A Universal Formulation of Uncertainty Relation for Error and Disturbance

Jaeha Lee[1,*] and Izumi Tsutsui[2,†]

[1] *Institute of Industrial Science, The University of Tokyo, Chiba 277-8574, Japan.*
[2] *Theory Center, Institute of Particle and Nuclear Studies, High Energy Accelerator Research Organization (KEK), Ibaraki 305-0801, Japan.*

(Dated: April 14, 2020)

We present a universal formulation of uncertainty relation valid for any conceivable quantum measurement and their observer effects of statistical nature. Owing to its simplicity and operational tangibility, our general relation is also experimentally verifiable. Our relation violates the naïve bound $\hbar/2$ for the position-momentum measurement while respecting Heisenberg's original philosophy of the uncertainty principle. Our error-disturbance relation is found to be a corollary to our relation for errors, the latter of which also entails the standard Kennard-Robertson (Schrödinger) relation as a special case; this attains a unified picture of the three orthodox realms of uncertainty regarding quantum indeterminacy, measurement, and observer effect within a single framework.

I. INTRODUCTION

Recently, we have seen a consistent — some are rapid while others are more steady — progress of quantum information technologies, and it should now be evident that our future society thrives upon these technologies we are going to develop in coming years. All of the technologies are made available by the application of quantum mechanics, which was established nearly a century ago but still defies our deeper understanding in many respects, such as the non-local correlation arising from quantum entanglement and the non-causal change inherent in quantum measurement. The crux of the matter behind these phenomena is arguably the renunciation of local reality, and in this regard the uncertainty principle has been deemed as the basis to guide us to the proper comprehension of the quantum world. Nevertheless, we have also been aware that the uncertainty principle, which was originally introduced by Heisenberg [1], is too vague to deduce rigorously testable statements, except for providing intuitive arguments which are helpful yet sometimes deceptive and misleading.

The earliest attempt to remove the vagueness was made by Kennard [2] who provided a mathematical formulation of the uncertainty principle in terms of the standard deviations for the pair of position and momentum observables, giving the familiar lower bound $\hbar/2$ for their product. Subsequently, its generalization for arbitrary observables A and B was given by Robertson [3] with the lower bound expressed by the expectation value of the commutator $|\langle[A, B]\rangle_\rho|/2$ for the state ρ under consideration. On account of the mathematical clarity, the Kennard-Robertson inequality became a standard textbook material as an exposition of the uncertainty principle, despite that it has little to do with measurement to which Heisenberg attributed the cause of the uncertainty in his discourse [1].

In fact, as typically exemplified by the famous gamma-ray microscope Gedankenexperiment, for devising his uncertainty principle Heisenberg considered the error in the position measurement and the subsequent disturbance occurred in the momentum, and argued that the product of the two has a lower bound of the order of $\hbar$. It was thus clear that this line of thought, which captures the essence of the quantum 'indeterminatenes' [4], should be followed in order to establish a more genuine uncertainty relation which governs the quantum phenomena on a rigorous basis. This was achieved by Ozawa [5, 6] who adopted the indirect measurement scheme where one considers explicitly the system of an ancillary meter device in addition to the quantum system of interest, realizing the concepts of error and disturbance in concrete terms. His uncertainty relation, obtained for an arbitrary pair of observables A and B, contains not only the product term of the error and the disturbance but also two other terms involving the standard deviations. Whereas these additional terms offer the possibility of 'breaking' the lower bound stated by Heisenberg, they also obscure Heisenberg's original spirit instilled in his principle [7]. Another shortcoming is that it assumes objects that are generally unobtainable from the measurement outcomes [8, 9].

In this paper, we present a novel uncertainty relation for error and disturbance associated with measurement in a conceivably most universal formulation, following our earlier work [10] on the uncertainty relation for errors where (part of) the formulation is adopted. To be more explicit, our formulation assumes only the space of quantum states before measurement and the space of probability distributions obtained by the measurement outcomes, together with the space of quantum states describing the change induced by the measurement, which are seemingly the least possible set of presumptions needed to discuss the uncertainty relation. In particular, it does not require any model of measurements such as the indirect measurement scheme and deals directly with the measurement outcomes. Besides, our formulation allows us to focus exclusively on the error and disturbance to make the uncertainty relation more in line with Heisenberg's original idea, rendering the result free from the shortcomings mentioned above. The universality also ensures the general validity of the resultant uncertainty relation for any measurements including the standard POVM measurements extensively used in quantum information. The most notable benefit of adopting the universal formulation is that it enables the cause of quantum uncertainty clearer; it is the existence of (joint) distributions that constrains the possi-

* lee@iis.u-tokyo.ac.jp
† izumi.tsutsui@kek.jp

ble form of change of states after measurement, and the uncertainty relation arises as the tradeoff relation between the measurement error and the inevitable disturbance.

This paper is organized as follows. After this introduction, in sec.II we present the geometric framework which forms the basis of our formulation. The subsequent three sections are devoted to provide the necessary tools for our formulation, where the general notion of measurements and that of processes describing change of states are spelled out for both classical and quantum systems and also between them. Our error and disturbance are defined in sec.VI, followed by the key observation as to how the tradeoff relation between the errors of two measurements can arise in sec.VII. Our main result, the uncertainty relation for error and disturbance, is then presented in sec.VIII and its affinity with Heisenberg's original idea is discussed in sec.IX. Finally, we provide our conclusion and discussions in sec.X.

II. GEOMETRIC FRAMEWORK

We first briefly summarize the geometric framework proposed in our previous paper [10] that forms the basis of our study.

Let $Z(\mathcal{H})$ be the state space of a quantum system given by the convex set of all the density operators on a Hilbert space $\mathcal{H}$, and also let $S(\mathcal{H})$ be the linear space of quantum observables given by the self-adjoint operators on $\mathcal{H}$. For an observable A, each quantum state $\rho \in Z(\mathcal{H})$ furnishes a seminorm $\|A\|_\rho$ on $S(\mathcal{H})$ inherited from the one defined for linear operators by $\|A\|_\rho := \sqrt{\langle A^\dagger A \rangle_\rho}$, where we have introduced

$$\langle X \rangle_\rho := \mathrm{Tr}[X\rho] \tag{1}$$

for any pair of a linear operator X on $\mathcal{H}$ and $\rho \in Z(\mathcal{H})$. The seminorm induces an equivalence relation $A \sim_\rho B \iff \|A - B\|_\rho = 0$ over $S(\mathcal{H})$, which results in the partitioning of the observables into equivalence classes. These equivalence classes of quantum observables collectively form a quotient space $S(\mathcal{H})/{\sim_\rho}$, the completion of which we denote it by $S_\rho(\mathcal{H})$; this can be visualized as a 'tangent space' attached to the point $\rho \in Z(\mathcal{H})$ of the state space. As commonly practiced, we make an abuse of notation to denote the equivalence class with one of its representatives. The space $S_\rho(\mathcal{H})$ bears a unique inner product $\langle A, B \rangle_\rho := \langle \{A, B\} \rangle_\rho / 2$ characterized by the anti-commutator $\{A, B\} := AB + BA$ of observables, which is compatible with the quotient norm in the sense of $\|A\|_\rho^2 = \langle A, A \rangle_\rho$. We call the resultant bundle consisting of the base space $Z(\mathcal{H})$ with the fibres $S_\rho(\mathcal{H})$ attached to each of the points $\rho \in Z(\mathcal{H})$ a *quantum system*. Needless to say, our target system for which a measurement will be made is a quantum system.

The classical counterpart can be constructed in a parallel manner. For this, let $W(\Omega)$ be the convex set of all the probability distributions p on a sample space Ω, and let $R(\Omega)$ be the linear space of all the real functions on Ω. Then one has the seminorm $\|f\|_p$ for each $f \in R(\Omega)$ inherited from that defined for complex functions by $\|f\|_p := \sqrt{\langle f^\dagger f \rangle_p}$, where, analogously to (1), we have introduced

$$\langle z \rangle_p := \int_\Omega z(\omega) p(\omega)\, d\omega \tag{2}$$

for any pair of a complex function z on Ω and $p \in W(\Omega)$. Like before, the seminorm induces an equivalence relation $f \sim_p g \iff \|f - g\|_p = 0$ over the space $R(\Omega)$ which, in turn, induces a quotient space $R(\Omega)/{\sim_p}$. Its completion yields a tangent space at each $p \in W(\Omega)$ which will be denoted by $R_p(\Omega)$. It is easy to check that the quotient norm admits a unique inner product $\langle f, g \rangle_p := \langle fg \rangle_p$ that satisfies $\|f\|_p^2 = \langle f, f \rangle_p$. We call the resultant bundle consisting of the base space $W(\Omega)$ with the fibres $R_p(\Omega)$ attached to each of the points $p \in W(\Omega)$ a *classical system*, which in the present paper will be used to represent the probability space of measurement outcomes.

III. PROCESSES

To describe the measurement that assigns a probability distribution to a given quantum state and also the change of states it induces, our primary objects to be used are affine maps between state spaces, which we call *processes* in general terms. To our purposes, we are specifically interested in three types, namely, those from quantum state spaces to classical state spaces $M : Z(\mathcal{H}) \to W(\Omega)$, as well as those between quantum state spaces $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ and those between classical state spaces $K : W(\Omega_1) \to W(\Omega_2)$. Among these, the first quantum-to-classical (Q-C) process M will be called a *quantum measurement*, whereas the second quantum-to-quantum (Q-Q) process Θ will be simply called a *quantum process* for obvious reasons (see FIG. 1). The third classical-to-classical (C-C) process K, which will be used for examining compatibility of two measurements in this paper, is called a *classical process*, which may be thought of as a classical measurement describing the statistical outcome of measurement on a classical system.

The archetype of quantum measurements is the projection measurement associated with a quantum observable $\hat{M}$. Assuming for simplicity that $\hat{M}$ is non-degenerate in a finite N-dimensional space $\mathcal{H}$, the spectral decomposition $\hat{M} = \sum_{i=1}^{N} m_i \, |m_i\rangle\langle m_i|$ induces the affine map

$$M : \rho \mapsto (M\rho)(m_i) := \mathrm{Tr}\big[|m_i\rangle\langle m_i|\rho\big]. \tag{3}$$

To this map one can associate the interpretation that $p(m_i) = (M\rho)(m_i)$ gives the probability distribution of possible measurement outcomes m_i in the sample space $\Omega := \{m_1, \ldots, m_N\}$ of all the eigenvalues of $\hat{M}$ in accordance with the Born rule. Throughout this paper, the reader may safely assume that quantum measurement M to be that of the familiar projection measurement without missing much of the essence of the subject, although our M is by no means restricted to that particular class. In fact, the map M should in general be regarded as completely independent from the physical observable one wishes to measure, which is required typically when one measures two distinct observables A and B with a single quantum measurement process M.

Analogously, the archetype of quantum processes we are interested in is the 'wave-function collapse' which is usually associated with the projection measurement. This is characterized under the projection postulate [11, 12] such that the initial quantum state $|\psi\rangle$ over which the measurement of $\hat{M}$ is performed collapses to one of its eigenvectors $|m_i\rangle$ with probability $|\langle m_i|\psi\rangle|^2$. The projection postulate uniquely extends to mixed quantum states ρ with the probability $p(m_i)$ mentioned above, resulting in the affine map

$$\Theta : \rho \mapsto \Theta\rho := \sum_{i=1}^{N} \mathrm{Tr}\big[|m_i\rangle\langle m_i|\rho\big] \cdot |m_i\rangle\langle m_i|, \quad (4)$$

from the quantum state space $Z(\mathcal{H})$ to itself. Again, throughout this paper, the reader may safely assume that the quantum process Θ to be that induced by the wave-function collapse, although our map Θ may describe state changes much more general than that.

In fact, the sole constraint we impose on the map is affineness, *i.e.*, the map preserves the structure of the probabilistic mixture

$$\Theta(\lambda\rho_1 + (1-\lambda)\rho_2) = \lambda\Theta\rho_1 + (1-\lambda)\Theta\rho_2 \quad (5)$$

for $\rho_1, \rho_2 \in Z(\mathcal{H})$, $0 \le \lambda \le 1$, which is indispensable for the self-consistent statistical interpretation of density operators, ensuring that the resultant quantum state $\kappa = \Theta\rho$ be invariant under every (pure-state) decomposition of a mixed quantum state ρ. In other words, our Θ effectively belongs to the broadest class of maps between quantum state spaces describing the most general change of states preserving statistical nature, including, *e.g.*, unitary evolution of closed quantum systems, non-unitary evolution of open quantum systems, quantum decoherence, observer effects (which is our main interest here), quantum channels and gates; we also note that even the completely positivity of the (adjoint of the) quantum process need not be assumed.

IV. ADJOINT OF PROCESS

In what follows, we focus our attention primarily to quantum processes in order to avoid unnecessary repetition. The same properties and facts described hereafter are also valid for any other processes, including quantum measurements [10] as well as classical processes, all of which can be demonstrated through a similar line of reasoning.

Now, an important observation is that a quantum process $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ uniquely induces a map Θ' between operator spaces. This dual notion of a quantum process, termed its adjoint, is uniquely characterized by the relation

$$\langle \Theta' X \rangle_\rho = \langle X \rangle_{(\Theta\rho)} \quad (6)$$

valid for all operators X on $\mathcal{K}$ and quantum states ρ on $\mathcal{H}$, where we have used (1) defined analogously for the two cases. Again, the wave-function collapse (4) under the projection postulate provides a prime example, the adjoint of which can be confirmed to read

$$\Theta' : X \mapsto \Theta' X = \sum_{i=1}^{N} \mathrm{Tr}\big[|m_i\rangle\langle m_i|X\big] \cdot |m_i\rangle\langle m_i|, \quad (7)$$

which fulfills (6). The projection postulate is convenient in that they admit concrete expressions both for the quantum process (4) and its adjoint (6) using familiar notions, allowing for the verification of the various claims in this paper by means of direct computation.

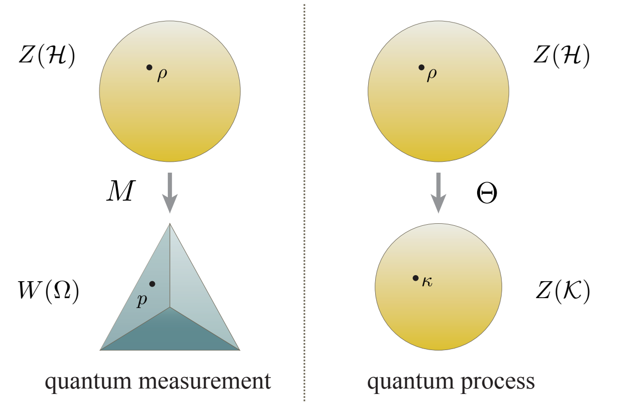

FIG. 1. Our basic premise of quantum measurements and quantum processes. The space of quantum states (density operators) $Z(\mathcal{H})$ is depicted as a sphere, while the space of probability distributions $W(\Omega)$ is represented by a tetrahedron. In general, a quantum measurement M can be regarded as a map $M : Z(\mathcal{H}) \to W(\Omega)$ attaching a probability distribution $p \in W(\Omega)$ to a given quantum state $\rho \in Z(\mathcal{H})$. Similarly, a quantum process Θ can be regarded as a map $M : Z(\mathcal{H}) \to Z(\mathcal{K})$ attaching a quantum state $\kappa \in Z(\mathcal{K})$ to a given quantum state $\rho \in Z(\mathcal{H})$.

V. PUSHFORWARD AND PULLBACK

In our framework, it is important to recognize that every process, which is a global map between state spaces, gives rise to an adjoint pair of *local* (*i.e.*, state-dependent) maps. To see this, we first note that regarding quantum processes $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ we have the inequality

$$\|A\|_{\Theta\rho} \ge \|\Theta' A\|_\rho \quad (8)$$

for any quantum observable A on $\mathcal{K}$ and quantum state ρ on $\mathcal{H}$. This can be understood as a corollary to the Kadison-Schwarz inequality [13], which is in a sense a generalization of the Cauchy-Schwarz inequality to C*-algebras. Indeed, its application to the adjoint Θ' yields the evaluation $\Theta'(N^\dagger N) \ge (\Theta' N)^\dagger(\Theta' N)$ valid for any normal operator N on $\mathcal{K}$. Since self-adjoint operators are normal operators, we have (8).

An immediate consequence of (8) is the implication $A \sim_{\Theta\rho} B \implies \Theta' A \sim_\rho \Theta' B$. This allows the adjoint Θ', which was initially introduced as a map between Hilbert space operators,

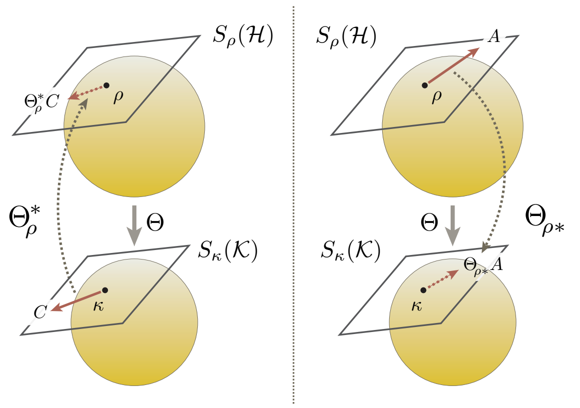

FIG. 2. The pullback and the pushforward of the quantum process. (Left) A quantum process Θ entails the pullback Θ_ρ^* from $S_\kappa(\mathcal{K})$ to $S_\rho(\mathcal{H})$, each of which is attached to the respective points $\kappa = \Theta\rho \in Z(\mathcal{K})$ and $\rho \in Z(\mathcal{H})$ of the corresponding state spaces. (Right) Conversely, Θ also entails the pushforward that maps in the opposite direction. The pullback and the pushforward are dual to each other through the relation (12), and both of them are contractions, that is, the norm decreases (or remains unchanged) under each of the maps.

to be passed to the map between their equivalence classes. We call the resultant map

$$\Theta_\rho^* : S_{(\Theta\rho)}(\mathcal{K}) \to S_\rho(\mathcal{H}) \tag{9}$$

the *pullback of the quantum process* Θ over the quantum state ρ. In concrete terms, this implies that, given (the equivalence class of) an operator $A \in S_{(\Theta\rho)}(\mathcal{K})$, we have (that of) a corresponding operator $\Theta_\rho^* A \in S_\rho(\mathcal{H})$. Among the various properties of the pullback, the contractivity $\|A\|_{\Theta\rho} \geq \|\Theta_\rho^* A\|_\rho$ as well as the composition law described shortly are of our particular interest. In fact, contractivity is trivial by construction. For the latter, let $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ and $\Phi : Z(\mathcal{K}) \to Z(\mathcal{L})$ be two quantum processes with $\mathcal{L}$ being a Hilbert space. Since affineness is closed under map composition, the composite map $\Phi \circ \Theta : Z(\mathcal{H}) \to Z(\mathcal{L})$ is itself a quantum process. Then, the pullback of the composite process can be shown to satisfy the composition law

$$(\Phi \circ \Theta)_\rho^* = \Theta_\rho^* \circ \Phi_{(\Theta\rho)}^* \tag{10}$$

by means of straightforward computation.

It now remains to introduce the dual notion of the pullback, which we call the *pushforward of the quantum process* Θ. The pushforward

$$\Theta_{\rho*} : S_\rho(\mathcal{H}) \to S_{(\Theta\rho)}(\mathcal{K}) \tag{11}$$

of the quantum process Θ over the quantum state ρ is defined as the adjoint of the pullback (9) regarding the inner products on the localized observable spaces. More explicitly, the pushforward (11) is uniquely characterized by the relation

$$\left\langle A, \Theta_\rho^* C \right\rangle_\rho = \left\langle \Theta_{\rho*} A, C \right\rangle_{\Theta\rho} \tag{12}$$

valid for any choices of $A \in S_\rho(\mathcal{H})$ and $C \in S_{(\Theta\rho)}(\mathcal{K})$ on their respective systems. Since the pullback is a contraction, its adjoint, *i.e.*, the pushforward, is also a contraction $\|A\|_\rho \geq \|\Theta_{\rho*} A\|_{\Theta\rho}$. It is also easy to confirm the validity of the composition law

$$(\Phi \circ \Theta)_{\rho*} = \Phi_{(\Theta\rho)*} \circ \Theta_{\rho*} \tag{13}$$

of the pushforward (see FIG. 2).

We note that, by induction, the composition laws of both the pullback (10) and the pushforward (13) admit generalizations to the composition of any number of arbitrary processes in an obvious manner.

VI. ERROR AND DISTURBANCE

Armed with our geometric framework, earlier we have introduced [10] our definition of (quantum) error by the amount of contraction induced by the pushforward of the measurement M,

$$\varepsilon_\rho(A; M) := \sqrt{\|A\|_\rho^2 - \|M_{\rho*} A\|_{M\rho}^2} \tag{14}$$

for the observable A and the quantum state ρ. In a parallel manner, we now introduce our definition of disturbance

$$\eta_\rho(A; \Theta) := \sqrt{\|A\|_\rho^2 - \|\Theta_{\rho*} A\|_{\Theta\rho}^2} \tag{15}$$

associated with a quantum process Θ with respect to an observable A over ρ. Here, our error and disturbance are defined in such a way that both are of essentially the same nature, each representing the amount of 'loss' induced by the respective processes. It goes without saying that, in general, our method of quantifying the loss by the amount of contraction is applicable to any types of processes, including not only quantum measurements and quantum processes as we have seen above, but also to classical processes.

Before proceeding, we note here that the non-negativity $\eta_\rho(A; M) \geq 0$ of the disturbance is guaranteed by the contractivity of the pushforward mentioned before. It is also easy to check the homogeneity $\eta_\rho(tA; M) = |t|\, \eta_\rho(A; M), \forall t \in \mathbb{R}$ and the subadditivity $\eta_\rho(A; M) + \eta_\rho(B; M) \geq \eta_\rho(A+B; M)$ of the disturbance. These show that the disturbance furnishes a seminorm on the state-dependent spaces of quantum observables (the same properties hold verbatim for the 'loss' of any types of processes).

Just as is the case for our error [10], our disturbance also admits an operational interpretation. To expound on this, let $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ be a quantum process, and let $L : Z(\mathcal{K}) \to W(\Omega)$ be any quantum measurement performed on the resultant space $Z(\mathcal{K})$. Since affineness is closed under map compositions, the composite map $L \circ \Theta : \rho \mapsto L(\Theta\rho)$ is again an affine map from $Z(\mathcal{H})$ to $W(\Omega)$; in other words, the composite map is a quantum measurement. One then readily confirms that the square of the error of the composite quantum measurement $L \circ \Theta$ admits a decomposition

$$\begin{aligned}
\varepsilon_\rho(A; L \circ \Theta)^2 &= \|A\|_\rho^2 - \|(L \circ \Theta)_* A\|_{(L\circ\Theta)\rho}^2 \\
&= \|A\|_\rho^2 - \|\Theta_* A\|_{\Theta\rho}^2 \\
&\quad + \|\Theta_* A\|_{\Theta\rho}^2 - \|L_*(\Theta_* A)\|_{L(\Theta\rho)}^2 \\
&= \eta_\rho(A; \Theta)^2 + \varepsilon_{(\Theta\rho)}(\Theta_* A; L)^2
\end{aligned} \tag{16}$$

into the sum of the square of the disturbance of the initial process Θ and that of the error of the secondary measurement L. Here, we have used the composition law (13) of the pushforward in the second equality adopting the abbreviated notation $\Theta_* = \Theta_{\rho *}$ for brevity. In what follows, we shall also use $\Theta^* = \Theta_\rho^*$ (and similarly $M_* = M_{\rho *}$ and $M^* = M_\rho^*$ for quantum measurement) for all occasions for the same reason.

Due to the fact that every element of the space $S_\rho(\mathcal{H})$ admits a (sequence of) measurement(s) that is capable of measuring it errorlessly (in the limit), the decomposition (16) allows for an operational characterization of the disturbance as the infimum

$$\eta_\rho(A; \Theta) = \inf_L \varepsilon_\rho(A; L \circ \Theta) \tag{17}$$

of the error of the composite measurement, along with the interpretation of the pushforward $\Theta_* A$ as the indicator of the locally optimal choices of the (sequence of) secondary measurement(s) L that attains the infimum (in the limit).

By the same reasoning with the help of induction, we note that the decomposition law (16) admits obvious generalizations to the composition of any number of arbitrary processes. This reveals that our error (or 'loss' induced by general processes) reflects the order structure of 'informativeness' of the quantum measurements (or general processes). To be more explicit, for example, given two quantum measurements $M : Z(\mathcal{H}) \to W(\Omega_1)$ and $N : Z(\mathcal{H}) \to W(\Omega_2)$, one may argue that the former is more informative than the latter if there exists a classical process $K : W(\Omega_1) \to W(\Omega_2)$ with which the behavior of the latter can be fully described by the former in the sense of $N = K \circ M$, which we may write $M \succeq N$. The decomposition law (16) then dictates that the error of the composite measurement is never less than the error of the component, which leads to $\varepsilon_\rho(A; N) \geq \varepsilon_\rho(A; M)$ for any observable A over every state ρ. Thus, our error (or 'loss' induced by general processes) is an (in the current form, reversed) order-homomorphism (order-preserving/isotone map) reflecting the (partial) order between quantum measurements (or general processes) regarding 'informativeness'. The full discussion on this topic is beyond the scope of this paper, and thus shall be elaborated elsewhere in one of our subsequent papers.

VII. UNCERTAINTY RELATION FOR ERRORS IN MEASUREMENTS ADMITTING A JOINT DISTRIBUTION

Prior to the introduction of our uncertainty relation for error and disturbance, we present a refined form of our previous inequality [10] for quantum measurements involving errors, as it serves as a basis for deriving our inequality involving error and disturbance. Let $M : Z(\mathcal{H}) \to W(\Omega_1)$ and $N : Z(\mathcal{H}) \to W(\Omega_2)$ be two quantum measurements chosen independently. Suppose that the two measurements admit a joint description in the sense that there exists an affine map $J : Z(\mathcal{H}) \to W(\Omega_1 \times \Omega_2)$ from which both the distributions $M\rho$ and $N\rho$ are retrieved as marginals from the distribution $J\rho$. More explicitly, this implies that one has the classical processes $\pi_1 : W(\Omega_1 \times \Omega_2) \to W(\Omega_1)$ and $\pi_2 : W(\Omega_1 \times \Omega_2) \to W(\Omega_2)$ that project the joint probability distributions to their respective marginals

$$(\pi_1 p)(\omega_1) := \int_{\Omega_2} p(\omega_1, \omega_2)\, d\omega_2 \tag{18}$$

$$(\pi_2 p)(\omega_2) := \int_{\Omega_1} p(\omega_1, \omega_2)\, d\omega_1 \tag{19}$$

satisfying $M = \pi_1 \circ J$ and $N = \pi_2 \circ J$. Here, since the adjoints of (18) and (19) respectively read $(\pi_1' f)(x, y) = f(x)$ and $(\pi_2' g)(x, y) = g(y)$, their pullbacks are isometries, *i.e.*, $\|f\|_{\pi_1 p} = \|\pi_1^* f\|_p$ and $\|g\|_{\pi_2 p} = \|\pi_2^* g\|_p$. In other words, this allows for the identification of the function spaces $R_{M\rho}(\Omega_1)$ and $R_{N\rho}(\Omega_2)$ regarding each of the measurements with their images under the pullbacks π_1^* and π_2^*, which are in turn subspaces of the larger space $R_{J\rho}(\Omega_1 \times \Omega_2)$ of the joint measurement.

Now that we have equipped ourselves with the necessary concepts and facts, we present our result. Let A and B be quantum observables, and ρ be a quantum state on $\mathcal{H}$. Then, for any pair of quantum measurements M and N admitting a joint description, the inequality

$$\varepsilon_\rho(A; M)\, \varepsilon_\rho(B; N) \geq \sqrt{\mathcal{R}^2 + \mathcal{I}^2} \tag{20}$$

holds, where

$$\begin{aligned} \mathcal{R} := \left\langle \frac{\{A, B\}}{2} \right\rangle_\rho - \langle M_* A, M_* B \rangle_{M\rho} \\ - \langle N_* A, N_* B \rangle_{N\rho} + \langle M_* A, N_* B \rangle_{J\rho} \end{aligned} \tag{21}$$

and

$$\mathcal{I} := \left\langle \frac{[A, B]}{2i} \right\rangle_\rho - \left\langle \frac{[M^* M_* A, B]}{2i} \right\rangle_\rho - \left\langle \frac{[A, N^* N_* B]}{2i} \right\rangle_\rho, \tag{22}$$

with the commutator $[A, B] := AB - BA$. Here, we have also introduced the abbreviated notation

$$\langle M_* A, N_* B \rangle_{J\rho} := \langle \pi_1^* M_* A, \pi_2^* N_* B \rangle_{J\rho} \tag{23}$$

under the identifications $f \simeq \pi_1^* f$ and $g \simeq \pi_2^* g$ mentioned above.

The proof of the inequality (20) is actually quite simple: it is just a direct corollary of the Cauchy-Schwarz inequality. A quick way to see this is to first introduce the semi-inner product

$$\begin{aligned} \langle (X, f), (Y, g) \rangle := \langle X^\dagger Y \rangle_\rho + \langle f^\dagger g \rangle_{J\rho} \\ - \langle K' f^\dagger K' g \rangle_\rho \end{aligned} \tag{24}$$

defined on the product space of Hilbert space operators and complex functions, as well as the seminorm $p(X, f) := \sqrt{\langle (X, f), (X, f) \rangle}$ that it induces. Noticing the equalities

$$\varepsilon_\rho(A; M) = p(X_A, f_A), \quad \varepsilon_\rho(B; N) = p(Y_B, g_B), \tag{25}$$

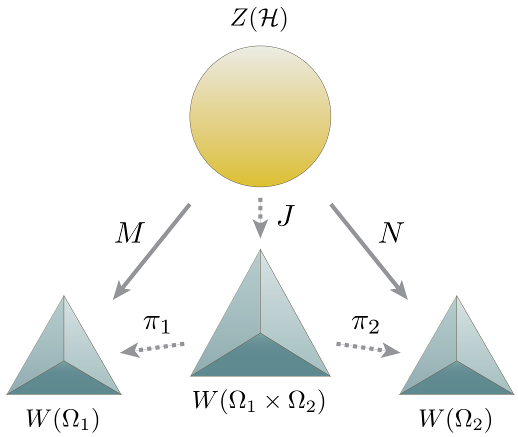

FIG. 3. The basic structure of measurements leading to the uncertainty relation for errors associated with two measurement maps M and N. The cause of the uncertainty lies in the presumption that there exists a quantum measurement J that jointly describes the given two measurements M and N.

where

$$\begin{aligned} X_A &:= A - M^* M_* A, \quad & f_A &:= \pi_1^* M_* A, \\ Y_B &:= B - N^* N_* B, \quad & g_B &:= \pi_2^* N_* B, \end{aligned} \tag{26}$$

we find that the Cauchy-Schwartz inequality for the product of $p(X_A, f_A)$ and $p(Y_B, g_B)$ becomes

$$\varepsilon_\rho(A; M)\, \varepsilon_\rho(B; N) \geq |\langle (X_A, f_A), (Y_B, g_B) \rangle|. \tag{27}$$

The semi-inner product appearing in the R.H.S. of (27) is a complex number $\langle (X_A, f_A), (Y_B, g_B) \rangle = \mathcal{R} + i\,\mathcal{I}$ whose real part $\mathcal{R}$ is given by (21) while the imaginary part $\mathcal{I}$ by (22). This completes our proof of the inequality (20).

The terms $\mathcal{R}$ and $\mathcal{I}$ have their own meanings; the former represents the bound that is shared in common with the classical case ('classical bound'), whereas the latter provides an additional contribution to the former preexisting bound, which is characteristic of quantum measurements ('quantum bound'). This is supported by the observation that, if the lower bound is considered for a pair of classical processes $K_i : W(\Omega) \to W(\Omega_i)$, $i = 1, 2$, by an analogous argument, then one finds that $\mathcal{I}$ vanishes and only the real part $\mathcal{R}$ contributes to the bound.

Before proceeding further, we stress that behind the appearance of the lower bound lies the existence of a joint distribution map J for the two given measurements M and N, which is not at all taken for granted in general. An elementary situation where such a map J is guaranteed to exist occurs when one performs a pair of projection measurements M and N respectively associated with observables $\hat{M}$ and $\hat{N}$ on $\mathcal{H}$ that are commutative with one another as operators, *i.e.*, $[\hat{M}, \hat{N}] = 0$. In such case, the quantum measurement $J : \rho \mapsto (J\rho)(m_i, n_j) = \mathrm{Tr}\big[|m_i\rangle\langle m_i||n_j\rangle\langle n_j|\rho\big]$ induced by combining the spectral decompositions $\hat{M} = \sum_i m_i\, |m_i\rangle\langle m_i|$ and $\hat{N} = \sum_j n_i\, |n_j\rangle\langle n_j|$ of the two operators provides an obvious example of the joint distribution map.

Although the current form of presentation suffices for our main purpose, we note that the condition of joint describability of two quantum measurements M and N admits obvious generalization; the essence is the existence of a quantum measurement $J : Z(\mathcal{H}) \to W(\Omega)$ mediating the behaviors of the two measurements M and N with classical processes $\pi_i : W(\Omega) \to W(\Omega_i)$, $i = 1, 2$, *i.e.*, $M = \pi_1 \circ J$ and $N = \pi_2 \circ J$. If the pullbacks of both classical processes π_i happen to be isometries over a certain state, which is always the case globally when both π_i are projections to the marginals as exemplified above, the same line of arguments leads to the same uncertainty relation as (20) valid over the said state. One of the elementary situations in which this is the case is when one measures two observables with a single measurement M; this amounts to the special case $M = N$ which admits the trivial joint description by $J = M$ and trivial classical processes $\pi_i = \mathrm{Id}$, which subsequently yields the uncertainty relation for errors given in [10]. For general cases (in which the pullbacks of π_i are not necessarily isometries), the uncertainty relation still holds in the form of (20) with due modification to the term $\mathcal{R}$.

A proper and comprehensive treatment of such general cases necessitates more elaborate mathematical tools, which is beyond the scope of this paper, and will thus be given elsewhere; there the potential non-uniqueness of the classical bound $\mathcal{R}$, which is dependent to the generally non-unique choice of the map J, will also be investigated. In view of this, we may just consider the simplified version of our inequality (20) which is free from the choice of J and contains only the quantum bound:

$$\varepsilon_\rho(A; M)\, \varepsilon_\rho(B; N) \geq |\mathcal{I}|. \tag{28}$$

VIII. UNCERTAINTY RELATION FOR ERROR AND DISTURBANCE

We are now ready to introduce our uncertainty relation for error and disturbance. Let $M : Z(\mathcal{H}) \to W(\Omega_1)$ be a quantum measurement and let $\Theta : Z(\mathcal{H}) \to Z(\mathcal{K})$ denote its observer effect, *i.e.*, the inevitable quantum process the measurement M causes on the system $\mathcal{H}$ to be measured. Here, we note again that our process Θ need not be confined to those that end in the same quantum system $\mathcal{H}$ as the initial system, but also admit those that result in different systems $\mathcal{K}$. This allows for the description of the measurement process as well as the quantitative evaluation of the resultant disturbance that were previously not quite viable, *e.g.*, measurement through high energy collision involving particle decays. Our disturbance thus entertains deeper implication on the conception of the observer effect than is commonly conceived.

Let us now consider a sequential measurement in which the primary measurement M is followed by a secondary measurement L which is meant to measure the observable B. As such, the secondary measurement is performed after the system underwent the process Θ caused by the former, and hence the corresponding map is given by $L : Z(\mathcal{K}) \to W(\Omega_2)$ where $W(\Omega_2)$ is the space of probability distributions on the outcomes of the measurement. Viewed from the original state

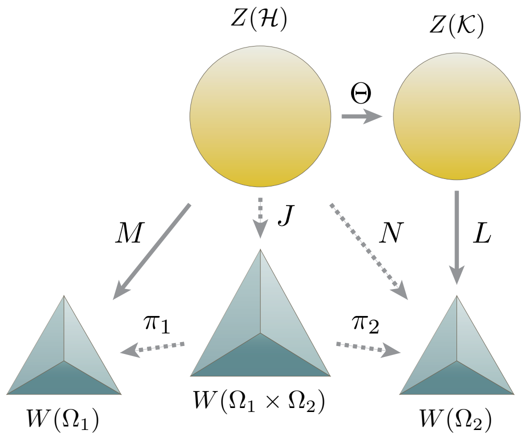

FIG. 4. The structure of generating error and disturbance associated with a measurement. The primary measurement, which is meant to measure the observable A in the system $Z(\mathcal{H})$, induces a map M to the space of probability distributions $W(\Omega_1)$ describing its measurement outcomes. At the same time, the measurement induces a map Θ to the space of quantum states $Z(\mathcal{K})$ describing the resultant states caused by the effect of the measurement. The secondly measurement, which is meant to measure the observable B in the system $Z(\mathcal{K})$, induces a map L to the space of probability distributions $W(\Omega_2)$ describing its measurement outcomes. The map M and the composite map $L \circ \Theta$ provide the pair of maps M and N and their joint distribution $W(\Omega_1 \times \Omega_2)$ discussed before (see FIG. 3).

space $Z(\mathcal{H})$, this sequential measurement by M and L naturally defines a joint measurement $J : Z(\mathcal{H}) \to W(\Omega_1 \times \Omega_2)$ describing both the primary $M = \pi_1 \circ K$ and the secondary composite $L \circ \Theta = \pi_2 \circ K$ measurements with the projections (18) and (19) introduced earlier. Note that the quantum process Θ must be constrained strictly by the measurement map M in order for the composite $L \circ \Theta$ to possess a joint measurement J along with the primary measurement M, which should be the case even for an arbitrary choice of the secondary measurement L.

At this point, we may choose L in such a way that the disturbance of the observable B caused by the process Θ coincides with the error of the composite measurement, that is,

$$\eta_\rho(B; \Theta) = \varepsilon_\rho(B; L \circ \Theta). \tag{29}$$

In fact, we have seen previously in (17) that this choice is not only possible (in the sense of limit, if necessary), but also supported from the ground that it minimizes the error of the composite measurement $L \circ \Theta$ for all possible L. Then, by substituting $N = L \circ \Theta$ in (20), one immediately obtains the inequality for error and disturbance, with $\mathcal{R}$ and $\mathcal{I}$ being respectively given by (21) and (22) with due substitutions for N.

To be more explicit, if we are content with the simplified version (28), then our uncertainty relation for error and disturbance reads

$$\varepsilon_\rho(A; M)\, \eta_\rho(B; \Theta) \geq |\mathcal{I}| \tag{30}$$

with

$$\mathcal{I} = \left\langle \frac{[A, B]}{2i} \right\rangle_\rho - \left\langle \frac{[M^* M_* A, B]}{2i} \right\rangle_\rho - \left\langle \frac{[A, \Theta^* \Theta_* B]}{2i} \right\rangle_\rho . \tag{31}$$

To confirm this, let us first note that the condition (29) is equivalent to $\varepsilon_{(\Theta\rho)}(\Theta_* B; L) = 0$ by (16), which is in turn equivalent to $\Theta_* B = L^* L_* (\Theta_* B)$ as shown in [10]. We then find

$$\begin{aligned} N^* N_* B &= (L \circ \Theta)^* (L \circ \Theta)_* B \\ &= \Theta^* L^* L_* \Theta_* B = \Theta^* \Theta_* B, \end{aligned} \tag{32}$$

where we have used the composition laws of the pullback (10) and the pushforward (13) in the second equality. This allows us to rewrite (22) into (31), which also shows that the lower bound is independent of the choice of L as expected from the definition (29).

IX. AFFINITY WITH HEISENBERG'S ORIGINAL IDEA

Just as with our uncertainty relation for errors [10] (as well as that under joint measurability (20)), our relation for error and disturbance (30) also implies a potential violation of the non-commutativity bound $|\langle [A, B] \rangle_\rho / 2i|$ for certain choices of quantum measurements and their observer effects. It is to be emphasized, however, that even though the product of the error and the disturbance may overcome the non-commutativity bound quantitatively, Heisenberg's original idea of the uncertainty principle remains valid; errorless measurement of an observable A is impossible without disturbing another observable B when $\langle [A, B] \rangle_\rho \neq 0$. We shall now argue why this is the case.

For this, we need to discuss the two separate situations where the measurement M and its observer effect Θ respectively become free from the error and disturbance. The situation for the errorless measurement M is discussed in detail in [10], where we have defined that M performs an errorless measurement of A over ρ, if the error $\varepsilon_\rho(A; M)$ in (14) vanishes. In the same vein, we say that Θ does not disturb A over ρ, if the disturbance $\eta_\rho(A; \Theta)$ in (15) vanishes. Analogous to the case of measurement, several characterizations of the disturbance-less process are possible, and here we note the equivalence of the following three conditions:

(a) $\eta_\rho(A; \Theta) = 0$,

(b) $A = \Theta^* \Theta_* A$,

(c) $\|A\|_\rho = \|\Theta_* A\|_{\Theta\rho} = \|\Theta^* \Theta_* A\|_\rho$.

In fact, (c) $\implies$ (a) is trivial by definition, (a) $\implies$ (b) is an immediate consequence of (16) by choosing L in such a way that $\varepsilon_\rho(\Theta_* A; L) = 0 \iff \Theta_* A = L^* L_* \Theta_* A$ (in the sense of limit, if necessary), and finally (b) $\implies$ (c) is due to the contractivity $\|A\|_\rho \geq \|\Theta_* A\|_{\Theta\rho} \geq \|\Theta^* \Theta_* A\|_\rho = \|A\|_\rho$ of both the pullback and the pushforward.

An immediate corollary of this is that, for a non-commuting pair of observables A and B, there is no quantum measurement that is capable of measuring one of the observables errorlessly without causing disturbance to the other over ρ for which the non-commutativity term $\langle [A,B]\rangle_\rho$ is non-vanishing. Indeed, if there were such a measurement, our uncertainty relation (30) combined with the equivalence (a) $\iff$ (b) would lead us to the contradiction $0 \geq \sqrt{|0|^2 + |\langle [A,B]\rangle_\rho/2i|^2}$, which is easily confirmed by plugging $M^*M_*A = A$ and $\Theta^*\Theta_*B = B$ (or $M^*M_*B = B$ and $\Theta^*\Theta_*A = A$, depending on the choice of the observables concerned) into the lower bound in (30) to find $\mathcal{R} = 0$ while $\mathcal{I} = -\langle [A,B]\rangle_\rho/2i$. One thus finds that for non-trivial (*i.e.*, $\dim(\mathcal{H}) \geq 2$) quantum systems, there exists no quantum measurement without any observer effect.

Note, however, that our formulation does not necessarily prohibit either of the error or disturbance from vanishing. For instance, when $\varepsilon_\rho(A;M) = 0$ and $\eta_\rho(B;\Theta) \neq 0$, one maintains $M^*M_*A = A$ and hence $\mathcal{R} = 0$, and further reveals $\langle [A, \Theta^*\Theta_*B]\rangle_\rho = 0$, which ensures $\mathcal{I} = 0$ and consequently yields no contradiction. This additional property is a consequence of the fact that our quantum process Θ is highly constrained by the measurement M as mentioned before. One may convince oneself that this is indeed the case by looking at the case of M being the projection measurement of A, for which we have $\hat{M} = A$ and the map M reads

$$M : \rho \mapsto (M\rho)(a_i) := \mathrm{Tr}\big[|a_i\rangle\langle a_i|\rho\big] \tag{33}$$

under the spectral decomposition $A = \sum_{i=1}^N a_i\,|a_i\rangle\langle a_i|$. Clearly, this fulfills $M^*M_*A = A$ ensuring the errorless measurement $\varepsilon_\rho(A;M) = 0$, whereas under the projection postulate the map Θ' in (7) becomes

$$\Theta' : X \mapsto \Theta' X = \sum_{i=1}^N \mathrm{Tr}\big[|a_i\rangle\langle a_i|X\big] \cdot |a_i\rangle\langle a_i|, \tag{34}$$

for any operator X on $\mathcal{K}$. One then finds that the pullback of a self-adjoint operator $C \in S(\mathcal{K})$ by Θ^*, which represents a map between the equivalent classes under the local norms, admits the form $\Theta^*C = \Theta'C + D$ with some self-adjoint operator $D \in S(\mathcal{H})$ satisfying $\|D\|_\rho = 0$. It is now immediate to see that $[A, \Theta'C] = 0$ and $\langle [A,D]\rangle_\rho = 0$, which implies $\langle [A, \Theta^*C]\rangle_\rho = 0$ and hence by putting $C = \Theta_*B$, one obtains the aforementioned additional property as promised.

X. CONCLUSION AND DISCUSSIONS

In this paper, we have presented a general uncertainty relation for error and disturbance based on a universal formulation of measurements. We have seen that, given a pair of observables A and B with a state ρ, the product of the measurement error of A and the disturbance caused on B is bounded from below, and that the bound contains a (quantum) part given by the expectation value of the familiar commutator of A and B along with other commutators of operators characterized by the measurement. This structure of inequality allows us to argue that our uncertainty relation maintains Heisenberg's original idea of the uncertainty principle, albeit in a more involved manner than usually recognized.

The fundamental premise we adopted in our formulation is that any measurement yields probability distributions representing its outcomes and induces a change of quantum states, and that each of them forms an affine map between relevant spaces. The former is ensured for any measurement of statistical nature with the possibility of affecting the target system, while the latter is required for consistency in describing the corresponding transition from the state (quantum probability) space to the state/classical probability space. Because of this simplicity and universality, our uncertainty relation is expected to be valid for all conceivable measurements and also verifiable experimentally without any conditions for operational implementability.

In fact, since our uncertainty relation for error and disturbance can be understood as a special case of that for errors, in this general sense we can say that our relation also entails the standard Kennard-Robertson (Schrödinger) relation regarding quantum indeterminacy expressed by standard deviations when the measurement is non-informative [10]. It is interesting to see that, all the three orthodox relations regarding quantum indeterminacy, measurement error, and observer effect, each of which governs the seemingly distinct realms in which the uncertainty principle manifests itself, are in fact all one.

It should be also worthwhile to consider whether our observations can shed some light on other notable uncertainty relations mentioned in the Introduction. For this, we first note that our framework naturally encompasses the indirect measurement scheme adopted by several alternative formulations including Ozawa's, for every quantum measurement employing detector systems also preserves the structure of probabilistic mixture. We find it assuring in this respect that our relation reduces Ozawa's relation [5] to one of its corollaries. This can be observed from the fact that our inequality is tighter than his; one finds that Ozawa's error ε and disturbance η is never less than ours ε_ρ and η_ρ, respectively, and further reveals $\varepsilon(A)\eta(B) \geq \varepsilon_\rho(A)\eta_\rho(B) \geq \sqrt{\mathcal{R}^2 + \mathcal{I}^2} \geq |\mathcal{I}| \geq |\langle [A,B]\rangle_\rho|/2 - \varepsilon(A)\sigma(B) - \sigma(A)\eta(B)$. Here, note that the inequality of the left- and right-most hand sides is equivalent to Ozawa's inequality. Details on this topic will be reported in our subsequent papers.

ACKNOWLEDGMENTS

The authors thank Prof. Naomichi Hatano for fruitful discussions and insightful comments. This work was supported by JSPS Grant-in-Aid for Scientific Research (KAKENHI), Grant Numbers JP18K13468 and JP18H03466.

[1] W. K. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Z. Phys. **43**, 172 (1927).
[2] E. H. Kennard, Zur Quantenmechanik einfacher Bewegungstypen, Z. Phys. **44**, 326 (1927).
[3] H. P. Robertson, The Uncertainty Principle, Phys. Rev. **34**, 163 (1929).
[4] W. K. Heisenberg, *Die physikalischen Prinzipien der Quantentheorie* (S. Hirzel Verlag, Leipzig, 1930).
[5] M. Ozawa, Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement, Phys. Rev. A **67**, 042105 (2003).
[6] M. Ozawa, Uncertainty relations for joint measurements of noncommuting observables, Phys. Lett. A **320**, 367 (2004).
[7] R. F. Werner and T. Farrelly, Uncertainty from Heisenberg to Today, Found. Phys. **49**, 460 (2019).
[8] R. F. Werner, The uncertainty relation for joint measurement of postion and momentum, Quant. Inf. Comput. **4**, 546 (2004).
[9] K. Koshino and A. Shimizu, Quantum Zeno effect by general measurements, Phys. Rep. **412**, 191 (2005).
[10] J. Lee and I. Tsutsui, A Universal Formulation of Uncertainty Relation for Errors, arXiv:2002.04008 (2020).
[11] J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer Verlag, 1932).
[12] G. Lüders, Über die Zustandsänderung durch den Meßprozeß, Ann. Phys. **8**, 322 (1951).
[13] R. V. Kadison, A Generalized Schwarz Inequality and Algebraic Invariants for Operator Algebras, Ann. Math. **56**, 494 (1952).