KYBERNETIKA — VOLUME 14 (1978), NUMBER 3

On the Evaluation of Properties of the Sequential Probability Ratio Test for Statistically Dependent Observations

IVAN VRANA, MOHAMED MAHMOUD EL-HEFNAWI

The paper deals with the method for the evaluation of performance parameters of the sequential probability ratio test (SPRT) when this test is applied to statistically dependent observations. The effective Monte-Carlo simulation method is proposed for this purpose. Some basic properties of the SPRT are shown when used to the detection of a signal with a random phase in a coloured Gaussian noise environment. These results were obtained by the proposed method.

1. INTRODUCTION

The sequential probability ratio test (SPRT) is the optimum sequential test in many decision problems [1]. From the practical point of view it is sometimes useful to know the properties of the SPRT also for decision problems in which the optimum sequential test is not yet known. Such a problem is e.g. a radar decision whether the target is present or absent. In this paper we shall show an effective method of evaluating the performance of the SPRT for a decision problem of the detection of a signal with a random phase in a coloured Gaussian noise environment. We shall also introduce some numerical results obtained by this method. We shall deal only with the case when the test will be terminated with the unit probability (for both hypotheses) after the finite number of sequential steps, i.e. with the case which has a practical importance.

Let us assume the decision problem

$$H_k : Y_m = N_m + kS_m(\varphi)\,, \quad k = 0, 1\,. \tag{1}$$

Let us denote by X^T the transpose of matrix X and m will be called the number of observations. Then the column complex vectors Y_m, N_m and $S_m(\varphi)$ have the following interpretation

$$Y_m = (y_1, y_2, \ldots, y_m)^T = ((y'_1 + jy''_1), \ldots, (y'_m + jy''_m))^T \tag{2}$$

 is the vector of observations with complex elements $y_i = (y'_i + jy''_i)$.

(3) $$S_m(\varphi) = (s_1(\varphi), \ldots, s_m(\varphi))^T = ((s'_1(\varphi) + js''_1(\varphi)), \ldots, (s'_m(\varphi) + js''_m(\varphi)))^T$$

is the vector of a determined signal with a random parameter φ. This vector has complex elements

(4) $$s_i(\varphi) = s'_i(\varphi) + js''_i(\varphi) = \bar{s}_i e^{j\varphi}$$

and φ is the random variable uniformly distributed in the interval $\langle 0, 2\pi)$.

(5) $$N_m = (n_1, \ldots, n_m)^T = ((n'_1 + jn''_1), \ldots, (n'_m + jn''_m))^T$$

is the noise vector with complex elements $n_i = n'_i + jn''_i$. We shall assume that the complex elements of the noise vector (the complex samples of noise) have the following properties for all $i, k \in \langle 1, m \rangle$:

(6) $$\begin{aligned} \mathrm{E}(n'_i) &= \mathrm{E}(n''_i) = 0 ; \\ \mathrm{E}(n'_i n''_k) &= 0 , \\ \mathrm{E}(n'_i n'_k) &= \mathrm{E}(n''_i n''_k) = c_{ik} ; \end{aligned}$$

E denotes an averaging operator, c_{ik} are elements of the positive-definite covariance matrix C_m of the vector $\mathrm{Re}\,(N^T_m) = (n'_1, \ldots, n'_m)$ and of the vector $\mathrm{Im}\,(N^T_m) = (n''_1, \ldots, n''_m)$ respectively. The matrix C_m is of the order m and it has real elements. It follows from (6) that

(7) $$\begin{aligned} \mathrm{E}(N_m) &= 0 \\ \mathrm{E}(N_m N^T_m) &= 0 \\ \mathrm{E}(N_m N^*_m) &= 2C_m \end{aligned}$$

where the symbol N^*_m denotes conjugate transpose of the matrix N_m. The term quadrature components of the complex number n_i will denote its real and imaginary parts n'_i and n''_i. The joint probability density function of the quadrature components of the noise vector $w_m(n'_1, \ldots, n'_m, n''_1, \ldots, n''_m)$ is Gaussian and we shall formally denote it by $w_m(N_m)$. Then

(8) $$w_m(N_m) = \frac{1}{(2\pi)^m \det C_m} \exp\left(-\tfrac{1}{2} N^*_m C^{-1}_m N_m\right)$$

where C^{-1}_m is an inverse of the matrix C_m and $\det C_m$ denotes the determinant of C_m. Analogically, we shall denote the joint probability density function of quadrature components of the observation vector $w_m(y'_1, \ldots, y'_m, y''_1, \ldots, y''_m)$ by $w_m(Y_m)$.

The sequential probability ratio test (SPRT) calculates the likelihood ratio

$$\Lambda(\boldsymbol{Y}_m) = \frac{\mathrm{E}_{\varphi}(w_m(\boldsymbol{Y}_m \mid H_1))}{w_m(\boldsymbol{Y}_m \mid H_0)} \tag{9}$$

sequentially for $m = 1, 2, \ldots$. The test is either terminated for the given m by accepting one of the hypotheses

$$\begin{aligned} &H_0 \quad \text{if} \quad \Lambda(\boldsymbol{Y}_m) \leqq B \\ &H_1 \quad \text{if} \quad \Lambda(\boldsymbol{Y}_m) \geqq A \end{aligned} \tag{10}$$

or the test continues in observation by increasing m by 1 if $B < \Lambda(Y) < A$. A, B are constant upper and lower thresholds of the SPRT, E_{φ} denotes an averaging operator over the random parameter φ, $w_m(\boldsymbol{Y}_m \mid H_K)$ are conditional distributions of the observation vector under hypothesis H_K.

Four performance parameters characterize the performance of the SPRT. They are probability of type 1 and type 2 errors and the average numbers of observations under both hypotheses. Let us denote

$$Z_m = \ln \Lambda(\boldsymbol{Y}_m) \tag{11}$$

and let

$$\gamma \equiv (\ln B,\ \ln A) \equiv (B', A')$$

be an open interval. Then the probability of accepting H_1 when H_K is true will be

$$\begin{aligned} P_K(H_1) = P_K(Z_1 \geqq A') + P_K(Z_1 \in \gamma)\, P_K(Z_2 \geqq A \mid Z_1 \in \gamma) + \ldots \\ \ldots + P_K(Z_1 \in \gamma)\, P_K(Z_2 \in \gamma \mid Z_1 \in \gamma) \times \ldots \\ \ldots \times P_K(Z_{n-1} \in \gamma \mid Z_1 \in \gamma, \ldots, Z_{n-2} \in \gamma)\, P_K(Z_n \geqq A' \mid Z_1 \in \gamma, \ldots Z_{n-1} \in \gamma) + \ldots \end{aligned} \tag{12}$$

Probabilities of errors are evidently $P_0(H_1)$ and $P_1(H_0)$. The probability of false alarm $P_f = P_0(H_1)$ and the probability of correct detection $P_d = P_1(H_1) = 1 - P_1(H_0)$. The average test length is given by equations

$$\mathrm{E}(m \mid H_K) = \sum_{i=1}^{\infty} i P_K(m = i) \tag{13}$$

where the probability density of m is given by the equation

$$P_K(m = i) = P_K(Z_i \notin \gamma,\ Z_{i-1} \in \gamma, \ldots, Z_1 \in \gamma)\,. \tag{14}$$

Substituting into (13) and after arrangement we obtain

$$\mathrm{E}(m \mid H_K) = P_K(Z_1 \notin \gamma) + \sum_{i=2}^{\infty} i P_K(Z_i \notin \gamma \mid Z_{i-1} \in \gamma, \ldots, Z_1 \in \gamma) \times \tag{15}$$

$$\times P_K(Z_{i-1} \in \gamma \mid Z_{i-2} \in \gamma, \ldots, Z_1 \in \gamma) \times \ldots$$
$$\ldots \times P_K(Z_2 \in \gamma \mid Z_1 \in \gamma) P_K(Z_1 \in \gamma) .$$

Equations (12) and (15) give us general relations for the calculation of all performance parameters of the SPRT. Wald [2] introduces the following equation for the average test length

$$\text{(16)} \qquad \mathrm{E}(m \mid H_K) = \frac{P_K(H_1) A' + (1 - P_K(H_1)) B'}{\mathrm{E}\left(\ln \frac{w_1(y \mid H_1)}{w_1(y \mid H_0)} \middle| H_K \right)}$$

where $w_1(y \mid H_K)$ denotes the conditional joint probability density function $w(y_i', y_i'' \mid H_K)$ of one pair of quadrature components y_i', y_i''. Since (16) is valid only for independent and identically distributed observations, i.e. if it holds,

$$\text{(17)} \qquad Z_m = \sum_{i=1}^{m} \ln \frac{w_1(y_i \mid H_1)}{w_1(y_i \mid H_0)} = m \ln \frac{w_1(y \mid H_1)}{w_1(y \mid H_0)}$$

it is clear that for our decision problem (16) cannot be used and it is necessary to use the general equation (15).

It is necessary to integrate the joint probability density functions of $Z_1, Z_2, \ldots, Z_m$ over the corresponding complex domains of these events for calculation of probabilities of events appearing in (15). Unfortunately it is very difficult to express the m-th joint probability density function of $Z_1, Z_2, \ldots, Z_m$ in the closed form and hence it is difficult to obtain the above mentioned probabilities. But it is possible to evaluate performance of the SPRT with the aid of the Monte-Carlo simulation method. This method gives us an unbiased estimate [3] for all four performance parameters describing properties of the SPRT. These estimates are random variables and we can control the accuracy of the method by guaranteeing that the true performance parameters will lie, with the given probability, inside the confidence interval of the chosen width. The Monte-Carlo simulation method usually requires a large amount of calculations. This implies high requirements for an operational speed of the computer used and high requirements for the computer time consumption. Further, we shall show how to simplify significantly the Monte Carlo simulation method in our case. Using this method we shall determine some basic properties of the SPRT for the case when the SPRT is applied to radar observations.

2. SIMPLIFICATION OF THE MONTE-CARLO SIMULATION METHOD

The Monte-Carlo simulation method will require the simulation of the whole hypothesis testing problem, i.e. to simulate both signal and noise processes and to simulate the sequential processor. Increasing the accuracy of the estimated perfor-

mance parameters will require an increasing number of times of performing the test.
For the proposed method to be effective (from the point of view of the computer time consumption) it is necessary to minimize the number of operations as much as possible.

The likelihood ratio $\Lambda(\boldsymbol{Y}_m)$, according to (1), (8) and (9), will be for our case

$$(18)\quad \Lambda(\boldsymbol{Y}_m) = \mathrm{E}_\varphi(\exp(\tfrac{1}{2}\boldsymbol{Y}_m^*\boldsymbol{C}_m^{-1}\boldsymbol{Y}_m - \tfrac{1}{2}((\boldsymbol{Y}_m - \boldsymbol{S}_m(\varphi))^*\,\boldsymbol{C}_m^{-1}(\boldsymbol{Y}_m - \boldsymbol{S}_m(\varphi)))))$$
$$= \mathrm{E}_\varphi(\exp(-\tfrac{1}{2}\,\boldsymbol{S}_m^*(\varphi)\,\boldsymbol{C}_m^{-1}\,\boldsymbol{S}_m(\varphi) + \tfrac{1}{2}\,\boldsymbol{S}_m^*(\varphi)\,\boldsymbol{C}_m^{-1}\boldsymbol{Y}_m + \tfrac{1}{2}\boldsymbol{Y}_m^*\boldsymbol{C}_m^{-1}\,\boldsymbol{S}_m(\varphi)))\,.$$

Let

$$(19)\qquad \bar{\boldsymbol{S}}_m = \boldsymbol{S}_m(\varphi)\,.\,\mathrm{e}^{-\mathrm{j}\varphi}$$

be a complex vector which, according to (4), has elements $\bar{s}_i$ not depending on φ. Then the likelihood ratio (18) can be arranged to the form

$$(20)\quad \Lambda(\boldsymbol{Y}_m) = \exp(-\tfrac{1}{2}\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m)\,\mathrm{E}_\varphi(\exp(\tfrac{1}{2}(\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\boldsymbol{Y}_m\mathrm{e}^{-\mathrm{j}\varphi} + \boldsymbol{Y}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m\mathrm{e}^{\mathrm{j}\varphi})))$$

and by averaging over φ and by taking the logarithm of both sides of the equation we shall get

$$(21)\qquad Z_m = \ln\Lambda(\boldsymbol{Y}_m) = -\tfrac{1}{2}\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m + \ln\mathrm{I}_0(|\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\boldsymbol{Y}_m|)$$

where I_0 is the modified Bessel function of the zero order and the symbol $|x|$ denotes the modulus of a complex variable x.

The SPRT processor compares the value Z_m to the two constant thresholds $A' = \ln A$ and $B' = \ln B$. Taking into account that the term $\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m$ is a constant for the given m and the $\ln\mathrm{I}_0$ is a strictly monotonic function, then the inequalities (10) are evidently equivalent to the inequalities

$$(22)\qquad \begin{aligned} |Q_m| &\leqq \mathscr{B}_m \quad \text{accept } H_0\\ |Q_m| &\geqq \mathscr{A}_m \quad \text{accept } H_1\\ \mathscr{B}_m < |Q_m| &< \mathscr{A}_m \quad \text{continue the test}\end{aligned}$$

where

$$(23)\qquad |Q_m| = |\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\boldsymbol{Y}_m|$$

and $\mathscr{B}_m$, $\mathscr{A}_m$ are modified lower and upper thresholds, respectively, given as follows

$$(24)\qquad \begin{aligned}\mathscr{B}_m &= (\ln\mathrm{I}_0(\ln B + \tfrac{1}{2}\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m))^{-1}\,,\\ \mathscr{A}_m &= (\ln\mathrm{I}_0(\ln A + \tfrac{1}{2}\bar{\boldsymbol{S}}_m^*\boldsymbol{C}_m^{-1}\bar{\boldsymbol{S}}_m))^{-1}\,,\end{aligned}$$

$(\ln\mathrm{I}_0)^{-1}$ is the inverse function of $\ln\mathrm{I}_0$.

Because the thresholds $\mathscr{A}_m$, $\mathscr{B}_m$ are constants for a given m (they do not depend on the observed signal), it is possible to calculate them only once for each m regardless on the number of repetitions of the test. Thus the SPRT processor operating according to (22) will calculate the function $(\ln I_0)^{-1}$ only twice for each m, whereas the processor according to (10) will calculate the function $\ln I_0$ for the given m each time the test is repeated when the test length is not less than m. In most cases, due to this fact, the processor operating according to (22) has a significant advantage compared to the processor operating according to (10).

For further reduction in the number of operations, it is very important to arrange the equation (23) for $|Q_m|$ in such a manner, which simplifies the simulation of both, signal and noise processes. $|Q_m|$ should be simulated as follows

$$(25)\qquad |Q_m| = |\bar{\mathbf{S}}_m^* \mathbf{C}_m^{-1}(\mathbf{N}_m + K e^{j\varphi} \bar{\mathbf{S}}_m)| = |\bar{\mathbf{S}}_m^* \mathbf{C}_m^{-1} \mathbf{N}_m + K e^{j\varphi} \bar{\mathbf{S}}_m^* \mathbf{C}_m^{-1} \bar{\mathbf{S}}_m| \; ; \quad K = 0 \; ; 1$$

where K denotes the true hypothesis. It is seen that $|Q_m|$ has two parts; due to signal and noise components of observation. In the Monte-Carlo method we can separately simulate both the signal and noise processes and then calculate $|Q_m|$ directly using (25). But by a simple arrangement and by factorisation of the matrix $\mathbf{C}_m^{-1}$ we can reach a considerable simplification.

Since the covariance matrix $\mathbf{C}_m$ and its inverse $\mathbf{C}_m^{-1}$ are both positive-definite, we can factorize them as a product of a lower triangular matrix and its conjugate transpose, i.e.

$$(26)\qquad \mathbf{C}_m = \mathbf{G}_m \mathbf{G}_m^* ,$$

$$(27)\qquad \mathbf{C}_m^{-1} = \mathbf{D}_m^* \mathbf{D}_m$$

where

$$(28)\qquad \mathbf{D}_m = \mathbf{G}_m^{-1} .$$

Let us multiply Q_m by $e^{-j\varphi}$, take the modulus of the resultant and use (27). Then

$$(29)\qquad |Q_m| = |(\mathbf{D}_m \bar{\mathbf{S}}_m)^* (\mathbf{D}_m \mathbf{N}_m e^{-j\varphi} + K \mathbf{D}_m \bar{\mathbf{S}}_m)|$$

Let us denote

$$(30)\qquad {}^1\mathbf{N}_m = \mathbf{D}_m \mathbf{N}_m .$$

The quadrature components of the vector ${}^1\mathbf{N}_m$ have again evidently the Gaussian probability density function $w_m({}^1\mathbf{N}_m)$ and they have the following properties

$$(31)\qquad \mathrm{E}({}^1\mathbf{N}_m) = \mathrm{E}(\mathbf{D}_m \mathbf{N}_m) = \mathbf{D}_m \, \mathrm{E}(\mathbf{N}_m) = 0 ,$$

$$(32)\qquad \mathrm{E}({}^1\mathbf{N}_m {}^1\mathbf{N}_m^T) = \mathrm{E}(\mathbf{D}_m \mathbf{N}_m \mathbf{N}_m^T \mathbf{D}_m^T) = \mathbf{D}_m \, \mathrm{E}(\mathbf{N}_m \mathbf{N}_m^T) \, \mathbf{D}_m^T = 0 ,$$

$$(33)\qquad \mathrm{E}({}^1\mathbf{N}_m {}^1\mathbf{N}_m^*) = \mathrm{E}(\mathbf{D}_m \mathbf{N}_m \mathbf{N}_m^* \mathbf{D}_m^*) = \mathbf{D}_m \, \mathrm{E}(\mathbf{N}_m \mathbf{N}_m^*) \, \mathbf{D}_m^* = 2\mathbf{D}_m \mathbf{C}_m \mathbf{D}_m^* = 2\mathbf{D}_m \mathbf{G}_m \mathbf{G}_m^* \mathbf{D}_m^* = 2\mathbf{I}_m ,$$

where $\boldsymbol{I}_m$ is the unit matrix of the order m. Then

$$w_m({}^1\boldsymbol{N}_m) = \frac{1}{(2\pi)^m} \exp\left(-\tfrac{1}{2}{}^1\boldsymbol{N}_m^{*1}\boldsymbol{N}_m\right). \tag{34}$$

Quadrature components of the elements of vector ${}^1\boldsymbol{N}_m$ are thus independent and they are zero-mean. Let

$$ {}^2\boldsymbol{N}_m = {}^1\boldsymbol{N}_m e^{-j\varphi} = \boldsymbol{D}_m \boldsymbol{N}_m e^{-j\varphi} \tag{35}$$

Transformation (35) is identical to that defined by (A.2) in Appendix and the random variable φ is statistically independent on the elements of a vector ${}^1\boldsymbol{N}_m$ and it has the probability density function

$$\begin{aligned} w(\varphi) &= \frac{1}{2\pi} \quad \text{for} \quad \varphi \in \langle 0, 2\pi), \\ &= 0 \quad \text{for} \quad \varphi \notin \langle 0, 2\pi). \end{aligned} \tag{36}$$

Then according to the theorem, proven in Appendix, it holds

$$w_m({}^2\boldsymbol{N}_m) = w_m({}^1\boldsymbol{N}_m). \tag{37}$$

The vector ${}^2\boldsymbol{N}_m$ has thus elements with independent quadrature components, which are zero-mean and have the Gaussian joint probability density function.

Equation (29) will be rewritten to the form

$$|Q_m| = |\mathbf{D}_m^*({}^2N_m + K\mathbf{D}_m)| \tag{38}$$

where

$$\mathbf{D}_m = \boldsymbol{D}_m \bar{\boldsymbol{S}}_m \tag{39}$$

is the column vector representing the signal part of $|Q_m|$. Vector $\mathbf{D}_m$ coincides with the vector $\mathbf{D}_{m+1}$ in all the first columns. From (38) it follows that during simulation it is simply sufficient to generate two mutually independent white Gaussian sequences with zero means and unit variances and to use them as the quadrature components of the vector ${}^2\boldsymbol{N}_m$, instead of the simulation of the coloured Gaussian complex process with the covariance matrix of the quadrature components $\boldsymbol{C}_m$.

The new arrangement of $|Q_m|$, as given by (38), will suppress the necessity for simulating the signal process. Since the only random variable in the signal process is φ and φ does not appear in the signal part of $|Q_m|$ as given by (38), then there is no need for generating the random variable φ and the signal part $\mathbf{D}_m$ should be calculated only once for the given m during the whole simulation runs of the SPRT.

Thus it is clearly seen that arrangement (38) of $|Q_m|$ will simplify the generation of both the signal and the noise processes and will simplify the calculation of $|Q_m|$ compared to the direct method via (25).

Note. It is seen that the SPRT, designed for the detection of the signal with the random phase in the Gaussian noise environment, has identical statistical properties also for the detection of the completly known signal in the Gaussian noise environment. In other words, the probabilities of errors and the average test lengths are the same for the both above mentioned cases.

3. MISMATCHED OBSERVATIONS

In technical applications both signal and noise are described by some set of parameters ψ, which define their concrete properties. Such parameters are e.g. the amplitude, the rate of phase changes, the variance, the spectrum width, etc. Let the assumed signal be defined by a set of nonrandom parameters ϑ_S and the assumed noise be defined by a set of nonrandom parameters ϑ_N. Let the processor be designed according to the values of parameters of the above mentioned sets. These values of parameters will be called the design parameters. The values of parameters of the actually observed signal and the accompanying noise will be called the actual parameters. Actual parameters can generally differ from the prespecified design parameters. This deviation of parameters of the actually observed signal from the corresponding design parameters will be called the case of mismatched observations. To differentiate between the actual and design parameters we shall denote the former by the symbol $\hat{\psi}$. i.e. $\hat{\psi} = (\hat{\vartheta}_S, \hat{\vartheta}_N)$.

From the point of view of technical applications it is important to know the performance of the SPRT also in the case of mismatched observations, i.e. to know the probabilities of errors and the average test lengths under both hypotheses. Equations (12) and (15), given in Introduction for the probabilities of errors and the average test lengths, hold also for the case of mismatched observations, but the statistics of the log-likelihood ratios $Z_1, Z_2, \ldots, Z_m$ will differ from that of matched case. For mismatched case the joint probability density function of the random variables $Z_1, Z_2, \ldots, Z_m$ will be also calculated from the probability density function of the observation $y_1, y_2, \ldots, y_m$, but with actual parameters $\hat{\psi}$ replacing the design parameters ψ. This results in a different joint probability density function of $Z_1, Z_2, \ldots, Z_m$, compared to the matched case. The direct calculation of the performance parameters of the SPRT in a mismatched case is practically impossible for the same reasons like those for the matched case. But these performance parameters could be again evaluated by the Monte-Carlo simulation method described in the preceding section. Further we shall show, how to simplify this method also for the case of mismatched observations.

The mismatched signal parameters can be easily introduced into (29) and (38) by assigning values $\hat{\vartheta}_S$ different from ϑ_S. Then

$$(40) \qquad |Q_m| = |(\boldsymbol{D}_m \bar{\boldsymbol{S}}_m(\vartheta_S))^* ({}^2\boldsymbol{N}_m + K\boldsymbol{D}_m\bar{\boldsymbol{S}}_m(\hat{\vartheta}_S))| = \\ = |(\mathscr{D}_m(\vartheta_S))^* ({}^2\boldsymbol{N}_m + K\,\mathscr{D}_m(\hat{\vartheta}_S))|$$

where evidently we understand $\boldsymbol{D}_m = \boldsymbol{D}_m(\vartheta_N)$, $\mathscr{D}_m(\vartheta_S) = \mathscr{D}_m(\vartheta_N, \vartheta_S)$ and $\mathscr{D}_m(\hat{\vartheta}_S) = \mathscr{D}_m(\vartheta_N, \hat{\vartheta}_S)$. The matched case is a special case of (40) if $\hat{\vartheta}_S = \vartheta_S$.

Equation (40) should be slightly changed for introducing mismatched noise parameters. Let $\boldsymbol{R}_m$ be the positive-definite covariance matrix of the actual coloured noise (of its quadrature components) and $\boldsymbol{R}_m^{-1}$ be the inverse of $\boldsymbol{R}_m$. Then (33) will not be valid, because $\boldsymbol{C}_m$ is not the covariance matrix of $\boldsymbol{N}_m$ yet and thus the quadrature components of ${}^1\boldsymbol{N}_m$ are not yet independent. To transform the vector $\boldsymbol{N}_m$ to a vector with independent quadrature components, let us arrange (25) as follows

(41) $$|Q_m| = |\bar{\boldsymbol{S}}_m^*(\vartheta_S)\, \boldsymbol{D}_m^* \boldsymbol{D}_m \boldsymbol{R}_m \boldsymbol{R}_m^{-1}(\boldsymbol{N}_m + K e^{j\varphi}\, \bar{\boldsymbol{S}}_m(\hat{\vartheta}_m))|$$

By factorizing the matrix $\boldsymbol{R}_m^{-1}$ to a product of triangular matrices $\boldsymbol{W}_m^*$ and $\boldsymbol{W}_m$ and by defining the square matrix

(42) $$\boldsymbol{L}_m = \boldsymbol{D}_m \boldsymbol{R}_m \boldsymbol{W}_m^*$$

and introducing the noise vector

(43) $${}^3\boldsymbol{N}_m = \boldsymbol{W}_m \boldsymbol{N}_m e^{-j\varphi}$$

we shall obtain, after an arrangement

(44) $$|Q_m| = |(\mathscr{D}_m(\vartheta_S))^* (\boldsymbol{L}_m {}^3\boldsymbol{N}_m + K\, \mathscr{D}_m(\hat{\vartheta}_S))| \,.$$

In equations (42) through (44) we shall evidently understand that $\boldsymbol{R}_m = \boldsymbol{R}_m(\hat{\vartheta}_N)$, $\boldsymbol{L}_m = \boldsymbol{L}_m(\vartheta_N, \hat{\vartheta}_N)$. When the noise parameters are matched, i.e. $\boldsymbol{R}_m = \boldsymbol{C}_m$, then $\boldsymbol{L}_m = \boldsymbol{I}_m$.

Analogically like in the previous section, we can show that the noise vector ${}^3\boldsymbol{N}_m$ has independent quadrature components with zero mean and with a Gaussian joint probability density function.

Performance of the SPRT for the case of mismatched noise parameters could be evaluated merely by introducing the matrix $\boldsymbol{L}_m$ into the simulation process. The advantage of this arrangement is that the matrix $\boldsymbol{L}_{m+1}$ coincides with the matrix $\boldsymbol{L}_m$ in the first m rows and m columns. Thus the matrix $\boldsymbol{L}_m$ is needed to calculate only once regardless on the number of the test repetitions. This evidently leads to decreasing the number of operations and memory requirements during simulation compared to the direct calculation via equation (41).

4. APPLICATION OF THE SPRT TO THE DETECTION OF A RADAR SIGNAL

In this section we shall show an example, when our decision problem (1) will be the detection of a radar signal. For this concrete decision problem we shall show what form the equations used for simulation have and further we shall show some basic properties of the SPRT when used for detection of a radar signal.

In the case of a radar signal we shall deal with the case where the complex elements of the signal vector $S_m(\varphi)$ have the form

$$s_i(\varphi) = X \exp(j(iF + \varphi)) \tag{45}$$

and the elements of the covariance matrix of the noise quadrature components are

$$c_{ik} = \delta_{ik} + \lambda \exp(-\Omega(i-k)^2) \tag{46}$$

where X is the amplitude and F is the phase modulation of the signal, λ is a variance and Ω is a positive constant characterizing the spectrum width of the coloured noise component and δ_{ik} is the Cronecker delta. In this case obviously

$$\vartheta_S \equiv (X, F); \quad \vartheta_N \equiv (\lambda, \Omega). \tag{47}$$

Let us denote

$$\bar{S}'_m(F) = \frac{1}{X} \bar{S}_m(F, X). \tag{48}$$

Then equation (44) will have the form

$$|Q_m| = X|(D_m \bar{S}'_m(F))^* (L_m{}^3 N_m + K\hat{X} D_m \bar{S}_m(\hat{F}))|. \tag{49}$$

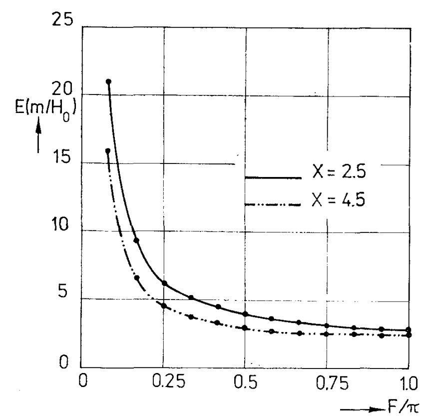

Fig. 1. The dependence of the average test length $E(m | H_0)$ on the phase modulation F. Matched case.
$P_f = 10^{-3}$, $P_d = 0.9$, $\Omega = 5.3 \cdot 10^{-4}$, $\lambda = 10^4$.

For illustration we shall further introduce some typical properties of the SPRT, evaluated by the described Monte-Carlo simulation method. More detailed results are in [3]. For each set of design and actual parameters the simulation consisted of 1000 runs when evaluating the average test lengths $E(m \mid H_0)$ and $E(m \mid H_1)$ and it consisted of 50 000 runs when evaluating both the actual probability of detec-

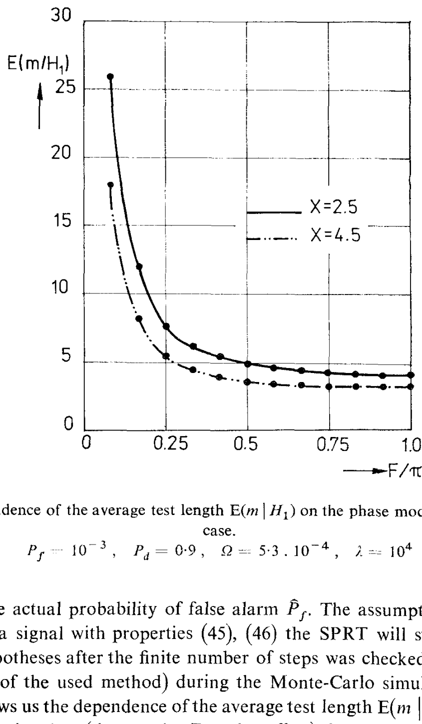

Fig. 2. The dependence of the average test length $E(m \mid H_1)$ on the phase modulation F. Matched case.

$$P_f = 10^{-3}, \quad P_d = 0{\cdot}9, \quad \Omega = 5{\cdot}3 \,.\, 10^{-4}, \quad \lambda = 10^4$$

tion $\hat{P}_d$ and the actual probability of false alarm $\hat{P}_f$. The assumptions that for the given type of a signal with properties (45), (46) the SPRT will stop by accepting one of two hypotheses after the finite number of steps was checked (with the statistical accuracy of the used method) during the Monte-Carlo simulations.

Figure 1 shows us the dependence of the average test length $E(m \mid H_0)$ on the phase modulation of signal F (due to the Doppler effect) for the matched case. Since $E(m \mid H_K)_F = E(m \mid H_K)_{2\pi - F}$, Fig. 1 shows only a part of these characteristics for $F \in \langle 0, \pi \rangle$. Analogically Fig. 2 shows us the dependence of the $E(m \mid H_1)$ on phase modulation F for the matched case.

The basic property of the SPRT is the fact that a total probability of an erroneous decision will not exceed the prespecified value. Thus even for very bad conditions like small phase modulation F or small amplitude X, the SPRT will achieve the desired error probabilities if it is made matched to the above mentioned signal conditions. But, as it is seen from Figs. 1 and 2, the small probabilities of error are paid by a relatively large average test length in these cases.

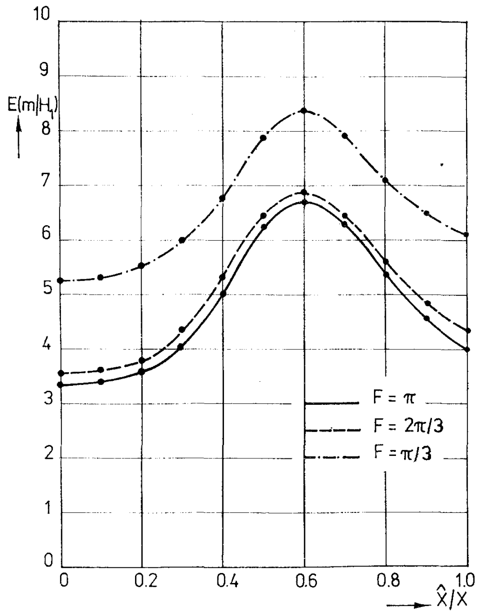

Fig. 3. The average test length $E(m | H_1)$ versus the actual signal amplitude $\hat{X}$. Matched noise parameters $\hat{\vartheta}_N = \vartheta_N$.
$P_f = 10^{-3}$, $P_d = 0{\cdot}9$, $X = 2{\cdot}5$, $\hat{F} = F$, $\Omega = 5{\cdot}3 \,.\, 10^{-4}$, $\lambda = 10^4$.

Mismatching of the signal parameters ϑ_S results in a change of both the probability of detection $\hat{P}_d$ and the average test length $E(m \mid H_1)$ while both values $\hat{P}_f$ and $E(m \mid H_0)$ remain unchanged. Figure 3 shows us an influence of mismatching of the amplitude X to the average test length $E(m \mid H_1)$. It is seen from these curves that their tops lie roughly near the value $\hat{X} = X/2$. The curves have the shape similar to those introduced in [4] for the case of the SPRT used for a completely known signal in a white Gaussian noise.

Figure 4 shows us, what influence has mismatching of the signal amplitude X on the actual probability of detection $\hat{P}_d$. The influence of mismatching the phase modulation F on the actual probability of detection $\hat{P}_d$ is shown in Fig. 5. This dependence is called the speed characteristics of the SPRT processor. For comparison purposes, the characteristic which corresponds to the dependence $\hat{P}_d$ on $\hat{F} = F$ in the matched case is also drawn.

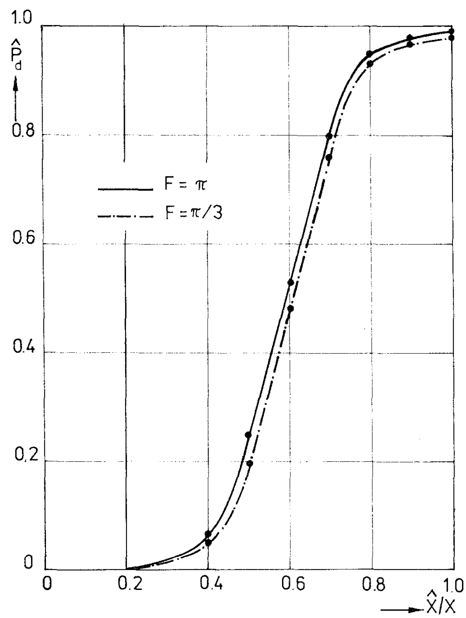

Fig. 4. The actual probability of detection $\hat{P}_d$ versus the actual signal amplitude $\hat{X}$. Matched noise parameters $\hat{\vartheta}_N = \vartheta_N$.

$P_f = 10^{-3}$, $P_d = 0{\cdot}9$, $X = 2{\cdot}5$, $\hat{F} = F$, $\Omega = 5{\cdot}3 \,.\, 10^{-4}$, $\lambda = 10^4$.

The influence of mismatching of the noise parameters λ, Ω can be briefly summarized as follows. The processor SPRT operates quite well under such noise conditions, if $\hat{\lambda} \leqq \lambda$, $\hat{\Omega} \leqq \Omega$. But if either $\hat{\lambda} > \lambda$ or $\hat{\Omega} > \Omega$ (i.e. when values of actual parameters exceed the values of their corresponding design parameters), then the performance of the SPRT will quickly deteriorate. From the class of the SPRT processors satisfying the design requirments on their error probabilities, the processor which is matched to the actual noise parameters will reach the minimum average test length.

For the radar signal the work [3] compares the performance of the optimum fixed-length test (the likelihood-ratio test [5]) to the performance of the SPRT.

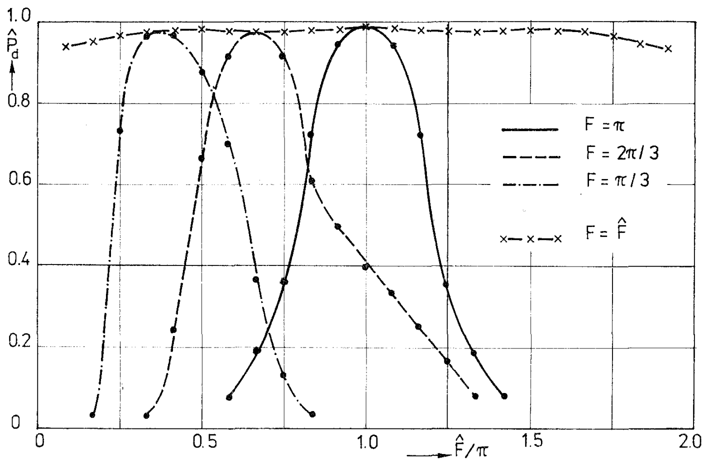

Fig. 5. The speed characteristics of the SPRT. Matched noise parameters $\hat{\vartheta}_N = \vartheta_N$. $P_f = 10^{-3}$, $P_d = 0{\cdot}9$, $X = 2{\cdot}5$, $\Omega = 5{\cdot}3 \cdot 10^{-4}$, $\lambda = 10^4$.

This comparison covers a wide range of changes of parameters ϑ_S, ϑ_N. The results of these comparisons can be summarized as follows. To reach the same probabilities of errors and the same average test lengths $E(m \mid H_1)$ by both tests, the SPRT requires a 0·8 times smaller value of the signal amplitude $\hat{X}$ than the likelihood-ratio test. Under the same values of the signal amplitude and the same error probabilities, the test length of the likelihood-ratio test is approximately 1·4 times longer than $E(m \mid H_1)$ of the SPRT and 1·8 times longer than $E(m \mid H_0)$ of the SPRT. Thus, it is clear that the SPRT is more advantageous than the optimum fixed-length test from all the assumed points of view.

Theorem. Let $U = (\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n, \beta_n, \varphi)$ be the random vector with real elements and with the joint probability density function of its elements

$$(A.1) \qquad w(U) = w(\alpha_1, \beta_1, \ldots, \alpha_n, \beta_n, \varphi) = \frac{1}{(2\pi)^{n+1}} \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(\alpha_i^2 + \beta_i^2)\right).$$

Further, let $V = (a_1, b_1, a_2, b_2, \ldots, a_n, b_n, \varphi)$ be a random vector with real elements a_i, b_i, φ which originated from the elements $\alpha_i, \beta_i, \varphi$ of the vector U by the functional transform

$$(A.2) \qquad \begin{aligned} a_k &= \alpha_k \cos\varphi - \beta_k \sin\varphi , \\ b_k &= \alpha_k \sin\varphi + \beta_k \cos\varphi \end{aligned}$$

for all $k \in \langle 1, n \rangle$. Then the joint probability density function of the elements of the vector V is

$$(A.3) \qquad W(V) = W(a_1, b_1, \ldots, a_n, b_n, \varphi) = \frac{1}{(2\pi)^{n+1}} \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(a_i^2 + b_i^2)\right)$$

Proof. We can easy show that the vector V originated from the vector U by the one-one functional transform $V = f(U)$. It holds evidently for all $k \in \langle 1, n \rangle$

$$(A.4) \qquad \begin{aligned} a_k &= f_1(\alpha_k, \beta_k, \varphi) , \\ b_k &= f_2(\alpha_k, \beta_k, \varphi) \end{aligned}$$

and

$$(A.5) \qquad \begin{aligned} \alpha_k &= g_1(a_k, b_k, \varphi) , \\ \beta_k &= g_2(a_k, b_k, \varphi) , \end{aligned}$$

where functions g_1, g_2 are given by equations

$$(A.6) \qquad \begin{aligned} \alpha_k &= a_k \cos\varphi + b_k \sin\varphi , \\ \beta_k &= -a_k \sin\varphi + b_k \cos\varphi . \end{aligned}$$

Then for the joint probability density function $W(V)$ it holds

$$(A.7) \qquad W(V) = (\det J)\, w(U)$$

 where J is a Jacobian of the transform $U = f^{-1}(V)$. Evidently

(A.8) $$J = \frac{\partial(U)}{\partial(V)} = \begin{bmatrix} N_1 & \emptyset & \emptyset & \dots & \emptyset & M_1 \\ \emptyset & N_2 & \emptyset & & \emptyset & M_2 \\ \emptyset & \emptyset & N_3 & & \emptyset & M_3 \\ \vdots & & & \ddots & & \vdots \\ \emptyset & \emptyset & \emptyset & & N_{n-1} & M_{n-1} \\ \emptyset' & \emptyset' & \emptyset' & \dots & \emptyset' & N' \end{bmatrix}.$$

where $\emptyset$ are square 2×2 fields with zero elements, $\emptyset'$ are 3×2 fields with zero elements and, for all $k \in \langle 1, n - 1 \rangle$, N_k are the square 2×2 fields given as follows

(A.9) $$N_k = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} = N$$

and M_k are some 2×3 fields. The field N' has the structure

(A.10) $$N' = \begin{bmatrix} \cos\varphi & \sin\varphi & l \\ -\sin\varphi & \cos\varphi & k \\ 0 & 0 & 1 \end{bmatrix}$$

where l and k are some functions of a_n, b_n, φ. The matrix J is then divided into fields with the square fields in the main diagonal. Its determinant can be easily evaluated by applying a generalized elimination method [6] and it is

(A.11) $$\det J = \det N_1 \det N_{2,1} \det N_{3,2} \dots \det N_{(n-1),(n-2)} \det N'.$$

Due to the structure of the matrix J according to (A.8), where the first column consists only of zero fields except of the first row, the leading fields are

(A.12) $$N_{k,(k-1)} = N_k = N.$$

Then

(A.13) $$\det J = (\det N)^{n-1} \det N'.$$

Since $\det N = \cos^2\varphi + \sin^2\varphi = 1$ and $\det N' = 1$, then

(A.14) $$\det J = 1.$$

Using (A.6) we can show that

(A.15) $$\alpha_k^2 + \beta_k^2 = a_k^2 + b_k^2.$$

Substituting (A.14), (A.1) and (A.15) into (A.7) we get

(A.16) $$W(V) = w(V).$$

The theorem is proven.

(Received December 1, 1977.)

[1] J. Cochlar, I. Vrana: On the optimum sequential test of two hypotheses for statistically dependent observations. Kybernetika *14* (1978), 1, 57—69.

[2] A. Wald: Sequential analysis. Wiley, New York 1947.

[3] M. M. El-Hefnawi: Sequential processing of radar signals in coloured noise environment. Candidate disertation work, VAAZ, Brno 1977.

[4] J. V. DiFranco and W. L. Rubin: Radar detection. Prentice-Hall, Englewood Cliffs, N. J., 1968.

[5] I. Vrana: On the optimum decision rule for the radar signal processing. Kybernetika *10* (1974), 3, 258—271.

[6] D. K. Faddějev, V. N. Faddějevová: Numerické metody lineární algebry, SNTL, Praha 1964.

Ing. Ivan Vrana, CSc., elektrotechnická fakulta ČVUT, katedra radioelektronických zařízení a soustav (Czech Technical University — Department of radioelectronic devices and systems), Suchbátarova 2, 166 27 Praha 6. Czechoslovakia.

PhDr. Mohamed Mahmoud El-Hefnawi, 4 Emara El-Yamani st. Zamalek, Cairo. Egypt.