OPTIMUM ESTIMATION IN QUANTUM CHANNELS BY
THE GENERALIZED HEISENBERG INEQUALITY METHOD

V. P. Belavkin and B. A. Grishanin

The optimization of indirect quantum measurement at the output of a quantum channel is investigated for a quadratic figure of merit. A lower bound compatible with the Heisenberg uncertainty principle is found for the estimation risk. This lower bound is used in the Gaussian case to establish the linearity of optimum estimation and to give a specific description of an appropriate optimum indirect measurement procedure.

§1. Introduction

In the context of the theory developed in [1, 2] real quantum communication channels are described by the specification of a family $\rho(\vartheta)$ of density matrices determining the state of the channel as a function of certain classical parameters ϑ. In general it is necessary to indicate at the channel output an optimum (by a suitable figure of merit) procedure for the measurement of those parameters, whose distribution function $P(d\vartheta)$ is considered to be known in the Bayes formulation discussed here.

In the interest of somewhat greater generality and continuity of the present article relative to [3, 4] (to which it is essentially a sequel) we replace $\vartheta^T = (\vartheta_1, \ldots, \vartheta_r)$, as the variables to be estimated, with quantum (in general) physical quantities described by Hermitian operators $\mathbf{x}^T = (\mathbf{x}_1, \ldots, \mathbf{x}_r)$ in $\Phi_\mathbf{x}$. Accordingly, instead of the family $\rho(\vartheta)$ and distribution function $P(d\vartheta)$ we discuss their surrogate density matrix ρ, which is an operator on the tensor product $\Phi = \Phi_\mathbf{x} \otimes \Phi_\mathbf{y}$. Here $\Phi_\mathbf{x}$ is a Hilbert space describing the channel input subsystem X, which is inaccessible to the observer and to which the variables x refer, and $\Phi_\mathbf{y}$ is a space describing the output subsystem, which has quantum coordinates y and on which arbitrary quantum measurements can be performed.

The problem of finding commutative operators $\mathbf{u}^T = (\mathbf{u}_1, \ldots, \mathbf{u}_r) = \gamma(\mathbf{y})$ in $\Phi_\mathbf{y}$ such that a direct measurement of the corresponding physical observables will yield a result closest to the variables $\mathbf{x}$ in the sense of the quadratic figure of merit (risk)

$$R = \mathbf{M}(\mathbf{u} - \mathbf{x})^\tau G(\mathbf{u} - \mathbf{x}) \equiv \mathrm{Sp}(\mathbf{u} - \mathbf{x})^\tau G(\mathbf{u} - \mathbf{x})\rho, \qquad (1.1)$$

where G is a real symmetric positive definite (numerical) $r \times r$ matrix, has been investigated in [4]. A similar formulation is given (without analysis) in [5]. Here we consider the more general case of indirect measurement, in which case $\mathbf{u}$ represents operators in the extended space $\Phi_\mathbf{y} \otimes \Phi_\zeta$ describing a primary observable system Y augmented with a certain independent auxiliary quantum system Z with state ρ_ζ in the space Φ_ζ. This generalization corresponds to the inclusion of quantum randomized decision rules $\mathbf{u} = \gamma(\mathbf{y}, \zeta)$ in the class of estimation algorithms. Under such randomization of the measurement process the figure (1.1) is replaced by the more general expression

$$R = \mathrm{Sp}(\mathbf{u} - \mathbf{x})^\tau G(\mathbf{u} - \mathbf{x})\rho \otimes \rho_\zeta, \qquad (1.2)$$

which is to be minimized by the optimal selection of the auxiliary space Φ_ζ, the commutative operators $\mathbf{u}$, and the state ρ_ζ. As a result of this optimization we should obtain better numerical estimation than in [4, 5], and that estimation, in accordance with the postulates of quantum measurement, can be realized by an arbitrarily precise measurement of the commutative operators $\mathbf{u}$. In particular it is expected that the optimum estimation for indirect measurement will go over in the classical limit $\hbar \to 0$ to optimum estimation based on the results of measurements of classical (commutative) observables, a statement

that cannot be made about estimation in direct measurement; for example, as shown in [4], the optimum estimation of the mathematical expectations $x^T = (\vartheta_p, \vartheta_q)$ of the momentum p and coordinate q is degenerate for an infinitely small commutator $[p, q]_-$ and reduces to a measurement of just one optimum combination of observables $y^T = (p, q)$.

§2. Lower Bound for the Mean-Square Estimation Risk

The method used here to minimize the risk (1.2) is the following. Expression (1.2) can be represented in the form of a sum:

$$R(u) = R_0(u_y) + \Delta R(u), \tag{2.1}$$

where

$$u_y = M[u|y] \equiv Sp_t u\rho_t \tag{2.2}$$

is the result of partial averaging of the commutative observables u over the auxiliary variables,

$$R_0(u_y) = Sp(u_y - x)^\tau G(u_y - x)\rho \tag{2.3}$$

represents the "irreducible" risk taken as the primary figure of merit in [3], and

$$\Delta R(u) = Sp(u - u_y)^\tau G(u - u_y)\rho \otimes \rho_t. \tag{2.4}$$

is the mean-square error of indirect measurement of the noncommutative quantum variables u_y as realized by direct measurement of the variables u. The measurement error $\Delta R(u)$ cannot be reduced to zero, because of the noncommutativity of the operators u_y. We can readily obtain for it a lower bound compatible with the generalized Heisenberg inequality

$$K_\zeta \geqslant \pm \tfrac{1}{2}C_\zeta, \tag{2.5}$$

where

$$K_\zeta = \tfrac{1}{2}M[\zeta, \zeta^\tau]_+ \equiv \|\tfrac{1}{2}M(\zeta_i\zeta_j + \zeta_j\zeta_i)\| \tag{2.6}$$

is the correlation matrix of the difference

$$\zeta^\tau = (u - u_y)^\tau = (\zeta_1, \ldots, \zeta_r), \tag{2.7}$$

and C_ζ is its averaged commutator:

$$C_\zeta = M[\zeta, \zeta^\tau]_- \equiv \|M(\zeta_i\zeta_j - \zeta_j\zeta_i)\|. \tag{2.8}$$

The sign $\geqslant$ in (2.5) is understood to mean nonnegative definiteness of the Hermitian matrices $K_\zeta + C_\zeta/2$, $K_\zeta - C_\zeta/2$, which is easily proved.

The use of the Heisenberg inequality (2.5) yields the following inequality for the error (2.4):

$$\Delta R(u) = Sp\,GK_\zeta \geqslant \tfrac{1}{2}\,Sp\,|GC_\zeta|, \tag{2.9}$$

in which $|GC_\zeta|$ is the modulus of the matrix GC_ζ and the symbol Sp, unlike $\mathbf{Sp}$, denotes the trace on non-quantum indices, i.e., the trace of numerical matrices GK_ζ and $|GC_\zeta|$.

For the proof of inequality (2.9) it is sufficient to use only the Hermitian property of the imaginary antisymmetric matrix C_ζ and the real symmetric matrix G, along with the positive definiteness of the latter. Taking into account that the product GC_ζ of any Hermitian matrices G and C_ζ, one of which is positive definite, is similar [6] to the real diagonal matrix $T^{-1}GC_\zeta T = \|c_j\delta_{ij}\|$, we go over to the representation $\tilde\zeta = T^+\zeta$, in which the matrix GC_ζ is diagonal. For definiteness we put $T = \sqrt{G}U$, where U is a unitary matrix whose columns are the eigenvectors of the imaginary Hermitian matrix $\sqrt{G}C_\zeta\sqrt{G}$. Inasmuch as the trace $Sp\,GK_\zeta$ is equal in the given representation to the trace of the correlation matrix $\tilde K \equiv 1/2M[\tilde\zeta, \tilde\zeta^+]_+ = T^+K_\zeta T$ of the complex variables $\tilde\zeta = T^+\zeta$:

$$Sp\,GK_\zeta = Sp\,T^{-1}GT^{+-1}T^+K_\zeta T = Sp\,\tilde K, \tag{2.10}$$

while the commutation matrix $\tilde C \equiv M[\tilde\zeta, \tilde\zeta_+]_- = T^+C_\zeta T$ is diagonal: $\tilde C = \|c_j\delta_{ij}\|$, inequality (2.5) only has to be applied to the diagonal elements k_{jj} of the matrix $\tilde K = \|k_{ij}\|$:

$$Sp\,\tilde K \equiv \sum_j k_{jj} \geqslant \sum_j \max(+\tfrac{1}{2}c_j, -\tfrac{1}{2}c_j) \equiv Sp|\tilde C|. \tag{2.11}$$

In the primary representation $\zeta = T^{-1+}\tilde{\zeta}$ the inequality (2.11) found above assumes, with regard for (2.10), the linearly invariant form (2.9), where the modulus $|GC_\zeta|$ of the product of the matrices G and C_ζ is interpreted according to the definition $f(A) = T \| f(\lambda_j)\delta_{ij} \| T^{-1}$ of the function $f(\cdot)$ of the matrix $A = T \| \lambda_j\delta_{ij} \| \cdot T^{-1}$, which is similar to the real diagonal matrix $\| \lambda_j\delta_{ij} \|$:

$$|GC_\zeta| = T \| |c_j|\delta_{ij} \| T^{-1}.$$

We now take into account the fact that the commutator averaged over the auxiliary variables, $[\zeta, \zeta^T]$, of the difference $\zeta = u - u_y$ coincides (by the commutativity of the operators u) with the commutator $[u_y, u_y^T]$ of the operators (2.2):

$$\mathrm{Sp}_\zeta[\zeta, \zeta^\tau]_{-\rho_\zeta} = [u_y, u_y^\tau]_- - \mathrm{Sp}_\zeta([u_y, u^\tau]_+ + [u, u_y^\tau]_-) = -[u_y, u_y^\tau]_-.$$

This result gives the lower bound

$$\Delta \underline{R}(u_y) = \tfrac{1}{2}\mathrm{Sp}\,|CG| \tag{2.12}$$

for the error (2.4) of indirect measurement of the operators u_y.

Here

$$C = M[u_y, u_y^\tau]_- = \mathrm{Sp}[u_y, u_y^\tau]_-\rho \tag{2.13}$$

is the average commutation matrix of the variables u_y, which determines the minimum possible value (2.12) of the error of indirect measurement of those variables. The lower bound corresponding to the bound (2.12) for the total estimation risk (2.1) has the form

$$\underline{R}(u_y) = R_0(u_y) + \Delta \underline{R}(u_y) = \mathrm{Sp}(KG + \tfrac{1}{2}|CG|), \tag{2.14}$$

where

$$K = \tfrac{1}{2}M[u_y - x, u_y^\tau - x^\tau]_+ = \tfrac{1}{2}\mathrm{Sp}[u_y - x, u_y^\tau - x^\tau]_+\rho \tag{2.15}$$

is the matrix of average deviations of the variables u_y from the estimated variables x.

The bound (2.14) for the figure of merit of the estimation u of the variables x depends only on the operators u_y obtained from u by the averaging operation (2.2) and clearly corresponds to the optimum indirect measurement procedure. The latter is interpreted in the following sense: in the space Φ_y of observable variables certain operator "estimates" are chosen for the variables x, and then those (not necessarily commutative) estimates u_y are subjected to indirect measurement in such a way as to bring $\Delta R(u)$ as close as possible to its lower bound (2.12). The latter cannot be attained in general, and the optimization of the indirect measurement of the selected operator estimates u_y yields losses

$$\min_{u|u_y} R(u) \geqslant \underline{R}(u_y). \tag{2.16}$$

The minimization of the right-hand side of inequality (2.16) with respect to all possible operators u_y in Ψ_y yields the lower bound for the mean-square optimum estimation risk.

We now derive an equation for the determination of the "optimum" operators u_y^0 minimizing the indicated lower bound. Varying the function (2.14) with respect to matrices K and C depending on u_y, we readily obtain

$$\delta \underline{R}(u_y) = \mathrm{Sp}\,G(\delta K + \tfrac{1}{2}(\mathrm{sgn}\,CG)\delta C),$$

where

$$\mathrm{sgn}\,CG = CG / |CG| \equiv (T \| \mathrm{sgn}\,(c_j)\delta_{ij} \| T^{-1})^+ \tag{2.17}$$

(the matrices G and C are presumed to be nondegenerate). Further variation with regard for the forms (2.15) and (2.13) of the matrices K and C yields

$$\delta \underline{R}(u_y) = \mathrm{Sp}\,G\,\mathrm{Sp}[(u_y - x)\rho + \rho(u_y - x) + \mathrm{sgn}\,GC(\rho u_y - u_y\rho)]\delta u_y^\tau. \tag{2.18}$$

Setting the variation (2.18) equal to zero and making use of the independence of the variations δu_y, we deduce the following equation for the optimum selection of u_y^0:

$$(u_y - m_y)\rho_y + \rho_y(u_y - m_y) + \mathrm{sgn}\, CG(\rho_y u_y - u_y \rho_y) = 0, \tag{2.19}$$

where $\rho_y = \mathrm{Sp}_x \rho$ is the partial density matrix describing the state of the subsystem Y and is found by taking the partial trace of the matrix ρ with respect to the space Φ_X, and m_y represents Hermitian operators comprising the solution of the equation

$$m_y \rho_y + \rho_y m_y = \mathrm{Sp}_x(x\rho + \rho x). \tag{2.20}$$

This equation has been derived in [3] for the determination of operator estimates $u_y^0 = m_y$ optimum by the "irreducible" figure (2.3).

Multiplying Eq. (2.19) on the right by the factor u_y^T and on the left by the matrix G and substituting the result into (2.14) with regard for (2.15) and (2.13), we at once obtain the following value of the lower bound for the optimum estimation risk (1.2):

$$\underline{R}^0 = \min_{u_y} \underline{R}(u_y) = \mathrm{Sp}\, G(K_x - K^0_{u_y x}), \tag{2.21}$$

where

$$K_x = \tfrac{1}{2}M[x, x^\tau]_+ = \tfrac{1}{2}M\|x_i x_j + x_j x_i\|,$$
$$K^0_{u_y x} = \tfrac{1}{2}M[u_y^0, x^\tau]_+ = \tfrac{1}{2}M\|u_{yi}^0 x_j + x_j u_{yi}^0\|,$$

and u_y^0 is determined from Eq. (2.19).

Equation (2.19) found for the determination of the optimum operators u_y^0 realizing the lower bound (2.21) is nonlinear, because it includes the commutation matrix C, which is equal to the average commutator (2.13) of the operators u_y. It is linearized, however, by the assumption of weak noncommutativity of the operators ρ_y and m_y:

$$[m_y, \rho_y]_- = m_y \rho_y - \rho_y m_y = \frac{\varepsilon}{i}v_y + o(\varepsilon), \tag{2.22}$$

where ε is a small parameter and v_y represents unknown Hermitian operators. Now the indicated equation (2.19) has a unique asymptotic solution

$$u_y^0 = m_y - \varepsilon\, \mathrm{sgn}\,(S_m G)d_y + o(\varepsilon), \tag{2.23}$$

in which m_y is the solution of Eq. (2.20); d_y is the solution of the equation

$$d_y \rho_y + \rho_y d_y = v_y = \lim_{\varepsilon \to 0}\frac{i}{\varepsilon}[m_y, \rho_y]_-, \tag{2.24}$$

and S_m is the antisymmetric Hermitian matrix

$$S_m = \lim_{\varepsilon \to 0}\frac{1}{2\varepsilon}M[m_y, m_y^\tau]_-, \tag{2.25}$$

which is presumed to be nondegenerate.

The lower bound (2.21) of the optimum estimation risk is reached only if it is possible to state an optimum procedure for indirect measurement of the extremal operators u_y^0 determined by the solution of Eq. (2.19), which gives a measurement error (2.4) coinciding with its lower bound: $\min \Delta R(u) = \Delta \underline{R}(u_y^0)$.

Such a measurement procedure, specified by optimum measurable operators u^0 in the extended space $\Phi_y \otimes \Phi_\zeta^0$ and by the optimum state ρ_ζ^0 in the auxiliary system Z described by the optimum accessory space Φ_ζ^0, can be stated in the Gaussian case treated below, provided that the solution of Eq. (2.19) is assumed to be unique (this assumption is guaranteed in the above-described asymptotic case of weak noncommutativity).

§3. Estimation of Gaussian Variables

Let the estimated quantities x and the observables y be Gaussian. This means that the operators $\xi^T = (x^T, y^T)$ are, repetitively, commuting with the commutation properties

$$[\xi, \xi^\tau] = C1 = \left\|\begin{matrix} C_x & C_{xy} \\ -C_{xy}^\tau & C_y \end{matrix}\right\| 1. \tag{3.1}$$

Here 1 is the unit operator in the space $\Phi_x \otimes \Phi_y$; C_x, C_y are the specified numerical commutation matrices of the respective operators x and y, and the cross-commutator $[x, y^T]_- = C_{xy} 1$ is equal to zero because the operators x and y act (see §1) in different spaces Ψ_x and Φ_y. The state of the Gaussian variables ξ is described by a Gaussian density matrix ρ having the form

$$\rho = \exp\{-\gamma - \xi^\tau Q \xi\}, \tag{3.2}$$

where Q is a positive definite matrix specifying the correlation properties of the operators $\xi^T = (x^T, y^T)$:

$$K = {}^1/_2 M[\xi, \xi^\tau]_+ = \left\| \begin{matrix} K_x & K_{xy} \\ K_{xy}^\tau & K_y \end{matrix} \right\| = {}^1/_2 (\operatorname{cth} CQ) C \tag{3.3}$$

(it is assumed that the expectations $M\xi = 0$) and $\gamma = -1/2 \operatorname{Sp} \ln |2 \operatorname{sh} 2CQ|$ is a normalization constant.

In this case the partial trace on the right-hand side of Eq. (2.20) is computed in explicit form by a method used for the same purpose in [3]. We obtain as a result

$$m_y \rho_y + \rho_y m_y = K_{xy} K_y^{-1} (y \rho_y + \rho_y y), \text{ i.e., } m_y = K_{xy} K_y^{-1} y. \tag{3.4}$$

Here ρ_y is the partial density operator of the Gaussian operators y and is obtained by taking the partial trace of the operator ρ with respect to x:

$$\rho_y = \operatorname{Sp}_x \rho = \exp\{-\gamma_y - y^\tau Q_y y\}. \tag{3.5}$$

The constant γ_y and the matrix Q_y have the same sense as in (3.2), the matrix Q_y being determined by a correlation matrix K_y and commutation matrix C_y that are submatrices of the matrices (3.3) and (3.1), according to the formula

$$C_y Q_y = \operatorname{arc th} {}^1/_2 C_y K_y^{-1} = {}^1/_2 \ln \frac{1 + {}^1/_2 C_y K_y^{-1}}{1 - {}^1/_2 C_y K_y^{-1}}. \tag{3.6}$$

Taking (3.4) into account, we see at once that Eq. (2.19) admits the following linear solution in the given situation:

$$u_y{}^o = A^o y, \tag{3.7}$$

where A^0 is a numerical matrix. Thus, substituting $u_y = Ay$ into (2.19), we have

$$(A - K_{xy} K_y^{-1})(y \rho_y + \rho_y y) + (\operatorname{sgn} CG) A (\rho_y y - y \rho_y) = 0. \tag{3.8}$$

Then the use of the Gaussian operator relation

$$\rho_y y - y \rho_y = \operatorname{th} C_y Q_y (\rho_y y + y \rho_y) = {}^1/_2 C_y K_y^{-1} (y \rho_y + \rho_y y), \tag{3.9}$$

which has been derived in [4], enables us to go from the operator equation (3.8) to the numerical matrix equation

$$A - K_{xy} K_y^{-1} + {}^1/_2 (\operatorname{sgn} CG) A C_y K_y^{-1} = 0, \tag{3.10}$$

in which

$$C = A C_y A^\tau. \tag{3.11}$$

The solution of (3.10) for A then determines the optimum linear transformation of (3.7). Equation (3.10) is nonlinear in A, because it includes the function (2.17) of the matrix CG, where the commutation matrix C of the operators $u_y = Ay$ depends on A. The lower bound (2.21) for the corresponding losses has the value

$$\underline{R}^o = \operatorname{Sp} G(K_x - A^o K_{xy}), \tag{3.12}$$

where, in the event of several solutions of Eq. (3.10); A^0 is to be taken as the solution yielding the smallest value of the bound (3.12). If the linear solutions (3.7) exhaust all the possible solutions of Eq. (2.19), then (3.7) and (3.12) determine, respectively, the optimum operators u^0 and the corresponding lower bound of the risk.

Equation (3.10) has been derived and investigated in [4], except that in the latter the matrix $G^{-1}\Lambda$ appears instead of $1/2 \operatorname{sgn} CG$, where the matrix of indeterminate multipliers Λ was determined from the accessory commutativity conditions $C = AC_y A^T = 0$. Also, since we have introduced the concept of indirect measurement into the discussion, the operator estimates $u_y = Ay$ do not necessarily commute.

Moreover, it was assumed in the derivation of Eq. (3.10) that the matrix $C = AC_y A^T$ is degenerate, because otherwise the matrix sgn CG occurring in that equation becomes indeterminate [the definition (2.17) of the matrix sgn CG is correct only for degenerate C and G]. Not to be rejected, however, is the possibility that the optimum operators $u_y^0 = A^0 y$ have a degenerate commutation matrix $C^0 = A^0 C_y A^{0T}$. In this case the optimum matrix A^0 also satisfies Eq. (3.10) if the function sgn $(\cdot)$ is understood in the generalized sense. In particular, the case of total degeneracy $C^0 = 0$ is possible, whereupon the optimum solutions u_y^0 commute. Those solutions coincide, obviously, with the ones found in [4], and the indirect measurement of u_y^0 reduces to direct measurement thereof. An alternative, more universal method of derivation of the optimum matrices A^0, suitable for cases of partial degeneracy, entails the approximation of the modulus $|CG|$ of the matrix describing the measurement-related part of the risk (2.14) by a scalar analytic function $f_\theta(CG) = CG \coth 1/\theta CG$, whose limit $\lim f_\theta(\cdot)$ with respect to the parameter $\theta \to 0$ is equal to the modulus $|\cdot| \equiv f_0(\cdot)$. Here, in Eqs. (2.19) and (3.10) obtained by differentiation of the risk (2.4) sgn (CG) is replaced by the derivative $f_\theta'(CG)$, which is automatically defined, being analytic also at zero. The indicated approximation of the equation has a unique solution A^θ coinciding in the limit $\theta \to 0$ with the required optimum solution $A^0 = K_{xy}(K_y + 1/2 | C_y H | H^{-1})^{-1}$, where $H = A^{0T} G A^0$.

The commutation matrix $C^0 = A^0 C_y A^{0T}$ is clearly nondegenerate in the case of weak noncommutativity for a nondegenerate matrix $A_0 C_y A_0^T$, where $A_0 = K_{xy} K_y^{-1}$ is the matrix describing the optimum transformation $\gamma_0(y) = A_0 y$ according to the irreducible criterion (2.3).

It is readily shown that the case of weak noncommutativity (2.22) of the operators $m_y = A_0 y$ with a density matrix ρ_y of the Gaussian form (3.5), guaranteeing uniqueness of the solution (3.7) of Eq. (2.19), corresponds to the quasiclassical approximation $K_y \gg \pm C_y/2$. Here the solution of Eq. (3.10) can be found by the method of successive approximations, taking $A_0 = K_{xy} K_y^{-1}$ as the zeroth matrix. Setting $1/2 K_y^{-1} C_y K_y^{-1} = \varepsilon S$ for definiteness, we have in the first approximation

$$A^0 = A_0 - \varepsilon \operatorname{sgn}(K_{xy} S K_{yx} G) K_{xy} S + O(\varepsilon^2), \tag{3.13}$$

where the matrix S is related to the matrix (2.25) by the expression $K_{xy} S K_{xy}^T = S_m$.

Substituting (3.13) into (3.12), we obtain the following expression for the lower bound of the optimum estimation risk in the given quasiclassical approximation:

$$\underline{R} = \operatorname{Sp} G(K_x - K_{xy} K_y^{-1} K_{xy}^\tau) + \varepsilon \operatorname{Sp} |K_{xy} S K_{xy}^\tau G| + O(\varepsilon^2). \tag{3.14}$$

The first term of Eq. (3.14) coincides with the minimum irreducible estimation risk considered in [3]. The second term is the minimum asymptotic (with respect to ε) error, compatible with the Heisenberg inequality (2.5), of indirect measurement of the optimum operators $A_0 y = K_{xy} K_y^{-1} y$ according to the irreducible criterion (2.3).

§4. Physical Procedure for Indirect Measurement of

Optimum Gaussian Noncommutative Estimates

All that remains for the final solution of the optimum estimation problem is to describe the optimum procedure for indirect measurement of the optimum operators u_y^0. This can be done in the Gaussian case by virtue of the fact that the optimum estimates u_y^0 are linear combinations $A^0 y$ of the Gaussian repetitively commutative variables themselves. On the other hand, indirect measurement of the repetitively commutative variables $A^0 y$ can be realized by direct measurement of the commutative linear combinations

$$u^0 = A^0 y + \zeta, \tag{4.1}$$

where the operators ζ are repetitively commutative, independent of y, and have a commutation matrix C_ζ of the form

$$C_\zeta = -C^0 = -A^0 C_y A^{0\tau} \tag{4.2}$$

and zero expectations: $M\zeta = 0$. It is readily verified that, given this choice of commutation matrix, the commutativity condition for the operators (4.1) is satisfied and that the averaging (2.2) of those operators with respect to the auxiliary variables ζ does in fact coincide with the operators $A^0 y$.

To fully specify the optimum indirect measurement procedure it is required, along with the measurable physical variables, to indicate the optimum state ρ_ζ^0 of an auxiliary physical system that can be identified with the measurement apparatus. In the given Gaussian case, as it turns out, there is a state for which the measurement error $\Delta R(u^0)$ attains its lower bound $\Delta \underline{R}(u_y^0)$, which is compatible with the

214

Heisenberg uncertainty principle, so that the optimum estimation risk coincides with its lower bound (3.12).

States of this kind can be represented as equilibrium thermodynamic states $\rho_\zeta = \exp\{-\gamma(\beta) - \beta H\}$ with a quadratic Hamiltonian:

$$H = \zeta^\tau G_\zeta \zeta, \qquad G_\zeta \geqslant 0, \tag{4.3}$$

and zero temperature $T = 1/\beta = 0$.

A system described by a Hamiltonian of the type (4.3) is linear, and its equilibrium states are clearly Gaussian. At zero temperature the Gaussian density matrix ρ_ζ becomes totally degenerate, and the state of the "apparatus" is pure. This pure state is an eigenstate of the Hamiltonian operator (4.3) and corresponds to the minimum eigenvalue of that operator; it is called a "vacuum" state.

Comparing the correlation matrices

$$K_\zeta = \lim_{\beta \to \infty} {}^{1}/_{2} (\operatorname{cth} \beta C_\zeta G_\zeta) C_\zeta = {}^{1}/_{2} |C_\zeta G_\zeta| G_\zeta^{-1}$$

for the vacuum states with the matrix

$$K_\zeta^0 = {}^{1}/_{2} |C_\zeta G| G^{-1}, \tag{4.4}$$

for which inequality (2.9) goes over to an equality, we see at once that the optimum state of the "apparatus" is the vacuum state for $G_\zeta = G$:

$$\rho_\zeta^0 = \lim_{\beta \to \infty} \exp\{-\gamma(\beta) - \beta \zeta^\tau G \zeta\}. \tag{4.5}$$

The correlation matrix $K_\zeta^0 = 1/2 M[\zeta, \zeta^T]_+$ in this case has the form (4.4), and the indirect measurement error attains its lower bound (2.12).

Thus, the optimum auxiliary system, or "apparatus," can be represented in the Gaussian case for a quadratic figure of merit as a certain linear physical system with Hamiltonian $H^0 = \zeta^T G \zeta$, existing in the equilibrium state at zero temperature. The phase coordinates ζ of this system are described by repetitively commutative operators in the space Φ_ζ^0, which is a copy of the space $\Phi_{A^0 y}$ in which the optimum operators $u_y^0 = A^0 y$ act, and they have the commutation matrix (4.2). The optimum commutative operators (4.1), which are measured to give the required optimum estimates of the variables x, are linear combinations of the Gaussian coordinates (y, ζ) of the aggregate observed system. This condition serves as the analog of linearity of the optimum (in the sense of a quadratic figure of merit) estimation (filtration) of classical Gaussian random variables; unlike the classical case, however, the resulting optimum estimation algorithm (4.1) is randomized on the vacuum noise ζ.

The authors are indebted to R. L. Stratonovich for a discussion and valuable comments.

LITERATURE CITED

1. Takahashi, "Application of information theory to quantum-mechanical channels," in: Statistical Communication Theory and Its Application [Russian translation], Mir, Moscow (1967), pp. 160-257.
2. C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum-mechanical communication theory," Proc. IEEE, 58, No. 10, 1578-1598 (1970).
3. B. A. Grishanin and R. L. Stratonovich, "Optimum filtering of quantum variables for a quadratic figure of merit," Probl. Peredachi Inform., 6, No. 3, 15-23 (1970).
4. V. P. Belavkin and B. A. Grishanin, "Optimum measurement of quantum variables," Probl. Peredachi Inform., 8, No. 3, 103-109 (1972).
5. S. D. Personick, Efficient Analog Communication over Quantum Channels, MIT Tech. Rep. 477, Cambridge, Mass. (1970).
6. R. Bellman, Introduction to Matrix Analysis, McGraw–Hill, New York (1960).