Radio Eng. Electron. Phys., 19(7), p??, 1974.
[transl. from: Radiotekhnika i Electronika, 1974, 19(7): 1391-1395.]

3. Ya. M. Dorfman, Radiotekhnika i Elektronika, 1972, 17, No. 11, 2423 [Radio Eng. Electron. Phys., 17, No. 11 (1972)].
4. L. S. Benenson (Ed.), Antennye reshetki (Antenna Arrays), Sovetskoye Radio Press, 1966.
5. E. S. Wentzel, Probability Theory (in Russian transl.), Nauka Press, 1969.
6. Ya. N. Feld and L. S. Benenson, Antenna-fidernye ustroistva, Ch. II (Antenna Feeders), N. E. Zhukovskii VVIA, 1959.

Submitted June 5, 1973

□♦♦♦□

Sufficient Conditions of Optimality of Processing Quantum Signals

V. P. BELAVKIN AND A. G. VANTSYAN

Sufficient conditions of optimality of reception of quantum signals are derived from both Bayes and information criteria. It is shown that for a linear channel with Gaussian boson noise these conditions are satisfied by coherent measurement of the received superposition.

* * *

1. NECESSARY AND SUFFICIENT CONDITIONS OF OPTIMALITY

In [1, 2], dealing with optimization of the reception of quantum signals, the search for the necessary conditions of optimality in the class of randomized strategies based on indirect measurements was the main concern. According to the universal approach discussed in these studies and used for detection and discrimination of nonorthogonal quantum signals of arbitrary nature (in particular, in processing electromagnetic signals in the optical or any other range), we shall specify the randomized strategies by the operator probability measures $\Pi(d\beta)$ ($\Pi(d\beta) \geqslant 0$, $\int \Pi(d\beta) = \hat{1}$). The equations derived in [1, 2] have the same form for both Bayes and information criteria of optimality:

$$(R(\beta)-\Lambda)\Pi(d\beta)=0, \quad \Lambda=\int R(\beta)\Pi(d\beta), \tag{1}$$

where $R(\beta) = \int C(\beta, \vartheta)\rho(\vartheta)P(d\vartheta)$ is the a posteriori risk operator. Here $\rho(\vartheta)$ is the family of density operators describing the state of the quantum channel depending on the transmitted information ϑ with a priori distribution $P(d\vartheta)$ and $C(\beta, \vartheta)$ is a given penalty function in Bayes case and random information (with opposite sign)

$$C(\beta,\vartheta)=-\ln\frac{p(\beta|\vartheta)}{\int p(\beta|\vartheta)P(d\vartheta)}, \quad p(\beta|\vartheta)=\mathrm{Tr}\,\Pi(\beta)\rho(\vartheta) \tag{2}$$

for optimization from the criterion of maximum Shannon information content

$$\mathcal{I}_{\beta\vartheta}=\left\langle \ln\frac{p(\beta|\vartheta)}{\int p(\beta|\vartheta)P(d\vartheta)}\right\rangle.$$

It is obvious that operators $\Pi^0(d\beta)$ satisfying Eq. (1) are degenerate (provided $R(\beta) - \Lambda \neq 0$) and for each β have a range of values belonging to the zero eigensubspace of the difference $R(\beta) - \Lambda$. If operators $B(\beta) = R(\beta) - \Lambda$ have a unique eigenvector φ_β for each β, corresponding to the zero eigenvalue, then the operator measure $\Pi^0(d\beta)$ is proportional to the projection operators $\Pi^0(d\beta)=\varphi_\beta^0\varphi_\beta^{0*}d\beta$. In the general case when degeneration of the zero eigenvalue of operators $B(\beta)$ is possible: $B(\beta)\varphi_{\beta\nu}=0$, $\nu\in N(\beta)$, each expansion of unity $\int\Pi(d\beta) = \hat{1}$ satisfying Eq. (1) may be included in some more detailed expansion $\int \varphi_\gamma\varphi_\gamma^* d\gamma=\hat{1}$, $\gamma=(\beta, \nu)$. For different γ the vectors φ_γ describing the "elementary" measurements need not necessarily be orthogonal: $\varphi_\gamma^*\varphi_{\gamma'}\neq 0$.

In Bayes case the sufficient conditions for optimality are very simple: the operators $\Pi^0(d\beta)$ satisfying Eq. (1) minimize the average risk

$$R = \langle C(\beta,\vartheta)\rangle = \mathrm{Tr}\int R(\beta)\,\Pi(d\beta) \tag{3}$$

if and only if the condition of nonnegative definiteness

$$B(\beta) \equiv R(\beta) - \int R(\beta')\,\Pi^0(d\beta') \geqslant 0 \tag{4}$$

is satisfied for all β. Actually, for any other operator measure $\Pi(d\beta) \neq \Pi^0(d\beta)$ the difference $R - R^0 = \mathrm{Tr}[\int R(\beta)\Pi(d\beta) - \int R(\beta)\Pi^0(d\beta)] = \mathrm{Tr}\int B(\beta)\Pi(d\beta)$ is nonnegative since it is the trace of a sum of products of nonnegative operators $B(\beta)$, $\Pi(d\beta)$.

Conditions (1) and (4) are applicable also for the optimization of the processing of quantum signals according to the maximum likelihood criterion. For this it is sufficient to consider that this criterion can be formally taken as Bayes criterion with uniform (unnormalized) a priori distribution $P(d\vartheta) = d\vartheta$ and a simple penalty function $C(\beta, \vartheta) = -\delta(\vartheta - \beta)$. This means that the a posteriori risk operator $R(\beta)$ should in this case be replaced by the density operator $\rho(\vartheta)$ at the estimate point $\vartheta = \beta$. We shall call the quantum strategies $\Pi^0(d\beta)$ satisfying conditions (1) and (4) for $R(\beta) = -\rho(\beta)$ optimum with respect to the maximum likelihood criterion. We shall give the solution of the problem of the discrimination of nonorthogonal signals for the following simplest case.

In [3] the concept of coherent processing of a boson* signal $b = \{b_\nu\}$, i.e., of indirect linear measurement realized by measuring the superposition $b + a_0^*$, where a_0 is vacuum boson noise, was introduced. The question of the physical realization of this measurement was discussed at the Third All-Union Conference on the Physical Principles of Information Transmission by Laser Radiation. In particular it was shown [4] that coherent measurement of a narrowband optical signal can be realized by using an ideal heterodyne reception (ideal count of photons at different points of superposition of the received and the reference waves). The backward vacuum wave radiated by an ideally matched receiver into the communication line plays the role of noise a_0.

In [3] the quality of such processing was also defined from the maximum likelihood criterion for the case where b is the superposition of a coherent signal $\vartheta = \{\vartheta_\nu\}$ and a Gaussian boson noise a. The use of Eqs. (1) and (4) and a suitable representation of the density operator makes it obvious that the processing described by the coherent projectors

$$\Pi(d\beta) = |\beta\rangle\langle\beta|\,d\mu(\beta), \qquad d\mu(\beta) = \prod_\nu \frac{1}{\pi}\,d\,\mathrm{Re}\,\beta_\nu\,d\,\mathrm{Im}\,\beta_\nu \tag{5}$$

is optimum. Such a suitable representation of the density operator $\rho(\vartheta)$ of the displaced Gaussian state $b = \vartheta + a$ is the representation in the form of the expression

$$\rho(\vartheta) = |L|^{-1} \colon \exp\{-(b-\vartheta)^+ L^{-1}(b-\vartheta)\} \colon \tag{6}$$

normally ordered with respect to the operators b^*, b. Here $L\|\langle a_\nu a_{\nu'}^*\rangle\|$ is the correlation matrix of the noise a; the colon :: denotes that the operators a* operate to the left of the operators a; $|L| = \det L$.

Putting $R(\beta) = -\rho(\beta)$, $\Lambda = -|L^{-1}|\hat{1}$, and considering the well-known [5] properties $\int |\beta\rangle\langle\beta|\,d\mu(\beta) = 1$, $:p(b^*, b): |\beta\rangle = p(b^*, \beta)\,|\beta\rangle$ of coherent vectors, we at once find that Eq. (1) has a unique solution coinciding with (5). The nonnegative definiteness of the operator

$$B(\beta) = |L^{-1}|\,(1 - \colon \exp\{-(b-\vartheta)^+ L^{-1}(b-\vartheta)\}\colon)$$

is beyond doubt.

We recall that we are giving the name boson signal to the quantum signal described by the operators $\{a_\nu, a_\nu^*\}$ satisfying the commutation relations $a_\nu a_{\nu'} - a_{\nu'} a_\nu = 0$, $a_\nu a_{\nu'}^* - a_{\nu'}^* a_\nu = \delta_{\nu\nu'}$. In particular, the optical signal is described by the photon annihilation and creation operators a and a respectively.

If the signals $\{\vartheta_i\}$ are known apart from the phase, the coherent processing is no longer optimum; however, it remains quasioptimum if the dimensionality of ϑ is large.

2. SUFFICIENT CONDITIONS OF OPTIMALITY ACCORDING TO INFORMATION CRITERION

In the case of the information criterion it is not possible to indicate a global criterion of optimality of the solutions $\Pi^0(d\beta)$ of Eq. (1) minimizing the "Shannon" risk $R = -\mathcal{I}_{\beta\vartheta}$ in view of its nonlinear dependence on $\Pi(d\beta)$ [the "penalty function" (2) also depends on $\Pi(d\beta)$]. Therefore we give a differential criterion of optimality.

We shall restrict the discussion to operator probabilities of degenerate form $\Pi(d\beta) = \varphi_\beta\varphi_\beta^* d\beta$, where $\{\varphi_\beta\}$ is the complete family of vectors $\int \varphi_\beta\varphi_\beta^* d\beta = \hat{1}$. It can be shown that this is sufficient for the verification of local optimality $\delta^2 R > 0$ of the degenerate solutions $\Pi^0(d\beta) = \varphi_\beta^0\varphi_\beta^{0*} d\beta$ of Eq. (1).

Making use of the dependence of the variations $\delta\varphi_\beta = \varphi_\beta - \varphi_\beta^0$

$$\int (\varphi_\beta^0 \delta\varphi_\beta^* + \delta\varphi_\beta\varphi_\beta^{0*} + \delta\varphi_\beta\delta\varphi_\beta^*)\, d\beta = 0,$$

it is not difficult to find the increment $\Delta R = R - R^0$ of Shannon risk (3) (with the penalty function (2) depending on φ_β) at the stationary point φ_β^0 with an accuracy up to second-order terms in $\delta\varphi_\beta$:

$$\Delta R \simeq \iint \Big\{ C(\beta,\vartheta)\,\delta p(\beta|\vartheta) P(d\vartheta) - \tag{7}$$
$$- \frac{1}{2}[(\delta \ln p(\beta|\vartheta))^2 - (\delta \ln p(\beta))^2] p(\beta|\vartheta) P(d\vartheta) \Big\} =$$
$$= \int \Big\{ \delta\varphi_\beta^* B(\beta)\,\delta\varphi_\beta - \frac{1}{2}\Big[\int (\delta\varphi_\beta^*\psi_\beta(\vartheta) + \psi_\beta^*(\vartheta)\,\delta\varphi_\beta)^2 p(\beta|\vartheta) P(d\vartheta) -$$
$$- (\delta\varphi_\beta^*\psi_\beta + \psi_\beta^*\delta\varphi_\beta)^2 p(\beta) \Big]\Big\} d\beta,$$

where $B(\beta)$ is the difference (4),

$$\psi_\beta(\vartheta) = \frac{\rho(\vartheta)\varphi_\beta^0}{p(\beta|\vartheta)}, \quad p(\beta|\vartheta) = \varphi_\beta^{0*}\rho(\vartheta)\varphi_\beta^0,$$

and the vector $\psi_\beta = \rho\varphi_\beta^0/p(\beta)$ $(p(\beta) = \int p(\beta|\vartheta)P(d\vartheta))$ is the vector $\psi_\beta(\vartheta)$ averaged with the conditional density $p(\vartheta|\beta) = p(\beta|\vartheta)p(\vartheta)/\int p(\beta|\vartheta)p(d\vartheta)$.

A simple analysis of the positiveness $\delta^2 R > 0$ of the variation (7) of Shannon risk shows that in contrast to the Bayes case the nonnegativeness of the operators $B(\beta) \geqslant 0$ is necessary, but not sufficient for the local optimality of the solutions φ_β^0 of the equation $B(\beta)\varphi_\beta = 0$: the additional term in (7) (in square brackets) has the meaning of a posteriori variance of the real random quantity $2\mathrm{Re}\,\delta\varphi_\beta^*\psi_\beta(\vartheta)$ and is generally positive.

Let, for example, the density operator have the form (6) and the a priori distribution $p(d\vartheta)$ be Gaussian in the multidimensional space of the information parameters $\vartheta = \{\vartheta_\nu\}$:

$$P(d\vartheta) = |S|^{-1} \exp\{-\vartheta^+ S^{-1}\vartheta\} d\mu(\vartheta), \quad d\mu(\vartheta) = \prod_\nu \frac{1}{\pi} d\,\mathrm{Re}\,\vartheta_\nu d\,\mathrm{Im}\,\vartheta_\nu.$$

We shall check the local optimality of the coherent solutions (5) of Eq. (1) in the boson Gaussian case according to the information criterion. As shown in [2], the coherent vectors (5) satisfy Eq. (1) and the operator $B(\beta)$ has a quadratic Gaussian form:

$$B(\beta) = (b-\beta)^+ H\rho(b-\beta), \ \rho = |L+S|^{-1} : e^{-b^+(L+S)^{-1}b} :,$$

where the matrix $H = L^{-1} - (S + L)^{-1}$ is not larger than unity in accordance with the inequalities $S \geqslant 0$, $L \geqslant 1$, $0 \leqslant H \leqslant 1$. Considering the analytic dependence of the function

$$\psi_\beta(\vartheta) \equiv \frac{\rho(\vartheta)|\beta\rangle}{\langle\beta|\rho(\vartheta)|\beta\rangle} = e^{-(b-\beta)^+ L^{-1}(\beta-\vartheta)}|\beta\rangle$$

on ϑ (i.e., the lack of dependence on ϑ^*) and carrying out conditional averaging over ϑ in (7) (i.e., the integration $\int (\delta\varphi_\beta^* \psi_\beta(\vartheta) + \psi_\beta^*(\vartheta)\delta\varphi_\beta)^2 p(\vartheta|\beta)\, d\mu(\vartheta)$ with density $p(\vartheta|\beta) = |M| \exp\{-(\vartheta - A\beta)^+ M(\vartheta - A\beta)\}$, $A = S(S + L)^{-1}$, $M = S^{-1} + L^{-1}$), we find that in the Gaussian case the variation $\delta^2 R$ has the form

$$\delta^2 R = \int \delta\varphi_\beta^* (B(\beta) - D(\beta))\,\delta\varphi_\beta\, d\mu(\beta),$$
$$d\mu(\beta) = \prod_\nu \frac{1}{\pi}\, d\,\mathrm{Re}\,\beta_\nu d\,\mathrm{Im}\,\beta_\nu,$$

where

$$D(\beta) = :p(b)\,[e^{-(b-\beta)^+(1-H)(b-\beta)} - e^{-(b-\beta)^+(b-\beta)}] : \geqslant 0,$$
$$p(b) = |S+L|^{-1} \exp\{-b^+(S+L)^{-1}b\}.$$

Thus, in order to prove the optimality of coherent measurement from the information criterion one should verify the operator inequality $B(\beta) - D(\beta) \geqslant 0$ or

$$p(\beta)\, e^{-b^+(S+L)^{-1}\beta} : [b^+ H e^{-b^+(S+L)^{-1}b} b - (e^{-b^+(H-1)b} - e^{-b^+b})] : e^{-\beta^+(S+L)^{-1}b} \geqslant 0 \quad (8)$$

(here the change $b - \beta \to b$ has been carried out). The operator occurring on the left-hand side of (8) has the structure $A_\beta^* : [\cdot] : A_\beta$ and is positive only if the operator in the square brackets is positive. Considering that the operator inequality $e^{-b^+(S+L)^{-1}b} : \geqslant : e^{-b^+(1-H)b}:$ is satisfied by virtue of the matrix inequality $(S + L)^{-1} = L^{-1} H \leqslant 1 - H$, we find that inequality (8) is satisfied if

$$: b^+ H e^{b^+(H-1)b} b : \geqslant : e^{b^+(H-1)b} - e^{-b^+b} : \quad (9)$$

Inequality (9) becomes obvious in the diagonal representation in the occupation numbers n_ν:

$$\prod_\nu h_\nu^{n_\nu} \sum_\nu n_\nu \geqslant \prod_\nu h_\nu^{n_\nu} \quad \text{for} \quad \sum_\nu n_\nu \neq 0,$$

where h_ν are the eigenvalues of the matrix $H = L^{-1} - (S + L)^{-1}$. For $\sum_\nu n_\nu = 0$ both the left- and the right-hand sides of inequality (9) vanish.

REFERENCES

1. R. L. Stratonovich, The quantum generalization of optimal statistical estimation and hypothesis testing, J. of Stochastics, 1973, 1, 87-126.
2. V. P. Belavkin and R. L. Stratonovich, Radiotekhnika i Elektronika, 1973, 19, 9, 1839 [Radio Eng. Electron. Phys., 19, 9 (1973)].
3. V. P. Belavkin, Radiotekhnika i Elektronika, 1972, 17, No. 12, 2533 [Radio Eng. Electron. Phys. 17, No. 12 (1972)].
4. V. P. Belavkin, Coherent measurement of optical signals, Proc. of the Third All-Union Conference on the Physical Principles of Recording and Processing Information by Laser Radiation, Kiev, 1973, I, p. 7.
5. J. Klauder and E. Sudershan, Fundamentals of Quantum Optics (in Russian transl.), Mir Press, 1970.

Submitted July 6, 1973

□♦♦♦□