

Uncertainty Relations for Approximation and Estimation

Jaeha Lee^{1*} and Izumi Tsutsui^{1,2†}

¹*Department of Physics, University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

²*Theory Center, Institute of Particle and Nuclear Studies,
High Energy Accelerator Research Organization (KEK),
1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan*

(Dated: November 26, 2015)

Abstract

We present a versatile inequality of uncertainty relations which are useful when one approximates an observable and/or estimates a physical parameter based on the measurement of another observable. It is shown that the optimal choice for proxy functions used for the approximation is given by Aharonov's weak value, which also determines the classical Fisher information in parameter estimation, turning our inequality into the genuine Cramér-Rao inequality. Since the standard form of the uncertainty relation arises as a special case of our inequality, and since the parameter estimation is available as well, our inequality can treat both the position-momentum and the time-energy relations in one framework albeit handled differently.

*Email: jlee@post.kek.jp

†Email: izumi.tsutsui@kek.jp

I. INTRODUCTION

Uncertainty relations lie undoubtedly at the heart of quantum mechanics, characterizing the indeterministic nature of microscopic phenomena which stem from the incompatibility of simultaneous measurement of two non-commuting observables, as typically exemplified by position and momentum. Soon after the celebrated exposition of Heisenberg's trade-off relation between error and disturbance [1], there appeared a revised form called the Robertson-Kennard (RK) inequality [2, 3] which refers to the relation in standard deviation in independently performed measurements on the two observables. Because of its mathematical clarity and universal validity, the latter has now become a standard textbook material.

Later, these relations were elaborated from operational viewpoints by taking account of the measurement device, and this has yielded, *e.g.*, the Arthurs-Goodman inequality [4] and the Ozawa inequality [5] which concerns a mixed relation among error, disturbance and standard deviation. Apart from these, the uncertainty relations on error and disturbance have also been analyzed in quantum estimation theory [6].

On the other hand, the uncertainty relation between time and energy has to be dealt with quite independently from these, due to the lack of a genuine time operator conjugate to the Hamiltonian. For this, several ingenious frameworks have been proposed, including the one devised by Mandelshtam-Tamm [7] and that by Helstrom [8], where the uncertainty relation is shown to be identified with a quantum version of the Cramér-Rao inequality [9] in estimation theory.

In this paper, we present a novel inequality of uncertainty relations for analyzing the error of approximating an observable based on the measurement of another observable through an appropriate choice of proxy functions. Since the standard deviation may be regarded as a special case of our approximation error, our inequality can formally be considered as an extension of the RK inequality. Moreover, instead of approximating an observable, we may also choose to estimate a physical parameter pertinent to the observable, so that the time-energy relation can be treated along with the position-momentum relation.

Interestingly, in both approximation and estimation, Aharonov's weak value [10] of the concerned observable arises as a key geometric ingredient, deciding the optimal choice for the proxy functions. We shall also find in the context of parameter estimation that the

weak value determines the classical Fisher information and turns our inequality into the Cramér-Rao inequality.

II. UNCERTAINTY RELATION FOR APPROXIMATION

Before presenting our uncertainty relation, let us recall the most familiar form of the relations, *i.e.*, the Robertson-Kennard (RK) inequality,

$$\|A - \langle A \rangle\| \cdot \|B - \langle B \rangle\| \geq \frac{1}{2} |\langle [A, B] \rangle|, \quad (1)$$

valid for two observables A, B . Here, $\langle A \rangle = \langle \psi | A | \psi \rangle$ is the expectation value of A under a given (normalized) state $|\psi\rangle$, $[A, B] = AB - BA$ is the commutator, and $\|X\| = \sqrt{\langle X^2 \rangle}$ is the operator seminorm defined for a self-adjoint operator X . The RK inequality gives the lower bound for the product of the standard deviation $\sigma(X) = \sqrt{\text{Var}[X]} = \|X - \langle X \rangle\|$ for the two observables A and B . Needless to say, the lower bound of the RK inequality (1) takes the state-independent value $\hbar/2$ when A and B are canonically conjugate to each other, $[A, B] = i\hbar$.

Now consider, instead of (1), a more versatile form of the inequality given by

$$\|A - f(B)\| \cdot \|g(B)\| \geq \frac{1}{2} |\langle [A, g(B)] \rangle|, \quad (2)$$

valid for arbitrary self-adjoint operator functions $f(B)$ and $g(B)$ of B . More explicitly, the operator $f(B)$ is defined from a function $f(b)$ of b through the spectral decomposition,

$$f(B) = \int f(b) |b\rangle \langle b| db, \quad (3)$$

where $\{|b\rangle\}$ is the basis set of the eigenstates $B|b\rangle = b|b\rangle$, and the integral in (3) is understood to imply summation when the eigenvalues are discrete. The operator $g(B)$ can be defined analogously from a function $g(b)$. In our discussion, $f(b)$ and $g(b)$ are assumed real so that both $f(B)$ and $g(B)$ are self-adjoint, but the definition (3) can be applied to any normal operators using complex functions.

The proof of the inequality (2) goes precisely the same way as that of the RK inequality. Namely, given two self-adjoint operators X, Y , we have $\|X\|^2 \cdot \|Y\|^2 \geq |\langle XY \rangle|^2$ by the Cauchy-Schwarz (CS) inequality. We also have $|\langle XY \rangle|^2 = |\langle [X, Y]/2 + \{X, Y\}/2 \rangle|^2$ where $\{X, Y\} = XY + YX$, but since $\langle [X, Y]/2 \rangle$ is purely imaginary whereas $\langle \{X, Y\}/2 \rangle$ is real,

we obtain $|\langle XY \rangle|^2 = |\langle [X, Y]/2 \rangle|^2 + |\langle \{X, Y\}/2 \rangle|^2 \geq |\langle [X, Y]/2 \rangle|^2$. Combining these, and taking the square root of the two sides, we arrive at $\|X\| \cdot \|Y\| \geq \frac{1}{2} |\langle [X, Y] \rangle|$. Since X and Y are arbitrary, we may put $X = A - f(B)$ and $Y = g(B)$ to obtain our inequality (2). The RK inequality (1) arises from (2) by letting $f(B) = \langle A \rangle$ (realized by the constant function $f(b) = \langle A \rangle$) and $g(B) = B - \langle B \rangle$.

Although our inequality is merely a generalization of the RK inequality, the acquired form (2) allows for a novel viewpoint on the uncertainty relation. Specifically, noting that $\|A - f(B)\|$ gives a measure for the ‘distance’ between the two observables A and $f(B)$, we may regard (2) as an inequality giving the lower bound for the distance under the choice of $f(B)$ and $g(B)$. This will be made more apparent by introducing $\bar{g}(B) = g(B)/\|g(B)\|$ to rewrite (2) as

$$\min_f \|A - f(B)\| \geq \max_{\bar{g}} \frac{1}{2} |\langle [A, \bar{g}(B)] \rangle|. \quad (4)$$

This indicates that the minimal distance between A and $f(B)$, or the minimal error in the approximation of A in terms of proxy functions $f(B)$, is dictated by the maximal degree of non-commutativity of A with respect to the family of all self-adjoint operators $\bar{g}(B)$ normalized as $\|\bar{g}(B)\| = 1$.

Clearly, our inequality will be useful in the operational context in which one measures only B and approximates A out of the measurement result by choosing the proxy function $f(B)$ properly. In this context, the choice $f(B) = \langle A \rangle$, which makes the distance into the standard deviation $\sigma(A)$, is far from the optimal one, having only its expectation value right. In fact, we shall see shortly that the optimal choice for $f(b)$ is provided explicitly by the real part of the weak value $A_w(b)$ which is defined in the quantum process specified by the initial state $|\psi\rangle$ and the final state $|b\rangle$ (see (6)). Under this optimal choice together with $g(B) = B - \langle B \rangle$, our inequality (2) yields an uncertainty relation stricter than the RK inequality. The freedom of choice for $g(b)$ may further be exploited for considering parameter estimation, that is, for estimating a parameter t that specifies the state, as exemplified later by the situation in which the state varies unitarily with the generator A . In this context, the optimal choice for $g(b)$ turns out to be given by the imaginary part of the weak value $A_w(b)$.

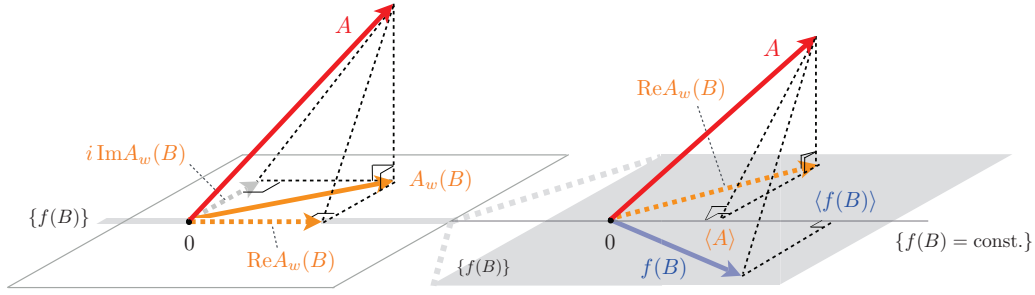


FIG. 1: Geometric relations among the operators involved. The left illustrates how the operator A is projected onto the subspace of normal operators generated by B , with the center line representing the space of self-adjoint operators $\{f(B)\}$. The right elaborates the projection onto the space $\{f(B)\}$, where now the center line represents the space of constant functions $\{f(B) = \text{const.}\}$ (more precisely, functions proportional to the identity operator I) including $f(B) = \langle A \rangle$.

III. OPTIMAL CHOICE AND THE WEAK VALUE

To see these, let us first note the identity,

$$A|\psi\rangle = \int |b\rangle\langle b|A|\psi\rangle db = \int A_w(b)|b\rangle\langle b|\psi\rangle db = A_w(B)|\psi\rangle, \quad (5)$$

where $A_w(B)$ is the operator function defined in (3) with $f(b)$ replaced by the weak value [14],

$$A_w(b) = \frac{\langle b|A|\psi\rangle}{\langle b|\psi\rangle}. \quad (6)$$

From (5), we have

$$\langle f(B)A \rangle = \langle f(B)A_w(B) \rangle = \langle A_w(B)f(B) \rangle, \quad (7)$$

for any self-adjoint $f(B)$ and, in particular,

$$\langle A \rangle = \langle A_w(B) \rangle = \langle \text{Re}A_w(B) \rangle, \quad (8)$$

since $\langle \text{Im}A_w(B) \rangle = 0$. Another consequence of (5) is

$$\|A\|^2 = \|\text{Re}A_w(B)\|^2 + \|\text{Im}A_w(B)\|^2. \quad (9)$$

These statistical properties on average and correlation suggest that the operator $\text{Re}A_w(B)$ may furnish the optimal proxy function for A minimizing the distance $\|A - f(B)\|$. That

this is indeed the case can be confirmed at once from the identity which follows from (7):

$$\|A - f(B)\|^2 = \|A - \text{Re}A_w(B)\|^2 + \|\text{Re}A_w(B) - f(B)\|^2. \quad (10)$$

We thus see that the optimal choice for $f(B)$ is made by

$$f_{\text{opt}}(B) = \text{Re}A_w(B) \quad (11)$$

for which we have $\langle f_{\text{opt}}(B) \rangle = \langle A \rangle$ as expected. As for the optimal choice for $\bar{g}(B)$ attaining the maximal value for the commutator in the r.h.s. of (4), one readily learns from the equality condition of the CS inequality and (7) that $\bar{g}_{\text{opt}}(B) = \text{Im}A_w(B)/\|\text{Im}A_w(B)\|$ provides the answer. It is then obvious that these optimal choices, $f_{\text{opt}}(B)$ and $\bar{g}_{\text{opt}}(B)$, realize the equality in (4).

IV. RK INEQUALITY REVISITED

Now, under the optimal choice (11) for $f(B)$, one may put $g(B) = B - \langle B \rangle$ in (2) to obtain

$$\|A - \text{Re}A_w(B)\| \cdot \|B - \langle B \rangle\| \geq \frac{1}{2} |\langle [A, B] \rangle|. \quad (12)$$

Recalling that the RK inequality arises at the non-optimal choice $f(B) = \langle A \rangle$, we see that, apart from the trivial case where the l.h.s. vanishes, the inequality (12) is tighter than the RK inequality (1). It is also evident that (12) reduces to the RK inequality if

$$\text{Re}A_w(B)|\psi\rangle = \langle A \rangle|\psi\rangle, \quad (13)$$

in which case the covariance,

$$\text{Cov}[A, B] = \frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle = \langle (\text{Re}A_w(B) - \langle A \rangle)(B - \langle B \rangle) \rangle, \quad (14)$$

vanishes identically.

An elementary example to illustrate our point is provided by the 1-qubit system with $A = \sigma_x$, $B = \sigma_z$. Writing

$$|\psi\rangle = \begin{pmatrix} \cos(\theta/2) \\ e^{i\varphi} \sin(\theta/2) \end{pmatrix}, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi \leq 2\pi, \quad (15)$$

in the Bloch sphere representation, one obtains

$$\|\text{Re}A_w(B) - \langle A \rangle\| = |\cos \theta \cos \varphi|, \quad \|B - \langle B \rangle\| = |\sin \theta|. \quad (16)$$

One thus finds that generically (as long as $\theta \neq 0, \pi/2, \pi$ and $\varphi \neq \pi/2, 3\pi/2$) our inequality (12) gives a tighter relation than the RK inequality.

Now, notice that applying the CS inequality to (14) yields another inequality,

$$\|\text{Re}A_w(B) - \langle A \rangle\| \cdot \|B - \langle B \rangle\| \geq \left| \frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle \right|. \quad (17)$$

We then see, from (7) and (14), that the lower bound is equal to $\text{Cov}[\text{Re}A_w(B), B]$, and hence (17) is nothing but the classical covariance inequality $\sigma(X)\sigma(Y) \geq \text{Cov}[X, Y]$. This should be the case, because the operators appearing in (17) are all generated by B and, accordingly, they are simultaneously measurable.

From this observation we learn that, while the inequality (12) gives a purely quantum lower bound for the product of error in approximating A and the standard deviation of B , the inequality (17) gives a classical lower bound given by the covariance of the two observables. These two may be regarded as complementary to each other in view of the fact that, if we sum them after squaring the both, and use the triangle inequality, we find the Schrödinger inequality [11],

$$\|A - \langle A \rangle\|^2 \cdot \|B - \langle B \rangle\|^2 \geq \frac{1}{2} |\langle [A, B] \rangle|^2 + \frac{1}{2} |\langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle|^2, \quad (18)$$

which is a tightened version of the RK inequality.

In passing, we observe from the equality condition of the CS inequality and the identity

$$\|A - \text{Re}A_w(B)\| = \|\text{Im}A_w(B)\|, \quad (19)$$

which follows from (5) that the equality in (12) holds if

$$\text{Im}A_w(B)|\psi\rangle = \lambda(B - \langle B \rangle)|\psi\rangle, \quad (20)$$

for some real λ , whereas the equality in (17) holds if

$$(\text{Re}A_w(B) - \langle A \rangle)|\psi\rangle = \mu(B - \langle B \rangle)|\psi\rangle, \quad (21)$$

for some real μ . Combining (20) and (21), and using (5) again, we obtain

$$(A - \langle A \rangle)|\psi\rangle = \beta(B - \langle B \rangle)|\psi\rangle \quad (22)$$

with $\beta = \mu + i\lambda$. Obviously, when position and momentum are considered for the observables $A = Q$, $B = P$, the condition (22) reduces to the standard minimal uncertainty condition, in which β is related to the parameter characterizing the squeezed coherent states.

V. PARAMETER ESTIMATION AND TIME-ENERGY UNCERTAINTY RELATION

Returning to the original form (2), we consider a family of states $|\psi(t)\rangle = e^{-itA/\hbar}|\phi\rangle$ generated by A with a real parameter t for a fixed $|\phi\rangle$. Suppose that our aim is to find the best function $g(B)$ to estimate the parameter t around a certain time $t = t_0$ by looking at the expectation value $\langle g(B) \rangle = \langle \psi(t) | g(B) | \psi(t) \rangle$. In more technical terms, we wish to find the locally unbiased estimator $g(B)$ fulfilling

$$\langle g(B) \rangle|_{t=t_0} = t_0, \quad \frac{d}{dt} \langle g(B) \rangle|_{t=t_0} = 1, \quad (23)$$

such that the variance $\text{Var}[g(B)]$ becomes minimal.

At this point, it is important to recognize from (19) that the optimal value $\|\text{Im}A_w(B)\|^2$ of the minimal (squared) error in the approximation of A that arises under $f_{\text{opt}}(B)$ is nothing but the Fisher Information associated with the probability density $p(b, t) = |\langle b | \psi(t) \rangle|^2$. Indeed, when the distribution $p(b, t)$ is regarded as the likelihood function for t , the corresponding Fisher Information reads

$$I(t) = \int \left[\frac{d}{dt} \ln p(b, t) \right]^2 p(b, t) db = \left(\frac{2}{\hbar} \right)^2 \|\text{Im}A_w(B)\|^2. \quad (24)$$

This prompts us to put $g(B) \rightarrow g(B) - \langle g(B) \rangle$ in (2) at the optimality (11) and see that our inequality turns into the Cramér-Rao inequality [9],

$$\text{Var}[g(B)] \geq \frac{\left(\frac{d}{dt} \langle g(B) \rangle \right)^2}{I(t)}, \quad (25)$$

on account of the identity $\frac{d}{dt} \langle g(B) \rangle = \frac{i}{\hbar} \langle [A, g(B)] \rangle$. The connection between the uncertainty relation and a quantum counterpart of the Cramér-Rao inequality has been pointed out earlier in estimation theory [8], but here we find that the precise connection between our uncertainty relation and the classical Cramér-Rao inequality holds when the optimal choice is made. The optimal choice of the efficient estimator $g(B)$ fulfilling (23) and attains the lower bound is now readily given by

$$g_{\text{opt}}(B) = -\frac{2}{\hbar I(t_0)} \text{Im}A_w(B) + t_0, \quad (26)$$

which is well-defined as long as the Fisher Information is nonvanishing $I(t_0) \neq 0$.

Let us specialize to the case where the unitary family of states is given by the time evolution generated by the Hamiltonian H . The parameter t , which is to be estimated by $g(B)$, is then the time parameter of the evolution. Namely, we are estimating the ‘energy’ of the system through $f(B)$ and the ‘time’ of the system through $g(B)$, both based on the measurement of B . Plugging $A = H$ and letting $g(B) \rightarrow g(B) - t_0$ in our inequality (2), and assuming that $g(B)$ is a locally unbiased estimator, we obtain the time-energy uncertainty relation

$$\|H - f(B)\| \cdot \|t_0 - g(B)\| \geq \frac{\hbar}{2}. \quad (27)$$

The equality in (27) holds for the optimal choice of $f(B)$ and $g(B)$ given respectively in (11) and (26) for $A = H$.

Before closing we note that since the latter condition in (23) is equivalent to $\langle [g(B), H] \rangle = i\hbar$, a locally unbiased estimator $g(B)$ is required to be canonically conjugate to H in the expectation value at least around $t = t_0$. The existence of such $g(B)$ is ensured from (26) with $A = H$, which implies that an admissible form of such estimators is provided by $g(B) = g(B)_{\text{opt}} + X$ for a time independent self-adjoint operator X with $\langle X \rangle = 0$. Of course, in the actual estimation we know neither H nor t_0 , but at least we know that there are a host of estimators that meet our requirements.

VI. SUMMARY AND REMARKS

To summarize, we have presented a novel inequality of uncertainty relations for approximation and/or estimation errors. The minimal uncertainty is determined by the weak value, and in the context of estimation our inequality reduces to the Cramér-Rao inequality. Since our inequality contains the RK inequality as a special case, it can treat both the position-momentum and the time-energy uncertainty relations in one formula, even though they have to be handled differently. The appearance of the weak value in the relations is not accidental. In fact, behind it lies the more fundamental notion of quasi-probability [12, 13], whose significance in quantum mechanics should be worth exploring further.

This work was supported in part by JSPS KAKENHI No. 25400423, No. 26011506, and

- [1] W. Heisenberg, *Z. Phys.* **43**, 172 (1927).
- [2] E. H. Kennard, *Z. Phys.* **44**, 326 (1927).
- [3] H. P. Robertson, *Phys. Rev.* **34**, 163 (1929).
- [4] E. Arthurs, and M. S. Goodman, *Phys. Rev. Lett.* **60**, 2447 (1988).
- [5] M. Ozawa, *Phys. Rev. A* **67**, 042105 (2003); *Phys. Lett. A* **320**, 367 (2004).
- [6] Y. Watanabe, T. Sagawa, and M. Ueda, *Phys. Rev. Lett.* **104**, 020401 (2010); *Phys. Rev. A* **84**, 042121 (2011).
- [7] L. I. Mandelshtam, and I. E. Tamm, *Izv. Akad. Nauk SSSR (ser. Fiz.)* **9**, 122 (1945) (English translation: *J. Phys. (USSR)* **9**, 249 (1945)).
- [8] C. W. Helstrom, “Quantum Detection and Estimation Theory”, Academic Press, (1976).
- [9] H. Cramér, “Mathematical Methods of Statistics”, Princeton Univ. Press, (1946).
- [10] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).
- [11] E. Schrödinger, *Phys.-Math. Kl.* **14**, 296 (1930).
- [12] J. Lee, and I. Tsutsui, *Prog. Theor. Exp. Phys.*, to appear.
- [13] T. Mori, and I. Tsutsui, *Quantum Stud.: Math. Found.* **2**, 371 (2015); *Prog. Theor. Exp. Phys.*, 043A01 (2015).
- [14] We assume here for simplicity that $\langle b|\psi\rangle \neq 0$ for all $|b\rangle$, which can always be ensured if one chooses B appropriately with respect to the given state $|\psi\rangle$. This restriction can be lifted if we adopt the definition of the weak value being a family of functions with ambiguity at the singular points $\langle b|\psi\rangle = 0$, for which our argument below goes through without an essential change.