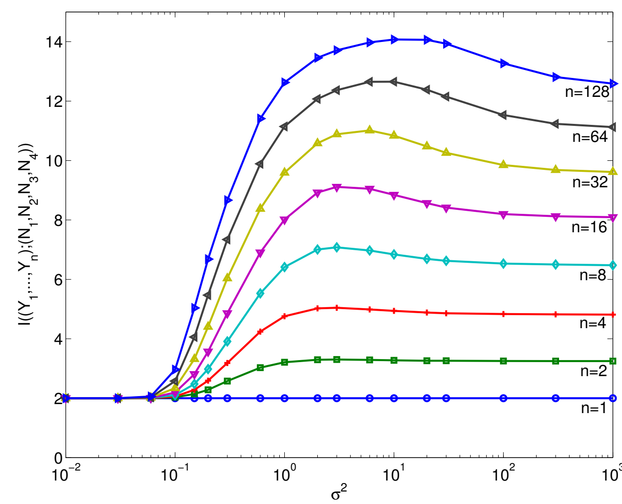

Fig. 3. Mutual information $I\big((Y_1, \dots, Y_n); (N_1, \dots, N_4)\big)$ as a function of σ^2 for $n = 1, 2, 4, \dots, 128$. Stochastic resonance is observed.

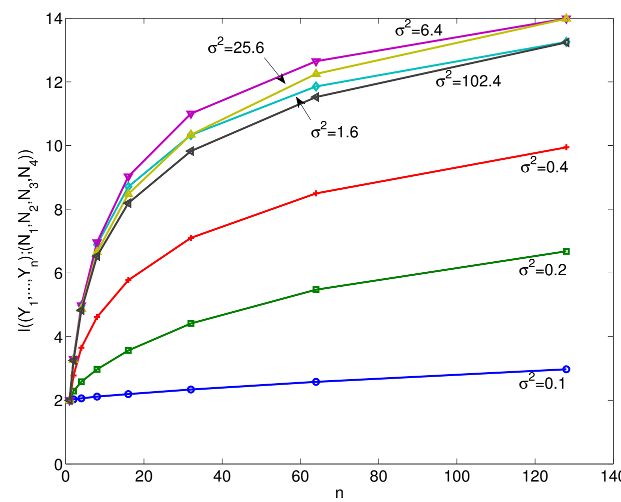

Fig. 4. Mutual information $I\big((Y_1, \dots, Y_n); (N_1, \dots, N_4)\big)$ as a function of n for $\sigma^2 = 0.1, 0.2, 0.4, 1.6, 6.4, 25.6, 102.4$.

VII. Numerical Tests for 4-QAM

We consider modulation alphabet 4-QAM with $d = 1$, equal symbol probabilities $p_1 = p_2 = p_3 = p_4 = 0.25$ and different numbers n of subchannels. The noise vectors W_i are assumed to be stochastically independent and follow a two-dimensional Gaussian with expectation 0 and covariance matrix $\sigma^2 \boldsymbol{I}_2$. The values a, b, c in (2) are easily computed and given in Table I, which gives $p_{\ell 1}, p_{\ell 2}, p_{\ell 3}, p_{\ell 4}$ in (3) and (4).

The probability mass function of (N_1, N_2, N_3, N_4) from (3) and the corresponding entropy are computed numerically for σ^2 between 10^{-2} and 10^{3} and $n = 2^k$, $k = 0, 1, \dots, 7$. Mutual information $I\big((Y_1, \dots, Y_n); (N_1, \dots, N_4)\big)$ is depicted in Fig. 3 as a function of σ^2 for different numbers of subchannels n. If σ^2 is close to zero, no errors occur and the

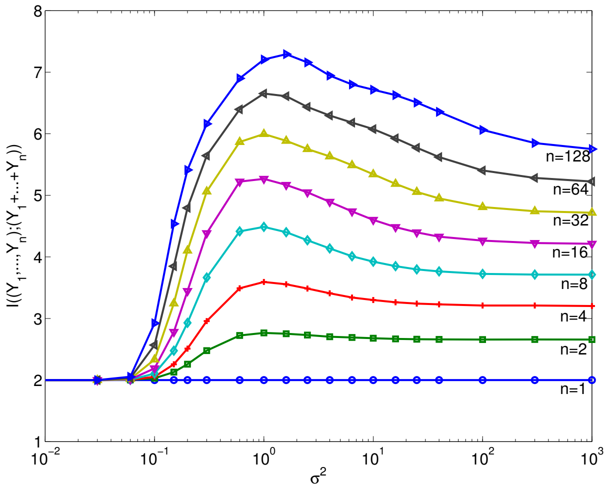

Fig. 5. Mutual information $I\big((Y_1,\dots,Y_n);Y_1+\cdots+Y_n\big)$ as a function of σ^2 for $n = 1,2,4,\dots,128$. Stochastic resonance is observed.

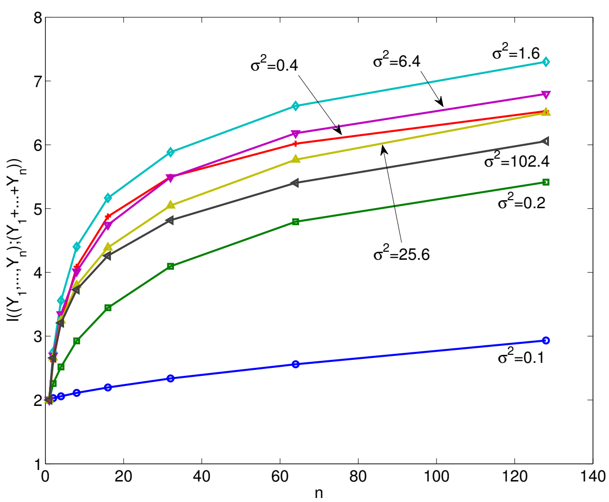

Fig. 6. Mutual information $I\big((Y_1,\dots,Y_n);Y_1+\cdots+Y_n\big)$ as a function of n for $\sigma^2 = 0.1, 0.2, 0.4, 1.6, 6.4, 25.6, 102.4$.

mutual information amounts to two bits per channels use. With increasing variance stochastic resonance may be observed in the sense that larger noise helps conveying information in a many-channel system, see [6] for a similar effect. Mutual information of the channel, $I\big((Y_1,\dots,Y_n);(N_1,\dots,N_4)\big)$, first increases to a certain maximum and decreases thereafter. In the limit as $\sigma^2 \to \infty$, I tends to the entropy of a multinomial distribution with parameters n and $(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})$.

To quantify the influence of the number of channels, in Fig. 4 mutual information is represented as a function of n for moderate σ^2 values. Because of the logarithmic shape of mutual information as a function of the number of subchannels n, failure of a few channels out of many will only marginally deteriorate the quality of information transfer. However, if merely a few channels are available, then each one counts.

σ^2	$\Phi_\sigma(-1)$	a	b	c
0.5	0.0786	0.8489	0.0725	0.0062
1.0	0.1587	0.7079	0.1335	0.0252
1.5	0.2071	0.6287	0.1642	0.0429
2.0	0.2398	0.5780	0.1823	0.0575

TABLE I

NUMERICAL VALUES $p_{\ell k}$ FOR 4-QAM.

Failure of only one will already drastically degrade system performance.

Finally, information from the individual detectors is compressed into a single real value by $U = \sum_{i=1}^{n} Y_i$. Similar to the above, mutual information $I\big((Y_1,\dots,Y_n);U\big)$ is depicted in Fig. 5 as a function of σ^2 for $n = 1,2,4,\dots,128$. Although mutual information is lower than in the previous case, stochastic resonance is more prominent here. Analogously, mutual information is shown as a function of the number of channels n for moderate values of σ^2 in Fig. 6.

VIII. CONCLUSIONS

We think that the model used in this paper is able to quantify certain aspects of cooperative communication in many-channel systems. It can certainly be extended to cover a wider range of situations, although at the price of increased analytical and numerical complexity. The main contribution of this work is a model for cooperative channel systems, providing the mathematical background for determining mutual information and quantization stagewise, and deriving explicit results for quadrature amplitude modulation with four signaling points.

Acknowledgement. This work was supported by the UMIC Research Center at RWTH Aachen University. The authors would like to thank Christoph Schmitz for carrying out the numerical computations.

REFERENCES

[1] S. Zozor, P.-O. Amblard, and C. Duchene, “On pooling networks and fluctuation in suboptimal detection framework,” *Fluctuation and Noise Letters*, vol. 7, no. 1, pp. L39–L60, 2007.

[2] M. D. McDonnell, P.-O. Amblard, and N. G. Stocks, “Stochastic pooling networks,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 1, pp. 1–18, January 2009.

[3] ——, “Bio-inspired communication: Performance limits for information transmission and compression in stochastic pooling networks with binary quantizing nodes,” *Journal of Computational and Theoretical Nanoscience*, vol. 7, no. 5, pp. 876–883, May 2010.

[4] K. Wiesenfeld and F. Moss, “Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs,” *Nature*, vol. 373, pp. 33–36, jan 1995.

[5] F. Moss, L. M. Ward, and W. G. Sannita, “Stochastic resonance and sensory information processing: a tutorial and review of application,” *Clinical Neurophysiology*, vol. 115, pp. 267–281, 2004.

[6] N. Stocks, “Suprathreshold stochastic resonance in multilevel threshold systems,” *Physical Review Letters*, vol. 84, no. 11, pp. 2310–2313, March 2000.

[7] G. Fabeck and R. Mathar, “Performance evaluation of sensor fusion with side information,” in *IEEE ISSPA 2010*, Kuala Lumpur, Malaysia, May 2010.

[8] C. Novak, P. Fertl, and G. Matz, “Quantization for soft-output demodulators in bit-interleaved coded modulation systems,” in *International Symposium on Information Theory (ISIT)*, Seoul, Korea, July 2009, pp. 1070–1074.

[9] J. Michalowicz, J. Nichols, and F. Buchholz, “Calculation of differential entropy for a mixed Gaussian distribution,” *Entropy*, vol. 10, pp. 200–206, 2008.