Error–Disturbance Relation in Stern–Gerlach Measurements

Yuki Inoue[1] and Masanao Ozawa[1,2]

[1]Graduate School of Informatics, Nagoya University, Chikusa-ku, Nagoya, 464-8601, Japan*

[2]College of Engineering, Chubu University, 1200 Matsumoto-cho, Kasugai-shi, Aichi, 487-8501, Japan†

Although "Heisenberg's uncertainty principle" is represented by a rigorously proven relation about intrinsic uncertainties in quantum states, Heisenberg's error–disturbance relation (EDR) has been commonly believed as another aspect of the principle. Based on the recent development of universally valid reformulations of Heisenberg's EDR, we study the error and disturbance of Stern–Gerlach measurements of a spin-1/2 particle. We determine the range of the possible values of the error and disturbance for arbitrary Stern–Gerlach apparatuses with the orbital degree prepared in an arbitrary Gaussian state. We show that their error–disturbance region is close to the theoretical optimal and actually violates Heisenberg's EDR in a broad range of experimental parameters. We also show the existence of orbital states in which the error is minimized by the screen at a finite distance from the magnet, in contrast to the standard assumption.

I. INTRODUCTION

A fundamental feature of quantum measurement is nontrivial error–disturbance relations (EDRs), first found in 1927 by Heisenberg [1], who, using the famous γ-ray microscope thought experiment, derived the relation

$$\varepsilon(Q)\eta(P) \geq \frac{\hbar}{2} \tag{1}$$

between the position measurement error, $\varepsilon(Q)$, and the momentum disturbance, $\eta(P)$, thereby caused. His formal derivation of this relation from the well-established relation

$$\sigma(Q)\sigma(P) \geq \frac{\hbar}{2} \tag{2}$$

for standard deviations $\sigma(Q)$ and $\sigma(P)$, due to Heisenberg [1] for the minimum uncertainty wave packets and Kennard [2] for arbitrary wave functions, needs an additional assumption on the state change caused by the measurement [3].

Nowdays, the state change caused by a measurement is generally described by a completely positive (CP) instrument, a family of CP maps summing to a trace-preserving CP map [4]. In such a general description of quantum measurements, Heisenberg's EDR (1) loses its universal validity, as revealed in the debate in the 1980s on the sensitivity limit for gravitational wave detection derived by Heisenberg's EDR (1), but settled questioning the validity of Heisenberg's EDR [5–10]. A universally valid error–disturbance relation for arbitrary pairs of observables was derived only recently by one of the authors [11–13], and has recently received considerable attention. The validity of this relation, as well as a stronger version of this relation [14–17], were experimentally tested with neutrons [18–21] and with photons [22–26]. Other approaches generalizing Heisenberg's original relation (1) can be found, for example, in [27–29], apart from the information theoretical approach [30, 31].

Stern–Gerlach measurements [32–34] are among the most important quantum measurements, and a number of theoretical analyses have been and are being published by many authors. In his famous textbook, Bohm [35, p. 596] derived the wave function of a spin-$1/2$ particle that has passed through the Stern–Gerlach apparatus. In his argument, he assumed that the magnetic field points in the same direction everywhere and varies in strength linearly with the z-coordinate of the position as

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ B_0 + B_1 z \end{pmatrix}. \tag{3}$$

However, as Bohm [35, p. 594] pointed out, such a magnetic field does not satisfy Maxwell's equations. Theoretical studies [36–38] of Stern–Gerlach measurements with the magnetic field

$$\mathbf{B} = \begin{pmatrix} -B_1 x \\ 0 \\ B_0 + B_1 z \end{pmatrix} \tag{4}$$

satisfying Maxwell's equations were performed only recently. According to these studies, if the magnetic field in the center of the beam is sufficiently strong, the precession of the spin component to be measured becomes small, and hence Bohm's approximation (3) holds.

Home et al. [39] investigated the error of Stern–Gerlach measurements with respect to the distinguishability of apparatus states. As an indicator of the operational distinguishability of apparatus states, they used the error integral, which is equal to the probability of finding the particle in the spin-up state on the lower half of the screen. They analyzed the error integral in the case where the spin state of the particle just before the measurement is the eigenstate, $|\uparrow\rangle_z$, of σ_z corresponding to the eigenvalue $+1$. Nevertheless, the trade-off between the error and disturbance in Stern–Gerlach measurements has not been studied in the literature, even though the subject would elucidate the fundamental limitations of measurements in quantum theory, as Heisenberg did with the γ-ray microscope thought experiment.

In this paper, we determine the range of the possible values of the error and disturbance for arbitrary Stern–Gerlach

* inoue.y.at@gmail.com

† ozawa@is.nagoya-u.ac.jp

apparatuses, based on the general theory of the error and disturbance, which has recently been developed to establish universally valid reformulations of Heisenberg's uncertainty relation. Throughout this paper, we consider an electrically neutral particle with spin-$1/2$. Following Bohm [35], we assume that the magnetic field of a Stern–Gerlach apparatus is represented by Eq. (3), which is assumed to be sufficiently strong. The particle is assumed to stay in the magnet from time 0 to time Δt. Only the one-dimensional orbital degree of freedom along the z-axis is considered. The kinetic energy is not neglected. The particle having passed through the magnetic field is assumed to evolve freely from time Δt to $\Delta t+\tau$. The initial state of the spin of the particle is assumed to be arbitrary. The initial state of the orbital degree of freedom is such that mean values of the position and momentum are both 0. We study the error, $\varepsilon(\sigma_z)$, in measuring σ_z with a Stern–Gerlach apparatus and the disturbance, $\eta(\sigma_x)$, caused thereby on σ_x for the orbital degree of freedom to be prepared in a Gaussian pure state [42] in detail. We obtain the EDR

$$\left|\frac{\eta(\sigma_x)^2-2}{2}\right| \le \exp\left[-\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2-2}{2}\right)^2\right], \qquad (5)$$

for Stern–Gerlach measurements, where erf^{-1} represents the inverse of the error function $\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x \exp(-s^2)ds$. We compare the above EDR with Heisenberg's EDR for spin measurements

$$\varepsilon(\sigma_z)^2\eta(\sigma_z)^2 \ge 1, \qquad (6)$$

which holds for measurements with statistically independent error and disturbance [11, 13]. We show that Stern–Gerlach measurements violate Heisenberg's EDR in a broad range of experimental parameters. We also compare it with the EDR

$$\left|\frac{\eta(\sigma_x)^2-2}{2}\right| \le 1-\left(\frac{\varepsilon(\sigma_z)^2-2}{2}\right)^2, \qquad (7)$$

that holds for improperly directed projective measurements experimentally tested with neutron spin measurements conducted by Hasegawa and co-workers [18, 19], and the tight EDR for the range of $(\varepsilon(\sigma_z),\eta(\sigma_x))$ values of arbitrary qubit measurements obtained by Branciard and one of the authors [14–16] (see Eq. (20) below).

In section II, the general theory of the error and disturbance is reviewed and Stern–Gerlach measurements are investigated in the Heisenberg picture in detail. In sections III and IV, the error and disturbance of Stern–Gerlach measurements are derived. In section V, the EDR for Stern–Gerlach measurements is derived. In section VI, our research is compared with the previous research conducted by Home et al. [39]. Section VII presents the conclusion for this paper.

II. MEASURING PROCESS

For general theory of quantum measurements and their EDRs, we refer the reader to Appendix A.

A. Spin measurements

We consider measurements for a spin-1/2 particle, $\mathbf{S}$, and investigate the EDR for the measurements of the z-component, $A=\sigma_z$, and the disturbance of the x-component, $B=\sigma_x$, of the spin. We suppose that the measurement is carried out by the interaction between the the system $\mathbf{S}$ prepared in an arbitrary state ρ and the probe $\mathbf{P}$ prepared in a fixed vector state $|\xi\rangle$ from time 0 to time t_0 and ends up with the subsequent reading of the meter observable M of the probe $\mathbf{P}$. We assume the meter M has the same spectral with the measured observable σ_z. The measuring process, $\mathbf{M}$, determines the time evolution operator, U, of the composite system of $\mathbf{S}$ plus $\mathbf{P}$. In the Heisenberg picture we have the time evolution of the observables,

$$\begin{array}{rclrcl}\sigma_z(0) &=& \sigma_z\otimes\mathbb{1}, & \sigma_z(t_0) &=& U^\dagger\sigma_z(0)U,\\ \sigma_x(0) &=& \sigma_x\otimes\mathbb{1}, & \sigma_x(t_0) &=& U^\dagger\sigma_x(0)U,\\ M(0) &=& \mathbb{1}\otimes M, & M(t_0) &=& U^\dagger M(0)U.\end{array} \qquad (8)$$

The POVM Π of the measuring process $\mathbf{M}$ is given by

$$\Pi(m)=\langle\xi|P^{M(t_0)}(m)|\xi\rangle. \qquad (9)$$

The *non-selective operation*, T, of the measuring process $\mathbf{M}$ is given by

$$T(\rho)=\mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes|\xi\rangle\langle\xi|)U^\dagger] \qquad (10)$$

for any state ρ of $\mathbf{S}$, where $\mathrm{Tr}_{\mathcal{K}}$ is the partial trace over the Hilbert space $\mathcal{K}$ of the probe $\mathbf{P}$.

The quantum root-mean-square (q-rms) error, $\varepsilon(\sigma_z)=\varepsilon(\sigma_z,\mathbf{M},\rho)$, is defined by

$$\varepsilon(\sigma_z)=\mathrm{Tr}[(M(t_0)-\sigma_z(0))^2\rho\otimes|\xi\rangle\langle\xi|]^{1/2}. \qquad (11)$$

The q-rms error $\varepsilon(\sigma_z)$ has the following properties [17].

(i) *Operational definability.* The q-rms error $\varepsilon(\sigma_z)$ is definable by the POVM Π of $\mathbf{M}$ with the observable σ_z to be measured and the initial state ρ of the measured system $\mathbf{S}$.

(ii) *Correspondence principle.* In the case where $\sigma_z(0)$ and $M(t_0)$ commute in $\rho\otimes|\xi\rangle\langle\xi|$, the relation

$$\varepsilon(\sigma_z)=\varepsilon_G(\mu) \qquad (12)$$

holds for the joint probability distribution μ of $\sigma_z(0)$ and $M(t_0)$ in $\rho\otimes|\xi\rangle\langle\xi|$, where $\varepsilon_G(\mu)$ is the classical rms error defined by μ.

(iii) *Soundness.* If $\mathbf{M}$ accurately measures σ_z in ρ then $\varepsilon(\sigma_z)$ vanishes, i.e., $\varepsilon(\sigma_z)=0$.

(iv) *Completeness.* If $\varepsilon(\sigma_z)$ vanishes then $\mathbf{M}$ accurately measures σ_z in ρ.

It is known that the completeness property may not hold in the general case [40], but for any dichotomic measurements such that $A(0)^2=M(t_0)^2=\mathbb{1}$ holds for the measured observable A and the mete observable M as in the case of the present investigation, the completeness property holds [17]. Thus, the q-rms error $\varepsilon(\sigma_z)$ satisfies all the properties required for any reliable quantum generalizations of the classical rms error, i.e., (i) operational definability, (ii) correspondence principle, (iii) soundness, and (iv) completeness; see Appendix A for further discussions.

The quantum root-mean-square (q-rms) disturbance, $\eta(\sigma_x) = \varepsilon(\sigma_x, \mathbf{M}, \rho)$, is defined by

$$\eta(\sigma_x) = \mathrm{Tr}[(\sigma_x(t_0) - \sigma_x(0))^2 \rho \otimes |\xi\rangle\langle\xi|]^{1/2}. \tag{13}$$

The q-rms disturbance $\eta(\sigma_x)$ has properties analogous to the q-rms error as follows.

(i) *Operational definability.* The q-rms disturbance $\eta(\sigma_x)$ is definable by the non-selective operation T of $\mathbf{M}$ with the observable σ_x to be disturbed, and the initial state ρ of the measured system $\mathbf{S}$.

(ii) *Correspondence principle.* In the case where $\sigma_x(0)$ and $\sigma_x(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$, the relation

$$\eta(\sigma_x) = \varepsilon_G(\mu) \tag{14}$$

holds for the joint probability distribution μ of $\sigma_x(0)$ and $\sigma_x(t_0)$ in $\rho \otimes |\xi\rangle\langle\xi|$.

(iii) *Soundness.* If $\mathbf{M}$ does not disturb σ_x in ρ then $\eta(\sigma_x)$ vanishes.

(iv) *Completeness.* If $\eta(\sigma_x)$ vanishes then $\mathbf{M}$ does not disturb σ_x in ρ.

It is known that the completeness property may not hold in the general case [41, p. 750], but for any dichotomic observables such that $B^2 = \mathbb{1}$ to be disturbed as in the case of the present investigation the completeness property always holds [17]. Thus, the q-rms disturbance $\eta(\sigma_x)$ satisfies all the properties required for any reliable quantum generalizations of the classical rms change of observable B from time 0 to t_0, i.e., (i) operational definability, (ii) correspondence principle, (iii) soundness, and (iv) completeness; see Appendix A for further discussions.

Since $\sigma_z^2 = \sigma_x^2 = \mathbb{1}$ and $M^2 = \mathbb{1}$, from Eq. (A32) in Appendix A we obtain

$$\hat{\varepsilon}(\sigma_z)^2 + \hat{\eta}(\sigma_x)^2 + 2\hat{\varepsilon}(\sigma_z)\hat{\eta}(\sigma_x)\sqrt{1 - D_{\sigma_z\sigma_x}^2} \geq D_{\sigma_z\sigma_x}^2, \tag{15}$$

where

$$D_{\sigma_z\sigma_x} = \mathrm{Tr}(|\sqrt{\rho}\sigma_y\sqrt{\rho}|), \tag{16}$$

$$\hat{\varepsilon}(\sigma_z) = \sqrt{1 - \left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)^2}, \tag{17}$$

$$\hat{\eta}(\sigma_x) = \sqrt{1 - \left(\frac{\eta(\sigma_x)^2 - 2}{2}\right)^2} \tag{18}$$

from the EDR obtained by Branciard [14] for pure states and extended to mixed states by one of the authors [16].

In the case where

$$\langle\sigma_z\rangle_\rho = \langle\sigma_x\rangle_\rho = 0, \tag{19}$$

relation (15) is reduced to the tight relation

$$\left(\varepsilon(\sigma_z)^2 - 2\right)^2 + \left(\eta(\sigma_x)^2 - 2\right)^2 \leq 4 \tag{20}$$

as depicted in FIG. 1 (see Eq. (A34) in Appendix A).

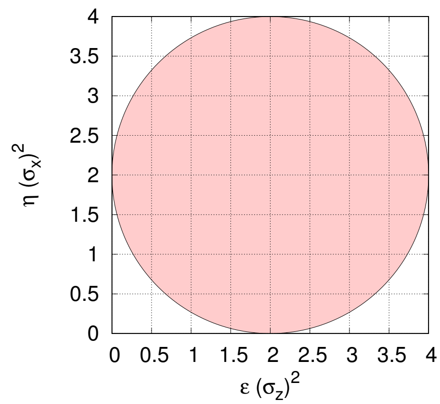

FIG. 1. $\varepsilon(\sigma_z)^2$-$\eta(\sigma_x)^2$ plot of tight EDR (15) for spin measurements in the state satisfying Eq. (19).

Lund and Wiseman [21] proposed a measurement model measuring Pauli $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ observable of an abstract qubit described by the Hilbert space $\mathcal{H} = \mathbb{C}^2$ with a computational basis $\{|0\rangle, |1\rangle\}$. The probe is another qubit prepared in the state $|\xi(\theta)\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$ and the meter observable M is chosen as Pauli σ_z observable of the probe. The measuring interaction is described by the unitary operator U_{CNOT} on $\mathbb{C}^2 \otimes \mathbb{C}^2$ performing the controlled-NOT (CNOT) operation controlled on the measured qubit. Thus, the measuring process is specified as $\mathbf{M}(\theta) = (\mathbb{C}^2, |\xi(\theta)\rangle, U_{\mathrm{CNOT}}, \sigma_z)$. Then, for the system state $|\psi\rangle = |\sigma_y = +1\rangle = (1/\sqrt{2})(|0\rangle + i|1\rangle)$, which satisfies condition (19) for $\rho = |\psi\rangle\langle\psi|$, the measurement error $\varepsilon(\sigma_z)$ of $\mathbf{M}(\theta)$ for $A = \sigma_z$ and the disturbance $\eta(\sigma_x)$ of $\mathbf{M}(\theta)$ for $B = \sigma_x$ is given by

$$\varepsilon(\sigma_z) = 2|\sin\theta|, \tag{21}$$

$$\eta(\sigma_x) = \sqrt{2}|\cos\theta - \sin\theta|. \tag{22}$$

Thus, the error $\varepsilon(\sigma_z)$ and disturbance $\eta(\sigma_x)$ satisfy the relation

$$(\varepsilon(\sigma_z)^2 - 2)^2 + (\eta(\sigma_x)^2 - 2)^2 = 4, \tag{23}$$

and attain the bound for the EDR (15). Experimental realizations of this EDR for optical polarization measurements were reported by Rozema et al. [22] and others [23, 25, 26].

In this paper, we consider another type of measurement model, known as Stern–Gerlach measurements, measuring the Pauli σ_z observable of a concrete qubit, the z-component of the spin of a spin 1/2 particle, and investigate the admissible region of the error $\varepsilon(\sigma_z)$ for σ_z and the disturbance $\varepsilon(\sigma_x)$ for σ_x, obtained from Gaussian orbital states.

B. Stern–Gerlach measurements

Let us consider a measurement of the spin component of an electrically neutral spin-1/2 particle with a Stern–Gerlach apparatus. A particle moving along the y-axis passes through an inhomogeneous magnetic field, and then the orbit is deflected depending on the spin component of the particle along the direction of the magnetic field. This situation is illustrated in FIG. 2. To analyze this measurement, we make the following assumptions.

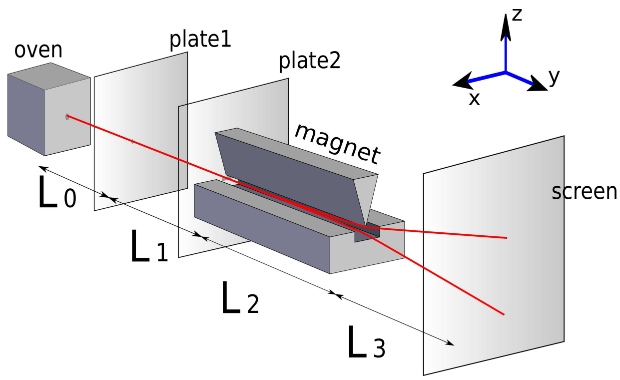

FIG. 2. Illustration of the experimental setup for a Stern–Gerlach measurement. The relations between the length and the time interval are $L_2 = v_y \Delta t$ and $L_3 = v_y \tau$.

(i) The magnetic field points everywhere in the z-axis.

(ii) The strength of the magnetic field increases proportional to the z-coordinate,

$$B_z = B_0 + B_1 z, \tag{24}$$

where B_0 and B_1 are real numbers representing the value at the origin and the gradient of B_z, respectively.

(iii) The velocity, v_y, in the y-direction is large in comparison with the motion in the x–z plane as well as the length L_2 is large in comparison with the separation of the pole faces. Thus we can treat the times $\Delta t = L_2/v_y$ and $\tau = L_3/v_y$ as deterministic for our purpose, because the determination of the spin does not depend in a sensible way on the precise evaluation of Δt and τ [35, pp. 595–596].

To describe the measuring process, $\mathbf{M}$, of a Stern–Gerlach measurement, the measured system $\mathbf{S}$ is taken as the spin degree of freedom described by the two-dimensional state space, $\mathcal{H}$, with the Pauli operators, $\sigma_x, \sigma_y, \sigma_z$, describing the x-, y- and z-components of the spin, respectively, of the spin 1/2 particle. The probe system $\mathbf{P}$ is taken as the orbital degree of freedom in the z-direction described by the Hilbert space, $\mathcal{K}$, of wave functions with position Z and momentum P satisfying the canonical commutation relation

$$[Z, P] = i\hbar. \tag{25}$$

The particle enters the magnetic field at time 0, emerges out of the magnetic field at time Δt, and freely evolves until time $\Delta t + \tau$ at which the particle reaches the screen and the observer can measure the meter observable, M, that assigns $+1$ or -1 depending on the particle z-coordinate, Z, as $M = f(Z)$, with function f such that

$$f(z) = \begin{cases} -1 & \text{if } z \geq 0, \\ +1 & \text{otherwise.} \end{cases} \tag{26}$$

Thus, the measuring process starts at time 0, when the system $\mathbf{S}$ is in any input state ρ and the probe $\mathbf{P}$ is prepared in the fixed state $|\xi\rangle$, and ends up at time $t_0 = \Delta t + \tau$. The time evolution operator, $U = U(\Delta t + \tau)$, of the composite system $\mathbf{S} + \mathbf{P}$ during the measurement is determined by the time-dependent Hamiltonian, $H(t)$, of the particle given by

$$H(t) = \begin{cases} \mu\sigma_z \otimes (B_0 + B_1 Z) + \dfrac{1}{2m}\mathbb{1} \otimes P^2 & (0 \leq t \leq \Delta t), \\ \dfrac{1}{2m}\mathbb{1} \otimes P^2 & (\Delta t \leq t \leq \Delta t + \tau), \end{cases} \tag{27}$$

where μ denotes the magnetic moment of the particle and m denotes the mass of the particle. By solving the Schrödinger equation, we obtain the time evolution operator, $U(t)$, of $\mathbf{S}+\mathbf{P}$ for $0 \leq t \leq \Delta t + \tau$ by

$$U(t) = \begin{cases} \exp\left\{ \dfrac{t}{i\hbar} \left[\mu\sigma_z \otimes (B_0 + B_1 Z) + \dfrac{1}{2m}\mathbb{1} \otimes P^2 \right] \right\} \\ \hfill (0 \leq t \leq \Delta t), \\ \exp\left[\dfrac{t - \Delta t}{2i\hbar m} \mathbb{1} \otimes P^2 \right] \\ \times \exp\left\{ \dfrac{\Delta t}{i\hbar} \left[\mu\sigma_z \otimes (B_0 + B_1 Z) + \dfrac{1}{2m}\mathbb{1} \otimes P^2 \right] \right\} \\ \hfill (\Delta t \leq t \leq \Delta t + \tau). \end{cases} \tag{28}$$

To describe the time evolution of composite system $\mathbf{S} + \mathbf{P}$ in the Heisenberg picture, we introduce Heisenberg operators for $0 \leq t \leq \Delta t + \tau$ as

$$Z(0) = \mathbb{1} \otimes Z, \quad Z(t) = U(t)^\dagger Z(0) U(t), \tag{29}$$

$$P(0) = \mathbb{1} \otimes P, \quad P(t) = U(t)^\dagger P(0) U(t), \tag{30}$$

$$\sigma_j(0) = \sigma_j \otimes \mathbb{1}, \quad \sigma_j(t) = U(t)^\dagger \sigma_j(0) U(t), \tag{31}$$

where $j = x, y, z$. For the relation between the time evolution operators in the Heisenberg picture and the Schrödinger picture, we refer the reader to Appendix D.

We also use the matrix representations of Pauli operators as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{32}$$

By solving Heisenberg equations of motion for $Z(t)$, $P(t)$,

$\sigma_x(t)$, $\sigma_y(t)$, and $\sigma_z(t)$, as shown in Appendix E, we have

$$Z(\Delta t+\tau) = Z(0) + \frac{\Delta t+\tau}{m}P(0) - \frac{\mu B_1\Delta t}{m}\left(\tau+\frac{\Delta t}{2}\right)\sigma_z(0), \tag{33}$$

$$P(\Delta t+\tau) = P(0) - \mu B_1 \Delta t \sigma_z(0), \tag{34}$$

$$\sigma_x(\Delta t+\tau) = \begin{pmatrix} 0 & \exp iS(\Delta t) \\ \exp -iS(\Delta t) & 0 \end{pmatrix}, \tag{35}$$

$$\sigma_y(\Delta t+\tau) = \begin{pmatrix} 0 & -i\exp iS(\Delta t) \\ i\exp -iS(\Delta t) & 0 \end{pmatrix}, \tag{36}$$

$$\sigma_z(\Delta t+\tau) = \sigma_z(0), \tag{37}$$

where

$$S(\Delta t) = \frac{2\mu\Delta t}{\hbar}\left[B_0 + B_1\left(Z(0)+\frac{\Delta t}{2m}P(0)\right)\right]. \tag{38}$$

III. ERROR

Let us consider the quantum rms error of a Stern–Gerlach measurement, $\mathbf{M}$, of the z-component, $\sigma_z(0)$, of the spin at time 0 using the meter observable,

$$M(\Delta t+\tau) = f(Z(\Delta t+\tau)), \tag{39}$$

introduced in Section II. The noise operator, N, of this measurement is given by

$$N = M(\Delta t+\tau) - \sigma_z(0). \tag{40}$$

Initial state ρ of the spin $\mathbf{S}$ is supposed to be an arbitrary state with the matrix

$$\rho = \frac{1}{2}(\mathbb{1} + n_x\sigma_x + n_y\sigma_y + n_z\sigma_z) \tag{41}$$

where $n_x, n_y, n_z \in \mathbb{R}$ and $n_x^2+n_y^2+n_z^2 \le 1$, so that the initial state of the composite system $\mathbf{S}+\mathbf{P}$ is given by $\rho\otimes|\xi\rangle\langle\xi|$, where $|\xi\rangle$ is a fixed but arbitrary wave function describing the initial state of the orbital degree of freedom, $\mathbf{P}$. Then, the error, namely, the quantum rms error, of this measurement of σ_z is given by

$$\varepsilon(\sigma_z) = \sqrt{\langle N^2\rangle_{\rho\otimes|\xi\rangle\langle\xi|}}, \tag{42}$$

where we abbreviate $\mathrm{Tr}[A\rho]$ as $\langle A\rangle_\rho$ for observable A and density operator ρ. We will give an explicit formula for $\varepsilon(\sigma_z)$, which eventually show that the error depends only on the parameter n_z in Eq. (41).

Let

$$U_t = \exp\left[\frac{t}{2i\hbar m}P^2\right], \tag{43}$$

$$\tilde{U}_t = \mathbb{1}_{\mathbf{S}}\otimes U_t, \tag{44}$$

$$g_0 = \frac{\mu B_1\Delta t}{m}\left(\tau+\frac{\Delta t}{2}\right). \tag{45}$$

From Eq. (33), we have

$$Z(\Delta t+\tau) = \tilde{U}^\dagger_{\Delta t+\tau}\begin{pmatrix} Z-g_0 & 0 \\ 0 & Z+g_0 \end{pmatrix}\tilde{U}_{\Delta t+\tau}. \tag{46}$$

Thus, we have

$$N = f(Z(\Delta t+\tau)) - \sigma_z(0) = 2\tilde{U}^\dagger_{\Delta t+\tau}\begin{pmatrix} -\chi_+(Z-g_0) & 0 \\ 0 & \chi_-(Z+g_0) \end{pmatrix}\tilde{U}_{\Delta t+\tau}, \tag{47}$$

where

$$\chi_+(z) = \begin{cases} 1 & \text{if } z\ge 0, \\ 0 & \text{otherwise,} \end{cases} \tag{48}$$

$$\chi_-(z) = 1-\chi_+(z), \tag{49}$$

$$f(z) = 1-2\chi_+(z). \tag{50}$$

It follows that

$$N^2 = 4\tilde{U}^\dagger_{\Delta t+\tau}\begin{pmatrix} \chi_+(Z-g_0) & 0 \\ 0 & \chi_-(Z+g_0) \end{pmatrix}\tilde{U}_{\Delta t+\tau}. \tag{51}$$

Therefore, we have

$$\begin{aligned}\varepsilon(\sigma_z)^2 &= \langle N^2\rangle_{\rho\otimes|\xi\rangle\langle\xi|} \\ &= \left\langle\xi|\mathrm{Tr}_{\mathbf{S}}[N^2\rho]|\xi\right\rangle \\ &= 2(1+n_z)\langle\xi|U^\dagger_{\Delta t+\tau}\chi_+(Z-g_0)U^\dagger_{\Delta t+\tau}|\xi\rangle \\ &\quad + 2(1-n_z)\langle\xi|U^\dagger_{\Delta t+\tau}\chi_-(Z-g_0)U^\dagger_{\Delta t+\tau}|\xi\rangle.\end{aligned} \tag{52}$$

Consequently, we have

$$\varepsilon(\sigma_z)^2 = 2(1+n_z)\int_{g_0}^{\infty}|U_{\Delta t+\tau}\xi(z)|^2dz + 2(1-n_z)\int_{-\infty}^{-g_0}|U_{\Delta t+\tau}\xi(z)|^2dz. \tag{53}$$

IV. DISTURBANCE

Let us consider the quantum rms disturbance, $\eta(\sigma_x)$, for the x-component of the spin in Stern–Gerlach measurements. The disturbance operator, σ_x, is given by

$$D = \sigma_x(\Delta t+\tau) - \sigma_x(0). \tag{54}$$

From Eq. (35),

$$D = \begin{pmatrix} 0 & \exp iS(\Delta t)-1 \\ \exp -iS(\Delta t)-1 & 0 \end{pmatrix}. \tag{55}$$

Consequently, we have

$$D^2 = \mathbb{1} \otimes (2 - 2\cos S(\Delta t)) . \tag{56}$$

Thus, we have

$$\eta(\sigma_x)^2 = 2-2\left\langle \cos\left\{\frac{2\mu\Delta t}{\hbar}\left[B_0 + B_1\left(Z + \frac{\Delta t}{2m}P\right)\right]\right\}\right\rangle_\xi . \tag{57}$$

V. ERROR AND DISTURBANCE FOR GAUSSIAN STATES

Let us consider the error and disturbance in Stern–Gerlach measurements under the condition that the orbital state of the particle is in the family $\mathcal{G}$ of Gaussian states given by

$$\mathcal{G} = \left\{ \xi_\lambda \in L^2(\mathbb{R}) \;\middle|\; \begin{array}{l} \xi_\lambda(z) = A\exp[-\lambda z^2], \\ \int_{-\infty}^{\infty} |\xi_\lambda(z)|^2 dz = 1, \\ \lambda \in \mathbb{C}, \mathrm{Re}\,(\lambda) > 0 \end{array} \right\}. \tag{58}$$

This family of states consists of all Gaussian pure states [42], whose mean values of the position and momentum are both 0. For simplicity, it is assumed that the spin state of the particle is in the eigenstate of the spin component, σ_y. It is easy to minimize the error of the measurement with respect to the mean values of the position and momentum. In particular, $\mathcal{G}$ is the family of optimal states for the measurement among the Gaussian pure states if the spin state of the particle is the eigenstate of σ_y. We remark that the equality in the Schrödinger inequality (see Eq. (B1) in Appendix B 1) holds for any state ξ in $\mathcal{G}$, i.e.,

$$\langle Z^2\rangle_\xi \langle P^2\rangle_\xi - \frac{1}{4}\langle\{Z,P\}\rangle_\xi^2 = \frac{\hbar^2}{4}. \tag{59}$$

Here, we use the abbreviation $\langle A\rangle_\xi = \langle\xi|A|\xi\rangle$. The converse also holds, that is, any state ξ satisfying $\langle P\rangle_\xi = \langle Z\rangle_\xi = 0$ and Eq. (59) belongs to $\mathcal{G}$.

Let us consider the range of the error and disturbance of Stern–Gerlach measurements. Let

$$V(\psi, t) = \left\langle \left(Z + \frac{t}{m}P\right)^2 \right\rangle_\psi \tag{60}$$

for any orbital state ψ. For disturbance $\eta(\sigma_x)$, from Eq. (57),

$$\begin{aligned} &\eta(\sigma_x)^2 \\ &= 2 - 2\left\langle \cos\left[\frac{2\mu\Delta t}{\hbar}(B_0 + B_1 Z)\right]\right\rangle_{U_{\Delta t/2}\xi_\lambda} \\ &= 2 - \frac{2}{\sqrt{2\pi V(\xi_\lambda, \Delta t/2)}} \\ &\times \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2V(\xi_\lambda,\Delta t/2)}\right)\cos\left[\frac{2\mu\Delta t}{\hbar}(B_0 + B_1 z)\right] dz \\ &= 2 - 2\exp\left(-\frac{2\mu^2 B_1^2 \Delta t^2}{\hbar^2} V(\xi_\lambda, \Delta t/2)\right)\cos\frac{2\mu\Delta t B_0}{\hbar}. \end{aligned} \tag{61}$$

From the above formula, the disturbance is determined by $V(\xi_\lambda, \Delta t/2)$ and the parameters of the magnet if the orbital state is in $\mathcal{G}$. Now, for a fixed constant, v, let us find the error for state ξ_λ in $\mathcal{G}$ and time interval Δt satisfying $V(\xi_\lambda, \Delta t/2) = v$. In the following, we fix the time interval, Δt.

From Eq. (53), we have

$$\begin{aligned} \varepsilon(\sigma_z)^2 &= 4\int_{g_0}^{\infty} |U_{\Delta t+\tau}\xi_\lambda(z)|^2 dz \\ &= \frac{4}{\sqrt{\pi}}\int_{g_0/\sqrt{2V(\xi_\lambda,\Delta t+\tau)}}^{\infty} \exp(-w^2)dw. \end{aligned} \tag{62}$$

Here, we use the relation $n_z = 0$, which is obtained from the assumption that the mean value of the z-component of the spin of the particle is 0. Equation (62) shows that the error is minimized by maximizing the lower limit of the integration $g_0/\sqrt{2V(\xi_\lambda, \Delta t+\tau)}$. First, we fix state ξ_λ and focus on time interval τ. Let $W_{\xi_\lambda}(\tau) = g_0/\sqrt{2V(\xi_\lambda, \Delta t+\tau)}$. From now on, we suppose $B_1 \le 0$. As shown in Appendix F, if

$$m\langle\{Z,P\}\rangle_{\xi_\lambda} + \langle P^2\rangle_{\xi_\lambda}\Delta t < 0 \tag{63}$$

holds, then $W_{\xi_\lambda}(\tau)$ assumes maximum value

$$W_{\xi_\lambda}(\tau_0) = \frac{\sqrt{2V(\xi_\lambda, \Delta t/2)}\mu B_1 \Delta t}{\hbar} \tag{64}$$

at

$$\begin{aligned} \tau =& \tau_0 \\ =& -\frac{4m^2\langle Z^2\rangle_{\xi_\lambda} + 3m\langle\{Z,P\}\rangle_{\xi_\lambda}\Delta t + 2\langle P^2\rangle_{\xi_\lambda}\Delta t^2}{2\left(m\langle\{Z,P\}\rangle_{\xi_\lambda} + \langle P^2\rangle_{\xi_\lambda}\Delta t\right)}. \end{aligned} \tag{65}$$

On the other hand, if condition (63) does not hold, the supremum of $W_{\xi_\lambda}(\tau)$ is given by

$$\sup_{\tau\ge 0} W_{\xi_\lambda}(\tau) = \lim_{\tau\to\infty} W_{\xi_\lambda}(\tau) = \frac{\mu B_1 \Delta t}{\sqrt{2}}\langle P^2\rangle_{\xi_\lambda}^{-1/2}. \tag{66}$$

Now, let us consider the maximization of $W_{\xi_\lambda}(\tau)$ with respect to state ξ_λ. For any pair of orbital states, ψ and ϕ, in $\mathcal{G}$ satisfying $V(\psi, \Delta t/2) = v$ and $V(\phi, \Delta t/2) = v$, respectively, if ψ satisfies condition (63), then

$$W_\psi(\tau_0) \ge \lim_{\tau\to\infty} W_\phi(\tau) \tag{67}$$

holds, since $W_\psi(\tau_0)/\lim_{\tau\to\infty} W_\phi(\tau) \ge 1$ by the Kennard inequality (2). Therefore, we obtain the supremum of $W_{\xi_\lambda}(\tau)$ with respect to state ξ_λ and time interval τ as

$$\sup_{\mathrm{Re}(\lambda)>0, \tau\ge 0} W_{\xi_\lambda}(\tau) = \frac{\sqrt{2v}\mu B_1 \Delta t}{\hbar}. \tag{68}$$

See Appendix F for the detailed derivation.

Although the above argument is for finding the range of the error and disturbance that Stern–Gerlach measurements can assume, it contains one more important assertion. That is, the calculation suggests that the error of Stern–Gerlach measurements is minimized by placing the screen at a finite distance from the magnet under the condition represented by (63), in contrast to the conventional assumption that the error is minimized by placing the screen at infinity. If a state in $\mathcal{G}$ satisfies condition (63), then the *correlation term* [7], $\left\langle\left\{Z-\langle Z\rangle_{\xi_\lambda}, P-\langle P\rangle_{\xi_\lambda}\right\}\right\rangle_{\xi_\lambda}$, is negative, and this leads to a narrowing of the standard deviation of the position of the particle during the free evolution (see Appendix B 4). Such a class of states is introduced by Yuen [7] and they are known as *contractive states*.

Let us return to the problem of finding the range of the values of the error and disturbance that Stern–Gerlach measurements can assume. Now, setting $W_0 = \sqrt{2v}\mu B_1 \Delta t/\hbar$, the disturbance and the infimum of the error under the condition that $V(\lambda, \Delta t/2) = v$ for fixed Δt and v are

$$\eta(\sigma_x)^2 = 2 - 2\exp\left(-W_0^2\right)\cos\frac{2\mu\Delta t B_0}{\hbar}, \quad (69)$$

$$\inf_{\lambda,T}\varepsilon(\sigma_z)^2 = \frac{4}{\sqrt{\pi}}\int_{W_0}^{\infty}\exp(-w^2)dw, \quad (70)$$

respectively. By varying the parameter of the magnet, B_0, we obtain the range of the disturbance as

$$2 - 2\exp\left(-W_0^2\right) \le \eta(\sigma_x)^2 \le 2 + 2\exp\left(-W_0^2\right). \quad (71)$$

We obtain the range of the disturbance and the infimum of the error of Stern–Gerlach measurements for each constant, v. By varying v, we obtain the range of the error and disturbance as the inequalities

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \le \exp\left\{-\left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)\right]^2\right\}, \quad (72)$$

$$0 \le \varepsilon(\sigma_z)^2 \le 2, \quad (73)$$

where erf^{-1} represents the inverse of the error function $\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x \exp(-s^2)ds$. The square of the error varies from 0 to 2 since W_0 is positive.

We now remove the constraint $B_1 \le 0$. For $B_1 \ge 0$, similarly to the above discussion, we have

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \le \exp\left\{-\left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)\right]^2\right\}, \quad (74)$$

$$2 \le \varepsilon(\sigma_z)^2 \le 4. \quad (75)$$

Therefore we have

$$\left|\frac{\eta(\sigma_x)^2 - 2}{2}\right| \le \exp\left\{-\left[\mathrm{erf}^{-1}\left(\frac{\varepsilon(\sigma_z)^2 - 2}{2}\right)\right]^2\right\}. \quad (76)$$

The plot of this region is shown in FIG. 3.

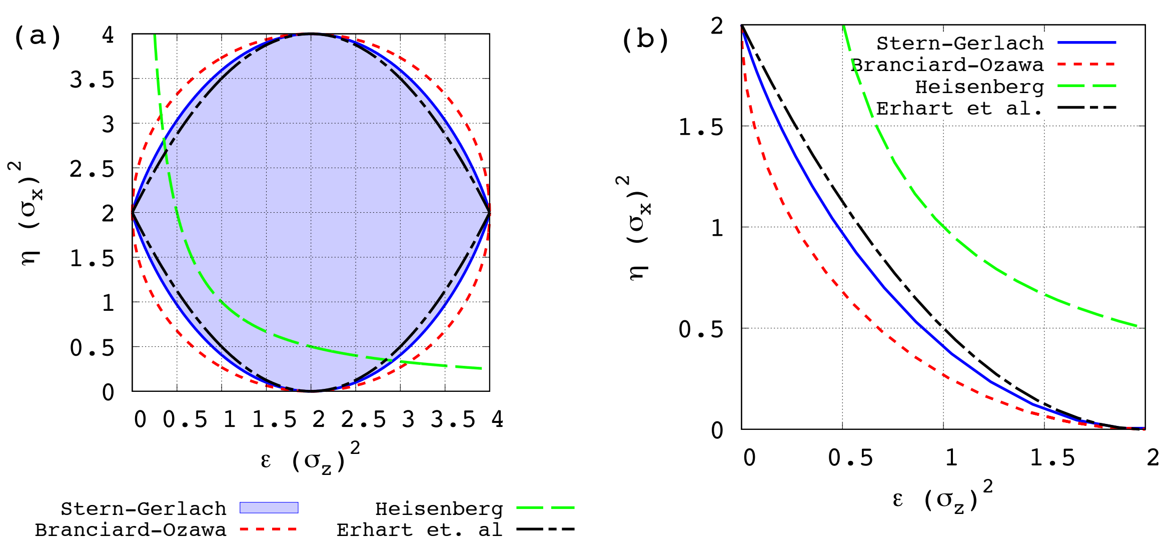

FIG. 3. (a) Range of the error and disturbance for Stern–Gerlach measurements. Blue region: the region (76) that Stern–Gerlach measurements can achieve. Red dotted line: the boundary of the Branciard–Ozawa tight EDR (20). Green dashed line: the boundary of Heisenberg's EDR (6). Black dotdash line: the theoretical boundary (7) of the EDR of the experiment conducted by Hasegawa and co-workers [18, 19]. The error–disturbance region of Stern-Gerlach measurements is close to the theoretical optimal given by the Branciard–Ozawa tight EDR (20), and actually violates Heisenberg's EDR (6) in a broad range of experimental parameters. (b) The enlarged plot for the part $[0,2]\times[0,2]$.

For comparison, the figure shows the plot of the boundary of the Branciard-Ozawa tight EDR (20) for general spin measurement. From this plot, we conclude that the range of the error and disturbance for Stern–Gerlach measurements considered in this paper is close to the theoretical optimal given by the Branciard–Ozawa tight EDR (20). Here, the range of the error and disturbance for Stern–Gerlach measurements is also compared with Heisenberg's EDR (6) (green line) and the EDR (7) for the neutron experiment [18, 19] (black line). We conclude that Stern-Gerlach measurements actually violate Heisenberg's EDR (6) in a broad range of experimental parameters.

Roughly speaking, the parameter v represents the spread of the wave packet of the particle in the Stern–Gerlach magnet. The reason why v appears in the formula of the disturbance is that the particle in the Stern–Gerlach magnet is exposed to the inhomogeneous magnetic field and its spin is processed in an uncontrollable way. This uncontrollable precession occurs because the position of the particle is uncertain while the magnetic field is inhomogeneous and hence depends on the position. The disturbance of the spin along the x-axis is caused by this uncontrollable precession around z-axis. This is why v appears in the formula of the disturbance. On the other hand, the error in our Stern–Gerlach setup comes from the non-zero dispersion of the z-component of the particle position when the particle has reached the screen. The smaller the dispersion of the particle position when the particle has reached the screen, the greater the dispersion of the z-component of the particle position in the Stern–Gerlach magnet. This is why v appears in the formula of the error.

VI. COMPARISON WITH "ASPECTS OF NONIDEAL STERN–GERLACH EXPERIMENT AND TESTABLE RAMIFICATIONS"

Home et al. [39] discussed the same error of Stern–Gerlach measurements as our paper does for similar conditions. Therefore, their paper is among the papers preceding ours. We consider in what sense their paper is related to ours, and we compare its results with ours. They derived the wave function of a particle in the Stern–Gerlach apparatus under the following conditions.

(i) The magnetic field is oriented along the z-axis everywhere, and the gradient of the z-component of the magnetic field is non-zero only in the z-direction.

(ii) The initial orbital state is a Gaussian state whose mean values of the position and momentum, and the correlation term of the particle in the wave function are all zero.

(iii) Unlike Bohm's discussion [35], the kinetic energy of the particle in the magnetic field is not neglected.

Based on their argument, they discussed the distinguishability of the value of the measured observable by observing the probe system directly in Stern–Gerlach measurements. To consider this problem, they introduced the two indices,

$$I := \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_+^*(\mathbf{x},\tau)\psi_-(\mathbf{x},\tau) d\mathbf{x} \right|, \qquad (77)$$

$$E(t) := \int_{-\infty}^{0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi_+(\mathbf{x},t)|^2 \, dxdydz, \qquad (78)$$

where $\psi_\pm$ are the wave functions of the particle in the Schrödinger picture whose spin z-components are $\pm 1/2$, respectively. The origin of time is taken to be the moment when the particle enters the Stern–Gerlach magnet. τ is the time at which the particle emerges from the Stern–Gerlach magnet (τ corresponds to Δt in our notation), and t is any time after emerging from the Stern–Gerlach magnet (t corresponds to $\Delta t + \tau$ in our notation). Namely, they adopted the inner product, I, of the two wave functions with different spin directions, and the probability, $E(t)$, of finding the particle with the spin z-components of $+1/2$ and $-1/2$ within the lower and upper half planes, respectively, at time t. They concluded that I always vanishes whenever $E(t)$ vanishes, but that $E(t)$ does not necessarily vanish even when I vanishes.

We discuss the relation between their paper and ours. The relation between the quantities $E(t)$ and $\varepsilon(\sigma_z)$ is

$$\varepsilon(\sigma_z)^2 = 4E(t). \qquad (79)$$

This relation, model-dependent though, bridges the two approaches and will enforce a theoretical background for our definition of a sound and complete quantum generalization of the classical root-mean-square error [17].

We compare their research with ours as follows.

(i) Their set up and approximation are the same as ours and they used the same Hamiltonian as in our research.

(ii) In both papers, the orbital state of the particle is assumed to be the pure state where the mean values of its position and momentum are zero. We assume that the correlation term of a Gaussian pure state is not necessarily zero, whereas they assumed that the orbital state is a Gaussian pure state with no correlation.

(iii) We evaluate the tradeoff between the error and disturbance, whereas they compared the error with the inner product, I, of the emerging wave functions expressing formal distinguishability. In addition, we obtain the range of the error and disturbance under the condition that the orbital state is a Gaussian pure state whose correlation term is not necessarily zero.

VII. CONCLUSION

Stern–Gerlach measurements, originally performed by Gerlach and Stern in 1922 [32–34], have been discussed for long as a typical model or a paradigm of quantum measurement [35]. As Heisenberg's uncertainty principle suggests, Stern–Gerlach measurements of one spin component inevitably disturb its orthogonal component, and Heisenberg's EDR (6) has been commonly believed as its precise quantitative expression. However, general quantitative relations between error and disturbance in arbitrary quantum measurements was extensively investigated in the last two decades and universally valid EDRs were obtained to reform Heisenberg's original EDR (see e.g., [11, 14, 17, 18, 28] and references therein).

Here, we investigated the EDR for this familiar class of measurements in the light of the general theory leading to the universally valid EDR relations. We have determined the range of the possible values of the error and disturbance achievable by arbitrary Stern–Gerlach apparatuses, assuming that the orbital state is a Gaussian state. Our result is depicted in Fig. 3 and the boundary of the error–disturbance region is given in Eq. (76) as a closed formula. The result shows that the error–disturbance region of Stern–Gerlach measurements occupies a near-optimal subregion of the universally valid error–disturbance region for arbitrary measurements. It can be witnessed that one of the earliest methods of quantum measurement violates Heisenberg's EDR (6) in a broad range of experimental parameters. Furthermore, we found a class of initial orbital states in which the error can be minimized arbitrarily small at the screen in a finite distance from the magnet in contrast to the conventional assumption that the error decreases aymptotically.

The relation for the general class of states beyond Gaussian states is left to the future study. In addition, we also leave it to the future research to analyze more realistic models; for example, a model described by the magnetic field satisfying Maxwell's equations [37, 38] or a model considering the decoherence of the particle during the measuring process [43].

Our results will contribute to answer the question as to how various experimental parameters can be controlled to achieve the ultimate limit. We expect that the present study will provoke further experimental studies.

ACKNOWLEDGMENTS

The authors thank Kazuya Okamura for helpful discussions. This work was partially supported by JSPS KAKENHI, Nos. 26247016 and 17K19970, and the IRI-NU collaboration.

Appendix A: Error and disturbance in quantum measurements

In this section, we review the general theory of error and disturbance in quantum measurements developed in [13, 17].

1. Classical root-mean-square error

Let us consider the classical case first. Recall the *root-mean-square (rms) error* introduced by Gauss [44]. Consider a measurement of the value x of a quantity, X, by actually observing the value y of a meter quantity, Y. Then the error of this measurement is given by $y-x$. If these quantities obey a joint probability distribution, $\mu(x,y)$, then the rms error, $\varepsilon_G(\mu)$, is defined as

$$\varepsilon_G(\mu) = \left(\sum_{x,y}(y-x)^2\,\mu(x,y)\right)^{1/2}. \qquad \text{(A1)}$$

2. Quantum measuring processes

We consider a quantum system, $\mathbf{S}$, described by a finite dimensional Hilbert space, $\mathcal{H}$. We assume that every measuring apparatus for the system $\mathbf{S}$ has its own output variable, $\mathbf{x}$. The statistical properties of the apparatus, $\mathbf{A}(\mathbf{x})$, having the output variable $\mathbf{x}$ are determined by (i) the probability distribution, $\Pr\{\mathbf{x}=m\|\rho\}$, of $\mathbf{x}$ for the input state ρ, and (ii) the output state, $\rho_{\{\mathbf{x}=m\}}$, given the outcome $\mathbf{x}=m$.

A *measuring process* of the apparatus $\mathbf{A}(\mathbf{x})$ measuring $\mathbf{S}$ is specified by a quadruple, $\mathbf{M}=(\mathcal{K},|\xi\rangle,U,M)$, consisting of a Hilbert space, $\mathcal{K}$, describing the probe system $\mathbf{P}$, a state vector, $|\xi\rangle$, in $\mathcal{K}$ describing the initial state of $\mathbf{P}$, a unitary operator, U, on $\mathcal{H}\otimes\mathcal{K}$ describing the time evolution of the composite system $\mathbf{S}+\mathbf{P}$ during the measuring interaction, and an observable, M, called the *meter observable*, of $\mathbf{P}$ describing the meter of the apparatus.

The *instrument* of the measuring process $\mathbf{M}$ is defined as a completely positive map valued function, $\mathcal{I}$, given by

$$\mathcal{I}(m)\rho = \mathrm{Tr}_{\mathcal{K}}[(\mathbb{1}\otimes P^M(m))U(\rho\otimes|\xi\rangle\langle\xi|)U^\dagger] \qquad \text{(A2)}$$

for any state ρ and real number m. The statistical properties of the apparatus $\mathbf{A}(\mathbf{x})$ are determined by the instrument $\mathcal{I}$ of $\mathbf{M}$ as

$$\Pr\{\mathbf{x}=m\|\rho\} = \mathrm{Tr}[\mathcal{I}(m)\rho], \qquad \text{(A3)}$$

$$\rho_{\{\mathbf{x}=m\}} = \frac{\mathcal{I}(m)\rho}{\mathrm{Tr}[\mathcal{I}(m)\rho]}. \qquad \text{(A4)}$$

The *non-selective operation* T of $\mathbf{M}$ is defined by

$$T = \sum_{m\in\mathbb{R}}\mathcal{I}(m). \qquad \text{(A5)}$$

Then, we have

$$T(\rho) = \mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes|\xi\rangle\langle\xi|)U^\dagger]. \qquad \text{(A6)}$$

See [4, 10, 13] for the detailed descriptions on measuring processes and instruments.

3. Heisenberg picture

In the measuring process $\mathbf{M}$, we suppose that the measuring interaction is turned on from time $t=0$ to time $t=t_0$. Then, the outcome $\mathbf{x}=m$ of the apparatus $\mathbf{A}(\mathbf{x})$ described by the measuring process $\mathbf{M}$ is defined as the outcome m of the meter measurement at time $t=t_0$. To describe the time evolution of the composite system $\mathbf{S}+\mathbf{P}$ in the Heisenberg picture, let

$$\begin{array}{rclrcl} A(0) &=& A\otimes 1\!\!1, & A(t_0) &=& U^\dagger A(0)U,\\ B(0) &=& B\otimes 1\!\!1, & B(t_0) &=& U^\dagger B(0)U,\\ M(0) &=& 1\!\!1\otimes M, & M(t_0) &=& U^\dagger M(0)U, \end{array} \tag{A7}$$

where A, B are observables of $\mathbf{S}$.

Then, the *POVM* Π of $\mathbf{M}$ is defined as

$$\Pi(m)=\langle\xi|P^{M(t_0)}(m)|\xi\rangle, \tag{A8}$$

and satisfies

$$\Pr\{\mathbf{x}=m\|\rho\}=\mathrm{Tr}[\Pi(m)\rho]. \tag{A9}$$

The *n-th moment operator* of Π for $n=1,\ldots,n$ is defined by

$$\hat{\Pi}^{(n)}=\langle\xi\,|M(t_0)^n|\,\xi\rangle\,. \tag{A10}$$

The *dual non-selective operation* T^* of $\mathbf{M}$ is defined by

$$T^*(B)=\langle\xi|B(t_0)|\xi\rangle \tag{A11}$$

for any observable B of $\mathbf{S}$ and satisfies

$$\mathrm{Tr}[(T^*(B))\rho]=\mathrm{Tr}[B(T(\rho))] \tag{A12}$$

for any observable B and state ρ.

4. Measurement of observables

If the observables $A(0)$ and $M(t_0)$ commute in the initial state $\rho\otimes|\xi\rangle\langle\xi|$, that is,

$$[P^{A(0)}(a),P^{M(t_0)}(m)](\rho\otimes|\xi\rangle\langle\xi|)=0 \tag{A13}$$

for all $a,m\in\mathbb{R}$, then their *joint probability distribution*, $\mu(a,m)$, is defined as

$$\mu(a,m)=\mathrm{Tr}[P^{A(0)}(a)P^{M(t_0)}(m)\,(\rho\otimes|\xi\rangle\langle\xi|)] \tag{A14}$$

and satisfies

$$\mathrm{Tr}[f(A(0),M(t_0))(\rho\otimes|\xi\rangle\langle\xi|)]=\sum_{a,m}f(a,m)\,\mu(a,m) \tag{A15}$$

for any polynomial $f(A(0),M(t_0))$ of $A(0)$ and $M(t_0)$.

We say that the measuring process $\mathbf{M}$ *accurately measures* the observable A in a state ρ if $A(0)$ and $M(t_0)$ are perfectly correlated in the state $\rho\otimes|\xi\rangle\langle\xi|$ [17, 41, 45]; namely, one of the following two equivalent conditions holds:

(S) $A(0)$ and $M(t_0)$ commute in $\rho\otimes|\xi\rangle\langle\xi|$ and their joint probability distribution μ satisfies

$$\sum_{a,m:a=m}\mu(a,m)=1. \tag{A16}$$

(W) For any $a,m\in\mathbb{R}$ with $a\neq m$,

$$\mathrm{Tr}\left[\Pi(m)P^A(a)\,\rho\right]=0. \tag{A17}$$

Note that $\nu(a,m):=\mathrm{Tr}\left[\Pi(m)P^A(a)\,\rho\right]$, called the *weak joint distribution* of $A(0)$ and $M(t_0)$, always exists and is operationally accessible by weak measurement and post-selection [21, 46], but possibly takes negative or complex values. Since $\nu(a,m)$ is operationally accessible, our definition of accurate measurements is operationally accessible.

5. Quantum root-mean-square error

The *noise operator* $N(A,\mathbf{M})$ of the measuring process $\mathbf{M}$ for measuring A is defined as

$$N(A,\mathbf{M})=M(t_0)-A(0). \tag{A18}$$

The *(noise-operator based) quantum root-mean-square (q-rms) error* $\varepsilon_{\mathrm{NO}}(A,\mathbf{M},\rho)$ for measuring A in ρ by $\mathbf{M}$ is defined as the root-mean-square of the noise operator, i.e.,

$$\varepsilon_{\mathrm{NO}}(A,\mathbf{M},\rho)=\mathrm{Tr}[N(A,\mathbf{M})^2(\rho\otimes|\xi\rangle\langle\xi|)]^{1/2}. \tag{A19}$$

To argue the reliability of the error measure $\varepsilon_{\mathrm{NO}}$ defined above, we consider the following requirements for any reliable error measures ε generalizing the classical root mean square error ε_G to quantify the mean error $\varepsilon(A,\mathbf{M},\rho)$ of the measurement of an observable A in a state ρ described by a measuring process $\mathbf{M}$ [17].

(i) *Operational definability.* The error measure ε should be definable by the POVM Π of the measuring process $\mathbf{M}$ with the observable A to be measured and the initial state ρ of the measured system $\mathbf{S}$.

(ii) *Correspondence principle.* In the case where $A(0)$ and $M(t_0)$ commute in $\rho\otimes|\xi\rangle\langle\xi|$, the relation

$$\varepsilon(A,\mathbf{M},\rho)=\varepsilon_G(\mu) \tag{A20}$$

holds for the joint probability distribution μ of $A(0)$ and $M(t_0)$ in $\rho\otimes|\xi\rangle\langle\xi|$.

(iii) *Soundness.* If $\mathbf{M}$ accurately measures A in ρ then ε vanishes, i.e., $\varepsilon(A,\mathbf{M},\rho)=0$.

(iv) *Completeness.* If ε vanishes then $\mathbf{M}$ accurately measures A in ρ.

It is shown in [17] that the noise-operator-based q-rms error $\varepsilon=\varepsilon_{\mathrm{NO}}$ satisfies requirements (i)–(iii), so that it is a sound generalization of the classical rms error. However, as pointed out by Busch, Heinonen, and Lahti [40], $\varepsilon=\varepsilon_{\mathrm{NO}}$ may not satisfy the completeness requirement (iv) in general. To improve this point, in Ref. [17] a modification of the noise-operator based q-rms error $\varepsilon_{\mathrm{NO}}$ is introduced to satisfy all the requirements (i)–(iv) as follows. The *locally uniform q-rms error* $\overline{\varepsilon}$ is defined by

$$\overline{\varepsilon}(A,\mathbf{M},\rho)=\sup_{t\in\mathbb{R}}\varepsilon_{\mathrm{NO}}(A,\mathbf{M},e^{-itA}\rho e^{itA}). \tag{A21}$$

Then, $\varepsilon=\overline{\varepsilon}$ satisfies all the requirements (i)–(iv) including completeness. In addition to (i)–(iv), the new error measure $\overline{\varepsilon}$ has the following two properties:

(v) *Dominating property.* The error measure $\overline{\varepsilon}$ dominates $\varepsilon_{\mathrm{NO}}$, i.e., $\varepsilon_{\mathrm{NO}}(A,\mathbf{M},\rho)\le\overline{\varepsilon}(A,\mathbf{M},\rho)$.

(vi) *Conservation property for dichotomic measurements.* The error measure $\overline{\varepsilon}$ coincides with $\varepsilon_{\mathrm{NO}}$ for dichotomic measurements, i.e., $\overline{\varepsilon}(A,\mathbf{M},\rho)=\varepsilon_{\mathrm{NO}}(A,\mathbf{M},\rho)$ if $A(0)^2=M(t_0)^2=1\!\!1$.

By property (v) the new error measure $\overline{\varepsilon}$ maintains the previously obtained universally valid EDRs [11, 14, 16]. In this paper, we consider the measurement of a spin component σ_z of a spin-$1/2$ particle using a dichotomic meter observable M, i.e., $M^2 = \mathbb{1}$, so that by property (vi) of $\overline{\varepsilon}$, we conclude that the noise-operator-based q-rms error $\varepsilon_{\rm NO}$ satisfies all the requirements (i)–(iv) for our measurements under consideration without modifying it to be $\overline{\varepsilon}$.

As shown in Eq. (47), in our model of the Stern-Gerlach measurement, the Heisenberg observables $A(0)$ and $M(t_0)$ commute , so that the error measure satisfying (i) and (ii) are uniquely determined as the (noise-operator) based q-rms error.

Busch, Lahti, and Werner [28] criticized the use of the noise-operator based q-rms error, by comparing it with the error measure based on the Wasserstein 2-distance, another error measure defined as the Wasserstein 2-distance between the probability distributions of $A(0)$ and $M(t_0)$. As shown in Ref. [17] the error measure based on the Wasserstein 2-distance, or based on any distance between he probability distributions of $A(0)$ and $M(t_0)$, satisfies (i) and (iii) but does not satisfy (ii) nor (iv), so that the discrepancies between those two measures do not leads to the conclusion that the noise-operator based q-rms error is less reliable than the error measured based on the Wasserstein 2-distance or based on any distance between probability distributions of of $A(0)$ and $M(t_0)$.

In what follows, we shall write $\varepsilon(A) = \varepsilon_{NO}(A)$ for brevity, where no confusion may occur.

6. Disturbance of observables

We say that the measuring process $\mathbf{M}$ *does not disturb* the observable B in a state ρ if $B(0)$ and $B(t_0)$ are perfectly correlated in the state $\rho \otimes |\xi\rangle\langle\xi|$ [41, 45, 47]; namely, one of the following two equivalent conditions holds:

(S) $B(0)$ and $B(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$ and their joint probability distribution μ satisfies

$$\sum_{b,b':b=b'} \mu(b, b') = 1. \tag{A22}$$

(W) For any $b, b' \in \mathbb{R}$ with $b \neq b'$,

$$\mathrm{Tr}\left[P^{B(t_0)}(b')P^{B(0)}(b)\rho \otimes |\xi\rangle\langle\xi|\right] = 0. \tag{A23}$$

Note that the left-hand side of Eq. (A23) is called the *weak joint distribution* of $B(0)$ and $B(t_0)$ and always exists possibly taking negative or complex values. The weak joint distribution is operationally accessible by weak measurement of $B(0)$ and post-selection for $B(t_0)$ [21, 46]. Thus, our definition of non-disturbing measurement is operationally accessible.

7. Quantum root-mean-square disturbance

For any observable B of the system $\mathbf{S}$, the *disturbance operator* $D(B, \mathbf{M})$ for the measuring process $\mathbf{M}$ causing on the observable B is defined as the change of the observable B during the measurement, i.e.,

$$D(B, \mathbf{M}) = B(t_0) - B(0). \tag{A24}$$

Similarly to the q-rms error, the *q-rms disturbance* $\eta(B, \mathbf{M}, \rho)$ of B in ρ caused by $\mathbf{M}$ is defined as the rms of the disturbance operator, i.e.,

$$\eta(B, \mathbf{M}) = \mathrm{Tr}[D(B, \mathbf{M})^2(\rho \otimes |\xi\rangle\langle\xi|)]^{1/2}. \tag{A25}$$

The q-rms disturbance η has properties analogous to the (noise-operator-based) q-rms error as follows.

(i) *Operational definability.* The q-rms disturbance η is definable by the non-selective operation T of the measuring process $\mathbf{M}$, the observable B to be disturbed, and the initial state ρ of the measured system $\mathbf{S}$.

(ii) *Correspondence principle.* In the case where $B(0)$ and $B(t_0)$ commute in $\rho \otimes |\xi\rangle\langle\xi|$, the relation

$$\eta(B, \mathbf{M}, \rho) = \varepsilon_G(\mu) \tag{A26}$$

holds for the joint probability distribution μ of $B(0)$ and $B(t_0)$ in $\rho \otimes |\xi\rangle\langle\xi|$.

(iii) *Soundness.* If $\mathbf{M}$ does not disturb B in ρ then η vanishes.

(iv) *Completeness for dichotomic observables.* In the case where $B^2 = \mathbb{1}$, if η vanishes then $\mathbf{M}$ does not disturb B in ρ.

Korzekwa, Jennings, and Rudolph [48] criticized the use of the operator-based q-rms disturbance relying on their definition of non-disturbing measurements. They define non-disturbing measurements in a system state ρ as measurements satisfying that $B(0)$ and $B(t_0)$ have the identical probability distributions for the initial state $\rho \otimes |\xi\rangle\langle\xi|$. They claimed that the operator-based q-rms disturbance does not satisfy the soundness requirement based on their definition of non-disturbing measurements. However, the conflict can be easily reconciled, since their definition of non-disturbing measurement is not strong enough, i.e., they call a measurement non-disturbing even when the disturbance is operationally detectable. In fact, they supposed that the projective measurement of $A = \sigma_z$ of a spin 1/2 particle in the state $|\sigma_z = +1\rangle$ does not disturb the observable $B = \sigma_x$. However, this measurement really disturbs the observable $B = \sigma_x$. In fact, we have

$$\begin{aligned}&\langle\psi, \xi|P^{B(t_0)}(b')P^{B(0)}(b)|\psi, \xi\rangle\\ &\quad = |\langle\sigma_z = +1|\sigma_x = b'\rangle|^2|\langle\sigma_z = +1|\sigma_x = b\rangle|^2.\end{aligned}$$

Thus, $B(0)$ and $B(t_0)$ have the same probability distribution, i.e.,

$$\langle\psi, \xi|P^{B(t_0)}(b)|\psi, \xi\rangle = \langle\psi, \xi|P^{B(0)}(b)|\psi, \xi\rangle, \tag{A27}$$

but the weak joint distribution operationally detects the disturbance on B, i.e.,

$$\langle\psi,\xi|P^{B(t_0)}(-1)P^{B(0)}(+1)|\psi,\xi\rangle = 1/4. \tag{A28}$$

In this case, we have $\eta(B,\mathbf{M},\rho) = \sqrt{2} \neq 0$ [49, p. S680]. However, this does not mean that η does not satisfy the soundness requirement, since $\mathbf{M}$ disturbs B in ρ according to Eq. (A28).

8. Universally valid error–disturbance relations

In the following, we abbreviate $\varepsilon(A,\mathbf{M},\rho)$ as $\varepsilon(A)$ and $\eta(B,\mathbf{M},\rho)$ as $\eta(B)$ where no confusion may occur.

In 2003, one of the authors [11] derived the relation

$$\varepsilon(A)\eta(B)+\varepsilon(A)\sigma(B)+\sigma(A)\eta(B) \ \geq\ \frac{1}{2}|\mathrm{Tr}([A,B]\rho)| \tag{A29}$$

holding for any pair of observables A, B, state $|\psi\rangle$, and measuring process $\mathbf{M}$. Later, Brancirard [14] and one of the authors [16] obtained a stronger EDR given by

$$\begin{aligned}&\varepsilon(A)^2\sigma(B)^2+\sigma(A)^2\eta(B)^2\\ &+2\varepsilon(A)\eta(B)\sqrt{\sigma(A)^2\sigma(B)^2-D_{AB}^2} \geq D_{AB}^2,\end{aligned} \tag{A30}$$

where

$$D_{AB} = \frac{1}{2}\mathrm{Tr}(|\sqrt{\rho}[A,B]\sqrt{\rho}|). \tag{A31}$$

In the case where $A^2 = B^2 = \mathbb{1}$ and $M^2 = \mathbb{1}$, the above relation can be strengthened as [14, 16]

$$\hat{\varepsilon}(A)^2+\hat{\eta}(B)^2+2\hat{\varepsilon}(A)\hat{\eta}(B)\sqrt{1-D_{AB}^2} \geq D_{AB}^2, \tag{A32}$$

where $\hat{\varepsilon}(A) = \epsilon(A)\sqrt{1-\frac{\epsilon(A)^2}{4}}$ and $\hat{\eta}(B) = \eta(B)\sqrt{1-\frac{\eta(B)^2}{4}}$. In the case where

$$A=\sigma_z,\ B=\sigma_x,\ \langle\sigma_z(0)\rangle_\rho = \langle\sigma_x(0)\rangle_\rho = 0, \tag{A33}$$

the above inequality (A32) is reduced to the tight relation [14, 16]

$$\left(\varepsilon(\sigma_z)^2-2\right)^2+\left(\eta(\sigma_x)^2-2\right)^2 \leq 4 \tag{A34}$$

as depicted in FIG 1.

Appendix B: Gaussian wave packets

In this appendix, we review relations between Gaussian states and inequalities. Let Z and P be the canonical position and momentum observables, respectively, of a one-dimensional quantum system. These observables satisfy the usual canonical commutation relation, $[Z,P]=i\hbar$. Here, we only consider a vector state denoted by ψ. However, some of the results in this appendix can easily be generalized to mixed states.

1. Schrödinger inequality

For the variances of the position and momentum, the following inequality holds [50]:

$$\mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P) \geq \frac{\left(\langle\{Z,P\}\rangle_\psi - 2\langle Z\rangle_\psi\langle P\rangle_\psi\right)^2+\hbar^2}{4}. \tag{B1}$$

Inequality (B1) is known as the Schrödiner inequality. The proof proceeds as follows. First, we consider the case $\langle Z\rangle_\psi = \langle P\rangle_\psi = 0$. Then, we have

$$\mathrm{Im}\,\langle Z\psi, P\psi\rangle = \frac{1}{2i}\langle[Z,P]\rangle_\psi = \hbar/2 \tag{B2}$$

$$\mathrm{Re}\,\langle Z\psi, P\psi\rangle = \frac{1}{2}\langle\{Z,P\}\rangle_\psi. \tag{B3}$$

Consequently, we have

$$|\langle Z\psi, P\psi\rangle|^2 = \frac{\left(\langle\{Z,P\}\rangle_\psi\right)^2+\hbar^2}{4}. \tag{B4}$$

On the other hand, according to the Cauchy–Schwarz inequality,

$$|\langle Z\psi, P\psi\rangle|^2 \leq \langle Z^2\rangle_\psi\langle P^2\rangle_\psi = \mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P). \tag{B5}$$

Hence, the Schrödinger inequality (B1) holds if $\langle Z\rangle_\psi = \langle P\rangle_\psi = 0$ holds. We can obtain the proof for the general case by substituting Z and P into $Z-\langle Z\rangle_\psi$ and $P-\langle P\rangle_\psi$, respectively. This concludes the proof.

The equation in this inequality holds if and only if

$$\left(Z-\langle Z\rangle_\psi\right)\psi = c\left(P-\langle P\rangle_\psi\right)\psi \tag{B6}$$

for some complex number c. From the condition above, we obtain the differential equation for the wave function as

$$\frac{d}{dz}\psi(z) = -2k\left[z-\left(\langle Z\rangle_\psi+\frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]\psi(z), \tag{B7}$$

where k is a complex number. Therefore, we have

$$\psi(z) = A\exp\left(-k\left[z-\left(\langle Z\rangle_\psi+\frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right), \tag{B8}$$

where A is a constant. Since the wave function should be normalizable, constant k must satisfy $\mathrm{Re}\,k>0$.

2. Kennard inequality

The inequality, which is known as the Kennard inequality [2],

$$\mathrm{Var}_\psi(Z)\mathrm{Var}_\psi(P) \geq \hbar^2/4, \tag{B9}$$

can be derived from the Schrödinger inequality (B1). The equality in Eq. (B9) holds if and only if $2i\hbar k\left(Z-\langle Z\rangle_\psi\right)\psi =$

$(P - \langle P\rangle_\psi)\,\psi$ for some positive real number k. A wave function ψ satisfies the equality in the Kennard inequality (B9) if and only if ψ has the form

$$\psi(z) = A\exp\left(-k\left[z - \left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right) \quad \text{(B10)}$$

for some positive real number k. This wave function has the same form as that of Eq. (B8) except for the condition of the constant k. i.e., the constant k in Eq. (B8) is a complex number with a positive real part whereas the constant k in Eq. (B10) is a positive real number. The state in Eq.(B10) is known as the minimal uncertainty state.

3. Squeezed state

For any two complex numbers, μ and ν, satisfying $|\mu|^2 - |\nu|^2 = 1$, squeezed operator $c_{\mu,\nu}$ is defined as

$$c_{\mu,\nu} := \mu a + \nu a^\dagger, \quad \text{(B11)}$$

where a and $a^\dagger$ are the annihilation and creation operators, respectively.

$$a := \sqrt{\frac{m\omega}{2\hbar}}Z + i\sqrt{\frac{1}{2\hbar m\omega}}P. \quad \text{(B12)}$$

Here, m and ω are the mass and angular frequency of the corresponding harmonic oscillator, respectively. A coherent state [51] is defined as the eigenstate of the annihilation operator, a, in Eq. (B12). A squeezed state [52] is defined as the eigenstate of squeezed operator $c_{\mu,\nu}$,

$$c_{\mu,\nu}\psi = \lambda\psi. \quad \text{(B13)}$$

By this definition, the wave function of every squeezed state satisfies the differential equation,

$$\left[(\mu+\nu)\sqrt{\frac{m\omega}{2\hbar}}z + (\mu-\nu)\sqrt{\frac{\hbar}{2m\omega}}\frac{d}{dz}\right]\psi(z) = \lambda\psi(z). \quad \text{(B14)}$$

The solution of this differential equation is

$$\psi(z) := A\exp\left[-\frac{m\omega}{2\hbar}\frac{\mu+\nu}{\mu-\nu}\left(z - \sqrt{\frac{2\hbar}{m\omega}}\frac{\lambda}{\mu-\nu}\right)^2\right]. \quad \text{(B15)}$$

Hence, the equality in the Schrödinger inequality (B1) holds for squeezed states.

Next, let us consider the relation between these parameters and the mean values of the position and momentum. By comparing the two formulas, (B8) and (B15), we have

$$\langle Z\rangle_\psi + \frac{i}{m\omega}\frac{\mu-\nu}{\mu+\nu}\langle P\rangle_\psi = \sqrt{\frac{2\hbar}{m\omega}}\frac{\lambda}{\mu-\nu}. \quad \text{(B16)}$$

Taking the imaginary part, we have

$$\langle P\rangle_\psi = \sqrt{2\hbar m\omega}|\mu+\nu|^2 \mathrm{Im}\left(\frac{\lambda}{\mu-\nu}\right), \quad \text{(B17)}$$

$$\langle Z\rangle_\psi = \sqrt{\frac{2\hbar}{m\omega}}\mathrm{Re}\left(\frac{(\mu+\nu)(\mu^*-\nu^*)}{\mu-\nu}\lambda\right). \quad \text{(B18)}$$

Next, let us calculate the variances of the position and momentum and the correlation $\langle\{Z,P\}\rangle_\psi$. Setting $\tilde{z} = z - \langle Z\rangle_\psi$, we have

$$\begin{aligned}
&\mathrm{Var}(Z)\\
&= |A|^2\int_{-\infty}^{\infty}\tilde{z}^2\exp\Bigg(-\frac{m\omega}{\hbar}\\
&\times \mathrm{Re}\left[\frac{\mu-\nu}{\mu+\nu}\left(\frac{\mu+\nu}{\mu-\nu}\tilde{z}+\frac{i}{m\omega}\langle P\rangle_\psi\right)^2\right]\Bigg)d\tilde{z}\\
&= \frac{\hbar}{2m\omega}|\mu-\nu|^2. \qquad \text{(B19)}
\end{aligned}$$

To calculate the variance of the momentum, it is convenient to obtain the Fourier transform of the wave function, $\tilde{\psi}(\tilde{z}) := \psi(\tilde{z} + \langle Z\rangle_\psi)$,

$$\begin{aligned}
\hat{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\tilde{\psi}(\tilde{z})\exp(ip\tilde{z}/\hbar)d\tilde{z}\\
&= \hat{A}\exp\left[-\frac{1}{2\hbar m\omega}\frac{\mu-\nu}{\mu+\nu}\left(p - \langle P\rangle_\psi\right)^2\right], \qquad \text{(B20)}
\end{aligned}$$

where $\hat{A}$ is the normalization constant. Consequently, we have

$$\begin{aligned}
\mathrm{Var}(P) &= \langle(P - \langle P\rangle_\psi)^2\rangle_\psi\\
&= |A|^2\int_{-\infty}^{\infty}\tilde{p}^2\exp\left[-\frac{1}{\hbar m\omega}\mathrm{Re}\left(\frac{\mu-\nu}{\mu-\nu}\right)\tilde{p}^2\right]d\tilde{p}\\
&= \frac{\hbar m\omega}{2}|\mu+\nu|^2. \qquad \text{(B21)}
\end{aligned}$$

Finally, we calculate the correlation term,

$$\begin{aligned}
&\langle\{Z - \langle Z\rangle_\psi, P - \langle P\rangle_\psi\}\rangle_\psi\\
&= \langle\{Z - \langle Z\rangle_\psi, P\}\rangle_\psi\\
&= 2\mathrm{Re}\,\langle\tilde{Z}\psi, P\psi\rangle\\
&= 2\mathrm{Re}\Bigg(|A|^2 im\omega\\
&\times\int_{-\infty}^{\infty}\frac{\mu+\nu}{\mu-\nu}\tilde{z}^2\exp\left[-\frac{m\omega}{\hbar}\mathrm{Re}\left(\frac{\mu+\nu}{\mu-\nu}\tilde{z}^2\right)\right]d\tilde{z}\Bigg)\\
&= 2\hbar\mathrm{Im}\,(\mu^*\nu). \qquad \text{(B22)}
\end{aligned}$$

The coherent state is defined as the eigenstate of the annihilation operator. Using the results of the calculation above, the corresponding wave function is

$$\psi(z) = A\exp\left[-\frac{m\omega}{2\hbar}\left(z - \sqrt{\frac{2\hbar}{m\omega}}\lambda\right)^2\right], \quad \text{(B23)}$$

where λ is the corresponding eigenvalue of the annihilation operator. Thus, every coherent state satisfies the equation in the Schrödinger inequality (B1) and the Kennard inequality (B9).

Since $\dfrac{\mu+\nu}{\mu-\nu}$ moves all over the right half-plane of the complex plane as μ and ν move all over the complex plane satisfying $|\mu|^2-|\mu|^2=1$, the union of all squeezed states and coherent states coincides with the states that satisfy the Schrödinger inequality (B1), namely, $\mathcal{G}$.

4. Contractive state

The contractive state is introduced by Yuen [7] as a squeezed state whose correlation term is negative. This state contracts during some period of time if it evolves freely. To see this, let us calculate the variance of the position in the Heisenberg picture. The position operator $Z(t)$ at time t in the Heisenberg picture is

$$\begin{aligned} Z(t) &= \exp\left[-\frac{t}{2i\hbar m}P(t)^2\right] Z(0) \exp\left[\frac{t}{2i\hbar m}P(t)^2\right] \\ &= Z(0) + \frac{t}{m}P(0). \end{aligned} \tag{B24}$$

Hence, we have

$$\begin{aligned} &\mathrm{Var}_\psi\left(Z(t)\right) \\ &= \left\langle \left(Z(0) + \frac{t}{m}P(0) - \langle Z(0) + \frac{t}{m}P(0)\rangle_\psi\right)^2 \right\rangle_\psi \\ &= \frac{t^2}{m^2}\mathrm{Var}_\psi\left(P(0)\right) + \mathrm{Var}_\psi\left(P(0)\right) \\ &+ \frac{t}{m}\left\langle \{Z(0)-\langle Z(0)\rangle_\psi, P(0)-\langle P(0)\rangle_\psi\}\right\rangle_\psi. \end{aligned} \tag{B25}$$

Therefore, if the state is a contractive state, the variance of the position contracts until the time

$$t = -\frac{m\langle\{Z(0)-\langle Z(0)\rangle_\psi, P(0)-\langle P(0)\rangle_\psi\}\rangle_\psi}{2\langle P(0)^2\rangle_\psi}. \tag{B26}$$

5. Covariance matrix formalism

Recently, the covariace matrix is used in order to characterize Gaussian states [53]. For a single-mode Gaussian state,

$$\psi(z) = A\exp\left(-k\left[z-\left(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi\right)\right]^2\right), \tag{B27}$$

the covariance matrix, V, is defined as

$$\begin{aligned} V &= \begin{pmatrix} \mathrm{Var}_\psi\left(Z\right) & \mathrm{Cor}_\psi(Z,P) \\ \mathrm{Cor}_\psi(Z,P) & \mathrm{Var}_\psi\left(P\right) \end{pmatrix} \\ &= \begin{pmatrix} \left(4\mathrm{Re}(k)\right)^{-1} & -\frac{\hbar\mathrm{Im}(k)}{\mathrm{Re}(k)} \\ -\frac{\hbar\mathrm{Im}(k)}{\mathrm{Re}(k)} & \frac{\hbar^2|k|^2}{\mathrm{Re}(k)} \end{pmatrix}. \end{aligned} \tag{B28}$$

Here, we used the abbreviation,

$$\mathrm{Cor}_\psi(Z,P) = \langle\{Z-\langle Z\rangle_\psi, P-\langle P\rangle_\psi\}\rangle_\psi. \tag{B29}$$

6. Summary

We have discussed the relation between the inequalities and the subclasses of Gaussian states whose wave functions are of the form

$$\psi(z) = A\exp\left(-k\left[z-(\langle Z\rangle_\psi + \frac{i}{2\hbar k}\langle P\rangle_\psi)\right]^2\right) \tag{B30}$$

and obtained the relations shown in Table. I. FIG. 4 represents the inclusion relation between the subsets of the set of Gaussian wave packets.

TABLE I. Classification of Gaussian states in terms of parameter k.

k	Name of state	inequality whose equality holds
$\mathrm{Re}\,k>0$	Squeezed	Schrödinger
$\mathrm{Re}\,k>0$ and $\mathrm{Im}\,k>0$	Contractive	Schrödinger
$\mathrm{Re}\,k>0$ and $\mathrm{Im}\,k=0$	Minimal uncertainty	Kennard

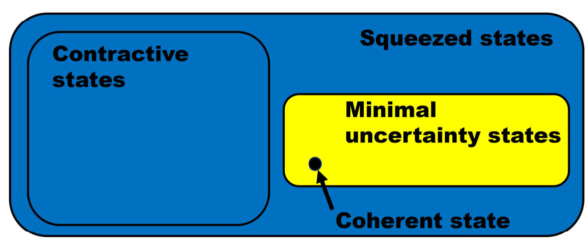

FIG. 4. Inclusion relation of the subsets of wave functions. A wave function is in the yellow region if and only if the equality in the Kennard inequality holds. A wave function is in the blue or yellow region if and only if the equality in the Schrödinger inequality holds.

Appendix C: Time evolution of Gaussian wave packets

In this appendix, we discuss the time evolution of the probability density of a Gaussian wave packet during free evolution.

The wave function under consideration is the Gaussian wave packet derived in the previous section,

$$\psi(z) := A \exp\left[-kz^2\right], \tag{C1}$$

where k is a complex number with a positive real part. For simplicity, we consider only the case in which the mean values of the position and momentum are zero. applying the Fourier transform $\mathfrak{F}$ successively, we obtain

$$\begin{aligned}
&\exp\left(\frac{t}{2i\hbar m}P^2\right)\psi(z) \\
&= \mathfrak{F}^{-1}\exp\left(\frac{t}{2i\hbar m}p^2\right)\hat{A}\exp\left(-\frac{p^2}{4k\hbar^2}\right) \\
&= \frac{\hat{A}}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\exp\left[\left(\frac{t}{2i\hbar m}-\frac{1}{k\hbar^2}\right)p^2 - ipz/\hbar\right]dp \\
&= N\exp\left[-\frac{z^2}{\left(k^{-1}-\frac{2\hbar t}{im}\right)}\right],
\end{aligned} \tag{C2}$$

where N is the normalization constant. Thus, the probability density, $\Pr(z)$, at time t has the form

$$\Pr(z) = |N|^2 \exp\left(-rz^2\right) \tag{C3}$$

for some positive real number r. That is, we have again obtained a Gaussian distribution. Since the variance of the Gaussian distribution is

$$\left\langle Z(t)^2\right\rangle_\psi = \left\langle\left(Z(0)+\frac{t}{m}P(0)\right)^2\right\rangle_\psi, \tag{C4}$$

we have

$$\Pr(z) = |N|^2\exp\left(-\frac{z^2}{2\langle\left(Z(0)+\frac{t}{m}P(0)\right)^2\rangle_\psi}\right). \tag{C5}$$

Appendix D: Relationship between the Heisenberg picture and the Schrödinger picture

Let us consider the relation between the Heisenberg picture and the Schrödinger picture. Consider the time evolution of quantum system $\mathbf{S}$ described by $\mathcal{H}$. Let A be an observable of system $\mathbf{S}$ and state ψ. Denote by $E(A,\psi,t)$ the expectation value of the outcome of the measurement of observable A at time t, provided that system $\mathbf{S}$ is in state ψ at time 0. In the Schrödinger picture, state $\psi(t)$ evolves in time t as a solution of the Schrödinger equation by the time evolution operator, $U(t)$, as $\psi(t) = U(t)\psi$ with the initial condition, $U(0) = \mathbb{1}$, so that $E(A,\psi,t) = \langle\psi(t), A\psi(t)\rangle$ holds. The unitary operator $U^{\mathrm{S}}(t_2,t_1)$ describing the time evolution from time $t = t_1$ to $t = t_2$ $(t_1 \le t_2)$ in the Schrödinger picture is defined by

$$U^{\mathrm{S}}(t_2,t_1) = U(t_2)U^\dagger(t_1). \tag{D1}$$

Then, we have

$$U^{\mathrm{S}}(t_2,t_1)\psi(t_1) = \psi(t_2), \tag{D2}$$

$$U^{\mathrm{S}}(t_3,t_2)U^{\mathrm{S}}(t_2,t_1) = U^{\mathrm{S}}(t_3,t_1). \tag{D3}$$

In the Heisenberg picture, observable $A(t)$ evolves in time t by the time evolution operator $U(t)$ as $A(t) = U(t)^\dagger AU(t)$, so that $E(A,\psi,t) = \langle\psi, A(t)\psi\rangle$ holds. The unitary operator, $U^{\mathrm{H}}(t_2,t_1)$, describing the time evolution from time $t = t_1$ to $t = t_2$ $(t_1 \le t_2)$ in the Heisenberg picture is defined by

$$U^{\mathrm{H}}(t_2,t_1) = U^\dagger(t_1)U(t_2). \tag{D4}$$

Then, we have

$$U^{\mathrm{H}}(t_2,t_1)^\dagger A(t_1)U^{\mathrm{H}}(t_2,t_1) = A(t_2), \tag{D5}$$

$$\alpha^{\mathrm{H}}(t_3,t_2)\alpha^{\mathrm{H}}(t_2,t_1) = \alpha^{\mathrm{H}}(t_3,t_1), \tag{D6}$$

where

$$\alpha^{\mathrm{H}}(t_2,t_1)A = U^{\mathrm{H}}(t_2,t_1)^\dagger AU^{\mathrm{H}}(t_2,t_1). \tag{D7}$$

We have the following relations between the Schrödinger picture and the Heisenberg picture.

$$U(t) = U^{\mathrm{S}}(t,0) = U^{\mathrm{H}}(t,0). \tag{D8}$$

$$U^{\mathrm{H}}(t_2,t_1) = U(t_1)^\dagger U^{\mathrm{S}}(t_2,t_1)U(t_1). \tag{D9}$$

Let $f(A_1,\dots,A_n,t,s)$ be a function of observables $A_1,\dots,A_n$ and real numbers t,s. If

$$U^{\mathrm{S}}(t_2,t_1) = f(A_1,\dots,A_n,t_1,t_2), \tag{D10}$$

then

$$U^{\mathrm{H}}(t_2,t_1) = f(A_1(t_1),\dots,A_n(t_1),t_1,t_2). \tag{D11}$$

Appendix E: Solutions of Heisenberg equations of motion for $Z(t)$, $P(t)$, $\sigma_x(t)$, $\sigma_y(t)$, and $\sigma_z(t)$

To consider the time evolution from time $t = \Delta t$ to time $\Delta t + \tau$, suppose $\Delta t \le t \le \Delta t + \tau$. By the Heisenberg equation of motion, position operator $Z(t)$ satisfies

$$\frac{d}{dt}Z(t) = \frac{1}{i\hbar}[Z(t), \frac{1}{2m}P(t)^2] = \frac{1}{m}P(t). \tag{E1}$$

Thus, we have

$$Z(t) = Z(\Delta t) + \frac{1}{m}\int_{\Delta t}^{t} P(t')dt'. \tag{E2}$$

In contrast, $P(t)$ does not change since $[P(t), H(t)] = 0$. Consequently, we have

$$Z(t) = Z(\Delta t) + \frac{t-\Delta t}{m}P(\Delta t), \tag{E3}$$

$$P(t) = P(\Delta t). \tag{E4}$$

Since $\sigma_z(t)$ and $\sigma_x(t)$ commute with $H(t)$, we have

$$\sigma_z(t) = \sigma_z(\Delta t), \quad \sigma_x(t) = \sigma_x(\Delta t). \tag{E5}$$

To describe the observables at time $t = \Delta t$ in terms of the observables at time $t = 0$, suppose $0 \le t \le \Delta t$. With the Heisenberg equations of motion, we obtain

$$\frac{d}{dt}Z(t) = \frac{1}{i\hbar}[Z(t), H(t)] = \frac{1}{m}P(t), \tag{E6}$$

and

$$Z(\Delta t) = Z(0) + \frac{1}{m}\int_0^{\Delta t} P(t)dt. \tag{E7}$$

On the other hand, we have

$$\frac{d}{dt}P(t) = \frac{1}{i\hbar}[P(t), H(t)] = -\mu B_1 \sigma_z(t). \tag{E8}$$

Now, $\sigma_z(t)$ commutes with Hamiltonian $H(t)$. Hence, we have

$$\sigma_z(t) = \sigma_z(0). \tag{E9}$$

Consequently, we have

$$P(t) = P(0) - \mu B_1 t \sigma_z(0), \tag{E10}$$

$$Z(t) = Z(0) + \frac{t}{m}P(0) - \frac{\mu B_1 t^2}{2m}\sigma_z(0). \tag{E11}$$

Therefore, we have

$$\begin{aligned} Z(\Delta t+\tau) &= Z(0) + \frac{\Delta t+\tau}{m}P(0) \\ &\quad - \frac{\mu B_1 \Delta t}{m}\left(\tau + \frac{\Delta t}{2}\right)\sigma_z(0), \end{aligned} \tag{E12}$$

$$P(\Delta t+\tau) = P(0) - \mu B_1 \Delta t \sigma_z(0), \tag{E13}$$

$$\sigma_z(\Delta t+\tau) = \sigma_z(0). \tag{E14}$$

Next, we calculate the x- and y-components of the spin of the particle at time $t = \Delta t + \tau$. Since Hamiltonian $H(t)$ from time $t = \Delta t$ to time $\Delta t + \tau$ commutes with $\sigma_x(t)$ and $\sigma_y(t)$, we have

$$\sigma_x(t) = \sigma_x(\Delta t), \tag{E15}$$

$$\sigma_y(t) = \sigma_y(\Delta t) \tag{E16}$$

if $\Delta t \le t \le \Delta t + \tau$, and it suffices to calculate $\sigma_x(\Delta t)$ and $\sigma_y(\Delta t)$.

Suppose $0 \le t \le \Delta t$. By the Heisenberg equations of motion, we have

$$\begin{aligned} \frac{d}{dt}\sigma_x(t) &= \frac{1}{i\hbar}\left[\sigma_x(t), H(t)\right] \\ &= \frac{1}{i\hbar}\left[\sigma_x(t), \frac{P(t)^2}{2m} + \mu\left(B_0 + B_1 Z(t)\right)\sigma_z(t)\right] \\ &= \frac{\mu}{i\hbar}\left(B_0 + B_1 Z(t)\right)\left(-2i\sigma_y(t)\right) \\ &= -\frac{2\mu}{\hbar}\left(B_0 + B_1 Z(t)\right)\sigma_y(t). \end{aligned} \tag{E17}$$

Similarly, we have

$$\begin{aligned} \frac{d}{dt}\sigma_y(t) &= \frac{1}{i\hbar}\left[\sigma_y(t), H(t)\right] \\ &= \frac{1}{i\hbar}\left[\sigma_y(t), \frac{P(t)^2}{2m} + \mu\left(B_0 + B_1 Z(t)\right)\sigma_z(t)\right] \\ &= \frac{\mu}{i\hbar}\left(B_0 + B_1 Z(t)\right)\left(2i\sigma_x(t)\right) \\ &= \frac{2\mu}{\hbar}\left(B_0 + B_1 Z(t)\right)\sigma_x(t). \end{aligned} \tag{E18}$$

Now, let us introduce σ_+ and σ_- by

$$\sigma_+(t) = \frac{1}{\sqrt{2}}\left(\sigma_x(t) + i\sigma_y(t)\right), \tag{E19}$$

$$\sigma_-(t) = \frac{1}{\sqrt{2}}\left(\sigma_x(t) - i\sigma_y(t)\right). \tag{E20}$$

From Eqs. (E17) and (E18), we have

$$\frac{d}{dt}\sigma_\pm(t) = \pm\frac{2\mu i}{\hbar}\left[B_0 + B_1\left(U^\dagger(t)Z(0)U(t)\right)\right]\sigma_\pm(t). \tag{E21}$$

Let

$$\gamma_\pm(t) = U(t)\sigma_\pm(t) = \exp\left[\frac{H(0)}{i\hbar}t\right]\sigma_\pm(t). \tag{E22}$$

The left-hand side (LHS) and right-hand side (RHS) of Eq. (E21) satisfy

$$\begin{aligned} \mathrm{LHS} &= \frac{d}{dt}U(-t)\gamma_\pm(t) \\ &= -\frac{H(0)}{i\hbar}U(-t)\gamma_\pm(t) + U(-t)\frac{d}{dt}\gamma_\pm(t), \end{aligned} \tag{E23}$$

$$\begin{aligned} \mathrm{RHS} &= \pm\frac{2\mu i}{\hbar}U^\dagger(t)\left(B_0 + B_1 Z(0)\right)U(t)U^\dagger(t)\gamma_\pm(t) \\ &= \pm\frac{2\mu i}{\hbar}U(-t)\left(B_0 + B_1 Z(0)\right)\gamma_\pm(t). \end{aligned} \tag{E24}$$

Hence, we have

$$\frac{d}{dt}\gamma_\pm(t) = \left[\frac{H(0)}{i\hbar} \pm \frac{2\mu i}{\hbar}\left(B_0 + B_1 Z(0)\right)\right]\gamma_\pm(t). \tag{E25}$$

The solution of the above differential equation is given by

$$\gamma_\pm(t) = \exp\left\{\frac{it}{\hbar}\left[-H(0) \pm 2\mu\left(B_0 + B_1 Z(0)\right)\right]\right\}\gamma_\pm(0). \tag{E26}$$

Since $\gamma_\pm(0) = \sigma_\pm(0)$, we have

$$\begin{aligned} \sigma_\pm(t) &= \exp\left(\frac{it}{\hbar}H(0)\right) \\ &\times \exp\left\{\frac{it}{\hbar}\left[-H(0) \pm 2\mu\left(B_0 + B_1 Z(0)\right)\right]\right\}\sigma_\pm(0). \end{aligned} \tag{E27}$$

Using the Baker-Campbell-Hausdorff formula [54] we have

$$\begin{aligned} \exp A \exp B &= \exp\Big[(A+B) + \frac{1}{2}\left[A, B\right] \\ &\quad + \frac{1}{12}\left(\left[\left[A,B\right],B\right] - \left[\left[A,B\right],A\right]\right) + \cdots\Big]. \end{aligned} \tag{E28}$$

Hence, for

$$A = \frac{it}{\hbar}H(0), \tag{E29}$$

$$B = \frac{it}{\hbar}\left[-H(0) \pm 2\mu\left(B_0 + B_1 Z(0)\right)\right], \tag{E30}$$

we have

$$[A,B]=\left[\frac{it}{\hbar}H(0),\frac{it}{\hbar}\left(-H(0)\pm 2\mu\left(B_0+B_1Z(0)\right)\right)\right]$$
$$=-\frac{t^2}{\hbar^2}\left[\frac{1}{2m}P(0)^2,\pm 2\mu\left(B_0+B_1Z(0)\right)\right]$$
$$=\pm\frac{2i\mu B_1t^2}{m\hbar}P(0), \tag{E31}$$

$$[[A,B],A]=\left[\pm\frac{2i\mu B_1t^2}{m\hbar}P(0),\frac{it}{\hbar}H(0)\right]$$
$$=\mp\frac{2\mu B_1t^3}{m\hbar^2}\left[P(0),\mu\left(B_0+B_1Z(0)\right)\sigma_z(0)\right]$$
$$=\pm\frac{2i\mu^2B_1^2t^3}{m\hbar}\sigma_z(0), \tag{E32}$$

$$[[A,B],B]$$
$$=\left[\pm\frac{2i\mu B_1t^2}{m\hbar}P(0),\frac{it}{\hbar}\left[-H(0)\pm 2\mu\left(B_0+B_1Z(0)\right)\right]\right]$$
$$=\mp\frac{2i\mu^2B_1^2t^3}{m\hbar}\sigma_z(0)$$
$$\mp\frac{2\mu B_1t^3}{m\hbar^2}\left[P(0),\pm 2\mu\left(B_0+B_1Z(0)\right)\sigma_z(0)\right]$$
$$=\frac{2i\mu^2B_1^2t^3}{m\hbar}\left(2\mp\sigma_z(0)\right). \tag{E33}$$

The commutators of the higher orders "$\cdots$" in Eq. (E28) are 0 since the third commutators, $[[A,B],A]$ and $[[A,B],B]$, commute with A and B, respectively.

Let

$$R(t)=\frac{\mu^2B_1^2t^3}{3m\hbar}, \tag{E34}$$
$$S(t)=\frac{2\mu t}{\hbar}\left[B_0+B_1\left(Z+\frac{t}{2m}P\right)\right]. \tag{E35}$$

We have

$$\sigma_\pm(t)=\exp i\left([R(t)\pm S(t)]\mathbb{1}\mp R(t)\sigma_z(0)\right)\sigma_\pm(0). \tag{E36}$$

Since

$$\sigma_+(0)=\frac{1}{\sqrt{2}}\left(\sigma_z(0)+i\sigma_y(0)\right)=\begin{pmatrix}0&\sqrt{2}\\0&0\end{pmatrix}, \tag{E37}$$
$$\sigma_-(0)=\frac{1}{\sqrt{2}}\left(\sigma_z(0)-i\sigma_y(0)\right)=\begin{pmatrix}0&0\\\sqrt{2}&0\end{pmatrix}, \tag{E38}$$

we have

$$\sigma_+(t)$$
$$=\begin{pmatrix}\exp iS(t)&0\\0&\exp i(S(t)+2R(t))\end{pmatrix}\begin{pmatrix}0&\sqrt{2}\\0&0\end{pmatrix}$$
$$=\exp iS(t)\sigma_+(0), \tag{E39}$$

$$\sigma_-(t)$$
$$=\begin{pmatrix}\exp i(-S(t)+2R(t))&0\\0&\exp -iS(t)\end{pmatrix}\begin{pmatrix}0&0\\\sqrt{2}&0\end{pmatrix}$$
$$=\exp -iS(t)\sigma_-(0). \tag{E40}$$

Therefore, $\sigma_x(t)$ and $\sigma_y(t)$ from time $t=0$ to time $t=\Delta t$ are

$$\sigma_x(t)=\frac{1}{\sqrt{2}}\left(\sigma_+(t)+\sigma_-(t)\right)$$
$$=\begin{pmatrix}0&\exp iS(t)\\\exp -iS(t)&0\end{pmatrix}, \tag{E41}$$

$$\sigma_y(t)=-\frac{i}{\sqrt{2}}\left(\sigma_+(t)-\sigma_-(t)\right)$$
$$=\begin{pmatrix}0&-i\exp iS(t)\\i\exp -iS(t)&0\end{pmatrix}. \tag{E42}$$

Appendix F: Supremum of the function $W_\lambda(t)$

Let us consider the supremum of the function in section V,

$$W_{\xi_\lambda}(\tau)=\alpha\left(\tau+\frac{\Delta t}{2}\right)\left[a+b(\Delta t+\tau)+c(\Delta t+\tau)^2\right]^{-1/2}. \tag{F1}$$

Here we put $\alpha=\frac{\mu B_1\Delta t}{\sqrt{2}m}$, $a=\langle Z^2\rangle$, $b=\frac{\langle\{Z,P\}\rangle}{m}$, and $c=\frac{\langle P^2\rangle}{m^2}$. The derivative of function $W_{\xi_\lambda}(\tau)$ is

$$\frac{d}{d\tau}W_\lambda(\tau)$$
$$=\frac{\alpha}{4}\left[a+b(\Delta t+\tau)+c(\Delta t+\tau)^2\right]^{-3/2}$$
$$\times\left[2(b+c\Delta t)(\Delta t+\tau)+4a+b\Delta t\right]. \tag{F2}$$

Hence, $W_{\xi_\lambda}(t)$ assumes the maximum value at $\tau=\tau_0=-\frac{4a+3b\Delta t+2c\Delta t^2}{2(b+c\Delta t)}\geq 0$ if the following conditions hold.

(i) $W'(0)>0$.

(ii) $2b+2c\Delta t<0$.

Condition (i) holds automatically. In fact, (i) is equivalent to condition

$$4a+3b\Delta t+2c\Delta t^2\geq 0. \tag{F3}$$

Now let us consider function

$$f(t)=4a+3bt+2ct^2. \tag{F4}$$

This function assumes the minimum value at $t=-\frac{3b}{4c}$,

$$\begin{aligned} f(t) &\geq f\left(-\frac{3b}{4c}\right) \\ &= \frac{32ac-9b^2}{8c} \\ &= \frac{9}{8c}(4ac-b^2) - \frac{4ac}{8c} \\ &\geq \frac{9\hbar^2}{8cm^2} - \frac{\hbar^2}{8cm^2} \\ &= \frac{\hbar^2}{cm^2} \\ &> 0. \end{aligned} \tag{F5}$$

Therefore, condition (i) is satisfied automatically. Here, we use Schrödinger inequality (B1). Hence, if condition (ii) holds, function $W_\lambda(\tau)$ assumes the maximum value at $\tau = \tau_0 \geq 0$. The maximum value of $W_{\xi_\lambda}(\tau)$ for $\tau \geq 0$ is

$$\begin{aligned} W_{\xi_\lambda}(\tau_0) &= -\alpha \frac{4a+2b\Delta t + c\Delta t^2}{2(b+c\Delta t)} \\ &\quad \times \left[a + b(\Delta t + \tau_0) + c(\Delta t+\tau_0)^2\right]^{-1/2} \\ &= \alpha \left(4a+2b\Delta t+c\Delta t^2\right)^{1/2}\left(4ac-b^2\right)^{-1/2} \\ &= \frac{2\alpha m}{\hbar}\left[a + b\frac{\Delta t}{2} + c\left(\frac{\Delta t}{2}\right)^2\right]^{1/2} \\ &= \frac{\sqrt{2}\mu B_1 \Delta t}{\hbar}\left\langle \left(Z + \frac{\Delta t}{2m}P\right)^2\right\rangle_{\xi_\lambda}^{1/2}. \end{aligned} \tag{F6}$$

If condition (ii) does not hold, function $W_{\xi_\lambda}(\tau)$ increases monotonically and we have

$$\sup_{\tau\geq 0} W_{\xi_\lambda}(\tau) = \lim_{\tau\to\infty} W_{\xi_\lambda}(\tau) = \frac{\mu B_1 \Delta t}{\sqrt{2\langle P^2\rangle_{\xi_\lambda}}}. \tag{F7}$$

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927), [English translation by J. A. Wheeler and W. H. Zurek, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton UP, Princeton, NJ, 1983), pp. 62–84].
[2] E. Kennard, Z. Phys. **44**, 326 (1927).
[3] M. Ozawa, Curr. Sci. **109** (2015).
[4] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[5] V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne, Science **209**, 547 (1980).
[6] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, Rev. Mod. Phys. **52**, 341 (1980).
[7] H. P. Yuen, Phys. Rev. Lett. **51**, 719 (1983).
[8] C. M. Caves, Phys. Rev. Lett. **54**, 2465 (1985).
[9] M. Ozawa, Phys. Rev. Lett. **60**, 385 (1988).
[10] M. Ozawa, in *Squeezed and nonclassical light*, edited by P. Tombesi and E. Pike (Plenum Press, New York, 1989) pp. 263–286, arXiv:1505.01083 [quant-ph].
[11] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).
[12] M. Ozawa, Phys. Lett. A **318**, 21 (2003).
[13] M. Ozawa, Ann. Phys. **311**, 350 (2004).
[14] C. Branciard, Proc. Natl. Acad. Sci. **110**, 6742 (2013).
[15] C. Branciard, Phys. Rev. A **89**, 022124 (2014).
[16] M. Ozawa, "Error–disturbance relations in mixed states," (2014), arXiv:1404.3388[quant-ph].
[17] M. Ozawa, npj Quantum Inf. **5**, 1 (2019).
[18] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Nat. Phys. **8**, 185 (2012).
[19] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, Phys. Rev. A **88**, 022110 (2013).
[20] B. Demirel, S. Sponar, G. Sulyok, M. Ozawa, and Y. Hasegawa, Phys. Rev. Lett. **117**, 140402 (2016).
[21] A. P. Lund and H. M. Wiseman, New J. Phys. **12**, 093011 (2010).
[22] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Phys. Rev. Lett. **109**, 100404 (2012).
[23] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, Sci. Rep. **3**, 2221 (2013).
[24] M. M. Weston, M. J. W. Hall, M. S. Palsson, H. M. Wiseman, and G. J. Pryde, Phys. Rev. Lett. **110**, 220402 (2013).
[25] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, Phys. Rev. Lett. **112**, 020402 (2014).
[26] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Phys. Rev. Lett. **112**, 020401 (2014).
[27] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. Lett. **111**, 160405 (2013).
[28] P. Busch, P. Lahti, and R. F. Werner, Rev. Mod. Phys. **86**, 1261 (2014).
[29] X.-M. Lu, S. Yu, K. Fujikawa, and C. H. Oh, Phys. Rev. A **90**, 042113 (2014).
[30] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, Phys. Rev. Lett. **112**, 050401 (2014).
[31] G. Sulyok, S. Sponar, B. Demirel, F. Buscemi, M. J. W. Hall, M. Ozawa, and Y. Hasegawa, Phys. Rev. Lett. **115**, 030401 (2015).
[32] W. Gerlach and O. Stern, Z. Phys. **8**, 110 (1922).
[33] W. Gerlach and O. Stern, Z. Phys. **9**, 349 (1922).
[34] W. Gerlach and O. Stern, Z. Phys. **9**, 353 (1922).
[35] D. Bohm, *Quantum theory* (Prentice-Hall, Englewood Cliffs, NJ, 1951).
[36] M. O. Scully, W. E. Lamb, and A. Barut, Found. Phys. **17**, 575 (1987).
[37] S. Cruz-Barrios and J. Gómez-Camacho, Phys. Rev. A **63**, 012101 (2000).
[38] G. Potel, F. Barranco, S. Cruz-Barrios, and J. Gómez-Camacho, Phys. Rev. A **71**, 052106 (2005).
[39] D. Home, A. K. Pan, M. M. Ali, and A. S. Majumdar, J. Phys. A: Math. Theor. **40**, 13975 (2007).
[40] P. Busch, T. Heinonen, and P. Lahti, Phys. Lett. A **320**, 261 (2004).
[41] M. Ozawa, Ann. Phys. **321**, 744 (2006).
[42] B. L. Schumaker, Phys. Rep. **135**, 317 (1986).

[43] M. Devereux, Can. J. Phys. **93**, 1382 (2015).

[44] C. F. Gauss, *Theory of the Combination of Observations Least Subject to Errors, Part One, Part Two, Supplement* (Society for Industrial and Applied Mathematics, 1995) [Originally published: *Theoria Combinationis Observationum: Erroribus Minimis Obnoxiae*, Pars prior, Pars posterior, Supplementum, (Societati Regiae Exhibita, 1821)].

[45] M. Ozawa, Phys. Lett. A **335**, 11 (2005).

[46] R. Jozsa, Phys. Rev. A **76**, 044103 (2007).

[47] M. Ozawa, Proc. SPIE **6244**, 62440Q (2006).

[48] K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. A **89**, 052108 (2014).

[49] M. Ozawa, J. Opt. B: Quantum Semiclass. Opt. **7**, S672 (2005).

[50] E. Schrödinger, Proceedings of The Prussian Academy of Sciences Physics–Mathematical Section XIX **19**, 296 (1930), [English translation is available at arXiv:9903100v1[quant-ph]].

[51] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[52] H. P. Yuen, Phys. Rev. A **13**, 2226 (1976).

[53] C. Weedbrook, S. Pirandola, R. G. Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S Lloyd, Rev. Mod. Phys. **84**, 621 (2012).

[54] H. F. Baker, Proc. London Math. Soc. (2) **3**, 24 (1905).