Revisiting the asymmetric binary channel: joint noise-enhanced detection and information transmission through threshold devices

Pierre-Olivier Amblard\textsuperscript{a}, Olivier J.J Michel\textsuperscript{b} and Savérion Morfu\textsuperscript{c}

\textsuperscript{a}Laboratoire des Images et des Signaux, Groupe Non-Linéaire
UMR CNRS 5083, Grenoble,
ENSIEG, BP 46, 38 402 Saint Martin d’Hères cedex, France
\textsuperscript{b}Laboratoire Universitaire d’Astrophysique de Nice, équipe MATIS
UMR CNRS 6525, Nice, Université de NICE-Sophia Antipolis
Parc Valrose, 06 108 Nice cedex 2, France
\textsuperscript{c}Laboratoire d’Électronique, Informatique et Image
UMR CNRS 5158, Dijon, UFR Sciences et Techniques, Université de Bourgogne
9, allée Alain Savary, BP 47870, 21078 DIJON Cedex , France

ABSTRACT

In this paper we revisit the asymmetric binary channel from the double point of view of detection theory and information theory. We first evaluate the capacity of the asymmetric binary channel as a function of the probabilities of false alarm and of detection. This sets the \textit{a priori} probabilities of the hypotheses and couples the two points of view. We then study the simple realization of the asymmetric binary channel using a threshold device. We particularly revisit noise-enhanced processing for subthreshold signals using the aforementioned parametrization of the capacity, and we report an paradoxical effect: using the channel at its capacity forbids in general an optimal detection.

Keywords: noise-enhanced processing, asymmetric binary channel, mutual information, capacity, detection

1. MOTIVATION AND AIMS

The aim of this work is (re)examine the asymmetric binary channel from the double perspective of information theory and detection theory. One instance of this channel is a threshold device attacked by a noisy binary input. This system has been shown to have the noise-enhanced processing property, and furthermore, this channel can be seen as one of the easiest ways to model the behaviour of real neurons.

The asymmetric binary channel is depicted in figure (1) when considered as a communication channel. The channel aims at transmitting a binary random variable $x$, taking the values 0 or 1 with the respective probabilities $\pi_0$ and $\pi_1 = 1 - \pi_0$. The channel is completely characterized by the probabilities of the output conditioned by the input, \textit{i.e.} by the knowledge of $P_{i,j} = \text{Prob}(y = i | x = j), i = 0, 1, j = 0, 1$. The probability of receiving 0 (resp. 1) when 0 (resp. 1) is emitted is denoted $\alpha$ (resp. $\beta$). The channel is called asymmetric whenever $\alpha \neq \beta$.

Another way of looking at the channel is to consider it as an instance of a binary detection problem (see figure 2). In that point of view, the problem is to decide which hypothesis among $H_0$ and $H_1$ is true, the decision being performed after the observation of a function of these hypotheses. In figure (2), we have specialized this to the observation of a noisy constant, $s_0$ in the case $H_0$ is true or $s_1$ if $H_1$ is true. A detector transforms the observation into a value which is finally compared to a threshold $\eta$ to take a decision. In the bottom of figure (2), we have depicted the parametrization of this binary detection problem as an asymmetric binary channel. In this view, we have $\alpha = 1 - P_f$, where $P_f = \text{Prob}(\text{decide } 1 | H_0 \text{ true})$ is called the probability of false alarm, and $\beta = P_d = \text{Prob}(\text{decide } 1 | H_1 \text{ true})$ is called the probability of detection. In a binary hypothesis testing framework, these two quantities are often referred to as the level and the power of the detector respectively.
Figure 1. The asymmetric binary channel when considered as a communication channel. It is completely specified by the a priori probabilities $\pi_{0,1}$ and by the matrix of transition probabilities $P(y|x)$. It is asymmetric if $\alpha \neq \beta$.

Figure 2. The asymmetric binary channel when considered as a detection scheme.

Note that $P_m = 1 - P_d$ is the second kind of probability of error and is called the probability of miss, that the probability of deciding that $H_0$ is true when it is not.

Finally, figure (3) shows a realization of the asymmetric binary channel using an additive noisy signal which is compared to a threshold. This simple channel is of course well known in information theory, but it has seen a regain in interest for about ten years now as a prototype of channel that depicts noise-enhanced processing properties. The story of stochastic resonance in threshold devices dates back to the middle of the nineties,\textsuperscript{1–6} even if the real story begins with the works of engineers in the late forties\textsuperscript{7} and in the sixties\textsuperscript{8} (more about these pioneering works on quantization may be found in the recent review by Gray&Neuhoff\textsuperscript{9}). The works on stochastic resonance and threshold devices was mainly motivated by neurobiology, since neurons are basically threshold devices and since the brain is known to be noisy. If the first works used the usual setting of stochastic resonance,\textsuperscript{1,2} it was rapidly realized that the noise-enhanced processing property of thresholds is nondynamical,\textsuperscript{5} and that the use of the words stochastic resonance is improper. Rapidly also, the analysis turned to the problem of transmission of information in these devices, and the measures used to quantify the noise-enhanced processing properties were chosen in information-theoretic measures.\textsuperscript{4,6} The use of information measures to study noise-

Further author information: (Send correspondence to P.O. Amblard)
e-mails: bidou.amblard@lis.inpg.fr, olivier.michel@unice.fr, smorfu@u-bourgogne.fr
enhanced transmission has been further developed since then in single threshold elements,\textsuperscript{10–12} in array of threshold devices,\textsuperscript{13–15} in the usual setting of stochastic resonance.\textsuperscript{16} In this last work, the point of view was that of detection theory. It is the aim of the paper to offer a fresh look at threshold devices from the double point of view of transmission of information and detection theory.

In section 2, we study the asymmetric binary channel from the double point of view of detection theory and information theory. This is done using a parametrization of the capacity of the channel as a function of the probabilities of false alarm and of detection, hence coupling the information theory point of view of the channel to the detection theory point of view. Section 3 is dedicated to the case of the simple threshold devices, a particular instance of the asymmetric binary channel. We revisit for this threshold device the noise-enhanced processing property, and examine also whether the mutual information and the probability of error can be simultaneously extremized.

2. CAPACITY AND PROBABILITY OF ERROR

2.1. Using the detection parametrization

In this section, we study the capacity of the channel and the probability of error in the same parametrization, namely, as function of the probability of false alarm $P_f$ and of the probability of detection $P_d$. We are therefore in the point of view of figure (2). The transmission of information through a channel is quantitatively measured by the Shannon mutual information (or transinformation) between the input and the output. This measure quantifies the information contained in the output $y$ of the channel, once the information or randomness introduced by the channel noise and perturbations have been taken into account. Actually, in the absence of any distortion or noise in the channel, the conditional entropy of $y$ for a given input $x$ is expected to be zero. Mathematically, the mutual information reads

$$I(X,Y) = H(Y) - H(Y|X) = -E_Y[\log p_Y(y)] + E_X[H(Y|(X=x))]$$

where $E_U[.]$ stands for the mathematical expectation over $U$, and where $E_U[.|V]$ is the conditional expectation of $U$ conditioned on $V$. Thus, $H(Y|X = x)$ is the entropy of the output random variable $Y$ when the output takes value $X = x$. The base of the logarithm may be 2 or $e$ (information is then measured in bits or nats).

It is clear from the definition of the mutual information that it explicitly depends upon both the probability distribution of the input and the perturbations introduced by the channel. Therefore, the mutual information is not a characteristic of the channel only. To get such a characteristic, a possibility is to consider the maximum amount of transinformation over the probability distributions of the input. This leads to the notion of rate or equivalently of capacity of the channel. This quantity, formally defined as

$$C = \sup_{p_X} I(X,Y)$$

is fundamental in communication theory since it provides an upper bound to the information that one could transfer through the channel in such a way that the receiver can instantaneously recover the emitted message with an arbitrarily small error.\textsuperscript{17} For the setting given in figure (2), we see that $Y$ is a Bernouilli random variable taking the value 1 with probability $\text{Proba}(Y=1) = \pi_1 P_d + \pi_0 P_f$ and the value 0 with $\text{Proba}(Y=0) = 1 - \text{Proba}(Y=1)$. The entropy $H(Y) = -E_Y[\log p_Y(y)]$ of a Bernouilli random variable of parameter $p = \text{Proba}(Y=1)$ reads

$$h(p) \overset{df}{=} -p \log p - (1-p) \log(1-p)$$

Furthermore, from figure (2), we deduce that $Z = Y|X$ is a Bernouilli random variable, with parameter $P_d$ if $X = 1$ and with parameter $P_f$ if $X = 0$. Therefore, the conditional entropy $H(Y|X)$ is given by $H(Y|X) = \pi_1 h(P_d) + \pi_0 h(P_f)$. Finally, the mutual information of the asymmetric binary channel parametrized with the probability of detection and the probability of false alarm reads

$$I(X,Y) = h(\pi_1 P_d + \pi_0 P_f) - \pi_1 h(P_d) - \pi_0 h(P_f)$$

(1)
This quantity is always positive (or null); this can easily be seen in our case as $I(X,Y)$ is nothing but the $\alpha$-Jensen concavity measure on $h^*$ Maximizing the mutual information $I(X,Y)$ with respect to the probability distributions of the input is easy in this problem since the possible inputs are binary, and hence their probability distribution is governed by parameter $\pi_1$. The value $\pi^*$ of $\pi_1$ that maximizes the mutual information is given by

$$\pi^*(P_d, P_f) = \frac{1 - \beta(P_d, P_f)P_f}{\beta(P_d, P_f)(P_d - P_f)}$$

where $\beta(P_d, P_f) = 2^{-\frac{h(P_d) - h(P_f)}{P_d - P_f}} + 1$ (2)

if the logarithm is in base 2. The capacity is then obtained by inserting $\pi_1 = \pi^*$ in the expression of the mutual information given in equation (1), and it becomes a function $C(P_f, P_d)$ of the two probabilities of interest.

Furthermore, once $\pi_1$ is set to the optimal value $\pi^*$, the probability of error reads

$$P_e^* = (1 - \pi_1)P_f + \pi_4(1 - P_d)_{\pi_1=\pi^*}$$

$$= (1 - \pi^*(P_d, P_f))P_f + \pi^*(P_d, P_f)(1 - P_d)$$

(3)

2.2. Discussion

The above formulation of the detection problem allows to separate the influence of the input probability law and the influence of the noise model in the computation of maximum information transferred through the channel. Actually, $P_f$ and $P_d$ are conditional probability values (computed under the condition that input are $s_0$ and $s_1$ respectively), and do not depend upon the input probability density functions. Consequently, the noise or perturbation model which contains our knowledge about the physical properties of the channel is entirely contained in the values of $P_f$ and $P_d$. However, the probability of error ($P_e = \pi_0, P_f + \pi_1(1 - P_d)$) and the mutual information $I(X,Y)$ are functions of both the channel perturbations and the input probability law. Both may be used to characterize the quality of the detector. This raises the following question: are they equivalent, in the sense the $I(X,Y)$ is maximal when $P_e$ is minimal?

If one focuses on the system by considering it as a detection device intended to decide whether $s_0$ or $s_1$ was emitted, the pertinent quantity that has to be considered is the probability of wrong decision, or probability of error $P_e$. Orsak & Paris \(^5\) have established that if $P_e$ is minimal with respect to some parameters of the problem (e.g. the threshold value in threshold detection device), it can be expressed as a function of a Csiszar-Ali-Silvey (CAL) divergence \(^6\) between the conditional probability density functions (pdfs) at the detector input, $P_i(X)$ and $P_0(X)$ associated with the assumptions that emitted signal is $s_1$ and $s_0$ respectively:

$$P_{e,MIN} = \frac{1}{2} - \frac{1}{2} E_0 [\pi_1 L(x) - \pi_0] = \frac{1}{2} - \frac{1}{2} E_0 [g(L(x))]

$$

where $E_0$ is the expectation with respect to the input probability function under the assumption that the transmitted signal is $s_0$, $L(x)$ is the likelihood ratio (ratio of the likelihood for each hypothesis) of the input signal, and $g(x) = \pi_1 x - \pi_0$. This CAL-divergence is referred to as the Kolmogorov divergence. Conversely, they proved that for suboptimal detectors (exhibiting $P_e > P_{e,MIN}$), $P_e$ cannot be expressed as a divergence in general. This implies that a design of a suboptimal detection strategy based on the maximization of the divergence between the input pdfs may lead to inconsistent solutions.

If on the contrary, the device is seen as a transmission device devoted to transfer information from the input toward the output, the mutual information seems the natural quantity which has to be optimized. The mutual information function $I(X,Y)$ can easily be expressed in terms of an information divergence:

$$I(X,Y) = E_{(X,Y)} \left[ -\log \left( \frac{P(X)P(Y)}{P(X,Y)} \right) \right] = D_{KL}(P(X,Y)||P(X)P(Y))$$

\(^*\)For any concave function $h$ and any $0 \leq \alpha \leq 1$, $J_{\alpha,h}(x,y) \overset{def}{=} h(\alpha x + (1 - \alpha)y) - \alpha h(x) - (1 - \alpha)h(y) \geq 0$
This again expresses an expectation of a likelihood, but this time it measures a divergence between the joint probability density function of the input and output $P(X,Y)$ with the probability density function that could be observed if the output $Y$ and the input $X$ were statistically independent.

These results illustrates that both mutual information or probability of decision error may be interpreted in terms of CAL divergences. However, for an optimal decision device, $P_e$ is expressed in terms of a divergence involving input statistics only, whereas $I(X,Y)$ involves both input and output statistics. This illustrates the fact that optimizing the device with respect to an information channel point of view or with respect to a decision making device approach, do not lead in general to equivalent solutions. This implies in particular that $P_e^*$ in equation (3) is in general strictly larger to $P_{e\text{MIN}}$.

3. THE PARTICULAR CASE OF THE THRESHOLD DETECTOR

We now turn to the case of the threshold device. Two parameters rules out the behavior of the probabilities of false alarm and detection, for a given normalized probability density function (pdf) of the input noise: The threshold $\eta$ and the power of the input noise $\sigma^2$. Consequently, the channel capacity for a given normalized noise pdf is expressed in terms of these two parameters only. In this section, we illustrate the noise enhanced performance of the nonlinear threshold devices, by considering either the mutual information or the decision error probability. This allows to shed a new light on the noise enhanced behavior of threshold devices. The values $(P_f, P_d)$ defines the position of a point on the channel capacity surface $C(P_f, P_d)$. Hence, varying $\eta$ and/or $\sigma^2$ amounts to moving the point on the capacity surface. This original approach allows a new description of noise enhanced detection, by investigating the meaning of this motion. We will put forward some known results concerning the threshold device. Therefore we first need to evaluate $P_f$ and $P_d$.

**Evaluation of the probabilities.** The notation are those given in figure (3). Let $f_n(x)$ be the normalized probability density function of the noise $n$ (i.e. when this one has power unity and zero mean). Let then $F_n(x) = \int_{-\infty}^{\eta} f_n(\mu) d\mu$. Hence, the probability of false alarm is given by

$$P_f = \text{Proba} (s_0 + n \geq \eta) = 1 - F_n \left( \frac{\eta - s_0}{\sigma} \right)$$

and the probability of detection by

$$P_d = \text{Proba} (s_1 + n \geq \eta) = 1 - F_n \left( \frac{\eta - s_1}{\sigma} \right)$$

3.1. Channel Capacity and minimal probability of error

From the preceding expressions of $P_f(\eta, \sigma)$ and $P_d(\eta, \sigma)$, and using equations (1) and (2), the capacity $C(\eta, \sigma)$ of the channel is easily obtained. Similarly, for each couple $\eta, \sigma$, $P_e$ can be computed; for sake of consistency, we chose to compute $P_e^*$ (equation (3)), the decision error probability obtained when the source is adapted to the channel in the sense that $I(X,Y)$ reaches the capacity. Note that from equation (2), $\pi^*$ is not defined for $P_f = P_d$. In this case, we chose $P_d = P_f = 1/2$ by continuity.

Figures (4) and (5) show the behavior of the channel capacity and the decision error probability in the $(\eta, \sigma)$ plane for the Gaussian channel and the channel corrupted by uniform noise respectively. For these simulations, we set $s_0 = 0$, $s_1 = 2$. Both simulations illustrate that noise enhanced performances can be attained if the threshold does not belong to the interval $[s_0, s_1]$; this case is referred to as a subthreshold situation. Note that this may also be easily obtained by computing the optimal value (in the $P_e$ sense) of $\sigma$:

$$\sigma^* = \text{Arg}_\sigma \left[ \frac{dP_e}{d\sigma} = 0 \right]$$
Figure 4. (a): Capacity for the Gaussian channel, computed as a function of the detector threshold \( \eta \) and the noise rms amplitude \( \sigma \). The dotted lines represent the trajectory of the maximal capacity wrt \( \eta \) and \( \sigma \). The lower plot shows the capacity as a function of \( \sigma \) for \( \eta = \frac{1}{2} \) (overthreshold case) and \( \eta = 2.5 > s_1 \) (subthreshold case). The left plot shows the capacity wrt \( \eta \) for \( \sigma = 1 \). (b): \( P_e \) as a function of \( \eta \) and \( \sigma \). Dotted lines represent the trajectory of the minimal \( P_e^* \) wrt \( \eta \) and \( \sigma \). The dotted line on the lower plot indicates the minimum \( P_e \) that could be obtained by adjusting \( \eta \) for each value of \( \sigma \), when all other parameters are those computed for maximizing the capacity.

Solving the above equation for the Gaussian channel leads to

\[
\sigma^* = \sqrt{\frac{2\eta(s_1 - s_0) + s_0^2 - s_1^2}{2 \ln \left( \frac{1 - \pi_1(\eta - s_0)}{\pi_1(\eta - s_1)} \right)}}
\]

which requires that \( \eta > s_1 \) or \( \eta < s_0 \) to be defined. A similar result holds for the channel corrupted by uniform noise; a noise enhanced capacity can be obtained for subthreshold input signals only.

Now, we focus on the departure between the optimality considered with respect to \( P_e \) and the optimality with respect to the channel capacity \( C \). In the preceding figures, it appears that the additive Gaussian noise rms amplitude \( \sigma^*_C \) which allows to maximize \( C \) is larger than the rms amplitude of the noise \( \sigma^*_P \) leading to the lowest \( P_e \). On the contrary, for uniform noise, the \( P_e \) optimal value that can be reached by adjusting the threshold is not better that the \( P_e \) value obtained by adjusting \( \sigma \). (see figure 5-(b), lower plot). Figure (6) and (7) illustrates these remarks, see also figure 5-(b), lower plot.

3.2. Capacity and modified ROC curves

If we set the power of the noise to a particular value and let the threshold \( \eta \) varies, the motion of the point with coordinates \((P_f(\eta), P_d(\eta))\) describes a curve in the \((P_f, P_d)\) plane. This curve is the so-called Receiver Operational Characteristics or ROC curves. The parametric equation of the ROC is easily obtained: Precisely, we extract \( \eta \) from \( P_f \) by \( \eta = s_0 + \sigma F_n^{-1}(1 - P_f) \) and insert this expression into the probability of detection to obtain

\[
P_d = 1 - F_n \left( \frac{s_0 - s_1}{\sigma} + F_n^{-1}(1 - P_f) \right)
\]
Figure 5. (a): Capacity for the channel corrupted by uniformly distributed noise, computed as a function of the detector threshold \( \eta \) and the noise rms amplitude \( \sigma \). The lines represent the trajectory of the maximal capacity wrt \( \eta \) (dotted line) and \( \sigma \) (●). The lower plot shows the capacity as a function of \( \eta = 2.5 > s_1 \) (subthreshold case), and the capacity that could be reached by adjusting \( \eta \) wrt \( \sigma \), all other parameters being kept the same. The left plot shows the capacity wrt \( \eta \) for \( \sigma = 1.44 \). (b): \( P_c^* \) as a function of \( \eta \) and \( \sigma \). The superimposed lines represent the trajectory of the minimal \( P_c^* \) wrt \( \eta \) (dotted line) and \( \sigma \) (●). The dotted line on the lower plot indicates the minimum \( P_c \) that could be obtained by adjusting \( \eta \), for each \( \sigma \), when all other parameters are computed for maximizing the capacity.

Figure 6. (a): Gaussian noise rms amplitude value \( \sigma_{c}^* \); maximizing \( C \) as a function of the Gaussian noise rms amplitude value \( \sigma_{c}^* \), minimizing \( P_c \). The straight line of equation \( \sigma_{c}^* = \sigma_{c}^* \), provides a reference. (b): Gaussian channel; \( P_c^* \) and \( C \) wrt \( \sigma \), for \( \eta = 3.2 > s_1 = 2 \).

**Modified ROC curves.** On the contrary, if the threshold is set to a particular value and the power of the noise is varied, we obtain another motion whose coordinates are given by

\[
P_d = 1 - F_{\eta} \left( \frac{\eta - s_1}{\eta - s_0} F_{\eta}^{-1}(1 - P_f) \right)
\]
Figure 7. Left: $P_e^*$ and $C$ for a channel corrupted by uniform noise, wrt the rms amplitude of the noise $\sigma$. Notice that optimal behavior is obtained for the same $\sigma$ for both optimization criteria. For this simulation, $\eta = 2.99 > s_1$. Right: uniform noise rms amplitude value $\sigma_C^*$ maximizing $C$ as a function of the uniform noise rms amplitude value $\sigma_{P_e}$ minimizing $P_e$. This allows to construct for any threshold detector, its trajectory in the $(P_f, P_d)$ plane as the rms amplitude of the additive noise varies. This is appealing as the channel capacity may be expressed in the same plane. Figure 8 depicts the trajectory expressed by equation (4) superposed onto the capacity surface, computed for the Gaussian channel. The lower triangle, associated to biased detectors ($P_f > P_d$) has not been investigated as a simple coin flipping detector to decide wether $s_0$ or $s_1$ was emitted, would lead to better performance in this case. An interesting feature of this graph is the concavity of the modified ROCs inside the upper left square defined by $(P_f < .5, P_d > .5)$. Inside this square, starting from the point $[P_f = 0, P_d = 1]$ and describing the trajectory, one observes that the trajectory is concave and that increasing $P_f$ leads to decrease $C$. On the contrary, outside
this area, the trajectory is convex and increasing $P_f$ from the point $[P_f = .5 = P_d]$, the trajectory goes through increasing values of $C$, then through decreasing values of $C$. Recall that each of these trajectory is computed by letting $\sigma$ vary for a fixed threshold value $\eta$. Thus it is straightforward to conclude that increasing the rms amplitude of the Gaussian noise will first improve the behavior of the threshold detector, go through an optimum, then decrease. Furthermore, it can be shown that the horizontal and vertical straight line trajectories are obtained for $\eta = s_0$ and $\eta = s_1$ respectively. This is a new illustration that noise-enhanced performance in threshold detectors are possible for subthreshold scenarios only. The existence of a noise-enhanced behavior

![Figure 9](image-url)

**Figure 9.** Derivative of the Gaussian channel capacity with respect to the rms amplitude of the noise $\sigma$. Each curve is associated to a fixed value $\eta$ of the threshold, larger than $s_1$ (subthreshold case) on the right hand side plot or in the interval $[s_0, s_1]$ (over-threshold case) on the left hand side plot.

![Figure 10](image-url)

**Figure 10.** Left : Gaussian channel capacity as a function of $\sigma$ for fixed values of the threshold $\eta > s_1$. Best capacity are reached for threshold values that are not too far from $s_1$. Right: Derivative of the channel capacity with respect to $\sigma$ for the curves shown on the left hand side plot.

for this subthreshold Gaussian channel is confirmed by the computation of the derivative $\frac{dC}{d\sigma} = f(\sigma)$ for a fixed value of $\eta$. Figure (9) shows the values of the derivative for the suprathreshold scenario and for the subthreshold scenario respectively. It appears that the equation $\frac{dC}{d\sigma} = 0$ admits a solution (unique) only if the threshold takes values larger than $s_1$. Notice that a symmetric behavior can be observed for $\eta < s_0$. On the contrary, for suprathreshold situations, the derivative is always negative, thus indicating that the performances of the threshold detector worsen as the rms amplitude of the noise increases.

The plots in figure (10) shows the channel capacity as a function of $\sigma$ for subthreshold situations. The trajectory indicates the maxima of the channel capacity when $\eta$ varies. The right hand side plot shows the derivative of the curves $C = f(\sigma)$ drawn on the left plot.
4. CONCLUSION

In this paper, the asymmetric binary channel has been studied from the double point of view of detection theory and information theory. Parametrizing the capacity of the channel as a function of the probabilities of false alarm and of detection allows to couple the points of view. Furthermore, this allows to superimpose onto the capacity surface the characteristic curves of detection theory. This parametrization is very interesting since it allows to see these characteristics as motions on the capacity surface, motions parametrized by some internal parameters of the channel. For example, we describe the realization of the asymmetric binary channel by a simple threshold device acting on a noisy signal. In this example, the internal parameters are the threshold $\eta$ and the noise rms amplitude $\sigma$. We have analytically established the parameter range in the $(\eta, \sigma)$ plane and in the $(P_d, P_f)$ plane where the well-known noise-enhanced processing effect takes place. This marginal range of parameter, corresponding to a non optimal setting of the device from both detection and information point of view, can be less marginal in the case of array of threshold devices where the optimal thresholds distribution cannot be optimally set.\textsuperscript{13,14}

Another question also studied in this paper has been the ability of a channel to be simultaneously optimal for detection and transmission of information. This question has been answered for the threshold device. In general, the answer to the question is negative, and this mean that the device is optimized to transfer as much as information as possible, it will not be able to perform a detection in an optimal fashion. This fact, also pointed out by Morse\&Stocks,\textsuperscript{20} is given here a theoretical basis, and could be extended to more complex devices such as nonlinear dynamical systems.

REFERENCES


