SPEEDIER SEQUENTIAL TESTS VIA STOCHASTIC RESONANCE

Marco Guerriero, Peter Willett

University of Connecticut
ECE Department
Storrs, U-2157, CT 06269 USA

Stefano Marano, Vincenzo Matta

University of Salerno
DIIIE
Fisciano (SA), Italy

ABSTRACT

Stochastic Resonance (SR) is a phenomenon long investigated by physicists that has recently attracted some interest in the signal processing literature. In this paper, we explore the potential benefits of the SR effect for shift-in-mean detection problems, specifically focusing on sequential decision rules. Amenable formulas for the optimal distribution of the SR noise, as well as an asymptotic comparison with the traditional Neyman-Pearson approach are obtained.

Index Terms— Stochastic Resonance, Sequential Detection.

1. INTRODUCTION AND MOTIVATION

Roughly speaking, Stochastic Resonance (SR) basically arises when, by injecting noise at the input of a (nonlinear) device, an increase of the signal-to-noise ratio at the corresponding output is observed. The physics literature is rich with contributions about the subject, since the pioneering work by Benzi *et al.* at the beginning of the '80s [1]. With no pretence of completeness, we refer the interest reader to, *e.g.*, [2, 3] as first entry-points.

Even if interesting works on the subject have been proposed to the signal processing community (see, *e.g.*, [4]), fuller connections with detection theory have been properly elucidated only recently, starting from the provoking question posed by Kay [5], continuing with the analysis in [6], and culminating into [7], where Chen *et al.* established a mathematical theory to deal with the SR effect in the context of detection.

Motivated by the latter works, in this paper we focus on a binary hypothesis test with the basic novelty of studying the SR effect in *sequential* detectors [8], and, as far as we can tell, this is the first investigation along this direction.

Aside from the theoretical aspects, the proposed strategy can be useful in several practical scenarios. Just to get a flavor of the possible applications, consider a Wireless Sensor Network engaged in a detection task. As usual, the nodes are severely constrained in terms of energy, computational capacity, *etc.*, and are accordingly forced to deliver some nonlinear transformation $t(\cdot)$ of the original observations. A remote fusion center is demanded to implement a sequential test in order to take the final decision. Can we improve on such scheme?

Let us allow the sensors to add some random noise to their own observations, *before* applying the nonlinearity. Perhaps surprisingly, this strategy may furnish a positive answer to the above question about performance improvements as we are promptly going to illustrate.

The paper is organized as follows. The addressed problem is formalized and the proposed solution is shown in Sect. 2. Section 3 is devoted to explore some examples in order to validate the theoretical analysis. In Sect. 4 we summarize.

2. PROBLEM STATEMENT AND MAIN RESULTS

We consider a shift-in-mean binary hypothesis test in the form:

$$\begin{aligned} \mathcal{H}_0 &: \ X_i \sim f_X(x + A), \\ \mathcal{H}_1 &: \ X_i \sim f_X(x - A), \end{aligned}$$

where $i = 1, 2, \ldots, \infty$, the X_i's are independent and identically distributed (iid), and A is a known location parameter. The probability density function (pdf) $f_X(x)$ is assumed to be an even function, *i.e.*, $f_X(x) = f_X(-x)$. The corresponding Cumulative Distribution Function (CDF) will be denoted by $F_X(x)$.

Inspired by the SR effect already studied in similar problems, we *contaminate* the observations with some additive noise, that is, for $i = 1, 2, \ldots, \infty$, we let $Y_i = X_i + W_i$, where the W_i's are iid and further independent of all the X_i's. Again, we consider even symmetry: $f_W(w) = f_W(-w)$. Since $f_Y(y) = f_X(x) * f_W(w)$, it is immediate to recast the test as

$$\begin{aligned} \mathcal{H}_0 &: \ Y_i \sim f_Y(y + A), \\ \mathcal{H}_1 &: \ Y_i \sim f_Y(y - A), \end{aligned} \tag{1}$$

where it is not difficult to recognize that, again, $f_Y(y) = f_Y(-y)$.

The basic question posed in this paper is whether a nonzero W_i exists giving an improvement of the detector performances on the scheme operating with uncontaminated data. In addition, we would like to select the best noise pdf f_W as the one yielding the speedest sequential test.

To answer the question, let us introduce the form of the decision statistic to be implemented. Given the sequential nature of the detection procedure, the system is constrained to employ an additive statistic in the form $T_n = \sum_{i=1}^{n} t(y_i)$. In order to meet the regularity conditions needed in the following, and in complying with the problem symmetries, the nonlinearity $t(y)$ is chosen as a bounded, non-decreasing odd function. Furthermore, with the aim of simplifying the analysis since it is actually unnecessary, we fix both the false dismissal probability and the false alarm probability of the test to one and the same value, hereafter denoted by p_e.

We are now in the position of defining the sequential decision rule: for $n = 1, 2, \dots$

$$\begin{cases} T_n \geq \gamma & \Rightarrow \quad \text{decide } \mathcal{H}_1, \\ T_n \leq -\gamma & \Rightarrow \quad \text{decide } \mathcal{H}_0, \\ \text{otherwise} & \Rightarrow \quad \text{continue collecting}, \end{cases}$$

where the choice of symmetric thresholds γ and $-\gamma$ is obvious. In the following, we assume the functional form of the detector *fixed*, while the threshold can be tuned by the user as the employed noise pdf f_W changes. Depending on the particular application, it could be also of interest to consider a threshold independent of the noise distribution, as well as a variable detector structure.

As standard in sequential analysis [8], the system performances are summarized in the error probability p_e and the Average Sample Number (ASN) $E[N]$ (which by symmetry is the same under both the hypotheses). In the case that the decision statistic is just the log-likelihood ratio of the observations, it is relatively easy to compute approximate analytical formulas for these two performance figures, see, *e.g.* [8].

On the other hand, in our scenario, due to the arbitrary shape of the nonlinearity, we must use a different analysis. To this aim, we shall resort to some known results about a random walk with two barriers reported, *e.g.*, in [9]. Working under $\mathcal{H}_1$, and using the large-deviation bound provided in [9], we get $p_e \leq e^{-\gamma r^*}$, where r^* is the solution of the equation $E\left[e^{t(X+W)\, r^*}; \mathcal{H}_0\right] = 1$. This gives a relationship between the error probability and the threshold.

Switching now to the ASN evaluation, by Wald's equality [9] we can write

$$E[N] \approx \frac{-\log p_e}{E[t(X+W); \mathcal{H}_1]\, r^*}, \tag{2}$$

where the last formula is obtained by using the above large-deviation approximation for the error probability, and in the regime of p_e small enough, which we refer to.

By defining the auxiliary functions

$$h_1(w) = E_X[t(X+w); \mathcal{H}_1]; \quad h_2(w) = E_X[e^{t(X+w)\, r}; \mathcal{H}_0], \tag{3}$$

we have the more compact formulas:

$$E[N] \approx \frac{-\log p_e}{E_W[h_1(W)]\, r^*}, \qquad r^* : E_W[h_2(W, r^*)] = 1. \tag{4}$$

The best noise pdf f_W , namely that yielding the smallest ASN among all sequential tests with error probability not exceeding p_e, is thus related to the following optimization:

$$\max_{f_W} \{E_W[h_1(W)]\, r^*\}\,. \tag{5}$$

We elaborate on the above in the next section.

Before ending this section, it is of interest to introduce a way to compare the proposed sequential strategy with a traditional Neyman-Pearson (NP) test. While it is well known that the (average) sample number required by the optimal Wald's SPRT is, for the same error probabilities, smaller than that required by the (fixed-sample-size) Neyman-Pearson test, the result of a comparison between a *suboptimal* sequential strategy and the optimal NP test is not clear in advance. To summarize the relative merits of a sequential detector with respect to a NP detector, a classical choice is to work in terms of the asymptotic relative efficiency defined as [10]

$$\text{ARE} = \lim_{p_e \to 0} \lim_{A \to 0} \frac{E[N]}{N_f}, \tag{6}$$

where N_f is the number of samples of the NP test. By working in the limit of vanishing A, the central limit theorem allows approximating the pertinent sample number N_f as $\left[Q^{-1}(p_e)\right]^2 \ A/I_f$, see [11], where $Q(\cdot)$ is the unit Gaussian exceedance function, and I_f is the Fisher information for location [11]. Depending on the particular scenario considered, one has to substitute the pertinent expression for $E[N]$ in eq. (6), and compute

$$\text{ARE} = \lim_{p_e \to 0} \lim_{A \to 0} \frac{E[N]\ I_f A^2}{[Q^{-1}(p_e)]^2}. \tag{7}$$

3. APPLICATIONS

As a possible example of application, we consider the sign nonlinearity $t(x) = \text{sign}(x)$, which was subject to considerable attention in earlier works about the stochastic resonance, see, *e.g.*, [5]. To carry out the optimization in (5), we now evaluate the functions $h_1(w)$ and $h_2(w, r)$ in eq. (3). By definition we obviously get $h_1(w) = 2F_X(A + w) - 1$, yielding $E_W[h_1(W)] = 2p - 1$, where $p = \Pr[X + W > 0; \mathcal{H}_1] = E_W[F_X(A + W)]$. On the other hand, it is straightforward to compute $E_W[h_2(W, r)] = e^r - 2p \sinh(r)$, whence, solving the second of eq. (4) yields $r^* = \log \frac{p}{1-p}$. Thus, the optimization problem (5) amounts to the maximization of

$$E_W[h_1(W)]\, r^* = (2p - 1) \log \frac{p}{1 - p}. \tag{8}$$

The above is an increasing function of p, revealing that we have simply to find the $f_W(w)$ maximizing p. To this aim, let us observe that, thanks to the symmetry of $f_W(w)$, we can rewrite

$$p = \int_0^\infty [F_X(A - w) + F_X(A + w)] f_W(w) dw. \tag{9}$$

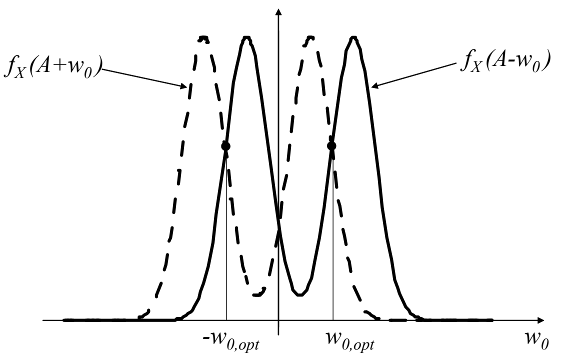

Fig. 1. Pictorial sketch of the optimization for the sign detector. According to eq. (11), the shifted versions of the pdf $f_X(\cdot)$ are depicted, and the optimal solution $w_{0,opt}$ is shown at the intersection between the two curves.

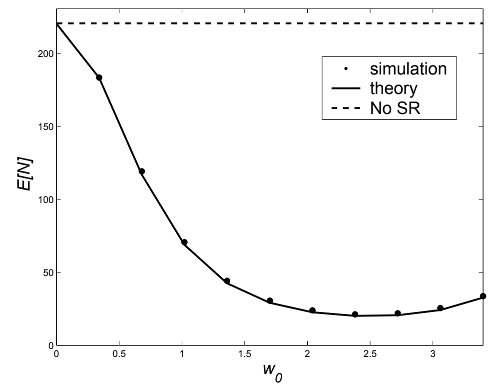

Fig. 2. ASN of the proposed sequential strategy for the sign nonlinearity, as a function of the noise level w_0. The relevant parameters are $p_e = 10^{-2}$, $A = 1$ and $\mu = 5$. Simulation points are obtained by 10^4 Monte Carlo trials.

Under regularity conditions (essentially $F_X(x)$ must be continuously differentiable), by the mean-value theorem it is immediate to recognize that, for a fixed $f_W(w)$, there always exists a value $w_0 \geq 0$ such that

$$p = \frac{1}{2}[F_X(A - w_0) + F_X(A + w_0)] := p(w_0), \qquad (10)$$

and the optimal solution $w_0 = w_{0,opt}$ obeys[1]

$$p'(w_0) = 0 \Longleftrightarrow f_X(A + w_0) = f_X(A - w_0). \qquad (11)$$

Remarkably, the above analysis also tells that the optimum noise pdf can be chosen in the class of the *coin flipping* distribution of strength w_0

$$f_W(w) = \frac{1}{2}[\delta(w - w_0) + \delta(w + w_0)], \qquad (12)$$

where $\delta(\cdot)$ is the Dirac delta. The relevant implication is that the optimal SR solution amounts to adding or subtracting, choosing which of these completely at random, a certain deterministic value $w_{0,opt}$ to the system input.

It is also of interest to compute the ARE for the described sequential sign detector in the presence of the stochastic resonance effect. Let us first observe that, for each value of A, a different injected noise $w(A) = w_{0,opt}$ is obtained, and thus the explicit expression of p in eq. (10) gives:

$$\frac{1}{2}[F_X(A - w(A)) + F_X(A + w(A))] := \psi(A). \qquad (13)$$

Using eq. (8), the ASN can be written as $E[N] = \frac{-\log p_e}{g(A)}$, where we define $g(A) := (2\psi(A)-1)\log\frac{\psi(A)}{1-\psi(A)}$. According to eq. (15), the ARE is thus given by

$$\text{ARE} = \frac{I_f}{2}\lim_{A\to 0}\frac{A^2}{g(A)} = \frac{I_f}{g''(0)}, \qquad (14)$$

where we used the fact that $\lim_{p_e\to 0}\log p_e/[Q^{-1}(p_e)]^2 = -1/2$ and it is implicitly assumed that $g(A)$ can be Taylor expanded around $A = 0$, with $g'(0) = 0$. By simply using the definition of $w(A)$ in eq. (13), it is straightforward to compute $g''(0) = 16[f_X(w(0))]^2$. A further simplification is obtained by noting that $f_X(w(0))$ is just a peak[2] of $f_X(x)$. The performances in the absence of stochastic resonance can be derived in a similar way. Summarizing:

$$\text{ARE} = \frac{I_f}{16[\max f_X(x)]^2}, \qquad \text{ARE}_{noSR} = \frac{I_f}{16[f_X(0)]^2}. \qquad (15)$$

We are now ready to explore the potential benefits of the stochastic resonance effect. Following the same line of previous works (see, *e.g.*, [7]), let us consider a Gaussian mixture $f_X(x) = \frac{1}{2}\left[\mathcal{N}(x;\mu/2) + \mathcal{N}(x;-\mu/2)\right]$, where $\mathcal{N}(x;a)$ is our shortcut for a Gaussian pdf of argument x, with mean a and unit variance. It could be useful to get a pictorial interpretation of eq. (11) for this specific pdf. This is provided in Fig. 1, where it is emphasized that the optimal solution requires looking for a symmetry of the pdf $f_X(\cdot)$ with respect to the value A. Clearly, a solution is here achieved thanks to the peaks of the *bimodal* pdf $f_X(\cdot)$, depending on the parameters A and μ.

Let us now focus on the performances of the presented strategy. Accordingly, in Fig. 2 we display the ASN needed for reaching a decision with an error probability $p_e = 10^{-2}$, as a function of the injected noise strength w_0: an optimal value $w_{0,opt}$ exists that minimizes the ASN, and the SR clearly arises. The theoretical values of the ASN predicted by the first of eq. (4) are also compared with the results of 10^4 Monte Carlo simulations. Note the good match between the

[1]Of course, also the sign of the second derivative should be checked.

[2]In fact, by rewriting, thanks to the even symmetry of f_X, eq. (11) as $f_X(A + w(A)) = f_X(w(A) - A)$, and simply taking the first derivative of both sides thereof, we get $f'_X(w(0)) = 0$.

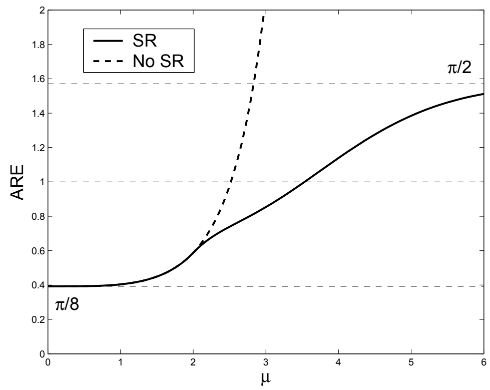

Fig. 3. Asymptotic relative efficiencies (15) of the proposed detectors, with and without the SR effect, displayed vs. the Gaussian mixture parameter μ.

simulated ASN and the values dictated by the theory. In addition, it has been checked that the error probabilities estimated via simulation are close to the large-deviation bound.

Before concluding, let us examine the asymptotic relative efficiencies obtained for this example. First, consider $\mu = 0$, so that the problem is purely Gaussian. In this case $I_f = 1$, and $w(A) = 0 \ \forall A$, yielding ARE $=$ ARE$_{noSR} = \pi/8 < 1$, as is known from [10]. It is worth noting that, since ARE < 1, the optimal NP detector is here outperformed by the sequential sign test. On the other hand, for $\mu \to \infty$, it can be verified that $I_f \approx 1$, and $\max f_X(x) \approx \frac{1}{2\sqrt{2\pi}}$, yielding ARE $=$ $\frac{\pi}{2} > 1$, which reveals that, as $\mu \to \infty$, the optimal NP test outperforms the SR sequential detector. These two opposite tendencies stimulate a deeper analysis of the ARE behavior as a function of the parameter μ, as shown in fig. 3, where three different regimes can be identified. For $\mu \leq 2$ there is no room for stochastic resonance. Indeed, in that region it can be shown that the pdf $f_X(x)$ is still unimodal. Furthermore, note that the sequential detection is advantageous, being ARE < 1.

For $\mu > 2$, and up to $\mu \approx 3.5$, the SR effect gives benefits with respect to the NP test (ARE < 1), and to the plain sequential detector (ARE $<$ ARE$_{noSR}$). Inside this interval, for values greater than $\mu \approx 2.5$, the ARE of the SR detector is less than unity *and* that of the sequential detector in the absence of stochastic resonance is not. Beyond $\mu \approx 3.5$, the SR sequential detector is worse than the NP one, and the maximum ARE is found to be $\pi/2$. However, orders of magnitude are gained with respect to the absence of stochastic resonance as μ grows.

4. CONCLUSIONS

We investigated the SR effect in the context of *sequential* detection, for shift-in-mean binary detection problems and assuming some symmetries which can be reasonably encountered in practical problems. The optimal injected noise distribution is characterized in terms of the closed-form optimization problem formalized in (5).

As an example, the sequential sign detector is considered: The system optimization is carried out in its full generality and the best enhancing noise is shown to be a *coin flipping* of appropriate strength $w_{0,opt}$. The comparison between the proposed sequential strategy and the optimal Neyman-Pearson fixed test is also addressed, in terms of asymptotic relative efficiency. The analysis emphasizes that there exist situations where the combined effect of the stochastic resonance and the sequential detection yields some gain, even with respect to the optimal NP test.

5. REFERENCES

[1] R. Benzi, A. Sutera, and A. Vulpiani, “The mechanism of stochastic resonance,” *J. Phys. A: Math. General*, vol. 14, pp. L453–L457, 1981.

[2] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, “Stochastic resonance,” *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223–287, Jan. 1998.

[3] M. E. Inchiosa and A. R. Bulsara, “Signal detection statistics of stochastic resonators,” *Phys. Rev. E*, vol. 53, no. 3, pp. R2021–R2024, Mar. 1996.

[4] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, “Sequential signal encoding from noisy measurements using quantizers with dynamic bias control,” *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[5] S. M. Kay, “Can detectability be improved by adding noise?,” *IEEE Trans. Signal Processing*, vol. 7, no. 1, pp. 8–10, Jan. 2000.

[6] S. M. Kay, J. H. Michels, H. Chen, and P. K. Varshney, “Reducing probability of decision error using stochastic resonance,” *IEEE Signal Processing Lett.*, vol. 13, no. 11, pp. 695–698, Nov. 2006.

[7] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, “Theory of the stochastic resonance effect in signal detection: Part I-Fixed detectors,” *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.

[8] A. Wald, *Sequential Analysis*, Dover, New York, 1947.

[9] R. G. Gallager, *Discrete Stochastic Processes*, Kluwer Academic Publisher, 1996.

[10] S. Tantaratana and J. B. Thomas, “On sequential sign detection of a constant signal,” *IEEE Trans. Inform. Theory*, vol. 23, no. 3, pp. 304–315, May 1977.

[11] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*, Springer-Verlag, 1987.