

Inaccurate weak measurements classical and quantum

D. Sokolovski ^{1,2,3,*} D. Alonso ⁴ and S. Brouard ⁴

¹*Departamento de Química-Física, Universidad del País Vasco, UPV/EHU, 48940 Leioa, Spain*

²*IKERBASQUE, Basque Foundation for Science, Plaza Euskadi 5, 48009 Bilbao, Spain*

³*EHU Quantum Center, Universidad del País Vasco, UPV/EHU, 48940 Leioa, Spain*

⁴*Departamento de Física y IUEA, Universidad de La Laguna, La Laguna, Tenerife, Spain*



(Received 30 March 2026; revised 9 July 2026; accepted 12 August 2026;
published 24 August 2026)

We consider highly inaccurate measurements made on classical stochastic and quantum systems. In the quantum case, such a “weak” measurement preserves coherence between the system’s alternatives. We demonstrate that in both classically and quantumly the information about the scenario realized in each individual trial is lost. However, ensemble parameters such as classical path probabilities and quantum quasiprobabilities can be extracted from the obtained statistics. Both classically and quantumly, causality ensures that additional postselection only redistributes individual outcomes between the system’s final states. Quantum quasiprobabilities may change sign, which allows for anomalously large meter’s (pointer’s) reading for some final states. These, we show, result from mere “re-shaping” of a broad distribution obtained earlier and provide no “experimental evidence” of quantum variables taking, albeit on rare occasions, exceptionally large or otherwise anomalous values.

DOI: [10.1103/t6gz-b6vt](https://doi.org/10.1103/t6gz-b6vt)

I. INTRODUCTION

The problem of establishing intermediate condition of a quantum system making a transition between given initial and final states has been a source of contradiction for several decades (see, e.g., Ref. [1]). In order to alter the transition probability as little as possible, one can try to use a highly inaccurate pointer, such that the uncertainty in its initial position greatly exceeds the displacements it may undergo in the course of the measurement.

The difficulty with such an approach is well known in the text book quantum theory of Feynman *et al.* [2]. The system can be seen as reaching its final state via virtual (Feynman) paths alongside which the quantity of interest takes definite values. An inaccurate pointer is designed to leave interference between them intact. The uncertainty principle [2] forbids knowing the path chosen by the system and, by extension, the value of the variable has taken. Among the well-known quotes on the subject are Bohm’s mention of “absurd results,” which can occur, should interference not be destroyed [3], and Feynman’s “blind alley” warning against rationalizing, in classical-like terms, the manner in which a quantum system travels interfering routes [4].

It must be noted that a different approach has been proposed in Ref. [5]. In the conventional narrative, quantum mechanics calculates the transition amplitudes as matrix elements of unitary

*Contact author: dgsokol15@gmail.com

operators between the states of interest, so one does not have to worry about the collapse of the wave function at the end of the experiment. The authors of Ref. [5] literally doubled down on the role of the wave function, by adding to the description the backward-evolving final state. The method was followed on under the name time-symmetric quantum mechanics (TSQM) (see, e.g., Ref. [6]). Interpretational value (no new results are predicted by the TSQM [5]) of this picture was questioned in Ref. [7]. More recent critique was given in Ref. [8], where the reader will also find further relevant references from the extensive literature existing on the subject. We share the concerns of Robertson [7,8], and will henceforth remain within the conventional framework of Feynman *et al.* [2] outlined in more detail in Ref. [9].

Having thus stated our preference, we proceed to the purpose of this paper. Consider a pointer employed to measure a variable $\hat{B}$ of a quantum system, previously prepared (preselected) and later found (postselected) in the known initial and final states. If the initial (Gaussian) distribution of the pointer's positions is sufficiently broad, the final distribution of its readings is given by a similar Gaussian, displaced as a whole by the real part of a complex quantity known as “the weak value (WV) of $\hat{B}$ ” [1,10]. It is easy to show that one can always find a suitable pair of the system's initial and final states so as to give the WV of $\hat{B}$ and, therefore, to the pointer's shift any desired value [9]. Earlier, we related the controversies surrounding the weak values to the predictive power erroneously attributed to quantum probability amplitudes (see, e.g., Ref. [11]). Here, we consider the problem from a different perspective and will address two particular questions.

(1) Can the shift of an inaccurate pointer be interpreted as a unique well-defined value of $\hat{B}$, occurring each time the system makes the transition? The question is not new. A historic perspective on the single system vs ensemble property controversy can be found in Ref. [8].

(2) For an unlikely transition, the shift can be very large. Does this mean that, on rare occasions, quantum mechanics allows a variable to take anomalously large, or otherwise unusual, values? The authors of Ref. [1] exercise caution and talk about the result of measuring spin-1/2 “turning out to be 100.” Reference [6] speaks about the discovery of “new quantum effects” related to this large value. Plainly, more clarity is desirable, and we will seek it in what follows.

In doing so, it will be useful to consider first a classical statistical model, in order to focus on the essentially quantum aspects of the problem later. The lack of clarity about the origin and status of quantum WV has lead in the past the authors of Ref. [12] to an erroneous [13] claim that anomalous WVs may exist in a purely classical environment. We will revisit this particular controversy in due course.

The rest of the paper is organized as follows. In Sec. II, we will review a useful mathematical property. Sections III and IV discuss highly inaccurate classical measurements with and without postselection, one's ability to reconstruct the path probabilities, and the classical causality principle. Sections V and VI contain a similar analysis of inaccurate (weak) quantum measurements. In Sec. VII, we ascribe quasiprobabilities to the system's virtual paths. Section VII contains a simple example. Section IX briefly discusses the meaning and use of negative probabilities in quantum theory. In Sec. X, we analyze the reshaping mechanism responsible for the anomalous readings obtained in the quantum case and Sec. XI gives another simple example. Our conclusions and discussion are in Sec. XII.

II. AN ASYMPTOTIC PROPERTY

We begin with a curious mathematical property. For a smooth real function $G(x)$, such that $-\infty < \partial_x G(x) < \infty$, the following limit is reached as the scaling parameter Δx becomes large compared to the largest of $|B_i|$ (we will write $\Delta x \rightarrow \infty$ to simplify the notations):

$$\sum_{i=1}^N A_i G\left(\frac{x - B_i}{\Delta x}\right) \xrightarrow{\Delta x \rightarrow \infty} \left[\sum_{i=1}^N A_i \right] G\left(\frac{x - z}{\Delta x}\right), \quad (1)$$

$$z \equiv \frac{\sum_{i=1}^N A_i B_i}{\sum_{i=1}^N A_i}, \quad (2)$$

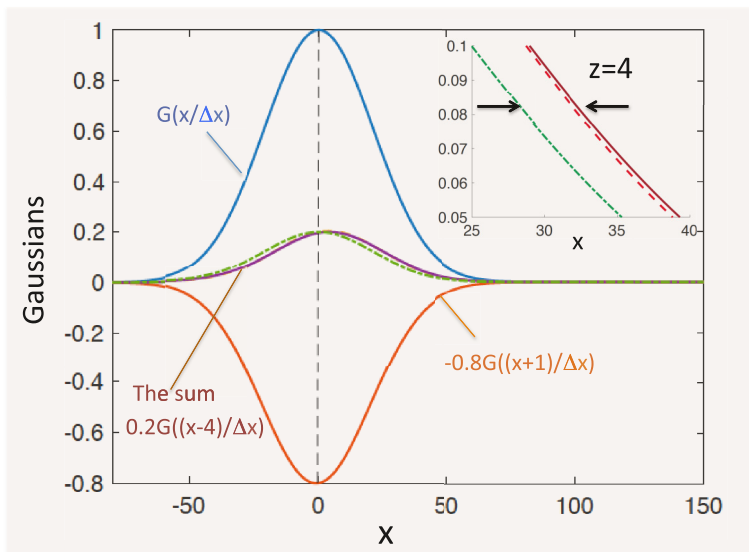


FIG. 1. Two broad Gaussians (5) in the left-hand side of Eq. (1), $\Delta x = 30$, $A_1 = 1$, $B_1 = 0$, and $A_2 = -0.8$, $B_2 = -1$, add up to a smaller Gaussian shifted to the right by approximately 4 units. The dot-dashed line shows the same Gaussian centered at $x = 0$. In the inset, the added dashed line shows, for comparison, the Gaussian in the right-hand side of Eq. (1).

where A_i and B_i , $i = 1, \dots, N$, are real, possibly negative numbers (see Appendix A). The convergence is pointwise [14], as illustrated in Fig. 1.

It follows further that for complex valued A_i the absolute square of the sum in Eq. (1) tends to a similar limit (see Appendix A)

$$\left| \sum_{i=1}^N A_i G\left(\frac{x - B_i}{\Delta x}\right) \right|^2 \xrightarrow{\Delta x \rightarrow \infty} \left| \sum_{i=1}^N A_i \right|^2 G^2\left(\frac{x - Z}{\Delta x}\right), \quad (3)$$

where

$$Z \equiv \text{Re} \left[\frac{\sum_{i=1}^N A_i B_i}{\sum_{i=1}^N A_i} \right]. \quad (4)$$

To our knowledge, these limits were first reported in Ref. [1] for the Gaussians,

$$G(x) = \exp\left(-\frac{x^2}{\Delta x^2}\right). \quad (5)$$

An early critique of Ref. [1] suggested that whenever two Gaussians are added, it should always be possible to recognize in the result the presence of the two constituent peaks [15]. This, however, does not seem to be the case (cf. Fig. 1; for relation to the Catastrophe Theory [16], see Ref. [14]). Next, we analyze the importance of Eqs. (1)–(4) for highly inaccurate measurements made on both classical and quantum systems.

III. TWO-POINT CLASSICAL MEASUREMENTS

We start with the classical case. Consider a simple setup where a classical particle (one can think, if it helps, of a little ball rolling down a system of tubes or slides), introduced at I into the system of routes shown in Fig. 2(a), is directed to the i th node, $i = 1, 2, \dots, N$ with probability $P(i \leftarrow I)$.

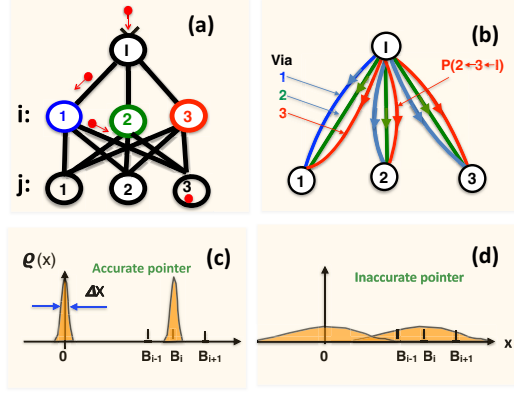


FIG. 2. (a) Alice’s classical setup ($N = 3$). The particle passes through a state i on the way to its final state j . (b) Each of the nine paths available to the particle is equipped with the probability in Eq. (6). (c) Initial uncertainty of an accurate pointer’s position, Δx , is small compared to the differences $\max|B_i|$. (d) For an inaccurate pointer, one has $\Delta x > \max|B_i|$.

From there, it is sent to a receptacle (final state) j , $j = 1, 2, \dots, N$ with a probability $P(j \leftarrow i)$. Thus, N paths connect the initial state I with a final state j [see Fig. 2(b)]. A path $j \leftarrow i \leftarrow I$ is traveled with probability

$$P(j \leftarrow i \leftarrow I) = P(j \leftarrow i)P(i \leftarrow I), \quad (6)$$

so the net probability to reach the state j is

$$P(j \leftarrow I) = \sum_{i=1}^N P(j \leftarrow i \leftarrow I). \quad (7)$$

Alice-the-experimenter (**A**) wants to measure a variable (a “property” [17]) B that takes the value B_i if the particle crosses the i th node. For this purpose, she uses a pointer with position x_1 that moves by B_i units whenever the i th intermediate state is crossed. The pointer’s initial position, not known exactly, is distributed around zero according to

$$\rho^I(x_1) = \pi^{-1/2} \Delta x_1^{-1} \exp(-x_1^2 / \Delta x_1^2). \quad (8)$$

The distribution of its final positions (readings), which can be determined accurately, is, therefore, given by

$$\rho(x_1) = \sum_{i=1}^N P(i \leftarrow I) \rho^I(x_1 - B_i). \quad (9)$$

If the pointer is accurate, $\Delta x_1 \rightarrow 0$, $\rho(x_1) = \sum_{j=1}^N P(i \leftarrow I) \delta(x_1 - B_i)$, where $\delta(x)$ is the Dirac delta [see Fig. 2(c)]. Thus, with all B_i ’s distinct, **A** always knows the ball’s final position.

We are, however, mostly interested in the case where **A** has at her disposal only a highly inaccurate pointer with $\Delta x_1 \gg \max|B_i|$ (we will write simply $\Delta x_1 \rightarrow \infty$). If so, by Eq. (1), the distribution (9) is a single broad Gaussian

$$\rho(x_1) \xrightarrow{\Delta x_1 \rightarrow \infty} \rho^I(x_1 - y(B)), \quad y(B) \equiv \sum_{i=1}^N B_i P(i \leftarrow I). \quad (10)$$

With all the Gaussians $\rho^I(x - B_i)$ in Eq. (9) overlapping, **A** cannot know the particle’s location. However, after many trials, she can evaluate the average value of the variable B , $y(B)$. If $B \equiv \pi_{i_0}$

probes the particle's presence in the i_0 th node, $B_i = \delta_{i i_0}$, $\mathbf{A}$ estimates, in this manner, the probability $P(i_0 \leftarrow I)$.

Finally, Eq. (10) may look as if a measurement has been made on a system where B has a unique well-defined value $y(B)$. This is, of course, not the case, since $y(B)$ is an ensemble average, whose evaluation requires many trials.

IV. INACCURATE CLASSICAL MEASUREMENTS WITH “POSTSELECTION”

Alice can extend her experiment and use a pointer with position x_2 to measure a variable F , which takes a value F_j if the particle ends up in its j th final state. The final distribution of the pointers' readings is now given by

$$\rho(x_1, x_2) = \sum_{i,j=1}^N P(j \leftarrow i \leftarrow I) \rho_1^I(x_1 - B_i) \rho_2^I(x_2 - F_j), \quad (11)$$

where $\rho_k^I(x_k) = \pi^{-1/2} \Delta x_k^{-1} \exp(-x_k^2 / \Delta x_k^2)$, $k = 1, 2$. Again, with both pointers accurate, $\Delta x_1, \Delta x_2 \rightarrow 0$, $\rho_{1,2}^I(x_{1,2})$ tend to $\delta(x_{1,2})$ and any pair of readings $(x_1 \approx B_i, x_2 \approx F_j)$ allow $\mathbf{A}$ to identify the path $j \leftarrow i \leftarrow I$ traveled in the trial.

As before, we are interested in the case $\Delta x_1 \rightarrow \infty$, $\Delta x_2 \rightarrow 0$, whereby, in a given trial, the detailed route remains unknown. By Eq. (1), we have

$$\begin{aligned} \rho(x_1, x_2) &\xrightarrow{\Delta x_1 \rightarrow \infty, \Delta x_2 \rightarrow 0} \sum_{j=1}^N P(j \leftarrow I) \rho_1^I(x_1 - z_j(B)) \delta(x_2 - F_j), \\ z_j(B) &\equiv \frac{\sum_i B_i P(j \leftarrow i \leftarrow I)}{\sum_i P(j \leftarrow i \leftarrow I)}. \end{aligned} \quad (12)$$

Alice may choose to select only the cases where $x_2 \approx F_j$ and the particle is “postselected” in the final state j . If so, the shift (position of the peak, or of the centroid) of the broad Gaussian in Eq. (12) coincides with the mean value of B , $z_j(B)$, calculated with the probabilities (6) of the paths ending in the j th final state. For the special choice $B_i = \pi_{i_0}$,

$$z_j(\pi_{i_0}) = \frac{P(j \leftarrow i_0 \leftarrow I)}{P(j \leftarrow I)} \quad (13)$$

yields the probability of passing through i_0 on the way to the final state j .

We note (this will be important for future discussion of the quantum case) that Alice, who can determine both $z_j(\pi_i)$ [by measuring the shift of the Gaussian in Eq. (12)] and $P(j \leftarrow I)$ (by counting the arrivals in j), is able to recover the probabilities $P(j \leftarrow i \leftarrow I)$ for all paths in Fig. 2(b),

$$P(j \leftarrow i \leftarrow I) = z_j(\pi_i) P(j \leftarrow I). \quad (14)$$

Finally, since $\sum_j P(j \leftarrow I) = 1$, integration of Eq. (11) over dx_2 recovers the distribution (9) obtained in the experiment in which the future state of the system is not known, or not controlled,

$$\int \rho(x_1, x_2) dx_2 = \sum_{i=1}^N P(i \leftarrow I) \rho_1^I(x_1 - B_i). \quad (15)$$

Equation (15) illustrates the causality principle, built into classical statistics. Summation over the final readings *must* recover the distribution (9) obtained in the shorter measurement of the previous Section. Were this not the case, $\mathbf{A}$'s future decision to postselect could be known in advance, and this should not be possible.

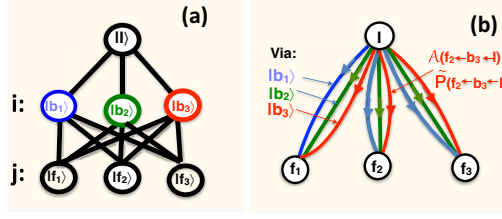


FIG. 3. (a) Alice's quantum setup ($N = 3$). The system passes through a state $|b_i\rangle$ on the way to its final state $|f_j\rangle$. (b) Each of the nine virtual paths available to the system is equipped with the amplitude in Eq. (16). Quasiprobabilities (32) can also be ascribed without destroying coherence between the paths (see Sec. VI).

V. TWO-POINT QUANTUM MEASUREMENTS

While classical statistic requires no introduction, the purpose of quantum theory needs to be stated with sufficient clarity. We take the main points from Ref. [2] (for more detail, see also Ref. [9]).

- (i) For each possible scenario, the theory provides a complex valued probability amplitude, whose absolute square gives the probability of the scenario to occur.
- (ii) If two scenarios leading to the same outcome cannot be distinguished, their amplitudes must be added. Otherwise, one must add the corresponding probabilities.
- (iii) Amplitudes of scenarios leading to different final outcomes are never added.

Consider a quantum system, whose distinguishable conditions are in correspondence with orthogonal states in an N -dimensional Hilbert space. For such a system, a possible scenario is to start in a state $|I\rangle$ at $t = 0$, be in one of the states $|b_i\rangle$, $\langle b_i|b_{i'}\rangle = \delta_{ii}$, $i = 1, \dots, N$, at $t_1 > 0$, and end up in one of the orthonormal states $|f_j\rangle$, $j = 1, \dots, N$ at $t_2 > t_1$ (see Fig. 3). For each scenario, the theory provides a probability amplitude (we use $\hbar = 1$)

$$\begin{aligned} \mathcal{A}^S(f_j \leftarrow b_i \leftarrow I) &= \mathcal{A}^S(f_j \leftarrow b_i) \mathcal{A}^S(b_i \leftarrow I) \\ &\equiv \langle f_j | \hat{U}^S(t_2 - t_1) | b_i \rangle \langle b_i | \hat{U}^S(t_1) | I \rangle, \end{aligned} \quad (16)$$

where $\hat{U}^S(t) \equiv \exp(-i\hat{H}^S t)$ and $\hat{H}^S$ is the system's Hamiltonian, chosen, for simplicity, to be time independent.

The simplest question **A** may ask is “What is the probability of a system prepared in $|I\rangle$ to be found in $|b_i\rangle$ at $t = t_1$?” By (i), there are N scenarios (paths) $b_i \leftarrow I$, which lead to distinguishable final outcomes. By (iii), the paths can be endowed with probabilities,

$$P^S(b_i \leftarrow I) = |\mathcal{A}^S(b_i \leftarrow I)|^2, \quad i = 1, 2, \dots, N. \quad (17)$$

As in the classical case, the existence of the transition probabilities (17), which follows from the rules (i)–(iii), does not seem to depend on the presence of a measuring device.

However, Alice, unable to observe the system directly, may need to rely on an instrument, accessible by her senses. Thus, she employs a von Neumann pointer (P) [18] whose position, x_1 , can be determined accurately just after it couples to the system at $t = t_1$. (Technically speaking, she also needs another pointer to record the system's initial state $|I\rangle$, but there is no need to mention it explicitly.) The pointer, designed to measure an operator

$$\hat{B} = \sum_{i=1}^N |b_i\rangle B_i \langle b_i|, \quad (18)$$

is prepared in a Gaussian state $|G_1\rangle$,

$$\langle x_1 | G_1 \rangle = \pi^{-1/4} \Delta x_1^{-1/2} \exp\left(-\frac{x_1^2}{2\Delta x_1^2}\right) \equiv G_1(x_1), \quad (19)$$

so its initial distribution,

$$\rho_1^I(x_1) = |G_1(x_1)|^2, \quad (20)$$

is a Gaussian of width Δx_1 . Now, **A** has to apply the rules (i)–(iii) to the composite system + pointer. The composite can make transitions from $|I\rangle|G_1\rangle$ to one of the distinguishable (orthogonal) states $|b_i\rangle|x_1\rangle$. Such is the property of the von Neumann pointer (see Appendix B) that the corresponding amplitude is given by

$$\mathcal{A}^{S+P}(b_i, x_1 \leftarrow I, G) = \mathcal{A}^S(b_i \leftarrow I)G(x_1 - B_i). \quad (21)$$

Again, by (iii), one can assign probabilities $P_i^{S+P}(x_1) = |\mathcal{A}^{S+P}(b_i, x_1 \leftarrow I, G_1)|^2$. To find the observed distribution of the pointer's readings, **A** needs to sum $P_i^{S+P}(x_1)$ over unobserved system's states $|b_i\rangle$. This yields expressions remarkably similar to its classical analog, Eq. (9),

$$\rho(x_1, t_1) = \sum_{i=1}^N P^S(b_i \leftarrow I) \rho^I(x_1 - B_i), \quad (22)$$

and, by Eq. (1),

$$\rho(x_1, t_1) \xrightarrow{\Delta x_1 \rightarrow \infty} \rho^I(x_1 - Y(B)), \quad (23)$$

where

$$Y(B) \equiv \sum_{i=1}^N B_i P^S(b_i \leftarrow I) = \langle I(t_1) | \hat{B} | I(t_1) \rangle \quad (24)$$

and $|I(t_1)\rangle \equiv \hat{U}^S(t_1)|I\rangle$. Thus, $Y(B)$ is the usual average of the operator $\hat{B}$ in the evolved initial state $|I(t)\rangle$. The rest of the discussion can follow that of Sec. III. One notes the consistency of the Feynman's rules (i)–(iii). Whenever the theory is able to predict probabilities for the system in isolation [cf. Eq. (17)], the same probabilities appear when the rules are applied to the composite system + measuring device [cf. Eq. (22)].

VI. QUANTUM MEASUREMENTS WITH POSTSELECTION

The principal difference between the classical and quantum cases becomes evident when one considers scenarios where a quantum system arrives in a final state $|f_j\rangle$ via an intermediate state $|b_i\rangle$. Whereas classical probabilities (6) are always available, quantum theory predicates their existence on one's ability to “distinguish” between the alternatives.

In practical terms, Alice, interested in the system making a three-point transition $|f_j\rangle \leftarrow |b_i\rangle \leftarrow |I\rangle$, needs to employ two von Neumann pointers (P_1 and P_2), designed to measure the system's operators

$$\hat{B} = \sum_{i=1}^N |b_i\rangle B_i \langle b_i| \quad \text{and} \quad \hat{F} = \sum_{j=1}^N |f_j\rangle F_j \langle f_j|, \quad (25)$$

at $t = t_1$ and $t = t_2$, respectively. The pointers are prepared in the Gaussian states [cf. Eq. (19)].

Now, the amplitude of the composite $S + P_1 + P_2$ making a two-point transition $|f_j\rangle|x_1\rangle|x_2\rangle \leftarrow |I\rangle|G_1\rangle|G_2\rangle$ is given by (see Appendix B)

$$\mathcal{A}^{S+P_1+P_2}(f_j, x_1, x_2 \leftarrow I, G_1, G_2) = \sum_{i=1}^N \mathcal{A}^S(f_j \leftarrow b_i \leftarrow I) G_1(x_1 - B_i) G_2(x_2 - F_j). \quad (26)$$

By rule (iii), the transition probability exists in the form

$$P^{S+P_1+P_2}(f_j, x_1, x_2) = |\mathcal{A}^{S+P_1+P_2}(f_j, x_1, x_2 \leftarrow I, G_1, G_2)|^2. \quad (27)$$

It is easy to check that with two accurate pointers, $\Delta x_1, \Delta x_2 \rightarrow 0$, Eq. (27) reduces to $\sum_{i,j=1}^N |\mathcal{A}^S(f_j \leftarrow b_i \leftarrow I)|^2 \delta(x_1 - B_i) \delta(x_2 - F_j)$. With all B_i and F_j distinct, inspection of a pair of readings $x_1 \approx B_i$, $x_2 \approx F_j$ allows **A** to identify the system's path $|f_j\rangle \leftarrow |b_i\rangle \leftarrow |I\rangle$. By (2), the path may be endowed with probability, $P^S(f_j \leftarrow b_i \leftarrow I) = |\mathcal{A}^S(f_j \leftarrow b_i \leftarrow I)|^2$. The rest of the discussion can follow that of Sec. IV, and we leave it at that.

Once again, we are interested in the case where, as in Sec. IV, Alice employs one accurate, $\Delta x_2 \rightarrow 0$, and one highly inaccurate, $\Delta x_1 \rightarrow \infty$, pointer. In Eq. (26), the Gaussians with different shifts B_i are almost identical, and no reading x_1 will help **A** to identify the path chosen by the system. The rules (i)–(iii) give no recipe for assigning probabilities to the paths, so what, if anything, can be learnt about the system's condition at $t = t_1$?

Summation of Eq. (27) over j yields the quantum analog of the classical Eq. (11) [we used Eqs. (3) and (26)]

$$\rho(x_1, x_2) \xrightarrow{\Delta x_1 \rightarrow \infty, \Delta x_2 \rightarrow 0} \sum_{j=1}^N P^S(f_j \leftarrow I) \rho_1^I(x_1 - Z_j(\hat{B})) \delta(x_2 - F_j), \quad (28)$$

where

$$P^S(f_j \leftarrow I) = |\langle f_j | \hat{U}^S(t_2) | I \rangle|^2 \quad (29)$$

and

$$Z_j(\hat{B}) \equiv \text{Re} \left[\frac{\sum_{i=1}^N B_i \mathcal{A}^S(f_j \leftarrow b_i \leftarrow I)}{\sum_{i=1}^N \mathcal{A}^S(f_j \leftarrow b_i \leftarrow I)} \right]. \quad (30)$$

As in the classical case, the distribution of the first pointer's readings is a single Gaussian, displaced by some sort of average, obtained not with the path probabilities (which, by (ii), need not exist), but with the path amplitudes (which are always available).

This result was first obtained in Ref. [1], in a different form. For $\hat{H}^S = 0$, or $0 < t_1 < t_2$, the complex quantity in the square brackets of Eq. (30) can be written as

$$\frac{\sum_i \langle f_j | b_i \rangle B_i \langle b_i | I \rangle}{\sum_i \langle f_j | b_i \rangle \langle b_i | I \rangle} = \frac{\langle f_j | \hat{B} | I \rangle}{\langle f_j | I \rangle}, \quad (31)$$

in which one recognizes the “weak value of the operator $\hat{B}$ for a system pre- and postselected in the states $|I\rangle$ and $|f_j\rangle$,” respectively (see, e.g., Ref. [10]). We will discuss the meaning of the weak values after taking the analogy with the classical case one step further.

VII. QUASIPROBABILITIES FOR VIRTUAL PATHS

Quantum mechanics has no recipe for ascribing probabilities to interfering (virtual) alternatives, but one may try to obtain them by the method that worked in the classical case of Sec. IV. For $\hat{B} = |b_i\rangle \langle b_i| \equiv \hat{\pi}_i$, one can measure the shift $Z_j(\hat{\pi}_i)$, as well as the net probability of the system reaching $|f_j\rangle$ at t_2 , $P^S(f_j \leftarrow I)$. Their product should give something akin to the classical path probability in Eq. (14). Indeed path “probabilities,” defined in this manner,

$$\tilde{P}^S(f_j \leftarrow b_i \leftarrow I) \equiv \text{Re} \left[\mathcal{A}^S(f_j \leftarrow b_i \leftarrow I) \sum_{i=1}^N \mathcal{A}^{S*}(f_j \leftarrow b_i \leftarrow I) \right], \quad (32)$$

share various properties with their classical counterparts. In particular, they add up to one,

$$\sum_{i,j=1}^N \tilde{P}^S(f_j \leftarrow b_i \leftarrow I) = \sum_{j=1}^N P^S(f_j \leftarrow I) = \sum_{i=1}^N P^S(b_i \leftarrow I) = 1, \quad (33)$$

their absolutes do not exceed unity, $|\tilde{P}^S(f_j \leftarrow b_i \leftarrow I)| \leq 1$, and the term multiplying $\delta(x_2 - F_j)$ in Eq. (28) can be written as in the classical case [cf. Eqs. (11) and (12), $\Delta x_1 \rightarrow \infty$]

$$P^S(f_j \leftarrow I) \rho_1^I(x_1 - Z_j(\hat{B})) = \sum_{i=1}^N \tilde{P}^S(f_j \leftarrow b_i \leftarrow I) \rho_1^I(x_1 - B_i). \quad (34)$$

However, unlike their classical counterparts, quasiprobabilities in Eq. (32) can take negative values. Distributions that yield correct quantum mechanical averages, but may change sign, are often used in quantum theory, the best known example being the Wigner functions [19]. It was Feynman who suggested that quantum mechanics can be formulated as a classical-like statistical theory where probabilities are not restricted to be non-negative [20]. Equations (32) and (33) clearly provide a simple example of this approach, and next we consider a simple example, in order to find out how common may be the occurrence of negative path “probabilities.”

VIII. THE DOUBLE-SLIT EXAMPLE

Consider the “double-slit” case, $N = 2$, where the states $|b_{1,2}\rangle$ and $|f_{1,2}\rangle$ play the roles of the “slits” and “points on the screen,” respectively. The initial state can be chosen directed along the z axis on the Bloch sphere, $|I\rangle = |z_\uparrow\rangle$. The other four are the states polarized up and down the directions n and n' , whose azimuthal and polar angles are (ϕ, θ) and (ϕ', θ') , respectively. For simplicity, we assume that the system has no own dynamics, and put $\hat{H}^S = 0$. The four paths are equipped with quasiprobabilities (32),

$$\begin{aligned} \tilde{P}^S(n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow) &= \cos^2(\theta'/2) \cos^2(\theta/2) + \cos(\phi - \phi') \sin(\theta) \sin(\theta')/4 \equiv \tilde{P}_1, \\ \tilde{P}^S(n'_\uparrow \leftarrow n_\downarrow \leftarrow z_\uparrow) &= \cos^2(\theta'/2) \sin^2(\theta/2) - \cos(\phi - \phi') \sin(\theta) \sin(\theta')/4 \equiv \tilde{P}_2, \\ \tilde{P}^S(n'_\downarrow \leftarrow n_\uparrow \leftarrow z_\uparrow) &= \sin^2(\theta'/2) \cos^2(\theta/2) - \cos(\phi - \phi') \sin(\theta) \sin(\theta')/4 \equiv \tilde{P}_3, \\ \tilde{P}^S(n'_\downarrow \leftarrow n_\downarrow \leftarrow z_\uparrow) &= \sin^2(\theta'/2) \sin^2(\theta/2) + \cos(\phi - \phi') \sin(\theta) \sin(\theta')/4 \equiv \tilde{P}_4, \end{aligned} \quad (35)$$

at least one of which is likely to be negative.

The regions where this happens are mapped in Fig. 4 for $\phi' = 0$, $\theta' = 0.95\pi$ and various choices of the intermediate states $|n_\uparrow\rangle$ and $|n_\downarrow\rangle$. In particular, for

$$\phi = \pi, \quad \theta = \pi/2 \quad (36)$$

(the red dot in Fig. 4), one finds

$$\begin{aligned} \tilde{P}_1 &= -0.0360 \quad (\text{from } z_\uparrow \text{ to } n'_\uparrow \text{ via } n_\uparrow), \\ \tilde{P}_2 &= 0.0422 \quad (\text{from } z_\uparrow \text{ to } n'_\uparrow \text{ via } n_\downarrow), \\ \tilde{P}_3 &= 0.5630 \quad (\text{from } z_\uparrow \text{ to } n'_\downarrow \text{ via } n_\uparrow), \\ \tilde{P}_4 &= 0.4578 \quad (\text{from } z_\uparrow \text{ to } n'_\downarrow \text{ via } n_\downarrow), \end{aligned} \quad (37)$$

where $\tilde{P}_1 + \tilde{P}_2 = \cos^2(0.475\pi) = P(n'_\uparrow \leftarrow z_\uparrow)$ and $\tilde{P}_3 + \tilde{P}_4 = \sin^2(0.475\pi) = P(n'_\downarrow \leftarrow z_\uparrow)$.

In the next section, we briefly review the meaning and some properties of the path “probabilities” (32).

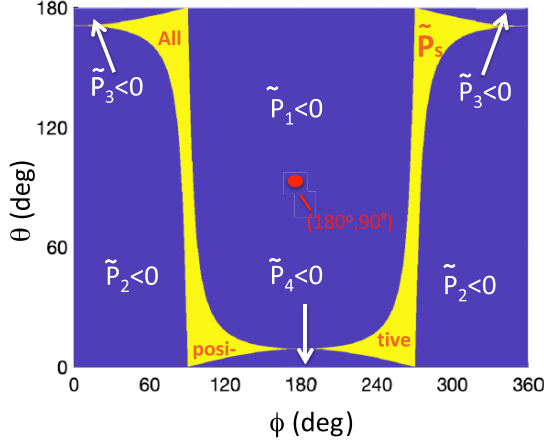


FIG. 4. The direction along which the intermediate states $|b_{1,2}\rangle = |n_{\uparrow,\downarrow}\rangle$ are quantized is parametrized by its azimuthal and polar angles, ϕ and θ . The system ($N = 2$) is prepared in $|z_{\uparrow}\rangle$ and postselected in $|n'_{\uparrow}\rangle$ or $|n'_{\downarrow}\rangle$, where $\phi' = 0$ and $\theta' = 0.95\pi$. Inside the yellow region, all four path “probabilities” in Eqs. (35) are positive. Elsewhere, at least one of them has a negative value.

IX. INTERPRETATION AND NONLOCALITY

There appears to be a consensus about using negative conditional probabilities as intermediaries in calculating non-negative frequencies of observable physical events [20,21]. Such are the calculations of the observable probabilities $P^S(f_j \leftarrow I)$ and $P^S(b_i \leftarrow I)$ in Eqs. (33) and (34) (see also Appendix C).

Recently, there have been attempts to give negative quasiprobabilities an operational meaning [22,23], and we will briefly discuss the problem in the context of this work. Indeed, **A** can also measure $\hat{\pi}_i$ and observe the shift $\tilde{P}^S(f_j \leftarrow b_i \leftarrow I) / \sum_{i'} \tilde{P}^S(f_j \leftarrow b_{i'} \leftarrow I) < 0$ [cf. Eqs. (2) and (34)]. This, however, should not be taken as an “experimental proof” that the probability of quantum system traveling a particular path can be negative.

One well-known objection is that a negative quantity cannot be a probability, since it cannot be related to the number of case in which the desired scenarios is realized [23]. This is the main reason why a quantum system cannot be simulated by a probabilistic classical computer [24] (see also Appendix D).

Second, the “probabilities” in Eq. (32) are *nonlocal* quantities in the following sense. Quasiprobability $\tilde{P}^S(f_j \leftarrow b_i \leftarrow I)$ depends [through the term $\sum_i A^{S*}(f_j \leftarrow b_i \leftarrow I)$] on the amplitudes of all paths, connecting $|I\rangle$ with $|f_j\rangle$. Thus, one hoping to probe the state of affairs in a particular path $f_j \leftarrow b_i \leftarrow I$ by using an inaccurate (or, which is the same, weakly coupled) pointer, discovers that his outcome depends on “what happens” in all paths interfering with the one he is interested in (see also Appendix E). Next, with the help of the “probabilities” (32), we analyze the mechanism that produces apparently “anomalous” pointer readings in inaccurate quantum measurements.

X. ANOMALOUS WEAK VALUES AND RESHAPING

Thus, Alice has a set of “probabilities” that can change sign, each referring to the entire (sub)set of the virtual paths leading to the same final outcome. As in the classical case [cf. Eq. (12)], the observed shift of a highly inaccurate pointer can now be written as an “average,”

$$Z_j(B) = \frac{\sum_{i=1}^N B_i \tilde{P}^S(f_j \leftarrow b_i \leftarrow I)}{\sum_{i=1}^N \tilde{P}^S(f_j \leftarrow b_i \leftarrow I)}. \quad (38)$$

For an unlikely transition, $P^S(j \leftarrow I) = \sum_i \tilde{P}^S(f_j \leftarrow b_i \leftarrow I) \ll 1$, the shift (38) can be made much larger than all the B_i 's. It can be shown (see, e.g., Ref. [9]) that one can always find the states $|I\rangle$ and $|f_j\rangle$ so as to give $Z_j(B)$ any desired value between $-\infty$ and ∞ . This, however, should not be taken as an “experimental proof” that in the quantum world the value of any quantity can be, say, arbitrary large.

One only needs to look in detail at the manner in which such values are obtained. In the classical case, integration of Eq. (12) over dx_2 in the limit $\Delta x_1 \rightarrow \infty$ yields [cf. Eqs. (1) and (10)]

$$\sum_{j=1}^N P(j \leftarrow I) \rho_1^I(x_1 - z_j(B)) = \rho_1^I(x_1 - Y(B)), \quad (39)$$

where $z_j(B)$ is given by Eq. (12). This, as we mentioned earlier, ensures that postselection cannot alter the outcomes of an earlier measurement.

It is easy to check that the causality principle (14) holds also in the quantum case, where one finds [cf. Eqs. (1) and (38)]

$$\sum_{j=1}^N P^S(f_j \leftarrow I) \rho_1^I(x_1 - Z_j(B)) = \rho_1^I(x_1 - Y(B)). \quad (40)$$

In both cases, the right-hand side of Eqs. (39) and (40) gives the distribution, obtained in a measurement without postselection. In the left-hand side, this distribution is partitioned into subdistributions, conditional on postselection in one of the final states. The only difference is that classically the shift $z(B)$ always lies between $\min(B_i)$ and $\max(B_i)$, whereas in the quantum case $Z(B)$ are not restricted to this interval. By postselecting in a state $|f_j\rangle$, **A** can separate the corresponding term in the left-hand side of Eq. (40), a Gaussian with a large shift $Z_j(B)$. The Gaussian is duly reduced by the factor $P^S(f_j \leftarrow I)$, which is smaller the larger is $Z_j(B)$, since the term must always fit under the tail of a larger Gaussian $\rho_1^I(x_1 - Y(B))$.

One misunderstanding may arise if one assumes that larger values of x_1 correspond also to larger values of B . This is, indeed, true in an accurate measurement, where the initial distribution $\rho^I(x_1)$ is narrow. This is no longer true if $\Delta x_1 \gg \max(B_i)$, since a large reading $x_1 \gg \max(B_i)$ is there not because the pointer has moved by that much [it has moved, we recall, by $Y(B)$, $B_{\min} \leq Y(B) \leq B_{\max}$] but because the initial distribution is very broad.

The same manipulation of the data can be performed without also relying on postselection. One can “carve” a reduced shifted copy from a larger Gaussian $\rho_1^I(x_1 - Y(B))$ by first performing a two-point measurement of Sec. V, and then accepting a reading x_1 with a probability

$$\omega_j(x_1) = \frac{P^S(j \leftarrow I) \rho^I(x_1 - Z_j(B))}{\sum_{j=1}^N P^S(j \leftarrow I) \rho^I(x_1 - Z_j(B))}, \quad (41)$$

so that

$$\omega_j(x_1) \rho_1^I(x_1 - Y(B)) = P^S(j \leftarrow I) \rho_1^I(x_1 - Z_j(B)). \quad (42)$$

With all P^S and ρ_1^I non-negative, one finds $0 \leq \omega_j(x_1) \leq 1$ so such a *reshaping* is always possible. Clearly, no anomalously large values of B are selected or detected by either method. Before concluding, we illustrate what has just been said on a simple numerical example.

XI. THE DOUBLE-SLIT EXAMPLE (CONTINUED)

As an example, consider a measurement of a projection of a spin-1/2,

$$\hat{B} = |n_\uparrow\rangle\langle n_\uparrow| - |n_\downarrow\rangle\langle n_\downarrow|, \quad B_1 = -B_2 = 1, \quad (43)$$

on the direction $n_\uparrow$, $\phi = \pi$, $\theta = \pi/2$, postselected along the direction $n'_\uparrow$, $\phi' = 0$, $\theta' = 0.95\pi$, discussed in Sec. VIII. We have [cf. Eqs. (32) and (38)]

$$\begin{aligned} P^S(n'_\uparrow \leftarrow z_\uparrow) &= \tilde{P}_1 + \tilde{P}_2 = 0.0062, \\ Z_1 &= (\tilde{P}_1 - \tilde{P}_2)/(\tilde{P}_1 + \tilde{P}_2) = -12.7062, \\ P^S(n_\uparrow \leftarrow z_\uparrow) &= \tilde{P}_1 + \tilde{P}_3 = 0.5, \\ P^S(n_\downarrow \leftarrow z_\uparrow) &= \tilde{P}_2 + \tilde{P}_4 = 0.5, \\ B_1 P^S(n_\uparrow \leftarrow z_\uparrow) + B_2 P^S(n_\downarrow \leftarrow z_\uparrow) &= 0. \end{aligned} \tag{44}$$

Results for $\Delta x_1/(B_1 - B_2) = 0.02$ are shown in Fig. 4. The distribution of the pointer's readings without postselection is a broad Gaussian centered at $x_1 = 0$. The distribution, conditioned upon postselection in $|n'_\uparrow\rangle$ is a much smaller Gaussian, centered at $x_1 = -12.7062$. It fits, as it must, under the larger curve, and can be obtained by “reshaping,” e.g., by retaining only a fraction of the results, with the acceptance probability (41) given by $[x_1 \leq 2009.5 \text{ so } \omega_1(x_1) \leq 1]$ [25],

$$\omega_1(x_1) \approx 0.0062 \exp(-0.0025x - 0.0161). \tag{45}$$

XII. CONCLUSIONS

To conclude, we address the questions posed earlier in the Introduction. In order to draw the parallels and highlight the distinctions between the classical and quantum cases, we found it useful to introduce quasiprobabilities, $\tilde{P}$ in Eq. (32), for the virtual paths connecting the initial and final states of a quantum system. Now, both statistical ensembles consist of sets of paths, equipped with probabilities in the classical, and with quasiprobabilities in the quantum case.

First, in a quantum measurement with postselection, can the shift of an inaccurate pointer in taken as a proof that a variable B takes a unique definite value? The classical case is quite clear. With more than one path connecting the initial and final states, the shift of a classical pointer yields a conventional ensemble average (12). With more than one virtual path involved, the shift of a quantum pointer in Eq. (38) is also an ensemble average, but for a bizarre ensemble, where the (quasi)probabilities are not required to be non-negative. (This is why the system cannot be seen as predestined to travel one of the virtual paths. Blocking a path endowed with a negative “probability” would *increase* the number of successful postselections, instead of reducing it [2].) In particular, for a quantity B taking value 1 on the chosen path, and 0 elsewhere, inaccurate classical measurement recovers the relative (i.e., normalized to a unit sum) path probability [cf. Eq. (12)]. A similar quantum measurement of such a quantity yields the relative quasiprobability or, which is the same, the real part of the relative path amplitude [cf. Eqs. (30) and (38)]. Thus, both classical and quantum inaccurate pointers loose the information about what happens in each individual trial. They allow one, instead, to evaluate the parameters of the corresponding statistical ensemble, combinations of the path probabilities (12) in the classical case, and of the path amplitudes (30) quantumly.

We also asked whether in unlikely quantum transitions, where the probability of postselection is small, a variable B does indeed take anomalous values in some meaningful sense. The argument for it could be based on large shifts, observed wherever the denominators in Eqs. (30) and (38) become small, while their numerators remain finite. The argument against is as follows. The pointer measuring B moves only once, and its position cannot be altered by whatever might be done to the system at a later time. Suppose the experiment is stopped immediately after, and no postselection is made. The distribution of a highly inaccurate pointer's readings is the initial broad Gaussian, shifted by an amount between the smallest and the largest values of B , $B_{\min}$ and $B_{\max}$. This is true both in the classical and quantum cases, as was discussed in Secs. III and V. If the postselection is made, the already existing readings can only be redistributed between the system's final states. It is the property of broad Gaussians (or, indeed, of all reasonably smooth functions; see Sec. II) that a distribution, conditional on arriving in a given final state, is also a reduced shifted copy of

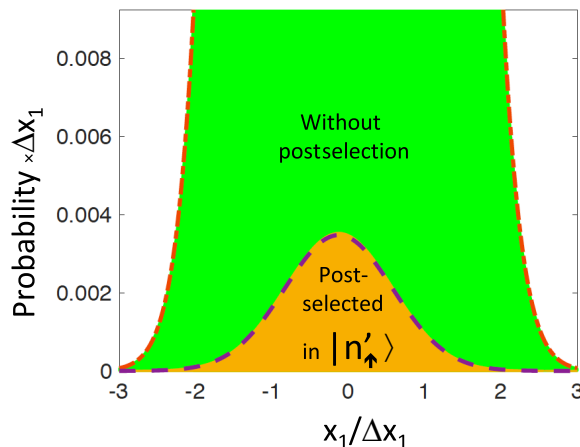


FIG. 5. Distributions of the readings of a highly inaccurate pointer, $\Delta x_1/(B_2 - B_1) = 50$ measuring spin's projection onto the direction $\phi = \pi$, $\theta = \pi/2$. The larger filled curve, centered at $x_1 = 0$, is the distribution without postselection (22), with the dot-dash line showing the approximation (10). The smaller filled curve, centered at $x_1 = -12.6051$ [cf. Eq. (44)], is the unnormalized distribution [cf. Eq. (27)] conditioned on postselection in the state $|n'_\uparrow\rangle$ polarized along the direction $\phi' = 0$, $\theta' = 0.95\pi$. The dashed line is the approximation (28).

the initial Gaussian. Adding up all such conditional distributions, one must, and does, recover the (unconditional) distribution obtained without postselection made.

The only difference between classical and quantum is that in the former case all path probabilities are non-negative, and all conditional shifts in Eq. (12) lie between $B_{\min}$ and $B_{\max}$. Quantally, the mathematics is the same, but the path probabilities are replaced by quasiprobabilities that may change sign. As a result, by choosing a suitable set of final states one can give the conditional shift in Eq. (38) any desired value, e.g., one which is much larger than $B_{\max}$, or much less than $B_{\min}$. (This is why “anomalous weak values” erroneously predicted in Ref. [12] cannot exist in classical statistics. They require path probabilities of opposite signs, and these do not exist in the classical case of Secs. III and IV.) In the example of Sec. XI, the experimenter, recording the pointer's readings only if the system is postselected in the chosen state, would obtain a reduced copy of the initial broad distribution, shifted to the left, as shown in Fig. 5. He/she would be ill advised to conclude that, by virtue of postselection, the spin projection has acquired the value of -12.7 , and the pointer has moved accordingly. Rather, since postselection in a given state implies retaining some readings, while discarding others, the unconditional distribution in Fig. 5 is merely “reshaped” into a smaller conditional Gaussian. In fact, the pointer has not moved at all [cf. the last of Eq. (44)], and the large negative readings, retained in this manner, present there from the beginning, have little to do with large negative projections of the measured spin. Exactly the same distribution can be obtained without postselection, by measuring the spin as in Sec. V, and accepting each reading with the probability in Eq. (41). The fact that each conditional distribution has to fit under the tail of the unconditional one makes such a procedure always possible. The notion of uncommonly large, or otherwise “unusual,” physical values lurking in quantum transitions where the interference between the relevant pathways is not destroyed is, therefore, spurious.

Finally, we note that the idea of “reshaping,” employed here in the context of inaccurate or “weak” measurements, is widely used to explain the related phenomenon of the apparent “superluminality” observed in quantum tunneling (see, e.g., Refs. [26,27]). There, the strongly reduced transmitted wave packet builds up by adding to the free state a number of its delayed copies [28]. Destructive interference ensures that the wave packet lies ahead of the freely propagating one creating the impression of a particle crossing the barrier infinitely fast. However, one does not

have to worry about violation of Einstein's relativity [29], just as one does not have to worry about the unusually large spin of Sec. XI. Both indicate the expected failure of a measurement protocol designed to circumvent the Uncertainty Principle [2].

ACKNOWLEDGMENTS

D.S. acknowledges financial support by the Grant No. PID2021-126273NB-I00 funded by MICINN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe," as well as by the Basque Government Grant No. IT1470-22.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: EQUATIONS (1) AND (3)

A change of variables $y = x/\Delta x$ in Eq. (1) yields

$$\begin{aligned} \sum_{i=1}^N A_i G(y - B_i/\Delta x) &\approx \sum_{i=1}^N A_i [G(y) - \partial_y G(y) B_i/\Delta x] \\ &\approx \sum_{i=1}^N A_i \left[G(y) - \partial_y G(y) \frac{\sum_{i=1}^N A_i B_i/\Delta x}{\sum_{i=1}^N A_i} \right] \\ &\approx \sum_{i=1}^N A_i \times G(y - z/\Delta x) = \sum_{i=1}^N A_i \times G\left(\frac{x - z}{\Delta x}\right). \end{aligned} \quad (\text{A1})$$

Note that the third inequality holds if $|z|/\Delta x \ll 1$, which would require a large Δx if A_i have opposite signs, $\sum_i A_i$ is small, and $|z|$ is large.

Furthermore, with complex A_i one finds $[G(B_i, \Delta x) \equiv G(\frac{x-B_i}{\Delta x})]$

$$\left| \sum_{i=1}^N A_i G(B_i, \Delta x) \right|^2 = \left(\sum_{i=1}^N \text{Re}[A_i] G(B_i, \Delta x) \right)^2 + \left(\sum_{i=1}^N \text{Im}[A_i] G(B_i, \Delta x) \right)^2. \quad (\text{A2})$$

Applying Eq. (3) to each sum and adding up the results yields

$$\left| \sum_{i=1}^N A_i G(B_i, \Delta x) \right|^2 \xrightarrow{\Delta x \rightarrow \infty} \left| \sum_{i=1}^N A_i \right|^2 G^2\left(x, \frac{\text{Re}[\sum_{i=1}^N A_i^* \times \sum_{i=1}^N A_i B_i]}{|\sum_{i=1}^N A_i|^2}, \Delta x\right). \quad (\text{A3})$$

A simple calculation shows that the second argument of the G in Eq. (A3) reduces to $\text{Re}[\sum_{i=1}^N A_i B_i / \sum_{i=1}^N A_i]$.

APPENDIX B: VON NEUMANN POINTERS

Alice can monitor a quantum system by means of pointers (P), heavy particles one-dimensional particles, whose positions she can observe directly. The pointers are briefly coupled to the system, at $t = t_1$, and just before $t = t_2$. The Hamiltonian, therefore, is ($\hbar = 1$)

$$\hat{H}^{S+P_1+P_2}(t) = \hat{H}^S + \hat{H}^{S+P_1}(t) + \hat{H}^{S+P_2}(t). \quad (\text{B1})$$

The pointers are designed to measure operators $\hat{B}$ and $\hat{F}$, respectively,

$$\hat{H}^{S+P_{1,2}}(t) = \hat{p}_{1,2} \times (\hat{B}, \hat{F}) \times \delta(t - t_{1,2}), \quad (\text{B2})$$

where $\hat{p}_l = \int dp_l |p_l\rangle p_l \langle p_l|$ and $\langle x_l | p_l \rangle = \exp(ip_l x_l)$, $l = 1, 2$, is the l th pointer's momentum operator. (There is also a third pointer, by means of which A establishes the system's initial state $|I\rangle$, and which we omit.) For any state $|\psi\rangle$, one finds

$$\exp\left[-i \int \hat{H}^{S+P_1}(t) dt\right] |\psi\rangle = \sum_{i=1}^N \langle b_i | \psi \rangle |G_1(B_i)\rangle |b_i\rangle, \\ |G_1(B_i)\rangle \equiv \int dx_1 \langle x_1 - B_i | G_1 \rangle |x_1\rangle \quad (\text{B3})$$

(and similarly for the second pointer). Thus, initial state of the composite $|\Phi(0)\rangle = |G_2\rangle |G_1\rangle |I\rangle$ evolves into

$$|\Phi(t_2)\rangle = \sum_{i,j=1}^N |f_j\rangle |G_2(F_j)\rangle |G_1(B_i)\rangle \langle f_j | \hat{U}^S(t_2 - t_1) | b_i \rangle \langle b_i | \hat{U}^S(t_1) | I \rangle, \quad (\text{B4})$$

and projecting it onto $|x_2\rangle |x_1\rangle |f_j\rangle$ yields Eq. (26).

APPENDIX C: NEGATIVE PROBABILITY AND OBSERVABLES

In Alice's experiment with a highly inaccurate pointer, $\Delta x_1 \rightarrow \infty$, there are three directly observable probabilities. One is defined by the rule (2) for the system in isolation [cf. Eq. (33)],

$$P^S(f_j \leftarrow I) = \sum_i \tilde{P}_{ij}^S = |\langle f_i | \hat{U}^S(t) | I \rangle|^2 \geq 0, \quad (\text{C1})$$

where we write $\tilde{P}_{ij}^S = \tilde{P}^S(f_j \leftarrow b_i \leftarrow I)$. The probability to find a reading x_2 in the vicinity of F_j agrees with it ($\epsilon \rightarrow 0$),

$$P(x_2 = F_j) \equiv \int_{F_j-\epsilon}^{F_j+\epsilon} dx_2 \int dx_1 \rho(x_1, x_2) = \sum_i \tilde{P}_{ij}^S \geq 0. \quad (\text{C2})$$

Finally, there is the probability to find a reading x_1 given that x_2 is in the vicinity of F_j

$$P(x_1, x_2 F_j) \equiv \int_{F_j-\epsilon}^{F_j+\epsilon} dx_2 \rho(x_1, x_2) = \sum_i \rho_1^I(x_1 - B_i) \tilde{P}_{ij}^S \geq 0. \quad (\text{C3})$$

Furthermore, negative probability of a scenario may mean that the scenario cannot be verified directly [20]. Indeed, in neither of the three cases, the system is seen to travel one path and not the other.

Finally, the same is true even when quasiprobabilities of all paths are non-negative (e.g., in the light-colored area in Fig. 4). One still cannot have the probabilities $\tilde{P}^S(f_j \leftarrow b_i \leftarrow I) = \text{Re}[\mathcal{A}^S(f_j \leftarrow b_i \leftarrow I) \mathcal{A}^{S*}(f_j \leftarrow I)]$ and know the path traveled by the system. An attempt to distinguish between the paths by using an accurate pointer would create different probabilities, $P^S(f_j \leftarrow b_i \leftarrow I) = |\mathcal{A}^S(f_j \leftarrow b_i \leftarrow I)|^2$.

APPENDIX D: AN EXAMPLE

Here is a simple illustration. With the classical path probabilities $P(j \leftarrow i \leftarrow I)$ known, A can simulate, e.g., the $P(i = 3 \leftarrow I)$ in Eq. (6). This can be done by splitting interval $[0, 1]$ into N^2 segments, with each length proportional to $P(j \leftarrow i \leftarrow I)$. Whenever random number generator yields a number within a segment, the ball travels the corresponding path. Dividing the number of cases in which the ball was seen, e.g., in the node 3, N_3 , by the total number of trials $N \gg 1$, Alice estimates $P(3 \leftarrow I) \approx N_3/N$. In the quantum case, where at least one $\tilde{P}^S(F_j \leftarrow b_i \leftarrow I) < 0$, the method fails, since instruction to sample with a negative probability is obviously meaningless.

APPENDIX E: NONLOCALITY OF QUASIPROBABILITIES

In the simplest classical case $N = 2$, Alice can alter the probability $P(1 \leftarrow 2 \leftarrow I)$, while keeping the $P(1 \leftarrow 1 \leftarrow I)$ intact simply by altering the branching probability at the second node, $P(1 \leftarrow 2) \rightarrow P'(1 \leftarrow 2)$. Now, using the pointer that monitors the path $1 \leftarrow 1 \leftarrow I$, $B_i = \delta_{i1}$, and multiplying its shift by $P(1 \leftarrow 1 \leftarrow I) + P'(1 \leftarrow 1 \leftarrow I)$, she will always recover the same path probability $P(1 \leftarrow 2 \leftarrow I)$. We call such measurement “local,” since the result only depends on “what happens” in the path $1 \leftarrow 1 \leftarrow I$.

This is no longer true in the quantum case, where a path is described by an amplitude, rather than a path probability. In the double-slit case of Sec. VIII $N = 2$, with $\phi = \phi'$, the amplitudes for the paths leading to $|n'_\uparrow\rangle$ are given by

$$\begin{aligned}\mathcal{A}(n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow) &= \cos(\theta/2) \cos(\theta/2 - \theta'/2), \\ \mathcal{A}(n'_\uparrow \leftarrow n_\downarrow \leftarrow z_\uparrow) &= \sin(\theta/2) \cos(\theta/2 - \theta'/2), \\ \mathcal{A}(n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow) + \mathcal{A}(n'_\uparrow \leftarrow n_\downarrow \leftarrow z_\uparrow) &= \cos(\theta'/2).\end{aligned}\tag{E1}$$

For $\beta < 1$ and a fixed $0 < \theta < 2 \arccos(\beta/2)$, the choice of the final state

$$\theta' = \theta + 2 \arccos(\beta / \cos(\theta/2))\tag{E2}$$

keeps $\mathcal{A}(n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow) = \beta$ constant, while the second amplitude on Eq. (E1) and

$$\tilde{P}^S(n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow) = \beta \cos(\theta'/2)\tag{E3}$$

both change. By using a highly inaccurate pointer, which moves only when the system follows the path $n'_\uparrow \leftarrow n_\uparrow \leftarrow z_\uparrow$, A recovers a quantity that depends on the amplitudes of all paths leading to the same outcome. In an optical realization of the double-slit experiment, the paths in different arms of a Mach-Zehnder interferometer can be well separated in space. Alice would not be wrong to describe this behavior as “nonlocal.”

APPENDIX F: CAUSALITY IN THE QUANTUM CASE

Distribution of the first pointer’s readings, regardless of what reads the second pointer, is [cf. Eqs. (26) and (27)]

$$\begin{aligned}\rho_1(x_1) &= \sum_{j=1}^N \int dx_2 |G_2(x_2 - F_j)|^2 \left| \sum_{i=1}^N \mathcal{A}^S(f_j \leftarrow b_i \leftarrow I) G_1(x_1 - B_i) \right|^2 \\ &= \sum_{i'=1}^N \left[\sum_{j=1}^N \mathcal{A}^{S*}(f_j \leftarrow b_{i'}) \mathcal{A}^S(f_j \leftarrow b_i) \right] \mathcal{A}^{S*}(b_{i'} \leftarrow I) \mathcal{A}^S(b_i \leftarrow I) G^*(x_1 - B_{i'}) G(x_1 - B_i).\end{aligned}\tag{F1}$$

Since $\sum_{j=1}^N \mathcal{A}^{S*}(f_j \leftarrow b_{i'}) \mathcal{A}^S(f_j \leftarrow b_i) = \langle b_{i'} | b_i \rangle = \delta_{i'i'}$, this reduces to

$$\sum_{i=1}^N |\mathcal{A}^S(b_i \leftarrow I)|^2 |G(x_1 - B_i)|^2 = \sum_{i=1}^N P^S(b_i \leftarrow I) \rho_1^I(x_1 - B_i),\tag{F2}$$

which is the distribution (22) one obtains if no postselection is made.

- [1] Y. Aharonov, D. Z. Albert, and L. Vaidman, How the result of a measurement of a component of a spin-1/2 particle can turn out to be 100, *Phys. Rev. Lett.* **60**, 1351 (1988).
- [2] R. P. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics III* (Dover Publications, New York, 1989).
- [3] D. Bohm, *Quantum Theory* (Dover Publications, NY, 1989), p. 600.
- [4] R. P. Feynman, *The Character of Physical Law* (MIT Press, London, 1985).
- [5] Y. Aharonov and L. Vaidman, Complete description of a quantum system at a given time, *J. Phys. A: Math. Gen.* **24**, 2315 (1991).
- [6] Y. Aharonov, S. Popescu, and J. Tollaksen, A time-symmetric formulation of quantum mechanics, *Phys. Today* **63**(11) (2010).
- [7] K. Robertson, Can the two-time interpretation of quantum mechanics solve the measurement problem? *Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys.* **58**, 54 (2017).
- [8] J. A. Barandes, The trouble with weak values, [arXiv:2602.09380](https://arxiv.org/abs/2602.09380).
- [9] D. Sokolovski, D. A. Ramirez, and S. B. Martin, Speakable and unspeakable in quantum measurements, *Ann. Phys.* **535**, 2300261 (2023).
- [10] J. Dressel, M. Malik, F. M. Miatto, N. Jordan, and R. W. Boyd, *Colloquium: Understanding quantum weak values: Basics and applications*, *Rev. Mod. Phys.* **86**, 307 (2014).
- [11] D. Sokolovski, Weak measurements measure probability amplitudes (and very little else), *Phys. Lett. A* **380**, 1593 (2016).
- [12] C. Ferry and J. Combes, How the result of a single coin toss can turn out to be 100 heads, *Phys. Rev. Lett.* **113**, 120404 (2014).
- [13] A. Brodutch, Comment on “How the result of a single coin toss can turn out to be 100 heads”, *Phys. Rev. Lett.* **114**, 118901 (2015).
- [14] A. Uranga, E. Akhmatskaya, and D. Sokolovski, Quantum weak values and the question, *Entropy* **27**, 259 (2025).
- [15] A. Peres, Quantum measurements with postselection, *Phys. Rev. Lett.* **62**, 2326 (1989).
- [16] E. C. Zeeman, *Catastrophe Theory: Selected Papers* (Addison-Wesley Educational Publishers, Inc., Boston, MA, 1977).
- [17] S. N. Sahoo, S. Chakraborti, S. Kanjilal, S. R. Behera, D. Home, A. Matzkin, and U. Sinha, Unambiguous joint detection of spatially separated properties of a single photon in the two arms of an interferometer, *Commun. Phys.* **6**, 203 (2023).
- [18] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955), Chap. VI, pp. 183–217.
- [19] E. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.* **40**, 749 (1932).
- [20] R. P. Feynman, Negative probability, in *Quantum Implications: Essays in Honour of David Bohm*, edited by F. D. Peat and B. Hiley (Routledge & Kegan Paul Ltd., London, 1987), p. 235.
- [21] M. S. Bartlett, Negative probability, *Math. Proc. Cambridge Philos. Soc.* **41**, 71 (1945).
- [22] B. L. Higgins, M. S. Palsson, G. Y. Xiang, H. M. Wiseman, and G. J. Pryde, Using weak values to experimentally determine in a two-photon state with Bell correlations, *Phys. Rev. A* **91**, 012113 (2015).
- [23] J. Ryu, S. Hong, J.-S. Lee, K. H. Seol, J. Jae, J. Lim, J. Lee, K.-G. Lee, and J. Lee, Optical experiment to test negative probability in context of quantum measurement selection, *Sci. Rep.* **9**, 19021 (2019).
- [24] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [25] Or just take a pair of scissors and cut along the dashed line in Fig. 4.
- [26] R. S. Dumont, T. Rivlin, and E. Pollak, The relativistic tunneling flight time may be superluminal, but it does not imply superluminal signaling, *New J. Phys.* **22**, 093060 (2020).
- [27] M. Yu, K. Liu, M. Li, J. Yan, C. Cao, J. Tan, J. Liang, K. Guo, W. Cao, P. Lan, *et al.*, Full experimental determination of tunneling time with attosecond-scale streaking method, *Light: Sci. Appl.* **11**, 215 (2022).
- [28] D. Sokolovski and E. Akhmatskaya, Superluminal paradox’ in wave packet propagation and its quantum mechanical resolution, *Ann. Phys.* **339**, 307 (2013).
- [29] D. Sokolovski and E. Akhmatskaya, No time at the end of the tunnel, *Commun. Phys.* **1**, 47 (2018).