Quantum Stud.: Math. Found. (2023) 10:373–397
https://doi.org/10.1007/s40509-023-00301-8

CHAPMAN UNIVERSITY | INSTITUTE FOR QUANTUM STUDIES

REGULAR PAPER

Uncertainty from the Aharonov–Vaidman identity

Matthew S. Leifer

Received: 25 January 2023 / Accepted: 2 March 2023 / Published online: 24 April 2023
© The Author(s) 2023

Abstract In this article, I show how the Aharonov–Vaidman identity $A|\psi\rangle = \langle A\rangle|\psi\rangle + \Delta A\left|\psi_A^\perp\right\rangle$ can be used to prove relations between the standard deviations of observables in quantum mechanics. In particular, I review how it leads to a more direct and less abstract proof of the Robertson uncertainty relation $\Delta A \Delta B \geq \frac{1}{2}|\langle [A, B]\rangle|$ than the textbook proof. I discuss the relationship between these two proofs and show how the Cauchy–Schwarz inequality can be derived from the Aharonov–Vaidman identity. I give Aharonov–Vaidman based proofs of the Maccone–Pati uncertainty relations and show how the Aharonov–Vaidman identity can be used to handle propagation of uncertainty in quantum mechanics. Finally, I show how the Aharonov–Vaidman identity can be extended to mixed states and discuss how to generalize the results to the mixed case.

Keywords Quantum mechanics · Uncertainty relations · Error propagation · Aharonov–Vaidman identity

1 Introduction

Let A be a Hermitian operator on a Hilbert space $\mathcal{H}$. Then, for any (not necessarily normalized) vector $|\psi\rangle \in \mathcal{H}$,

$$A|\psi\rangle = \langle A\rangle|\psi\rangle + \Delta A\left|\psi_A^\perp\right\rangle, \tag{1}$$

where $\langle A\rangle = \langle\psi|A|\psi\rangle / \langle\psi|\psi\rangle$ is the expectation value of A, $\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}$ is its standard deviation, and $\left|\psi_A^\perp\right\rangle$ is a vector that is orthogonal to $|\psi\rangle$, has equal norm $\left\langle\psi_A^\perp\middle|\psi_A^\perp\right\rangle = \langle\psi|\psi\rangle$, and depends on the operator A.

M. S. Leifer (✉)
Institute for Quantum Studies, Schmid College of Science and Technology, Chapman University, One University Drive, Orange, CA 92866, USA
e-mail: leifer@chapman.edu

Equation (1) is the *Aharonov–Vaidman identity*, which first appeared in [1]. Yakir Aharonov has stated that he "[does not] understand why it doesn't appear in every quantum book" [2]. The main purpose of this article is to explain why it should appear in undergraduate quantum mechanics textbooks.[1]

The uncertainty relation that is proved most often in quantum mechanics classes and textbooks is the Robertson relation [4]:

$$\Delta A \Delta B \geq \frac{1}{2} \left| \langle [A, B] \rangle \right|, \tag{2}$$

where $[A, B] = AB - BA$ is the commutator.

As pointed out by Schrödinger [5], the Robertson relation can be extended to

$$(\Delta A)^2 (\Delta B)^2 \geq \left| \frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle \right|^2 + \left| \frac{1}{2} \langle [A, B] \rangle \right|^2, \tag{3}$$

where $\{A, B\} = AB + BA$ is the anti-commutator.

Although not often emphasized in quantum mechanics classes, the Schrödinger relation is not harder to prove than the Robertson relation. In fact, the standard textbook proof of the Robertson relation effectively proves the Schrödinger relation and then throws away the anti-commutator term.

The proof almost universally adopted in textbooks is based on the Cauchy–Schwarz inequality. While this proof is elementary for those familiar with the mathematics of Hilbert spaces, it can be daunting for undergraduate physics students, who are likely encountering Hilbert spaces for the first time along with quantum mechanics.

In this article, I will review more direct proofs of (2) and (3) from the Aharonov–Vaidman identity that only make use of basic properties of complex numbers and inner products. These proofs previously appeared in [6] and the proof of the Robertson relation is also problem 3.10 in Aharonov and Rohrlich's book "Quantum Paradoxes" [7]. The proof of the Aharonov–Vaidman identity itself uses similar ideas to one of the standard proofs of the Cauchy–Schwarz identity, but is perhaps more memorable to undergraduate physics students because it uses concepts that have a physical meaning, i.e. expectation values and standard deviations. The proof of the Robertson and Schrödinger relations so obtained is not independent of the standard Cauchy–Schwarz based proof. I shall discuss their relationship and show that the Cauchy–Schwarz inequality can itself be derived from (1). The main virtue of using the Aharonov–Vaidman based proof of the uncertainty relation is that it is more direct and involves fewer abstractions.

To be clear, I am not against using or teaching the Cauchy–Schwarz inequality. It has been called "one of the most widely used and important inequalities in all of mathematics" [8]. In fact, the Aharonov–Vaidman based proof still uses one instance of the Cauchy–Schwarz inequality, namely that if $|\psi\rangle$ and $|\phi\rangle$ are unit vectors then $\left|\langle\phi|\psi\rangle\right| \leq 1$. But this is easily motivated by the idea that $\langle\phi|\psi\rangle$ is a generalization of the cosine of an angle, and it is used in a more direct way than in the standard proof. Students of quantum mechanics also need to know the Cauchy–Schwarz inequality to prove that the Born rule always yields well-defined probabilities. Physics students should learn the Cauchy–Schwarz inequality. I just think it should be used in a less abstract way where possible.

Besides the Robertson and Schrödinger relations, many other uncertainty relations are known. Indeed, since uncertainty relations have found applications in quantum information science [9–15] and quantum foundations [16,17], proving new ones has become something of a sport. The two most common classes of uncertainty relations are those based on entropy [18] and those based on standard deviations [4,5,19]. Many of the standard deviation based relations can be derived from the Aharonov–Vaidman relation. I include a proof of the Maccone-Pati uncertainty relations [20] to illustrate this. While these are not the most recent or tightest known uncertainty relations, I include

[1] Other demonstrations of the usefulness of the Aharonov–Vaidman identity include its use in the proof that, for any state $|\psi\rangle$ and any observable A, $|\psi\rangle^{\otimes n}$ is an approximate eigenstate of the observable $\bar{A} = \frac{1}{n}\sum_{j=1}^{n} A_j$ for large n, where A_j refers to A acting on the jth subsystem [1], and its use in deriving the minimum time required for evolution to an orthogonal quantum state [3].

them because they have a simple and elegant Aharonov–Vaidman based proof. For more recent work on standard deviation uncertainty relations, see [21–41].

Another place where relationships between standard deviations are important is in the propagation of uncertainty. In classical statistics, if random variables $X_1, X_2, \ldots, X_n$ have standard deviations $\Delta X_1, \Delta X_2, \ldots \Delta X_n$, then a function of them $f(X_1, X_2, \ldots, X_n)$ has standard deviation Δf that is a function of $\Delta X_1, \Delta X_2, \ldots \Delta X_n$ (and their correlations if the variables are not independent). Formulas for the propagation of uncertainty tell us how to compute this function and are commonly used to estimate experimental errors. In quantum mechanics, similar formulas can be derived relating the standard deviations of observables. They differ from their classical counterparts due to the fact that quantum observables do not commute, but provided this is taken care of they can be derived by the same methods as in the classical case. However, they can alternatively be derived from the Aharonov-Vaidman identity, as I shall explain.

Although the Aharonov–Vaidman identity is usually discussed for pure quantum states, it can be extended to mixed states, either by use of purification or an equivalent concept called an *amplitude operator*. Relations between standard deviations can be extended to mixed states, but obtaining tight bounds is sometimes more difficult than in the pure case due to the need to optimize over all purifications or amplitude operators that can represent a given mixed state.

The remainder of this article is structured as follows. Section 2 gives the proof of the Aharonov–Vaidman identity and a corollary that is useful for understanding the equality conditions in uncertainty relations. Section 3 presents the proof of the Robertson and Schrödinger relations based on the Aharonov–Vaidman identity. Section 4 explains the relationship with the standard textbook proof of the Robertson relation and explains how the Cauchy–Schwarz inequality can be derived from the Aharonov–Vaidman identity. Section 5 comments on the effective teaching of the Robertson uncertainty relations via the Aharonov–Vaidman identity. Section 6 presents Aharonov–Vaidman based proofs of the Maccone–Pati uncertainty relations. Section 7 describes how to use the Aharonov–Vaidman identity to derive formulas for the propagation of quantum uncertainty. Section 8 explains how to generalize the Aharonov–Vaidman relation to mixed states using amplitude operators. (The relationship between amplitude operators and purifications is discussed in Appendix A.) Finally, Sect. 9 presents the summary and conclusions.

I intend this article to be pedagogical and self-contained, so as to be accessible to undergraduate students and anyone teaching introductory quantum mechanics.

2 Proof of the Aharonov Vaidman identity

Sometimes, it is useful to generalize the Aharonov–Vaidman identity to non-Hermitian operators, so we prove the more general version here.

Proposition 2.1 (The Aharonov–Vaidman Identity) *Let A be a linear operator on a Hilbert space $\mathcal{H}$ and let $|\psi\rangle$ be a (not necessarily normalized) vector in $\mathcal{H}$. Then,*

$$A|\psi\rangle = \langle A\rangle |\psi\rangle + \Delta A \left|\psi_A^{\perp}\right\rangle, \tag{4}$$

where $\langle A\rangle = \langle\psi|A|\psi\rangle / \langle\psi|\psi\rangle$, $\Delta A = \sqrt{\langle A^\dagger A\rangle - |\langle A\rangle|^2}$, and $\left|\psi_A^{\perp}\right\rangle$ is a vector orthogonal to $|\psi\rangle$ that depends on both $|\psi\rangle$ and A and satisfies $\langle\psi_A^{\perp}|\psi_A^{\perp}\rangle = \langle\psi|\psi\rangle$.

Note that, if A is Hermitian, then this reduces to (1), where $\langle A\rangle$ and ΔA are the expectation value and standard deviation. In general, $\langle A\rangle$ is a complex number, but ΔA is always real and non-negative.

For most of what we need to do, it is sufficient to consider the case where $|\psi\rangle$ is a unit vector, in which case $\left|\psi_A^{\perp}\right\rangle$ is also a unit vector. The exception is the proof of the Cauchy–Schwarz inequality (Proposition 4.1 in Sect. 4), which uses the identity with an unnormalized vector.

Proof Given a vector $|\psi\rangle \in \mathcal{H}$, any other vector $|\phi\rangle \in \mathcal{H}$ can be written as $|\phi\rangle = \alpha |\psi\rangle + \beta \left|\psi^\perp\right\rangle$, where α and β are complex coefficients and $\left|\psi^\perp\right\rangle$ is some vector that is orthogonal to $|\psi\rangle$. By an appropriate rescaling of β, we can ensure that $\left\langle\psi^\perp\middle|\psi^\perp\right\rangle = \left\langle\psi\middle|\psi\right\rangle$. Applying this to $|\phi\rangle = A|\psi\rangle$ gives

$$A|\psi\rangle = \alpha|\psi\rangle + \beta\left|\psi^\perp\right\rangle. \tag{5}$$

To determine α, take the inner product of (5) with $|\psi\rangle$, which gives

$$\left\langle\psi\middle|A\middle|\psi\right\rangle = \alpha\left\langle\psi\middle|\psi\right\rangle. \tag{6}$$

Rearranging this gives $\alpha = \langle A\rangle$.

To determine β, substitute $\alpha = \langle A\rangle$ into (5) and take the inner product of $A|\psi\rangle$ with itself to obtain

$$\begin{aligned}\langle\psi|A^\dagger A|\psi\rangle &= |\langle A\rangle|^2\left\langle\psi\middle|\psi\right\rangle + |\beta|^2\left\langle\psi^\perp\middle|\psi^\perp\right\rangle \\ &= |\langle A\rangle|^2\left\langle\psi\middle|\psi\right\rangle + |\beta|^2\left\langle\psi\middle|\psi\right\rangle,\end{aligned}$$

where we have used $\left\langle\psi^\perp\middle|\psi^\perp\right\rangle = \left\langle\psi\middle|\psi\right\rangle$.

Rearranging and using $\left\langle A^\dagger A\right\rangle = \left\langle\psi\middle|A^\dagger A\middle|\psi\right\rangle / \left\langle\psi\middle|\psi\right\rangle$ gives

$$|\beta|^2 = \left\langle A^\dagger A\right\rangle - |\langle A\rangle|^2 = (\Delta A)^2. \tag{7}$$

This means that $\beta = (\Delta A)e^{i\theta}$ for some phase angle θ. If we define $\left|\psi_A^\perp\right\rangle = e^{i\theta}\left|\psi^\perp\right\rangle$, then $\left|\psi_A^\perp\right\rangle$ is still orthogonal to $|\psi\rangle$, its norm is unchanged, and we have (4). □

The following corollary is useful for finding the conditions for equality in uncertainty relations.

Corollary 2.2 *In general, for two operators A and B, and for a unit vector* $|\psi\rangle$*,*

$$\left\langle\psi_A^\perp\middle|\psi_B^\perp\right\rangle = \frac{\left\langle A^\dagger B\right\rangle - \langle A\rangle^*\langle B\rangle}{\Delta A\Delta B}. \tag{8}$$

Proof From Proposition 2.1, we have

$$A|\psi\rangle = \langle A\rangle|\psi\rangle + \Delta A\left|\psi_A^\perp\right\rangle, \tag{9}$$

$$B|\psi\rangle = \langle B\rangle|\psi\rangle + \Delta B\left|\psi_B^\perp\right\rangle. \tag{10}$$

Taking the inner product of these gives

$$\langle\psi|A^\dagger B|\psi\rangle = \langle A\rangle^*\langle B\rangle + \Delta A\Delta B\left\langle\psi_A^\perp\middle|\psi_B^\perp\right\rangle. \tag{11}$$

Rearranging gives the desired result. □

Note that if A and B are Hermitian, then we have

$$\left\langle\psi_A^\perp\middle|\psi_B^\perp\right\rangle = \frac{\langle AB\rangle - \langle A\rangle\langle B\rangle}{\Delta A\Delta B}. \tag{12}$$

If it is also the case that $[A, B] = 0$, then (12) is the *correlation*, denoted $\mathrm{corr}_{A,B}$, that would be obtained from a joint measurement of A and B. The correlation is a well-known statistical measure of how two random variables are related to one another. (12) is a formal generalization of the correlation, so we will also denote it $\mathrm{corr}_{A,B}$. However, if A and B do not commute, then $\mathrm{corr}_{A,B}$ is generally a complex number, there is no joint measurement of A and B of which $\mathrm{corr}_{A,B}$ could be the correlation, and AB is not an observable.

The real and imaginary parts of $\mathrm{corr}_{A,B}$ are

$$\mathrm{Re}\left(\mathrm{corr}_{A,B}\right) = \frac{1}{2}\left(\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle + \left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle\right) = \frac{\frac{1}{2}\langle\{A, B\}\rangle - \langle A\rangle\langle B\rangle}{\Delta A \Delta B}, \tag{13}$$

$$\mathrm{Im}\left(\mathrm{corr}_{A,B}\right) = \frac{1}{2i}\left(\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle - \left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle\right) = \frac{\langle[A, B]\rangle}{2i\Delta A \Delta B}. \tag{14}$$

The real part is also a formal generalization of the correlation, in that it reduces to the classical formula when A and B commute. We denote it $\mathrm{Rcorr}_{A,B}$.

3 The Robertson and Schrödinger uncertainty relations

We are now in a position to prove the Robertson and Schrödinger uncertainty relations.

Proposition 3.1 (The Robertson uncertainty relation) *Let A and B be two Hermitian operators on a Hilbert space $\mathcal{H}$. Then, for any unit vector $|\psi\rangle \in \mathcal{H}$*

$$\Delta A \Delta B \geq \frac{1}{2}\left|\langle[A, B]\rangle\right|. \tag{15}$$

Proof From the Aharonov–Vaidman identity, we have

$$A|\psi\rangle = \langle A\rangle|\psi\rangle + \Delta A\left|\psi_A^{\perp}\right\rangle, \tag{16}$$

$$B|\psi\rangle = \langle B\rangle|\psi\rangle + \Delta B\left|\psi_B^{\perp}\right\rangle. \tag{17}$$

Taking the inner product of these two equations and its complex conjugate gives

$$\left\langle\psi\middle|AB\middle|\psi\right\rangle = \langle A\rangle\langle B\rangle + \Delta A\Delta B\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle, \tag{18}$$

$$\left\langle\psi\middle|BA\middle|\psi\right\rangle = \langle A\rangle\langle B\rangle + \Delta A\Delta B\left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle. \tag{19}$$

Subtracting these two equations gives

$$\left\langle\psi\middle|(AB - BA)\middle|\psi\right\rangle = \Delta A\Delta B\left(\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle - \left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle\right), \tag{20}$$

or

$$\langle[A, B]\rangle = \Delta A\Delta B\left(\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle - \left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle\right). \tag{21}$$

Since $\left\langle\psi_B^{\perp}\middle|\psi_A^{\perp}\right\rangle$ is the complex conjugate of $\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle$, we can rewrite this as

$$\langle[A, B]\rangle = 2i\Delta A\Delta B\,\mathrm{Im}\left(\left\langle\psi_A^{\perp}\middle|\psi_B^{\perp}\right\rangle\right). \tag{22}$$

Taking the absolute value of both sides and rearranging it gives

$$\Delta A \Delta B \left|\mathrm{Im}\left(\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle\right)\right| = \frac{1}{2} \left|\langle [A, B] \rangle\right| . \tag{23}$$

Because $\left|\psi_A^{\perp}\right\rangle$ and $\left|\psi_B^{\perp}\right\rangle$ are unit vectors, $0 \leq \left|\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle\right|^2 \leq 1$, and hence the absolute value of the imaginary part of $\left\langle \psi_B^{\perp} \middle| \psi_A^{\perp} \right\rangle$ is also bounded between 0 and 1. Hence, we have

$$\Delta A \Delta B \geq \frac{1}{2} \left|\langle [A, B] \rangle\right| . \tag{24}$$

□

The condition for equality in the Robertson relation is $\left|\mathrm{Im}\left(\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle\right)\right| = 1$ or, equivalently, $\mathrm{corr}_{A,B} = \pm i$. States that saturate the inequality are called *(Robertson) intelligent states*. The condition $\mathrm{corr}_{A,B} = \pm i$ can be used to find intelligent states, although this is not easier than solving for equality in the Robertson relation directly.

Proposition 3.2 (The Schrödinger uncertainty relation) *Let A and B be two Hermitian operators on a Hilbert space $\mathcal{H}$. Then, for any unit vector $|\psi\rangle \in \mathcal{H}$,*

$$(\Delta A)^2 (\Delta B)^2 \geq \left|\frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle\right|^2 + \left|\frac{1}{2} \langle [A, B] \rangle\right|^2 . \tag{25}$$

Proof Taking the sum of (18) and (19) gives

$$\langle \{A, B\} \rangle = 2 \langle A \rangle \langle B \rangle + \Delta A \Delta B \left(\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle + \left\langle \psi_B^{\perp} \middle| \psi_A^{\perp} \right\rangle\right), \tag{26}$$

or

$$\langle \{A, B\} \rangle - 2 \langle A \rangle \langle B \rangle = \Delta A \Delta B \left(\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle + \left\langle \psi_B^{\perp} \middle| \psi_A^{\perp} \right\rangle\right). \tag{27}$$

Adding this to (21) gives

$$\langle \{A, B\} \rangle - 2 \langle A \rangle \langle B \rangle + \langle [A, B] \rangle = 2 \Delta A \Delta B \left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle, \tag{28}$$

or

$$\Delta A \Delta B \left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle = \frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle + \frac{1}{2} \langle [A, B] \rangle . \tag{29}$$

Now, because A and B are Hermitian, $\{A, B\}$ is Hermitian and $[A, B]$ is anti-Hermitian. Therefore, $\langle \{A, B\} \rangle$ is real and $\langle [A, B] \rangle$ is imaginary. Further, $\langle A \rangle$, $\langle B \rangle$, ΔA and ΔB are real. Therefore, taking the modulus squared of (29) gives

$$(\Delta A)^2 (\Delta B)^2 \left|\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle\right|^2 = \left|\frac{1}{2} \langle \{A, B\} \rangle - \langle A \rangle \langle B \rangle\right|^2 + \left|\frac{1}{2} \langle [A, B] \rangle\right|^2 . \tag{30}$$

Finally, because $\left|\psi_A^{\perp}\right\rangle$ and $\left|\psi_B^{\perp}\right\rangle$ are unit vectors, we have $0 \leq \left|\left\langle \psi_A^{\perp} \middle| \psi_B^{\perp} \right\rangle\right|^2 \leq 1$, from which the result follows. □

The condition for equality in the Schrödinger relation is $\left|\langle\psi_A^\perp|\psi_B^\perp\rangle\right|^2 = \left|\mathrm{corr}_{A,B}\right|^2 = 1$. States that saturate the inequality are called *(Schrödinger) intelligent states*. The condition $\left|\mathrm{corr}_{A,B}\right|^2 = 1$ can be used to find intelligent states, although this is not easier than solving for equality in the Schrödinger relation directly.

4 The textbook proof and the Cauchy–Schwarz inequality

The textbook proofs of the Robertson and Schrödinger uncertainty relations are based on the Cauchy–Schwarz inequality

$$\left|\langle f|g\rangle\right|^2 \leq \langle f|f\rangle\langle g|g\rangle. \tag{31}$$

Note that the proofs given in Sect. 3 also make use of a special case of this inequality: that for unit vectors $\left|\langle f|g\rangle\right|^2 \leq 1$. This is applied to $|f\rangle = \left|\psi_A^\perp\right\rangle$, $|g\rangle = \left|\psi_B^\perp\right\rangle$. My aim is not to eliminate any use of the Cauchy–Schwarz inequality, but just to argue that the proof is more memorable if the inequality is applied in a different way than in the standard proof.

In the standard proof, the Cauchy–Schwarz inequality is applied to the two vectors $|f\rangle = (A - \langle A\rangle)\,|\psi\rangle$ and $|g\rangle = (B - \langle B\rangle)\,|\psi\rangle$ to obtain

$$\left|\langle\psi|(A - \langle A\rangle)(B - \langle B\rangle)|\psi\rangle\right|^2 \leq \left\langle\psi|(A - \langle A\rangle)^2|\psi\right\rangle\left\langle\psi|(B - \langle B\rangle)^2|\psi\right\rangle. \tag{32}$$

A few lines of messy algebra and cancellations, which I will spare you the details of, yield

$$(\Delta A)^2\,(\Delta B)^2 \geq \left|\frac{1}{2}\,\langle\{A, B\}\rangle - \langle A\rangle\,\langle B\rangle + \frac{1}{2}\,\langle[A, B]\rangle\right|^2, \tag{33}$$

from which we can derive the Schrödinger and Robertson relations by recognizing the real and imaginary parts of the right hand side.

As physics students do not often see the Cauchy–Schwarz inequality prior to their first course on quantum mechanics, most textbooks include a proof of this as well. One of the common proofs uses reasoning similar to that which we used to establish the Aharonov–Vaidman identity. It starts by recognizing that $|g\rangle$ can be written as

$$|g\rangle = \alpha\,|f\rangle + \beta\left|f^\perp\right\rangle, \tag{34}$$

where $\left|f^\perp\right\rangle$ is a unit vector that is orthogonal to $|f\rangle$. To find α, take the inner product of this with $|f\rangle$, which yields $\alpha = \langle f|g\rangle / \langle f|f\rangle$. Substituting this back into (34) and then taking the inner product of $|g\rangle$ with itself give

$$\langle g|g\rangle = \frac{\left|\langle f|g\rangle\right|^2}{\langle f|f\rangle} + |\beta|^2. \tag{35}$$

The Cauchy–Schwarz inequality follows from this by recognizing that $|\beta|^2$ is real and non-negative.

Summarizing, the standard proof of the Robertson inequality consists of: proving the Cauchy–Schwarz inequality and then finding convenient vectors to insert into the inequality that will yield terms involving ΔA and ΔB after some algebra. From the Aharonov–Vaidman identity, we can see that the reason the choice $|f\rangle = (A - \langle A\rangle)\,|\psi\rangle$ and $|g\rangle = (B - \langle B\rangle)\,|\psi\rangle$ is guaranteed to work is that $|f\rangle = \Delta A\left|\psi_A^\perp\right\rangle$ and $|g\rangle = \Delta B\left|\psi_B^\perp\right\rangle$.

After inserting these choices, one has to multiply out and simplify the expressions in the Cauchy–Schwarz inequality. This involves recognizing things like $\langle A\rangle\langle\psi|A|\psi\rangle = \langle A\rangle^2$ and then canceling several terms. It is

difficult for students to follow the full details of this in a lecture. In the approach using the Aharonov–Vaidman relation, we already have expressions involving ΔA and ΔB, so it is easier to see how to get an expression involving $\Delta A \Delta B$. This expression has fewer terms and there is less cancellation to do.

Although the approach using the Aharonov–Vaidman identity uses the Cauchy–Schwarz inequality in a less convoluted way, it uses similar mathematical ideas. For vectors $|f\rangle$ and $|g\rangle$, we can write $|g\rangle$ in terms of $|f\rangle$ and an orthogonal vector, as in the proof of Cauchy–Schwarz, or we can write both vectors in terms of a third vector $|h\rangle$ as

$$|f\rangle = \alpha_1 |h\rangle + \beta_1 \left|h_f^\perp\right\rangle, \tag{36}$$

$$|g\rangle = \alpha_2 |h\rangle + \beta_2 \left|h_g^\perp\right\rangle, \tag{37}$$

where $\left|h_f^\perp\right\rangle$ and $\left|h_g^\perp\right\rangle$ are (generally different) vectors orthogonal to $|h\rangle$ and $\alpha_1, \beta_1, \alpha_2, \beta_2$ are complex coefficients. This is what we do in the proof of the Aharonov–Vaidman identity with the choices $|f\rangle = A|\psi\rangle$, $|g\rangle = B|\psi\rangle$ and $|h\rangle = |\psi\rangle$. The advantage of this approach is that it immediately yields expressions involving the expectation values and standard deviations of the observables, with which it is easy to see what to do to get the uncertainty relations. From this point of view, the standard proof looks like shoehorning something into the Cauchy–Schwarz inequality that will yield standard deviations, and then backtracking to a point more easily obtained from the Aharonov–Vaidman identity. At the end of the day, both approaches use the same mathematics, but the Aharonov–Vaidman approach does so in a simpler and more direct way.

I would go so far as to say that whenever you are tempted to use the Cauchy–Schwarz inequality to prove a relationship between standard deviations of observables in quantum mechanics, you will have an easier time working from the Aharonov–Vaidman identity (and the special case $\left|\langle f|g\rangle\right|^2 \leq 1$ of the Cauchy–Schwarz inequality for unit vectors) instead. Sections 6 and 7 give more examples of this.

I end this section by showing that you can prove the Cauchy–Schwarz inequality from the Aharonov–Vaidman identity. I include this not because I think it is the best way to prove the Cauchy–Schwarz inequality, but because finding alternative proofs of the Cauchy–Schwarz inequality is the mathematician's equivalent of the sport of finding new uncertainty relations in quantum mechanics. It also shows that, in principle, there is nothing that can be proved using the Cauchy–Schwarz inequality that could not be proved using the Aharonov–Vaidman identity. Of course, outside the context of standard deviations in quantum mechanics, using the Aharonov–Vaidman identity instead of the Cauchy–Schwarz inequality is unlikely to yield a better proof.

Proposition 4.1 (Cauchy–Schwarz Inequality) *Let $|f\rangle$ and $|g\rangle$ be two vectors in a Hilbert space $\mathcal{H}$. Then*

$$\left|\langle f|g\rangle\right|^2 \leq \langle f|f\rangle\langle g|g\rangle. \tag{38}$$

Proof First, note that the inequality trivially holds whenever $\langle f|g\rangle = 0$ and that $\langle f|f\rangle = 0$ implies $\langle f|g\rangle = 0$. Therefore, we can assume that both $\langle f|g\rangle \neq 0$ and $\langle f|f\rangle > 0$.

Let $P = |g\rangle\langle g|$. Note this is not necessarily a projector because $|g\rangle$ does not have to be normalized, but it is a Hermitian operator. Applying the Aharonov–Vaidman identity to P and $|f\rangle$ gives

$$P|f\rangle = \langle P\rangle |f\rangle + \Delta P \left|f_P^\perp\right\rangle, \tag{39}$$

or equivalently,

$$|g\rangle\langle g|f\rangle = \frac{\langle f|g\rangle\langle g|f\rangle}{\langle f|f\rangle}|f\rangle + \Delta P \left|f_P^\perp\right\rangle. \tag{40}$$

Taking the inner product with $\left|f_P^{\perp}\right\rangle$ gives

$$\left\langle f_P^{\perp}|g\right\rangle\langle g|f\rangle = \Delta P\,\langle f|f\rangle, \tag{41}$$

where we used the fact that $\left\langle f_P^{\perp}|f_P^{\perp}\right\rangle = \langle f|f\rangle$. Rearranging and taking the complex conjugate give

$$\left\langle g|f_P^{\perp}\right\rangle = \frac{\Delta P\,\langle f|f\rangle}{\langle f|g\rangle}. \tag{42}$$

Now, taking the inner product of (40) with $|g\rangle$ gives

$$\langle g|g\rangle\langle g|f\rangle = \frac{\langle f|g\rangle\langle g|f\rangle}{\langle f|f\rangle}\langle g|f\rangle + \Delta P\left\langle g|f_P^{\perp}\right\rangle. \tag{43}$$

Multiplying both sides by $\langle f|f\rangle / \langle g|f\rangle$ gives

$$\langle f|f\rangle\langle g|g\rangle = \langle f|g\rangle\langle g|f\rangle + \frac{\Delta P\,\langle g|f_P^{\perp}\rangle\langle f|f\rangle}{\langle g|f\rangle}. \tag{44}$$

Substituting (42) into this gives

$$\langle f|f\rangle\langle g|g\rangle = \langle f|g\rangle\langle g|f\rangle + \frac{(\Delta P)^2\,|\langle f|f\rangle|^2}{\langle f|g\rangle\langle g|f\rangle}, \tag{45}$$

or

$$\langle f|f\rangle\langle g|g\rangle = |\langle f|g\rangle|^2 + \frac{(\Delta P)^2\,|\langle f|f\rangle|^2}{|\langle f|g\rangle|^2}. \tag{46}$$

Now, the terms ΔP, $\langle f|f\rangle$ and $|\langle f|g\rangle|$ are all real and non-negative. Hence,

$$\langle f|f\rangle\langle g|g\rangle \geq |\langle f|g\rangle|^2. \tag{47}$$

□

5 Pedagogical notes

To teach the Robertson uncertainty relation via the Aharonov–Vaidman identity, you first have to establish the Aharonov–Vaidman identity. For the purposes of proving the Robertson uncertainty relation, it is sufficient to restrict the operator in the identity to be Hermitian and the vector $|\psi\rangle$ to be a unit vector, as I shall in this section.

In my experience, not all students immediately understand why, given a unit vector $|\psi\rangle$, any other unit vector $|\phi\rangle$ can be written as

$$|\phi\rangle = \alpha\,|\psi\rangle + \beta\left|\psi^{\perp}\right\rangle, \tag{48}$$

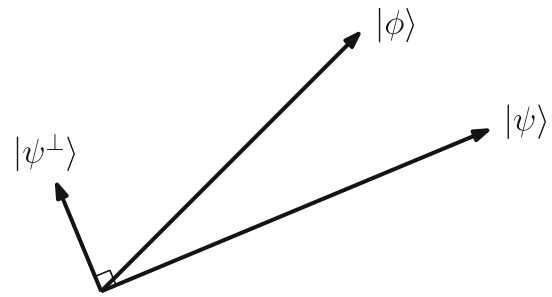

Fig. 1 Diagram showing that there exists a unit vector $|\psi^\perp\rangle$ such that $|\psi\rangle$ and $|\psi^\perp\rangle$ form an orthogonal basis for the two dimensional subspace of $\mathcal{H}$ spanned by $|\psi\rangle$ and $|\phi\rangle$

where $|\psi^\perp\rangle$ is a unit vector orthogonal to $|\psi\rangle$. They will probably have seen Gram–Schmidt orthogonalization in a linear algebra class, but may have difficulty using that knowledge here due to the jump to abstract Hilbert spaces and Dirac notation. To aid intuition, I remark that $|\psi\rangle$ and $|\phi\rangle$ span a two-dimensional subspace of $\mathcal{H}$ and show them Fig. 1.

By the process of Gram–Schmidt orthogonalization, we can construct an orthornormal basis for this subspace consisting of $|\psi\rangle$ and

$$|\psi^\perp\rangle = \frac{1}{\sqrt{1-|\langle\phi|\psi\rangle|^2}}\left(|\phi\rangle - |\psi\rangle\langle\psi|\phi\rangle\right), \tag{49}$$

from which we have (48) with $\alpha = \langle\psi|\phi\rangle$ and $\beta = \sqrt{1-|\langle\phi|\psi\rangle|^2}$.

In my quantum mechanics classes, I set students in-class activities that involve things like deriving important equations or making order of magnitude estimates. These take about 5–10 min each and are done in pairs. I usually do two or three such activities per class. I believe this increases active engagement and retention of the main principles. I try to reduce the number of long derivations that I do myself on the board, because I think they cause confusion about what the most important equations are and the derivations are rarely remembered by the students. However, I also do not want to set the students a long and complicated derivation to do themselves in class, so I try to find shorter derivations that they can do with guidance instead. The proof of the Robertson relation from the Aharonov–Vaidman relation is better suited to this approach than the standard proof.

After establishing (48), I set students the following activity.

In class activity

Given that $A|\psi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle$, *find* α *and* β *in terms of the expectation value* $\langle A\rangle$ *and standard deviation* ΔA *of* A *in the state* $|\psi\rangle$.

Although some students can do this straight away, most need some help. During the course of the activity, I walk around the class to get an idea of how they are doing. When it seems like many students are stuck, I reveal the following three hints in sequence.

Hints

1. *Try taking the inner product of* $A|\psi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle$ *with other states.*
2. *Try taking the inner product of* $A|\psi\rangle$ with $|\psi\rangle$.
3. *Try taking the inner product of* $A|\psi\rangle$ with itself.

Although most students can get $\alpha = \langle A\rangle$ either straight away or after the first hint, $|\beta| = \Delta A$ is more challenging. After taking the inner product with $|\psi\rangle$, the obvious instinct is to take the inner product with $|\psi^\perp\rangle$, which does not help, so the third hint is usually needed. After this, it is a short hop to the Robertson relation via the proof given in Sect. 3.

I think it would be more difficult to teach the standard proof in this way. One would either have to ask the students to derive the Cauchy–Schwarz inequality for themselves or derive the Robertson relation from Cauchy–Schwarz. The former is a bit abstract for a quantum mechanics class and the latter involves a lot of algebra and cancellations

with a high potential for making mistakes. Both would require a large number of hints. In contrast, the proof of the Aharonov–Vaidman identity is relatively short, and I think that students who retain the identity are more likely to be able to reconstruct the proof of the Robertson relation for themselves.

6 Other uncertainty relations for standard deviations

Despite the ubiquity of the Schrödinger–Robertson uncertainty relations in quantum mechanics classes, there are good reasons to go beyond them. For example, consider a spin-1/2 particle with spin operators S_x, S_y and S_z. For this case, the Robertson uncertainty is $\Delta S_x \Delta S_y \geq \hbar\,|\langle S_z\rangle|$. Let $|x+\rangle$ be the spin-up state in the x direction. For this state we have $\langle S_z\rangle = 0$, which is perfectly valid because $|x+\rangle$ is an eigenstate of S_x and hence $\Delta S_x = 0$. However, because $[S_x, S_y] \neq 0$ there is necessarily some uncertainty in S_y and in fact $\Delta S_y = \hbar/2$. The Schrödinger relation also yields $\Delta S_x \Delta S_y \geq 0$. So the Schrödinger–Robertson relations do not capture all uncertainty trade-offs that necessarily exist in quantum mechanics.

More generally, for bounded operators A and B, any uncertainty relation of the form $\Delta A \Delta B \geq f\,(A, B, |\psi\rangle)$ for some function f must necessarily have $f\,(A, B, |\psi\rangle) = 0$ whenever $|\psi\rangle$ is an eigenstate of A or B. For this reason, it makes sense to seek uncertainty relations that bound the sum of standard deviations $\Delta A + \Delta B$, the sum of variances $(\Delta A)^2 + (\Delta B)^2$, or more exotic combinations. We shall discuss the Maccone–Pati relations, and some simple generalizations, in this section.

Uncertainty relations are classified as either *state dependent* or *state independent*, depending on whether the right hand side of the inequality depends on the state $|\psi\rangle$. For two observables A and B, a state dependent uncertainty relation is of the form $f(\Delta A, \Delta B) \geq g(A, B, |\psi\rangle)$, where f and g are specified functions, whereas a state independent uncertainty relation would be of the form $f(\Delta A, \Delta B) \geq g(A, B)$, noting that g is no longer allowed to depend on $|\psi\rangle$.

On the face of it, a state-dependent uncertainty relation is a strange idea, since, for any given normalized state $|\psi\rangle$, we can always just calculate the uncertainties ΔA and ΔB and get the exact value of $f(\Delta A, \Delta B)$. Therefore, bounds on uncertainty that apply to all states seem more useful.

However, a state-dependent uncertainty relation can be a useful step in deriving a state independent one. This can happen in two ways. First, it may happen that, for a particular choice of the observables A and B, the function $g(A, B, |\psi\rangle)$ turns out not to depend on $|\psi\rangle$. For example, the Robertson relation $\Delta A \Delta B \geq \frac{1}{2}\left|\langle\psi\left|[A, B]\right|\psi\rangle\right|$ is state dependent, but if we choose $A = x$, $B = p$, then $\left|\langle\psi\left|[A, B]\right|\psi\rangle\right| = 1$ and so we get the Heisenberg relation $\Delta x \Delta p \geq \frac{\hbar}{2}$, which is state independent. Since the main point of proving the Robertson uncertainty relation in a quantum mechanics class is to give a rigorous derivation of the Heisenberg relation, its state dependence does no harm. However, the utility of the Robertson relation for other classes of observable, such as spin components, is more questionable. Despite the fact that I have asked students to compute it for states of a spin-1/2 particle as a homework problem, I do not think there is ever a need to do this in practice, as it is just as easy to calculate the exact uncertainties.

The second way of obtaining a state independent uncertainty relation from a state dependent one is to optimize, i.e. if $f(\Delta A, \Delta B) \geq g(A, B, |\psi\rangle)$ then[2]

$$f(\Delta A, \Delta B) \geq \min_{|\psi\rangle} g(A, B, |\psi\rangle). \tag{50}$$

Of course, if $f(\Delta A, \Delta B) = \Delta A \Delta B$ and A and B are bounded operators, then this leads to the trivial relation $\Delta A \Delta B \geq 0$ because we can choose $|\psi\rangle$ to be an eigenstate of either A or B. However, for sums and more general combinations of observables, optimization can lead to a nontrivial relation.

Further, if we are considering a set of experiments that can only prepare a subset of the possible states, then we can get an uncertainty relation that applies to those states by optimizing over the subset. An example might

[2] The minimum in (50) may have to be replaced by an infimum, depending on the Hilbert space that the observables are defined on.

be experiments in which we can only prepare the system in a Gaussian state. Although this does not yield a state independent uncertainty relation, it is more useful than a completely state dependent one, as it allows us to bound the possible uncertainties for a class of relevant states.

To summarize, state-dependent uncertainty relations are a strange idea, and I am not sure whether they would ever have been considered had not Robertson introduced one as a way-point in proving the Heisenberg relation. However, they can be useful in proving more generally applicable uncertainty relations. The relations that we discuss here are state dependent.

The remainder of this section is structured as follows. In Sect. 6.1 we prove two propositions called the *sum relations* that will be used repeatedly using the Aharonov–Vaidman identity. In Sect. 6.2, we give an Aharonov–Vaidman based proof of the Maccone-Pati uncertainty relations, and in in Sect. 6.3 we give some simple generalizations.

6.1 The sum relations

Proposition 6.1 *Let A and B be linear operators acting on $\mathcal{H}$. Then, for any $|\psi\rangle \in \mathcal{H}$,*

$$\Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle = \Delta A\left|\psi^{\perp}_{A}\right\rangle + \Delta B\left|\psi^{\perp}_{B}\right\rangle.$$

Proof Apply the Aharonov–Vaidman identity to $A+B$ in two different ways. The first way is

$$\begin{aligned}(A+B)|\psi\rangle &= \langle A+B\rangle|\psi\rangle + \Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle \\ &= (\langle A\rangle + \langle B\rangle)|\psi\rangle + \Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle, \end{aligned} \tag{51}$$

and the second is

$$\begin{aligned}(A+B)|\psi\rangle &= A|\psi\rangle + B|\psi\rangle \\ &= (\langle A\rangle + \langle B\rangle)|\psi\rangle + \Delta A\left|\psi^{\perp_A}\right\rangle + \Delta B\left|\psi^{\perp}_{B}\right\rangle. \end{aligned} \tag{52}$$

Subtracting (52) from (51) and rearranging give the desired result. □

The next proposition comes from [19]. Here, the proof relies on Poposition 6.1 and so is based on the Aharonov–Vaidman relation. The original proof uses a different method and is a little more complicated.

Proposition 6.2 (The sum relation) *Let A and B be two linear operators acting on a Hilbert space $\mathcal{H}$. Then,*

$$\Delta(A+B) \leq \Delta A + \Delta B.$$

Proof Let $|\psi\rangle$ in Proposition 6.1 be a unit vector. Then, starting from $\Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle = \Delta A\left|\psi^{\perp}_{A}\right\rangle + \Delta B\left|\psi^{\perp}_{B}\right\rangle$ and taking the inner product with $\left|\psi^{\perp}_{A+B}\right\rangle$ gives

$$\Delta(A+B) = \Delta A\left\langle\psi^{\perp}_{A+B}\middle|\psi^{\perp}_{A}\right\rangle + \Delta B\left\langle\psi^{\perp_{A+B}}\middle|\psi^{\perp}_{B}\right\rangle.$$

The left hand side of this equation is a real number, so the right hand side must be too. Therefore, we can take the real part of each term to give

$$\Delta(A+B) = \Delta A\mathrm{Re}\left(\left\langle\psi^{\perp}_{A+B}\middle|\psi^{\perp}_{A}\right\rangle\right) + \Delta B\mathrm{Re}\left(\left\langle\psi^{\perp_{A+B}}\middle|\psi^{\perp}_{B}\right\rangle\right),$$

but the real part of an inner product between two unit vectors is ≤ 1, so we have

$$\Delta(A+B) \leq \Delta A + \Delta B.$$

□

From the proof, we see that the equality condition for the sum relation is

$$\mathrm{Rcorr}(A+B, A) = \mathrm{Rcorr}(A+B, B) = 1.$$

Remark 6.3 For a set of linear operators $A_1, A_2, \ldots, A_n$ on a Hilbert space $\mathcal{H}$, Proposition 6.1 is easily generalized to

$$\Delta\left(\sum_{j=1}^{n} A_j\right)\left|\psi^{\perp}_{\sum_{j=1}^{n} A_j}\right\rangle = \sum_{j=1}^{n} \Delta A_j \left|\psi^{\perp}_{A_j}\right\rangle, \tag{53}$$

by applying the Aharonov–Vaidman identity to $\sum_{j=1}^{n} A_j$. Similarly, Proposition 6.2 is easily generalized to

$$\Delta\left(\sum_{j=1}^{n} A_j\right) \leq \sum_{j=1}^{n} \Delta A_j, \tag{54}$$

by taking the inner product of (53) with $\left|\psi^{\perp}_{\sum_{j=1}^{n} A_j}\right\rangle$. We will also refer to the generalization in (54) as the sum relation.

6.2 The Maccone–Pati uncertainty relations

Between the time of Robertson's uncertainty relation and now, there has always been some literature on uncertainty relations for variances and standard deviations. However, the field was reinvigorated in 2014, when Maccone and Pati [20] proved a pair of uncertainty relations for sums of variances, which always give a nontrivial bound, even in the case of an eigenstate of an observable.

Here, we give Aharonov–Vaidman based proofs of the Maccone–Pati relations.[3]

Theorem 6.4 (The first Maccone–Pati uncertainty relation) *Let A and B be Hermitian operators on a Hilbert space $\mathcal{H}$ and let $|\psi\rangle \in \mathcal{H}$ be a unit vector. Then,*

$$(\Delta A)^2 + (\Delta B)^2 \geq \pm i \langle [A, B] \rangle + \left|\left\langle \psi^{\perp} \middle| (A \mp iB) \middle| \psi \right\rangle\right|^2, \tag{55}$$

where $\left|\psi^{\perp}\right\rangle$ is any unit vector orthogonal to $|\psi\rangle$.

Proof We will prove $(\Delta A)^2 + (\Delta B)^2 \geq -i \langle [A, B] \rangle + \left|\left\langle \psi^{\perp} \middle| (A + iB) \middle| \psi \right\rangle\right|^2$ by applying the Aharonov–Vaidman identity to $(A + iB)$. The proof of the other inequality follows by replacing $A + iB$ with $A - iB$. Note that, even though A and B are Hermitian, $A + iB$ is not, so it is crucial that we previously generalized the Aharonov–Vaidman identity to arbitrary linear operators.

Applying the Aharonov–Vaidman identity to $A + iB$ gives

$$(A + iB)|\psi\rangle = (\langle A \rangle + i \langle B \rangle)|\psi\rangle + \Delta(A + iB)\left|\psi^{\perp}_{A+iB}\right\rangle.$$

Taking the inner product with any unit vector $\left|\psi^{\perp}\right\rangle$ orthogonal to $|\psi\rangle$ gives

$$\left\langle \psi^{\perp} \middle| (A + iB) \middle| \psi \right\rangle = \Delta(A + iB)\left\langle \psi^{\perp} \middle| \psi^{\perp}_{A+iB} \right\rangle,$$

[3] Although the Aharonov–Vaidman identity is used in [20], it is not used in the proofs of the uncertainty relations.

and taking the modulus squared of this gives

$$\left|\left\langle\psi^{\perp}\right|(A+iB)\left|\psi\right\rangle\right|^2 = (\Delta(A+iB))^2\left|\left\langle\psi^{\perp}\right|\left.\psi^{\perp}_{A+iB}\right\rangle\right|^2.$$

Now, $\left|\langle\psi^{\perp}|\psi^{\perp}_{A+iB}\rangle\right| \leq 1$, so

$$(\Delta(A+iB))^2 \geq \left|\left\langle\psi^{\perp}\right|(A+iB)\left|\psi\right\rangle\right|^2.$$

The result now follows by expanding $(\Delta(A+iB))^2$ as follows.

$$\begin{aligned}(\Delta(A+iB))^2 &= \langle(A-iB)(A+iB)\rangle - \langle A-iB\rangle\langle A+iB\rangle \\ &= \left\langle A^2\right\rangle + \left\langle B^2\right\rangle + i\langle[A,B]\rangle - \langle A\rangle^2 - \langle B\rangle^2 \\ &= (\Delta A)^2 + (\Delta B)^2 + i\langle[A,B]\rangle.\end{aligned}$$

□

Theorem 6.5 (The second Maccone–Pati uncertainty relation) *Let A and B be linear operators on a Hilbert space $\mathcal{H}$ and let $|\psi\rangle \in \mathcal{H}$ be a unit vector. Then,*

$$(\Delta A)^2 + (\Delta B)^2 \geq \frac{1}{2}\left|\left\langle\psi^{\perp}_{A+B}\right|(A+B)\left|\psi\right\rangle\right|^2. \tag{56}$$

Proof Applying the Aharonov–Vaidman identity to $A+B$ gives

$$(A+B)|\psi\rangle = (\langle A\rangle + \langle B\rangle)|\psi\rangle + \Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle.$$

Taking the inner product with $\left|\psi^{\perp}_{A+B}\right\rangle$ gives

$$\begin{aligned}\left\langle\psi^{\perp}_{A+B}\right|(A+B)\left|\psi\right\rangle &= \Delta(A+B) \\ &\leq \Delta A + \Delta B,\end{aligned}$$

where the second line follows from the sum relation.

We could stop here and regard $\Delta A + \Delta B \geq \left\langle\psi^{\perp}_{A+B}\right|(A+B)\left|\psi\right\rangle$ as an uncertainty relation, but Maccone and Pati wanted a relation in terms of variances to compare to their first result. To do this, we take the modulus squared of both sides to obtain

$$(\Delta A + \Delta B)^2 \geq \left|\left\langle\psi^{\perp}_{A+B}\right|(A+B)\left|\psi\right\rangle\right|^2.$$

The result now follows from the real number inequality $x^2 + y^2 \geq \frac{1}{2}(x+y)^2$ with $x = \Delta A$ and $y = \Delta B$. For completeness, this inequality is proved as follows.

$$\begin{aligned}&0 \leq (x-y)^2 = x^2 + y^2 - 2xy \\ &\Rightarrow x^2 + y^2 \geq 2xy \\ &\Rightarrow 2x^2 + 2y^2 \geq x^2 + y^2 + 2xy \\ &\Rightarrow 2x^2 + 2y^2 \geq (x+y)^2 \\ &\Rightarrow x^2 + y^2 \geq \frac{1}{2}(x+y)^2.\end{aligned}$$

□

6.3 Generalizations

Generalizations of the Maccone–Pati uncertainty relations can be obtained by applying the Aharonov–Vaidman identity to more general linear combinations $\alpha A + \beta B$, where $\alpha, \beta \in \mathbb{C}$. This gives

$$(\alpha A + \beta B)\,|\psi\rangle = (\alpha\,\langle A\rangle + \beta\,\langle B\rangle)\,|\psi\rangle + \Delta(\alpha A + \beta B)\left|\psi^{\perp}_{\alpha A+\beta B}\right\rangle. \tag{57}$$

Applying the strategy we used to prove Theorem 6.4, we can take the inner product of this with an arbitrary unit vector $\left|\psi^{\perp}\right\rangle$ that is orthogonal to $|\psi\rangle$, which gives

$$\left\langle\psi^{\perp}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle = \Delta(\alpha A + \beta B)\left\langle\psi^{\perp}\middle|\psi^{\perp}_{\alpha A+\beta B}\right\rangle.$$

We can now take the modulus squared of this and recognize that $0 \leq \left|\left\langle\psi^{\perp}\middle|\psi^{\perp}_{\alpha A+\beta B}\right\rangle\right|^2 \leq 1$ to obtain

$$\left|\left\langle\psi^{\perp}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle\right|^2 \leq \Delta(\alpha A + \beta B).$$

Next, we can expand $\Delta(\alpha A + \beta B)$ and rearrange to obtain

$$\begin{aligned}|\alpha|^2\,(\Delta A)^2 + |\beta|^2\,(\Delta B)^2 \geq\ & -\mathrm{Re}(\alpha^*\beta)\,(\langle\{A, B\}\rangle - 2\,\langle A\rangle\,\langle B\rangle) - i\mathrm{Im}\left(\alpha^*\beta\right)\langle[A, B]\rangle \\ & + \left|\left\langle\psi^{\perp}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle\right|^2.\end{aligned} \tag{58}$$

Substituting $\alpha = 1, \beta = i$ and $\alpha = 1, \beta = -i$ immediately yields the first Maccone–Pati uncertainty relation.

Alternatively, we can apply the strategy used to prove Theorem 6.5. Starting from (57), we can take the inner product with $\left|\psi^{\perp}_{\alpha A+\beta B}\right\rangle$ and rearrange to obtain

$$\Delta(\alpha A + \beta B) = \left\langle\psi^{\perp}_{\alpha A+\beta B}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle.$$

Using the sum relation, together with $\Delta(\alpha A) = |\alpha|\Delta A$ gives

$$|\alpha|\Delta A + |\beta|\Delta B \geq \left\langle\psi^{\perp}_{\alpha A+\beta B}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle.$$

Finally, squaring and using the inequality $x^2 + y^2 \geq \frac{1}{2}(x + y)^2$ gives

$$|\alpha|^2\,(\Delta A)^2 + |\beta|^2\,(\Delta B)^2 \geq \frac{1}{2}\left|\left\langle\psi^{\perp}_{\alpha A+\beta B}\middle|(\alpha A + \beta B)\middle|\psi\right\rangle\right|^2. \tag{59}$$

The inequalities (58) and (59) are related to some of the generalizations of the Maccone–Pati uncertainty relations that have previously appeared in the literature [21,28]. For example, (58) can be used to derive an uncertainty relation that has appeared in the literature under the name "weighted uncertainty relation" [28]. To do so, we set $\alpha = \sqrt{\lambda}$, $\beta = \pm i/\sqrt{\lambda}$ in (58), where $\lambda > 0$. This yields

$$\lambda\,(\Delta A)^2 + \frac{1}{\lambda}\,(\Delta B)^2 \geq \pm i\,\langle[A, B]\rangle + \frac{1}{\lambda}\left|\left\langle\psi^{\perp}\middle|(\lambda A \mp i B)\middle|\psi\right\rangle\right|^2.$$

This is an uncertainty relation in its own right, but the relation in [28] comes from adding this to (55), which yields

$$\begin{aligned}(1 + \lambda)\,(\Delta A)^2 + \left(1 + \frac{1}{\lambda}\right)(\Delta B)^2 \geq\ & \pm 2i\,\langle[A, B]\rangle\left|\left\langle\psi_1^{\perp}\middle|(A \mp i B)\middle|\psi\right\rangle\right|^2 \\ & + \frac{1}{\lambda}\left|\left\langle\psi_2^{\perp}\middle|(\lambda A \mp i B)\middle|\psi\right\rangle\right|^2,\end{aligned} \tag{60}$$

where $\left|\psi_1^{\perp}\right\rangle$ and $\left|\psi_2^{\perp}\right\rangle$ are (possibly different) unit vectors that are orthogonal to $|\psi\rangle$.

This is intended as a simple example of a generalization that is easily obtained from the Aharonov–Vaidman identity, but I expect many other uncertainty relations that are usually proved using the Cauchy–Schwarz inequality or the parallelogram law would also have simple Aharonov–Vaidman based proofs.

7 Quantum propagation of uncertainty

In this section, we develop generalizations of the classical formulas for the propagation of uncertainty. We start with the case of linear functions in Sect. 7.1, for which exact formulas are easy to obtain, before moving on to the general, possibly nonlinear, case in Sect. 7.2, for which we have to employ a Taylor series approximation.

7.1 Linear functions

We start with the simplest case: a sum of two observables. Classically, if A and B are random variables, then

$$[\Delta(A+B)]^2 = (\Delta A)^2 + (\Delta B)^2 + 2\Delta A \Delta B \,\mathrm{corr}_{A,B}. \tag{61}$$

Consider an experiment consisting of multiple runs. On each run, the quantities A and B are measured. These quantities are formalized as random variables because we assume that our experiments are subject to random statistical fluctuations, and that the "true" values that we are seeking are the means $\langle A\rangle$ and $\langle B\rangle$ of these random processes. We then use the average values calculated from the data as estimates of $\langle A\rangle$ and $\langle B\rangle$, and the standard deviations as a measure of the error in our experiment. If we are actually interested in the quantity $A+B$, then we would sum the averages to form our estimate of $\langle A+B\rangle$, and we would use (61) to determine the error in our estimate of $\langle A+B\rangle$. Using (61) in this way is called the *propagation of uncertainty* or *propagation of error*.

If the random variables, A and B, are independent, which would be the case if the randomness were due to independent statistical errors, then $\mathrm{corr}_{A,B} = 0$ and we would have

$$[\Delta(A+B)]^2 = (\Delta A)^2 + (\Delta B)^2,$$

which is the formula for propagation of uncertainty that is most commonly used in practice.

We now want to generalize these formulas by replacing classical random variables with quantum observables. The generalization of (61) is as follows.

Theorem 7.1 *Let A and B be Hermitian operators on a Hilbert space $\mathcal{H}$. Then,*

$$[\Delta(A+B)]^2 = (\Delta A)^2 + (\Delta B)^2 + 2\Delta A \Delta B \,\mathrm{Rcorr}_{A,B} \tag{62}$$

$$= (\Delta A)^2 + (\Delta B)^2 + \langle\{A, B\}\rangle - 2\langle A\rangle\langle B\rangle. \tag{63}$$

Proof Proposition 6.1 implies that, for any unit vector $|\psi\rangle \in \mathcal{H}$,

$$\Delta(A+B)\left|\psi^{\perp}_{A+B}\right\rangle = \Delta A\left|\psi^{\perp}_{A}\right\rangle + \Delta B\left|\psi^{\perp}_{B}\right\rangle.$$

Taking the inner product of this with itself gives

$$\begin{aligned}[\Delta(A+B)]^2 &= (\Delta A)^2 + (\Delta B)^2 + \Delta A\Delta B\left(\left\langle\psi^{\perp}_{A}\middle|\psi^{\perp}_{B}\right\rangle + \left\langle\psi^{\perp}_{B}\middle|\psi^{\perp}_{A}\right\rangle\right)\\ &= (\Delta A)^2 + (\Delta B)^2 + 2\Delta A\Delta B\,\mathrm{Re}\left(\left\langle\psi^{\perp}_{A}\middle|\psi^{\perp}_{B}\right\rangle\right).\end{aligned}$$

Applying (13) completes the proof. □

Remark 7.2 For operators $A_1, A_2, \cdots, A_n$ and real numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$, Theorem 7.1 is easily generalized to

$$\left[\Delta\left(\sum_{j=1}^{n}\alpha_j A_j\right)\right]^2 = \sum_{j=1}^{n}\alpha_j^2\left(\Delta A_j\right)^2 + \sum_{j\neq k}\alpha_j\alpha_k\Delta A_j\Delta A_k\,\mathrm{Rcorr}_{A_j,A_k}$$

$$= \sum_{j=1}^{n} \alpha_j^2 \left(\Delta A_j\right)^2 + \sum_{j \neq k} \alpha_j \alpha_k \left(\left\langle \{A_j, A_k\} \right\rangle - 2 \left\langle A_j \right\rangle \left\langle A_k \right\rangle \right).$$

Although Theorem 7.1 is a true theorem about quantum observables, it cannot be used to propagate uncertainty in the same way as its classical counterpart. Classically, we can measure A and B together in the same run of the experiment. We can then estimate $A+B$ by summing the average values of A and B that we found in the experiment. We also have all the information we need to calculate the uncertainty $\Delta(A+B)$, i.e., ΔA, ΔB, $\langle A \rangle$, $\langle B \rangle$ and $\langle AB \rangle$, so we can determine the uncertainty without doing any more experiments.

In quantum mechanics, this is not the case. When A and B do not commute, they cannot both be accurately measured on the same run of an experiment. We can still estimate their expectation values by measuring A on half of the runs of the experiment and B on the other half and taking averages. Since $\langle A+B \rangle = \langle A \rangle + \langle B \rangle$, summing these averages is still a way of estimating $\langle A+B \rangle$. However, we do not have enough information to calculate $\Delta(A+B)$. The reason is that $\Delta(A+B)$ is the uncertainty in a *direct* measurement of $A+B$. Since A and B do not commute, this requires a different experimental setup from a measurement of A and B alone.

If we wanted to use (62) to calculate $\Delta(A+B)$, we would also have to estimate $\langle \{A, B\} \rangle$. The most straightforward way of doing this would be to measure the observable $\{A, B\} = AB + BA$, but this requires yet another different experimental setup, and one that is likely to be at least as complicated as measuring $A+B$ directly.

An exception to this are cases where $\{A, B\} = cI$ for some constant c, in which case $\langle \{A, B\} \rangle = c$ regardless of the state. In particular, this is true of the Pauli observables $\sigma_x, \sigma_y, \sigma_z$ of a qubit for which $\{\sigma_j, \sigma_k\} = \delta_{jk} I$, where j and k run over x, y, z. Therefore, if we measure σ_x on many qubits prepared in the same way and σ_y on another set of such qubits, we can estimate $\langle \sigma_x + \sigma_y \rangle$ and $\Delta(\sigma_x + \sigma_y)$ without doing any further experiments using the formula

$$\left[\Delta\left(\sigma_x + \sigma_y\right)\right]^2 = (\Delta\sigma_x)^2 + \left(\Delta\sigma_y\right)^2 - 2\left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle.$$

When $\{A, B\} \neq cI$, I do not know of any situations in which (62) would be useful in practice, but from a theoretical point of view it is the appropriate generalization of (61) to quantum mechanics, and this bolsters the case that $\mathrm{Rcorr}_{A,B}$ is the appropriate quantum generalization of the classical correlation.

7.2 Nonlinear functions

For nonlinear functions $f(A, B)$ of two random variables A and B, it is common to use a first-order Taylor expansion of $f(A, B)$ about the point $f(\langle A \rangle, \langle B \rangle)$ to derive an approximation for the variance $[\Delta f(A, B)]^2$ to second order in ΔA and ΔB. This yields the formula

$$[\Delta f(A, B)]^2 \approx \left(\left.\frac{\partial f}{\partial A}\right|_{A=\langle A \rangle, B=\langle B \rangle} \right)^2 (\Delta A)^2 + \left(\left.\frac{\partial f}{\partial B}\right|_{A=\langle A \rangle, B=\langle B \rangle} \right)^2 (\Delta B)^2 + \left.\frac{\partial f}{\partial A}\right|_{A=\langle A \rangle, B=\langle B \rangle} \left.\frac{\partial f}{\partial B}\right|_{A=\langle A \rangle, B=\langle B \rangle} \Delta A \Delta B \, \mathrm{corr}_{A,B}.$$

To avoid cluttering notation, I will write $\bar{A}$ for $A = \langle A \rangle$, so that we can more compactly write

$$[\Delta f(A, B)]^2 \approx \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}}^2 (\Delta A)^2 + \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}}^2 (\Delta B)^2 + \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} \Delta A \Delta B \, \mathrm{corr}_{A,B}. \tag{64}$$

When A and B are independent, this reduces to

$$[\Delta f(A, B)]^2 \approx \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}}^2 (\Delta A)^2 + \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}}^2 (\Delta B)^2,$$

which is the most commonly used form.

The quantum generalization of (64) is as follows.

Theorem 7.3 *Let A and B be Hermitian operators on a Hilbert space $\mathcal{H}$ and consider a function $f : \mathfrak{H}(\mathcal{H}) \times \mathfrak{H}(\mathcal{H}) \to \mathfrak{H}(\mathcal{H})$, where $\mathfrak{H}(\mathcal{H})$ is the space of Hermitian operators on $\mathcal{H}$. Then,*

$$[\Delta f(A,B)]^2 \approx \left.\frac{\partial f}{\partial A}\right|^2_{\bar{A},\bar{B}} (\Delta A)^2 + \left.\frac{\partial f}{\partial B}\right|^2_{\bar{A},\bar{B}} (\Delta B)^2 + \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} \Delta A \Delta B \,\mathrm{Rcorr}_{A,B}, \tag{65}$$

where $\approx$ means equality to second order in ΔA and ΔB.

Proof Consider the first-order Taylor expansion of $f(A,B)$ about the point $f_0 = f(\langle A\rangle, \langle B\rangle)$,

$$f(A,B) \approx f_0 + \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} A + \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} B.$$

Applying Proposition 6.1 to this gives

$$[\Delta f(A,B)] \left|\psi^{\perp}_{f(A,B)}\right\rangle \approx \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} \Delta A \left|\psi^{\perp}_{A}\right\rangle + \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} \Delta B \left|\psi^{\perp}_{B}\right\rangle.$$

Taking the inner product of this with itself gives

$$\begin{aligned}[\Delta f(A,B)]^2 &\approx \left.\frac{\partial f}{\partial A}\right|^2_{\bar{A},\bar{B}} (\Delta A)^2 + \left.\frac{\partial f}{\partial B}\right|^2_{\bar{A},\bar{B}} (\Delta B)^2 + \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} \Delta A \Delta B \,\mathrm{Re}\left(\left\langle \psi^{\perp}_{A} \middle| \psi^{\perp}_{B} \right\rangle\right) \\ &= \left.\frac{\partial f}{\partial A}\right|^2_{\bar{A},\bar{B}} (\Delta A)^2 + \left.\frac{\partial f}{\partial B}\right|^2_{\bar{A},\bar{B}} (\Delta B)^2 + \left.\frac{\partial f}{\partial A}\right|_{\bar{A},\bar{B}} \left.\frac{\partial f}{\partial B}\right|_{\bar{A},\bar{B}} \Delta A \Delta B \,\mathrm{Rcorr}_{A,B}.\end{aligned}$$

□

Remark 7.4 For operators $A_1, A_2, \ldots, A_n$ and a function $f(A_1, A_2, \cdots, A_n)$, Theorem 7.3 is easily generalized to

$$[\Delta f(A_1, A_2, \cdots, A_n)]^2 \approx \sum_{j=1}^{n} \left.\frac{\partial f}{\partial A_j}\right|^2_{\bar{A}} (\Delta A_j)^2 + \sum_{j\neq k} \left.\frac{\partial f}{\partial A_j}\right|_{\bar{A}} \left.\frac{\partial f}{\partial A_k}\right|_{\bar{A}} \Delta A_j \Delta A_k \,\mathrm{Rcorr}_{A_j,A_k}, \tag{66}$$

where $\bar{A}$ is shorthand for $A_1 = \langle A_1\rangle, A_2 = \langle A_2\rangle, \cdots A_n = \langle A_n\rangle$.

As a formula for propagating uncertainty, (65) inherits all of the problems of (62), but the problems are compounded further by use of the first-order Taylor approximation. This approximation is valid when ΔA and ΔB are suitably small compared to $\langle A\rangle$, $\langle B\rangle$, $f(\langle A\rangle, \langle B\rangle)$ and the derivatives of $f(A,B)$ at $A = \langle A\rangle$, $B = \langle B\rangle$. This is often the case in classical experiments where everything can be measured with a small statistical error. However, in quantum mechanics, when A and B do not commute, the (various) uncertainty relations tell us that there is necessarily a trade-off between the size of ΔA and ΔB. If one of them is small, then the other might necessarily have

to be large. For example, for the Pauli observables σ_x and σ_y, at least one of the uncertainties must be comparable in size to 1, which is the largest possible value of $\langle\sigma_x\rangle$ or $\langle\sigma_y\rangle$.

A case where the formula will work well is for a continuous variable system where $\Delta x \sim \Delta p \sim \sqrt{\hbar}$, and $\langle x\rangle, \langle p\rangle$ are large compared to $\sqrt{\hbar}$. But this is a case where you would expect classical physics to be a good approximation anyway.

I do not know whether there is a practical use of (65), but it is nonetheless a correct formal generalization of (64).

8 Dealing with mixed states

So far, we have dealt exclusively with the case of pure state vectors $|\psi\rangle$. However, all of our results can be extended to more general density operators ρ, which can represent mixed states. The most familiar way to do this is to make use of the concept of a *purification* of a density operator. Given a density operator on a Hilbert space $\mathcal{H}_S$, where S stands for "system", we can always find a pure state vector $|\psi\rangle_{SE} \in \mathcal{H}_S \otimes \mathcal{H}_E$, where E is the "environment", such that

$$\rho_S = \mathrm{Tr}_E\left(|\psi\rangle\langle\psi|_{SE}\right),$$

and Tr_E is the partial trace over $\mathcal{H}_E$. You can then apply the Aharonov–Vaidman identity to operators of the form $A_S \otimes I_E$ acting on a purification to obtain results about the density operator ρ_S.

However, to make the parallels to the pure state case as close as possible, I prefer to use an equivalent concept, called an *amplitude operator*. The equivalence between amplitude operators and purifications is discussed in Appendix A.

Definition 8.1 Given a density operator ρ_S on a Hilbert space $\mathcal{H}_S$, an *amplitude operator* for ρ_S is a linear operator $L_S : \mathcal{H}_E \to \mathcal{H}_S$, where $\mathcal{H}_E$ is any Hilbert space, such that

$$\rho_S = L_S L_S^\dagger.$$

The reason for the name *amplitude operator* is that, in pure-state quantum mechanics, an amplitude is a complex number α such that $|\alpha|^2$ is a probability. A density operator is a non-commutative generalization of a probability distribution [42,43], and hence an amplitude operator ought to be an operator that "squares" to a density operator.

Given a density operator ρ_S, one obvious way of constructing an amplitude operator is to set $\mathcal{H}_E = \mathcal{H}_S$ and $L_S = \sqrt{\rho}_S$, but there are an infinite number of alternatives, as the following proposition shows

Proposition 8.2 *An operator $L_S : \mathcal{H}_E \to \mathcal{H}_S$ is an amplitude operator for ρ_S if and only if*

$$L_S = \sqrt{\rho}_S U_{S|E},$$

where $U_{S|E} : \mathcal{H}_E \to \mathcal{H}_S$ is a semi-unitary operator, i.e. it satisfies $U_{S|E} U_{S|E}^\dagger = I_S$

Proof An operator of the form $L_S = \sqrt{\rho}_S U_{S|E}$ obviously satisfies Definition 8.1. For the other direction, assume L_S is an amplitude operator. Like any operator, it may be decomposed in its polar decomposition $L_S = P_S U_{S|E}$ where P_S is a positive semi-definite operator on $\mathcal{H}_S$, and $U_{S|E} : \mathcal{H}_E \to \mathcal{H}_S$ is semi-unitary.[4] The definition of an amplitude operator then implies that $\rho_S = P_S U_{S|E} U_{S|E}^\dagger P_S = P_S^2$, so we must have $P_S = \sqrt{\rho}_S$. □

Going back to the analogy between amplitudes and amplitude operators, multiplying an amplitude α by a phase factor $e^{i\phi}$ does not change the probability it represents. Similarly, multiplying an amplitude operator L_S by a semi-unitary $V_{E|E'}$, i.e., an operator $V_{E|E'} : \mathcal{H}_{E'} \to \mathcal{H}_E$ satisfying $V_{E|E'} V_{E|E'}^\dagger = I_E$, on the right does not change the

[4] The polar decomposition is often only defined for square matrices, in which case $\mathcal{H}_E = \mathcal{H}_S$ and $U_{S|E}$ is unitary. Here, we use the generalization to non-square matrices (see e.g., [44]).

density operator it represents. Although one might think it desirable to work directly with probabilities or density operators to eliminate these ambiguities, the mathematical manipulations we need to do in quantum mechanics are often linear in the amplitudes or amplitude operators, but would be nonlinear if you used probabilities or density operators. Therefore, it is often more convenient to live with the ambiguity.

Since *every* operator has a polar decomposition, the only requirement for L_S to be an amplitude operator for *some* density operator is that $\mathrm{Tr}_S\left(L_S L_S^\dagger\right) = 1$. If we want to work with unnormalized density operators, i.e., any positive operator, then any operator $L_S : \mathcal{H}_E \to \mathcal{H}_S$ is the amplitude operator for some (possibly unnormalized) density operator. This is analogous to the fact that any vector in $\mathcal{H}_S$ represents a (possibly unnormalized) pure state.

The strategy for generalizing the Aharonov–Vaidman identity, and everything that follows from it, is to replace the state vector $|\psi\rangle_S$ with an amplitude operator L_S. The reason this works is that the space of linear operators mapping $\mathcal{H}_E$ to $\mathcal{H}_S$, which we denote $\mathfrak{L}_{S|E}$, is itself a Hilbert space with inner product $\langle L_S, M_S\rangle = \mathrm{Tr}_E\left(L_S^\dagger M_S\right)$, known as the *Hilbert–Schmidt* inner product.[5] Since the Aharonov–Vaidman identity is valid for any Hilbert space, it must be valid on $\mathfrak{L}_{S|E}$ as well.

Proposition 8.3 (The Aharonov–Vaidman Identity for operators) *Let A_S be a linear operator on a Hilbert space $\mathcal{H}_S$ and let $L_S : \mathcal{H}_E \to \mathcal{H}_S$. Then,*

$$A_S L_S = \langle A_S\rangle\, L_S + (\Delta A_S)\, L_{A_S}^\perp, \tag{67}$$

where $\langle A_S\rangle = \mathrm{Tr}_S\left(A_S L_S L_S^\dagger\right)/\mathrm{Tr}_S\left(L_S L_S^\dagger\right)$, $\Delta A = \sqrt{\left\langle A_S^\dagger A_S\right\rangle - |\langle A_S\rangle|^2}$, and $L_{A_S}^\perp : \mathcal{H}_E \to \mathcal{H}_S$ is an amplitude operator that is orthogonal to L_S, i.e., $\mathrm{Tr}_E\left(L_S^\dagger L_{A_S}^\perp\right) = 0$, satisfies $\mathrm{Tr}_S\left(L_{A_S}^\perp L_{A_S}^{\perp\dagger}\right) = \mathrm{Tr}_S\left(L_S L_S^\dagger\right)$, and depends on both L_S and A_S.

The proof of this proposition is essentially the same as the proof of the vector Aharonov–Vaidman identity (Proposition 2.1) with the standard inner product replaced by the Hilbert–Schmidt inner product. The only difference is that the cyclic property of the trace also needs to be used to write things in the exact form given in Proposition 8.3. I leave this as an exercise for the reader.

Since $\rho_S = L_S L_S^\dagger$ is always a (possibly unnormalized) density operator, we can write

$$\langle A_S\rangle = \frac{\mathrm{Tr}_S\left(A_S L_S L_S^\dagger\right)}{\mathrm{Tr}_E\left(L_S^\dagger L_S\right)} = \frac{\mathrm{Tr}_S\left(A_S \rho_S\right)}{\mathrm{Tr}_E\left(\rho_S\right)}.$$

We can also introduce the density operator $\rho_{A_S}^\perp = L_{A_S}^\perp L_{A_S}^{\perp\dagger}$, which will be normalized in the same way as ρ_S, i.e., $\mathrm{Tr}_S\left(\rho_{A_S}^\perp\right) = \mathrm{Tr}_S\left(\rho_S\right)$.

When L_S is normalized so that $\rho_S = L_S L_S^\dagger$ is a normalized density operator, i.e., $\mathrm{Tr}_S\left(L_S L_S^\dagger\right) = 1$, then $\rho_{A_S}^\perp$ is also normalized, i.e., $\mathrm{Tr}_S\left(\rho_{A_S}^\perp\right) = 1$.

As defined, $\rho_{A_S}^\perp = L_{A_S}^\perp L_{A_S}^{\perp\dagger}$ looks like it depends on the choice of the amplitude operator L_S. In fact, it does not. It only depends on ρ_S and A_S. To see this, rewrite the operator Aharonov–Vaidman identity as

$$L_{A_S}^\perp = \frac{1}{\Delta A_S}\left(A_S - \langle A_S\rangle\, I_S\right) L_S,$$

and then we have

$$\rho_{A_S}^\perp = L_{A_S}^\perp L_{A_S}^{\perp\dagger}$$

[5] By the cyclic property of the trace, we can also write $\langle L_S, M_S\rangle = \mathrm{Tr}_S\left(M_S L_S^\dagger\right)$.

$$= \frac{1}{(\Delta A_S)^2} (A_S - \langle A_S \rangle I_S) L_S L_S^\dagger \left(A_S^\dagger - \langle A_S \rangle^* I_S \right)$$
$$= \frac{1}{(\Delta A_S)^2} (A_S - \langle A_S \rangle I_S) \rho_S \left(A_S^\dagger - \langle A_S \rangle^* I_S \right),$$

which is clearly independent of the choice of L_S. Note that, although L_S and $L_{A_S}^\perp$ are Hilbert–Schmidt orthogonal, ρ_S and $\rho_{A_S}^\perp$ are generally not.

To generalize the results of this paper from state vectors to density operators, we replace the vector Aharonov–Vaidman identity with its operator counterpart applied to amplitude operators, and we replace the usual inner product with the Hilbert–Schmidt inner product. In many cases, the final result is independent of the amplitude operator used to represent the state. Although we use it in the proof, it drops out in the final result by only appearing in the combination $L_S L_S^\dagger$, as in the expression we derived for $\rho_{A_S}^\perp$. In fact, the final formulas are usually the same as in the pure state case, except that we have to interpret $\langle A_S \rangle$ as $\mathrm{Tr}_S (A_S \rho_S)$ rather than $\langle \psi | A_S | \psi \rangle$.

However, this is not true for the Maccone–Pati uncertainty relations and their generalizations, which do depend on the choice of amplitude operator L_S.

Theorem 8.4 (The first Maccone–Pati uncertainty relation for amplitude operators) *Let A_S and B_S be Hermitian operators on a Hilbert space $\mathcal{H}_S$ and let ρ_S be a normalized density operator on $\mathcal{H}_S$. Then,*

$$(\Delta A)^2 + (\Delta B)^2 \geq \pm i \langle [A, B] \rangle + \left| \mathrm{Tr}_E \left(L_S^{\perp\dagger} (A \mp i B) L_S \right) \right|^2, \tag{68}$$

where $L_S : \mathcal{H}_E \to \mathcal{H}_S$ is any amplitude operator for ρ_S, and $L_S^\perp : \mathcal{H}_E \to \mathcal{H}_S$ is any normalized amplitude operator orthogonal to L_S that has the same input space $\mathcal{H}_E$.

Note that, to obtain the tightest possible bound on $(\Delta A)^2 + (\Delta B)^2$, the right hand side of (68) should be maximized over all possible choices of L_S and $L_S^\perp$. To do this in practice, a bound on the largest dimension d_E required to obtain the maximum is needed. I conjecture that $d_E = 2d_S$ is sufficient because this allows L_S and $L_S^\perp$ to have orthogonal kernels on $\mathcal{H}_E$, but I do not have a proof of this.

Theorem 8.5 (The second Maccone–Pati uncertainty relation for amplitude operators) *Let A_S and B_S be linear operators on a Hilbert space $\mathcal{H}_S$ and let ρ_S be a normalized density operator on $\mathcal{H}_S$. Then,*

$$(\Delta A_S)^2 + (\Delta B_S)^2 \geq \frac{1}{2} \left| \mathrm{Tr}_E \left(L_{A_S+B_S}^{\perp\dagger} (A + B) L_S \right) \right|^2, \tag{69}$$

where L_S is any amplitude operator for ρ_S and

$$L_{A_S+B_S}^\perp = \frac{1}{\Delta(A_S + B_S)} (A_S + B_S - \langle A_S + B_S \rangle I_S) L_S.$$

In this case, to obtain the tightest bound, we have to maximize the right hand side over L_S. We do not have to separately optimize over $L_{A_S+B_S}^\perp$ because it is a function of L_S, A_S and B_S. However, its dependence on L_S makes the problem into a complicated nonlinear optimization.

9 Summary and conclusions

In this paper, I discussed how the standard textbook uncertainty relations of Robertson and Schrödinger can be derived from the Aharonov–Vaidman identity in a more direct way than the standard proof. I also demonstrated the identity's usefulness in proving other uncertainty relations, such as the Maccone–Pati relations, and the quantum formulas for propagation of uncertainty. Finally, I gave a mixed-state generalization of the Aharonov–Vaidman

identity in terms of amplitude operators. I hope that this has persuaded you that the Aharonov–Vaidman identity belongs in undergraduate textbooks and that it ought to be a first-line tool in proving relationships between standard deviations in quantum mechanics. I am sure there are other uncertainty relations that have an elegant Aharonov–Vaidman based proofs, and I hope to find new and useful uncertainty relations that have not been discovered before via this method.

The Aharonov–Vaidman identity naturally gives rise to two quantum generalizations of the correlation, $\mathrm{corr}_{A,B}$ and $\mathrm{Rcorr}_{A,B}$. It would be interesting to determine whether these quantities have an operational meaning in the case where A and B do not commute. On the more formal side, perhaps there is a pseudo-probability representation of quantum mechanics, such as the Wigner function [45–47] or the Kirkwood-Dirac distribution [48–50], for which these are the correlations for observables as defined on the appropriate phase space. This might help to find uses for the propagation of error formulas in cases where the observables do not commute.

Acknowledgements I would like to thank Yakir Aharonov for introducing me to the Aharonov–Vaidman identity and emphasizing its importance. I would like to acknowledge (but not thank) the role played by the COVID-19 pandemic shutdowns in giving me the opportunity to think about uncertainty relations and their pedagogy. This research was supported in part by the Fetzer Franklin Fund of the John E. Fetzer Memorial Trust and by grant number FQXi-RFPIPW-1905 from the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Alley Community Foundation. This research was also supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science, and Economic Development, and by the Province of Ontario through the Ministry of Colleges and Universities.

Funding Information Open access funding provided by SCELC, Statewide California Electronic Library Consortium

Data Availability Statement Not applicable.

Declarations

Conflict of interest The author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: Amplitude operators and purifications

Proposition A.1 *Given a density operator ρ_S on a Hilbert space $\mathcal{H}_S$, let $\mathcal{H}_{S'}$ be another copy of the same Hilbert space and let $\{|j\rangle\}$ be an orthonormal basis for $\mathcal{H}_S$ and $\mathcal{H}_{S'}$. Define the vector*

$$\left|\Phi^+\right\rangle_{SS'} = \sum_j |j\rangle_S |j\rangle_{S'} .$$

Let $L_S : \mathcal{H}_E \to \mathcal{H}_S$ be an amplitude operator for ρ_S and let $\{|k\rangle_E\}$ be an orthonormal basis for $\mathcal{H}_E$. Then $I_S \otimes L_{S'}^T \left|\Phi^+\right\rangle_{SS'}$ is a purification of ρ_S, where ^{T} denotes transpose in the $|j\rangle\langle k|_{SE}$ basis. Similarly, if $|\psi\rangle_{SE} \in \mathcal{H}_S \otimes \mathcal{H}_E$ is a purification of ρ_S, then $L_S = \langle\psi^*|_{S'E} \left|\Phi^+\right\rangle_{SS'}$ is an amplitude operator for ρ_S, where * denotes complex conjugation in the $|jk\rangle_{S'E}$ basis.

Proof If L_S is an amplitude operator for ρ_S, then $\rho_S = L_S L_S^\dagger$. We have to show that this implies that $\mathrm{Tr}_E \left(I_S \otimes L_{S'}^T \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|_{SS'} I_S \otimes \left(L_{S'}^T\right)^\dagger \right) = \rho_S$. Note that $\left(L_{S'}^T\right)^\dagger = L_{S'}^*$, where * denotes complex conjugate

in the $|j\rangle\langle k|_{SE}$ basis. Therefore, we have

$$\mathrm{Tr}_E\left(I_S\otimes L_{S'}^T\,|\Phi^+\rangle\langle\Phi^+|_{SS'}\,I_S\otimes L_{S'}^*\right)=\sum_{j,k}|j\rangle\langle k|_S\,\mathrm{Tr}_E\left(L_{S'}^T\,|j\rangle\langle k|_{S'}\,L_{S'}^*\right) \tag{A1}$$

$$=\sum_{j,k}|j\rangle\langle k|_S\left\langle k|L_S^*L_S^T|j\right\rangle_S, \tag{A2}$$

where we have changed the index S' to S because they refer to the same Hilbert space and $\langle k|L_S^TL_S^*|j\rangle_S$ is a scalar. Rearranging this, we have

$$\mathrm{Tr}_E\left(I_S\otimes L_{S'}^T\,|\Phi^+\rangle\langle\Phi^+|_{SS'}\,I_S\otimes L_{S'}^*\right)=\sum_{j,k}|j\rangle_S\left\langle k|L_S^*L_S^T|j\right\rangle_S\langle k|_S \tag{A3}$$

$$=\sum_{j,k}|j\rangle\langle j|_S\left(L_S^*L_S^T\right)^T|k\rangle\langle k|_S \tag{A4}$$

$$=\left(L_S^*L_S^T\right)^T=L_SL_S^\dagger=\rho_S. \tag{A5}$$

For the other direction, we have to prove that $L_SL_S^\dagger=\rho_S$, where $L_S=\langle\psi^*|_{S'E}\,|\Phi^+\rangle_{SS'}$ and $|\psi\rangle_{SE}$ is any purification of ρ_S, i.e., $\mathrm{Tr}_E\left(|\psi\rangle\langle\psi|_{SE}\right)=\rho_S$.

First, let $|\psi\rangle_{SE}=\sum_{jk}\alpha_{jk}\,|j\rangle_S\otimes|k\rangle_E$ be the decomposition of $|\psi\rangle_{SE}$ in the $|jk\rangle_{SE}$ basis. We have $|\psi^*\rangle_{SE}=\sum_{jk}\alpha_{jk}^*\,|j\rangle_S\otimes|k\rangle_E$ and the condition $\mathrm{Tr}_E\left(|\psi\rangle\langle\psi|_{SE}\right)=\rho_S$ is equivalent to $\sum_{j,k,l}\alpha_{jk}\alpha_{lk}^*\,|j\rangle\langle l|_S=\rho_S$. Note also that $\langle j|_{S'}\,|\Phi^+\rangle_{SS'}=|j\rangle_S$.

Hence, we have

$$L_SL_S^\dagger=\left(\langle\psi^*|_{S'E}\,|\Phi^+\rangle_{SS'}\right)\left(\langle\psi^*|_{S'E}\,|\Phi^+\rangle_{SS'}\right)^\dagger \tag{A6}$$

$$=\langle\psi^*|_{S'E}\,|\Phi^+\rangle_{SS'}\,\langle\Phi^+|_{SS'}\,|\psi^*\rangle_{S'E} \tag{A7}$$

$$=\sum_{jklm}\alpha_{jk}\,\langle j|_{S'}\,\langle k|_{E'}\,|\Phi^+\rangle\langle\Phi^+|_{SS'}\,\alpha_{lm}^*\,|l\rangle_{S'}\,|m\rangle_E \tag{A8}$$

$$=\sum_{jklm}\alpha_{jk}\alpha_{lm}^*\langle k|m\rangle_E\left(\langle j|_{S'}\,|\Phi^+\rangle_{SS'}\right)\left(\langle\Phi^+|_{SS'}\,|l\rangle_{S'}\right) \tag{A9}$$

$$=\sum_{jkl}\alpha_{jk}\alpha_{lk}^*\,|j\rangle\langle l|_S \tag{A10}$$

$$=\rho_S. \tag{A11}$$

□

References

1. Aharonov, Y., Vaidman, L.: Properties of a quantum system during the time interval between two measurements. Phys. Rev. A **41**(1), 11–20 (1990). https://doi.org/10.1103/PhysRevA.41.11
2. Aharonov, Y.: Visiting Researcher Presentation. Talk at Perimeter Institute to PSI Masters Students: comment is made at 41:16 (2011). https://pirsa.org/11080091
3. Vaidman, L.: Minimum time for the evolution to an orthogonal state. Am. J. Phys. **60**(2), 182 (1992). https://doi.org/10.1119/1.16940
4. Robertson, H.P.: The uncertainty principle. Phys. Rev. **34**(1), 163–164 (1929). https://doi.org/10.1103/PhysRev.34.163

5. Schrödinger, E.: Zum Heisenbergschen unschärfeprinzip. Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse **14**, 296–303 (1930)
6. Goldenberg, L., Vaidman, L.: Applications of a simple quantum mechanical formula. Am. J. Phys. **64**(8), 1059 (1996). https://doi.org/10.1119/1.18307. arXiv:quant-ph/9506030
7. Aharonov, Y., Rohrlich, D.: Quantum paradoxes: quantum theory for the perplexed. Wiley (2005). https://doi.org/10.1002/9783527619115
8. Steele, M.J.: 1. The Cauchy–Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities. Mathematical Association of America Problem Books, p. 1. Cambridge University Press (2004)
9. Fuchs, C.A., Peres, A.: Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. Phys. Rev. A **53**(4), 2038 (1996). https://doi.org/10.1103/PhysRevA.53.2038. arXiv:quant-ph/9512023
10. Hofmann, H.F., Takeuchi, S.: Violation of local uncertainty relations as a signature of entanglement. Phys. Rev. A **68**(3), 032103 (2003). https://doi.org/10.1103/PhysRevA.68.032103. arXiv:quant-ph/0212090
11. Gühne, O.: Characterizing entanglement via uncertainty relations. Phys. Rev. Lett. **92**(11), 117903 (2004). https://doi.org/10.1103/PhysRevLett.92.117903. arXiv:quant-ph/0306194
12. Koashi, M.: Unconditional security of quantum key distribution and the uncertainty principle. J. Phys. Conf. Ser. **36**, 98–102 (2006). https://doi.org/10.1088/1742-6596/36/1/016. arXiv:quant-ph/0505108
13. Berta, M., Christandl, M., Colbeck, R., Renes, J.M., Renner, R.: The uncertainty principle in the presence of quantum memory. Nat. Phys. **6**, 659–662 (2010). https://doi.org/10.1038/nphys1734. arXiv:0909.0950
14. Majumdar, A.S., Pramanik, T.: Some applications of uncertainty relations in quantum information. Int. J. Quantum Inf. **14**(6), 1640022 (2016). https://doi.org/10.1142/S0219749916400220. arXiv:1410.5974
15. Yunger Halpern, N., Bartolotta, A., Pollack, J.: Entropic uncertainty relations for quantum information scrambling. Commun. Phys. **2**, 92 (2019). https://doi.org/10.1038/s42005-019-0179-8. arXiv:1806.04147
16. Oppenheim, J., Wehner, S.: The uncertainty principle determines the nonlocality of quantum mechanics. Science **330**(6007), 1072–1074 (2010). https://doi.org/10.1126/science.119206. arXiv:1004.2507
17. Catani, L., Leifer, M., Scala, G., Schmid, D., Spekkens, R.W.: What is nonclassical about uncertainty relations? Phys. Rev. Lett. **129**(24), 240401 (2022). https://doi.org/10.1103/PhysRevLett.129.240401. arXiv:2207.11779
18. Coles, P.J., Berta, M., Tomamichel, M., Wehner, S.: Entropic uncertainty relations and their applications. Rev. Mod. Phys. **89**(1), 015002 (2017). https://doi.org/10.1103/RevModPhys.89.015002. arXiv:1511.04857
19. Pati, A.K., Sahu, P.K.: Sum uncertainty relation in quantum theory. Phys. Lett. A **367**(3), 177–181 (2007). https://doi.org/10.1016/j.physleta.2007.03.005. arXiv:quant-ph/0608092
20. Maccone, L., Pati, A.K.: Stronger uncertainty relations for all incompatible observables. Phys. Rev. Lett. **113**(26), 039902 (2015). https://doi.org/10.1103/PhysRevLett.113.260401. arXiv:1407.0338
21. Bannur, V.M.: General and stronger uncertainty relation (2015). https://doi.org/10.48550/arXiv.1503.00405
22. Li, J.L., Qiao, C.F.: Reformulating the quantum uncertainty relation. Sci. Rep. **5**, 12708 (2015). https://doi.org/10.1038/srep12708. arXiv:1502.06292
23. Yao, Y., Xiao, X., Wang, X., Sun, C.P.: Implications and applications of the variance-based uncertainty equalities. Phys. Rev. A **91**, 062113 (2015). https://doi.org/10.1103/PhysRevA.91.062113. arXiv:1503.00239
24. Abbott, A.A., Alzieu, P., Hall, M.J.W., Branciard, C.: Tight state-independent uncertainty relations for qubits. Mathematics **4**(1), 8 (2016). https://doi.org/10.3390/math4010008. arXiv:1512.02383
25. Chen, B., Cao, N., Fei, S., Long, G.: Variance-based uncertainty relations for incompatible observables. Quantum Inf. Process. **15**, 3909–3917 (2016). https://doi.org/10.1007/s11128-016-1365-1. arXiv:1608.06075
26. Qin, H., Fei, S., Li-Jost, X.: Multi-observable uncertainty relations in product form of variances. Sci. Rep. **6**, 31192 (2016). https://doi.org/10.1038/srep31192. arXiv:1608.03089
27. Song, Q., Qiao, C.: Stronger Schrödinger-like uncertainty relations. Phys. Lett. A **380**(37), 2925–2930 (2016). https://doi.org/10.1016/j.physleta.2016.06.054. arXiv:1504.01137
28. Xiao, Y., Jing, N., Li-Jost, X., Fei, S.: Weighted uncertainty relations. Sci. Rep. **6**, 23201 (2016). https://doi.org/10.1038/srep23201. arXiv:1603.01004
29. Mondal, D., Bagchi, S., Pati, A.K.: Tighter uncertainty and reverse uncertainty relations. Phys. Rev. A **95**(5), 052117 (2017). https://doi.org/10.1103/PhysRevA.95.052117. arXiv:1607.06712
30. Song, Q., Li, G., Peng, J., Qiao, C.: A stronger multi-observable uncertainty relation. Sci. Rep. **7**, 44764 (2017). https://doi.org/10.1038/srep44764. arXiv:1701.01072
31. Zhang, J., Zhang, Y., Yu, C.: Stronger uncertainty relations with improvable upper and lower bounds. Quantum Inf. Process. **16**, 131 (2017). https://doi.org/10.1007/s11128-017-1585-z. arXiv:1607.08223
32. Zheng, X., Zhang, G.: Variance-based uncertainty relation for incompatible observers. Quantum Inf. Process. **16**, 167 (2017). https://doi.org/10.1007/s11128-017-1619-6. arXiv:1705.07396
33. Dodonov, V.V.: Variance uncertainty relations without covariances for three and four observables. Phys. Rev. A **97**(2), 022105 (2018). https://doi.org/10.1103/PhysRevA.97.022105. arXiv:1711.04037
34. de Guise, H., Maccone, L., Sanders, B.C., Shukla, N.: State-independent uncertainty relations. Phys. Rev. A **98**(4), 042121 (2018). https://doi.org/10.1103/PhysRevA.98.042121. arXiv:1804.06794
35. Busch, P., Reaerdon-Smith, O.: On quantum uncertainty relations and uncertainty regions (2019). https://doi.org/10.48550/arXiv.1901.03695

36. Giorda, P., Maccone, L., Riccardi, A.: State-independent uncertainty relations from eigenvalue minimization. Phys. Rev. A **99**(5), 052121 (2019). https://doi.org/10.1103/PhysRevA.99.052121. arXiv:1810.09775
37. Zheng, X., Ma, S., Zhang, G., Fan, H., Liu, W.: Unified and exact framework for variance-based uncertainty relations. Sci. Rep. **10**, 150 (2020). https://doi.org/10.1038/s41598-019-56803-2. arXiv:1803.08720
38. Li, J., Qiao, C.: The generalized uncertainty principle. Ann. Phys. **533**(1), 2000335 (2021). https://doi.org/10.1002/andp.202000335. arXiv:2003.08705
39. Zhang, L., Luo, S., Fei, S., Wu, J.: Uncertainty regions of observables and state-independent uncertainty relations. Quantum Inf. Process. **20**, 357 (2021). https://doi.org/10.1007/s11128-021-03303-w. arXiv:2110.14134
40. Chiew, S., Gessner, M.: Improving sum uncertainty relations with the quantum Fisher information. Phys. Rev. Res. **4**(1), 013076 (2022). https://doi.org/10.1103/PhysRevResearch.4.013076. arXiv:2109.06900
41. Xiao, Y., Jing, N., Yu, B., Fei, S., Li-Jost, X.: Near-optimal variance-based uncertainty relations. Front. Phys. **10**, 846330 (2022). https://doi.org/10.3389/fphy.2022.846330. arXiv:1610.01692
42. Rédei, M., Summers, S.J.: Quantum probability theory. Stud. Hist. Philos. Mod. Phys. **38**(2), 390–417 (2007). https://doi.org/10.1016/j.shpsb.2006.05.006. arXiv:quant-ph/0601158
43. Leifer, M.S., Spekkens, R.W.: Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. Phys. Rev. A **88**(5), 052130 (2013). https://doi.org/10.1103/PhysRevA.88.052130. arXiv:1107.5849
44. Ben-Israel, A., Greville, T.N.: Chapter 6. Generalized Inverses: Theory and Applications. CMS Books in Mathematics, 2nd edn, pp. 220–221. Springer (2003). https://doi.org/10.1007/b97366
45. Wigner, E.: On the quantum correction for thermodynamic equilibrium. Phys. Rev. **40**(5), 749–759 (1932). https://doi.org/10.1103/PhysRev.40.749
46. Gross, D.: Hudson's theorem for finite-dimensional quantum systems. J. Math. Phys. **47**, 122107 (2006). https://doi.org/10.1063/1.2393152. arXiv:quant-ph/0602001
47. Curtright, T.L., Fairlie, D.B., Zachos, C.K.: A Concise Treatise on Quantum Mechanics in Phase Space. World Scientific (2014). https://doi.org/10.1142/8870
48. Kirkwood, J.G.: Quantum statistics of almost classical assemblies. Phys. Rev. **44**(1), 31–37 (1933). https://doi.org/10.1103/PhysRev.44.31
49. Dirac, P.A.M.: On the analogy between classical and quantum mechanics. Rev. Mod. Phys. **17**(2–3), 195–199 (1945). https://doi.org/10.1103/RevModPhys.17.195
50. Lostaglio, M., Belenchia, A., Levy, A., Hernández-Gómez, S., Fabbri, N., Gherardini, S.: Kirkwood-Dirac quasiprobability approach to quantum fluctuations: theoretical and experimental perspectives (2022). arXiv:2206.11783

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.