

Wigner-friend scenarios with noninvasive weak measurements

A. Matzkin

Laboratoire de Physique Théorique et Modélisation, Centre National de la Recherche Scientifique Unité 8089, CY Cergy Paris Université, 95302 Cergy-Pontoise, France

D. Sokolovski

Departamento de Química-Física, Universidad del País Vasco, UPV/EHU, Leioa, Spain and IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain

(Received 28 August 2020; accepted 16 November 2020; published 4 December 2020)

Wigner-friend scenarios—in which an external agent describes quantum mechanically a laboratory in which a friend is making a measurement—give rise to possible inconsistencies due to the ambiguous character of quantum measurements. In this paper, we investigate Wigner-friend scenarios in which the external agents can probe in a noninvasive manner the dynamics inside the laboratories. We examine probes that can be very weakly coupled to the systems measured by the friends, or to the pointers or environments inside the laboratories. These couplings, known as weak measurements, are asymptotically small and do not change the outcomes obtained by the friends or their probabilities. Within our scheme, we show that the weakly coupled probes indicate to the external agents how to obtain consistent predictions, irrespective of the possible inconsistencies of quantum measurement theory. These noninvasive couplings could be implemented with present-day technologies.

DOI: [10.1103/PhysRevA.102.062204](https://doi.org/10.1103/PhysRevA.102.062204)**I. INTRODUCTION**

Measurements are special in quantum mechanics. Indeed quantum systems evolve unitarily, except when a measurement takes place: a “special rule” needs to be enforced in order to account for the appearance of a definite measurement outcome out of the linear superposition of pointer-system states. Technically, this rule appears as a projection of the quantum state right before measurement to an eigenstate of the measured observable. Physically, there is no unambiguous and consistent manner of accounting for this rule [1]. Logically, this makes quantum theory inconsistent [2]. Several options are on the table, but none of them is entirely satisfactory and there is unsurprisingly no consensus on these issues, overall known as the celebrated measurement problem.

Could there be an underlying basis that would account for the projection that takes place when a measurement is completed—the “collapse” of the wave function? In a series of papers [3–5], Wigner assumed unitary evolution to be universal—implying it also applies to macroscopic bodies—but that linear superposition cannot be applied to conscious observers making a measurement. Hence if a Stern-Gerlach experiment takes place in a sealed laboratory L, an agent W outside the laboratory would describe the state of the atom having passed through the inhomogeneous magnetic field by a linear superposition of up and down spin along the field direction. Now if there is a friend F inside the sealed laboratory running the experiment, she would take the superposition to “collapse” upon completion of her measurement. But how should W describe the experiment?

If unitary evolution is universal, the answer would be that as long as W is unaware of the outcome, he will take the

quantum state of the laboratory to be in superposition even after F completed her measurement. Wigner was dubious that a superposition of different states of consciousness would make sense,¹ but concluded (at least in his 1962 paper [3]) that an experimental test attempting to distinguish a linear superposition of L from a mixed state would in practice be impossible. Still it is worthwhile to investigate whether such assumptions can be taken to be consistent.

Remarkably, in two recent works Brukner [7] and Frauchiger and Renner [8] have shown by employing an extended Wigner-friend scenario (EWFS) composed of two sets $i = 1, 2$ each involving a friend F_i inside a laboratory L_i and an agent W_i external to the laboratories L_1 and L_2 that assuming the friends get a definite outcome while the agents W_i describe these processes by unitary evolutions leads to a contradiction. The implications of this have sparked a wild interest [9–16], though its relevance depends on a set of specific assumptions that are explicitly or implicitly being made.

In this paper we will not be interested in these specific assumptions—in particular in those made in [7,8]—but focus instead on the basic hypothesis at the root of the Wigner-friend scenarios (WFSSs) according to which a completed measurement in a closed laboratory can be described by an external agent as a unitary evolution. We will introduce a procedure by which the agents W_i can acquire information on the state of

¹Actually Wigner’s position was that quantum mechanics does not apply to conscious observers. He speculated [3,5] that consciousness induces a nonlinear modification of linear quantum mechanics, though he will change his mind afterwards [6].

the system or the state of the pointer inside the laboratories L_i without perturbing in any essential way the measurements taking place in L_i , in particular without modifying any outcome probabilities. This information will be acquired with noninvasive probes through a minimal disturbing scheme known as weak measurements [17]. In general, a weak measurement involves a unitary interaction between a system and a very weakly coupled probe. The weakly coupled probe becomes entangled with the system, and it is only when the system is subsequently subjected to a measurement of another observable that the information can be retrieved from the probe.

Here our strategy will be to connect a noninvasive probe available to the agents W_i to elements inside the otherwise sealed laboratories L_i . For instance such probes may be coupled through a weak interaction to the spins inside L_i , to the pointers measuring the spins, or to the environment inside L_i by which the friends acquire information on the spin measurement outcomes. We will see that even in the original Wigner setup assuming unitary evolution leads to a contradiction as W would read from the weakly coupled probe that the spin (or the pointer that measured the spin, or the environment inside L) is in a state of superposition, while F announces she obtained a definite outcome. This contradiction subsists in the extended scenario: the friends announce they obtained a definite outcome but if the external agents assume unitary evolution this would not be reflected in the state of the probes. Actually if the agents W_i are allowed to communicate, they would realize their probe states act as a witness of entanglement between the spins (or the pointers, or the environments) of each laboratory. This contradiction is of course the counterpart of the contradiction uncovered in [7,8]. As expected, no contradiction occurs if the friends' measurements result in a global collapse: in this case the weakly coupled probes are in perfect agreement with the outcomes announced by F_1 and F_2 . Note that in this paper we will take ‘‘collapse’’ to represent the second step of the measurement process, recalled in Sec. II, without any commitment to an underlying picture or specific interpretation (we will sometimes recall this point by writing ‘‘possibly effective collapse’’).

In Sec. II we will introduce the original (WFS) and the extended (EWFS) setups we will be working with. Section III will deal with implementing different noninvasive measurements with weakly coupled probes in the laboratories for both the WFS and EWFS. We will assume the external agents W can compute the quantum state evolution either by following full unitary evolution or by applying globally the projection postulate, each case leading to different predictions for the weakly coupled probes. We will discuss our results in Sec. IV; in particular we will argue that although connecting a weak noninvasive probe to a laboratory breaks the independence of L from a logical viewpoint, it does not change the physical situation in the laboratories. We close with our conclusions in Sec. V.

II. WIGNER-FRIEND SETUPS

A. WFS: Single laboratory and observers

In the original WFS Wigner [3] introduces a laboratory L in which a friend F performs a Stern-Gerlach experiment

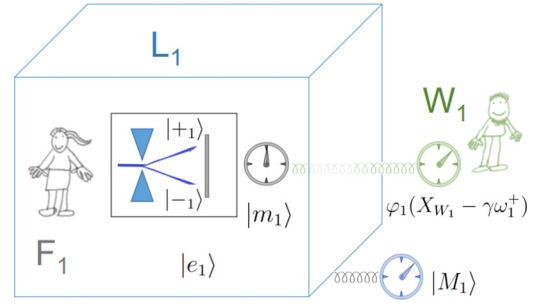


FIG. 1. Wigner-friend scenario. In the original WFS, the friend F_1 measures the spin of an atom with her pointer in state $|m_1\rangle$ inside the sealed laboratory L_1 . The external agent W_1 measures L_1 with his pointer in state $|M_1\rangle$. In this paper, W_1 has in addition a noninvasive probe (online color: in green), coupled here to the friend's pointer. After F_1 completes her measurement, the probe, initially in the quantum state $\varphi_1(X_{W_1})$, will have shifted to $\varphi_1(X_{W_1} - \gamma\omega_1^+)$ where ω_1^+ is the so-called weak value.

on an atomic spin, while an agent W is outside the isolated laboratory and ultimately measures the quantum state of the laboratory (see Fig. 1). We will introduce the following quantum states:

$$\begin{aligned} |\Psi(t=0)\rangle &= \underbrace{|s(t=0)\rangle|m(t=0)\rangle|e(t=0)\rangle}_{|L(t=0)\rangle} |M(t=0)\rangle \quad (1) \\ &\equiv |L(t=0)\rangle \otimes |M(t=0)\rangle. \quad (2) \end{aligned}$$

$|s\rangle$, $|m\rangle$, and $|e\rangle$ label, respectively, the spin, the pointer measuring the spin, and the environment of the isolated laboratory, whose overall quantum state is denoted by $|L\rangle$. $|M\rangle$ is the quantum state of W 's pointer, that is assumed to measure L .

A measurement in quantum theory is represented in two steps. First a unitary evolution

$$\exp\left(-i \int_0^t dt' H_{\text{int}}(t')/\hbar\right) \quad (3)$$

ouples the system observable \hat{A} that is being measured to a pointer observable \hat{P} (typically the momentum). The coupling Hamiltonian is $H_{\text{int}} = g(t)\hat{A}\hat{P}$ where $g(t)$ is the coupling strength. Before the measurement the overall quantum state can be put as $\psi(t=0) = \sum_k c_k |a_k\rangle|m\rangle$, where the system is expanded over the eigenstates $|a_k\rangle$ of \hat{A} . The pointer state $|m\rangle$ is typically a tightly localized function (e.g., a Gaussian), initially centered on $X = 0$. Unitary evolution leads to

$$\psi(t) = \exp(-ig\hat{A}\hat{P}/\hbar)\psi(t=0) \quad (4)$$

$$= \sum_k c_k |a_k\rangle m(X - ga_k) \quad (5)$$

where we have $\langle X|m_k\rangle = m(X - ga_k)$ and set $g \equiv \int dt' g(t')$ assuming \hat{A} is time independent and recalling that $\exp(-iga_k\hat{P}/\hbar)$ is a translation operator. Each term of the sum in Eq. (5) represents the eigenstate $|a_k\rangle$ correlated with a pointer state shifted by ga_k relative to the original position. For sufficiently strong couplings g the pointer wave functions

$m(X - ga_k)$ do not overlap in position space so measuring the pointer position directly yields the corresponding eigenvalue.

In the Wigner-friend scenario, F measures the spin component along some direction, say \mathbf{x} , hence \hat{A} is given by $\sigma_{\mathbf{x}}$ and there are two pointer states $m(X \mp g)$ correlated, respectively, with the spin $s_x = \pm 1$. The second step of the measurement is the projection that, as was first axiomatized by von Neumann [18], is necessary in order to account for the appearance of a definite outcome. Indeed including additional unitary interactions, e.g., between the pointer and the environment of the pointer (that could include the photons having interacted with the pointer traveling toward the friend's eye), would only create a growing chain of entangled states, the von Neumann chain, defined in the measurement basis; e.g., including the quantum states of the environment $|e\rangle$ is tantamount to replacing Eq. (5) by $\sum_k c_k |a_k\rangle |m_k\rangle |e_k\rangle$.

So as far as F measuring the spin is concerned, the initial state $|L(t=0)\rangle$ introduced in Eq. (2) goes into $|L(t)\rangle = \sum_{\pm} c_{\pm} |\pm\rangle |m_{\pm}\rangle |e_{\pm}\rangle$ in the first unitary step but once the measurement is completed at $t = t_f$ and an outcome, say $+$ (respectively, $-$), is obtained, the final state is the projection $|\mathcal{L}+\rangle \equiv |+\rangle \langle +| L(t_f)\rangle$ [respectively, $|\mathcal{L}-\rangle \equiv |-\rangle \langle -| L(t_f)\rangle$]. According to the ‘‘universality’’ assumption, W should treat L with F inside performing a measurement just as he would treat any other unmeasured quantum system, in which case the evolution in L is considered to be fully unitary and before W's measurement the quantum state is

$$|\Psi(t)\rangle = |L(t)\rangle |M(t=0)\rangle, \quad (6)$$

where we assumed W's pointer has no evolution previous to the coupling with L. Alternatively, if W discards the universality assumption and assumes instead that measurements are always ‘‘special’’ (as per von Neumann's rule, irrespective of where and how the measurement was made), he will take the friend's measurement to result in a (possibly effective) collapse of the global quantum state. In this case W would consider the state of L to be given not by $|\Psi(t)\rangle$ but by a statistical mixture of density matrices $|\mathcal{L}+\rangle \langle \mathcal{L}+|$ and $|\mathcal{L}-\rangle \langle \mathcal{L}-|$. Wigner in his paper [3] concludes that in practice, no experiment could discriminate between both alternatives. Indeed, for a genuine macroscopic size laboratory the interferences that need to be detected are so small that they become effectively undetectable [19,20], an argument that goes back to Bohm [21].

B. EWFS: Two sets of laboratories

Drawing on a version of the WFS proposed by Deutsch [2], Brukner [7] and Frauchiger and Renner [8] introduce an extended scenario composed of two sets of Wigner-friend setups sharing spins 1/2 in an entangled state (see Fig. 2). Our scheme will follow the latter proposal [8], in which a pair of spin-1/2 atoms is prepared in the nonmaximally entangled state

$$|\chi\rangle = \frac{1}{\sqrt{3}} |+\rangle |\downarrow_2\rangle + \frac{\sqrt{2}}{\sqrt{3}} |-\rangle |+\rangle_2 \quad (7)$$

where $\uparrow\downarrow$ refer to the eigenstates of the spin component along z , with $\sigma_z|\uparrow\rangle = |\uparrow\rangle$, $\sigma_z|\downarrow\rangle = -|\downarrow\rangle$, and $|\pm\rangle = (|\downarrow\rangle \pm |\uparrow\rangle)/\sqrt{2}$ are the eigenstates of σ_x . Employing the en-

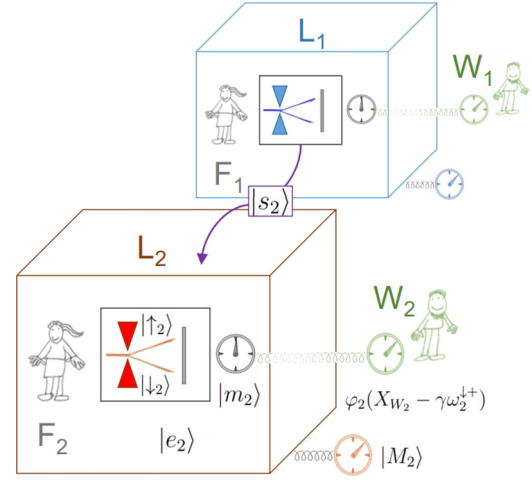


FIG. 2. Extended Wigner-friend scenario. In addition to the original WFS of Fig. 1, reproduced here in the top half of the figure, a friend F_2 in a laboratory L_2 receives a spin in state $|s_2\rangle$, entangled with the spin in L_1 when unitary evolution takes place. W_2 has his own noninvasive probe coupled to systems inside L_2 .

tangled state (7) leads to a Bell-type inequality with zero as the upper bound [22,23], instead of the usual bound of 2 obtained with maximally entangled states that Brukner uses in his scheme [7]. Conceptually, the advantage of having zero as an upper bound is that the violation of the inequality by the quantum version appears as logically impossible for a classical model in which all the measured observables have joint values (recall that when observables do not commute, the quantum formalism precludes the existence of joint probability distributions [24]).

In the narrative of [8], F_1 measures her spin first and then sends the other atom to L_2 in state $|\downarrow_2\rangle$ if the outcome was $+$ and in the state $|+\rangle_2$ if F_1 obtained $-$. This is equivalent to dealing with Eq. (7) from the outset but has the advantage of starting with a product state for the atoms, which at $t = 0$ are both inside L_1 . In this case however L_1 and L_2 are not isolated from each other, though our main requirement is that they should be isolated from the external agents W_1 and W_2 .

The initial state of L_1 is thus

$$|L_1(t=0)\rangle = |s_1\rangle |s_2\rangle |m_1(t=0)\rangle |e_1(t=0)\rangle, \quad (8)$$

where

$$|s_1\rangle = \left(\frac{1}{\sqrt{3}} |+\rangle + \frac{\sqrt{2}}{\sqrt{3}} |-\rangle \right) \quad (9)$$

and $|s_2\rangle$ can be any spin state. The first step of F_1 's measurement involves the unitary coupling (3) leading to

$$|L_1(0 < t < t_1)\rangle = \left(\frac{1}{\sqrt{3}} |+\rangle |m_1+\rangle + \frac{\sqrt{2}}{\sqrt{3}} |-\rangle |m_1-\rangle \right) \times |s_2\rangle |e_1(t=0)\rangle. \quad (10)$$

Upon completion of the measurement at $t = t_1$, an observer inside L_1 reads a definite outcome projecting the premeasurement state to

$$|\mathcal{L}_1(t_1)\rangle |s_2\rangle = \begin{cases} |\mathcal{L}_1+\rangle |\downarrow_2\rangle \equiv |+\rangle |m_1+\rangle |e_1+\rangle |\downarrow_2\rangle \\ |\mathcal{L}_1-\rangle |+\rangle_2 \equiv |-\rangle |m_1-\rangle |e_1-\rangle |+\rangle_2 \end{cases}; \quad (11)$$

we have assumed a gate coupling the two spins brings $|s_2\rangle$ to either $|\downarrow_2\rangle$ or $|+\rangle_2$ depending on whether $|+\rangle_1$ or $|-\rangle_1$ was obtained. Spin 2 is sent to L_2 at t'_1 whose state becomes

$$|\mathcal{L}_2(t'_1)\rangle = \begin{cases} |+\rangle_2|m_2\rangle|e_2+\rangle \\ |\downarrow_2\rangle|m_2\rangle|e_2-\rangle \end{cases}, \quad (12)$$

where $|m_2\rangle$ is F_2 's pointer in the ready state. F_2 measures her spin in the $|\uparrow\downarrow\rangle$ basis, again through the interaction (3), and after her measurement is complete at $t = t_2$ the state of L_2 becomes

$$|\mathcal{L}_2(t_2)\rangle = \begin{cases} |\mathcal{L}_2\uparrow\rangle \equiv |\uparrow_2\rangle|m_2\uparrow\rangle|e_2\uparrow\rangle \\ |\mathcal{L}_2\downarrow\rangle \equiv |\downarrow_2\rangle|m_2\downarrow\rangle|e_2\downarrow\rangle \end{cases}. \quad (13)$$

F_2 has at that point obtained a definite outcome.

Now if we assume that for the external agents W_1 and W_2 , L_1 and L_2 are ordinary (though macroscopic) quantum systems, the laboratories evolution is described unitarily and instead of Eq. (11) we have

$$|L_1(t_1)\rangle = \frac{1}{\sqrt{3}}|H_1\rangle|\downarrow_2\rangle|m_1+\rangle|e_1+\rangle + \frac{\sqrt{2}}{\sqrt{3}}|-\rangle_1|+\rangle_2|m_1-\rangle|e_1-\rangle. \quad (14)$$

Upon receiving atom 2, L_2 gets entangled with L_1 and the overall state at t_2 is

$$|L_{12}(t_2)\rangle = \frac{1}{\sqrt{3}}|\mathcal{L}_1+\rangle|\mathcal{L}_2\downarrow\rangle + \frac{1}{\sqrt{3}}|\mathcal{L}_1-\rangle[|\mathcal{L}_2\uparrow\rangle + |\mathcal{L}_2\downarrow\rangle], \quad (15)$$

where we have used the notation introduced in Eqs. (11) and (13). Finally W_1 measures L_1 in the basis $|\mathcal{L}_1\uparrow\downarrow\rangle$ and W_2 measures L_2 in the basis $|\mathcal{L}_2\pm\rangle$ defined as per the spin basis by

$$|\mathcal{L}_i\uparrow\downarrow\rangle = \frac{1}{\sqrt{2}}(|\mathcal{L}_i+\rangle \mp |\mathcal{L}_i-\rangle), \quad (16)$$

$$|\mathcal{L}_i\pm\rangle = \frac{1}{\sqrt{2}}(|\mathcal{L}_i\downarrow\rangle \pm |\mathcal{L}_i\uparrow\rangle). \quad (17)$$

It is useful to rewrite Eq. (15) in the W 's measurement basis:

$$|L_{12}(t_2)\rangle = \frac{1}{\sqrt{12}}(-|\mathcal{L}_1\uparrow\rangle|\mathcal{L}_2+\rangle + |\mathcal{L}_1\uparrow\rangle|\mathcal{L}_2-\rangle + 3|\mathcal{L}_1\downarrow\rangle|\mathcal{L}_2+\rangle + |\mathcal{L}_1\downarrow\rangle|\mathcal{L}_2-\rangle). \quad (18)$$

The contradiction obtained by Frauchiger and Renner [8] takes here the following form.

(i) If F_2 obtains the outcome \uparrow , then necessarily F_1 has obtained $-$ [by Eq. (11), with $|+\rangle = (|\downarrow\rangle + |\uparrow\rangle)/\sqrt{2}$].

(ii) If W_2 obtains $-$, then necessarily F_1 has obtained $+$ [from Eq. (18) with Eqs. (17) and (11)].

(iii) If W_1 obtains \uparrow , then necessarily F_2 has obtained \uparrow [from Eq. (18) with Eqs. (16) and (13)].

(iv) It follows from (i)–(iii) that the outcomes $W_1 = \uparrow$, $W_2 = -$ cannot be obtained jointly (indeed, $W_1 = \uparrow$ implies that $F_1 = -$, but then $W_2 = -$ cannot be obtained).

(v) From Eq. (18), the outcomes $W_1 = \uparrow$, $W_2 = -$ can be obtained with nonzero probability.

Statements (iv) and (v) are in direct contradiction. Technically, this follows from the fact that taken together, points (ii)–(iv) implicitly assume the existence of a joint probability distribution for the outcomes of σ_x and σ_z within each set

$\{L_i, W_i\}$. This is well known to be impossible, given that σ_x and σ_z do not commute: if only unitary evolution is assumed, the outcome probabilities are obtained by adding amplitudes that interfere as per (v). But here agents W_i still need to account for the outcomes obtained by the friends which is the reason points (ii)–(iv) appear to hold. Note that if global collapse occurs, then neither (ii) nor (iii) [and therefore (iv)] holds, since the correlations between the outcomes of W_1 and F_2 as those between W_2 and F_1 follow from Eqs. (11) and (13). We are only adding probabilities and no contradiction appears.

III. MONITORING THE LABORATORIES WITH NONINVASIVE MINIMALLY DISTURBING MEASUREMENTS

We work out here noninvasive minimally disturbing schemes that allow an external observer to gain information on the quantum states of laboratory systems or of a laboratory itself. In order to be noninvasive, the readout of the measurement cannot change the outcomes or the outcome probabilities of the friends or the external agents W . Any type of fuzzy meter [25] with a weak coupling strength could be considered for our scheme. For definiteness we will use weak measurements, a minimally disturbing scheme that has been widely investigated recently. The underlying conceptual issue—whether any type of interaction, be it noninvasive, fundamentally changes the closed character of the laboratories L_i —will be discussed below in Sec. IV.

A. Weak measurements

The scheme we use is based on weak measurements, originally introduced by Aharonov, Albert, and Vaidman [17] (see, e.g., [26,27] for an introduction to the main concepts). The idea is to extract information about a given property, represented by an observable \hat{A} on a system that evolves from a prepared initial state $|\zeta(t=0)\rangle$ towards the final eigenstate $|b_f\rangle$ obtained after measuring a different observable \hat{B} . This is done by creating at some time t_w a von Neumann interaction as per Eq. (3), but with a very weak coupling constant (that will be labeled γ instead of g) so that the ensuing evolution operator can be expanded to first order as $1 - i\gamma\hat{A}\hat{P}/\hbar$ (as above, \hat{P} is an observable of the probe to which \hat{A} is coupled to). Then we let the system, now effectively entangled with the weakly coupled probe, evolve up to t_f when a standard projective measurement of \hat{B} is made.

Up to some time $t_w < t < t_f$, the system-probe state evolves according to

$$U(t, t_w)(1 - i\gamma\hat{A}\hat{P})U(t_w, 0)|\zeta(t=0)\rangle|\varphi(t=0)\rangle \quad (19)$$

where $|\varphi\rangle$ denotes the quantum state of the probe and U is the system evolution operator (for simplicity we assume no evolution for the probe). Assume that the result of the \hat{B} measurement at t_f is b_f and $|b_f\rangle$ denotes the corresponding eigenstate. It is straightforward to show (see, e.g., [27]) that the weakly coupled probe state after b_f has been obtained is

$$|\varphi(t_f)\rangle = \langle b_f|\zeta(t_f)\rangle \exp(-i\gamma A_f^w \hat{P})|\varphi(t=0)\rangle \quad (20)$$

where

$$A_f^w = \frac{\langle b_f | U(t_f, t_w) \hat{A} | \zeta(t_w) \rangle}{\langle b_f | \zeta(t_f) \rangle} \quad (21)$$

is known as the weak value of \hat{A} . A_f^w can be complex. If $|\varphi(t=0)\rangle$ is a Gaussian in position space and \hat{P} is the momentum Eq. (20) states that the initial probe state is shifted by $\gamma \text{Re} A_f^w$. Since γ is small the shift will be very small—the initial and final pointer states substantially overlap and in practice the measurements must be repeated in order to gather enough statistics.

Note that through the weak value the final probe state depends on the initial quantum state of the system $|\zeta\rangle$ but also on the final projective measurement that is made on the system. This is a property that we will see to be useful in the Wigner-friend scenarios. Note also that the weak interaction modifies the evolution of the system state to first order in γ [compare Eq. (19) with the free evolution $|\zeta(t)\rangle = U(t, 0)|\zeta(t=0)\rangle$]. In particular the probabilities to obtain an outcome for any subsequent measurement are not modified to first order (they depend on γ^2). In this sense the weak measurement of \hat{A} is minimally disturbing.

B. Monitoring the atoms

1. WFS

Let us consider the original Wigner setup, with a laboratory L_1 , a friend F_1 , and an external agent W_1 , and for definiteness assume L_1 is initially in state $|L_1(t=0)\rangle$ as given by Eq. (8), except for the spin s_2 that plays here no role and will be left out. We further assume W_1 measures L_1 in the σ_z basis as in Sec. II B above, with a pointer initially in the ready state $|M_1(t=0)\rangle$. In addition W_1 has a probe with momentum variable \hat{P}_{W_1} , initially in state $\varphi_1(X_{W_1})$ that will be weakly coupled to the spin through an interaction Hamiltonian $\gamma(t)\hat{\omega}_1\hat{P}_{W_1}$ (see Fig. 1), where $\hat{\omega}_1$ is an observable related to the spin (e.g., a spin component along a conveniently chosen axis, or a spin projection operator). The weak interaction takes place before F_1 measures her spin in the σ_x basis, and the initial state evolves for $t < t_1$ to

$$|L_1(t)\rangle|\varphi_1(t)\rangle = e^{-ig\sigma_x\hat{P}/\hbar} e^{-i\gamma\hat{\omega}_1\hat{P}_{W_1}/\hbar} |s_1\rangle |m_1(t=0)\rangle \times |e_1(t=0)\rangle \varphi_1(X_{W_1}) \quad (22)$$

where $|s_1\rangle$ is given by Eq. (9). The unitary with the weak coupling is expanded to first order and the standard von Neumann term is expanded in the $|\pm_1\rangle$ basis, yielding

$$\left[\frac{1}{\sqrt{3}} |m_1+\rangle |+\rangle \varphi_1(X_{W_1} - \gamma\omega_1^+) + \frac{\sqrt{2}}{\sqrt{3}} |m_1-\rangle |-\rangle \varphi_1(X_{W_1} - \gamma\omega_1^-) \right] |e_1(t=0)\rangle. \quad (23)$$

We have not included the pointer state $|M_1\rangle$ that plays no role at this stage, and we have set $\gamma \equiv \int \gamma(t') dt'$ as in the usual von Neumann measurement [see below Eq. (5)]. The quantities

$$\omega_1^\pm \equiv \frac{\langle \pm | \hat{\omega}_1 | s_1 \rangle}{\langle \pm | s_1 \rangle} \quad (24)$$

are the weak values. In most instances the exact choice of $\hat{\omega}_1$ is not important, as long as it is known to W_1 so he can infer the weak value by measuring the probe. Indeed, the dependence of the weak value on the initial and final states $|s_1\rangle$ and $|\pm\rangle$ is what will matter most in the present context. However when the choice of $\hat{\omega}_1$ becomes significant we will mention it. Note that some specific choices can be particularly useful in some circumstances—for instance $\hat{\omega}_1 = \Pi_+ \equiv |+\rangle\langle +|$ will leave the probe untouched when the spin outcome is $-$, since by Eq. (24) we then have $\langle - | \Pi_+ | s_1 \rangle = 0$.

If we assume global collapse upon F_1 's measurement then Eq. (23) leads to

$$|\mathcal{L}_1(t_1)\rangle \varphi_1(t_1) = \begin{cases} |\mathcal{L}_1+\rangle \varphi_1(X_{W_1} - \gamma\omega_1^+) \\ |\mathcal{L}_1-\rangle \varphi_1(X_{W_1} - \gamma\omega_1^-) \end{cases} \quad (25)$$

with $|\mathcal{L}_1\pm\rangle = |\pm_1\rangle |m_1\pm\rangle |e_1\pm\rangle$. W_1 finally measures L_1 in the σ_z basis by coupling his pointer in the ready state $|M_1(t=0)\rangle$. This is a standard projective measurement that leaves the pointer in state $|M_1 \uparrow\rangle$ or $|M_1 \downarrow\rangle$. The probe states are shifted by the weak values ω_1^\pm correlated with the states $|\mathcal{L}_1\pm\rangle$. W_1 therefore obtains one of the following final states:

$$\begin{cases} |M_1 \uparrow\rangle \langle \mathcal{L}_1 \uparrow | \mathcal{L}_1+\rangle \varphi_1(X_{W_1} - \gamma\omega_1^+) \\ |M_1 \uparrow\rangle \langle \mathcal{L}_1 \uparrow | \mathcal{L}_1-\rangle \varphi_1(X_{W_1} - \gamma\omega_1^-) \\ |M_1 \downarrow\rangle \langle \mathcal{L}_1 \downarrow | \mathcal{L}_1+\rangle \varphi_1(X_{W_1} - \gamma\omega_1^+) \\ |M_1 \downarrow\rangle \langle \mathcal{L}_1 \downarrow | \mathcal{L}_1-\rangle \varphi_1(X_{W_1} - \gamma\omega_1^-) \end{cases}. \quad (26)$$

If instead unitary evolution is assumed, Eq. (23) becomes

$$\frac{1}{\sqrt{3}} |\mathcal{L}_1+\rangle \varphi_1(X_{W_1} - \gamma\omega_1^+) + \frac{\sqrt{2}}{\sqrt{3}} |\mathcal{L}_1-\rangle \varphi_1(X_{W_1} - \gamma\omega_1^-). \quad (27)$$

Note that the laboratory has become slightly entangled with the noninvasive probe [slightly in the sense that γ is small so that the zeroth-order term is the product state $(|\mathcal{L}_1+\rangle + \sqrt{2}|\mathcal{L}_1-\rangle)\varphi_1(X_{W_1})/\sqrt{3}$]. Finally W_1 measures L_1 in the σ_z basis, so writing Eq. (27) in that basis yields

$$|\mathcal{L}_1 \uparrow\rangle \left[\frac{1}{\sqrt{6}} \varphi_1(X_{W_1} - \gamma\omega_1^+) - \frac{1}{\sqrt{3}} \varphi_1(X_{W_1} - \gamma\omega_1^-) \right] + |\mathcal{L}_1 \downarrow\rangle \left\{ \frac{1}{\sqrt{6}} \varphi_1(X_{W_1} - \gamma\omega_1^+) + \frac{1}{\sqrt{3}} \varphi_1(X_{W_1} - \gamma\omega_1^-) \right\}. \quad (28)$$

Hence when W_1 obtains \uparrow (or \downarrow) the probe will be found in the states given by the term between $[\cdot]$ and $\{\cdot\}$, respectively, indicating in both cases a superposition of the spin in states \pm .

The first important result we have obtained is that a weakly coupled probe allows us to empirically discriminate the mixed pointer states obtained from Eq. (26) from the interfering patterns of Eq. (28), whereas as we have noted at the end of Sec. II A this is in practice impossible for macroscopically interfering states. That said, according to Eq. (26) the information W_1 gets from his probe φ_1 if global collapse is assumed matches the outcome that F_1 announces she obtained: there is perfect consistency. Instead if pure unitary evolution is assumed [Eq. (28)] the probe states read by W_1 indicate a superposition of the spin in states $+$ and $-$: this is inconsistent

with the fact that F_1 obtained a definite outcome. Therefore, if we accept the existence of a noninvasive coupling (this contentious point—under which conditions this coupling contradicts the closed character of L_1 , given that in practical terms it does not change the dynamical processes taking place inside L_1 —will be tackled in Sec. IV), the original Wigner-friend scenario leads to a contradiction. The origin of this contradiction is obvious: W_1 's probe is noninvasively coupled to the same spin whose state collapses when F_1 completes her measurement. Recall that a collapse on the entangled state (23) occurs when F_1 measures the position of her pointer $|m_{1\pm}\rangle$: depending on the value of X , either $m_{1+}(X)$ or $m_{1-}(X)$ will be nonzero. So if F_1 obtains an outcome, the superposition given by Eq. (27) cannot occur.

2. EWFS

We now apply the same method to the extended Wigner-friend scenario of Sec. II B. The external agents W_1 and W_2 couple a probe with a weak interaction to the atoms in laboratories L_1 and L_2 , respectively (see Fig. 2). For W_1 , the process is exactly the one we have described immediately above. W_2 also has a probe in state $\varphi_2(X_{W_2})$ that couples to the spin s_2 of atom 2 once it is received by F_2 .

For definiteness assume the coupling Hamiltonian to be given by $\gamma(t)\hat{\omega}_2\hat{P}_{W_2}$ (where the notation follows the pattern introduced above).

a. Global collapse upon the friend's measurement. After F_1 completes her measurement, a collapse of her meter $m_1(X)$ leads to a weakly coupled probe in states $\varphi_1(X_{W_1} - \gamma\omega_1^+)$ or $\varphi_1(X_{W_1} - \gamma\omega_1^-)$. Concurrently L_2 is in either of the states given by Eq. (12). With the weakly coupled probe initially in state $\varphi_2(X_{W_2})$, the same analysis² leading to Eq. (25) can be done for L_2 , yielding

$$\begin{aligned} & |\mathcal{L}_2(t_2)\rangle|\varphi_2(t_2)\rangle|\mathcal{L}_1(t_2)\rangle|\varphi_1(t_2)\rangle \\ &= \begin{cases} |\mathcal{L}_2\uparrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^+)|\mathcal{L}_1-\rangle\varphi_1(X_{W_1} - \gamma\omega_1^-) \\ |\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^+)|\mathcal{L}_1-\rangle\varphi_1(X_{W_1} - \gamma\omega_1^-) \\ |\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^\downarrow\downarrow)|\mathcal{L}_1+\rangle\varphi_1(X_{W_1} - \gamma\omega_1^+) \end{cases} \quad (29) \end{aligned}$$

with $|\mathcal{L}_2\uparrow\rangle = |\uparrow_2\rangle|m_2\uparrow\rangle|e_2\uparrow\rangle$ and where

$$\omega_v^{s_f, s_c} \equiv \frac{\langle s_f | \hat{\omega}_2 | s_c \rangle}{\langle s_f | s_c \rangle}. \quad (30)$$

$\hat{\omega}_2$ is a spin operator or projector related to atom 2, $|s_c\rangle = |\downarrow\rangle$ or $|+\rangle$ is the spin state of the atom received by F_2 , and $|s_f\rangle = |\downarrow\rangle$ or $|\uparrow\rangle$ is the outcome of F_2 's measurement. It can be checked that $\omega_2^{\downarrow\downarrow}$ is the expectation value $\langle \downarrow | \hat{\omega}_2 | \downarrow \rangle$. Finally the external observers W_i make their measurements. For W_1 the outcomes were given in Eq. (26). W_2 measures L_2 and the outcomes are $|\mathcal{L}_2\pm\rangle$ as given by Eq. (17). His pointer $|M_2\rangle$ is

thus correlated with

$$\begin{aligned} & \begin{cases} |M_2+\rangle\langle\mathcal{L}_2+\mid\mathcal{L}_2\uparrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+}) \\ |M_2+\rangle\langle\mathcal{L}_2+\mid\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow+}) \\ |M_2+\rangle\langle\mathcal{L}_2+\mid\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow\downarrow}) \end{cases} \\ & \times \begin{cases} |M_2-\rangle\langle\mathcal{L}_2-\mid\mathcal{L}_2\uparrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+}) \\ |M_2-\rangle\langle\mathcal{L}_2-\mid\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow+}) \\ |M_2-\rangle\langle\mathcal{L}_2-\mid\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow\downarrow}) \end{cases}. \quad (31) \end{aligned}$$

The joint outcomes for W_1 and W_2 are independent and can be obtained from Eqs. (26), (29), and (31). They can read from their probes φ_i the outcomes obtained by their respective friends. For example let us look at the case $W_1 = \uparrow$, $W_2 = -$ which as we saw lies at the basis of the contradiction [see points (iv) and (v) below Eq. (18)]. $W_1 = \uparrow$ can be obtained with either $F_1 = +$ or $-$ [left-hand side of Eq. (26)]. $W_2 = -$ can be obtained with either $F_2 = \uparrow$ or \downarrow with F_1 having obtained either $+$ or $-$. Hence none of the points (ii)–(iv) hold in this case. For instance if we take $\hat{\omega}_2 = |-\rangle\langle-|$, then $\omega_2^{\uparrow+} = \omega_2^{\downarrow+} = 0$ and W_2 will see a shift in the weak probe only when the state received by F_2 was $|\downarrow\rangle$ (and hence F_1 obtained $+$), while, for $\hat{\omega}_2 = |\uparrow\rangle\langle\uparrow|$, $\omega_2^{\downarrow+} = \omega_2^{\downarrow\downarrow} = 0$ only when F_2 received the atom in state $|+\rangle$ and obtained the outcome \uparrow .

b. Unitary evolution of the laboratories. If L_1 and L_2 are assumed to evolve unitarily (as per the ‘‘universality’’ assumption) the weak couplings with W_1 's and W_2 's noninvasive probes are taken into account by the evolution operator $e^{-i\gamma\hat{\omega}_2\hat{P}_{W_2}/\hbar}e^{-i\gamma\hat{\omega}_1\hat{P}_{W_1}/\hbar}$ so that the overall wave function given by Eq. (15) becomes with this coupling

$$\begin{aligned} |\Psi(t_2)\rangle &= \frac{1}{\sqrt{3}}|\mathcal{L}_1+\rangle|\mathcal{L}_2\downarrow\rangle\varphi_1(X_{W_1} - \gamma\omega_1^+)\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow\downarrow}) \\ &+ \frac{1}{\sqrt{3}}|\mathcal{L}_1-\rangle\varphi_1(X_{W_1} - \gamma\omega_1^-)[|\mathcal{L}_2\uparrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+}) \\ &+ |\mathcal{L}_2\downarrow\rangle\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow+})]. \quad (32) \end{aligned}$$

Let us look at the correlation between outcomes as by points (i)–(v) below Eq. (18).

(i') Point (i) [$F_2 = \uparrow \Rightarrow F_1 = -$] is verified by construction so it holds for all the cases we have examined so far.

(ii') If we write Eq. (32) in the basis $|\mathcal{L}_1\pm\rangle|\mathcal{L}_2\pm\rangle$, the term proportional to $|\mathcal{L}_2-\rangle$ between the square brackets [...] is $-\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+}) + \varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow+})$ for which the leading term vanishes. We thus neglect it and we are left with the sole term $|\mathcal{L}_1+\rangle|\mathcal{L}_2-\rangle\varphi_1(X_{W_1} - \gamma\omega_1^+)\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow\downarrow})$ coming from the upper line. This implies that (ii) holds [$W_2 = - \Rightarrow F_1 = +$] and the weak probe φ_2 further indicates to W_2 that F_2 received and measured the atom spin in state \downarrow .

(iii') We write Eq. (32) in the basis $|\mathcal{L}_1\uparrow\rangle|\mathcal{L}_2\uparrow\rangle$; to leading order, we are left with $|\mathcal{L}_1\uparrow\rangle|\mathcal{L}_2\uparrow\rangle\varphi_1(X_{W_1} - \gamma\omega_1^-)\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+})$. Hence (iii) holds [$W_1 = \uparrow \Rightarrow F_2 = \uparrow$] and we have the information through the weak probes that the outcome for F_1 was $-$.

(iv') The information carried by the weak probes in (ii') and (iii') strengthens the apparent contradiction (iv) pointed

²The term $\gamma|m\uparrow\rangle$ that appears due to the weak coupling is neglected since the probability to detect this pointer state is second order in γ .

out above: (ii') tells us that $W_2 = - \Rightarrow (F_1 = + \text{ and } F_2 = \downarrow)$ while from (iii') $W_1 = \uparrow \Rightarrow (F_1 = - \text{ and } F_2 = \uparrow)$.

(v') Eq. (32) is written in the basis $|\mathcal{L}_1 \uparrow \downarrow\rangle |\mathcal{L}_2 \pm\rangle$. The term yielding the outcomes $W_1 = \uparrow, W_2 = -$ is factored by

$$\begin{aligned} & |\mathcal{L}_1 \uparrow\rangle |\mathcal{L}_2 -\rangle \{ \varphi_1(X_{W_1} - \gamma \omega_1^+) \varphi_2(X_{W_2} - \gamma \omega_2^{\downarrow\downarrow}) \\ & + \varphi_1(X_{W_1} - \gamma \omega_1^-) \varphi_2(X_{W_2} - \gamma \omega_2^{\uparrow+}) \\ & - \varphi_1(X_{W_1} - \gamma \omega_1^-) \varphi_2(X_{W_2} - \gamma \omega_2^{\downarrow+}) \}. \end{aligned} \quad (33)$$

Here the terms that can be neglected to leading order in γ depend on the choice of $\hat{\omega}_1$ and $\hat{\omega}_2$. In typical cases, $\omega_2^{\downarrow\downarrow} \neq \omega_2^{\downarrow+}$ and the second line becomes negligible to leading order in γ leaving the term $\varphi_1(X_{W_1} - \gamma \omega_1^+) \varphi_2(X_{W_2} - \gamma \omega_2^{\downarrow\downarrow})$ corresponding to $F_1 = +$ and $F_2 = \downarrow$. However if $\omega_2^{\downarrow\downarrow} = \omega_2^{\downarrow+}$ [e.g., for $\hat{\omega}_2 = \sigma_z$] Eq. (33) becomes to leading order

$$\begin{aligned} & |\mathcal{L}_1 \uparrow\rangle |\mathcal{L}_2 -\rangle \frac{1}{2} [\varphi_1(X_{W_1} - \gamma \omega_1^+) \varphi_2(X_{W_2} - \gamma \omega_2^{\downarrow\downarrow}) \\ & + \varphi_1(X_{W_1} - \gamma \omega_1^-) \varphi_2(X_{W_2} - \gamma \omega_2^{\uparrow+})]. \end{aligned} \quad (34)$$

In this case, when $W_1 = \uparrow, W_2 = -$ is obtained, the probes are, like in Eq. (28) for the original setup, in a state of superposition of states in which $(F_1 = +, F_2 = \downarrow)$ and $(F_1 = -, F_2 = \uparrow)$. The probes do not indicate to the external observers W_i that a specific outcome has been obtained.

We will see below that if the probes are coupled to the pointer or the environment inside each L_i (hence after each friend's measurement has occurred), the behavior represented by Eq. (34) is the generic one. We therefore defer its discussion to the paragraph immediately below. We wish to emphasize here that we are combining projective measurements and weak measurements. The projective measurements carried out in points (ii'), (iii'), and (v') are not compatible (since they involve noncommuting observables). The weak measurements are noninvasive unitary interactions that give information on an observable. Therefore the problem of compatibility with the final projective measurement does not arise—this is actually one of the reasons weak measurements have been found to be useful—but it should be stressed that weak values are conditioned on how the amplitudes interfere upon a final projective measurement.

C. Monitoring inside the laboratory

We investigate the case in which the external agents W_i weakly couple an external probe to an element inside L_i rather than to the atoms after F_i 's measurement has occurred. By "element" we mean a system that carries information about the measurement, such as the pointers, or the environment inside L_i . In the version in which the friend's measurement induces an overall collapse a weak probe of this type would be redundant, as it would simply confirm the outcome obtained by each of the friends' projective measurements. So we will only consider weak coupling in the unitary evolution scheme that would follow from Wigner's "universality" assumption. We will simplify the notation by using \updownarrow to stand for " \uparrow or \downarrow ".

1. WFS

For definiteness, assume the weakly coupled probe placed outside L_1 interacts, through a minute hole, with the mea-

surement apparatus in the otherwise sealed laboratory. This interaction takes place after the spin interacts with the friend's pointer. Hence, with $|s_1\rangle$ yet again given by Eq. (9), the quantum state after the weak interaction is similar to Eq. (22),

$$\begin{aligned} |L_1(t)\rangle |\varphi_1(t)\rangle &= e^{-iy\hat{\omega}_1 \hat{p}_{W_1}/\hbar} e^{-ig\sigma_x \hat{p}/\hbar} |s_1\rangle m_1(X, t=0) \\ &\times |e_1(t=0)\rangle \varphi_1(X_{W_1}), \end{aligned} \quad (35)$$

but $\hat{\omega}_1$ is now an observable coupled to the measurement apparatus after the spin measurement has taken place and the apparatus states have become $|m_1 \pm\rangle$. W_1 measures L_1 in the $|\mathcal{L}_1 \updownarrow\rangle$ basis, leading to

$$\begin{aligned} & \left[\frac{1}{\sqrt{3}} \langle \mathcal{L}_1 \updownarrow | \mathcal{L}_1 + \rangle \varphi_1(X_{W_1} - \gamma \omega_1^{\uparrow+}) \right. \\ & \left. + \frac{\sqrt{2}}{\sqrt{3}} \langle \mathcal{L}_1 \updownarrow | \mathcal{L}_1 - \rangle \varphi_1(X_{W_1} - \gamma \omega_1^{\uparrow-}) \right] |M_1 \updownarrow\rangle. \end{aligned} \quad (36)$$

Since $\langle \mathcal{L}_1 \updownarrow | \mathcal{L}_1 \pm \rangle = \pm 1/\sqrt{2}$, we have essentially the same result as when the probe was weakly coupled to the spin [see Eq. (28)]. For both outcomes \updownarrow , the prediction according to unitary evolution is that W_1 will find his weakly coupled probe φ_1 as indicating a superposition of the spin in states \pm . So W_1 can again conclude there is an inconsistency between the information obtained by measuring φ_1 and the definite outcome F_1 announced she measured. Hence W_1 can discriminate empirically this prediction obtained by assuming unitary evolution from the mixed state expected if the friend's spin measurement induces a global collapse.

Note that in principle the weakly interacting probe can be coupled to the environment inside L_1 , or more precisely to the fraction of the environment affected by the friend's measurement (assuming we can factor the environment into an interacting part, described by $|e_1\rangle$ and a noninteracting part that plays no role), for example the photons that carry the measurement outcome information. Equation (35) still holds, with $\hat{\omega}_1$ now coupled to the environment that can quickly be described by states $|e_1 \pm\rangle$ correlated with the measurement apparatus states $|m_1 \pm\rangle$ (we are neglecting here the dynamics during the very short time it takes for the environment states to become orthogonal in the pointer state basis). While remaining orthogonal, the states $|e_1 \pm\rangle$ change rapidly with time (see [15] for a model of the environment in the WFS scenario), so that the weak values $\omega_1^{\uparrow\pm}$ will depend on the timing of the weak interaction. Hence while Eq. (36) still holds if unitary evolution is assumed and the conclusions for W_1 remain identical to the ones that hold for a probe coupled to the measurement apparatus, it is difficult to make general statements on the practical realizability of the weak probe reading without a specific model of the environment.

2. EWFS

Now both W_1 and W_2 couple a probe to the measuring apparatus inside L_1 and L_2 . The noninvasive probes interact with the respective measurement apparatus after the spin measurements have taken place, the apparatus being, respectively, in states $|m_1 \pm\rangle$ and $|m_2 \pm\rangle$. Assuming unitary evolution, the

quantum state reads

$$|\Psi(t_2)\rangle = e^{-i\gamma\hat{\omega}_1\hat{P}_{W_1}/\hbar} e^{-i\gamma\hat{\omega}_2\hat{P}_{W_2}/\hbar} |L_{12}(t_2)\rangle \varphi_1(X_{W_1}) \varphi_2(X_{W_2}) \quad (37)$$

where $|L_{12}\rangle$ is given by Eq. (15) or Eq. (18). The same projective measurements as in Sec. III B 2 b are performed by W_1 and W_2 on $|\Psi(t_2)\rangle$. The weak values are now defined relative to the pointer states $|m_1\rangle$ or $|m_2\rangle$:

$$\omega_i^{\beta\alpha} = \frac{\langle m_i\beta|\hat{\omega}_i|m_i\alpha\rangle}{\langle m_i\beta|m_i\alpha\rangle}, \quad (38)$$

where α labels the state in the measurement basis of the friend F_i and β labels a state in the basis in which W_i measures. Points (i'')–(iv'') are the same as in the case in which the noninvasive probe was coupled directly to the spins. Point (v'') is slightly different.

(i'') Again, this point [$F_2 = \uparrow \Rightarrow F_1 = -$] is verified by construction.

(ii'') If we take the scalar products $\langle \mathcal{L}_1 \pm | \langle \mathcal{L}_2 - |$ to leading order only the term $\langle \mathcal{L}_1 + | \langle \mathcal{L}_2 - | \Psi$ is nonzero with the probes left in states $\varphi_1(X_W - \gamma\omega_1^{\uparrow+})\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow-})$, implying [$W_2 = - \Rightarrow F_1 = +$ and $F_2 = \downarrow$], exactly as in (ii').

(iii'') Computing the scalar products $\langle \mathcal{L}_1 \uparrow | \langle \mathcal{L}_2 \downarrow | \Psi$ leaves the leading term with the probes in state $\varphi_1(X_W - \gamma\omega_1^{\uparrow-})\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+})$. Hence (iii'') holds [$W_1 = \uparrow \Rightarrow F_2 = \uparrow$ and $F_1 = -$].

(iv'') We have the same contradiction as in (iv').

(v'') The projection $\langle \mathcal{L}_1 \uparrow | \langle \mathcal{L}_2 - |$ leads to

$$\begin{aligned} & \varphi_1(X_W - \gamma\omega_1^{\uparrow+})\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow-}) + \varphi_1(X_W - \gamma\omega_1^{\uparrow-})\varphi_2 \\ & \times (X_{W_2} - \gamma\omega_2^{\uparrow+}) - \varphi_1(X_W - \gamma\omega_1^{\uparrow-})\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow-}). \end{aligned} \quad (39)$$

To lowest order in γ , this can be put under the form

$$\begin{aligned} & \frac{1}{2} [\varphi_1(X_W - \gamma\omega_1^{\uparrow+})\varphi_2(X_{W_2} - \gamma\omega_2^{\downarrow-}) \\ & + \varphi_1(X_W - \gamma\omega_1^{\uparrow-})\varphi_2(X_{W_2} - \gamma\omega_2^{\uparrow+})]. \end{aligned} \quad (40)$$

This is similar to Eq. (34) of point (v'), but now holds for any choice of $\hat{\omega}_1$ and $\hat{\omega}_2$ concerning the measurement pointers: the outcomes $W_1 = \uparrow$, $W_2 = -$ can be obtained with nonzero probability, and when this happens the probes φ_i indicate that the friends' pointers were in a state of superposition of outcomes ($F_1 = +$, $F_2 = \downarrow$) and ($F_1 = -$, $F_2 = \uparrow$). The probes act here as a witness of the entanglement of $|L_{12}(t_2)\rangle$ [Eq. (18)]. Since W_1 and W_2 can communicate, a clever choice of $\hat{\omega}_1$ and $\hat{\omega}_2$ would allow them to determine that their probes have been entangled by the unitary evolution. For the agents W_i such a finding would be inconsistent given that the friends announced they have obtained definite outcomes (in which case their states should not be entangled).

IV. DISCUSSION

A. Summary

Let us start by summarizing the results we have obtained.

In the original WFS, if the external agent W assumes L evolves unitarily, he will run into a contradiction: a noninvasive probe coupled (and therefore entangled with) the friend's atom or her pointer would display a signature of the

macroscopic superposition. Indeed W can measure the noninvasive probe irrespective of whether F tells him she obtained a definite outcome. If unitary evolution is assumed, quantum theory predicts that the probe will indicate a superposition of the atom spin or pointer states: this is incompatible with F having obtained a definite outcome. Instead if global collapse is assumed, quantum theory predicts that the weakly coupled probes will indicate the outcome obtained by F . Note that while measuring the state of an entire macroscopic laboratory in order to discriminate a mixed state from unitary evolution is experimentally out of reach, measuring noninvasive weakly coupled probes could be achieved with present-day technologies.

In the EWFS, assuming unitary evolution results in entangled probes. This is a signature that the atoms or the pointers to which the probes were coupled are entangled. But this is impossible since F_1 and F_2 have announced they have obtained definite outcomes. This inconsistency when unitary evolution is assumed is similar to the one for the original WFS, with the superposition in the extended version taking the specific form of entanglement. This inconsistency is similar to the logical contradiction obtained by Frauchiger and Renner [8] [see points (i)–(v) below Eq. (18)], except that here this contradiction appears as information the agents could read on their probes. The agents W_i would therefore conclude by reading their probes that they were wrong in assuming unitary evolution, as the probes are consistent with the friends' announcements only if global collapse is assumed.

B. Closed laboratories and noninvasive measurements

According to the summary we have just given, an external agent should apply the projection postulate (“global collapse”) and not assume unitary evolution for a closed laboratory in which a measurement takes place, as this will lead him to an observable contradiction. One may question however whether the laboratories can still be qualified as closed when a very weak interaction is allowed with the external world.

From a purely logical viewpoint, this objection makes sense: an interaction, however small, breaks the purported independence of the closed laboratory from the external world. The very idea of attempting to discriminate observer-dependent from observer-independent facts [7] seems indeed ruined. This is a crucial point in Deutsch's approach [2] that employs a notion called “kinematic independence” to assume two subsystems can evolve as independent worlds.

From the physical viewpoint we adopt here, however, this logical independence is a fiction: a laboratory can never be perfectly isolated. Operationally, a laboratory is said to be sealed if the inevitable interactions with the rest of the universe can be neglected relative to the effects that are investigated. Here, the sealed character of the laboratory implies that the operations carried out by the friends while running their experiment inside the laboratory cannot be detected by any type of measurement in the external universe, except through the channel opened for the noninvasive probe. The noninvasive interaction itself is asymptotically small and does not change in any essential way the dynamical evolution inside the laboratories: in particular the predictions made

inside the laboratories and the outcome probabilities are not modified. From this perspective, adding noninvasive probes does not fundamentally alter the physical interactions and processes inside the laboratories. Our operational assumption is that such a probe simply records the local value of the system observable it is interacting with, with minimal change to the system's state. This is how the effects of weak measurements are generally interpreted [27].

Another way to tackle this question, that we examined very recently [28], is to associate a measurement outcome with the existence of a stable material record. For practical or fundamental reasons, on which we will not speculate here, a stable material record is immune to linear superpositions (in Feynman's words, amplitudes are lost and play no further role [29]). We have argued [28] that the material record includes F's memory (or any printout that would record the measurement), so that if the stable records produced by F's measurement could be erased (through unitary evolution) by W's subsequent measurement, there is no way to ensure F's measurement has even existed. This is precisely the situation measured by the weakly coupled probes when unitary evolution is assumed: there is a lack of records that measurements should produce. Hence within this perspective, there can be no observer-dependent facts even if the sealed laboratory is assumed to evolve unitarily, unless one supposes that the presence of a noninvasive interaction prevents the formation of a stable material record for the friends that would take place for totally sealed laboratories (admittedly, this sounds quite unlikely given the properties of weak measurements).

C. Measurements: Quantum mechanics is consistent in its inconsistency

Wigner-friend scenarios are at the crossroads of two distinct but related problems: what constitutes a measurement, and whether quantum mechanics holds in the macroscopic limit. Any account of the WFS and the problems that it might raise will hinge on one's stance on these issues, which to a large extent depends on the interpretation one has in mind (this includes a set assumptions some of which might be clearly expressed explicitly, along with others that might be fuzzily implicitly presupposed).

Wigner, shortly after he published the WFS argument, endorsed more clearly the idea that quantum mechanics applied to macroscopic bodies but not to conscious observers, for whom unitary evolution "is not credible" (Sec. 12 of [5]). Obviously, this position conflicts with the idea that quantum mechanics should apply in a "universal" manner to macroscopic systems, the assumption that is at the root of the WFS and more particularly in the EWFS of [7,8]: the agents W_i would discard that the friends evolved unitarily once they completed their measurements (and therefore no contradiction will be obtained between the agents).

Purely unitary approaches (such as the many-worlds interpretation) uphold that the projection postulate makes quantum theory nonuniversal and logically inconsistent [2]. In such approaches, whether the friends and the agents W evolve or not in different worlds, observing or not different facts and how different facts can be reconciled when all the agents meet, depends on subtle assumptions concerning the branching

mechanism (for a recent discussion detailing how branching issues in Einstein-Podolsky-Rosen or Greenberger-Horne-Zeilinger setups might lead agents evolving independently to reconcile their viewpoints, see [30]).

In most other approaches, the projection postulate for all the agents must be implicitly or explicitly employed. It is then straightforward to see [10] that, while the application of the projection postulate remains ambiguous, all the agents involved in the EWFS can agree beforehand on how to apply it and avoid reaching contradictions. It has also been argued [16] that the sole unitary coupling is insufficient to define a measurement—unitary evolution leads at best to the pre-measurement step of the measurement process, but this does not suffice in order to obtain recorded outcomes. Then all the agents will agree on their measurements. Note that the collapse can be implicit, like when taking the partial trace over the environment in a decoherence based approach (a detailed decoherence based account of the WFS and EWFS has recently been given [15]), or in the projectors defining consistent histories (such an approach for the EWFS has also been given [13]).

For approaches in which the state vector is considered to embody an objective physical reality, global collapse upon measurement for all the agents is the only option. This is obvious in dynamical collapse theories [31] in which a nonlinear term is added to the Schrödinger equation, preventing unitary evolution. While we do not know of any detailed rendering of the WFS with dynamical collapse, once the measurement is completed our results assuming global collapse in Sec. III will be recovered. In the de Broglie-Bohm approach, the dual particle-wave ontology renders the effective collapse somewhat less obvious, as the particle is detected but the wave function—assumed to be unique for all the agents—presents (in most versions of this approach) empty branches that still evolve, creating an ambiguity over when they remain dynamically relevant [32,33]. While slightly different, the recent de Broglie-Bohm accounts of the WFS and EWFS [9,11,12] all agree that particle detection by the friends changes the Bohmian particle configurations and trajectories for the external agents W, so that no contradiction is obtained.

The conclusion we can draw is that quantum mechanics is consistent in its inconsistency: even if we have good reasons to believe that quantum mechanics should apply to macroscopic bodies, a special rule, involving a global effective collapse relying on an ambiguous "shifty split" [34], remains necessary. Quantum measurements are special and the theoretical accounts remain unsatisfactory, if not logically inconsistent [2,34], but operationally there is no inconsistency in accounting for the observed outcomes.

V. CONCLUSION

In this paper, we have investigated Wigner-friend scenarios in which the external agents make noninvasive measurements inside the otherwise sealed laboratories. In the original WFS, an external agent W who would want to test the friend's unitary evolution would not be able to infer anything by directly measuring the quantum states of the macroscopic laboratory. With a noninvasive probe weakly coupled to the friend's atom or measuring apparatus—with a coupling so weak that it

does not affect the outcomes or their probabilities inside the laboratory— W will be able to infer the state of the laboratory after the friend's measurement. Within the limitations of our scheme—in which the friend's laboratory cannot be said to be perfectly closed—unitary evolution can be empirically tested, and probably invalidated, since if this hypothesis is assumed, we are lead to a contradiction between the friend having obtained a definite outcome and the state of the probe.

For the extended version of the WFS recently introduced by Brukner [7] and Frauchiger and Renner [8], we have shown that the noninvasive probes given to each external agent would be entangled if the hypothesis held. But we

have seen that if the friends obtained definite outcomes, the probes cannot be entangled (this forms the basis of the Bell-type inequality that leads to the contradiction put forward in [8]). Again, in our framework with noninvasive probes, this becomes a question that can be given an experimental confirmation.

To conclude, we have investigated Wigner-friend scenarios with noninvasive probes. While there is room to support the idea that the measurement problem makes quantum theory logically inconsistent, we can rely on the noninvasive probes to avoid inconsistency at the operational level. This approach could moreover be confirmed by experiments.

-
- [1] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University, Princeton, NJ, 1983).
- [2] D. Deutsch, *Int. J. Theor. Phys.* **24**, 1 (1985).
- [3] E. P. Wigner, in *The Scientist Speculates*, edited by I. J. Good (Heinemann, London, 1962); for a short history of the paper, see R. S. Mackintosh, [arXiv:1903.00392](https://arxiv.org/abs/1903.00392).
- [4] H. Margenau and E. P. Wigner, *Philos. Sci.* **29**, 292 (1962).
- [5] E. P. Wigner, *Contemporary Physics* (Atomic Energy Commission, Vienna, 1969), Vol. 2, pp. 431–437.
- [6] E. P. Wigner, in *Science, Computers, and the Information Onslaught*, edited by D. M. Kerr *et al.* (Academic, New York, 1984), pp. 63–82.
- [7] C. Brukner, *Entropy* **20**, 350 (2018).
- [8] D. Frauchiger and R. Renner, *Nat. Commun.* **9**, 3711 (2018).
- [9] A. Sudbery, *Found. Phys.* **47**, 658 (2017).
- [10] F. Laloe, [arXiv:1802.06396](https://arxiv.org/abs/1802.06396).
- [11] A. Drezet, [arXiv:1810.10917](https://arxiv.org/abs/1810.10917).
- [12] D. Lazarovici and M. Hubert, *Sci. Rep.* **9**, 470 (2019).
- [13] M. Losada, R. Laura, and O. Lombardi, *Phys. Rev. A* **100**, 052114 (2019).
- [14] K.-W. Bong, A. Utreras-Alarcón, F. Ghafari, Y.-C. Liang, N. Tischler, E. G. Cavalcanti, G. J. Pryde, and H. M. Wiseman, *Nat. Phys.* (2020), doi: 10.1038/s41567-020-0990-x.
- [15] A. Relaño, *Phys. Rev. A* **101**, 032107 (2020).
- [16] M. Zukowski and M. Markiewicz, [arXiv:2003.07464](https://arxiv.org/abs/2003.07464).
- [17] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).
- [18] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University, Princeton, NJ, 1955).
- [19] N. Wiebe and L. E. Ballentine, *Phys. Rev. A* **72**, 022109 (2005).
- [20] A. Matzkin, *Phys. Rev. A* **84**, 022111 (2011).
- [21] D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1951).
- [22] G. C. Ghirardi and L. Marinatto, *Phys. Lett. A* **372**, 1982 (2008).
- [23] L. Hardy, *Phys. Rev. Lett.* **71**, 1665 (1993).
- [24] A. Fine, *Phys. Rev. Lett.* **48**, 291 (1982).
- [25] M. B. Mensky, *Quantum Measurements and Decoherence* (Kluwer, Dordrecht, 2000).
- [26] B. E. Y. Svensson, *Quanta* **2**, 18 (2013).
- [27] A. Matzkin, *Found. Phys.* **49**, 298 (2019).
- [28] A. Matzkin and D. Sokolovski, *Europhys. Lett.* **131**, 40001 (2020).
- [29] The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference, edited by C. M. DeWitt and D. Rickles, Edition Open Access (Berlin, Germany, 2017), see Feynman's transcripts in Ch. 23.
- [30] M. Waegell and K. J. McQueen, *Stud. Hist. Philos. Sci. B* **70**, 39 (2020).
- [31] G. C. Ghirardi, A. Rimini, and T. Weber, *Phys. Rev. D* **34**, 470 (1986).
- [32] P. R. Holland, in *Protective Measurement and Quantum Reality*, edited by S. Gao (Cambridge University, Cambridge, England, 2015), pp. 145–163.
- [33] A. Matzkin, *Found. Phys.* **39**, 903 (2009).
- [34] J. S. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University, Cambridge, England, 2004), Chap. 23.