Radio Eng. Electron. Phys., 19(9), p1349, 1973.
[trans. from: Radiotekhnika i Electronika, 1973, 19, 9, 1839-844]

Optimization of Processing of Quantum Signals according to an Information Criterion*

V. P. BELAVKIN AND R. L. STRATONOVICH

An equation is derived which determines the "excess" set of wave functions giving optimum decoding, according to the criterion of maximum of the amount of information, based on quasi-measurements. A solution of this equation in the Gaussian multimode case is given. A method of realization of the obtained maximum amount of information $J = \mathrm{Sp}\ln[1 + S/(N + 1)]$ is given.

* * *

INTRODUCTION

The use of power or phase-sensitive receivers for the reception of a coherently modulated quantized signal enables one to extract only about one half of the encoded information. The coherent reception of quantized signals, based on simultaneous direct measurement of amplitude and phase (or phase coordinates of each mode of the signal), is forbidden by the postulates of the quantum theory of measurements. Therefore, it is of interest to investigate the potential possibilities of decoding in quantized communication channels based on some indirect methods of measurement.

1. DECODING BASED ON QUASI-MEASUREMENT

The theory of "quasi-measurements of noncompatible observables" has been under development starting from the studies [1-3] to meet the demands of the quantum communication theory. Such measurements with values in a given (measurable) space $B \ni \beta$ are described by operator (nonnegative definite, denumerably additive) measure $\Pi(d\beta)$ normalized to unit operator 1 in Hilbert space H of quantum states:

$$\int \Pi(d\beta) = 1. \tag{1}$$

If the expansion of unity (1) is orthogonal; $\Pi(d\beta)\Pi(d\beta') = \Pi(d\beta \cap d\beta')$, then it describes the usual (direct) measurement of compatible observables of the form $b_j = \int \beta_j \Pi(d\beta)$. In the general case, however, the operators $\{b_j\}$ do not commute and it can be said that the nonorthogonal expansion (1) describes simultaneous quasi-measurement of the observables corresponding to it. The probabilistic operator measure $\Pi(d\beta)$ is an analog of classical randomized decision rules $P(d\beta | b)$.

The probability distribution $P(d\beta)$ of the results of quasi-measurements in the state specified by the statistical operator ρ is determined by the measure $\Pi(d\beta)$ in the same way as in the case of direct measurements:

$$P(d\beta) = \mathrm{Sp}\,\Pi(d\beta)\rho. \tag{2}$$

* Originally published in: ↑

Below we shall specify the operator $\Pi(d\beta)$ in the form

$$\Pi(d\beta) = \varphi_\beta \varphi_\beta^+ \mu(d\beta), \tag{3}$$

explicitly taking account of its nonnegative definite character and σ-additiveness. Here $\{\varphi_\beta, \beta \in B\}$ is a set of non-Hermitian operators operating from some space into the original space H and, together with the numerical measure $\mu(d\beta)$ on the set B, satisfy the condition of normalization and completeness (1). In particular, if $\{\varphi_\beta\}$ are operators operating from a unidimensional complex space into H, i.e., are a set of wave functions $\varphi_\beta \in H$ (in general nonorthogonal and unnormalized: $\varphi_\beta^+\varphi_{\beta'} \neq \delta_{\beta\beta'}$), then the quasi-measurement specified by them is called elementary. We note that any quasi-measurement can be realized by direct measurement of a certain set of compatible observables $\beta = \{\beta_i\}$ in an expanded quantum system consisting of the original system and an auxiliary system independent of the first. This follows from the fact that any operator measure $\Pi(d\beta)$ can be represented as partial averaging $\Pi(d\beta) = \mathrm{Sp}_{H_0}\, \Theta(d\beta)\rho_0$ of eigenorthoprojectors $\Theta(d\beta)$ of commuting operators $\{\beta_i\}$, operating into the expanded space $H \otimes H_0$ along the state ρ_0 of the additional system chosen in a certain manner, described by Hilbert space H_0. Such a realization of quasi-measurement is called indirect measurement.

The use of the concept of indirect measurement has made it possible to indicate for a Gaussian semiquantized communication channel the amount of information

$$J = \ln\left(1 + \frac{s}{n+1}\right),$$

transmitted by each mode of coherently modulated signal in using a quantized receiver with a large linear amplifier* (s, $\bar{n}$ are the energies of the corresponding modes of the classical useful signal and of the quantized additive noise expressed in units of $h\nu$). The same amount of information was deduced in [4], where a method of its decoding was indicated, leading to indirect measurement. The rate of information transmission by the coherent signal was also determined in an earlier work [5], where vectors of the φ_β coherent states were used as the operators characterizing the coherent reception. The problem of the realization of the coherent measurement and its optimality was not considered. However, it is easy to show that an ideal quasi-measurement specified by the coherent vectors

$$\varphi_\alpha = \exp\{-{}^1\!/_2\alpha^+\alpha + a^+\alpha\}\,|0\rangle = |\alpha\rangle, \tag{4}$$

can be realized by direct simultaneous measurement of the complex amplitudes

$$\alpha = a + a_0^+ = \{a_j + a_{0j}^+,\ j = 1, \ldots, r\}, \tag{5}$$

where $a = \{a_j\}$ are annihilation operators operating into the original Hilbert space H, while $a_0^+ = \{a_{0j}^+\}$ are creation operators of the additional system occurring in the vacuum state $|0\rangle_0 \in H_0$. For this purpose it is sufficient to note that the partial averaging over the vacuum state $|0\rangle_0$ amounts to the determination of the projections φ_α of the eigenvectors $\psi_\alpha \in H \otimes H_0$ of commuting operators $\alpha = a + a_0^+$ in the original space H. It is not difficult to verify that the vectors

$$\psi_\alpha = \exp\{-(a-\alpha)^+ a_0^+\}\,|0\rangle_0 \otimes |\alpha\rangle \tag{6}$$

satisfy the equations $\alpha\psi_\alpha = \alpha\psi_\alpha$, $\alpha^+\psi_\alpha = \alpha^*\psi_\alpha$, and form a complete orthonormalized set with the measure

$$d\mu(\alpha) = \prod_{j=1}^{r} \frac{1}{\pi}\, d\,\mathrm{Re}\,\alpha_j\, d\,\mathrm{Im}\,\alpha_j.$$

Multiplying these vectors from the left by ${}_0\langle 0|$ we find that the projections $\varphi_\alpha \in H$ are actually the coherent vectors $|\alpha\rangle$ and, together with the measure $\mu(d\alpha) = d\mu(\alpha)$, determine ideal measurement corresponding to the indirect measurement of noncommuting a.

---

*This result was presented by R. L. Stratonovich at the symposium on information theory (Dubna, June 15-25, 1969).

The question—whether the decoding based on the quasi-measurement described by the coherent vectors (4) is optimum and if so in what conditions—is of interest. In order to answer this question we derive the equation determining the operators φ_β, which are optimum according to the criterion of the maximum amount of information and specify the optimum quasi-measurement.

We shall assume that the family of statistical operators $\{\rho(\vartheta)\}$ defining the state of the quantized communication channel is given as a function of the random information parameters $\vartheta = \{\vartheta_j\}$ having the distribution $P(d\vartheta)$. The amount of Shannon information of the parameters ϑ and the results of quasi-measurements β is determined by the usual formula

$$J_{\vartheta\beta} = \iint \ln \frac{p(\beta|\vartheta)}{p(\beta|\vartheta)P(d\vartheta)} p(\beta|\vartheta)P(d\vartheta)\mu(d\beta),$$

where according to (2), (3) the density $p(\beta|\vartheta)$ has the form

$$p(\beta|\vartheta)\,\mathrm{Sp}\,\varphi_\beta\varphi_\beta^+\rho(\vartheta) = \mathrm{Sp}\,\varphi_\beta^+\rho(\vartheta)\varphi_\beta \tag{7}$$

and satisfies the normalization condition

$$\int p(\beta|\vartheta)\mu(d\beta) = 1.$$

Using the method of Lagrange we set up a function

$$\int\left(\int J(\beta,\vartheta)\,\mathrm{Sp}\,\varphi_\beta^+\rho(\vartheta)\varphi_\beta P(d\vartheta) - \mathrm{Sp}\,\varphi_\beta^+\lambda\varphi_\beta\right)\mu(d\beta),$$

where $J(\beta,\vartheta) = \ln[p(\beta|\vartheta)/\int p(\beta|\vartheta)P(d\vartheta)]$, and λ is an operator determined from the condition of completeness $\int \varphi_\beta\varphi_\beta^+\mu(d\beta) = 1$. Varying it over φ_β^+, it is not difficult to obtain the equation determining optimum φ_β:

$$\left(\int J(\beta,\vartheta)\rho(\vartheta)P(d\vartheta) - \lambda\right)\varphi_\beta = 0. \tag{8}$$

In the derivation of this equation we have made use of the fact that

$$\int p(\beta|\vartheta)\,\delta J(\beta,\vartheta)P(d\vartheta) = \int p(\beta|\vartheta)\,\delta\varphi_\beta^+\left(\rho(\vartheta)/p(\beta|\vartheta) - \right.$$
$$\left. - \int \rho(\vartheta)P(d\vartheta) \Big/ \int p(\beta|\vartheta)P(d\vartheta)\right)\varphi_\beta P(d\vartheta) = 0.$$

Multiplying Eq. (8) from the right by φ_β^+ and integrating with the measure $\mu(d\beta)$ we determine the operator λ:

$$\lambda = \iint \rho(\vartheta)J(\beta,\vartheta)\varphi_\beta\varphi_\beta^+\mu(d\beta)P(d\vartheta),$$

the trace of which $\mathrm{Sp}\lambda$ gives the maximum amount of information.

And, finally, eliminating the operator from Eq. (8) and multiplying it from the left by λ, we rewrite this equation in the following equivalent (because of the condition of completeness $\int \varphi_\beta\varphi_\beta^+\mu(d\beta) = 1$) form:

$$\int \varphi_\alpha^+\rho(\vartheta)\left(\varphi_\beta J(\beta,\vartheta) - \int \varphi_{\beta'}\varphi_{\beta'}^+\varphi_\beta J(\beta',\vartheta)\mu(d\beta')\right)P(d\vartheta) = 0. \tag{9}$$

The equation thus obtained is a complex nonlinear equation in φ_β and it is not possible to solve explicitly it in the general case. However, as will be shown in the next section, in the Gaussian case the coherent vectors (4) are optimum operators φ_β.

2. OPTIMUM DEQUANTIZATION IN A LINEAR BOSONIAN GAUSSIAN CHANNEL

Let the received signal b be a superposition $b = \vartheta + a$ of a complex vector $\vartheta = \{\vartheta_j,\ j = 1, \ldots, r\}$

and an independent Gaussian Boson noise $a = \{a_j\}$ (a_j are annihilation operators of the investigated modes: $[a_j, a_k^+] = \delta_{jk}$, having zero mathematical expectation $\langle a_j \rangle = 0$ and specified average number of quanta $\langle a_j^+ a_j \rangle = n_j$). In Glauber's representation [6] the statistical operator of this noise has the form

$$\rho = \int |\alpha\rangle\langle\alpha| \, |N|^{-1} \exp\{-\alpha^+ N^{-1}\alpha\} \, d\mu(\alpha), \tag{10}$$

where $|\alpha\rangle$ is the coherent vector (4); N is a matrix whose eigen values are n_j, and $|N| \equiv \det N = n_1 \ldots n_r$. In formula (10), as also in (4), the matrix form of writing scalar products is used: $a^+\alpha = \sum_j a_j^+ = \alpha_j$, $\alpha^+ N^{-1} \alpha = \sum_{j,k} \alpha_j^* (N^{-1})_{jk} \alpha_k$. The signal b after passing through the indicated linear channel is described by a family of operators $\{\rho(\vartheta)\}$ of the form

$$\rho(\vartheta) = \int |\alpha\rangle\langle\alpha| p(\alpha|\vartheta) \, d\mu(\alpha), \tag{11}$$

where $p(\alpha|\vartheta) = |N|^{-1} \exp\{-(\alpha - \vartheta)^+ N^{-1} (\alpha - \vartheta)\}$. The probability distribution $P(d\vartheta)$ of the transmitted signal ϑ, chosen from the condition of maximum entropy and finiteness of the energy $\langle \vartheta^+ S^{-1} \vartheta \rangle \leqslant$ (S is the given correlation matrix $\langle \vartheta_j^* \vartheta_k \rangle$), also has a Gaussian form:

$$P(d\vartheta) = |S|^{-1} \exp\{-\vartheta^+ S^{-1} \vartheta\} \, d\mu(\vartheta), \quad d\mu(\vartheta) = \prod_{j=1}^{r} \frac{1}{\pi} d \operatorname{Re} \vartheta_j d \operatorname{Im} \vartheta_j. \tag{12}$$

In the case under investigation it is natural to expect that the results of optimum quasi-measurement of the Gaussian observable b also have a Gaussian distribution; among the Gaussian operators φ_β, determining this distribution, it is natural to separate out the operators characterized by the minimum ratio of indeterminacies $\langle (b-\beta)_j^+ (b-\beta)_k \rangle = \delta_{jk}$. Such operators are coherent vectors.

Let us verify if vectors φ_β of form (4) satisfy Eq. (9) in the Gaussian case. Substituting (4), (11) into (7) and considering that $\langle \beta | \alpha \rangle = \exp\{\beta^+\alpha - (\alpha^+\alpha + \beta^+\beta)/2\}$ it is not difficult to find the conditional density

$$\begin{aligned} p(\beta|\vartheta) &= \int \langle\beta|\alpha\rangle\langle\alpha|\beta\rangle \, p(\alpha|\vartheta) \, d\mu(\alpha) = \\ &= |N+1|^{-1} \exp\{-(\beta-\vartheta)^+ (N+1)^{-1} (\beta - \vartheta)\}, \end{aligned} \tag{13}$$

and also the function $J(\beta, \vartheta)$ occurring in Eq. (9):

$$J(\beta, \vartheta) = \ln|1 + S(N+1)^{-1}| + \vartheta^+ S^{-1} \vartheta - (\vartheta - A\beta)^+ G (\vartheta - A\beta), \tag{14}$$

where $A = S(S + N + 1)^{-1}$, and $G = S^{-1} + (N + 1)^{-1}$. In (14) only the last term is significant for Eq. (9): the first two terms drop out after substitution into this equation. Substituting the last term $(A\beta - \vartheta)^+ G (A\beta - \vartheta) \equiv c(A\beta, \vartheta)$ of function (14) into Eq. (9) and using representation (11) of the operator $\rho(\vartheta)$, we verify that the coherent vector $\varphi_\beta = |\beta\rangle$ is a solution of Eq. (9). Thus the solution of the formulated problem is reduced, as also in the optimization according to the mean square criterion, to the verification of the identity

$$\begin{aligned} &\iint \langle\alpha|\alpha'\rangle\langle\alpha'|\beta\rangle \left(c(A\beta, \vartheta) - \int \frac{\langle\alpha'|\beta'\rangle\langle\beta'|\beta\rangle}{\langle\alpha'|\beta\rangle} c(A\beta', \vartheta) \, d\mu(\beta') \right) \times \\ &\times p(\alpha'|\vartheta) \, d\mu(\alpha') P(d\vartheta) = 0, \end{aligned} \tag{15}$$

where $p(\alpha|\vartheta)$, $P(d\vartheta)$ are defined in (11), (12).

Below we shall require the following formulas of integration in the complex r-dimensional space $C^r \ni z$:

$$\begin{aligned}
&\int \exp\{-(z-\alpha)^+Q(z-\beta)\}\,|Q|\,d\mu(z)=1,\\
&\int z\exp\{-(z-\alpha)^+Q(z-\beta)\}\,|Q|\,d\mu(z)=\beta,\\
&\int z^+\exp\{-(z-\alpha)^+Q(z-\beta)\}\,|Q|\,d\mu(z)=\alpha^+,\\
&\int (z-\alpha)^+H(z-\beta)\exp\{-(z-\alpha)^+Q(z-\beta)\}\,|Q|\,d\mu(z)=\mathrm{Sp}\,Q^{-1}H,
\end{aligned}\tag{16}$$

these formulas are valid for all vectors α, $\beta \in C^r$ and all positive definite $(r \times r)$-matrices Q, H. Considering that $\langle\alpha'|\beta'\rangle\langle\beta'|\beta\rangle(\langle\alpha'|\beta\rangle)^{-1}=\exp\{-(\beta'-\alpha')^+(\beta'-\beta)\}$, in identity (15) we carry out the integration over after making the substitution $A\beta' = z$ and making use of formula (16) for $Q = (AA^+)^{-1}$, $H = G$. As a result the expression in the parentheses in (15) becomes $(\beta - \alpha')^+A^+G(A\beta - \vartheta)\mathrm{Sp}AA^+G$. The integration of the last expression over α' with the density

$$\langle\alpha|\alpha'\rangle\langle\alpha'|\beta\rangle p(\alpha'|\vartheta) = \langle\alpha|\beta\rangle|N+1|^{-1}|\bar{Q}|\exp\{-(\vartheta-\alpha)^+(N+1)^{-1}(\vartheta-\beta)-(\alpha'-\bar{\alpha})^+\bar{Q}(\alpha'-\bar{\beta})\}$$

amounts to the use of the first and the third formulas of integration (16), where for Q, α, β we should take $\bar{Q} = N^{-1} + 1$, $\bar{\alpha} = (N + 1)^{-1}(\vartheta + N\alpha)$, $\bar{\beta} = (N + 1)^{-1}(\vartheta + N\beta)$. This gives

$$\begin{aligned}
&[(\beta-(N+1)^{-1}(\vartheta+N\alpha))^+A^+G(A\beta-\vartheta)-\mathrm{Sp}\,AA^+G]\langle\alpha|\beta\rangle|\times\\
&\times|N+1|^{-1}\exp\{-\vartheta-\alpha)^+(N+1)^{-1}(\vartheta-\beta)\}.
\end{aligned}\tag{17}$$

In order to complete the verification of identity (15) it only remains to carry out averaging with the distribution $P(d\vartheta)$. Rewriting the expression in the square brackets in (17) in the form

$$(\vartheta-(N+1)\beta+N\alpha)^+(N+1)^{-1}A^+G(\vartheta-A\beta)-\mathrm{Sp}\,AA^+G$$

and integrating it with respect to ϑ with the density

$$\begin{aligned}
&|N+1|^{-1}|S|^{-1}\exp\{-(\vartheta-\alpha)^+(N+1)^{-1}(\vartheta-\beta)-\vartheta^+S^{-1}\vartheta\}=\\
&=|G|\,|S+N+1|e^{-(\vartheta-A\alpha)^+G(\vartheta-A\beta)}
\end{aligned}$$

according to the formulas (16) for $Q = G$, $H = (N + 1)^{-1}A^+G$ we obtain

$$|S+N+1|^{-1}\mathrm{Sp}\,(G^{-1}(N+1)^{-1}A^+-AA^+)G.$$

This expression coincides with the left-hand side of (15) with an accuracy up to a nonzero factor $\langle\alpha|\beta\rangle$. Considering that

$$G^{-1}=(S^{-1}+(N+1)^{-1})^{-1}=S(S+N+1)^{-1}(N+1)=A(N+1),$$

as a consequence of which $\mathrm{Sp}(G^{-1}(N+1)^{-1}A^+ - AA^+)G = \mathrm{Sp}(AA^+ - AA^+)G = 0$, the validity of the verified identity is confirmed.

Thus in a linear Gaussian-Bosonian channel the optimum decoding is given by coherent vectors of the form (4). The ideal quasi-measurement, defined by the complete nonorthogonal (i.e., "excess") set $\{|\beta\rangle\}$ of coherent vectors, is realized by the measurement of complex amplitudes

$$\beta = b + a_0^+ = \vartheta + a + a_0^+ = \vartheta + \alpha,\tag{18}$$

in the expanded space these amplitudes are of a complete orthonormal set of eigen vectors. The measurement of simultaneous observables (18) is a linear indirect measurement of noncommuting $b = \{b_j\}$ investigated in [7]. The maximum amount of information decoded in optimum ideal quasi-measurement is easily found by averaging the function (14): $J = \ln|1 + S(N + 1)^{-1}| = \mathrm{Sp}\ln(1 + S(N + 1)^{-1})$. In the one-dimensional case the obtained amount of information coincides with that shown in [4-5].

In conclusion, we note that if a Gaussian stationary signal transmitted along the channel is mixed with stable thermal noise of temperature $\theta = kT$, then the rate information transmission in optimum decoding is determined by the information $\ln(1 + s_\nu / (n_\nu + 1))$ transmitted by each mode summed over the operating frequency range $N \ni \nu$:

$$J = \int_N \ln(1 + s_\nu (n_\nu + 1)^{-1})\, d\nu = \int_N \ln(1 + s_\nu (1 - e^{-h\nu/\theta}))\, d\nu. \tag{19}$$

Here s_ν is the spectral intensity of the transmitted signal expressed in units of $h\nu$, and $n_\nu = (e^{h\nu/\theta} - 1)^{-1}$ is the spectral intensity of the noise. In the classical limit $h\nu/\theta \ll 1$ the rate of information transmission goes over into the corresponding classical expression $J = \int \ln(1 + \sigma_\nu^2/\theta)\, d\nu$, where $\sigma_\nu^2 = s_\nu h\nu$. In the opposite case $h\nu/\theta \gg 1$ of low temperatures the transmission rate $J = \int \ln(1 + s_\nu)\, d\nu$ remains finite in contrast to the classical case. It is not difficult to see that the quantum corrections to (19) are significant only for weak (at the receiving end of the channel) signals $s_\nu \ll 1$, for which formula (19) can be written in the form $J \simeq \int (1 - e^{-h\nu/\theta})\, s_\nu\, d\nu$.

REFERENCES

1. J. P. Gordon and W. H. Louisell, in Physics of Quantum Electronics, P. L. Kelley et al. (eds.), McGraw-Hill Book Co., New York, 1966.
2. C. Y. She and H. Heffner, Phys. Rev., 1966, 152, 4, 1103.
3. K. Helstron, J. Liu and J. Gordon, TIIE'R, 1970, 58, 10, 186.
4. A. S. Kholevo, Problemy perdachi informatsii (Problems of information transmission), 1972, 8, 1, 62.
5. V. V. Mityugov and V. P. Morozov, Izv. vuzov MVSSO SSSR (Radiofizika), 1968, 11, 2, 260.
6. R. Glauber, in coll. Kvantovaya optika i kvantovaya radiofizika (Quantum optics and quantum radiophysics), Mir, 1965.
7. V. P. Belavkin, Radiotekhnika i Elektronika, 1972, 17, No. 12, 2533 [Radio Eng. Electron. Phys., 17, No. 12 (1972)].

Submitted November 17, 1972

□◆◆◆□

Design of Circuits Containing Lumped-Element Circulators

A. E. PANOV

A method is proposed of designing circuits containing Y-type lumped-element circulators using their rated parameters: wave admittance, operating frequency band and the forward loss at the mean frequency. The design procedure is illustrated by examples of an isolator based on the use of a circulator and a series-connected LCR circuit in the absorbing arm and of an isolator with two Y-type circulators.

* * *

In the design of microwave systems a circulator is usually described by the resistance, admittance of scattering matrix. In order to obtain the matrix of a circulator we have to know the characteristics of the ferromagnetic material, the value of the magnetizing field and other parameters which are often not known. Because of this fact and also in view of the laborious calculations the circulator is usually considered to be perfect. Below we discuss a simple approximate method of designing circuits with actual lumped-element circulators [1].

Analysis [2] shows that the characteristics of a lumped-element circulator which is inserted in an arbitrary section can be obtained from its two complex characteristic admittances Y and Y_{in}. The first admittance when connected to the absorbing arm of an isolator, constructed on the basis of the