Journal of Statistical Physics, Vol. 114, Nos. 5/6, March 2004 (© 2004)

An Informational Characterization of Schrödinger's Uncertainty Relations

Shunlong Luo[1] **and Zhengmin Zhang**[2]

Received May 20, 2003; accepted October 21, 2003

Heisenberg's uncertainty relations employ commutators of observables to set fundamental limits on quantum measurement. The information concerning incompatibility (non-commutativity) of observables is well included but that concerning correlation is missing. Schrödinger's uncertainty relations remedy this defect by supplementing the correlation in terms of anti-commutators. However, both Heisenberg's uncertainty relations and Schrödinger's uncertainty relations are expressed in terms of variances, which are not good measures of uncertainty in general situations (e.g., when mixed states are involved). By virtue of the Wigner–Yanase skew information, we will establish an uncertainty relation along the spirit of Schrödinger from a statistical inference perspective and propose a conjecture. The result may be interpreted as a quantification of certain aspect of the celebrated Wigner–Araki–Yanase theorem for quantum measurement, which states that observables not commuting with a conserved quantity cannot be measured exactly.

KEY WORDS: Quantum measurement; uncertainty relations; Fisher information; skew information; Wigner–Yanase correlation.

1. UNCERTAINTY RELATIONS: FROM HEISENBERG TO SCHRÖDINGER

Shortly after Heisenberg invented matrix mechanics and Schrödinger established wave mechanics around 1925, which constitute the new theory of quantum mechanics, Heisenberg discovered the uncertainty principle in 1927,[12] which was put into mathematical forms by Weyl[30] and

[1] Academy of Mathematics and System Sciences, Chinese Academy of Sciences, 100080 Beijing, People's Republic of China; e-mail: luosl@mail.amt.ac.cn

[2] School of Mathematics and Statistics, Carleton University, Ottawa, Ontario K1S 5B6, Canada; e-mail: zhangzm@math.carleton.ca

Robertson.[22] Now the standard forms of Heisenberg's uncertainty relations for any two observables A and B are usually expressed as

$$\mathrm{Var}_\rho A \cdot \mathrm{Var}_\rho B \geqslant \tfrac{1}{4} |\langle [A, B] \rangle_\rho|^2. \tag{1}$$

Here $\mathrm{Var}_\rho A = \mathrm{tr}\, \rho A^2 - (\mathrm{tr}\, \rho A)^2$ is the variance of A in the relevant quantum state ρ (a mixed state in general) and $\mathrm{Var}_\rho B$ is defined similarly. $\langle [A, B] \rangle_\rho = \mathrm{tr}\, \rho [A, B]$ is the average of the commutator $[A, B] = AB - BA$ in the state ρ. The conventional position-momentum uncertainty relation is a particular case.

It is remarkable that the commutator, which is so characteristic in quantum mechanics, makes its appearance here. Thus uncertainty relations in Heisenberg's form are intimately related to non-commutativity. The above inequality shows clearly the limitations in the possibility of simultaneously assigning exact values to two non-commuting observables.

In recent years, especially in the newly emerging field of quantum computation and quantum information, the phenomenon of entanglement is widely studied in connection with Bell's inequalities and teleportation,[20, 28] and with other foundational issues of quantum mechanics. Apart from the ubiquitous feature of non-commutativity, one comes to recognize that another distinguished feature of quantum mechanics is the strong correlation in quantum world that cannot be accounted for in classical mechanics. However, the uncertainty relations in Heisenberg's form do not encode any correlation between observables which is usually expressed in terms of anti-commutators. It is amazing that Schrödinger in 1930 already established the uncertainty relations taking into account of the correlation between observables,[23] though almost all quantum mechanics text books have ignored this, and consequently, Schrödinger's uncertainty relations are left out of the general awareness of physicists for many years. Only quite recently have Schrödinger's uncertainty relations been studied thoroughly by some authors.[25]

Schrödinger's uncertainty relations are expressed in both commutator (encoding non-commutativity) and anti-commutator (encoding correlation):

$$\mathrm{Var}_\rho A \cdot \mathrm{Var}_\rho B \geqslant \tfrac{1}{4} (|\langle [A, B] \rangle_\rho|^2 + |\langle \{A_0, B_0\} \rangle_\rho|^2). \tag{2}$$

Here $\langle \{A_0, B_0\} \rangle_\rho = \mathrm{tr}\, \rho \{A_0, B_0\}$ denotes the average of the anti-commutator $\{A_0, B_0\} = A_0 B_0 + B_0 A_0$, and $A_0 = A - \langle A \rangle_\rho$, $B_0 = B - \langle B \rangle_\rho$. The first term in the right hand side of (2) encodes incompatibility, while the second term encodes correlation, between the observables A and B. Indeed, the average of the anti-commutator is simply related to covariance as

$$\langle \{A_0, B_0\} \rangle_\rho = \mathrm{Cov}_\rho(A, B) + \mathrm{Cov}_\rho(B, A)$$

with the conventional covariance defined as

$$\mathrm{Cov}_\rho(A, B) = \mathrm{tr}\, \rho AB - \mathrm{tr}\, \rho A \cdot \mathrm{tr}\, \rho B. \tag{3}$$

The difference between Schrödinger's uncertainty relations and Heisenberg's uncertainty relations is fundamental. The contribution due to the correlation (expressed in terms of anti-commutators or covariance) should be put at least on an equal footing with that due to incompatibility (expressed in terms of commutators or incompatibility). In particular, the correlation term plays a crucial role in studying composite quantum systems. For example, consider a composite quantum system described by a tensor product Hilbert space $\mathscr{H}_1 \otimes \mathscr{H}_2$, where $\mathscr{H}_1$ and $\mathscr{H}_2$ are the Hilbert spaces of the two component systems. Let a and b be observables for the first and the second system, respectively. Define $A = a \otimes I_2$, $B = I_1 \otimes b$, which are observables on $\mathscr{H}_1 \otimes \mathscr{H}_2$ (here I_1 and I_2 denote the identity operators on $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively), then clearly A commutes with B, and consequently, Heisenberg's uncertainty relations (1) for A and B reduces to a trivial inequality $\mathrm{Var}_\rho\, A \cdot \mathrm{Var}_\rho\, B \geqslant 0$ which does not provide any useful information. But this is not the case for Schrödinger's uncertainty relations (2) since it is still a non-trivial inequality involving variances and anti-commutator:

$$\mathrm{Var}_\rho\, A \cdot \mathrm{Var}_\rho\, B \geqslant \tfrac{1}{4} |\langle \{A_0, B_0\} \rangle_\rho|^2.$$

The place of the uncertainty principle in quantum mechanics is unique and somewhat peculiar. To quote from the first section of the first chapter of Landau and Lischitz's celebrated quantum mechanics textbook:[(15)] *In that it rejects the ordinary ideas of classical mechanics, the uncertainty principle might be said to be negative in content. Of course, this principle in itself does not suffice as a basis on which to construct a new mechanics of particles. Such a theory must naturally be founded on some positive assertions.*

The main reason that such an impression prevails lies in the fact that most people regard the uncertainty principle as a principle imposing theoretical limitations of quantum measurement without precise quantitative characterizations, and therefore, the enthusiasm is mainly devoted into the physical interpretations and philosophical implications, rather than precise mathematical derivations and characterizations, of the uncertainty principle. The ubiquitous mathematical forms of uncertainty relations involving variance and covariance are not very useful except for giving some crude estimations and a rough idea. In particular, for mixed states, the conventional covariance has a lot of shortcomings in characterizing quantum correlation (entanglement).

Several new measures quantifying uncertainty have been proposed by many authors. A prominent measure is the Shannon entropy, and there are many mathematical characterizations of uncertainty and of the uncertainty principle by virtue of this information concept.[2, 8, 27] Recall how the scenarios of statistical physics and communication engineering have changed after the Shannon entropy is introduced to replace variance in characterizing uncertainty (indeed, the principle of maximum entropy and the informational characterizations of the second law of thermodynamics have generated profound implications in physics), one may expect that once the uncertainty principle is quantified more precisely, many more physical results and laws can be derived naturally via symmetry arguments and variational calculus from the uncertainty principle.

Another prominent measure related to uncertainty is the Fisher information arising in the theory of statistical inference,[7] and there are also many mathematical characterizations of uncertainty (or rather, information) and of the uncertainty principle by virtue of the Fisher information.[9, 10, 17, 18] Moreover, the Fisher information has many intrinsic relevances to physics.[9] For instance, Hall and Reginatto provided an interesting and stimulating derivation of Schrödinger's equation from an identity characterization of uncertainty relations by virtue of the Fisher information.[11] Luo presented an informational derivation of the sinusoidal law (Malus' law) for photon polarization based on the principle of minimum Fisher information.[19] There are also various notions of quantum Fisher information, which refine the notion of the conventional variance in quantum detection and estimation theory.[13, 14]

Before we can do more physics from the uncertainty principle, we have to establish more precise quantifications of uncertainty. In this paper, we will show that how a quantity introduced by Wigner and Yanase in 1963,[32] the skew information, arises naturally as a quantum generalization of the Fisher information, and can be used to characterize uncertainty relations. In fact, we will establish a new uncertainty relation of Schrödinger's type by virtue of the skew information. The result is not only stronger than the conventional Heisenberg's uncertainty relations (1), but also sheds considerable new light on the relationships between the theories of quantum measurement and statistical inference. Moreover, it is strictly stronger than the Schrödinger's uncertainty relations (2) at least for two-dimensional quantum systems, and for some cases as illustrated in a variety of examples. Whether this is true in general remains open (though we strongly believe so). The result is also intimately related to certain quantitative aspect of the Wigner–Araki–Yanase theorem concerning quantum measurement, which states that the presence of a conservation law imposes a limitation on the measurement of observables which are incompatible (not commuting) with the conserved quantity.[1, 33]

2. FISHER INFORMATION AND SKEW INFORMATION

Consider a natural problem in the theory of statistical estimation: Suppose $\{p_\theta : \theta \in R\}$ is a family of probability densities on R parameterized by θ, and we have observed samples $x_1, \ldots, x_n$, each is a random variable (independently) distributed according to p_θ for some fixed unknown θ, we want to estimate this θ as precise as possible by virtue of the available data.

In this context, the Fisher information defined by[6, 7]

$$I_F(p_\theta) = 4 \int \left(\frac{\partial \sqrt{p_\theta(x)}}{\partial \theta} \right)^2 dx \tag{4}$$

is a central concept. The celebrated Cramér–Rao inequality and the asymptotic normality of maximum likelihood estimation are both phrased in terms of the Fisher information.[6] Moreover, the Fisher information is intimately related to the Shannon entropy via the elegant de Bruijin identity,[24] plays a pivotal role in quantifying Heisenberg's uncertainty principle from a statistical inference perspective,[9, 10, 17, 18] and even has interesting applications in some probability problems.[3, 4]

It is remarkable that the probability amplitude $\sqrt{p_\theta(x)}$, so fundamental in quantum mechanics, manifests itself here. In particular, when $p_\theta(x) = p(x-\theta)$, by the translation invariance of the Lebesgue integral, we have

$$I_F(p_\theta) = 4 \int \left(\frac{\partial \sqrt{p(x)}}{\partial x} \right)^2 dx.$$

Thus in this circumstance, $I_F(p_\theta)$ is independent of θ, and we can denote $I_F(p_\theta)$ simply by $I_F(p)$. It is the Fisher information of p with respect to the location parameter, and is reminiscent of the kinetic energy functional.

Remark. The Fisher information can be rewritten as

$$I_F(p_\theta) = \int \left(\frac{\partial \log p_\theta(x)}{\partial \theta} \right)^2 p_\theta(x)\, dx. \tag{5}$$

In this expression, the statistical score function $\frac{\partial \log p_\theta(x)}{\partial \theta}$, so useful in statistical inference, makes its appearance. It is probably the coincidence of the Fisher information defined by Eq. (4), which involves the probability amplitude, and by Eq. (5), which involves the statistical score function, that renders the Fisher information so useful in studying certain quantum mechanical issues from a statistical inference perspective.

When we pass from classical theory to quantum mechanics, the integral is replaced by trace, the probability densities are replaced by density operators. Motivated by (4), for a family of quantum states ρ_θ, $\theta \in R$, let us define heuristically

$$I_F(\rho_\theta) = 4 \operatorname{tr} \left(\frac{\partial \sqrt{\rho_\theta}}{\partial \theta} \right)^2$$

which may be viewed as a generalization of the classical Fisher information to quantum case. Let us just call it quantum Fisher information. Of course, quantum generalizations of the Fisher information are not unique, and there are many others such as those defined via symmetric logarithmic derivative and via right logarithmic derivative which are related to the theory of quantum detection and quantum estimation,[13, 14] and those defined via general operator monotone functions,[21] but we will not pursue them here.

In particular, if ρ_θ satisfies the Landau–von Neumann equation (we put $\hbar = 1$)

$$i \frac{\partial \rho_\theta}{\partial \theta} = [A, \rho_\theta], \qquad \rho_0 = \rho$$

where $\theta \in R$ is a (temporal or spatial) parameter, and A may be interpreted as the generator of the temporal shift or the spatial displacement, then $\rho_\theta = e^{-i\theta A} \rho e^{i\theta A}$, and

$$\frac{\partial \sqrt{\rho_\theta}}{\partial \theta} = i e^{-i\theta A} [\rho^{1/2}, A]\, e^{i\theta A},$$

which in turn implies that

$$I_F(\rho_\theta) = 8 I(\rho, A)$$

with $I(\rho, A)$ defined as

$$I(\rho, A) = -\tfrac{1}{2} \operatorname{tr}[\rho^{1/2}, A]^2. \tag{6}$$

Remarkably, this $I(\rho, A)$ is precisely the skew information introduced by Wigner and Yanase as the amount of information on the values of observables not commuting with A.[32] Therefore, the skew information is essentially a particular kind of quantum Fisher information! With this observation in mind, it is not surprising that the skew information has so many nice properties expected from a measure of information, and is relevant to the theory of quantum measurement. In fact, Wigner and Yanase argued and proved that this quantity satisfies all the desirable intuitive requirements of an information measure,[32] among which the basic ones are:

1. *Invariance*

Let U be a unitary operator, then

$$I(U\rho U^{-1}, A) = I(\rho, U^{-1}AU).$$

In particular, if U commutes with A, then

$$I(U\rho U^{-1}, A) = I(\rho, A).$$

When the state changes according to the Landau–von Neumann equation generated by A, we have $U = e^{-i\theta A}$, which commutes with A, therefore the skew information remains constant for isolated systems. Moreover, $I(\rho, A + A_0) = I(\rho, A)$ for any A_0 commuting with ρ.

2. *Convexity*

Let ρ_1 and ρ_2 be two quantum states described by density operators, then

$$I(\lambda_1\rho_1 + \lambda_2\rho_2, A) \leqslant \lambda_1 I(\rho_1, A) + \lambda_2 I(\rho_2, A)$$

for any $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$. The above convexity means that the information content of mixing of two ensembles should be smaller than the average information content of the component ensembles, that is, the skew information decreases when two different ensembles are united, since by uniting, one "forgets" from which a particular sample stems.

3. *Additivity*

Let ρ_1 and ρ_2 be two density operators describing the first system and the second system respectively, A_1 and A_2 be the corresponding conserved quantities, then

$$I(\rho_1 \otimes \rho_2, A_1 \otimes I_2 + I_1 \otimes A_2) = I(\rho_1, A_1) + I(\rho_2, A_2).$$

Here $A_1 \otimes I_2$ denotes the tensor product of A_1 in the first system with the identity operator in the second system, and $I_1 \otimes A_2$ is defined similarly. The above identity means that the skew information is additive under tensor product, namely, the information content of a system composed of two independent parts is the sum of information of the parts.

The Wigner–Yanase skew information can be rewritten as

$$I(\rho, A) = \mathrm{tr}\, \rho A^2 - \mathrm{tr}\, \rho^{1/2} A \rho^{1/2} A. \tag{7}$$

In particular, if $\rho = |\Psi\rangle\langle\Psi|$ is a pure state, then

$$I(\rho, A) = \mathrm{Var}_\rho\, A = \langle\Psi|\, A^2 \,|\Psi\rangle - \langle\Psi|\, A \,|\Psi\rangle^2.$$

Therefore, $I(\rho, A)$ may also be regarded as a measure of uncertainty of A, and for pure states, it reduces to variance. This is in accordance with the intuition based on the principle of complementarity, since the larger the variance of A (which means the less information we can have about the value of A), the more information we can have about an observable which does not commute with (is skew to) A by complementarity, thus corresponding to larger skew information. Moreover, if the quantum state ρ is an eigenstate of A, then $I(\rho, A) = 0$. To gain a more intuitive insight of this relation, take A to be the position observable and ρ to be an eigenstate of A, then $I(\rho, A) = 0$, and this means that the skew information of any observable not commuting with A (e.g., momentum observable) is zero. Indeed, the value of the momentum observable is equally probable over the whole line (thus we have the least information about the value of momentum) when the quantum system is in a position eigenstate.

In summary, we have clarified the statistical idea underlying the Wigner–Yanase skew information, and have put it into the context of the theory of statistical estimation and have related it to variance.

Remark. The Wigner–Yanase skew information is later on generalized by Dyson to

$$I_\alpha(\rho, A) = -\tfrac{1}{2}\, \mathrm{tr}[\rho^\alpha, A][\rho^{1-\alpha}, A], \qquad 0 < \alpha < 1,$$

and the famous Wigner–Yanase–Dyson conjecture concerning the convexity of $I_\alpha(\rho, A)$ with respect to ρ is solved by Lieb.[16] Another generalization to the setting of operator algebras is made by Connes and Stormer,[5] and plays a crucial role when they proved the homogeneity of state spaces of III_1 von Neumann algebras.

Remark. The Wigner–Yanase skew information has its origin in the theory of quantum measurement. In quantum mechanics, it is typical that some observables are more difficult to measure than others. Wigner demonstrated that the observables which do not commute with an additive conserved quantity are more difficult to measure than those which commute with the conserved quantity.[31] This phenomenon is closely related, but not equivalent, to Heisenberg's uncertainty principle. Pursuing Wigner's idea further, Araki and Yanase established rigorously that observables not commuting with a conserved quantity cannot be measured exactly, only an approximate measurement is possible, and there is a trade-off between the "size" of the measuring apparatus and measuring accuracy.[1, 33] This is the celebrated Wigner–Araki–Yanase theorem for quantum measurement, and the Wigner–Yanase skew information is introduced to quantify certain aspect of this result. The observable A serves as a conserved quantity such as a Hamiltonian, a momentum, or other conserved quantities. Formally, $I(\rho, A)$ may also be interpreted as a measure of non-commutativity between ρ and A with asymmetric emphasis on the state ρ and on the conserved observable A.[5]

3. THE WIGNER–YANASE CORRELATION

Covariance is usually used to characterize the correlation between two observables in a given quantum state. Alternatively, it can also be used to characterize the intrinsic correlation of a quantum state, given the two observables fixed. That is, $\mathrm{Cov}_\rho(A, B)$ may be used as a measure to quantify the "correlation strength" of the state ρ. The observables A and B serve here as testing observables.

The distinction between classical and quantum correlation is fundamental and subtle, and it is a difficult and thorny problem as how to distinguish between them. In this respect, the conventional covariance often gives ambiguous results. Let us demonstrate this point by a simple example.

Example 1. Let C^2 be a qubit space with orthonormal base $\{|\uparrow\rangle, |\downarrow\rangle\}$ which are the two eigenvectors of the Pauli Spin-z operator $\sigma_z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ with eigenvalues $+1$ and -1, respectively. Let $\mathscr{H}_1 = C^2$, $\mathscr{H}_2 = C^2$, and form the composite quantum system $\mathscr{H}_1 \otimes \mathscr{H}_2 = C^2 \otimes C^2 = C^4$. Then $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$ constitute an orthonormal base for C^4. Take a quantum state

$$\rho = \tfrac{1}{2}\,(|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + |\downarrow\downarrow\rangle\langle\downarrow\downarrow|)$$

and

$$A = \sigma_z \otimes I_2, \qquad B = I_1 \otimes \sigma_z.$$

Then direct calculation leads to

$$\mathrm{Cov}_\rho(A, B) = 1.$$

On the other hand, if we take

$$\rho' = |\Psi\rangle\langle\Psi|$$

with $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$, then

$$\mathrm{Cov}_{\rho'}(A, B) = 1.$$

From the entanglement point of view, the two states ρ and ρ' are very different: The former is a mixture of two disentangled states, while the latter is a Bell state which is maximally entangled, and the covariance cannot distinguish between them. In this sense, the conventional covariance has a limited use in characterizing entanglement. We will introduce a new quantity which has some advantages in quantifying entanglement from an informational perspective.

Motivated by the definition of the Wigner–Yanase skew information, Eq. (7), we define a correlation measure for two observables A and B in any quantum state ρ as

$$\mathrm{Corr}_\rho(A, B) = \mathrm{tr}\, \rho AB - \mathrm{tr}\, \rho^{1/2} A \rho^{1/2} B. \tag{8}$$

We shall call this the Wigner–Yanase correlation, and compare it with the conventional covariance defined by Eq. (3). (More generally, we may define $\mathrm{Corr}_\rho(X, Y) = \mathrm{tr}\, \rho X^* Y - \mathrm{tr}\, \rho^{1/2} X^* \rho^{1/2} Y$ as an inner product on the space of all bounded linear operators). In particular,

$$\mathrm{Corr}_\rho(A, A) = \mathrm{tr}\, \rho A^2 - \mathrm{tr}(\rho^{1/2} A)^2 = -\tfrac{1}{2}\, \mathrm{tr}[\rho^{1/2}, A]^2$$

is exactly the skew information $I(\rho, A)$ introduced by Wigner and Yanase when they study the information contents of quantum states and the theory of quantum measurement.[32] This concept is closely related to the quantum theory of measurement. It can be easily shown that if $\rho = |\Psi\rangle\langle\Psi|$ is a pure state, then

$$\mathrm{Cov}_\rho(A, B) = \mathrm{Corr}_\rho(A, B)$$

for any observables A and B.

The Wigner–Yanase correlation has the following properties:

(1) $\mathrm{Corr}_\rho(A, B) = \overline{\mathrm{Corr}_\rho(B, A)}$.

(2) $\mathrm{Corr}_\rho(A - a, B - b) = \mathrm{Corr}_\rho(A, B)$ for any real constants a, b.

(3) $\mathrm{Corr}_\rho(aA, bB) = ab\, \mathrm{Corr}_\rho(A, B)$ for any real constants a, b.

(4) $\mathrm{Corr}_{U\rho U^{-1}}(A, B) = \mathrm{Corr}_\rho(U^{-1}AU, U^{-1}BU)$ for any unitary operator U.

From Example 1, we have seen that the conventional covariance has a limited use in characterizing quantum correlation. We now illustrate by an example that the Wigner–Yanase correlation has certain advantages in this respect.

Example 2. Let ρ, ρ', A, and B be the same as in Example 1. Then in the canonical base $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$, we can write

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \rho' = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Now direct calculation leads to

$$\mathrm{Corr}_\rho(A, B) = 0, \qquad \mathrm{Corr}_{\rho'}(A, B) = 1.$$

Thus the Wigner–Yanase correlation indeed distinguishes the states ρ and ρ'. This is in sharp contrast to the conventional covariance.

One may wonder what is the relationship between the covariance $\mathrm{Cov}_\rho(A, B)$ and the Wigner–Yanase correlation $\mathrm{Corr}_\rho(A, B)$. In this respect, we have $\mathrm{Cov}_\rho(A, B) \geqslant \mathrm{Corr}_\rho(A, B)$ when $A = B$, that is, $\mathrm{Var}_\rho\, A \geqslant I(\rho, A)$, as will be shown in the next section. But in general, there is no simple dominant relations between them, as illustrated by the following examples.

Example 3. Let us take a Hilbert space C^4. Take a quantum state to be

$$\rho = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

and two observables

$$A=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},\qquad B=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}.$$

The state ρ can be diagonalized as

$$\rho=U\begin{pmatrix}0&0&0&0\\0&\frac{1}{4}&0&0\\0&0&\frac{1}{4}&0\\0&0&0&\frac{1}{2}\end{pmatrix}U^{-1},\quad \text{with unitary } U=\frac{1}{\sqrt{2}}\begin{pmatrix}1&0&0&1\\0&1&1&0\\0&-1&1&0\\-1&0&0&1\end{pmatrix}.$$

From this we obtain

$$\rho^{1/2}=\frac{1}{2}\begin{pmatrix}\frac{1}{\sqrt{2}}&0&0&\frac{1}{\sqrt{2}}\\0&1&0&0\\0&0&1&0\\\frac{1}{\sqrt{2}}&0&0&\frac{1}{\sqrt{2}}\end{pmatrix}.$$

Now, by the definitions, Eqs. (3) and (8), simple calculation leads to

$$\mathrm{Cov}_\rho(A,B)=0,\qquad \mathrm{Corr}_\rho(A,B)=\tfrac{1}{2}.$$

Thus in this example, we have

$$|\mathrm{Cov}_\rho(A,B)|<|\mathrm{Corr}_\rho(A,B)|.$$

Example 4. For any p satisfying $0\leqslant p\leqslant 1$, take the quantum state to be

$$\rho=\frac{1}{4}\begin{pmatrix}1+p&0&0&2p\\0&1-p&0&0\\0&0&1-p&0\\2p&0&0&1+p\end{pmatrix}$$

and two observables A and B the same as in Example 3.

The state ρ is actually the Werner state introduced in quantum information theory,[20, 28] and it can be diagonalized as

$$\rho = V \begin{pmatrix} \frac{1-p}{4} & 0 & 0 & 0 \\ 0 & \frac{1-p}{4} & 0 & 0 \\ 0 & 0 & \frac{1-p}{4} & 0 \\ 0 & 0 & 0 & \frac{1+3p}{4} \end{pmatrix} V^{-1},$$

$$\text{with unitary} \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 & \sqrt{2} \\ \sqrt{2} & 1 & -1 & 0 \\ -\sqrt{2} & 1 & -1 & 0 \\ 0 & -1 & -1 & \sqrt{2} \end{pmatrix}.$$

From this we obtain

$$\rho^{1/2} = \frac{1}{4} \begin{pmatrix} \sqrt{1+3p}+\sqrt{1-p} & 0 & 0 & \sqrt{1+3p}-\sqrt{1-p} \\ 0 & 2\sqrt{1-p} & 0 & 0 \\ 0 & 0 & 2\sqrt{1-p} & 0 \\ \sqrt{1+3p}-\sqrt{1-p} & 0 & 0 & \sqrt{1+3p}+\sqrt{1-p} \end{pmatrix}.$$

Now, straightforward calculation leads to

$$\mathrm{Cov}_\rho(A, B) = p, \qquad \mathrm{Corr}_\rho(A, B) = \tfrac{1}{2}\left(1+p-\sqrt{(1+3p)(1-p)}\right).$$

Simple verification shows that in this case

$$|\mathrm{Cov}_\rho(A, B)| > |\mathrm{Corr}_\rho(A, B)|$$

when $0 < p < 1$.

In summary, the magnitude of the conventional covariance $\mathrm{Cov}_\rho(A, B)$ may be larger, or smaller, than the Wigner–Yanase correlation $\mathrm{Corr}_\rho(A, B)$.

4. SCHRÖDINGER'S UNCERTAINTY RELATIONS IN TERMS OF SKEW INFORMATION

Before we derive an uncertainty relation in the spirit of Schrödinger from a statistical inference perspective, we show that the Wigner–Yanase skew information is dominated by variance.

Theorem 1. Let ρ be a quantum state (pure or mixed), then

$$\mathrm{Var}_\rho\, A \geqslant I(\rho, A).$$

Furthermore, when A is a non-degenerate observable, the equality holds if and only if ρ is a pure state.

Proof. We will prove the result by writing ρ in spectral form and evaluating both $\mathrm{Var}_\rho\, A$ and $I(\rho, A)$ in the orthonormal base diagonalizing ρ.

Let ρ be written in spectral decomposition form

$$\rho = \sum_m \lambda_m \, |\psi_m\rangle\langle\psi_m|.$$

For simplicity, we assume that ρ is non-degenerate and thus $\{|\psi_m\rangle\}$ constitute an orthonormal base. Then

$$\mathrm{tr}\, \rho A = \sum_m \langle\psi_m|\, \rho A \,|\psi_m\rangle = \sum_m \lambda_m \, \langle\psi_m|\, A \,|\psi_m\rangle,$$

and

$$\begin{aligned}
\mathrm{tr}\, \rho A^2 &= \sum_{m,n} \lambda_m \, \langle\psi_m|\, A \sum_n |\psi_n\rangle\langle\psi_n|\, A \,|\psi_m\rangle \\
&= \sum_{m,n} \lambda_m \, |\langle\psi_m|\, A \,|\psi_n\rangle|^2 \\
&= \sum_{m,n} \frac{\lambda_m + \lambda_n}{2} \, |\langle\psi_m|\, A \,|\psi_n\rangle|^2,
\end{aligned}$$

therefore, the variance of A in the state ρ satisfies

$$\begin{aligned}
\mathrm{Var}_\rho\, A &= \mathrm{tr}\, \rho A^2 - (\mathrm{tr}\, \rho A)^2 \\
&= \sum_{m,n} \frac{\lambda_m + \lambda_n}{2} \, |\langle\psi_m|\, A \,|\psi_n\rangle|^2 - \left(\sum_m \lambda_m \, \langle\psi_m|\, A \,|\psi_m\rangle \right)^2 \\
&= \sum_{m \neq n} \frac{\lambda_m + \lambda_n}{2} \, |\langle\psi_m|\, A \,|\psi_n\rangle|^2 + \sum_m \lambda_m \, |\langle\psi_m|\, A \,|\psi_m\rangle|^2 \\
&\quad - \left(\sum_m \lambda_m \, \langle\psi_m|\, A \,|\psi_m\rangle \right)^2 \\
&\geqslant \sum_{m \neq n} \frac{\lambda_m + \lambda_n}{2} \, |\langle\psi_m|\, A \,|\psi_n\rangle|^2.
\end{aligned}$$

The last inequality follows from the Cauchy–Schwarz inequality. On the other hand,

$$
\begin{aligned}
I(\rho, A) &= \operatorname{tr} \rho A^2 - \operatorname{tr} \rho^{1/2} A \rho^{1/2} A \\
&= \sum_{m,n} \frac{\lambda_m + \lambda_n}{2} |\langle \psi_m | A | \psi_n \rangle|^2 - \sum_{m,n} \langle \psi_m | \rho^{1/2} A | \psi_n \rangle \langle \psi_n | \rho^{1/2} A | \psi_m \rangle \\
&= \sum_{m,n} \frac{\lambda_m + \lambda_n}{2} |\langle \psi_m | A | \psi_n \rangle|^2 - \sum_{m,n} \lambda_m^{1/2} \lambda_n^{1/2} |\langle \psi_m | A | \psi_n \rangle|^2 \\
&= \sum_{m \neq n} \frac{\lambda_m + \lambda_n - 2\lambda_m^{1/2} \lambda_n^{1/2}}{2} |\langle \psi_m | A | \psi_n \rangle|^2.
\end{aligned}
$$

Now the conclusion follows readily.

The following theorem provides a characterization of the uncertainty principle in the spirit of Schrödinger by virtue of the Wigner–Yanase skew information.

Theorem 2. Let A, B be two observables, and ρ a quantum state, then

$$I(\rho, A) \cdot I(\rho, B) \geqslant \tfrac{1}{4} |\langle [A, B] \rangle_\rho|^2 + \tfrac{1}{16} |I(\rho, A+B) - I(\rho, A-B)|^2.$$

In particular,

$$I(\rho, A) \cdot I(\rho, B) \geqslant \tfrac{1}{4} |\langle [A, B] \rangle_\rho|^2.$$

Proof. Thanks to the cyclic property of trace, the Wigner–Yanase correlation defines an inner product in the space of all bounded operators:

$$\operatorname{Corr}_\rho(X, Y) = \operatorname{tr} \rho X^* Y - \operatorname{tr} \rho^{1/2} X^* \rho^{1/2} Y.$$

Clearly, for self-adjoint operators (observables) A and B,

$$\operatorname{Corr}_\rho(A, A) = I(\rho, A), \qquad \operatorname{Corr}_\rho(B, B) = I(\rho, B).$$

From

$$
\begin{aligned}
\tfrac{1}{2}(I(\rho, A+B) - I(\rho, A-B)) &= \operatorname{Corr}_\rho(A, B) + \operatorname{Corr}_\rho(B, A) \\
\langle [A, B] \rangle_\rho &= \operatorname{tr} \rho [A, B] = \operatorname{Corr}_\rho(A, B) - \operatorname{Corr}_\rho(B, A),
\end{aligned}
$$

we obtain

$$2 \operatorname{Corr}_\rho(A, B) = \tfrac{1}{2}(I(\rho, A+B) - I(\rho, A-B)) + \langle [A, B] \rangle_\rho.$$

In the right hand side, the first term is real, while the second term is purely imaginary, therefore

$$4\,|\mathrm{Corr}_\rho(A,B)|^2 = |\langle [A,B]\rangle_\rho|^2 + \tfrac{1}{4}\,(I(\rho, A+B) - I(\rho, A-B))^2,$$

but by the Schwarz inequality,

$$|\mathrm{Corr}_\rho(A,B)|^2 \leqslant \mathrm{Corr}_\rho(A,A)\cdot \mathrm{Corr}_\rho(B,B) = I(\rho,A)\cdot I(\rho,B).$$

The conclusion follows.

Remark. The term $\langle [A,B]\rangle_\rho$ is related to the incompatibility (non-commutativity) between A and B, while the term

$$\tfrac{1}{2}\,(I(\rho, A+B) - I(\rho, A-B)) = \mathrm{Corr}_\rho(A,B) + \mathrm{Corr}_\rho(B,A)$$

is the correlation between A and B expressed in terms of the Wigner–Yanase skew information or the Wigner–Yanase correlation. When ρ is a pure state, it reduces to the conventional covariance.

In view of Theorem 1, we conclude that Theorem 2 implies the conventional Heisenberg's uncertainty relations (1) involving variances. When ρ is a pure state, it reduces to the original Schrödinger's uncertainty relations (2). When ρ is a mixed state, Theorem 2 establishes a new uncertainty relation in the spirit of Schrödinger in the sense that it encodes correlation information, and it is strictly stronger than Schrödinger's uncertainty relations (2) at least for two-dimensional quantum systems, and for the situations described by all examples in this paper. However, whether it is generally stronger remains open. We propose the following conjecture which we strongly believe to be true:

Conjecture. It holds that

$$\begin{aligned}&\mathrm{Var}_\rho\, A\cdot \mathrm{Var}_\rho\, B - \tfrac{1}{4}\,|\langle \{A_0, B_0\}\rangle_\rho|^2\\ &\qquad \geqslant I(\rho,A)\cdot I(\rho,B) - \tfrac{1}{16}\,|I(\rho, A+B) - I(\rho, A-B)|^2,\end{aligned}$$

or equivalently,

$$\begin{aligned}&\mathrm{Var}_\rho\, A\cdot \mathrm{Var}_\rho\, B - \tfrac{1}{4}\,|\mathrm{Cov}_\rho(A,B) + \mathrm{Cov}_\rho(B,A)|^2\\ &\qquad \geqslant I(\rho,A)\cdot I(\rho,B) - \tfrac{1}{4}\,|\mathrm{Corr}_\rho(A,B) + \mathrm{Corr}_\rho(B,A)|^2. \qquad (9)\end{aligned}$$

If this conjecture is true, then Theorem 2 will imply Schrödinger's uncertainty relations (2). The conjecture is supported by the following concrete cases:

(1) If either A or B commute with ρ, then clearly, $I(\rho, A) = 0$ or $I(\rho, B) = 0$, and $\mathrm{Corr}_\rho(A, B) = 0$, thus in this circumstance, the conjecture is true.

(2) It can be verified directly that the above conjecture holds when ρ, A, and B are taken as in Examples 1–4.

In fact, in Examples 1 and 2, we have

$$\begin{aligned}
\mathrm{Var}_\rho\, A = \mathrm{Var}_\rho\, B = 1, &\qquad \mathrm{Cov}_\rho(A, B) = \mathrm{Cov}_\rho(B, A) = 1, \\
I(\rho, A) = I(\rho, B) = 0, &\qquad \mathrm{Corr}_\rho(A, B) = \mathrm{Corr}_\rho(B, A) = 0. \\
\mathrm{Var}_{\rho'}\, A = \mathrm{Var}_{\rho'}\, B = 1, &\qquad \mathrm{Cov}_{\rho'}(A, B) = \mathrm{Cov}_\rho(B, A) = 1, \\
I(\rho', A) = I(\rho', B) = 1, &\qquad \mathrm{Corr}_{\rho'}(A, B) = \mathrm{Corr}_{\rho'}(B, A) = 1.
\end{aligned}$$

In Example 3, we have

$$\mathrm{Var}_\rho\, A = \mathrm{Var}_\rho\, B = 1, \qquad \mathrm{Cov}_\rho(A, B) = \mathrm{Cov}_\rho(B, A) = 0,$$
$$I(\rho, A) = I(\rho, B) = \mathrm{Corr}_\rho(A, B) = \mathrm{Corr}_\rho(B, A) = \tfrac{1}{2}.$$

In Example 4, we have

$$\mathrm{Var}_\rho\, A = \mathrm{Var}_\rho\, B = 1, \qquad \mathrm{Cov}_\rho(A, B) = \mathrm{Cov}_\rho(B, A) = p,$$
$$I(\rho, A) = I(\rho, B) = \mathrm{Corr}_\rho(A, B) = \mathrm{Corr}_\rho(B, A) = \tfrac{1}{2}\left(1 + p - \sqrt{(1+3p)(1-p)}\right).$$

We see readily the conjecture holds true for all these cases.

(3) When the quantum system Hilbert space is two dimensional, and thus ρ, A and B are all self-adjoint operators in C^2, the conjecture is true.

To prove this, due to the covariant transformation properties

$$\begin{aligned}
\mathrm{Var}_{U\rho U^{-1}}\, A = \mathrm{Var}_\rho(U^{-1}AU), &\quad \mathrm{Cov}_{U\rho U^{-1}}(A, B) = \mathrm{Cov}_\rho(U^{-1}AU, U^{-1}BU) \\
I(U\rho U^{-1}, A) = I(\rho, U^{-1}AU), &\quad \mathrm{Corr}_{U\rho U^{-1}}(A, B) = \mathrm{Corr}_\rho(U^{-1}AU, U^{-1}BU),
\end{aligned}$$

we may assume that ρ is diagonal (otherwise, diagonalizing ρ and absorbing the unitary transformation into the observables) and non-degenerate (otherwise ρ will be a pure state in C^2 and (9) becomes an equality since $\mathrm{Var}_\rho\, A = I(\rho, A)$, $\mathrm{Cov}_\rho(A, B) = \mathrm{Corr}_\rho(A, B)$, etc, for pure state ρ). Thus it suffices to check inequality (9) for

$$\rho = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad A = \begin{pmatrix} a_{11} & a \\ \bar{a} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b \\ \bar{b} & b_{22} \end{pmatrix}.$$

Here $\lambda_1+\lambda_2=1$, $\lambda_1>0$, $\lambda_2>0$, and $a_{11}, a_{22}, b_{11}, b_{22}$ are all real numbers, while a and b may be complex. Due to the invariance of all the quantities in (9) when A and B are translated by any constant real numbers, we may assume without loss of generality that $\operatorname{tr} \rho A=0$, $\operatorname{tr} \rho B=0$. These imply that

$$\lambda_1 a_{11}+\lambda_2 a_{22}=0, \qquad \lambda_1 b_{11}+\lambda_2 b_{22}=0.$$

From the above relations and noting the scaling property (inequality (9) is invariant when A or B is multiplied by any real number), if $a_{ii}\neq 0$, $b_{ii}\neq 0$ (the zero coefficients cases can be verified directly or proved by a continuity argument), we may further assume that

$$a_{11}=b_{11}=\lambda_2, \qquad a_{22}=b_{22}=-\lambda_1.$$

Now put $\Lambda=\sqrt{\lambda_1\lambda_2}$, we can easily compute

$$\operatorname{Var}_\rho A=\Lambda^2+|a|^2, \qquad \operatorname{Var}_\rho B=\Lambda^2+|b|^2$$
$$\operatorname{Cov}_\rho(A, B)+\operatorname{Cov}_\rho(B, A)=2\Lambda^2+2\operatorname{Re}(a\bar{b}),$$

and

$$I(\rho, A)=(1-2\Lambda)\,|a|^2, \qquad I(\rho, B)=(1-2\Lambda)\,|b|^2$$
$$\operatorname{Corr}_\rho(A, B)+\operatorname{Corr}_\rho(B, A)=(1-2\Lambda)\,2\operatorname{Re}(a\bar{b}).$$

Consequently, we have

$$\operatorname{Var}_\rho A\cdot\operatorname{Var}_\rho B-\tfrac{1}{4}\,|\operatorname{Cov}_\rho(A, B)+\operatorname{Cov}_\rho(B, A)|^2=\Lambda^2\,|a-b|^2+|\operatorname{Im}(a\bar{b})|^2,$$
$$I(\rho, A)\cdot I(\rho, B)-\tfrac{1}{4}\,|\operatorname{Corr}_\rho(A, B)+\operatorname{Corr}_\rho(B, A)|^2=(1-2\Lambda)^2\,|\operatorname{Im}(a\bar{b})|^2.$$

We have used Re and Im to denote the real and imaginary part, respectively, of a complex number. Since $0\leqslant\Lambda\leqslant 1/2$, inequality (9) follows.

5. DISCUSSIONS

The statistical origin and measurement-theoretic significance of the Wigner–Yanase skew information are investigated from a statistical inference perspective: The skew information is a kind of quantum Fisher information, and it is a useful notion in quantifying informational aspect of quantum measurement, in particular, the uncertainty relations.

We have introduced a new notion, the Wigner–Yanase correlation, to characterize correlation and entanglement. We have shown that it has certain advantages over the conventional covariance.

We have established an informational characterization of Schrödinger's uncertainty relations by virtue of the Wigner–Yanase skew information. The motivation for this characterization comes from the desire to put the uncertainty principle in a more fundamental place, and to make it more useful in a quantitative treatment of quantum measurement.

Among many informational approaches to physics, a particularly appealing and fruitful one is to consider quantum theory as a theory of statistical inference based on observed experimental data.[9, 26, 29] This approach depends crucially on various information quantities (such as the Shannon entropy, the Fisher information and quantum Fisher information) synthesizing the structural properties of the experimental data. Our informational characterization of Schrödinger's uncertainty relations sheds new light on the statistical inference aspect of quantum theory.

ACKNOWLEDGMENTS

The authors are very grateful to the referee for helpful suggestions and critical comments. Luo is supported by NSF Grant 10131040.

REFERENCES

1. H. Araki and M. M. Yanase, Measurement of quantum mechanical operators, *Phys. Rev.* **120**:622–626 (1960).
2. I. Bialynicki-Birula and J. Mycielski, Uncertainty relations for information entropy in wave mechanics, *Comm. Math. Phys.* **44**:129–132 (1975).
3. E. A. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.* **101**:194–211 (1991).
4. E. A. Carlen and A. Soffer, Entropy production by block variable summation and central limit theorems, *Comm. Math. Phys.* **140**:339–371 (1991).
5. A. Connes and E. Stormer, Homogeneity of the state space of factors of type III_1, *J. Funct. Anal.* **28**:187–196 (1978).
6. H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1974).
7. R. A. Fisher, Theory of statistical estimation, *Proc. Camb. Phil. Soc.* **22**:700–725 (1925).
8. G. B. Folland and A. Sitaram, The uncertainty principle: A mathematical survey, *J. Fourier Anal. Appl.* **3**:207–236 (1997).
9. B. R. Frieden, *Physics from Fisher Information, A Unification* (Cambridge University Press, Cambridge, 1998).
10. M. J. W. Hall, Exact Heisenberg uncertainty relations, *Phys. Rev. A* **64**:052103 (2001).
11. M. J. W. Hall and M. Reginatto, Schrödinger equation from an exact uncertainty principle, *J. Phys. A* **35**:3289–3303 (2002).

12. W. Heisenberg, Über den anschaulichen Inhalt der quantumtheoretischen Kinematik und Mechanik, *Z. Phys.* **43**:172–198 (1927).
13. C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
14. A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
15. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd edn. (Butterworth-Heinemann, Oxford, 1977), p. 2.
16. E. H. Lieb, Convex trace functions and the Wigner–Yanase–Dyson conjecture, *Adv. Math.* **11**:267–288 (1973).
17. S. L. Luo, Quantum Fisher information and uncertainty relations, *Lett. Math. Phys.* **53**:243–251 (2000).
18. S. L. Luo, Fisher information matrix of Husimi distribution, *J. Stat. Phys.* **102**:1417–1428 (2001).
19. S. L. Luo, Maximum Shannon entropy, minimum Fisher information, and an elementary game, *Found. Phys.* **32**:1757–1772 (2002).
20. N. A. Nielsen and I. L. Chung, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
21. D. Petz, Monotone metrics on matrix spaces, *Linear Alg. Appl.* **244**:81–96 (1996).
22. H. P. Robertson, The uncertainty principle, *Phys. Rev.* **34**:163–164 (1929).
23. E. Schrödinger, About Heisenberg uncertainty relation, *Proc. Prussian Acad. Sci., Phy-Math. Section* **XIX**:293–303 (1930). See also arXiv: quant-ph/9903100 for an English translation.
24. A. J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inf. Control* **2**:101–112 (1959).
25. A. D. Sukhanov, Schrödinger uncertainty relations and physical features of correlated-coherent states, *Theor. Math. Phys.* **132**:1277–1294 (2002).
26. J. Summhammer, Inference in quantum physics, *Found. Phys. Lett.* **1**:113–137 (1988).
27. A. Wehrl, General concept of entropy, *Rev. Modern Phys.* **50**:221–260 (1978).
28. R. F. Werner and M. M. Wolf, Bell inequalities and entanglement, quant-ph/0107093.
29. J. A. Wheeler, Information, physics, quantum: the search for links, in *Complexity, Entropy, and Physics of Information*, Z. H. Zurek, ed. (Addison–Wesley, MA, 1990), pp. 3–28.
30. H. Weyl, *Gruppentheorie und Quantenmechanik* (Leipzig, 1928), p. 272.
31. E. P. Wigner, Die Messung quantenmechanischer Operatoren, *Z. Phys.* **133**:101–118 (1952).
32. E. P. Wigner and M. M. Yanase, Information contents of distributions, *Proc. Nat. Acad. Sci. USA*, **49**:910–918 (1963).
33. M. M. Yanase, Optimal measuring apparatus, *Phys. Rev.* **123**:666–668 (1961).