

Multiple-time states and multiple-time measurements in quantum mechanics

Yakir Aharonov,^{1,2} Sandu Popescu,^{3,4} Jeff Tollaksen,² and Lev Vaidman¹

¹*Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, 69978 Tel Aviv, Israel*

²*Department of Physics, Computational Science, and Engineering, Schmid College of Science, Chapman University, 1 University Drive, Orange, California 92866, USA*

³*H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*

⁴*Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom*

(Received 13 December 2007; revised manuscript received 10 March 2009; published 12 May 2009)

We discuss experimental situations that consist of multiple preparation and measurement stages. This leads us to an alternative approach to quantum mechanics. In particular, we introduce the idea of multitime quantum states which are the appropriate tools for describing these experimental situations. We also describe multitime measurements and discuss their relation to multitime states. A consequence of our formalism is to put states and operators on an equal footing. Finally we discuss the implications of our approach to quantum mechanics for the problem of the flow of time.

DOI: [10.1103/PhysRevA.79.052110](https://doi.org/10.1103/PhysRevA.79.052110)

PACS number(s): 03.65.Ta

I. INTRODUCTION

The main aim of this paper is to introduce a type of quantum state, a “multiple-time state.” We will also discuss multiple-time measurements and introduce the notion of “multiple-time measurement states.”

The simplest situation, namely, two-time states (also called pre- and postselected states) was first discussed by Aharonov *et al.* [1] in 1964 and was extensively studied during the last two decades, both theoretically [2] and experimentally [3]. The idea of multitime measurements and the first steps toward multitime states were discussed by Aharonov and Albert in [4]. The present paper is based on ideas described in the (unpublished) Ph.D. theses of Vaidman and Popescu [5,6]. Similar questions were treated via different approaches by Griffiths [7], Gell-Mann and Hartle [8], Cramer [9], and Schulman [10].

From a mathematical point of view, the “state” of a physical system is nothing other than a compact description of all the relevant information we have about that system. The usual quantum state is perfectly suited for the simple situations studied routinely in quantum mechanics, namely, experiments that consist of a preparation stage followed by a measurement stage. The state $|\Psi\rangle$ (or the density matrix ρ , if appropriate) contains all the information. Based on it, we can predict the probabilities of any measurement. Of course, we may know much more about the preparation stage than what is encoded in the state, such as details about the measuring devices that were used or about the past history of the system, but as far as the measurement stage is concerned everything is encapsulated in $|\Psi\rangle$ (or ρ). It is in fact remarkable that for some systems only very few parameters are needed, such as three real numbers for a spin 1/2 particle, while we might know many more things about the preparation (such as the magnetic field that may have acted on the spin during its entire history).

In any case, while the usual quantum state is perfectly suitable for describing the standard experiment as discussed above, we can imagine more complex experiments that consist of many stages of preparation interspread with many

stages of measurement (Fig. 1). Multiple-time states refer to these situations.

To avoid any confusion, we want to emphasize from the outset that we do not want to modify quantum theory. Our results are totally and completely part of ordinary quantum mechanics. Furthermore, we want to make it clear that the ordinary formalism of quantum mechanics is perfectly capable of describing every experiment that we consider here, including experiments that consist of many preparation-measurement stages. The issue, however, is to get a convenient, compact and illuminating description; as we will show, multiple-time states are ideally suited tools for this purpose.

We can, of course, consider such complex experiments in classical physics as well. In that case however the experiment can always be decomposed into many elementary experiments, each involving a single preparation-measurement stage and there is effectively nothing interesting to note.

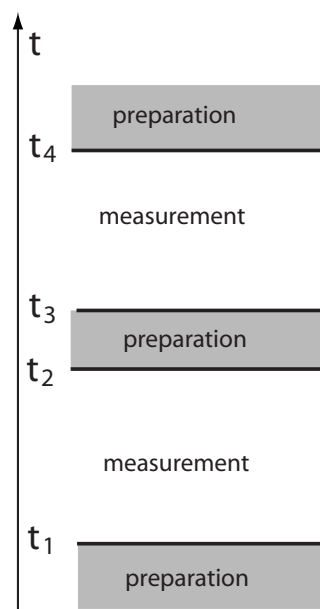


FIG. 1. An experiment consisting of three “preparation” stages and two “measurement” stages.

Quantum mechanically, however, the situation is far more interesting.

In the discussion above we referred to quantum states as being simply the mathematical tools for describing the system. However, states also have an ontological dimension. This is a highly debated issue, which even for the simple case of a standard state $|\Psi\rangle$ is very controversial. Does the state have a “reality” of its own or is it just a mathematical tool for making predictions? Does the state actually collapse or is the collapse simply our updating the mathematical description following the acquisition of new data (the results of new measurement). Is the state a physical entity (such as in Bohm’s pilot wave model)? Discussing the ontological status of multitime states is bound to be even more controversial. It is not our intention to dwell too much on this issue here. Our main focus is simply to find out what are the parameters that describe the system fully; the structure that we uncovered is independent from its interpretation. We will comment however in the conclusions on our world view in the light of the present results.

Coming now to measurements, we have two aims. The first is simply to discuss “multiple-time measurements.” These are measurements consisting of multiple measurement stages, but which cannot be decomposed into separate measurements, one for each time. Considering such measurements is natural in the context of multitime states. Such measurements were introduced in [4]. The second aim is to introduce the notion of “measurement state.” Traditionally, the idea of “state” is never associated with measurements; it makes however a lot of sense. Indeed, consider first the notion of the state of the system. As discussed above, the state is nothing other than a compact description of all the relevant information about a system, the totality of the parameters needed to deduce what will happen to the system in any conceivable situation. One may know much more about the system but this knowledge may be redundant. In a similar way, we can ask what are all the relevant parameters that describe a measurement; the totality of these parameters will then form a measurement state. For example, consider the usual von Neumann measurement. Suppose we measure an observable A . All the relevant information is encoded in the projectors P_n corresponding to its eigensubspaces. We may know, of course, much more about the measurement (detailed information about the measuring device for example) but this information is irrelevant. In fact, in theoretical discussions one very rarely discusses how such a measurement could be performed—the explicit von Neumann measuring formalism is mostly restricted to a few textbooks [11]. Then we can view each of the projectors P_n as a state describing the measurement, corresponding to the outcome n . While, of course, in this very simple example the notion of measurement state is trivial, its full force will become apparent when dealing with multitime measurements.

Again, one may ask what the ontological meaning of a measurement state is. This is, of course, a perfectly legitimate question. But whatever the ontological meaning is, from a formal point of view the set of projectors P_n are all that is needed to describe the von Neumann measurement, so they form a state. For the main part of this paper we will focus on the mathematical formalism and discuss possible interpretations later.

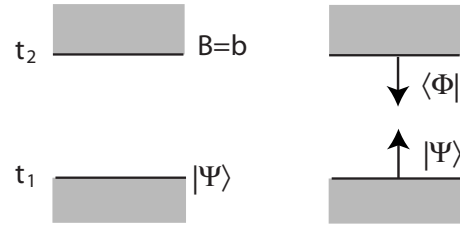


FIG. 2. (a) At time t_1 the system is prepared in state $|\Psi\rangle$ and at time t_2 an operator B is measured. The outcome of B happens to be b , the (nondegenerate) eigenvalue corresponding to the eigenstate $|\Phi\rangle$. (b) This situation we describe by the *two-time state* ${}_2\langle\Phi||\Psi\rangle_{t_1}$

II. SIMPLE TWO-TIME STATES: PRE- AND POSTSELECTION

We will start with some simple situations and set up the general formalism afterward. The physical situation, illustrated in Fig. 2, is the following. There are two preparation stages, one at t_1 and the other at t_2 , and one measurement stage that takes place between t_1 and t_2 . The system is prepared at time t_1 in some quantum state $|\Psi\rangle$. At a later time, t_2 the system is subjected to the measurement of an observable B and the result $B=b$ is obtained; suppose that b is a nondegenerate eigenvalue of B corresponding to the eigenstate $|\Phi\rangle$ (i.e., $B|\Phi\rangle=b|\Phi\rangle$).

More precisely, there are in fact two interesting different physical situations that we can consider. In the first case t_1 and t_2 are both in the past, t_1 in the remote past, and t_2 in a more recent past. In this case, the measurement of B has already been performed and the result b is the actual result of this measurement. The other case is when the second preparation did not yet take place. In this case we cannot guarantee that the result b will actually be obtained—the measurement might very well yield some other result b' . Then, if we are interested only in the case in which the second preparation stage yields b , we have no other option but discard the system and start all over again. This is called “postselection.”

In both the above cases, the entire information about the two preparation stages relevant for the physics in the time period $[t_1, t_2]$ is contained in the two states $|\Psi\rangle$ and $|\Phi\rangle$. We now define a two-time state corresponding to this situation by

$${}_2\langle\Phi||\Psi\rangle_{t_1}. \tag{1}$$

Note that expression (1) is *not* a scalar product (i.e., it is not the complex number $\langle\Phi||\Psi\rangle$) but a mathematical object which is comprised of a bra and a ket vector with an empty slot in between. In this slot we eventually insert information about the measurement period.

We use this state in the following way. Suppose that the particle evolves from t_1 to t according to the unitary operator $U(t, t_1)$ and from t to t_2 according to $U(t_2, t)$. Furthermore, suppose that at t the particle is subjected to a von Neumann measurement of an observable C . Let P_n be the projector associated to the eigenvalue c_n .

To obtain the probability $p(C=c_n)$ that the measurement of the observable C yields $C=c_n$ given the two-time state (1) we simply insert the “history” $U(t_2, t)P_nU(t, t_1)$ in the avail-

able slot, and make the contractions (i.e., we apply the operator $U(t, t_1)$ to $|\Psi\rangle_{t_1}$, then act with the projector P_n , etc.) to obtain the complex number

$$\langle \Phi | U(t_2, t) P_n U(t, t_1) | \Psi \rangle. \quad (2)$$

The probability $p(C=c_n)$ is then given by

$$p(C=c_n) = \frac{1}{N} |\langle \Phi | U(t_2, t) P_n U(t, t_1) | \Psi \rangle|^2. \quad (3)$$

The normalization constant N is given by

$$N = \sum_k |\langle \Phi | U(t_2, t) P_k U(t, t_1) | \Psi \rangle|^2 \quad (4)$$

and ensures that the probabilities for all possible outcomes add up to 1. This formula is known as the Aharonov-Bergman-Lebowitz (ABL) rule [1]. Note that the normalization constant N could not have been included in the definition of the state itself because its value depends not only on the state but also on the experiment to which the state is subjected.

The case of multiple measurements performed between t_1 and t_2 can also be dealt with easily [1]. Consider, for example, two von Neumann measurements of the observables C and D performed at t and t' , respectively, and let P_n be the projector corresponding to $C=c_n$ and Q_k be the projector associated to $D=d_k$. Then

$$p((C=c_n) \& (D=d_k)) = \frac{1}{N} |\langle \Phi | U(t_2, t') Q_k U(t', t) P_n U(t, t_1) | \Psi \rangle|^2 \quad (5)$$

with

$$N = \sum_{j,l} |\langle \Phi | U(t_2, t') Q_j U(t', t) P_l U(t, t_1) | \Psi \rangle|^2. \quad (6)$$

Finally, going beyond von Neumann measurements, general measurements can be described in the positive-operator-valued measure (POVM) formalism. Any measurement can be viewed as an interaction between the measured system and the measuring device, followed by “reading” the outcome indicated by the measuring device, i.e., by performing a von Neumann measurement on the measuring device itself. We will first discuss POVMs in the usual context of a one-time state.

Consider first a “detailed” POVM. In such a measurement we leave no information unread. That is, we subject the measuring device to a *complete* von Neumann measurement, i.e., a von Neumann measurement which is such that all the eigenvalues correspond to one-dimensional projectors. Following such a measurement, the system ends up in a *pure* state—it may first get entangled with the measuring device but then the entanglement is destroyed by reading the measuring device. (Note that as discussed below, following a general POVM, the system may remain entangled with the measuring device.)

A detailed POVM is described by the operators that describe the evolution of a quantum state due to the measurement. As noted above, under a detailed POVM pure states

evolve into pure states. Let A_k be the operator that describes the evolution given the measurement outcome k , i.e., the initial state $|\Psi\rangle$ evolves into the (unnormalized) state $A_k|\Psi\rangle$,

$$|\Psi\rangle \rightarrow A_k|\Psi\rangle. \quad (7)$$

The operators A_k are called Krauss operators. They are linear operators and they are arbitrary (not necessarily Hermitian), up to the normalization condition

$$\sum_k A_k^\dagger A_k = I, \quad (8)$$

where I is the identity. The probability of obtaining the result k is given by the norm of the postmeasurement state, namely,

$$p(k) = \langle \Psi | A_k^\dagger A_k | \Psi \rangle, \quad (9)$$

and the normalization condition ensures that the probabilities add up to 1.

Note that von Neumann measurements are particular cases of detailed POVMs in which the Krauss operators $A_k = P_k$ are projection operators. Time evolutions can also be easily included into the Krauss operators: $A_k = U(t_2, t) P_k U(t, t_1)$ describes a von Neumann measurement preceded and followed by unitary time evolutions. Furthermore, a series of von Neumann measurements is also a particular detailed POVM. Indeed the operator $U(t_2, t') Q_k U(t', t) P_n U(t, t_1)$ considered in Eq. (5) above can be viewed as a Krauss operator A_{nk} corresponding to the outcome given by the pair (k, n) . For simplicity, from now on, unless explicitly specified otherwise, we consider the Krauss operators to cover the entire measurement period they refer to.

Dealing with detailed POVMs in the context of pre- and postselected states is identical to the way in which we dealt with von Neumann measurements. We associate Krauss operators A_k with the entire experiment that takes place between t_1 and t_2 (considering all unitary evolutions as part of the measurement itself) and the probability of obtaining the result k is given by

$$p(k) = \frac{1}{N} |\langle \Phi | A_k | \Psi \rangle|^2, \quad (10)$$

where N is a normalization factor which ensures that $\sum_k p(k) = 1$.

A general POVM is different from a detailed POVM in that we do not perform a *complete* reading of the measuring device. To find the probabilities in this case, we can imagine that after finishing the original POVM we proceed to read the remaining information but this new information is simply disregarded. (Since the measuring device no longer interacts with the system, whether or not we make this supplementary reading of the measuring device makes no difference to the system.) In effect, what we now have is a detailed POVM which is such that to each outcome k of the original POVM correspond a number of different outcomes (k, μ) . All we have to do, then, is simply add the probabilities for the different outcomes of the detailed POVM corresponding to the same k . Formally, to each measurement outcome k of the original POVM correspond, in general, more Krauss opera-

tors $A_{k\mu}$ where the index k refers to the measurement outcome, and the index μ describes different results that could have been differentiated but are lumped together and associated to the overall outcome k . Again, the Krauss operators are arbitrary linear operators subject to the condition $\sum_{k\mu} A_{k\mu}^\dagger A_{k\mu} = 1$. For this POVM, the probability of obtaining the result k is given by

$$p(k) = \frac{1}{N} \sum_{\mu} |\langle \Phi | A_{k\mu} | \Psi \rangle|^2, \quad (11)$$

where N is a normalization factor that ensures that $\sum_k p(k) = 1$.

We thus conclude that *any* measurements performed on a pre- and postselected system can be described using the mathematical object $\langle \Phi | | \Psi \rangle$. We therefore are entitled to view $\langle \Phi | | \Psi \rangle$ as *the* state of the system.

Up to this point, however, the situation is rather trivial and can be handled quite simply with the standard formalism of quantum mechanics (in the manner indicated below). It suffices to consider the simplest case of a single von Neumann measurement discussed above; all other cases can be dealt with in a similar manner. For simplicity, we will write $U_1 = U(t, t_1)$ and $U_2 = U(t_2, t)$. In the usual formalism we say that the system starts in the state $|\Psi\rangle$ and evolves into $U_1|\Psi\rangle$ just prior to the measurement of C . The probability to obtain c_n is

$$\langle \Psi | U_1^\dagger P_n U_1 | \Psi \rangle \quad (12)$$

and the state after the measurement becomes

$$\frac{P_n U_1 | \Psi \rangle}{\sqrt{\langle \Psi | U_1^\dagger P_n U_1 | \Psi \rangle}}. \quad (13)$$

The probability to obtain b , the eigenvalue corresponding to $|\Phi\rangle$ when measuring B at t_2 is the absolute value square of the scalar product between $|\Phi\rangle$ and the state (13) after it undergoes propagation by U_2 , i.e.,

$$\left| \frac{\langle \Phi | U_2 P_n U_1 | \Psi \rangle}{\sqrt{\langle \Psi | U_1^\dagger P_n U_1 | \Psi \rangle}} \right|^2. \quad (14)$$

The overall probability to obtain c_n and then b is

$$\left| \frac{\langle \Phi | U_2 P_n U_1 | \Psi \rangle}{\sqrt{\langle \Psi | U_1^\dagger P_n U_1 | \Psi \rangle}} \right|^2 \langle \Psi | U_1^\dagger P_n U_1 | \Psi \rangle = |\langle \Phi | U_2 P_n U_1 | \Psi \rangle|^2. \quad (15)$$

This, however, is the probability to obtain c_n and then b . The *conditional* probability to obtain c_n given that the measurement of B obtained b is given by the usual conditional probability formula, by dividing the above probability by the overall probability to obtain b (given that we measured C at t), i.e.,

$$\frac{|\langle \Phi | U_2 P_n U_1 | \Psi \rangle|^2}{\sum_k |\langle \Phi | U_2 P_k U_1 | \Psi \rangle|^2}, \quad (16)$$

which yields our formula (3). The case of multiple measurements can be handled in a similar way.

Note however that in this standard way of computing we use the notion of “state” in an ontological way, not as a

repository of all the relevant information about the system. That is, we considered that the system actually *is* in a state $|\Psi\rangle$ at time t_1 , that the state evolves into $U_1|\Psi\rangle$ just prior to the measurement, that it then collapses into $P_n U_1|\Psi\rangle$ and so on. Of course, this usage may seem very appealing from the point of view of an intuition established in standard discussions about quantum experiments. Nevertheless, conceptually this is a very different usage of the notion of state. In the situation in which we are interested, when we have information about the system at two different times, the two-time state (1) is the *only* mathematical object that can be called a state in the sense of containing all the relevant information. The full power of this approach will become evident in Sec. III.

III. TWO-TIME STATES

Let us now return to the two-time state (1). Although the case of pre- and postselection described above can be dealt with relatively simply by the ordinary formalism, our formalism which uses the two-time state ${}_2\langle \Phi | | \Psi \rangle_{t_1}$ has advantages. Not only is it more compact, but it also leads us to ask new questions that could not be easily articulated in the old language.

The two-time state is a mathematical object living in a Hilbert space $\mathcal{H} = \mathcal{H}_{t_2}^\dagger \otimes \mathcal{H}_{t_1}$, where \mathcal{H}_{t_1} is the Hilbert space of the states at t_1 and $\mathcal{H}_{t_2}^\dagger$ is the Hilbert space for t_2 . The dagger indicates that $\mathcal{H}_{t_2}^\dagger$ is a space of bra vectors while \mathcal{H}_{t_1} is a space of ket vectors.

The remarkable thing about two-time states is that, similar to ordinary quantum states, we can form superpositions (which originally were named *generalized states* [12]). In other words, any vector in $\mathcal{H} = \mathcal{H}_{t_2}^\dagger \otimes \mathcal{H}_{t_1}$ is a possible state of the system.

Consider the state

$$\alpha_1 {}_2\langle \Phi_1 | | \Psi_1 \rangle_{t_1} + \alpha_2 {}_2\langle \Phi_2 | | \Psi_2 \rangle_{t_1}, \quad (17)$$

where $\langle \Phi_1 |$, $\langle \Phi_2 |$, $|\Psi_1\rangle$, and $|\Psi_2\rangle$ are arbitrary states. What does this state represent and how can we prepare it? The answer to this question is obtained by looking at the probabilities for different measurements when the system is in this state. Suppose that at time t , between t_1 and t_2 we measure an observable C and let the projection operator corresponding to the eigenvalue c_n be denoted by P_n . Applying the rule used for simple two-time states, the probability for obtaining c_n is given by

$$p(C = c_n) = \frac{1}{N} |\alpha_1 \langle \Phi_1 | U_2 P_n U_1 | \Psi_1 \rangle + \alpha_2 \langle \Phi_2 | U_2 P_n U_1 | \Psi_2 \rangle|^2. \quad (18)$$

One way to prepare a two-time state that leads to this result is the following. Consider our system and a supplementary particle, an ancilla. Consider now an ordinary pre- and postselection as described before, but this time let both the preselected state and the postselected state be entangled states between the system and the ancilla. Specifically, let the state at t_1 be

$$|\Psi_1\rangle_{t_1}^S |1\rangle_{t_1}^A + |\Psi_2\rangle_{t_1}^S |2\rangle_{t_1}^A \quad (19)$$

and the state at t_2 be

$$\alpha_1 |1\rangle_{t_2}^S \langle\Phi_1| + \alpha_2 |2\rangle_{t_2}^S \langle\Phi_2|, \quad (20)$$

where the indices S and A denote the system and the ancilla. The two-time state for the two particles is then

$$(\alpha_1 |1\rangle_{t_2}^S \langle\Phi_1| + \alpha_2 |2\rangle_{t_2}^S \langle\Phi_2|) (|\Psi_1\rangle_{t_1}^S |1\rangle_{t_1}^A + |\Psi_2\rangle_{t_1}^S |2\rangle_{t_1}^A). \quad (21)$$

Suppose now that we perform a measurement on the system while the ancilla is left completely undisturbed—no measurement is performed on it and its Hamiltonian is zero. Since neither the projection operator P_n associated with the measurement that is performed on the system nor the unitary evolutions U_1 and U_2 affect the ancilla, we obtain

$$\begin{aligned} p(C = c_n) &= \frac{1}{N} |\alpha_1 \langle 1 | \langle \Phi_1 | + \alpha_2 \langle 2 | \langle \Phi_2 | U_2 P_n U_1 (|\Psi_1\rangle |1\rangle + |\Psi_2\rangle |2\rangle)|^2 \\ &= \frac{1}{N} |\alpha_1 \langle \Phi_1 | U_2 P_n U_1 |\Psi_1\rangle + \alpha_2 \langle \Phi_2 | U_2 P_n U_1 |\Psi_2\rangle|^2. \end{aligned} \quad (22)$$

So as long as we are interested in the system alone and trace over the ancilla the system is described by Eq. (17). The state (17) is a *pure, entangled* two-time state. The entanglement is between the states of the system at the two different moments of time, more precisely between the “forward in time” propagating states prepared at t_1 and the “backward in time” propagating states prepared at t_2 .

Note that there are many other—in fact infinitely many other ways—in which the state (17) can be prepared. For example we can preselect

$$\sum_{i=1,2} \beta_i |\Psi_i\rangle_{t_1}^S |i\rangle_{t_1}^A \quad (23)$$

and postselect

$$\sum_{j=1,2} \gamma_j \langle\Phi_j|_{t_2}^A |j\rangle_{t_2}^S \quad (24)$$

with $\gamma_i \beta_i = \alpha_i$. This freedom in preparing the state (17) is similar to the freedom in the way in which an ordinary density matrix for a system can be obtained by entanglement with an ancilla—there are infinitely many pure entangled states that lead to the same reduced density matrix for the system.

The generalization of state (17) and of its method of preparation [Eqs. (23) and (24)] to a superposition with an arbitrary number of terms is obvious.

Yet another way to prepare arbitrary superpositions of two-time states is to put all information about the two-time state in the initial state of the system and ancilla and to use a standard postselection to transfer information from the ancilla onto the system (Fig. 3). The simplest way to describe this method is to use a decomposition of the desired two-time state using orthonormal basis vectors ${}_{t_2}\langle j || i \rangle_{t_1}$ in $\mathcal{H} = \mathcal{H}_{t_2}^\dagger \otimes \mathcal{H}_{t_1}$. Consider an arbitrary two-time state

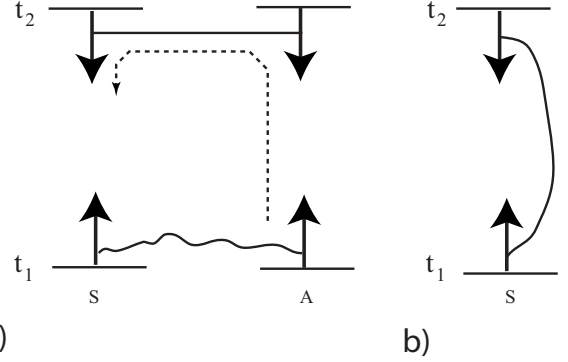


FIG. 3. (a) The system S interacts with the ancilla A . The arrows represent states propagating “forward” and “backward” in time, i.e., ket and bra vectors. The wiggled line connecting the forward in time propagating states (i.e., ket vectors) describes (arbitrary) entanglement. The continuous line connecting the backward in time propagating states (i.e., bra vectors) illustrates maximal entanglement. More precisely, since it refers to the bra vectors, it denotes post-selecting the maximally entangled state for the system and the ancilla. The dotted line illustrates how entanglement is transferred from the ancilla onto the system. The diagram (b) illustrates the same situation as (a) but from the point of view of the system alone.

$$\sum_{i,j} \alpha_{ij} {}_{t_2}\langle j || i \rangle_{t_1}. \quad (25)$$

To prepare this state we start with our quantum system and an ancilla in the preselected state

$$\sum_{i,j} \alpha_{ij} |i\rangle_{t_1}^S |j\rangle_{t_1}^A, \quad (26)$$

which is a “map” of the desired state. We then postselect the maximally entangled state

$${}_{t_2}^{SA}\langle \Phi^+ | = \sum_n {}_{t_2}^S \langle n | {}_{t_2}^A \langle n |. \quad (27)$$

This can be done, for example, by measuring the well known Bell operator and selecting the appropriate result. By postselecting the maximally entangled state $|\Phi^+\rangle$, we effectively transfer the state of the ancilla into the backward-in-time propagating state of the system. In other words, postselecting on the maximally entangled state $|\Phi^+\rangle$ acts as a channel by which a ket vector of the ancilla is transformed into a bra vector of the system (see Fig. 3).

Indeed, the pre- and postselected state of system+ancilla is

$$\sum_{ij} \alpha_{ij} {}_{t_2}^{SA}\langle \Phi^+ || i \rangle_{t_1}^S |j\rangle_{t_1}^A. \quad (28)$$

When only measurements on the system are concerned, we can contract the ancilla states obtaining

$$\begin{aligned}
 \sum_{ij} \alpha_{ij}^{SA} \langle \Phi^+ || i \rangle^S | j \rangle^A &= \sum_{ijn} \alpha_{ij}^A \langle n |^S \langle n | | i \rangle^S | j \rangle^A \\
 &= \sum_{ijn} \alpha_{ij}^S \langle n | | i \rangle^S \langle n | j \rangle^A \\
 &= \sum_{ijn} \alpha_{ij}^S \langle n | | i \rangle^S \delta_{nj} = \sum_{i,j} \alpha_{ij} \langle j | | i \rangle_{t_1},
 \end{aligned}
 \tag{29}$$

which is the desired state (25).

Until now, we discussed two-time states of a single quantum system. Of course, any number of particles can be grouped together into a single system, so the discussion was completely general. We may, however, find it convenient to describe different particles separately. Consider for example a quantum system composed of two particles, *A* and *B*. A general pure two-time state is

$$\sum_{ijkl} \alpha_{ijkl} \langle i |_{t_2}^B \langle j | | k \rangle_{t_1}^A | l \rangle_{t_1}^B.
 \tag{30}$$

In general, such a state is entangled both between the two particles, as well as between the two times. For example there are states in which the postselected state of particle *A* is entangled with the preselected state of particle *B*, etc.

Finally, we note that along with pure two-time states we can have *mixed* two-time states. A mixture arises when we prepare different pure two-time states with different probabilities.

IV. MULTIPLE-TIME STATES

The two-time states discussed above are just the simplest example of multiple-time states. They correspond to the situation in which there is one measurement stage sandwiched between two preparation stages, as illustrated in Fig. 2. Our formalism however applies equally well to situations consisting of multiple preparation and measurement stages.

Consider an experiment as illustrated in Fig. 4. To each time boundary between a preparation period followed by a measurement period we associate a Hilbert space of ket vectors and to each time boundary between a measurement period followed by a preparation period we associate a Hilbert space of bra vectors. The total Hilbert space is the tensor product of the Hilbert spaces for all the time boundaries, $\mathcal{H} = \mathcal{H}_{t_n}^{(\cdot)} \otimes \dots \otimes \mathcal{H}_{t_{k+1}}^{\dagger} \otimes \mathcal{H}_{t_k} \otimes \mathcal{H}_{t_{k-1}}^{\dagger} \otimes \dots \otimes \mathcal{H}_{t_1}^{(\cdot)}$. Note that the bra and ket Hilbert spaces alternate due to the alternation of preparation and measurement periods. Furthermore, note also that we marked the first and last Hilbert spaces with the index (\cdot) to denote the presence or absence of \dagger . This is because there are four different cases (Fig. 4) depending on whether the first and last Hilbert spaces are bra or ket spaces, i.e., whether the procedure starts (ends) with a preparation or measurement period. Which of these four cases occurs depends on whether the past and future are uncertain or well defined. We will discuss the significance of the difference between these four cases shortly.

We are now ready to state the basic result of our paper. Let Ψ denote a state in \mathcal{H} .

Theorem. In the case of multiple periods of preparation

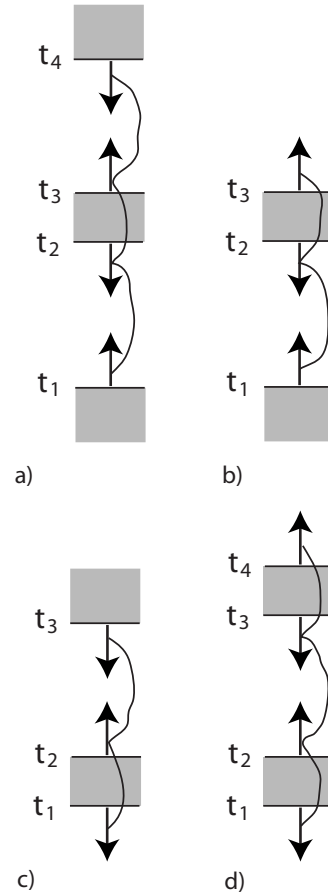


FIG. 4. Different multi-time situations. In (a) both the past and the future are well defined, i.e., they are part of “preparation” stages. In (b) the future is uncertain, in (c) the past is uncertain while in (d) both the future and the past are uncertain.

and measurements, any physical state of a quantum system can be described by a vector Ψ in \mathcal{H} or by mixtures of such vectors. Furthermore, to any vector or any mixture of vectors in \mathcal{H} corresponds a physical state of the system.

In the above, the word “mixture” is taken to have the same two different meanings as in standard quantum mechanics: (a) the preparer throws a die and prepares a different multitime state for each outcome; when the preparer gives us the state but does not inform us about the outcomes of the die, from our point of view we have a mixture and (b) the multitime state of the system is entangled with an ancilla. For simplicity, in the present paper we restrict ourselves to the case of “pure” multitime states. A density matrix formulation for mixed multitime states will be developed elsewhere.

It is important to note that, unlike in standard quantum theory, we do not require the multitime states to be normalized. This is because there is no advantage in normalizing the multitime states. Indeed, normalization of multitime states does not automatically imply normalization of the probabilities of measurement outcomes. Normalization of probabilities is an issue that can only be resolved when it is known what measurements were actually performed. For any given set of measurements the state Ψ prescribes only the *relative* probabilities of the different outcomes of the measurements

and the normalization of the probabilities is then calculated so that the total probability is 1. Ultimately this stems from the fact that the overall probability to prepare such a state depends on the probabilities of success of the different post-selections involved, and these probabilities depend not only on what happens during the preparation times but also on the measurements to which the state is subjected. This is different from the case of ordinary one-time states which are prepared in advance; the probability of preparation is in this case equal to 1 and it is independent of the measurement to which the system is thereafter subjected.

Finally, note that when discussing the case of multiple quantum systems, we may have a different number of preparation and measurement stages for each system. For example

$$|\Psi_3\rangle_{t_3}^{AA} \langle\Psi_2||\Psi_1\rangle_{t_1}^A |\Phi\rangle_{t_1}^B \quad (31)$$

represents a state of two quantum systems, A and B in which system A is subjected to two preparation stages, from $t=-\infty$ to t_1 and from t_2 to t_3 while system B is subjected to a single preparation stage, from $t=-\infty$ to t_1 . This idea generalizes easily for multiple particles and multiple times.

V. MEASUREMENT PROBABILITIES FOR MULTITIME STATES

As in the case of two-time states, the meaning of the multitime states is defined by the probabilities they yield when the system is subjected to measurements. The probabilities for the outcomes of different measurements are obtained from multitime states in a very similar way to that in which they are obtained from one- and two-time states. Consider first the case of detailed POVMs. To obtain the probability of a given outcome we must:

(i) Step 1. Act on the multitime state with the corresponding Krauss operators, i.e., insert the Krauss operators in the appropriate slots and make all the scalar products with the bra and ket vectors to which they apply. (Note that if in a certain measurement period nothing is done, this corresponds to a Krauss operator that is simply the identity)

(ii) Step 2. Compute the norm-squared of the resulting vector. Note that the four cases discussed in the previous section (i.e., uncertain or well-defined future and past) are slightly different: indeed, after acting with the Krauss operators we end either with a ket, a bra, a superposition of tensor products of a ket and a bra or just a complex number. Computing the norm has to be done in the appropriate way.

(iii) Step 3. Normalize the probabilities. That is, do steps 1 and 2 for each particular outcome—this will determine the relative probabilities of the outcomes. To obtain the absolute probabilities divide all the relative probabilities by their sum.

We now consider two examples. First consider the four-time state corresponding to the situation illustrated in Fig. 4(a) in which there is a well-defined past and future (determined by the initial preparation and final post-selection) and two measurement periods ($t_1 < t < t_2$) and ($t_3 < t < t_4$). The multitime state Ψ for this example is a vector in the Hilbert space $\mathcal{H} = \mathcal{H}_{t_4}^\dagger \otimes \mathcal{H}_{t_3} \otimes \mathcal{H}_{t_2}^\dagger \otimes \mathcal{H}_{t_1}$ and can be expanded in terms of basis states as

$$\sum_{ijkl} \alpha_{ijkl} t_4 \langle l||k\rangle_{t_3} t_2 \langle j||i\rangle_{t_1}. \quad (32)$$

Let us denote the Krauss operators acting in the first and second measurement periods by A_μ and B_ν , respectively, where μ and ν denote the corresponding results.

Acting on state (32) with the Krauss operators according to step 1 above, we obtain

$$\sum_{ijkl} \alpha_{ijkl} t_4 \langle l|B_\nu|k\rangle_{t_3} t_2 \langle j|A_\mu|i\rangle_{t_1} \quad (33)$$

which is a complex number. According to step 2, the relative probability to obtain the results μ and ν is the norm-squared of this complex number. Dividing these relative probabilities by their sum (step 3) we obtain the absolute probabilities of the results μ and ν ,

$$p(\mu, \nu) = \frac{1}{N} \left| \sum_{ijkl} \alpha_{ijkl} t_4 \langle l|B_\nu|k\rangle_{t_3} t_2 \langle j|A_\mu|i\rangle_{t_1} \right|^2, \quad (34)$$

where N is such that $\sum_{\mu, \nu} p(\mu, \nu) = 1$.

A second example corresponds to the situation illustrated in Fig. 4(d). In this example, both the future and the past are uncertain, i.e., they belong to the experimentalist who performs measurements not to the preparer. There are three measurement periods ($t < t_1$), ($t_2 < t < t_3$), and ($t_4 < t$). The four-time state corresponding to this situation is

$$\sum_{ijkl} \alpha_{ijkl} |l\rangle_{t_4} t_3 \langle k||j\rangle_{t_2} t_1 \langle i|. \quad (35)$$

Let A_μ , B_ν , and C_ξ denote the Krauss operators corresponding to the measurements performed in the three measurement periods and μ , ν , and ξ denote the corresponding results. Then the first step is to act on the state with the Krauss operators. The result of acting with the Krauss operator and making all the contractions (all the scalar products) is

$$\sum_{ijkl} \alpha_{ijkl} C_\xi^\dagger |l\rangle_{t_4} t_3 \langle k|B_\nu|j\rangle_{t_2} t_1 \langle i|A_\mu, \quad (36)$$

which is a tensor product between ket and bra vectors corresponding to the initial and final time, respectively. Indeed, note that in the above formula $t_3 \langle k|B_\nu|j\rangle_{t_2}$ is just a complex number while $C_\xi^\dagger |l\rangle_{t_4}$ and $t_1 \langle i|A_\mu$ are uncontracted vectors. According to step 2, the relative probability to obtain the results μ , ν , and ξ is the norm-squared of this vector. Dividing these relative probabilities by their sum (step 3) we obtain the absolute probabilities of the results

$$\begin{aligned} p(\mu, \nu, \xi) &= \frac{1}{N} \left\| \sum_{ijkl} \alpha_{ijkl} C_\xi^\dagger |l\rangle_{t_4} t_3 \langle k|B_\nu|j\rangle_{t_2} t_1 \langle i|A_\mu \right\|^2 \\ &= \frac{1}{N} \sum_{ijkl i' j' k' l'} \alpha_{i' j' k' l'}^* \alpha_{ijkl} t_4 \langle l'|C_\xi^\dagger C_\xi^\dagger |l\rangle_{t_4} t_2 \langle j'|B_\nu^\dagger|k'\rangle_{t_3} \\ &\quad \times t_3 \langle k|B_\nu|j\rangle_{t_2} t_1 \langle i|A_\mu A_\mu^\dagger|i'\rangle_{t_1}, \end{aligned} \quad (37)$$

where N is such that $\sum_{\mu, \nu, \xi} p(\mu, \nu, \xi) = 1$.

It is important to note that when dealing with multitime states that describe a situation with multiple measurement periods, in order to be able to predict the probabilities for the

outcomes of a given measurement we need, in general, information about *all* the measurement periods, not only the ones when the measurement takes place. Indeed, it is easy to see that what happens in other periods may influence the (relative) probabilities of the different outcomes. Consider for example the four-time state $\sum_{i_4} \langle \Phi | i \rangle_{t_3 t_2} \langle i | \Psi \rangle_{t_1}$ where the states $|i\rangle_{t_3}$ and ${}_{t_2} \langle i|$ form complete bases in their Hilbert spaces. This state describes a situation with two measurement stages, from t_1 to t_2 and from t_3 to t_4 . Suppose now that a measurement takes place from t_3 to t_4 , and suppose also that during the period from t_1 to t_2 some action is performed on the system, say a unitary evolution U . Then the probabilities $p(k)$ turn out to be

$$p(k) = \frac{1}{N} |\langle \Phi | A_k U | \Psi \rangle|^2, \quad (38)$$

where A_k are the corresponding Krauss operators. Clearly these probabilities depend on U . Similarly, also in the case when the measurement takes place first, (i.e., measurement between t_1 and t_2 and unitary evolution between t_3 and t_4 , the probabilities are also influenced by U . In this case

$$p(k) = \frac{1}{N} |\langle \Phi | U A_k | \Psi \rangle|^2. \quad (39)$$

Basically, what happens during one time period influences what happens during another time period via the correlations between the vectors associated with these periods. The only case when we do not need to know information about all the periods is when some periods effectively decouple from the rest, i.e, when the vectors that refer to these measuring periods are not entangled with vectors from any other measurement period (see Fig. 5). In this case we can reduce the multitime state to an effective state covering only the connected periods of interest.

Finally, the formalism can be made far more compact in the following way. When there are multiple measurement periods, each characterized by its Krauss operator, we can define a global Krauss operator as the tensor product of the individual operators corresponding to the different measurement periods. For example, when there are two measurement periods, such as in the first example above, one described by A_μ and one by B_ν , we can define the total Krauss operator $K_\lambda = A_\mu \otimes B_\nu$, where the index λ describes now the outcome of the two measurements and is, in this case, nothing other than the pair (μ, ν) . Then the probability formula is

$$p(\lambda) = \frac{1}{N} \|K_\lambda \cdot \Psi\|^2 \quad (40)$$

with N such that $\sum_\lambda p(\lambda) = 1$. Here by the dot product $K_\lambda \cdot \Psi$, we simply mean that every bra (ket) vector belonging to K_λ is contracted with the ket vector belonging to Ψ and corresponding to the same time and the contraction is the scalar product. This formula is the direct equivalent of the well-known formula for determining the probability of a von Neumann measurement in a standard one-time experiment,

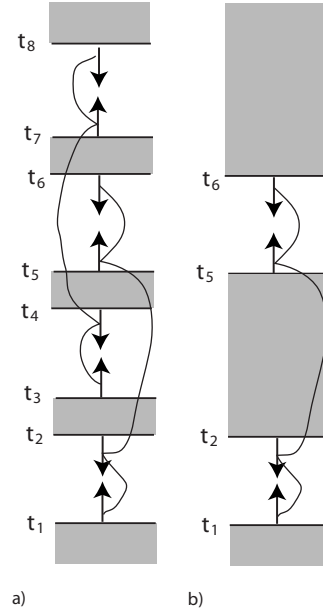


FIG. 5. (a) The second and the fourth “measurement” stages, i.e., from t_3 to t_4 and from t_7 to t_8 are correlated with each other but not with the other two “measurement” stages. If we are interested in measurements that occurred during the first and the third stages, the second and the fourth stages are irrelevant. (b) An effective state describing the periods t_1 to t_2 and t_5 to t_6 can be obtained simply by ignoring the other stages.

$$p(\lambda) = \frac{1}{N} \|P_\lambda |\Psi\rangle\|^2, \quad (41)$$

where P_λ is the projector associated to the eigenvalue λ of the measured observable and where $N = \langle \Psi | \Psi \rangle$.

VI. PREPARING MULTITIME STATES I

There are many (infinite) ways of preparing multitime states. Here we will present one particular method, which is a generalization of the last method of preparing two-time states presented in Sec. III. Yet another way to prepare two-time and multiple-time states is to perform multiple-time measurements which will be analyzed in Sec. IX [13].

In this section we discuss multiple-time states in which the first time corresponds to a ket vector, that is, in which the whole experiment starts with a preparation period. The cases that start with a measurement period are discussed in the next section. We exemplify our method for an arbitrary four-times state [Fig. 4(a)]; generalizations are obvious. The preparation procedure is illustrated in Fig. 6.

Consider the four-time state

$$\sum_{ijkl} \alpha_{ijkl} |l\rangle_{t_4} \langle j|_{t_3 t_2} \langle i|_{t_1}. \quad (42)$$

We start by using three ancillas and preparing at t_1 the state

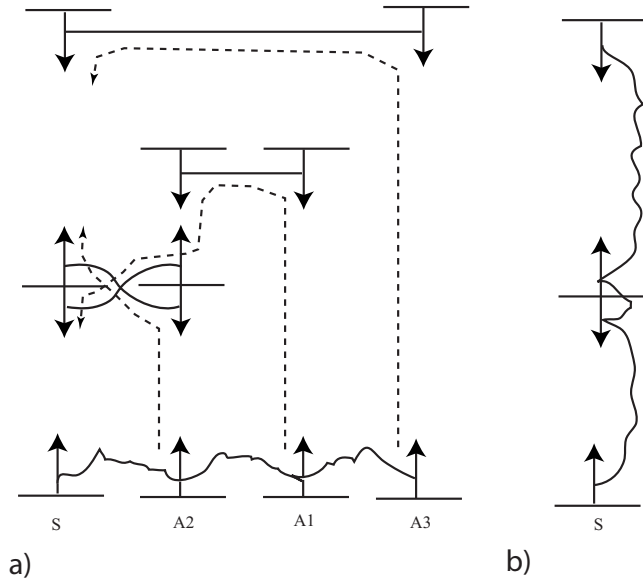


FIG. 6. (a) The system S and the three ancillas, A_1 , A_2 and A_3 start in an entangled state that gets transferred onto the system via the interactions of the system and the ancillas and via appropriate post-selections. The continuous line describes maximal entanglement while the wiggled line describes arbitrary entanglement and the arrows represent states propagating “forward” and “backward” in time, i.e., ket and bra vectors. Note the SWAP interaction between the system S and the ancilla A_2 ; it is nothing other than maximal entanglement between bra states of the system and ket states of the ancilla and vice versa. The dotted line illustrates how entanglement is transferred from the ancillas onto the system. The diagram (b) illustrates the same situation as (a) but from the point of view of the system alone.

$$\sum_{ijkl} \alpha_{ijkl} |i\rangle_S |j\rangle_{A1} |k\rangle_{A2} |l\rangle_{A3} \quad (43)$$

which is a map of the desired state (42). The ancillas are kept undisturbed except when we use them to transfer their states onto the system. The transfer is performed via postselection of maximally entangled states and SWAP operations.

The role of maximally entangled states as channels for transforming ket states of the ancilla into bra vectors of the system was discussed in Sec. III and illustrated in Fig. 3. The swap operation has a similar role. Indeed, the swap $S_{1,2}$ is a unitary operator that swaps the states of two quantum systems, S and A

$$S_{SA} = \sum_{ij} |j\rangle_S |i\rangle_A \langle j|_S \langle i|_A. \quad (44)$$

Note that this operator can also be written as

$$S_{SA} = \left(\sum_i |i\rangle_{AS} \langle i| \right) \left(\sum_j |j\rangle_{SA} \langle j| \right), \quad (45)$$

which is a product of two mathematical objects, each of them looking like a maximally entangled state, but one in which a ket is entangled with a bra. The swap operator then represents two entangled channels, one in which the forward-in-time propagating state of the system is entangled with the backward-in-time propagating state of the ancilla and one in

which the forward-in-time propagating state of the ancilla is entangled with the backward-in-time propagating state of the system. In particular the swap operator allows for the transfer of ket vectors of ancilla into ket vectors of the system and of bra vectors of the ancilla into bra vectors of the system.

The overall procedure for preparing state (42) is the following:

(i) At t_1 prepare the entangled state (43) of the system and of the ancillas.

(ii) At time t' , $t_2 < t' < t_3$ perform the swap operation $S_{S,A2}$ between the system and ancilla A_2 . The system is kept undisturbed at all other times between t_2 and t_3 .

(iii) At time t'' , $t' < t'' < t_4$ perform a Bell operator measurement on ancillas A_1 and A_2 and postselect the maximally entangled state $|\Phi^+\rangle_{A1,A2}$.

(iv) At t_4 perform a Bell operator measurement on the system S and ancilla A_3 and postselect the maximally entangled state $|\Phi^+\rangle_{S,A3}$.

The resulting state of the system and ancillas is

$${}_{t_4}^{S,A3} \langle \Phi^+ | {}_{t_1}^{A1,A2} \langle \Phi^+ | {}_{t'}^{S,A2} \sum_{ijkl} \alpha_{ijkl} |i\rangle_S^S |j\rangle_{t_1}^{A1} |k\rangle_{t_1}^{A2} |l\rangle_{t_1}^{A3}. \quad (46)$$

By contracting the states of the ancillas (i.e., by making the appropriate scalar products) and by propagating in time (without any change, since the system is undisturbed during these times) the state of the system, the bra from t' to t_2 and the ket from t' to t_3 we obtain the desired state (42). The procedure is illustrated in Fig. 6. There the transfer of the ancilla states onto the system can be seen clearly.

Preparing a state for the case when the first period is a preparation and the last period is a measurement period, i.e., a state in which both the first and the last vectors are kets is done by a simple modification of the procedure described above. Consider for example the three-time state

$$\sum_{ijk} \alpha_{ijk} |k\rangle_{t_3 t_2} \langle j| |i\rangle_{t_1}. \quad (47)$$

We prepare it in the same way as the four-time state above, only that the last ancilla, and therefore all the actions involving it, are missing. That is we start from the state

$$\sum_{ijk} \alpha_{ijk} |i\rangle_S |j\rangle_{A1} |k\rangle_{A2} \quad (48)$$

and we perform the exact procedure described above, except the final measurement at t_4 .

VII. PAST AND FUTURE BOUNDARY CONDITIONS

In the previous section we discussed experiments which start with a preparation stage. Correspondingly, the multitime states that describe them start with a ket vector. However, we mentioned in our general theorem that we can also consider experiments that start with a measurement stage, and thus the corresponding multitime states start with a bra vector. At first sight this seems puzzling. Indeed, there is always some state prepared in the remote past, either explicitly prepared by the experimentalist or naturally occurring. So it seems that we should always start with a ket vector. The key however is to

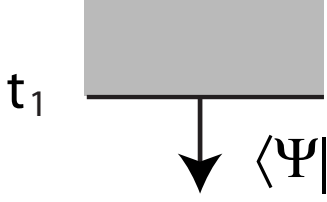


FIG. 7. A simple one-time state with uncertain past.

realize that this problem can be avoided if we make the past “neutral,” i.e., if we arrange a situation such that all states coming from the remote past toward our experiment are equally probable [14]. In other words, a neutral past is one in which the initial state of the system is not any pure state but an equal mixture of all possible states, i.e., (up to normalization) the identity density matrix. This can be done for example by actually starting the experiment with a preparation stage in the remote past, in which we maximally entangle the system with an ancilla.

An example suffices - all multitime states starting with a bra vector can be constructed in a similar way. Consider the one-time state ${}_{t_1}\langle\Psi|$ which is supposed to describe the situation illustrated in Fig. 7 where the experiment consists of a measurement period followed by a preparation.

According to our definitions, the meaning of this state is that if during the measurement period we perform a detailed POVM described by the Krauss operators A_k , the probability to obtain the outcome k is given, up to normalization, by the norm [15] of the vector ${}_{t_1}\langle\Psi|A_k$, i.e.,

$$p(k) = \frac{1}{N} \langle\Psi|A_k A_k^\dagger|\Psi\rangle. \quad (49)$$

A procedure for obtaining ${}_{t_1}\langle\Psi|$ is to prepare the pre- and postselected state of the system and ancilla

$${}_{t_1}^S \langle\Psi||\Phi^+\rangle_{t_0}^{SA}, \quad (50)$$

where $t_0 < t_1$ and where $|\Phi^+\rangle^{SA}$ is the maximally entangled state (27). Note that the ancilla A is then left unmeasured. One can explicitly see that this state is equivalent to ${}_{t_1}\langle\Psi|$. Indeed,

$$\begin{aligned} p(k) &= \frac{1}{N} {}_{t_0}^{SA} \langle\Phi^+|A_k^\dagger|\Psi\rangle_{t_1 t_1}^S \langle\Psi|A_k|\Phi^+\rangle_{t_0}^{SA} \\ &= \frac{1}{N} \sum_n {}_{t_0}^A \langle n|_0^S \langle n|A_k^\dagger|\Psi\rangle_{t_1 t_1}^S \langle\Psi|A_k \sum_m |m\rangle_{t_0}^S |m\rangle_{t_0}^A \\ &= \frac{1}{N} \sum_{nm} \delta_{nm} {}_{t_0}^S \langle n|A_k^\dagger|\Psi\rangle_{t_1 t_1}^S \langle\Psi|A_k|m\rangle_{t_0}^S \\ &= \frac{1}{N} \sum_n {}_{t_1}^S \langle\Psi|A_k|n\rangle_{t_0 t_0}^S \langle n|A_k^\dagger|\Psi\rangle_{t_1}^S \\ &= \frac{1}{N} {}_{t_1}^S \langle\Psi|A_k A_k^\dagger|\Psi\rangle^S. \end{aligned} \quad (51)$$

Incidentally, this means that we can view the state ${}_{t_1}\langle\Psi|$ both as a one-time pure state and as a two-time mixture, (in

which the kets at t_0 come with equal probability).

It is worth at this point looking in more detail at the “future boundary condition” as well. By analogy with the past boundary condition, we conclude that the future is akin to the postselection of the identity density matrix. In other words, we can view the standard one-time state $|\Psi\rangle_{t_1}$ either as a pure one-time state or as a two-time mixture (in which the bra vectors at $t_2 > t_1$ come with equal probability).

Finally, similar arguments show that we can prepare states in which a given time boundary does not have a corresponding bra or ket vector, just by making this boundary completely uncertain. For example, in the situation described in Fig. 4(b), to which we would generally associate a three-time state $\sum_{ijk} \alpha_{ijk} |k\rangle_{t_3 t_2} \langle j||i\rangle_{t_1}$ we can also prepare an effective two-time state of the form

$$\sum_{ik} \beta_{ik} |k\rangle_{t_3} |i\rangle_{t_1} \quad (52)$$

by making the “future” boundary condition at t_2 completely uncertain.

To conclude the last two sections, we showed that any multitime state can be prepared. There are many ways to prepare them, and the general method presented here may not be the most efficient, that is, the probability for the success of all the required post-selections may not be optimal. Indeed, we did not make an optimality study here. However, the main point, namely that all these states are possible, has been made.

VIII. PARTICULAR EXAMPLES OF MULTITIME STATES

An interesting case is the two-time state

$$\sum_{i=1}^n |i\rangle_{t_2 t_1} \langle i|, \quad (53)$$

where the vectors $|i\rangle_{t_2}$ and ${}_{t_1}\langle i|$, respectively, form complete orthonormal bases in \mathcal{H}_2 and \mathcal{H}_1^\dagger , respectively. Here the vectors propagating backward in time at t_1 are completely correlated with those propagating forward in time at t_2 , (i.e., the bra vectors at t_1 and the kets at t_2) are “maximally” entangled. In effect they form an identity operator. (Note that this is very similar to the ordinary entanglement of two particles in a singlet type state, but here it is entanglement between bra and ket vectors and represents total correlations in all possible bases while total correlations are impossible in the case of entanglement between two sets of ket states—the singlet state represents total anticorrelation not total correlation.) Most importantly, this state can be prepared by simply leaving the system unperturbed between t_1 and t_2 . In this case any information reaching t_1 is then propagated to t_2 . For example the state

$$\sum_{i=1}^n |i\rangle_{t_2 t_1} \langle i||\Psi\rangle_{t_0} \quad (54)$$

is (up to normalization) nothing other than the standard state $|\Psi\rangle_{t_0}$ as one can see by verifying that all the probabilities for all the possible measurements are the same for Eq. (54) and

for $|\Psi\rangle_{t_0}$. This example contains a most important message: a time interval when nothing happens, such as between t_1 and t_2 here, is equivalent to a preparation in which the backward-in-time and the forward-in-time propagating vectors emanating from this time interval are “maximally” entangled.

Another interesting state is

$$\sum_i \langle i|_2 \langle i|_1 \rangle_{t_1}. \tag{55}$$

Here the vectors propagating forward in time at t_1 are completely correlated (i.e., maximally entangled) with those propagating backward in time at t_2 . This state represents a “closed time loop”—any information that reaches time t_2 is “propagated” back to time t_1 .

IX. MULTIPLE-TIME MEASUREMENTS

Up to this point when we discussed measurements we considered the usual quantum mechanical measurements, such as measuring the observable C at time t . But such measurements are very simple in the sense that they are “one-time” measurements. One can consider far more complex measurements, namely multitime measurements [4]; it is very natural to consider such measurements here, when discussing multitime states.

A simple example of a two-time observable is $\sigma_x(t_1) - \sigma_x(t_2)$, the difference between the x component of the spin of a spin 1/2 particle at two different times [here the notation $\sigma_x(t_1)$ means the Heisenberg representation of the operator σ_x at time t_1]. We would like to emphasize that $\sigma_x(t_1) - \sigma_x(t_2)$ is a “two-time observable” in the sense that it refers to two different times, not in the sense that it is, or it has to be, described in a multitime formalism; nevertheless, as we show below, a multiple-state-like description is the most appropriate. The important thing to note is that this is an observable that gives the value zero in the case when the x component of the spin is the same at the two times, but does not offer any information about the actual value of the x component. Measuring this operator is therefore *not* equivalent to measuring the x component of the spin at t_1 , followed by another measurement at t_2 , and finally subtracting the values of the results. Indeed, such a measurement would yield too much information: it would tell the actual value of the spin at the both times, not only the difference. How to measure such an observable has been described in [4] and we describe it here for completeness.

Two ways to accomplish the above task are the following. In the first method we use a single measuring device that we couple to the spin twice, once at t_1 and once at t_2 . Following the von Neumann measuring procedure [11] we consider a measuring device consisting of a pointer whose position is denoted q and its conjugate momentum p . The initial state of the measuring device is the pointer indicating zero, i.e., $|q=0\rangle$. The measuring device interacts with the spin via the interaction Hamiltonian

$$H_{\text{int}} = \delta(t - t_1)p\sigma_x - \delta(t - t_2)p\sigma_x. \tag{56}$$

The first time the coupling is such as to shift the pointer’s position q by an amount proportional to σ_x and the second

time to shift it proportional to $-\sigma_x$. Indeed, the time evolution corresponding to the first interaction is $U(t_1) = e^{-ip\sigma_x}$ which is a shift operator shifting q by the value σ_x while the evolution corresponding to the second interaction is a shift operator $U(t_2) = e^{ip\sigma_x}$ representing a shift of q by $-\sigma_x$. Assuming that during the time interval between the two measurements the pointer is preserved in an undisturbed quantum state (i.e., the effective Hamiltonian of the measuring device is zero between the two interactions with the spin) the Heisenberg equations of motion show that [16]

$$q_{\text{final}} = q_{\text{initial}} + \sigma_x(t_1) - \sigma_x(t_2). \tag{57}$$

As the initial position of the pointer is known, $q_{\text{initial}} = 0$, the final value of q indicates $\sigma_x(t_1) - \sigma_x(t_2)$. Furthermore, note that since we did not read the position of the pointer after the first interaction, when the whole measurement is finished we no longer have the possibility of finding out what $\sigma_x(t_1)$ was. Also we cannot find out what $\sigma_x(t_2)$ was because we do not know the position of the pointer before the second interaction.

The second way to perform such a measurement relies on the method of nonlocal measurements [17]. It involves two independent measuring devices one interacting with the spin at t_1 and the other interacting at t_2 . Let the two pointers be described by q_1, p_1 and q_2, p_2 , respectively, and let the interaction Hamiltonian be

$$H_{\text{int}} = \delta(t - t_1)p_1\sigma_x - \delta(t - t_2)p_2\sigma_x. \tag{58}$$

To ensure that we do not get any information about the spin at t_1 and t_2 but only about the difference we prepare the pointers in the entangled state $|q_1 - q_2 = 0, p_1 + p_2 = 0\rangle$. In this state the initial position of each pointer is completely uncertain so by reading their indications after the measurement we cannot infer $\sigma_x(t_1)$ and $\sigma_x(t_2)$ separately, only their difference.

While the observable $\sigma_x(t_1) - \sigma_x(t_2)$ clearly represents the difference in the x component of the spin at times t_1 and t_2 , its actual significance in the context of the usual quantum setting is subtle and somewhat obscure; its significance becomes natural however in the context of multitime preparations.

In the usual setting one starts by preparing a state, say $|\Psi(t_0)\rangle$ at t_0 and then subjects the system to a measurement which, in our case, takes place at t_1 and t_2 . In this situation it is natural to want to relate the measured observable to the properties of the preparation that took place at t_0 . What happens however is that in general $\sigma(t_2)$ is *different* from what it would have been had the interaction with the measuring device at t_1 not occurred. Consequently the measured observable $\sigma_x(t_1) - \sigma_x(t_2)$ is also different from what it would have been, had we not performed the measurement but simply calculate it on the initial state. Indeed, let the spin evolve under the Hamiltonian H_0 . If no other interactions take place then the observables $\sigma_x(t_1)$ and $\sigma_x(t_2)$ are given in terms of the Schrodinger operators by

$$\sigma_x(t_1) = e^{iH_0/\hbar(t_1-t_0)}\sigma_x e^{-iH_0/\hbar(t_1-t_0)},$$

$$\sigma_x(t_2) = e^{iH_0/\hbar(t_2-t_0)} \sigma_x e^{-iH_0/\hbar(t_2-t_0)}. \quad (59)$$

On the other hand, if we do perform the measurement then $\sigma_x(t_1)$ is still the same as before but $\sigma_x(t_2)$, just before the second interaction, is now given by

$$\begin{aligned} \sigma_x(t_2) &= e^{iH_0/\hbar(t_2-t_1)} e^{i\sigma_x p/\hbar} e^{iH_0/\hbar(t_2-t_1)} \\ &\times \sigma_x e^{-iH_0/\hbar(t_2-t_1)} e^{-i\sigma_x p/\hbar} e^{-iH_0/\hbar(t_1-t_0)} \end{aligned} \quad (60)$$

When $e^{-iH_0/\hbar(t_2-t_1)}$ does not commute with $e^{-i\sigma_x p/\hbar}$ then $\sigma_x(t_2)$ is different from what it had been, had the interaction with the measuring device at t_1 not occurred. For example, let $H_0 = B\sigma_z$ which produces a precession of the spin around the z axis. Let $t_2 - t_1$ be such that during this time σ_y rotates into σ_x . In this case the observable $\sigma_x(t_2)$ is the same as $\sigma_y(t_1 + \epsilon)$. Now, if no measurement takes place at t_1 then $\sigma_y(t_1 + \epsilon) = \sigma_y(t_1 - \epsilon)$ but if we do perform the measurement then they are no longer equal.

In the context of multitime preparations however, the observable $\sigma_x(t_1) - \sigma_x(t_2)$ becomes very natural, because in this case we can directly prepare any state of the spin at times t_1 and t_2 , and the observable refers to these values. For example in the state $|\downarrow_x\rangle_{t_2-\epsilon} |\uparrow_x\rangle_{t_1-\epsilon}$ the value of $\sigma_x(t_1) - \sigma_x(t_2) = 2$ while in the entangled state $|\uparrow_x\rangle_{t_2-\epsilon} |\uparrow_x\rangle_{t_1-\epsilon} + |\downarrow_x\rangle_{t_2-\epsilon} |\downarrow_x\rangle_{t_1-\epsilon}$ we have $\sigma_x(t_1) - \sigma_x(t_2) = 0$.

Now, although in the discussion above we described in detail how to measure $\sigma_x(t_1) - \sigma_x(t_2)$ it is important to note that in order to predict the probabilities for the different outcomes we do not need to know the specific way in which the measurement is implemented; just knowing the state and the

observable itself is enough. This is similar to the case of ordinary one-time variables usually studied in quantum mechanics. For example when considering a von Neumann measurement of an observable C we do not need to describe the entire measuring procedure. We just use the state which is measured and the projectors on the different eigenvalues of C . To do this for a two-time observable such as $\sigma_x(t_1) - \sigma_x(t_2)$ we will now use a multi-time-state formalism.

The observable $\sigma_x(t_1) - \sigma_x(t_2)$ could yield three possible values: $+2$, 0 , and -2 . To each of these values we associate a multitime projector. The value $+2$ is obtained when σ_x is ‘‘up’’ at t_1 and ‘‘down’’ at t_2 . The corresponding multistate projector is

$$P_2 = |\downarrow_x\rangle_{t_2} |\uparrow_x\rangle_{t_1} \langle\downarrow_x|_{t_2} \langle\uparrow_x|_{t_1}, \quad (61)$$

where we denoted $t_1^- = t_1 - \epsilon$, $t_1^+ = t_1 + \epsilon$ and so on. The projector corresponding to -2 is

$$P_{-2} = |\uparrow_x\rangle_{t_2} |\downarrow_x\rangle_{t_1} \langle\uparrow_x|_{t_2} \langle\downarrow_x|_{t_1} \quad (62)$$

Finally, the projector corresponding to 0 is

$$P_0 = |\uparrow_x\rangle_{t_2} |\uparrow_x\rangle_{t_1} \langle\uparrow_x|_{t_2} \langle\uparrow_x|_{t_1} + |\downarrow_x\rangle_{t_2} |\downarrow_x\rangle_{t_1} \langle\downarrow_x|_{t_2} \langle\downarrow_x|_{t_1}. \quad (63)$$

The way to use these projectors is identical to the way the projectors for one-time measurements are used: we insert them into the state, in the corresponding slots and make the scalar products. Then the probability to obtain, say $\sigma_x(t_2) - \sigma_x(t_1) = 0$ when the spin is, say, in the four-time state ${}_{t_6} \langle \Psi | | \Phi \rangle_{t_4 t_3} \langle \Xi | | \Theta \rangle_{t_1}$ with $t_1 < t_2 < t_3 < t_4 < t_5 < t_6$ is

$$\begin{aligned} p(\sigma_x(t_2) - \sigma_x(t_1) = 0) &= \frac{1}{N} |_{t_6} \langle \Psi | U_{6,5+} |\uparrow_x\rangle_{t_5} \langle\uparrow_x| U_{5-,4} | \Phi \rangle_{t_4} |_{t_3} \langle \Xi | U_{3,2+} |\uparrow_x\rangle_{t_2} \langle\uparrow_x| U_{2-,1} | \Theta \rangle_{t_1} \\ &+ |_{t_6} \langle \Psi | U_{6,5+} |\downarrow_x\rangle_{t_5} \langle\downarrow_x| U_{5-,4} | \Phi \rangle_{t_4} |_{t_3} \langle \Xi | U_{3,2+} |\downarrow_x\rangle_{t_2} \langle\downarrow_x| U_{2-,1} | \Theta \rangle_{t_1} |^2, \end{aligned} \quad (64)$$

where $U_{2-,1}$ denotes the unitary operator describing the evolution of the spin between the times t_1 and t_2^- and so on.

X. PREPARING MULTITIME STATES II

In Sec. VI we presented a particular method (based on SWAPS and postselection of maximally entangled states) that allows the preparation of any arbitrary multitime state. It is important to note however that any measurement can be used to prepare multitime states. This is similar to the situation in the standard discussions of quantum measurements, but the multitime approach introduces a very important twist.

The usual case is the following. Suppose that the state of a system at time t_1 is $|\Psi\rangle$ and then a measurement is performed between t_1 and t_2 . When the measurement is a detailed POVM and the outcome k is observed, the state of the system at t_2 becomes (up to normalization) $|\Phi\rangle = A_k |\Psi\rangle$,

where A_k is the corresponding Krauss operator.

In the usual way of looking at preparations as described above, the role of the operator A_k , is to transform the initial state into the final state. However, as we will now show, this way of looking at the problem obscures the true role of A_k . The operator is not there in order to evolve the state, but it is *part of the state itself*. A few examples will make this situation clear.

Suppose a quantum system was prepared at time t_0 in the state $|\Psi\rangle_{t_0}$. Furthermore, suppose that between times t_1 and t_2 a measurement was performed and the outcome k (corresponding to A_k) was obtained. The result is the three-time state

$$A_k^{t_2, t_1} |\Psi\rangle_{t_0} \quad (65)$$

where we added upper indexes to the Krauss operator to denote the times between which it acts. To better understand

the meaning of the above state, note that any Krauss operator acting between t_1 and t_2 can be written as

$$A_k^{t_2,t_1} = \sum_{i,j} \alpha_{i,j} |j\rangle_{t_2} \langle i|_{t_1}. \quad (66)$$

Indeed, any linear operator acting on ket vectors at t_1 and transforming them into ket vectors at t_2 can be written in this form. Hence, explicitly written, state (65) is

$$\sum_{i,j} \alpha_{i,j} |j\rangle_{t_2} \langle i|_{t_1} |\Psi\rangle_{t_0}. \quad (67)$$

In the above we considered the POVM performed on a quantum system that was prepared at t_0 in state $|\Psi\rangle_{t_0}$. The effect of adding information from the POVM was to expand the state from a one-time state to a three-time state by simply adding the Krauss operator into the state. This procedure is however far more general: Whatever a multitime state is, if we are further told that a POVM was performed between t and t' , we simply add the corresponding Kraus operator into the description of the state (and therefore expand an n -time state into an $(n+2)$ -time state).

The true force of the formalism however only becomes clear when we consider multitime measurements such as those described in Sec. IX. A multitime measurement has no simple description in the standard quantum formalism. There is no ordinary Krauss operator that simply propagates an initial state into a final state, since there is no well defined “final” state. Indeed, the measurement takes place at many times, and there can be any other interactions in between. In the multi-time formalism however any multitime measurement can be described by Krauss operators—they are however multitime operators. An example is the spin measurement described in Sec. IX. The Kraus operators corresponding to this measurement are the multitime projectors (61)–(63). To obtain the state of the system given the outcome k of the POVM all we do is, again, just to insert the multitime Krauss operator into the original multitime state. Figure 8 illustrates the procedure.

XI. OPERATORS VERSUS STATES

One of the main advantages of the multitime formalism presented in this paper is to put states and operators on an equal footing. Indeed, to start with, operators and multitime states look formally identical—they are both just superposition of tensor products of bra and ket vectors at different times. But this similarity is by no means only superficial or coincidental. In the standard quantum mechanical formalism states are meant to describe how the system was prepared while operators are meant to describe measurements performed on the system. But physically preparations and measurements both involve exactly the same processes—interactions of the system of interest with other quantum systems and/or with measuring devices. The multistate formalism succeeds in making this explicit.

As we argued in Sec. I, the projector operators describing a von Neumann measurement (or indeed, more generally, the Krauss operators) can be viewed as “measurement states,” in

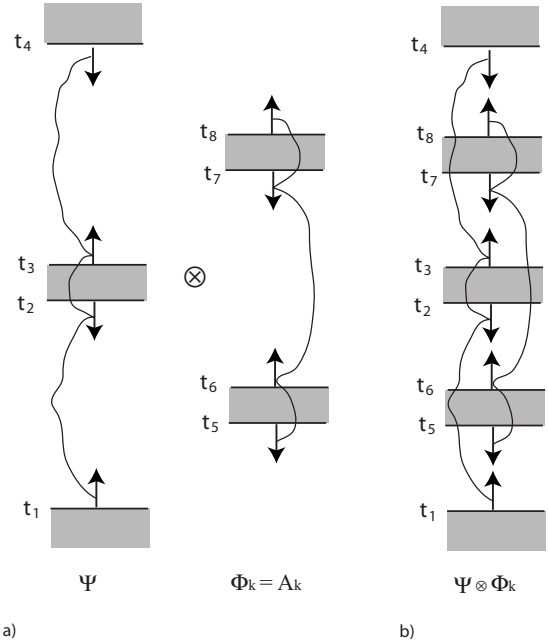


FIG. 8. Using measurements for preparation. To the original state Ψ we add the information that the result of a POVM performed between t_5 to t_6 and t_7 to t_8 yielded the outcome k corresponding to the Krauss operator A_k . The operator A_k can also be viewed as a multi-time state Φ_k . The new state that takes into account all the information is $\Psi \otimes \Phi_k$ which is simply the composition of Ψ and Φ_k , as illustrated in (b).

the sense that they encode all the relevant information about the measurement. But we find it now very useful to think of both the ordinary multitime states (that describe the way in which the system was prepared) and the measurement states (that describe the measurements) on equal footing. This view allows a lot of flexibility.

Let the state of the system be Ψ and let us denote the Krauss operators A_k that describe a given POVM by Φ_k to emphasize that each of them can be interpreted as a state. Now, if we use the measurement as part of the preparation, i.e., if in addition to the information that the system was prepared in the state Ψ we also are informed that we obtained the result k , then the new state of the system is simply the tensor product

$$\Psi \otimes \Phi_k, \quad (68)$$

where by the tensor product we mean combining the two states, as described in the previous section. What this formula tells us is that the total history is simply the combination of the two histories.

On the other hand, suppose that we want to use a POVM not to prepare a state but to test it. That is, suppose we ask, given the state Ψ what are the probabilities to obtain different outcomes k ? In general, of course, there is no definite answer—the answer may depend on other things that may occur to the system meanwhile. For example, suppose we are given the two-time state ${}_2\langle\Phi||\Psi\rangle_{t_1}$ and the POVM takes place between two intermediate times, t' and t'' , $t_1 < t' < t'' < t_2$. Then, the probabilities of the outcomes of the

POVM depend also on what happens between t_1 and t' and between t' and t_2 and therefore we cannot determine them unless we are given this supplementary information. But if we are given the whole information, that is, if in effect the POVM covers the whole measurement period from t_1 to t_2 than we can predict its results. In the operator language, as described in Eq. (40) we have to apply the different Krauss operators to the state and compute the norms of the resulting vectors. On the other hand, we can interpret the same formula as telling that the probability is given, (up to overall normalization), by the norm square of the scalar product between the two histories, the state of the system and the measurement state,

$$p(k) = \frac{1}{N} |\Phi_k \cdot \Psi|^2. \quad (69)$$

This formula generalizes for arbitrary multi-time states and measurements. Of course, in order for the probabilities to be well defined, the POVM must entirely cover all the measurement periods (or only some of the measurement periods, in the case they are disconnected from the rest—see the discussion at the end of Sec. V). In case when the POVM covers all the measurement periods, then we use in the probability formula (69) the full state Ψ ; otherwise we use the reduced state.

Finally note that depending on the past and future boundary conditions, the “scalar product” of the two histories is not always just a complex number but may also be a bra vector, a ket, or a superposition of bra and ket pairs (see the discussion in Sec. VII). In those cases the “norm square” of the scalar product is to be taken as the norm of the resulting vectors.

In any case, conceptually, what formula (69) does is to generalize the usual notion that when a system is in a state $|\Psi\rangle$, the probability of finding it in the state $|\Phi\rangle$ is the norm square of the scalar product between $|\Psi\rangle$ and the measured state $|\Phi\rangle$, i.e., $|\langle\Phi|\Psi\rangle|^2$.

XII. MEASUREMENTS—OPEN QUESTIONS

As far as the states of the system are concerned, the situation is completely solved: any superposition of products of bra and ket vectors is a legitimate state of the system. Coming now to measurements, there are open questions.

As discussed in previous sections, a measurement can be described in two different ways. One way is to say exactly how the measurement is performed. Of course, every measurement for which we are given the explicit recipe of how to implement can, in principle, be performed. The second way of describing measurements is via its Krauss operators. It is in connection with this latter way of describing measurements that there are very interesting open problems.

In the case of ordinary one-time measurements, any set of Krauss operators (provided they fulfill the normalization condition (8) represents a possible measurement. This is not the case for multitime measurements. In fact there are two questions here.

First, is it the case that any superposition of products of bra and ket vectors as discussed above represents a possible

Krauss operator? In other words, given a Krauss operator, can we always find some multitime measurement such that this operator describes a particular outcome of the measurement? Or, to put it in a yet other way, can every arbitrarily given measurement state be implemented by a measurement?

Second, a measurement is described not by a single Krauss operator but by a whole set of them. For example, a von Neumann measurement is characterized by a complete basis of orthogonal projectors. Then what are the conditions that a set of operators must satisfy in order to describe a measurement? That is, even if each Krauss operator in a set is legitimate, i.e., if each Krauss operator separately describes an outcome of a possible measurement, does the set of them describe a possible measurement?

One major issue here is that measurements must obey causality. That is, by acting on the system in the future we should not be able to change the probabilities of outcomes of measurements in the past. While this condition is obeyed automatically for measurements that are sequences of one-time measurements, it is not the case that any arbitrary set of legitimate histories obeys this constraint. A somewhat similar situation is encountered when dealing with instantaneous nonlocal measurements. Indeed, there are sets of legitimate Krauss operators that are not measurable because they would lead to superluminal signaling [18]. That is, each operator separately is legitimate in the sense that it describes an outcome of a possible measurement, but together they cannot be measured.

Furthermore, it is also conceivable that there are cases of sets of Krauss operators that do not lead to causality violations but still there is no actual way to implement them in quantum mechanics.

Again, a similar situation is encountered in the case of nonlocal measurements: there are known cases [19] when a set of Krauss operators is unmeasurable although such a measurement would not lead to superluminal communication; the reason they are unmeasurable is that this would allow for establishing of nonlocal correlations stronger than allowed by quantum mechanics (Popescu-Rohrlich-type correlations [20]). Finally there may be other cases of nonmeasurable sets of Krauss operators in which the reason for unmeasurability is different from the above. Coming back to multitime measurements, we expect to find similar behavior. Partial answers to the above questions and other related problems are discussed in [21].

XIII. DISCUSSION: THE FLOW OF TIME

So far in this paper we approached the idea of multiple-time states from a rather formal point of view and avoided questions of interpretation. That is, we considered physical situations in which a quantum system is subjected to multiple stages of preparation and measurement. We then asked, given the preparation, what is the set of parameters that are relevant for inferring as well as possible the results of the measurements. What we found is that these parameters can be expressed as vectors in a “multitime” Hilbert space (which is the tensor product of Hilbert spaces associated with each time boundary between preparation and measurement

stages). Each vector, or mixture of vectors, describes a possible physical situation, and each possible physical situation can be described in this way. Clearly this is a basic fact about the structure of quantum mechanics and it is here to stay, no matter what philosophical interpretation we may associate with these states. It is very tempting, however, to go further and ask what does this all mean.

As we mentioned in the introduction, trying to give a philosophical interpretation for multitime states is certainly not easy. Indeed, even the interpretation of the ordinary (one-time) quantum state is highly controversial. We ourselves do not have one preferred interpretation of multitime states—in fact we have two of them, and we find both these points of view useful. We will describe here one of these points of view while the other one, the “block-time universe,” is presented in a forthcoming paper [22].

It is quite usual when thinking about the ordinary quantum state, to regard it not just as a static collection of parameters associated to some preparation stage, but to think that at each moment in time the system is described by a “state,” i.e., by a ket vector, and that this state evolves in time, being affected by all the interactions the system has. On one hand, one can view this “evolution” as a simple mathematical procedure by which we transform the parameters given at the preparation time t_0 into a more convenient form for computing what happens at the moment of interest t . In effect, we simply interpret part of the measurement stage, namely the period from t_0 to t as being part of the preparation stage. On the other hand, one may view the state as a physical object that evolves in time, undergoes collapses, etc. Obviously, although the probabilities we compute using these two different notions of state are the same, there is a great conceptual difference here—the state being a simple mathematical recipe for computing probabilities versus the state having an objective physical existence.

But consider now the simple example illustrated in Fig. 2(b). As far as the preparation is concerned, the system is described by the two-time state ${}_2\langle\Phi|\Psi\rangle_{t_1}$. Suppose further that the moment of interest is some time t , $t_1 < t < t_2$. We can then mathematically “evolve” the vectors $|\Psi\rangle$ forward and $\langle\Phi|$ backward until they reach that moment, t . The (ket) vector $|\Psi\rangle$ originates at t_1 , it is determined by the time boundary condition in the past, and “evolves” toward the future. The (bra) vector $\langle\Phi|$ originates at t_2 , it is determined by the time boundary condition in the future and “evolves” toward the past. Again, in effect all we do is to include the period from t_1 to t and the period from t to t_2 into the preparation stage instead of in the global measurement stage. On the other hand, we could think of the vectors $|\Psi\rangle$ and $\langle\Phi|$ as having objective physical meaning. This view however implies a dramatic conceptual change, far greater than that related to the interpretation of the standard quantum state. Indeed, the issue now is no longer only whether or not the quantum state has objective meaning or is just a mathematical tool for computing probabilities. The issue is now that of the *flow of time*.

To start with, it is a quite trivial fact that if we acquire new information we can affect the probabilities of events that happened in the past. This happens not only in quantum mechanics but in ordinary classical probabilities as well. For example suppose we have a bag with an equal number of

white and black balls and extract one ball at random and put it, without looking, into a bag containing only black balls. The probability that the ball is white is $1/2$. But suppose we then extract a ball from the second bag and see that the ball is white. In the light of this new information we can now infer that in this situation the probability that a white ball was extracted from the first bag is actually 1 and not $1/2$. The future information affects our knowledge about the past, but there is nothing surprising about this. Similarly, there is nothing surprising about the fact that post-selection at t_2 affects the probabilities for events that happened at the earlier time t . So, as long as we view the vector $\langle\Phi|$ just as a mathematical tool for calculating probabilities, it is nothing surprising that it “evolves” backward in time. But if $\langle\Phi|$ has objective meaning, then we have to admit that it really propagates backward in time.

At first sight it appears that the idea of a state propagating backward in time is ridiculous and should be immediately abandoned. The example of the classical postselection described above seems to show that an attempt to interpret the change in the statistics of results of experiments due to post-selection as a true backward-in-time influence is trivially wrong. However, we do feel that the situation is far more interesting in quantum mechanics. Indeed, there is a fundamental difference between post-selection in the classical and quantum cases. In the classical case, probabilities are only due to our *subjective* lack of knowledge. In principle, we could have had *complete* information about the system from the initial moment, and then there is no issue of probabilities and a future measurement does not really provide new information. On the other hand, it is one of the most important aspects of quantum mechanics—perhaps *the* most important aspect—that even when we have whole information about the past (say, we know the state $|\Psi\rangle$ at t_1), in general we still cannot predict with certainty the result of a later measurement. The later measurement does therefore yield truly new information about the system. In other words, the future is not completely determined by the past. Hence the whole notion of past and future in quantum mechanics is fundamentally different than in classical mechanics, and the whole idea of time flow may need to be reconsidered.

Of course, as we emphasized from the very beginning of this paper, all our results are fully consistent with ordinary quantum mechanics. In particular they could all be obtained using the traditional view of a single quantum state evolving in time. But we personally found it very useful to think of states propagating forward and backward in time. In particular, during each measurement period we think of two vectors, a ket propagating from the past time-boundary condition toward the future and a bra propagating from the future time-boundary condition toward the past. Each moment of time is therefore described by these two vectors [23]. Of course, more generally each time moment can be described by entangled bra and ket vectors or mixtures of them.

Thinking of vectors propagating forward and backward in time opens many new possibilities that we found very intriguing. In particular, one can ask about the possibility of having such time flow consistent with freewill. As we show elsewhere, that is consistent [24]. It is also possible to take forward and backward in time propagation as a starting point

for possible modifications of quantum mechanics. Finally, it is tempting to try and apply the idea of multi-time states in cosmological context, in particular to speculate about the possibility that the universe has both an initial and a final state which are given independently of each other.

ACKNOWLEDGMENTS

Y.A. and L.V. acknowledge support from Grant No. 990/06 of the Israel Science Foundation. L.V. acknowledges

support in part by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract No. 015848 and S.P. acknowledges support from the U.K. EPSRC Grant No. GR/527405/01 and the U.K. EPSRC QIP IRC project. J.T. acknowledges support from Grant No. N00173-07-1-G008 of the Naval Research Laboratory; J.T. and Y.A. acknowledge support from Grant No. 70NANB7H6138 of the National Institutes of Standards and Technology.

-
- [1] Y. Aharonov, P. G. Bergmann, and J. Lebowitz, *Phys. Rev.* **134**, B1410 (1964).
- [2] For a comprehensive list see Y. Aharonov and L. Vaidman, *Lect. Notes Phys.* **734**, 399 (2008).
- [3] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, *Phys. Rev. Lett.* **66**, 1107 (1991); K. J. Resch, J. S. Lundeen, and A. M. Steinberg, *Phys. Lett. A* **324**, 125 (2004); G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, *Phys. Rev. Lett.* **94**, 220405 (2005); D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, *ibid.* **92**, 043601 (2004); Q. Wang, F. W. Sun, Y. S. Zhang, J.-Li, Y. F. Huang, and G. C. Guo, *Phys. Rev. A* **73**, 023814 (2006); O. Hosten and P. Kwiat, *Science* **319**, 787 (2008); K. Yokota *et al.*, *New J. Phys.* **11**, 033011 (2009); J. S. Lundeen and A. M. Steinberg, *Phys. Rev. Lett.* **102**, 020404 (2009).
- [4] Y. Aharonov and D. Z. Albert, *Phys. Rev. D* **29**, 223 (1984); Y. Aharonov, D. Z. Albert, and S. S. D'Amato, *ibid.* **32**, 1975 (1985).
- [5] L. Vaidman, Ph.D. thesis, Tel Aviv University, 1987.
- [6] S. Popescu, Ph.D. thesis, Tel Aviv University, 1991.
- [7] R. B. Griffiths, *J. Stat. Phys.* **36**, 219 (1984).
- [8] M. Gell-Mann and J. B. Hartle, in *Physical Origins of Time Asymmetry*, edited by J. Halliwell *et al.* (Cambridge University Press, Cambridge, 1994).
- [9] J. G. Cramer, *Rev. Mod. Phys.* **58**, 647 (1986).
- [10] L. S. Schulman, *Time's Arrows and Quantum Measurement* (Cambridge University Press, New York, 1997).
- [11] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).
- [12] Y. Aharonov and L. Vaidman, *J. Phys. A* **24**, 2315 (1991).
- [13] L. Vaidman and I. Nevo, *Int. J. Mod. Phys. B* **20**, 1528 (2006).
- [14] L. Vaidman, *J. Phys. A* **40**, 3275 (2007).
- [15] Incidentally, note that the normalization factor in this case is *not* equal to 1. Indeed, $N = \sum_k \langle \Psi | A_k A_k^\dagger | \Psi \rangle = \langle \Psi | \sum_k A_k A_k^\dagger | \Psi \rangle$ but in general $\sum_k A_k A_k^\dagger \neq 1$ —it is only the expression $\sum_k A_k^\dagger A_k$ which is equal to 1 due to the normalization of POVMs, not the expression $\sum_k A_k A_k^\dagger$.
- [16] Note that since σ_x commutes with the measurement Hamiltonian, we have $\sigma_x(t_1 - \epsilon) = \sigma_x(t_1 + \epsilon)$ and similarly $\sigma_x(t_2 - \epsilon) = \sigma_x(t_2 + \epsilon)$; this allows us to denote them simply by $\sigma_x(t_1)$ and $\sigma_x(t_2)$.
- [17] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. D* **34**, 1805 (1986).
- [18] S. Popescu and L. Vaidman, *Phys. Rev. A* **49**, 4331 (1994).
- [19] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **74**, 012305 (2006).
- [20] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
- [21] S. Popescu and A. Short (unpublished).
- [22] Y. Aharonov, S. Popescu, and J. Tollaksen (unpublished).
- [23] In fact, this two-vector description of a quantum system at a particular time is the basic concept of most of the previous work on pre- and postselection; see Ref. [2] above. Note that the term used in these works to denote the two vectors associated to the same moment of time is a *two-state vector* and a superposition of two-state vectors is named a *generalized two-state vector*.
- [24] Y. Aharonov and S. Popescu (unpublished).