Violation of Heisenberg's error-disturbance uncertainty relation in neutron-spin measurements

Georg Sulyok,[1] Stephan Sponar,[1] Jacqueline Erhart,[1] Gerald Badurek,[1] Masanao Ozawa,[2] and Yuji Hasegawa[1]

[1]*Institute of Atomic and Subatomic Physics, Vienna University of Technology, 1020 Vienna, Austria*

[2]*Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya, Japan*

In its original formulation, Heisenberg's uncertainty principle dealt with the relationship between the error of a quantum measurement and the thereby induced disturbance on the measured object. Meanwhile, Heisenberg's heuristic arguments have turned out to be correct only for special cases. An alternative universally valid relation was derived by Ozawa in 2003. Here, we demonstrate that Ozawa's predictions hold for projective neutron-spin measurements. The experimental inaccessibility of error and disturbance claimed elsewhere has been overcome using a tomographic method. By a systematic variation of experimental parameters in the entire configuration space, the physical behavior of error and disturbance for projective spin-$\frac{1}{2}$ measurements is illustrated comprehensively. The violation of Heisenberg's original relation, as well as the validity of Ozawa's relation become manifest. In addition, our results conclude that the widespread assumption of a reciprocal relation between error and disturbance is not valid in general.

PACS number(s): 03.65.Ta, 03.75.Dg, 42.50.Xa, 03.67.−a

I. INTRODUCTION

The uncertainty principle, proposed by Heisenberg [1] in 1927, ranks without doubt among the most famous statements of modern physics. The content of the principle is often explained by simply saying "The more precisely the position is determined, the less precisely the momentum is known, and conversely [1]." It is usually understood that this leads to the impossibility of simultaneous measurements of the position and momentum of a particle and consequently to the rejection of determinism of the Newtonian mechanics in determining the future motion from the past state. However, the quantitative formulation of the uncertainty principle has been a target of debate for a long time ([2–10], and references therein).

Heisenberg's original formulation [1,11] can be read as the relation

$$\epsilon(Q)\eta(P) \geqslant \frac{\hbar}{2} \quad (1)$$

for the root-mean-square error $\epsilon(Q)$ of a measurement of the position observable Q and the root-mean-square disturbance $\eta(P)$ of the momentum observable P induced by the position measurement. However, modern textbooks usually explain the "uncertainty principle" as the relation

$$\sigma(Q)\sigma(P) \geqslant \frac{\hbar}{2}, \quad (2)$$

originally proved by Kennard [12] in 1927, for the standard deviations $\sigma(Q)$ and $\sigma(P)$ of the position observable Q and the momentum observable P in an arbitrary state ψ where the standard deviation is defined by $\sigma(A)^2 = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2$ for an observable A and a state ψ. Experimental investigations of the above relation can be found in [13–16].

Heisenberg actually derived Kennard's relation, Eq. (2), for Gaussian wave functions ψ, applied this relation to the state just after the Q measurement with error $\epsilon(Q)$ and disturbance $\eta(P)$, and concluded relation, Eq. (1), from the additional, implicit assumptions $\epsilon(Q) \geqslant \sigma(Q)$ and $\eta(P) \geqslant \sigma(P)$. However, his assumption $\epsilon(Q) \geqslant \sigma(Q)$ holds only for a restricted class of measurements [17]. The assumption $\eta(P) \geqslant \sigma(P)$ holds if the initial state is the momentum eigenstate state [18], but it

does not hold generally. Thus, his argument did not establish the universal validity of Eq. (1).

In 1929, Robertson [19] extended Kennard's relation, Eq. (2), to an arbitrary pair of observables A and B as

$$\sigma(A)\sigma(B) \geqslant \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle|. \quad (3)$$

The generalized form of Heisenberg's original relation, Eq. (1), would read

$$\epsilon(A)\eta(B) \geqslant \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle|. \quad (4)$$

The validity of this relation is known to be limited to specific circumstances [2,20–22]. In 2003, Ozawa [17,23–25], one of the authors of the present paper, thus proposed a new error-disturbance uncertainty relation

$$\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B) \geqslant \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle| \quad (5)$$

and proved the universal validity in the general theory of quantum measurements, where $\epsilon(A)$ is the root-mean-square error of an arbitrary measurement for an observable A, $\eta(B)$ is the root-mean-square disturbance on another observable B induced by the measurement, and $\sigma(A)$ and $\sigma(B)$ stand for the standard deviations of A and B in the state ψ just before the measurement.

Questions appeared on the accessibility of errors and disturbances in practical experiments, e.g., [26,27]. This difficulty has been overcome in two ways: the "three-state method" and the weak-measurement method. The tomographic three-state method is based on a formula to statistically determine error and disturbance, respectively, in a given initial state from the experimentally accessible data from three different auxiliary input states generated from the initial state [24]. The three-state method has already been successfully implemented in our previous publication [28]. The weak-measurement method is proposed [29] to quantify the error and disturbance exploiting the weak-measurement technique [30] based on the relation between mean-square differences and weak joint distributions [31]; an optical experiment was carried out recently [32].

In this paper, we report a general test of the validity of Ozawa's formulation in the framework of projective spin-$\frac{1}{2}$ measurements. Polarized neutrons experience successive spin measurements A and B, where the former measurement is detuned on purpose thus causing the measurement error of the A measurement. Likewise, the disturbance on the observable B also changes when the A measurement is detuned. In [28], we focused solely on a small subset of the possible configurations in parameter space where Heisenberg's original relation, Eq. (4), is always violated and the typical reciprocal trade-off between the error and the disturbance is maintained. Here, we give a more complete view on how systematic variations of the observables, the settings of the measurement apparatus, and the premeasurement system state influence error, disturbance, and premeasurement standard deviations. By studying the whole configuration space the physical meaning of error and disturbance for projective spin-$\frac{1}{2}$ measurements can be illustrated quite comprehensively. Parameter regions where Heisenberg's original formulation is violated as well as the behavior of Ozawa's relation become manifest: Not only the validity but also the analytic dependence of two notions of inequalities, Eqs. (4) and (5), on the experimental parameters is clearly seen. In addition, investigating error and disturbance in the whole configuration space shows the restricted validity of a reciprocal relation between these quantities in the spin-$\frac{1}{2}$ system.

II. THEORY

In order to derive the universally valid error-disturbance uncertainty relation it is convenient to discuss general models of measuring processes described in the quantum-mechanical framework. For a detailed account, we refer the reader to [17,23–25]. Here, we will sketch the main results necessary for later discussions on our experimental work.

A. Indirect measurement models

In this paper, we consider only *finite-type quantum measurements* such that the measured system is described by a finite dimensional Hilbert space and that the apparatus has a finite number of possible outcomes. We assume that every measuring apparatus has its own output variable **x**. The apparatus **A(x)** having the output variable **x** is assumed to determine the probability distribution $\Pr\{\mathbf{x} = m \| \rho\}$ of **x** on every input state ρ, and to determine the output state $\rho_{\{\mathbf{x}=m\}}$ for every input state ρ conditional upon each possible output **x** $= m$.

An *indirect measurement model* of an apparatus **A(x)** measuring a system **S** described by a Hilbert space $\mathcal{H}$ is specified by a quadruple $(\mathcal{K}, |\xi\rangle, U, M)$ consisting of a Hilbert space $\mathcal{K}$ describing the probe system **P**, a state vector $|\xi\rangle$ in $\mathcal{K}$ describing the initial state of **P**, a unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ describing the time evolution of the composite system **S** + **P** during the measuring interaction, and an observable M, called the *meter observable*, of **P** describing the meter of the apparatus [33].

The class of indirect measurement models is a universal class of models of quantum measurement [24,33] in the sense that for any apparatus **A(x)** with the output variable **x** there is

an indirect measurement model $(\mathcal{K}, |\xi\rangle, U, M)$ that determines the statistical properties of **A(x)** by

$$\Pr\{\mathbf{x} = m \| \rho\} \rho_{\{\mathbf{x}=m\}}$$
$$= \mathrm{Tr}_{\mathcal{K}}\{U^{\dagger}[\mathbb{1} \otimes E^{M}(m)]U(\rho \otimes |\xi\rangle\langle\xi|)\}, \qquad (6)$$

where $\mathrm{Tr}_{\mathcal{K}}$ stands for the partial trace over $\mathcal{K}$ and $E^{M}(m)$ the spectral projection of M corresponding to the real number m. From Eq. (6) we have

$$\Pr\{\mathbf{x} = m \| \rho\} = \mathrm{Tr}\{U^{\dagger}[\mathbb{1} \otimes E^{M}(m)]U(\rho \otimes |\xi\rangle\langle\xi|)\}, \qquad (7)$$

$$\rho_{\{\mathbf{x}=m\}} = \frac{\mathrm{Tr}_{\mathcal{K}}\{U^{\dagger}[\mathbb{1} \otimes E^{M}(m)]U(\rho \otimes |\xi\rangle\langle\xi|)\}}{\mathrm{Tr}\{U^{\dagger}[\mathbb{1} \otimes E^{M}(m)]U(\rho \otimes |\xi\rangle\langle\xi|)\}}. \qquad (8)$$

B. Measurement operators

In this paper, we treat only the case where the meter observable M has nondegenerate eigenvalues. In this case, M has a spectral decomposition $M = \sum_{m} m|m\rangle\langle m|$, where m varies over eigenvalues of M. Then, the apparatus **A(x)** has a family $\{M_m\}$ of operators, called the *measurement operators*, defined by $M_m = \langle m|U|\xi\rangle$. We have

$$U|\psi\rangle|\xi\rangle = \sum_{m} M_m|\psi\rangle|m\rangle. \qquad (9)$$

This means that on an input vector state $|\psi\rangle$ the apparatus **A(x)** outputs the outcome **x** $= m$ with probability $\Pr\{\mathbf{x} = m \| |\psi\rangle\} = \|M_m|\psi\rangle\|^2$, where $\|\cdots\|$ denotes the Euclidean norm given by $\||\phi\rangle\| = \langle\phi|\phi\rangle^{1/2}$, and leaves **S** in the vector state

$$|\psi\rangle_{\{\mathbf{x}=m\}} = \frac{M_m|\psi\rangle}{\|M_m|\psi\rangle\|}. \qquad (10)$$

The *probability operator-valued measure (POVM)* of **A(x)** is the family $\{\Pi_m\}$ of operators defined by $\Pi_m = M_m^{\dagger}M_m$. Then, we have $\Pr\{\mathbf{x} = m \| |\psi\rangle\} = \|\Pi_m^{1/2}|\psi\rangle\|^2 = \langle\psi|\Pi_m|\psi\rangle$. The *nonselective operation* of **A(x)** is a trace-preserving completely positive map T defined by

$$T\rho = \sum_{m} M_m\rho M_m^{\dagger} \qquad (11)$$

for any state ρ; the state $T\rho$ is the state just after the measurement without postselection for the input state ρ.

C. Universal uncertainty principle

Let A, B be observables of **S**. We consider the measurement of the observable A carried out by the apparatus **A(x)** and the disturbance on B caused by this measurement. If the input state of **S** is $|\psi\rangle$, the *root-mean-square (rms) error* $\epsilon(A)$ of **A(x)** for measuring an observable A of **S** and the *rms disturbance* $\eta(B)$ of **A(x)** caused on an observable B of **S** are defined as

$$\epsilon(A) = \|[U^{\dagger}(\mathbb{1} \otimes M)U - A \otimes \mathbb{1}]|\psi\rangle|\xi\rangle\|, \qquad (12)$$

$$\eta(B) = \|[U^{\dagger}(B \otimes \mathbb{1})U - B \otimes \mathbb{1}]|\psi\rangle|\xi\rangle\|, \qquad (13)$$

where $\mathbb{1}$ stands for the identity operator on the respective space. The error $\epsilon(A)$ is the root mean square of the difference between the meter observable M after the interaction and the observable A before the interaction. The disturbance $\eta(B)$ is the root mean square of the change in the observable B

during the measuring interaction. Then, it is mathematically proved [23,24] that

$$\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B) \geqslant \tfrac{1}{2}|\langle\psi|[A,B]|\psi\rangle| \tag{14}$$

holds for any state $|\psi\rangle$ of **S** and any indirect measurement model $(\mathcal{K},|\xi\rangle,U,M)$.

We define the *kth moment output operator* by

$$O_A^{(k)} = \sum_m m^k \Pi_m = \sum_m m^k M_m^\dagger M_m. \tag{15}$$

Then the *k*th moment of the output variable **x** of **A**(**x**) on input $|\psi\rangle$ is given by

$$\mathrm{Ex}\{\mathbf{x}^k \| |\psi\rangle\} = \langle\psi|O_A^{(k)}|\psi\rangle, \tag{16}$$

where "Ex" abbreviates expectation value. We have [24]

$$\epsilon(A)^2 = \langle\psi|O_A^{(2)} - O_A^{(1)}A - AO_A^{(1)} + A^2|\psi\rangle. \tag{17}$$

We define the *postmeasurement kth moment operator* of the observable *B* by

$$O_B^{(k)} = \sum_m M_m^\dagger B^k M_m. \tag{18}$$

Then the *k*th moment of the observable *B* in the state on output from the apparatus **A**(**x**) on input $|\psi\rangle$ is given by

$$\mathrm{Ex}\{B^k \| T(|\psi\rangle\langle\psi|)\} = \langle\psi|O_B^{(k)}|\psi\rangle, \tag{19}$$

where *T* is the nonselective operation of **A**(**x**) [see Eq. (11)]. Then, we have [24]

$$\eta(B)^2 = \langle\psi|O_B^{(2)} - O_B^{(1)}B - BO_B^{(1)} + B^2|\psi\rangle. \tag{20}$$

Therefore, both error $\epsilon(A)$ and disturbance $\eta(B)$ are determined by the measurement operators of the apparatus **A**(**x**).

With the help of $\{M_m\}$ we can rewrite error and disturbance starting from their definitions, Eqs. (12) and (13), to

$$\epsilon(A)^2 = \sum_m \|M_m(m - A)|\psi\rangle\|^2, \tag{21}$$

$$\eta(B)^2 = \sum_m \|[M_m, B]|\psi\rangle\|^2. \tag{22}$$

Details of the calculation can be found in Sec. 4.4 of [25]. If the M_m are mutually orthogonal projection operators, using the Pythagorean theorem, the sum in Eq. (21) can be replaced by

$$\epsilon(A)^2 = \left\|\sum_m M_m(m - A)|\psi\rangle\right\|^2 = \|(O_A - A)|\psi\rangle\|^2, \tag{23}$$

where we used $\sum_m M_m = \mathbb{1}$ and defined the output operator $O_A = \sum_m m M_m$ being just the first moment output operator $O_A^{(1)}$ in cases of projective measurements. In an analogous manner, we will often abbreviate $O_B^{(1)}$ as O_B.

D. Three-state method for quantifying error and disturbance

The error $\epsilon(A)$ and the disturbance $\eta(B)$ have been defined through the noise operator $N(A) = U^\dagger(\mathbb{1} \otimes M)U - A \otimes \mathbb{1}$ and the disturbance operator $D(B) = U^\dagger(B \otimes \mathbb{1})U - B \otimes \mathbb{1}$,

respectively. However, given an apparatus **A**(**x**), those operators are usually unknown in practice. It is also impossible to measure $N(A)$ by measuring $A \otimes \mathbb{1}$ and $U^\dagger(\mathbb{1} \otimes M)U$ successively, since those two observables may not commute. Similarly, it is also impossible to measure $D(B)$ by successive measurement. The *three-state method for quantifying error and disturbance* is a method to measure $\epsilon(A)$ and $\eta(B)$ using the outcomes from the apparatus **A**(**x**) but without knowing either the noise operator $N(A)$ or the disturbance operator $D(B)$.

From Eq. (17) the error $\epsilon(A)$ can be written as [24]

$$\begin{aligned}\epsilon(A)^2 = {}& \langle\psi|A^2|\psi\rangle + \langle\psi|O_A^{(2)}|\psi\rangle + \langle\psi|O_A^{(1)}|\psi\rangle \\ & + \langle A\psi|O_A^{(1)}|A\psi\rangle - \langle(A + \mathbb{1})\psi|O_A^{(1)}|(A + \mathbb{1})\psi\rangle,\end{aligned} \tag{24}$$

where we use such abbreviations as $|A\psi\rangle \equiv A|\psi\rangle$, etc. Thus, $\epsilon(A)$ can be statistically estimated from the measurement outcomes from the apparatus **A**(**x**) in the three states $|\psi\rangle$, $A|\psi\rangle/\|A|\psi\rangle\|$, and $(A + \mathbb{1})|\psi\rangle/\|(A + \mathbb{1})|\psi\rangle\|$ on input. Similarly, from Eq. (20) the disturbance $\eta(B)$ can be written as

$$\begin{aligned}\eta(B)^2 = {}& \langle\psi|B^2|\psi\rangle + \langle\psi|O_B^{(2)}|\psi\rangle + \langle\psi|O_B^{(1)}|\psi\rangle \\ & + \langle B\psi|O_B^{(1)}|B\psi\rangle - \langle(B + \mathbb{1})\psi|O_B^{(1)}|(B + \mathbb{1})\psi\rangle.\end{aligned} \tag{25}$$

Thus, $\eta(B)$ can be statistically estimated from the measurement outcomes of the *B* measurement in the state just after passing through the the apparatus **A**(**x**) in the three states $|\psi\rangle$, $B|\psi\rangle/\|B|\psi\rangle\|$, and $(B + \mathbb{1})|\psi\rangle/\|(B + \mathbb{1})|\psi\rangle\|$ on the input of **A**(**x**). According to the theory of operations [34], for a general observable *X* the state $X|\psi\rangle$ can be prepared from the state $|\psi\rangle$ with success probability $\|X|\psi\rangle\|^2/\|X\|^2$ even without knowing the state $|\psi\rangle$. In fact, an apparatus **A**(**x**) with measurement operators $\{M_0, M_1\}$ with $M_0 = X/\|X\|$ and $M_1 = \sqrt{\mathbb{1} - |M_0|^2}$ produces the state $X|\psi\rangle/\|X|\psi\rangle\|$ on input state $|\psi\rangle$ with probability $\mathrm{Pr}\{\mathbf{x} = 0\||\psi\rangle\} = \|X|\psi\rangle\|^2/\|X\|^2$.

III. THEORY FOR SPIN MEASUREMENTS

We will now apply the general results of the previous section to projective spin-$\frac{1}{2}$ measurements. The observables under consideration are spins along two different directions $\vec{a}$ and $\vec{b}$, that is,

$$A = \vec{a} \cdot \vec{\sigma}, \quad B = \vec{b} \cdot \vec{\sigma}, \tag{26}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ denotes the vector of the Pauli matrices. The apparatus is supposed to carry out a projective spin measurement along a distinct axis $\vec{o}_a$. The output operator $O_A(=O_A^{(1)})$ and the family of measurement operators $\{M_m\}$ are then given explicitly by

$$O_A = \vec{o}_a \cdot \vec{\sigma} = M_{+1} - M_{-1}, \quad M_{\pm 1} = \tfrac{1}{2}(\mathbb{1} \pm \vec{o}_a \cdot \vec{\sigma}). \tag{27}$$

Inserting these expressions into Eqs. (21) and (22) we get for error and disturbance

$$\epsilon(A) = \sqrt{2 - 2\,\vec{a} \cdot \vec{o}_a} = 2\left|\sin\frac{\alpha}{2}\right|, \tag{28}$$

$$\eta(B) = \sqrt{2 - 2(\vec{b} \cdot \vec{o}_a)^2} = \sqrt{2}\,|\sin\beta|, \tag{29}$$
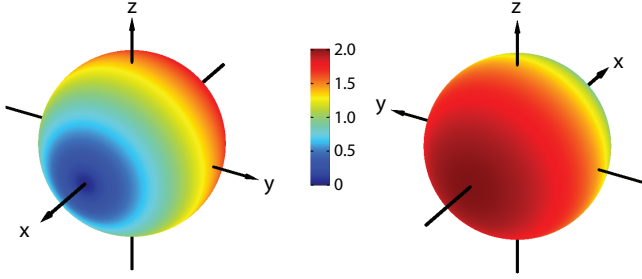
FIG. 1. (Color online) Error $\epsilon(A)$ of the A measurement: The direction $\vec{a}$ of A (A represents the observable to be measured) is fixed to be $(1,0,0)^T$ and the points on the Bloch sphere indicate the direction of O_A (that is the observable actually measured by the apparatus). The value of corresponding error for this combination of vectors $\vec{a}$ and $\vec{o}_a$ is color encoded resulting in concentric circles around the axis defined by $\vec{a}$.

where α denotes the angle between $\vec{a}$ and $\vec{o}_a$, β the angle between $\vec{b}$ and $\vec{o}_a$, and $|\cdots|$ the modulus. Both error and disturbance are thus independent of the initial system state; they are solely determined by the angle between the direction of the observable and the direction of the output operator, that is, the apparatus' measurement axis. The error vanishes if $O_A = A$ and reaches its maximal value ($\epsilon = 2$) if A and O_A point in opposite directions. The whole expression can be illustrated on the Bloch sphere (see Fig. 1).

The disturbance induced on B by the prior measurement of O_A is zero if B and O_A point in the same direction, but also if O_A points exactly in the opposite direction, that is, if $\vec{o}_a = -\vec{b}$. In both cases O_A and B have an identical set of eigenvectors leading to a vanishing disturbance [see Eq. (22)]. That is a notable difference between the physical concepts of error and disturbance. If the two end points of $\vec{b}$ and $-\vec{b}$ define the poles on the Bloch sphere, the disturbance reaches its maximal value ($\eta = \sqrt{2}$) if O_A lies on the corresponding equator (illustrated in Fig. 2 for $B = \sigma_y$).

Since error and disturbance are both independent from the initial system state their product $\epsilon(A)\eta(B)$ also depends only on the relative orientation of $\vec{a}$, $\vec{b}$, and $\vec{o}_a$ and can be depicted on the Bloch sphere for fixed $\vec{a}$ and $\vec{b}$ (see Fig. 3). In order to test if a generalized Heisenberg-like relation [Eq. (4)] is obeyed we also have to evaluate the state-dependent limit. By
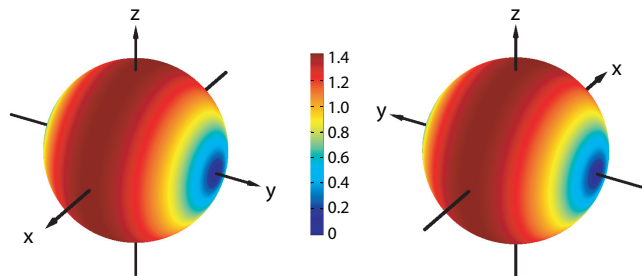

FIG. 2. (Color online) Disturbance $\eta(B)$ on observable $B = \vec{b} \cdot \vec{\sigma} = \sigma_y$ after the projective measurement of O_A. Every point on the Bloch sphere corresponds to a possible direction of O_A. The color of this point on the Bloch sphere indicates the value of the disturbance.
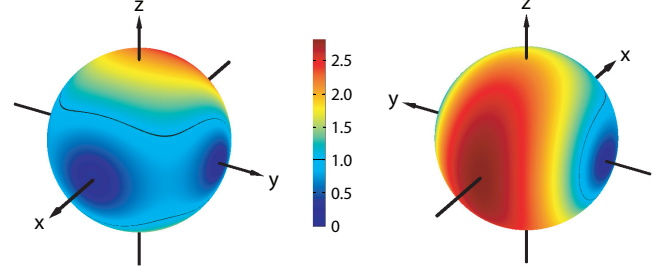

FIG. 3. (Color online) The product of error and disturbance for $A = \sigma_x$ and $B = \sigma_y$. Every point on the Bloch sphere corresponds to a possible direction of O_A and its color encodes the value of $\epsilon(A)\eta(B)$. The black line represents the maximal lower bound given by the commutator $\frac{1}{2}|\langle\psi|[A,B]|\psi\rangle| = 1$ for $|\psi\rangle = |+z\rangle$. Only if the apparatus is so strongly detuned that it measures O_A along a direction outside of the region enclosed by this contour line is Heisenberg's error-disturbance relation, Eq. (4), fulfilled.

assigning the vector $\vec{r}$ to the initial spin state using the Bloch representation of its corresponding density matrix $|\psi\rangle\langle\psi| = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma})$ we get

$$\frac{1}{2}|\langle\psi|[A,B]|\psi\rangle| = |\vec{r} \cdot (\vec{a} \times \vec{b})|. \qquad (30)$$

The limit is thus given by the volume of the parallelepiped spanned by the three vectors $\vec{a}$, $\vec{b}$, and $\vec{r}$ and ranges from 0 to 1. Therefore, except for the special cases where it vanishes, we can always find regions where the product of error and disturbance is below the limit. This is an example for a general result (Chap. 6 in [25]) stating that projective measurements of A violate a Heisenberg-like error-disturbance uncertainty relation provided that B is bounded and $\langle\psi|[A,B]|\psi\rangle \neq 0$.

For a complete investigation of Ozawa's error-disturbance uncertainty relation [Eq. (5)] we additionally need the standard deviations of A and B for an arbitrary initial state $|\psi\rangle$ given by

$$\sigma(A) = \sqrt{1 - (\vec{a} \cdot \vec{r})^2}, \quad \sigma(B) = \sqrt{1 - (\vec{b} \cdot \vec{r})^2}. \qquad (31)$$

For a graphical representation of the left-hand side of the Ozawa relation, Eq. (5), we have to fix the observables A and B and the state $|\psi\rangle$ and can then vary O_A over the Bloch sphere and indicate the resulting values of the sum by different colors. While the Heisenberg term $\epsilon(A)\eta(B)$ is independent from the initial state, the sum in Eq. (5) changes because of the standard deviations and always lies above the lower bound given by the commutator (see Fig. 4).

IV. MEASUREMENT CONCEPT

The expressions (21) and (22) are very convenient to calculate the explicit formulas for error and disturbance in spin measurements [Eqs. (28) and (29)], but they are not suitable for designing an experiment. The output operator O_A and the observable A cannot be measured simultaneously and therefore their difference is no detectable quantity. The same is true for the change of the observable B during the measurement. That is why error and disturbance have been claimed to be experimentally inaccessible [26,27]. But, this obstacle can be overcome by using the expressions (24)
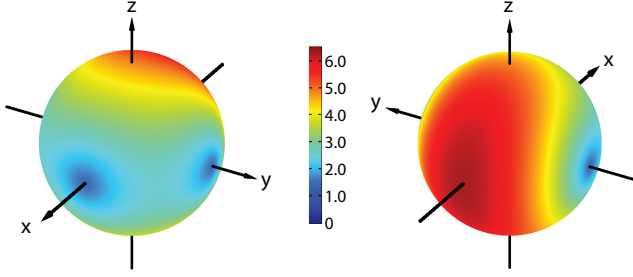
FIG. 4. (Color online) $\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B)$ for $|\psi\rangle = |+z\rangle$, $A = \sigma_x$, and $B = \sigma_y$. Every point on the Bloch sphere corresponds to a possible direction of O_A and its color encodes the value of the sum of the three terms lying everywhere above $\frac{1}{2}|\langle\psi|[A,B]|\psi\rangle| = 1$.

and (25) which read in our case of a projective spin-$\frac{1}{2}$ measurement,

$$\epsilon(A)^2 = 2 + \langle\psi|O_A|\psi\rangle + \langle A\psi|O_A|A\psi\rangle \\ - \langle(A + \mathbb{1})\psi|O_A|(A + \mathbb{1})\psi\rangle \quad (32)$$

and

$$\eta(B)^2 = 2 + \langle\psi|O_B|\psi\rangle + \langle B\psi|O_B|B\psi\rangle \\ - \langle(B + \mathbb{1})\psi|O_B|(B + \mathbb{1})\psi\rangle \quad (33)$$

since $A^2 = O_A^{(2)} = B^2 = O_B^{(2)} = \mathbb{1}$. Now, error and disturbance are solely determined by expectation values of experimentally accessible operators for various input states.

The expression for the error only contains the output operator O_A, which is the observable that is actually measured by the apparatus. The desired observable A occurs solely in two of the necessary three input states $|\psi\rangle, |A\psi\rangle$, and $|(A + \mathbb{1})\psi\rangle$. If we succeed in preparing all these known input states the error $\epsilon(A)$ of the A measurement can be determined from the expectation values of O_A.

For determining the disturbance we need the input states $|\psi\rangle, |B\psi\rangle$, and $|(B + \mathbb{1})\psi\rangle$ and the expectation values of O_B, which is just a shorthand notation for $O_B^{(1)}$. From Eq. (18) it is given by $\sum_m M_m^\dagger B M_m$, where M_m are the projection operators of the output operator O_A. Thus, if we measure B immediately after the measurement of O_A we actually perform the measurement of O_B.

After all, our experiment needs three connected components: a preparation stage for generating the input states $|\psi\rangle$, $|A\psi\rangle$, $|(A + \mathbb{1})\psi\rangle$, $|B\psi\rangle$, and $|(B + \mathbb{1})\psi\rangle$; a measurement apparatus **A1** carrying out the O_A measurement; and a second apparatus **A2** performing the B measurement (see Fig. 5).

As indicated in the measurement scheme (Fig. 5) the two projective measurements of O_A and B result in four possible outcomes since every spin operator can be decomposed into two projectors belonging to the eigenvalues ± 1. Explicitly, these projectors are given by $\frac{1}{2}(\mathbb{1} \pm \vec{o}_a \cdot \vec{\sigma})$ for O_A and by $\frac{1}{2}(\mathbb{1} \pm \vec{b} \cdot \vec{\sigma})$ for B. Equivalently, we can write the projector as ket-bra of the eigenvectors which we denote by $|+\vec{o}_a\rangle$ and $|-\vec{o}_a\rangle$ for the operator O_A and by $|+\vec{b}\rangle$ and $|-\vec{b}\rangle$ for the operator B, respectively. We will use this kind of notation throughout the paper for all spin states. The vector inside the ket indicates the reference axis and the sign denotes the
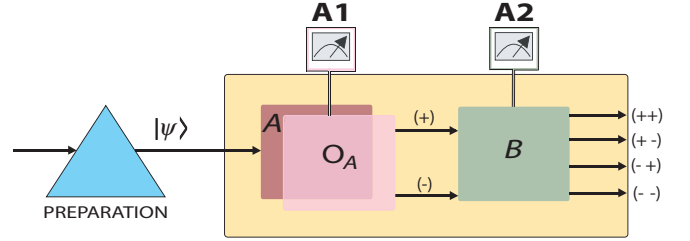

FIG. 5. (Color online) Experimental concept for the verification of the error-disturbance relation: In the first measurement, the apparatus **A1** is detuned in a way that it projectively measures the observable O_A instead of A thus causing an error ϵ in the A measurement. The subsequent measurement of observable B in the eigenstate of O_A, performed by apparatus **A2**, virtually modifies B to be O_B from whose expectation values the disturbance η on B can be determined.

positive or negative spin component. For directions along the Cartesian axes we omit the arrow (for example, $|-x\rangle$ stands for the negative spin component along the x axis).

From the spectral theorem, the operators O_A and B can be decomposed into their projection operators via

$$O_A = \vec{o}_a \cdot \vec{\sigma} = |+\vec{o}_a\rangle\langle+\vec{o}_a| - |-\vec{o}_a\rangle\langle-\vec{o}_a|, \quad (34)$$

$$B = \vec{b} \cdot \vec{\sigma} = |+\vec{b}\rangle\langle+\vec{b}| - |-\vec{b}\rangle\langle-\vec{b}|. \quad (35)$$

The indices $(++)$, $(+-)$, $(-+)$, and $(--)$ of the output ports in Fig. 5 indicate which projections have been carried out. From the intensities at the four possible output ports, denoted by I_{++}, I_{+-}, I_{-+}, and I_{--}, the expectation value of O_A in a state $|\psi\rangle$ is obtained from

$$\langle\psi|O_A|\psi\rangle = \frac{(I_{++} + I_{+-}) - (I_{-+} + I_{--})}{I_{++} + I_{+-} + I_{-+} + I_{--}}. \quad (36)$$

As already mentioned, the prior measurement of O_A modifies the measurement operator of **A2** from B to O_B. The expectation values of O_B necessary for the determination of the disturbance are obtained from

$$\langle\psi|O_B|\psi\rangle = \frac{(I_{++} + I_{-+}) - (I_{+-} + I_{--})}{I_{++} + I_{+-} + I_{-+} + I_{--}}. \quad (37)$$

All expectation values necessary for the determination of error ϵ and disturbance η can thus be derived from the output intensities when the input states $|\psi\rangle$, $|A\psi\rangle$, $|(A + \mathbb{1})\psi\rangle$, $|B\psi\rangle$, and $|(B + \mathbb{1})\psi\rangle$ are applied to the joint measurement apparatuses **A1** and **A2**.

Ozawa's relation [Eq. (5)] additionally contains the standard deviations of A and B in the state $|\psi\rangle$. For our measurement observables, the standard deviations reduce to

$$\sigma(A)^2 = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2 = 1 - \langle\psi|A|\psi\rangle^2 \quad (38)$$

and the corresponding expression for $\sigma(B)$. The expectation value of A can be obtained from our experimental setup by adjusting O_A to be equal to A and then using Eq. (36). For the expectation value of B, we have to turn off the measurement apparatus **A1** and apply $|\psi\rangle$ to **A2** which then projectively measures B (and not O_B). If we then denote the intensities of **A2**'s two output ports by I_+ and I_- the expectation value of B is, according to Eq. (35), given by $(I_+ - I_-)/(I_+ + I_-)$.

Obviously, for the expectation value of *A* we can also turn off **A2** and only use the two output ports of **A1** in the same manner. For the actual experiment we favor this method since increasing the number of involved components usually increases the experimental error.

A final remark concerns the principle aim of our experiment. We do not seek to minimize the measurement error $\epsilon(A)$ as one may assume at first sight. By the way, measuring the spin of a spin-$\frac{1}{2}$ particle along any direction is easily achievable with a high degree of precision. We are interested in a controlled variation of O_A and a systematic investigation of the resulting measurement error $\epsilon(A)$ and disturbance $\eta(B)$, which are given by Eqs. (32) and (33), in order to demonstrate the behavior of Heisenberg's relation [Eq. (4)] in comparison to Ozawa's relation [Eq. (5)].

V. NEUTRON SPIN MEASUREMENT SETUP

In the previous section, we have outlined the idea behind the experiment. We now want to turn to the actually performed measurements on neutron spins. The experiment was carried out at the tangential beam port of the research reactor facility TRIGA Mark II of the Vienna University of Technology. The whole setup is depicted in Fig. 6.

A monochromatic, thermal neutron beam with a mean wavelength of 1,96 Å and a cross section of about 10 (vertical) $\times$ 5 (horizontal) mm^2 propagates in the *y* direction. In the so-called preparation stage, it first crosses a bent Co-Ti supermirror resulting in an approximately 99% polarization in the $+z$ direction. For the further manipulation of the spin state we use magnetic fields which interact with the neutron via the Zeeman Hamiltonian $\mu\vec{\sigma}\vec{B}$. For thermal neutrons the effect of the magnetic fields on the spatial part of the wave function can be neglected [35] and the action of a static magnetic field pointing in direction $\vec{n}_B$ on the spin state is described by a unitary transformation U_R:

$$U_R = e^{i\alpha\vec{n}_B\cdot\vec{\sigma}}, \quad \alpha = \mu|B|T/\hbar, \tag{39}$$

where μ denotes the magnetic moment of the neutron, $|B|$ the modulus of the magnetic field strength, and *T* the time of flight through the field region. The magnetic field thus induces a rotation of the neutron polarization around the axis $\vec{n}_B$ with an angle α, commonly referred to as Larmor precession. In order to obtain any arbitrary spin state we combine the action of a 10 G guide field pointing in the *z* direction and so-called spin flippers. The spin flipper produces a field in the negative *z* direction which compensates the guide field, and a second field B_x that points in the *x* direction. When the *z*-polarized neutrons propagate through the spin flipper their polarization obtains a polar angle depending on the strength of B_x. In the guide field, Larmor precession around the *z* axis is induced. Since the strength of the guide field is fixed, the azimuthal
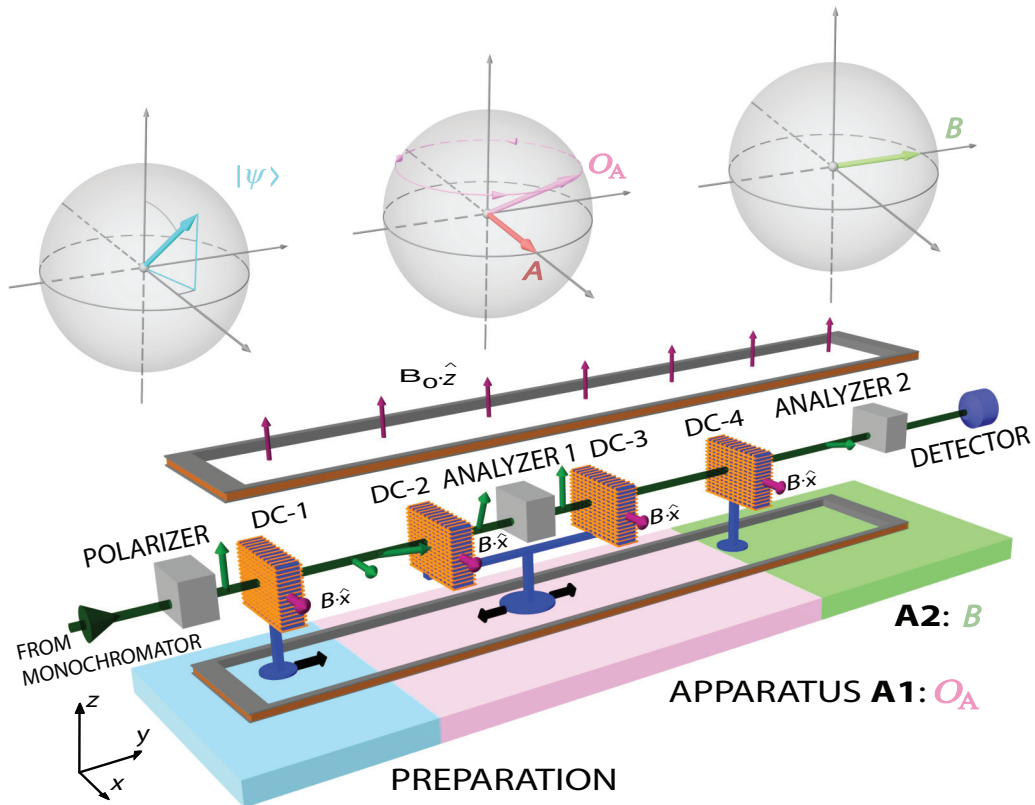

FIG. 6. (Color online) Illustration of the experimental apparatus. The setup is designed for the demonstration of error and disturbance in neutron spin measurements. The neutron optical setup consists of three stages: preparation (blue region), apparatus **A1** measuring O_A (pink region), and apparatus **A2** measuring *B* (green region). A monochromatic neutron beam is polarized in the $+z$ direction by passing through a supermirror spin polarizer. By combining the action of four spin-flipper coils, the magnetic guide field, and the analyzing supermirrors, the successive measurements of O_A and *B* are made for the required input states.

angle of the polarization is varied by changing the time of flight through the guide field. Thus, in order to get a certain azimuthal angle at the end of the preparation stage, we have to position the first spin flipper DC-1 properly between the polarizer and the end point of the preparation stage.

Carrying out the projective measurement of O_A consists of two steps. At first, we have to project the initially prepared state onto the eigenstates of O_A. To complete the measurement we then have to prepare the neutron spin in the eigenstates of O_A. For the projection, spin flipper DC-2 has to be positioned in the guide field such that the spin component along the axis to be measured is rotated into the $+z$ direction. For the eigenstate belonging to eigenvalue $+1$, we need the component along $\vec{o}_a$; for the eigenstate belonging to -1, we rotate the $-\vec{o}_a$ component in the $+z$ direction. The supermirror (analyzer) then selects only the $|+z\rangle$ part of the wave function. The projective measurement is completed by the preparation of the measured spin component with spin flipper DC-3. In an analogous manner to the preparation of the initial state, this is accomplished by properly positioning spin flipper DC-3 in the guide field, so that the desired state is generated at the exit of apparatus **A1**.

Note that the magnetic guide field strength and the dimensions of our experiment are chosen such that any desired direction for the output operator O_A and any initial state can be realized.

On the state after the O_A measurement, the B measurement is performed using the last spin flipper (DC-4) and the analyzer. A subsequent preparation of the eigenstates of B is not necessary since the counting detector is insensitive to the spin state.

In contrast to an ideal measurement, in our setup the four possible outcomes of the two projective spin measurements along $\vec{o}_a$ and $\vec{b}$ are not measured at the same time, but one after the other. At first, we adjust the apparatus to measure the projection onto the positive eigenstates $|+o_a\rangle$ and $|+b\rangle$ of O_A and B. The resulting intensity derived from the counts/600 s is consequently denoted by I_{++}. Afterwards, we change apparatus **A2** to constitute the projection operator $|-b\rangle\langle-b|$ yielding I_{+-}. Then we change apparatus **A1** to measure $|-o_a\rangle\langle-o_a|$ and repeat the procedure for both eigenstates of B to get I_{-+} and I_{--}. From the four output intensities, we obtain the expectation values of O_A and O_B as given by Eqs. (36) and (37). The measured expectation values are normalized taking the limiting efficiency of the entire apparatus ($\sim 96\%$) into account, so that expectation values are bounded between ± 1. In order to get results for error ϵ and disturbance η the expectation values have to be recorded for the different states $|\psi\rangle$, $A|\psi\rangle$, $(A + \mathbb{1})|\psi\rangle$, $B|\psi\rangle$, and $(B + \mathbb{1})|\psi\rangle$. In Fig. 7, we show an explicit example of related intensity sets.

For special configurations of A, B, and $|\psi\rangle$ some of these necessary states are equivalent since overall phase factors are irrelevant. For example, if $A = \sigma_x$, $B = \sigma_y$, and $|\psi\rangle = |+z\rangle$ we have $|A\psi\rangle = |-z\rangle$ and $|B\psi\rangle = i|-z\rangle$.

For the measurement of the standard deviation $\sigma(A)$, we chose O_A to be A and then project the initial state $|\psi\rangle$ onto the eigenstates of A. A subsequent preparation of the eigenstates and the B measurement are not necessary, which means the spin flippers DC-3 and DC-4 can be turned off. For the standard deviation $\sigma(B)$, we virtually remove apparatus **A1** by turning
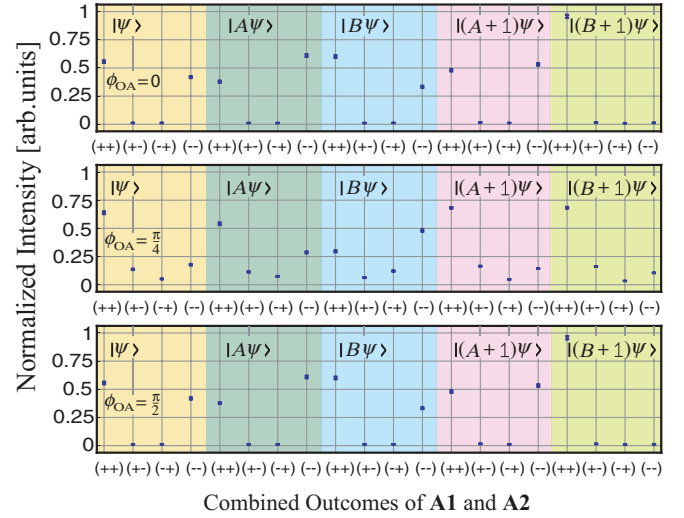

FIG. 7. (Color online) Normalized intensity of the successive measurements carried out by apparatuses **A1** and **A2**. The combined projective measurements of O_A and B have four outcomes, denoted as $(+ +)$, $(+ -)$, $(- +)$, and $(- -)$ and have to be recorded for each initial state, i.e., $|\psi\rangle$, $|A\psi\rangle$, $|B\psi\rangle$, $|(A + \mathbb{1})\psi\rangle$, and $|(B + \mathbb{1})\psi\rangle$. Here $A = \sigma_x$, $B = \sigma_y$ and $|\psi\rangle = |\theta_\psi = \pi/4, \phi_\psi = \pi/12\rangle$. O_A is varied within the xy plane with azimuthal angle given by $\phi_{OA} = 0$, $\pi/4$, and $\pi/2$. Error bars represent $\pm$ one standard deviation of the normalized intensities. Some error bars are at the size of the markers.

off the spin-flipper coils DC-1 and DC-2 and then use DC-3 (in combination with the guide field) to prepare the state $|\psi\rangle$ that is then applied to apparatus **A2** that projectively measures the expectation value of B.

With known error $\epsilon(A)$, disturbance $\eta(B)$, and standard deviations $\sigma(A)$ and $\sigma(B)$ we can examine the behavior of Heisenberg's [Eq. (4)] and Ozawa's [Eq. (5)] relations for varying O_A.

VI. EXPERIMENTAL RESULTS

A. Standard configuration

In the first experiment, we investigate the uncertainty relations in the so-called standard configuration where the Bloch vectors of the observables and the initial state ($\vec{a}$, $\vec{b}$, and $\vec{r}$) are orthogonal to each other:

$$A = \sigma_x, \quad B = \sigma_y, \quad |\psi\rangle = |+z\rangle. \tag{40}$$

Note that since all results only depend on the relative orientation of the involved quantities we can always fix one direction. Therefore, we will chose A to be σ_x in all experiments. For the standard configuration, the standard deviations and the expectation value of commutator between A and B become maximal [see Eqs. (30) and (31)]:

$$\sigma(A) = 1, \quad \sigma(B) = 1, \quad \tfrac{1}{2}\langle\psi|[A,B]|\psi\rangle = 1. \tag{41}$$

In Sec. III, we have depicted $\epsilon(A)$ (Fig. 1), $\eta(B)$ (Fig. 2), their product $\epsilon(A)\eta(B)$ (Fig. 3), and the sum $\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B)$ (Fig. 4) for $A = \sigma_x$ and $B = \sigma_y$ for all possible directions of O_A on the Bloch sphere. In our first experimental run, we vary O_A along the equator. O_A is thus parametrized by its azimuthal angle ϕ_{OA} yielding the direction
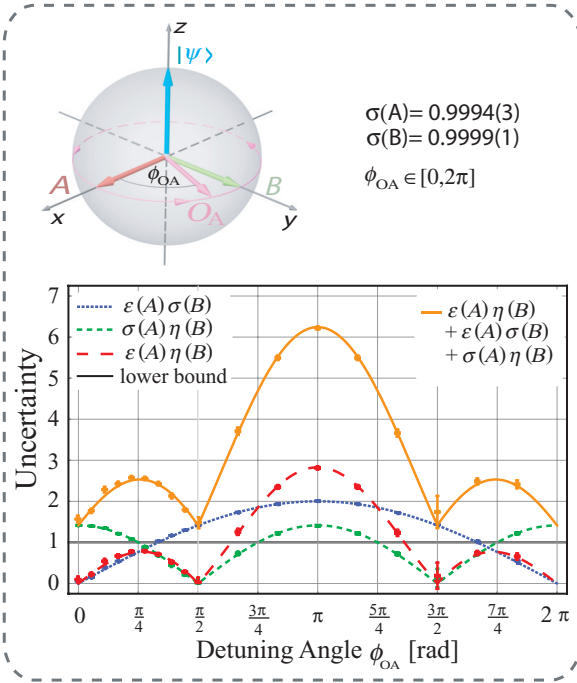
FIG. 8. (Color online) Experimentally determined values of error $\epsilon(A)$, disturbance $\eta(B)$, $\sigma(A)$, and $\sigma(B)$ plotted as $\epsilon(A)\sigma(B)$, $\sigma(A)\eta(B)$, and $\epsilon(A)\eta(B)$ [this term corresponds to the left-hand side of the Heisenberg relation (Eq. (4)] and as the sum $\epsilon(A)\sigma(B) + \sigma(A)\eta(B) + \epsilon(A)\eta(B)$ (corresponds to the left-hand side of the Ozawa's relation [Eq. (5)]) against the azimuthal angle of O_A. The observables A and B and the initial state $|\psi\rangle$ are depicted on the Bloch sphere, as well as the path along which O_A is varied. The respective theory curves are given by Eqs. (41) and (42).

$\vec{o}_a = (\cos\phi_{OA}, \sin\phi_{OA}, 0)$. The theory curves for $\epsilon(A)$ and $\eta(B)$ are then given by

$$\epsilon(A) = 2\sin\frac{\phi_{OA}}{2}, \quad \eta(B) = \sqrt{2}|\cos\phi_{OA}|, \quad (42)$$

which follows from Eqs. (28) and (29).

In the standard configuration, we have to prepare $|+z\rangle$, $|-z\rangle$, $|+x\rangle$, and $|+y\rangle$ as input states in order to obtain the necessary expectation values that determine error, disturbance, and standard deviations. The errors of the measured values for $\epsilon(A)$ and $\eta(B)$ are calculated using error propagation from the standard deviation of the count rates and considering inaccuracies of the Larmor precession angles ($\sim 1.5°$). The latter result mainly from inhomogeneity of the guide field along the beam.

In Fig. 8, we show the experimental outcomes for the three terms occurring in Ozawa's relation plotted against the azimuthal angle of O_A, which we call detuning angle since it also indicates the amount of deviation between A and O_A. The standard deviations are independent of O_A and their measured values are practically equal to $\sigma(A) = \sigma(B) = 1$, which is why we can also investigate the behavior of $\epsilon(A)$ from the $\epsilon(A)\sigma(B)$ curve and likewise the behavior of $\eta(B)$ from the $\sigma(A)\eta(B)$ curve.

Initially, for $\phi_{OA} = 0$, which is $O_A = A$, the measurement error $\epsilon(A)$ vanishes. When we move along the equator it

increases and reaches its maximal value for $O_A = -A(\phi_{OA} = \pi)$. Then, O_A approaches A again and $\epsilon(A)$ decreases.

The disturbance is maximal for $\phi_{OA} = 0$ and vanishes when $O_A = B$ ($\phi_{OA} = \pi/2$). It has a second maximum for $O_A = -A$ because a $(-A)$ measurement disturbs B as much as an A measurement. A and $-A$ have the same set of eigenstates and therefore the same set of projectors, which for the general formula Eq. (22), already leads to an equal expression for the disturbance. We can also verify that the disturbance vanishes again for $O_A = -B$ ($\phi_{OA} = 3\pi/2$), which follows from similar arguments for B and $-B$.

The famous trade-off relation stating that when one observable is measured more precisely, the other is more disturbed has to be treated with care. This reciprocity between error and disturbance only holds for $-\pi/2 \leqslant \phi_{OA} \leqslant \pi/2$. When we move away from $\phi_{OA} = \pi/2$ where $\eta(B) = 0$ to increasing ϕ_{OA} we again disturb the B measurement, but at the same time the error of the A measurement increases as well since we tend towards $-A$. Then, for $\pi \leqslant \phi_{OA} \leqslant 3\pi/2$ both decrease.

The product of error and disturbance $\epsilon(A)\eta(B)$ lies below the limit given by the commutator for the majority of ϕ_{OA} values revealing the violation of the generalized Heisenberg relation [Eq. (4)]. The strongest violation occurs around the regions of error-free ($\phi_{OA} = 0$) or disturbance-free ($\phi_{OA} = \pi/2, 3\pi/2$) measurements.

Contrary to that, the sum $\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B)$ is always above the expectation value of the commutator showing the validity of the Ozawa's relation [Eq. (5)].

B. Arbitrary direction of O_A

Though moving O_A along the equator of the Bloch sphere already reveals a lot of remarkable features about error, disturbance, and the related inequalities a systematic investigation requires arbitrary variations of O_A. The direction of O_A is in general given by an azimuthal angle ϕ_{OA} and a polar angle θ_{OA}, so that $\vec{o}_a = (\cos\phi_{OA}\sin\theta_{OA}, \sin\phi_{OA}\sin\theta_{OA}, \sin\theta_{OA})^T$ which yields for error and disturbance if still $A = \sigma_x$ and $B = \sigma_y$,

$$\epsilon(A) = \sqrt{2 - 2\cos\phi_{OA}\sin\theta_{OA}}, \quad (43)$$

$$\eta(B) = \sqrt{2 - 2\sin^2\phi_{OA}\sin^2\theta_{OA}}. \quad (44)$$

For the following experiments, we will fix the polar angle θ_{OA} and then vary the azimuthal angle ϕ_{OA} thereby realizing an evolution of O_A along circles of latitude on the Bloch sphere (see Fig. 9).

Neither error nor disturbance can then vanish completely since O_A never coincides with A, B, or $-B$ along these paths. The ϵ and η curves are as a whole shrunken from below. Note that we can deduce the behavior of $\epsilon(A)$ and $\eta(B)$ from Fig. 9 since $\sigma(A)$ and $\sigma(B)$ are still 1. The shrinking effect increases the more θ_{OA} differs from $\pi/2$ ending up in straight lines for the degenerate case $\theta_{OA} = 0$.

The trade-off effect between error and disturbance is weakened, but still valid for $-\pi/2 \leqslant \phi_{OA} \leqslant \pi/2$ if we move along the circles of latitude. If we move along meridians no trade-off exists at all since error and disturbance both increase. Thus, at least for spin measurements, the trade-off is not only
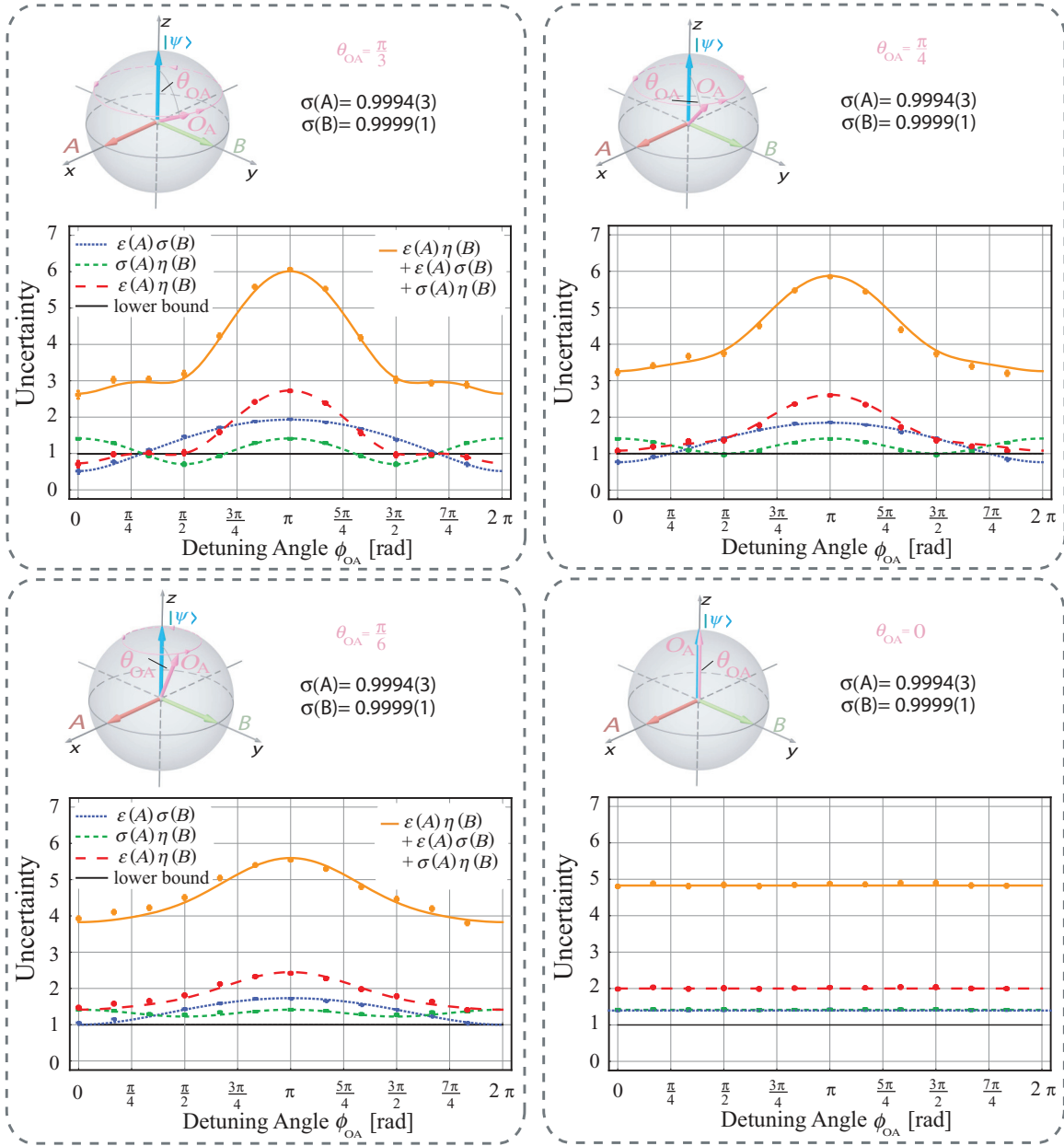
FIG. 9. (Color online) Standard configuration results for O_A outside the xy plane. The observables A and B, the initial state $|\psi\rangle$, and the chosen path for the output operator O_A are depicted on the Bloch sphere. The theoretically predicted values for $\sigma(A)$, $\sigma(B)$, and the lower limit are given by Eq. (41) and the theory curves for $\epsilon(A)$ and $\eta(B)$ by Eqs. (43) and (44), respectively.

restricted to certain regions but also to certain directions in the parameter space.

Since both error and disturbance increase their product increases as well and lies above the limit in larger and larger intervals. For $\theta_{OA} \leqslant \arcsin(0.75) \approx 48.6°$ the Heisenberg relation is always fulfilled (see, for example, top right panel of Fig. 9).

Ozawa's inequality is again always fulfilled; the sum $\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B)$ gets shifted far above the limit when O_A approaches the poles.

C. Varying the azimuthal angle of B

In the next step we quit the standard configuration and change the relative orientation of A and B by moving B along

the equator (see Fig. 10).

$$A = \sigma_x, \quad B = \sigma_x \cos \phi_B + \sigma_y \sin \phi_B, \quad |\psi\rangle = |+z\rangle, \quad (45)$$

which leads to

$$\sigma(A) = 1, \quad \sigma(B) = 1, \quad \tfrac{1}{2}\langle\psi|[A,B]|\psi\rangle = |\sin \phi_B|. \quad (46)$$

If we also vary O_A along the equator the error remains unaltered compared to the standard configuration but we expect the disturbance curve to be shifted an amount $\phi_B - \pi/2$,

$$\epsilon(A) = 2 \sin \frac{\phi_{OA}}{2}, \qquad \eta(B) = \sqrt{2}|\sin(\phi_{OA} - \phi_B)|. \quad (47)$$

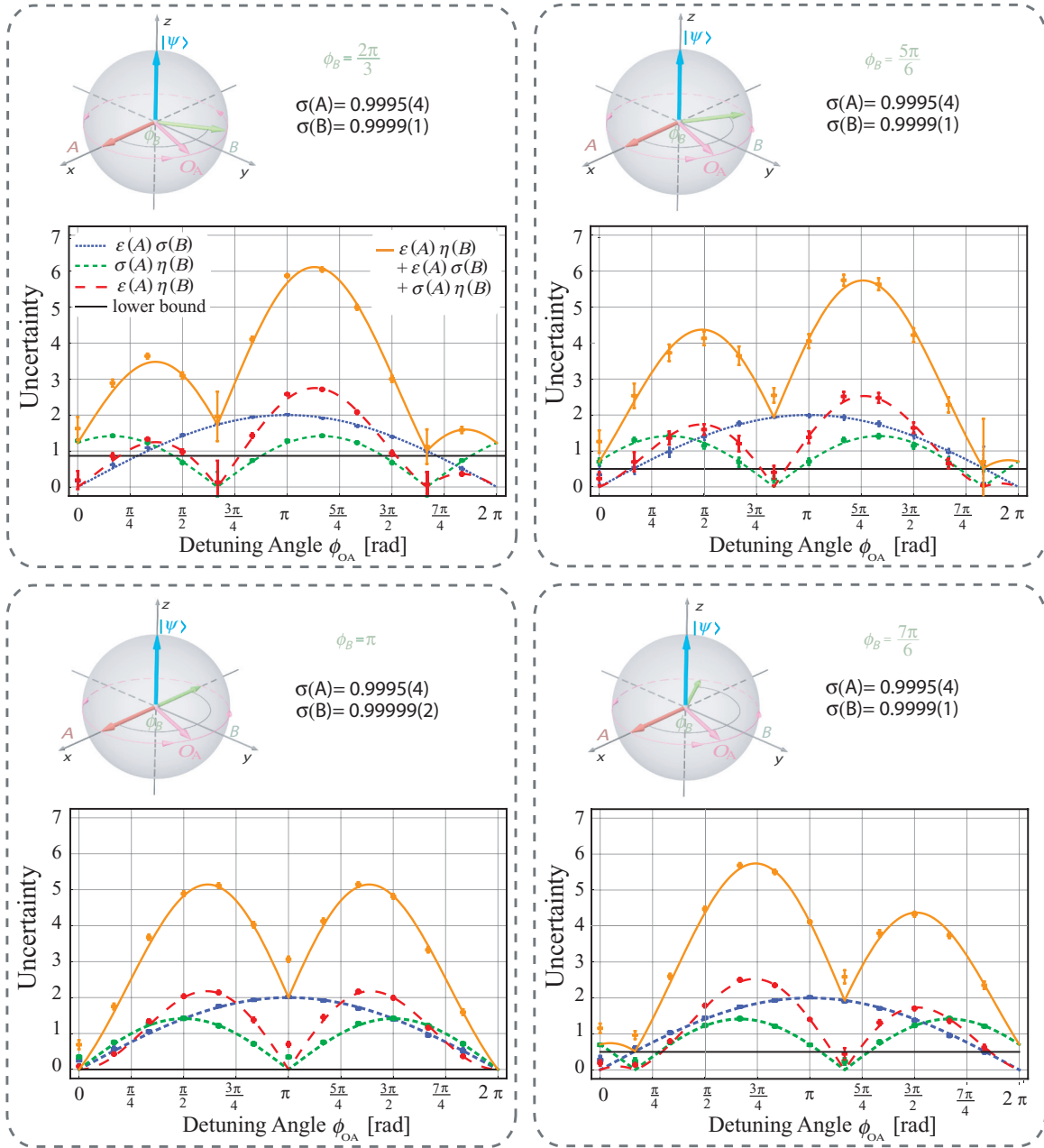Note that in the standard configuration $\phi_B = \pi/2$, which turns the sine dependence into a cosine.

FIG. 10. (Color online) Variations of the azimuthal angle ϕ_B of the observable B. The observables A and B, the initial state $|\psi\rangle$, and the chosen path for the output operator O_A are depicted on the Bloch sphere. The theoretically calculated values for $\sigma(A)$, $\sigma(B)$, and the lower limit are given by Eq. (46) and the theory curves for $\epsilon(A)$ and $\eta(B)$ by Eq. (47).

The error still vanishes for $O_A = A$ and has its maximum at $O_A = -A$; the disturbance vanishes for $O_A = B$ and $O_A = -B$ and has maxima in the middle between the two minima, which do not correspond to the direction of A or $-A$ now.

Due to the loss of symmetry, the product $\epsilon(A)\eta(B)$ changes considerably. Since the commutator between A and B additionally decreases, compared to the standard configuration, the regions where the Heisenberg error-disturbance relation is violated become smaller.

The expression $\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B)$ is always above the limit again demonstrating the validity of Ozawa's relation. However, in the standard configuration the sum never touched the limit, whereas for shifted B it

approaches the lower bound. Points associated with disturbance-free measurement come closest to the limit and finally touch it for the degenerate cases $B = A$ and $B = -A$.

In this context we want to mention an interesting aspect concerning the lower bound. For the premeasurement standard deviations in the state $|\psi\rangle$, the lower bound derived by Robertson [19] and given by Eq. (3) was soon refined by Schrödinger [36] to be

$$\sigma(A)^2\sigma(B)^2 \geqslant \left\langle \frac{1}{2}\{A - \langle A\rangle, B - \langle B\rangle\}\right\rangle^2 + \left\langle \frac{1}{2i}[A,B]\right\rangle^2,$$
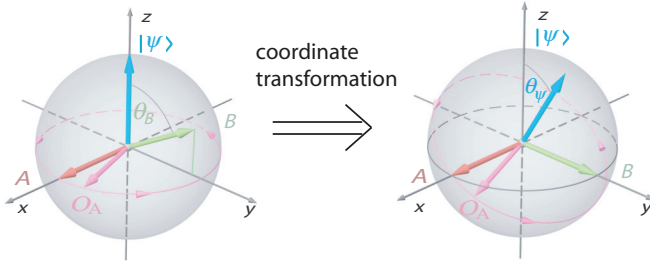
(48)

FIG. 11. (Color online) Since only the relative orientations of A, B, $|\psi\rangle$, and O_A determine the results, the experiments involving the polar angle θ_B can be equivalently seen as configurations with a changed initial state and an evolution of O_A along an inclined plane.

where $\langle \cdots \rangle$ abbreviates $\langle \psi | \cdots | \psi \rangle$ and $\{X,Y\} = XY + YX$ stands for the anticommutator. For our observables $A = \vec{a} \cdot \vec{\sigma}$, $B = \vec{b} \cdot \vec{\sigma}$, and the initial state $|\psi\rangle\langle\psi| = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma})$ the additional term explicitly reads

$$\frac{1}{2}\langle\{A - \langle A\rangle, B - \langle B\rangle\}\rangle = \vec{a} \cdot \vec{b} - (\vec{a} \cdot \vec{r})(\vec{b} \cdot \vec{r}). \quad (49)$$

It vanishes for the standard configuration, but for nonorthogonal A and B it contributes. The measured standard deviations in Fig. 10 of course obey the improved relation, Eq. (48). But for the error-disturbance relations, the anticommutator term cannot be added to the lower bound. This shows explicitly that the three terms of Ozawa's relation are not bounded from below by the product of the premeasurement standard deviations as one could naively conclude from the results of the standard configuration. Meanwhile, improvements to the tightness of the error-disturbance relation have been achieved in [37].

When moving O_A along the equator we find out that the trade-off relation between error and disturbance is valid in

unconnected intervals for ϕ_{OA} whose positions depend on the orientation of B.

D. Varying the polar angle of B

Now, we also introduce a polar angle for B while leaving its azimuthal angle to be $\frac{\pi}{2}$. The observable and the initial state are then given by

$$A = \sigma_x, \quad B = \sigma_y \sin\theta_B + \sigma_z \cos\theta_B, \quad |\psi\rangle = |{+}z\rangle, \quad (50)$$

which leads to

$$\sigma(A) = 1, \quad \sigma(B) = \sin\theta_B, \quad \frac{1}{2}\langle\psi|[A,B]|\psi\rangle = \sin\theta_B. \quad (51)$$

We vary O_A along the equator and along a circle of latitude determined by $\theta_{OA} = \frac{\pi}{3}$. The error $\epsilon(A)$ remains the same as in Eq. (43), whereas the expression for the disturbance written in spherical coordinates becomes rather lengthy and we refer to the general expression, Eq. (29).

Note that we can obviously perform a coordinate transformation so that A and B lie in the xy plane again (see Fig. 11). Then, the state $|\psi\rangle$ gets a polar angle and the path of O_A inclines likewise and consequently its parametrization becomes more elaborate. Nevertheless, one can also see the above experiments as a realization of this particular evolution of O_A for $A = \sigma_x$, $B = \sigma_y$, and a declined initial state.

For the O_A variation along the equator, the error curve does not change compared to the standard configuration; the term $\epsilon(A)\sigma(B)$ in Fig. 12 becomes smaller due to smaller $\sigma(B)$. The disturbance never vanishes totally since O_A never coincides with B or $-B$. The curve is still symmetric because the enclosed angles between O_A and B for $0 \leqslant \phi_{OA} < \pi$ are the same as the angles between O_A and $-B$ for $\pi \leqslant \phi_{OA} < 2\pi$.
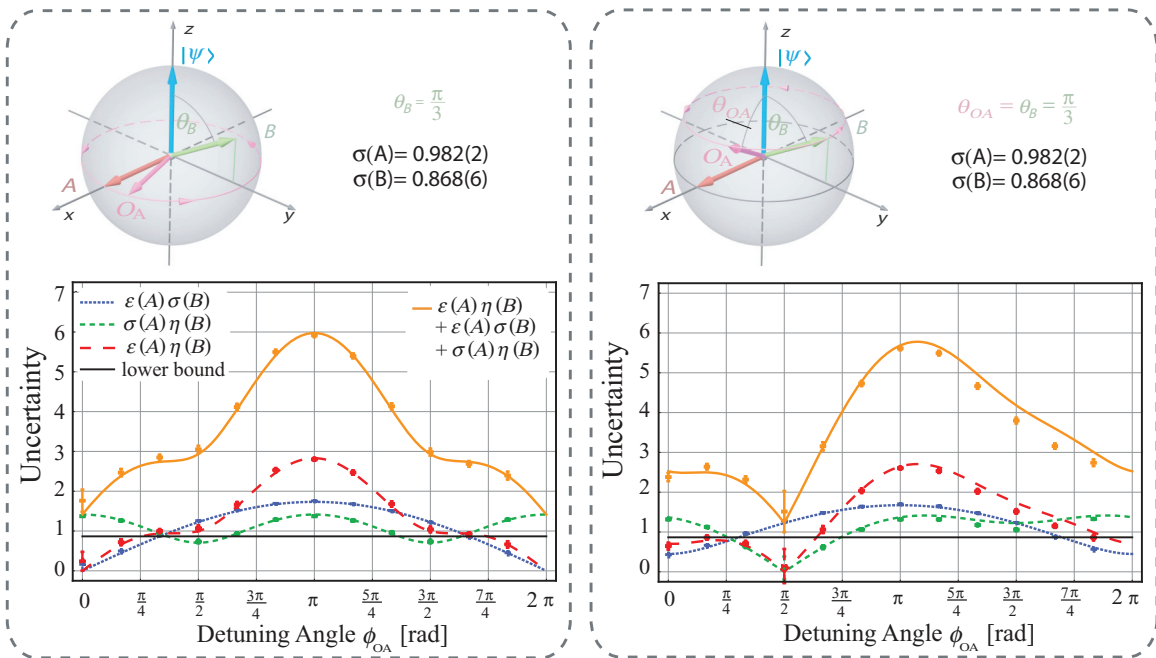

FIG. 12. (Color online) Variations of the polar angle θ_B of the observable B. The observables A and B, the initial state $|\psi\rangle$, and the chosen path for the output operator O_A are depicted on the Bloch sphere. The theoretically predicted values for $\sigma(A)$, $\sigma(B)$, and the lower limit are given by Eq. (51) and the theory curves for $\epsilon(A)$ and $\eta(B)$ by Eqs. (43) and (29), respectively.

When O_A varies along the circle of latitude given by $\theta_{O_A} = \theta_B$ the situation changes. The error cannot become zero since $O_A \neq A$ holds everywhere. The disturbance vanishes for $\phi_{O_A} = \frac{\pi}{2}$ where $O_A = B$ but the symmetry is lost. O_A does not approach $-B$ in the same manner as B and so the disturbance is hardly reduced for $\pi \leqslant \phi_{O_A} < 2\pi$.

E. Varying the initial state

Until now, we have varied O_A along different paths and changed the direction of B. The next thing to investigate is alterations of the initial state $|\psi\rangle$ for an already examined choice of A and B which should not, according to the theoretical predictions, affect error and disturbance. The
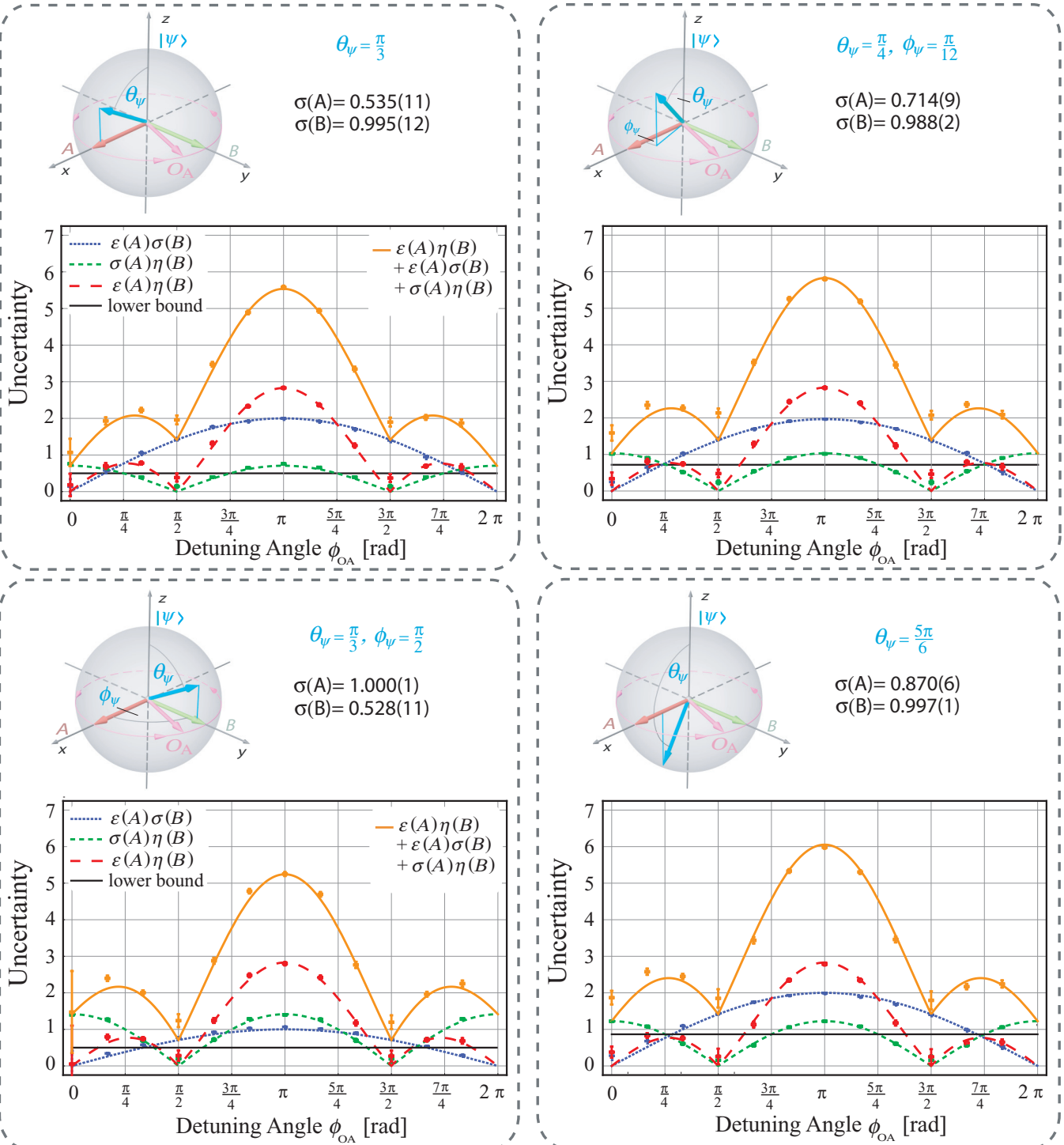

FIG. 13. (Color online) Variations of the initial state $|\psi\rangle$. The observables A and B, the initial state $|\psi\rangle$, and the chosen path for the output operator O_A are depicted on the Bloch sphere. The theoretically predicted values for $\sigma(A)$, $\sigma(B)$, and the lower limit are given by Eqs. (53)–(55) and the theory curves for $\epsilon(A)$ and $\eta(B)$ by Eq. (42).

parameters are given by

$$A = \sigma_x, \quad B = \sigma_y, \quad |\psi\rangle = \cos\frac{\theta_\psi}{2}|+z\rangle + e^{i\phi_\psi}\sin\frac{\theta_\psi}{2}|-z\rangle,$$
$$(52)$$

leading to

$$\sigma(A) = \sqrt{1 - \cos^2\phi_\psi \sin^2\theta_\psi}, \qquad (53)$$

$$\sigma(B) = \sqrt{1 - \sin^2\phi_\psi \sin^2\theta_\psi}, \qquad (54)$$

$$\tfrac{1}{2}\langle\psi|[A,B]|\psi\rangle = |\cos\theta_\psi|. \qquad (55)$$

Since O_A is varied along the equator no change for error and disturbance in comparison to the standard configuration is expected and their theory curves are thus given by Eq. (42). After dividing with the corresponding standard deviations this can be verified from the $\epsilon(A)\sigma(B)$ and $\sigma(A)\eta(B)$ curves in Fig. 13.

The plots show that the product of error and disturbance is indeed independent from the initial state. However, the lower limit changes and the regions where Heisenberg's error-disturbance relation is fulfilled become larger when $|\psi\rangle$ approaches the xy plane.

The left-hand side of Ozawa's relation also contains the standard deviations and therefore changes with the initial state and still lies above the limit.

VII. DISCUSSION

In our previous letter [28], we reported an experiment for the particular case where error and disturbance of a neutron's spin-component measurements obey Ozawa's new relation, Eq. (5), but violate the old one, Eq. (4), in the whole range of the experimental parameter. The old relation is extended from Heisenberg's original relation, Eq. (1), between the error of a position measurement and the induced disturbance on the momentum. Although mainly motivated by thought experiments and in its mathematical derivation based on unjustified assumptions, Heisenberg's old relation for measurement error and disturbance prevailed and was regarded as a peculiarity of quantum mechanics for a long time. The important step for the test of error-disturbance uncertainty relations is to give clear and consistent definitions of error and disturbance for quantum measurements. This is done by Eqs. (12) and (13). Here, we again investigate these relations in the $\frac{1}{2}$-spin system, but enlarge our parameter range in comparison to [28]. For projective spin-$\frac{1}{2}$ measurements, error and disturbance are determined by three (Bloch) vectors $\vec{a}$, $\vec{b}$, and $\vec{o}_a$, characterizing the spin observables A and B and the actual measurement operator O_A. In particular, the error (disturbance) only depends on the deviation angle between $\vec{a}$ ($\vec{b}$) and $\vec{o}_a$, which is described in Eq. (28) [Eq. (29)] and depicted in Fig. 1 (Fig. 2). It is worth noting here that error and disturbance are independent from the initial quantum state which follows straightforward from the properties of the Pauli matrices.

Our measurement strategy is based on the expansion of error and disturbance in terms of expectation values of outcomes in three different input states given by Eqs. (24)

and (25) reminiscent of a quantum process tomography [38,39]. There is another experimental strategy based on weak values [29,32,40]. They insist that "if the system is weakly measured before the measurement apparatus the precision (error) and disturbance can be *directly* observed in the resulting weak values" [32] and claim our method to be indirect. Although it is not clear in what sense the measurements are direct and indirect, we consider here possible arguments: The determination of error and disturbance is done (i) whether by using a single incident state or a combination of some states and/or (ii) whether for single events or for an ensemble. For point (i), we point out the fact that the incident state is affected by some operation before the measurements A and B in both cases. In the weak-value strategy, weak measurements actually demand interaction with the system, which inevitably makes backreaction on the system (however weak they may be), otherwise no information of the system will be derived: After interactions through needed weak-value measurements, the incident state is no longer the same as it was. In contrast, the measurements of A and B are done just after the incident state goes through the operation of $\mathbb{1}$, A, B, $A + \mathbb{1}$, or $B + \mathbb{1}$ in the tomographic strategy. The only difference between them is that the operation is weak or strong. For point (ii), as is stated in one of the above papers, "Experimentally, the problem with weak values is that they cannot be confirmed by precise measurements on individual systems" [40]: Experimental determination of the error and disturbance requires measurements on a set of an ensemble, which is a common feature of both approaches. This is a clear statement of abandonment of determining the error and disturbance for single events by using neither strategy. Another point to be mentioned is the following: In a comparison of the two graphs in Fig. 5 of [28] (respectively, the corresponding part of Fig. 8 in this work) and Fig. 4 of [32], the accuracy of the former is clearly much higher than the latter.

Conscientious readers may presume without difficulty that the new uncertainty relation could be reformulated as

$$[\epsilon(A) + \sigma(A)][\eta(B) + \sigma(B)] \geqslant |\langle\psi|[A,B]|\psi\rangle| \quad (56)$$

by just adding Robertson's and Ozawa's inequalities [Eqs. (3) and (5)]. Although it seems appealing to consider the left-hand side as the product of a newly defined "overall" error and "overall" disturbance, given by the sum of measurement-induced error for A (disturbance on B) and measurement-independent intrinsic fluctuation (standard deviation) of A (B), the "overall" lower bound is less tight than Eqs. (3) and (5) considered separately. Recently, a tight error-disturbance relation (for pure states) has been presented in [37].

Finally, we are aware that a completely different quantification of uncertainty relations, viz., in terms of entropy, has been developed. For instance, a relation for position and momentum was derived first [41], followed by a generalization for arbitrary pairs of observables [42]. An improved version suggested in [5] and proven in [43] has the form

$$H(A)H(B) \geqslant -\log 2 \max_{j,k}\|\langle a_j|b_k\rangle\|, \quad (57)$$

where $H(A)$ [$H(B)$] denotes the Shannon entropy of the probability distribution of the outcome and $|a_j\rangle$ () represent nondegenerating eigenstates of A (B). Recently, a stronger

entropic uncertainty relation for an entangled system was presented [44] and studied in an optical setup [45]. The entropic uncertainty relations describe informational contents and do not refer to the interaction of quantum measurements. They rather represent a generalization of Robertson's relation [Eq. (3)] for probability distributions for which the standard deviation has little physical meaning. Though, it would be interesting if the universally valid uncertainty relation of Ozawa [Eq. (5)] also has an entropic counterpart.

VIII. CONCLUSION

In Heisenberg's original formulation, the uncertainty principle refers to a relation for the error in the measurement of a certain observable and the thereby induced disturbance on another observable. However, his famous relation has been based on thought experiments or a rather heuristic argument with unsupported assumptions. A recent development of rigorous treatments of quantum measurement has enabled us to reformulate the error-disturbance relation and lead to a universally valid relation.

In the experiment presented here, we have investigated this relation for projective neutron spin measurements. The two noncommuting observables under consideration are spins along different directions which are measured successively. By using a tomographic procedure, based on applying three different states generated from the initial system state on the joint measurement apparatuses, the error of the first measurement and the disturbance induced on the second observable can be determined from the output intensities. By

detuning the measurement axis of the first apparatus we can study the variation of error and disturbance. Together with the Bloch vector of the initial system state, the directions of the two spin observables and of the measurement axis constitute the relevant parameters determining all quantities occurring in the uncertainty relations.

In our neutron-optical experimental runs, we have at first fixed the observables and the initial state, and then varied the detuned measurement axis along equilatitude circles on the Bloch sphere. For various representative configurations, the measured values for error and disturbance are in excellent agreement with the theory curves. They are independent of the initial state and are solely determined by the enclosed angle between the measurement axis and the first or second spin axis, respectively. The results demonstrate that Heisenberg's error-disturbance relation can always be violated for a nonvanishing lower limit. In contrast, Ozawa's alternative inequality always holds. Furthermore, we conclude that increasing error does not always lead to decreasing disturbance and vice versa in spin measurements. Such a reciprocal behavior occurs only in certain areas and along certain directions. Thus, our results give an experimental demonstration that the generalized form of Heisenberg's error-disturbance relation has to be abandoned.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927); English translation in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton Univ. Press, Princeton, NJ, 1984), p. 62.

[2] E. Arthurs and J. L. Kelly, Bell Syst. Tech. J. **44**, 725 (1965).

[3] L. E. Ballentine, Rev. Mod. Phys. **42**, 358 (1970).

[4] P. Busch, Int. J. Theor. Phys. **24**, 63 (1985).

[5] K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[6] J. Hilgevoord and J. Uffink, in *Sixty-Two Years of Uncertainty, Historical and Physical Inquiries into the Foundations of Quantum Mechanics*, edited by A. Miller (Plenum, New York, 1990), pp. 121–139.

[7] H. Martens and W. Muynck, Found. Phys. **20**, 357 (1990).

[8] H. Martens and W. M. de Muynck, J. Phys. A: Math. Gen. **25**, 4887 (1992).

[9] D. Appleby, Int. J. Theor. Phys. **37**, 1491 (1998).

[10] M. Ozawa, Phys. Lett. A **299**, 1 (2002).

[11] W. Heisenberg, *The Physical Principles of Quantum Mechanics* (University of Chicago Press, Chicago, IL, 1930).

[12] E. H. Kennard, Z. Phys. A: Hadrons Nucl. **44**, 326 (1927).

[13] C. G. Shull, Phys. Rev. **179**, 752 (1969).

[14] H. Kaiser, S. A. Werner, and E. A. George, Phys. Rev. Lett. **50**, 560 (1983).

[15] A. G. Klein, G. I. Opat, and W. A. Hamilton, Phys. Rev. Lett. **50**, 563 (1983).

[16] O. Nairz, M. Arndt, and A. Zeilinger, Phys. Rev. A **65**, 032109 (2002).

[17] M. Ozawa, Phys. Lett. A **318**, 21 (2003).

[18] H. Wiseman, Found. Phys. **28**, 1619 (1998).

[19] H. P. Robertson, Phys. Rev. **34**, 163 (1929).

[20] E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).

[21] S. Ishikawa, Rep. Math. Phys. **29**, 257 (1991).

[22] M. Ozawa, in *Quantum Aspects of Optical Communications*, edited by C. Bendjaballah, O. Hirota, and S. Reynaud, Lecture Notes in Physics Vol. 378 (Springer, Berlin, 1991), pp. 1–17.

[23] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).

[24] M. Ozawa, Ann. Phys. **311**, 350 (2004).

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

[26] R. F. Werner, Quantum Inf. Comput. **4**, 546 (2004).

[27] K. Koshino and A. Shimizu, Phy. Rep. **412**, 191 (2005).

[28] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Nat. Phys. **8**, 185 (2012).

[29] A. P. Lund and H. M. Wiseman, New J. Phys. **12**, 093011 (2010).

[30] R. Mir, J. S. Lundeen, M. W. Mitchell, A. M. Steinberg, J. L. Garretson, and H. M. Wiseman, New J. Phys. **9**, 287 (2007).

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

[32] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Phys. Rev. Lett. **109**, 100404 (2012).

[33] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[34] *States, Effects, and Operations Fundamental Notions of Quantum Theory*, edited by K. Kraus, A. Böhm, J. D. Dollard, and W. H. Wootters, Lecture Notes in Physics Vol. 190 (Springer-Verlag, Berlin, 1983).

[35] J. Klepp, S. Sponar, S. Filipp, M. Lettner, G. Badurek, and Y. Hasegawa, Phys. Rev. Lett. **101**, 150404 (2008).

[36] E. Schrödinger, Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse **14**, 296 (1930) [Bulg. J. Phys. **26**, 193 (1999)].

[37] C. Branciard, Proc. Natl. Acad. Sci. USA **110**, 6742 (2013).

[38] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[39] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).

[40] H. F. Hofmann, arXiv:1205.0073.

[41] I. Bialynicki-Birula and J. Mycielski, Commun. Math. Phys. **44**, 129 (1975).

[42] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[43] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).

[44] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nat. Phys. **6**, 659 (2010).

[45] C. F. Li, J. S. Xu, X. Y. Xu, K. Li, and G. C. Guo, Nat. Phys. **7**, 752 (2011).