## Estimating preselected and postselected ensembles

Serge Massar<sup>1</sup> and Sandu Popescu<sup>2,3</sup>

<sup>1</sup>Laboratoire d'Information Quantique, C.P. 225, Université libre de Bruxelles (U.L.B.), Av. F. D. Rooselvelt 50, B-1050 Bruxelles, Belgium <sup>2</sup>H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom <sup>3</sup>Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom

(Received 22 February 2011; published 11 November 2011)

In analogy with the usual quantum state-estimation problem, we introduce the problem of state estimation for a pre- and postselected ensemble. The problem has fundamental physical significance since, as argued by Y. Aharonov and collaborators, pre- and postselected ensembles are the most basic quantum ensembles. Two new features are shown to appear: (1) information is flowing to the measuring device both from the past and from the future; (2) because of the postselection, certain measurement outcomes can be forced never to occur. Due to these features, state estimation in such ensembles is dramatically different from the case of ordinary, preselected-only ensembles. We develop a general theoretical framework for studying this problem and illustrate it through several examples. We also prove general theorems establishing that information flowing from the future is closely related to, and in some cases equivalent to, the complex conjugate information flowing from the past. Finally, we illustrate our approach on examples involving covariant measurements on spin-1/2 particles. We emphasize that *all* state-estimation problems can be extended to the pre- and postselected situation. The present work thus lays the foundations of a much more general theory of quantum state estimation.

DOI: 10.1103/PhysRevA.84.052106

PACS number(s): 03.65.Ta, 03.65.Wj

# I. INTRODUCTION

Quantum mechanics is usually formulated in terms of initial conditions. The state  $|\psi_i\rangle$  is given at time  $t_i$  and then evolves according to the Schrödinger equation. However, it was realized in [1] that one could also use a time-symmetric formulation in which one imposes both the initial condition  $|\psi_i\rangle$  at the initial time  $t_i$  and the final condition  $\langle \psi_f |$  at final time  $t_f$ . For an exposition we refer to the review [2] and to [3] where the concept of pre- and postselection has been extended to multiple time states. Pre- and postselection gives rise to a number of paradoxes and surprising effects that do not occur in the standard formulation of quantum theory. Studying them is a worthy endeavor: Pre- and postselected ensembles are the most detailed quantum ensemble one can prepare, hence arguably they are the fundamental quantum ensembles.

Independently of the above line of work, the past decades have seen the development of quantum information theory and, in particular, an in-depth study of quantum state estimation (see, e.g., [4–11]). The general problem of state estimation is, given an unknown quantum state  $\psi$ , to devise the best procedure to estimate the state.

In the present paper we try to bring together these two lines of inquiry. We consider the problem of estimating an unknown ensemble, when both the pre- and the postselected states are unknown. This differs from the usual state-estimation problem because information is flowing to the observer both from the past and from the future. In the first part of the paper we will show how to formulate this problem. Two new features are shown to appear: (1) information is flowing to the measuring device both from the past and from the future; (2) because of the postselection, certain measurement outcomes can be forced never to occur. Due to these features, state estimation in such ensembles is very different from the case of ordinary, preselected-only ensembles. For instance, in the usual stateestimation problem in which information arrives only from the past, measurements are described by positive operator valued measures (POVMs), whereas when information arrives both from the past and from the future, measurements are described by Kraus operators. In a second part of the paper, we prove general theorems establishing that information flowing from the future and the complex conjugate information flowing from the past are closely related and, in some cases, equivalent. In the final part of the paper, we illustrate this formalism on examples involving covariant measurements on spin-1/2 particles.

Considerable work has already been devoted to studying measurements on pre- and postselected ensembles. These works have mainly focused on the counterintuitive results which can be exhibited by "weak measurements" carried out at an intermediate time, between fixed pre- and postselected states [12]. This approach has applications for understanding quantum paradoxes-see the theoretical and experimental studies of Hardy's paradox [13-15], of superluminal light propagation [18,19], of polarization mode dispersion effects in optical networks [20], of cavity QED [21], as well as the recent approaches for measuring wave functions and trajectories of quantum particles [16,17]. Other experimental investigations of weak measurements are reported in [22]. In addition, it was shown, following the initial suggestion of [23], that weak measurements can have applications for high-precision measurements. These include the first observation of the spin Hall effect [24] and the observation of small transverse deflections of a light beam [25]; see also the proposals for measurements of charge [26] and of imaginary phase shifts [27]. Note that, in all these works, the pre- and postselected states are kept fixed, and it is the effects of the measurement which are investigated.

Closely related to the present work is [28] where it was shown that, in the presence of a fixed postselected state, some (preselected) states can be estimated to extremely high precision, with as consequence that the computational power of pre- and postselected ensembles is equivalent to the complexity class probabilistic polynomial time (PP). This shows that the presence of a postselected state can dramatically change the state-estimation problem because certain measurement outcomes can be forced never to occur. Another well-known example which can be interpreted in the same way (see discussion below) is the unambiguous state-estimation (USE) problem [29–31].

Here we both formulate in full generality the problem of state estimation in the presence of a postselected state and introduce the new problem of estimating an unknown pre- and postselected ensemble. At this stage we do not know if this approach will have applications (e.g., for highprecision measurements); rather, in this first work we are interested in the conceptual issue of formulating this problem and understanding its relation to the usual state-estimation problem.

## **II. SETTING UP THE PROBLEM**

#### A. Standard state-estimation problem

It will be useful to view the standard state-estimation problem as a game played between Alice and Bob: Alice chooses a parameter  $\theta$  and Bob must try to guess the value of  $\theta$ , given access to a quantum state  $|\psi(\theta)\rangle$ . We call  $\tilde{\theta}$  Bob's guess, which should be as close as possible (according to some merit function *F*) to the true value  $\theta$ . We now define this game with precision. To this end we introduce several additional actors that follow the instructions of either Alice or Bob. The whole state estimation problem consists of the following steps [described graphically in panel (a) of Fig. 1]:

(1) Alice chooses a (multidimensional) parameter  $\theta$  taken from some set  $\Theta$  according to a probability distribution  $p(\theta)$ . The set  $\Theta$  and probability distribution  $p(\theta)$  are known to Bob.

(2) The first actor is the preparer. He receives from Alice the value of  $\theta$  and prepares a quantum state  $|\psi(\theta)\rangle$ . The dependency of the quantum state on the parameter  $\theta$  [i.e., the function  $|\psi(\theta)\rangle$ ] is known to Bob.

(3) The second actor is the measurer who carries out a measurement on the state provided by the preparer. The POVM is chosen by Bob. Denote the outcome of the measurement by k. The measurer sends the value of k to Bob.

(4) Finally, Bob outputs a guess  $\tilde{\theta}(k)$  which depends on the value of k. The quality of the guess is measured by some merit function  $F(\theta, \tilde{\theta}(k))$ .

The experiment is then repeated many times. Each time Alice chooses a new value for  $\theta$  according to the probability distribution  $p(\theta)$ . The quality of the state-estimation procedure is measured by the average of the merit function *F*.

The above scenario may seem overly complicated. However, the separation of the roles of the different actors will become important in the pre- and postselected case.

Note that here and throughout this paper we neglect the free (unitary) evolution between preparation and measurement. Any such free evolution is supposed to be known to the parties and can therefore be taken into account.

#### B. Estimating pre- and postselected ensemble

We now set up, in parallel with the standard state-estimation problem, the problem of estimating a pre- and postselected ensemble.



FIG. 1. State estimation. Panel (a) illustrates the usual formulation of state estimation. The parameter  $\theta$  is chosen by Alice. It determines the quantum state  $|\psi(\theta)\rangle$  which is prepared by the preparer. The state is then sent to the measurer who carries out a quantum measurement, obtaining outcome k. The outcome is then sent to Bob who, on the basis of the value k, tries to guess what was the value of the parameter  $\theta$ . His guess is denoted  $\tilde{\theta}(k)$ . Panel (b) illustrates the estimation of pre- and postselected ensemble. The parameter  $\theta$  is chosen by Alice. It determines the pre- and postselected states  $|\psi_i(\theta)\rangle$  and  $\langle\psi_f(\theta)|$ . These are prepared by the preand postselectors respectively. The measurer carries out a quantum measurement at an intermediate time, obtaining outcome k. The role of the gate is to check if the pre- and postselections succeeded. Only if the pre- and postselections succeeded is the outcome k of the measurement sent to Bob. Upon receiving k, Bob tries to guess what was the value of the parameter  $\theta$ . His guess is denoted  $\tilde{\theta}(k)$ .

First of all note that, although the aim of pre- and postselection is to have a formulation in which past and future play symmetric roles, it is often useful to rephrase the problem in the language of usual quantum mechanics, in which the past and future play nonequivalent roles. Then by imposing that the postselection succeeds, one recovers a time-symmetric formulation. In the following paragraphs we take this more traditional point of view.

Once more, we view the estimate of a pre- and postselected ensemble as a game played between Alice and Bob: Alice chooses a parameter  $\theta$  and Bob must try to guess the value of  $\theta$ , given access to a pre- and postselected ensemble  $\langle \psi_f(\theta) || \psi_i(\theta) \rangle$ . We call  $\tilde{\theta}$  Bob's guess, which should be as close as possible (according to some merit function *F*) to the true value  $\theta$ . We now define this game with precision. To this end we introduce several additional actors. The estimation problem consists of the following steps [described graphically in panel (b) of Fig. 1]:

(1) Alice chooses a (multidimensional) parameter  $\theta$  taken from some set  $\Theta$  according to a probability distribution  $p(\theta)$ . The set  $\Theta$  and probability distribution  $p(\theta)$  are known to Bob. She sends the value of  $\theta$  to the preselector and to the postselector (see steps 2 and 4).

(2) The first actor is the preselector. He prepares a quantum state  $|\psi_i(\theta)\rangle$ . The dependency of the quantum state on the parameter  $\theta$  [i.e., the function  $|\psi_i(\theta)\rangle$ ] is known to Bob.

(3) The second actor is the measurer who carries out a measurement on the state provided by the preselector. The actions of the measurer are chosen by Bob. The result of the measurement consists of two pieces. First of all is the classical data produced by the measurement. Call this classical information k. Second is the quantum state of the system, which is modified by the action of the measurement. After the measurement is finished, the measurer sends the classical information k to a logical gate (see step 5) and sends the quantum state of the system to the postselector (see step 4).

(4) The third actor is the postselector. He checks whether or not the state sent to him by the measurer is  $\langle \psi_f(\theta) \rangle$ . He does this by measuring an observable that has  $\langle \psi_f(\theta) \rangle$  as one of its nondegenerate eigenstates. He sends the result of his measurement to the gate (see step 5). The dependency of the quantum state on the parameter  $\theta$  [i.e., the function  $\langle \psi_f(\theta) \rangle$ ] is known to Bob.

(5) The gate receives the value k from the measurer and the information on whether the postselection succeeded from the postselector. If the postselection succeeded, then the gate sends the result k of the measurement to Bob. If the postselection failed, then the gate instructs the preselector, postselector, and measurer that they must start over at step 2, the value of  $\theta$  being kept fixed.

(6) Finally, Bob outputs a guess  $\tilde{\theta}(k)$  which depends on the value of k. The quality of the guess is measured by some merit function  $F(\theta, \tilde{\theta}(k))$ .

In the present work we are interested in the information contained in the pre- and postselected ensemble itself (i.e., in the conditional information), given that we succeeded to prepare the ensemble. We want to exclude that information about the probability to actually prepare the ensemble can be used to estimate the ensemble. The role of the gate in the above procedure is to make this condition explicit. Indeed, because of the gate, Bob only receives the result of the measurement if the pre- and postselection succeeded and does not have any information on how many times steps 2, 3, 4 must be repeated before the postselection succeeds.

Note that one can consider the case where the postselector postselects a fixed state  $\langle 0 |$  which does not depend on  $\theta$ , or a combination  $\langle \psi_f(\theta) | \langle 0 |$  of a state which depends on  $\theta$  and a state that does not. We refer to these situations as the cases where there is a "fixed postselected state."

Note that, although the above setup is described within the usual framework of quantum theory, with evolution going from the past to the future, the final expressions for the quality of the ensemble estimation by Bob will be time symmetric. The pre- and postselected states will play the same role. This will become apparent below.

## III. STATE ESTIMATION IN PRESENCE OF FIXED POSTSELECTED STATE

Before studying the general case, it is useful to consider a simple situation; namely, the case in which the postselected state is fixed (i.e., independent of  $\theta$ ). Indeed this case is closest to the usual state-estimation problem, and several interesting results have already been obtained in the literature which can help develop an intuition. For definiteness we denote the fixed postselected state (0). When the postselected state is fixed, no information flows to the measurer from the future—there simply is no information in the postselected state since there is no uncertainty about it. So naively, one would expect that, in this case, the estimation problem is identical to the standard preselected-only case. However, as we now show, the existence of a postselected state completely changes the state-estimation problem.

One way to interpret this situation is that the measurer can reject certain measurement outcomes for free. Namely, if the measurement provides a useful outcome, the measurer prepares the state  $|0\rangle$ , and the postselection will succeed. On the contrary, if the measurement outcome is not useful, the measurer prepares the state  $|1\rangle$ , the postselection will fail, and he will be allowed to begin the measurement anew on a fresh copy of the state.

In this context, a dramatic example is provided by the problem of unambiguous state estimation [29–31]. Suppose the preselector prepares one of two nonorthogonal states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ , while the postselector selects the fixed state  $\langle 0|$ . The task of Bob is to say either "the state is  $\psi_1$ ," or "the state is  $\psi_2$ ," or "I do not know." The constraint is that, if one says that the state is  $\psi_1$  ( $\psi_2$ ), then one cannot make a mistake.

As is well known, in the standard unambiguous state discrimination problem (i.e., without postselection), such discrimination is possible for all pairs of states  $\psi_1$  and  $\psi_2$ , but the probability of success goes to zero as the states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  get closer and closer  $|\langle\psi_1|\psi_2\rangle| \rightarrow 1$ . However, in the presence of postselection Bob can always succeed. The procedure is for the measurer to perform the standard (preselected-only) unambiguous state discrimination and then prepare the system in the state  $|0\rangle$  whenever the measurement indicates  $\psi_1$  or  $\psi_2$ , but prepare the system in state  $|1\rangle$  whenever the outcome is "I do not know." The "I do not know" cases will thus never pass postselection and will never be counted.

A second spectacular example is taken from [28]. Suppose that the preparer prepares *n* identical particles all prepared in the same state  $|\uparrow_{\theta}\rangle = \cos \theta/2|\uparrow\rangle + \sin \theta/2|\downarrow\rangle$ . We are promised that either  $\theta \in S_+ = [\epsilon, \pi - \epsilon]$  or  $\theta \in S_- = [-\pi + \epsilon, -\epsilon]$  with  $0 < \epsilon < \pi/2$ . The task is to distinguish whether  $\theta$  belongs to set  $S_+$  or to set  $S_-$ . We are allowed a small error probability [say *P*(error) < 1/3]. If the measurer is promised that there is a fixed postselected state  $\langle 0|$ , then this task can be solved with  $n = O((\ln 1/\epsilon) \ln(\ln 1/\epsilon))$  particles. On the other hand in the usual formulation of state estimation with no postselection one needs  $n = O(\epsilon^{-2})$  particles. This is a huge difference and has dramatic consequences: essentially the same state-estimation problem is used in [28] to show that a quantum computer with access to a postselected state  $\langle 0|$  can solve PP complete problems.

More generally, one can take any state-estimation problem in the standard formulation and inquire how the quality of the state-estimation changes if there is a fixed postselected state  $\langle 0 |$ . Does one always get dramatic improvements as in the above two examples? Below we will analyze the cases of covariant measurements on *n* spin-1/2 particles in the state  $|\uparrow_{\Omega}^{\otimes n}\rangle$ , and of covariant measurements when the *n* spins are in the state  $|\uparrow_{\Omega}^{\otimes n/2}\downarrow_{\Omega}^{\otimes n/2}\rangle$  (for *n* even). We will see that, in these cases, the presence of a fixed postselected state  $\langle 0 |$  can sometimes give a small increase in fidelity, but nothing as spectacular as in the above examples. Before presenting these results, we first give a general framework for describing state estimation in the pre- and postselected context.

### **IV. GENERAL FORMALISM**

We give here general expressions for state estimation in the presence of both pre- and postselection. In Sec. V we argue why these are the natural generalizations of the standard formalism.

## A. Standard state estimation

For definiteness we first recall standard state-estimation theory. In this case the most general measurement is a positive operator valued measure (POVM) described by operators  $M_k$ which are positive and sum to the identity:

$$M_k \ge 0, \quad \sum_k M_k = \mathbb{1}.$$
 (1)

The probability of finding outcome k if the state was  $|\psi\rangle$  is

$$P(k|\psi) = \langle \psi | M_k | \psi \rangle.$$
<sup>(2)</sup>

The average value of the merit function can then be expressed as

$$F = \int d\theta p(\theta) \sum_{k} \langle \psi(\theta) | M_k | \psi(\theta) \rangle F(\theta, \tilde{\theta}(k)).$$
(3)

We note the well-known fact that POVMs with rank-one operators are the most informative (for a proof see the argument at the end of Sec. IV B).

## B. State estimation with fixed postselected state

When there is a fixed postselected state (the situation discussed in Sec. III), the preceding formalism must be generalized. The probability of finding outcome k if the state was  $|\psi\rangle$  is

$$P_{M}(k|\psi) = \frac{\langle \psi | M_{k} | \psi \rangle}{\sum_{k'} \langle \psi | M_{k'} | \psi \rangle},\tag{4}$$

where the operators  $M_k$  are positive but no longer normalized (we say they are "subnormalized"):

$$M_k \ge 0, \quad \sum_k M_k \leqslant \mathbb{1}.$$
 (5)

The average value of the merit function can then be expressed as

$$F = \int d\theta p(\theta) \sum_{k} \frac{\langle \psi(\theta) | M_k | \psi(\theta) \rangle}{\sum_{k'} \langle \psi(\theta) | M_{k'} | \psi(\theta) \rangle} F(\theta, \tilde{\theta}(k)).$$
(6)

We now show that POVMs with rank-one operators are the most informative, whether or not there is a fixed postselected state. Consider an arbitrary POVM with elements  $M_k$  and associated estimator  $\tilde{\theta}(k)$ . Since the  $M_k \ge 0$  are positive operators, we can write them as  $M_k = \sum_j |m_{kj}\rangle \langle m_{kj}|$ , with  $|m_{kj}\rangle$  being unnormalized states. Consider the refined POVM with elements  $M_{kj} = |m_{kj}\rangle \langle m_{kj}|$ . If to the refined POVM element  $M_{kj}$  we associate the same estimator  $\tilde{\theta}(k)$  as for the original POVM, then the value of the merit function does

not change [to see this note that the denominator in Eq. (6) does not change when one replaces the original POVM by the refined POVM with elements  $M_{kj}$ ]. Hence the value of the merit functions for POVMs with rank-one elements are always at least as large as the merit functions for unrefined POVMs.

## C. Estimation of pre- and postselected ensembles

In the case of estimation of pre- and postselected ensembles, the measurement operators are no longer POVM elements, but Kraus operators. Kraus operators describe the most general evolution of an open quantum system:

$$\rho \to \sum_{k} A_k \rho A_k^{\dagger}, \tag{7}$$

and are normalized according to

$$\sum_{k} A_k^{\dagger} A_k = \mathbb{1}.$$
(8)

Kraus operators are the appropriate operators to describe interaction with a pre- and postselected ensemble because Kraus operators consist of a ket-bra which points both toward the past and toward the future:

$$A_k = \sum_{l} |\phi_k^l\rangle \langle \varphi_k^l|.$$
(9)

In addition, if there is a fixed postselected state, one must also modify the normalization condition and replace the equality in Eq. (8) by an inequality (we say the Kraus operators are "subnormalized").

We thus have for the probability of obtaining outcome k conditional on the pre- and postselected states  $\langle \psi_f || \psi_i \rangle$ :

$$P_A(k|\psi_f,\psi_i) = \frac{|\langle \psi_f | A_k | \psi_i \rangle|^2}{\sum_{k'} |\langle \psi_f | A_{k'} | \psi_i \rangle|^2},$$
(10)

with the normalization

$$\sum_{k} A_{k}^{\dagger} A_{k} = 1 \quad \text{no additional postselection,} \tag{11}$$

or

$$\sum_{k} A_{k}^{\dagger} A_{k} \leqslant \mathbb{1} \quad \text{fixed postselected state.}$$
(12)

Note that, in this expression,  $\langle \psi_f |$  need not belong to the same Hilbert space as  $|\psi_i\rangle$ , as Kraus operators allow the description of the evolution of a system belonging to one Hilbert space into a system belonging to another Hilbert space. The average value of the merit function can then be expressed as

$$F = \int d\theta p(\theta) \sum_{k} \frac{|\langle \psi_f(\theta) | A_k | \psi_i(\theta) \rangle|^2}{\sum_{k'} |\langle \psi_f(\theta) | A_{k'} | \psi_i(\theta) \rangle|^2} F(\theta, \tilde{\theta}(k)).$$
(13)

# V. INTERACTION BETWEEN SYSTEM AND MEASURING DEVICE

We now go back to the setups presented in Sec. II and argue why the expressions given in Secs. IV B and IV C are a natural generalization of the Born rule to the case of pre- and postselected ensembles. Note that we cannot provide a proof that they constitute the only possible generalization, but only plausibility arguments.

To derive the above expressions for the probability of obtaining the outcome k, we go back to the general setup described in Sec. II and Fig. 1(b) and describe it in the standard quantum formalism.

To simplify the problem, let us first note that, if there is a fixed postselected state, then we can, without loss of generality, take it to be a single qubit. Indeed, in this case the most general procedure is for the postselector to postselect that the state belongs to a subspace. Denote by  $\Pi$  the projector onto this subspace and by  $|\Psi\rangle$  the state just before projection onto  $\Pi$ . We now describe an equivalent postselection in which only a single qubit is used. First we add to the system space an ancillary qubit initially in the state  $|0\rangle$ . The state is thus  $|\Psi\rangle \otimes |0\rangle$ . Next we carry out the unitary

$$U = \Pi \otimes |0\rangle \langle 0| + (\mathbb{1} - \Pi) \otimes (|1\rangle \langle 0| + |0\rangle \langle 1|)$$

(i.e. a controlled-NOT gate), where the projection onto  $\Pi$  acts as control. Finally, we postselect that the qubit is in state  $|0\rangle$ . The probability of success of this postselection is exactly the same as the original one, hence the two methods are equivalent.

We now go back to the general setup described in Sec. II and Fig. 1(b). Let  $|\psi_i\rangle_S \in H_S$  be the initial state of the system. We adjoin to  $H_S$  two additional Hilbert spaces. First, there is the Hilbert space  $H_R$  of the measurement register. The initial state of the measurement register is  $|0\rangle_R$ . If the final state is  $|k\rangle_R$ , then the outcome of the measurement will be k. Second there is the Hilbert space of single qubit  $H_P$  which is used in case there is a fixed postselection. The initial state of this qubit is  $|0\rangle_P$ . The fixed postselection will succeed if the final state of this qubit is still  $|0\rangle_P$ . The initial state is thus

$$|\psi_i\rangle_S \otimes |0\rangle_R \otimes |0\rangle_P, \tag{14}$$

where the subscripts denote to which Hilbert space each state belongs. The action of the measurer can be described by a unitary evolution U that entangles the Hilbert spaces  $H_S$ ,  $H_R$ , and  $H_P$ . This yields the state

$$U|\psi_{i}\rangle_{S} \otimes |0\rangle_{R} \otimes |0\rangle_{P}$$

$$= \sum_{k} \sum_{x=0,1} [\mathbb{1}_{S} \otimes |k\rangle_{R} \langle k| \otimes |x\rangle_{P} \langle x|] U|\psi_{i}\rangle_{S} \otimes |0\rangle_{R} \otimes |0\rangle_{P}$$

$$= \sum_{k} \sum_{x=0,1} (A_{kx}|\psi_{i}\rangle_{S}) \otimes |k\rangle_{R} \otimes |x\rangle_{P}, \qquad (15)$$

where unitarity of U imposes that

$$\sum_{kx} A_{kx}^{\dagger} A_{kx} = \mathbb{1}_{\mathcal{S}}.$$
 (16)

Consider first the case where the only postselected state is the fixed state  $_P\langle 0|$ . The probability to find the register in state k and for the postselection to succeed is

$$P(k, x = 0 | \psi_i) = \langle \psi_i | A_{k0}^{\dagger} A_{k0} | \psi_i \rangle.$$
(17)

Because of the presence of the gate that checks that the postselection succeeded, the relevant quantity is the probability

to find the register in state k conditional on the postselection having succeeded. This is

$$P(k|\psi_i, x=0) = \frac{\langle \psi_i | A_{k0}^{\dagger} A_{k0} | \psi_i \rangle}{\sum_{k'} \langle \psi_i | A_{k'0}^{\dagger} A_{k'0} | \psi_i \rangle} = \frac{\langle \psi_i | M_k | \psi_i \rangle}{\sum_{k'} \langle \psi_i | M_{k'} | \psi_i \rangle},$$
(18)

where  $M_k = A_{k0}^{\dagger} A_{k0}$  are POVM elements. They are hermitian, positive  $M_k \ge 0$ , and subnormalized:  $\sum_k M_k \le 1$ . We thereby obtain the formalism of Sec. IV B.

Consider now the case where one postselects both that the final state is  $_{S}\langle\psi_{f}|$  and that there is the fixed postselected state  $_{P}\langle 0|$ . The amplitude of finding state  $_{S}\langle\psi_{f}\otimes_{R}\langle k|\otimes_{P}\langle 0|$  is  $\langle\psi_{f}|A_{k0}|\psi_{i}\rangle$ . The probability of this event is

$$P(k, x = 0, \psi_f | \psi_i) = |\langle \psi_f | A_{k0} | \psi_i \rangle|^2.$$
(19)

Because of the presence of the gate that checks that the postselection succeeded, the relevant quantity is the probability to find the register in state k conditional on the postselections having succeeded. This is

$$P(k|\psi_i, \psi_f, x = 0) = \frac{|\langle \psi_f | A_{k0} | \psi_i \rangle|^2}{\sum_{k'} |\langle \psi_f | A_{k'0} | \psi_i \rangle|^2},$$
(20)

where the operators  $A_{k0}$  are arbitrary, except for the condition  $\sum_{k} A_{k0}^{\dagger} A_{k0} \leq \mathbb{1}_{S}$ .

Note that, if there is no fixed postselection onto  $_P(0)$ , then the above calculation carries through with the Hilbert space  $H_P$ (and hence the index x) omitted. One then obtains the standard normalization for the Kraus operators  $\sum_k A_k^{\dagger} A_k = \mathbb{1}_S.^1$ 

Note that, if the postselected state  $\langle \psi_f |$  belongs to a different Hilbert space then the preselected state  $|\psi_i\rangle$ , then the above calculation carries through unchanged, except that one must enlarge the Hilbert space to contain both the spaces of the initial state and that of the final state.

The above analysis thus leads to the formalism of Sec. IV C.

## VI. INFORMATION FLOW FROM PAST AND FROM FUTURE

#### A. Two theorems

How well can we estimate the parameter  $\theta$  in the above situations? Obviously, the estimation can be done better in the pre- and postselected ensemble than if one is given the preselected state  $|\psi_i(\theta)\rangle$  only, since the postselected state provides additional information. But how much more information? We now show that the relevant comparison is with the preselected

<sup>&</sup>lt;sup>1</sup>In some cases, the postselection of a state  $\langle \psi_f |$  by itself implies the existence of a fixed postselection. For instance, suppose that  $\langle \psi_f | = \alpha \langle 0 | + \beta \langle 1 |$  belongs to a two-dimensional subspace of a three-dimensional space with basis  $\langle 0 |, \langle 1 |, \langle 2 |$ . Then whenever the measurer does not want an outcome to occur, he prepares the state  $|2\rangle$ , and the postselection never occurs. On the other hand, it may be that the Hilbert space to which  $\langle \psi_f |$  belongs is intrinsically two dimensional (e.g., polarization of a photon). In this case there is a difference between the presence or not of a fixed postselection. For this reason we keep the two notions distinct in the present paper.

tensor product state  $|\overline{\psi_f(\theta)}\rangle \otimes |\psi_i(\theta)\rangle$ , where  $|\overline{\psi}\rangle$  is the state obtained by complex conjugating the coefficients of  $|\psi\rangle$  in a basis:  $|\psi\rangle = \sum_k c_k |k\rangle \rightarrow |\overline{\psi}\rangle = \sum_k \overline{c_k} |k\rangle$ .

Some intuition for this mapping can be obtained by recalling that in a pre- and postselected ensemble, the preselected state arrives from the past, whereas the postselected state arrives from the future. It is thus natural that it behaves like the time reversal of a preselected state. And time reversal is realized mathematically by complex conjugation. Another motivation follows from the remark made in [3] that it is possible to realize a pre- and postselected ensemble  $\langle \psi_f(\theta) || \psi_i(\theta) \rangle$  by: (A) preselecting the tensor product state  $|\overline{\psi_f(\theta)} \rangle \otimes |\psi_i(\theta) \rangle$ , (B) postselecting the maximally entangled state  $\sum_k \langle k | \otimes \langle k |$  (where  $|k \rangle$  is the basis in which the complex conjugation is defined), and (C) at intermediate times acting only on the second system.

Thus, both lines of reasoning suggest that, to a postselected state  $\langle \psi_f |$ , we should associate the preselected complex conjugate state  $|\overline{\psi}_f \rangle$ . The following results put this intuition on a firm basis. To state them we use the following notation:

Denote by  $|\phi\rangle \in H^d$  and  $|\psi\rangle \in H^{d'}$  states belonging to Hilbert spaces of dimension *d* and *d'*, respectively. Denote by  $|\overline{\phi}\rangle$  the state obtained from  $|\phi\rangle$  by complex conjugation in a (fixed but arbitrary) basis. Consider a subnormalized POVM acting on the tensor product space  $H^d \otimes H^{d'}$  with rank-one elements:  $M_k = |m_k\rangle \langle m_k|, \sum_k M_k \leq 1$ . The probability of outcome *k* when the state is the tensor product  $|\overline{\phi}\rangle \otimes |\psi\rangle$  is given by Eq. (4):

$$P_M(k|\overline{\phi}\psi) = \frac{|\langle m_k|\phi\rangle\otimes|\psi\rangle|^2}{\sum_{k'}|\langle m_k|\overline{\phi}\rangle\otimes|\psi\rangle|^2}.$$
 (21)

Consider a subnormalized completely positive (CP) map described by Kraus operators  $A_k : H^{d'} \to H^d, \sum_k A_k^{\dagger} A_k \leq 1$ . The probability of finding outcome k using operators  $A_k$  in the pre- and postselected ensemble  $\langle \phi || \psi \rangle$  is given by Eq. (10):

$$P_A(k|\phi\psi) = \frac{|\langle\phi|A_k|\psi\rangle|^2}{\sum_k |\langle\phi|A_k|\psi\rangle|^2}.$$
(22)

Then we have:

Theorem 1. For any subnormalized rank-one POVM  $M_k$ , there exists a subnormalized CP map  $A_k$  such that  $P_A(k|\phi\psi) = P_M(k|\overline{\phi}\psi)$ . Conversely, for any subnormalized CP map  $A_k$ , there exists a subnormalized POVM  $M_k$ , such that  $P_M(k|\overline{\phi}\psi) = P_A(k|\phi\psi)$ .

This result, combined with the fact that rank-one POVM's are always the most informative (see end of Sec. IV B), shows that the problem of estimating the unknown preselected state  $|\overline{\phi}\rangle|\psi\rangle$  in the presence of a fixed postselected state is completely equivalent to estimating the pre- and postselected state  $\langle \phi ||\psi\rangle$  in the presence of a fixed postselected state.

In the case where there is no fixed postselected state, we have implication in one direction only:

Theorem 2. For any normalized rank-one POVM  $M_k$  $(\sum_k M_k = 1)$ , there exists a normalized CP map  $A_k$  $(\sum_k A_k^{\dagger} A_k = 1)$  such that  $P_A(k|\phi\psi) = P_M(k|\overline{\phi}\psi)$ .

This result shows that the problem of estimating the unknown preselected state  $|\overline{\phi}\rangle|\psi\rangle$  without any fixed postselection is always at least as hard as estimating the pre- and postselected state  $\langle \phi || \psi \rangle$  (without any fixed postselection).

One would expect that the converse of Theorem 2 should not hold, since the presence of some postselection should give additional discriminating power. Below we show that this intuition is correct and provide an example showing that the converse of Theorem 2 does not hold; that is, in some cases estimating the unknown preselected state  $|\overline{\phi}\rangle|\psi\rangle$  without any fixed postselection is harder than estimating the pre- and postselected state  $\langle \phi ||\psi \rangle$  without any fixed postselection.

## **B.** Proof of theorems

*Proof of Theorem 1.* Part (1): Consider the rank-one subnormalized POVM  $M_k = |m_k\rangle\langle m_k|$ . We will construct the Kraus operators  $A_k$  so that the probabilities of outcomes of measurement  $A_k$ ,  $P_A(k|\phi\psi)$ , are identical to the probabilities of outcomes of the measurement M:  $P_A(k|\phi\psi) = P_M(k|\overline{\phi}\psi)$ .

Let us rewrite

$$|\langle m^{k} | \overline{\phi} \rangle | \psi \rangle|^{2} = \left| \sum_{\alpha \beta} m^{k}_{\alpha \beta} \langle \alpha | \overline{\phi} \rangle \langle \beta | \psi \rangle \right|^{2}, \qquad (23)$$

where  $m_{\alpha\beta}^k$  are the coefficients of  $|m_k\rangle$  in basis  $|\alpha\rangle \otimes |\beta\rangle$ ,

$$m_{\alpha\beta}^{k} = \langle m_{k} | | \alpha \rangle \otimes | \beta \rangle, \qquad (24)$$

and  $|\alpha\rangle$  is the basis in which complex conjugation of  $|\phi\rangle$  is defined. Let us now consider the Kraus operators

$$A^{k} = \sum_{\alpha\beta} |\alpha\rangle\langle\beta|A^{k}_{\alpha\beta}, \qquad (25)$$

with the choice

$$A^k_{\alpha\beta} = m^k_{\alpha\beta} / \sqrt{d}, \qquad (26)$$

where *d* is the dimension of the Hilbert space of state  $|\phi\rangle$ . (The reason for this choice of normalization will appear below.) We then have

$$\left|\langle \phi | A_k | \psi \rangle\right|^2 = \langle \overline{\phi} | \otimes \langle \psi | M_k | \overline{\phi} \rangle \otimes | \psi \rangle / d.$$
(27)

Inserting this identity into Eqs. (21) and (22) proves the equality  $P_A(k|\phi\psi) = P_M(k|\overline{\phi}\psi)$ .

Note that we have  $\sum_{k\alpha} \overline{A}^{k}_{\alpha\beta'} A^{k}_{\alpha\beta} = \sum_{k\alpha} \overline{m}^{k}_{\alpha\beta'} m^{k}_{\alpha\beta} / d$ . Using the subnormalization  $\sum_{k} M_{k} \leq 1$ , and the fact that the partial trace preserves inequalities between matrices (i.e., if *A* and *B* act on  $H \otimes H'$ , and  $A \leq B$ , then  $\operatorname{Tr}_{H} A \leq \operatorname{Tr}_{H} B$ ), we have

$$\sum_{k\alpha} \overline{A}^{k}_{\alpha\beta'} A^{k}_{\alpha\beta} \leqslant \sum_{\alpha} \delta_{\beta\beta'} / d = \delta_{\beta\beta'}, \qquad (28)$$

where the inequality is taken to be a matrix inequality, not an inequality for each  $\beta\beta'$ . This implies that the Kraus operators are also subnormalized  $\sum_k A_k^{\dagger} A_k \leq 1$ .

Part (2): Consider the subnormalized Kraus operators  $A_k$ . We will construct a rank-one POVM  $M_k = |m_k\rangle\langle m_k|$  such that the probabilities of outcomes of measurement  $M_k$ ,  $P_M(k|\overline{\phi}\psi)$  are identical to the probabilities of outcomes of the measurement  $A_k$ :  $P_M(k|\overline{\phi}\psi) = P_A(k|\phi\psi)$ . The argument is essentially the reverse of the argument given in part (1). We write the Kraus operators and POVM elements using the

notation of Eqs. (24) and (25) and choose the  $m_{\alpha\beta}^k$  according to

$$m^k_{\alpha\beta} = c A^k_{\alpha\beta},\tag{29}$$

where c > 0 is a constant we will fix below. With this choice we have

$$\langle \overline{\phi} | \otimes \langle \psi | M_k | \overline{\phi} \rangle \otimes | \psi \rangle = c^2 | \langle \phi | A_k | \psi \rangle |^2$$
(30)

Inserting this identity into Eqs. (21) and (22) proves the equality  $P_A(k|\phi\psi) = P_M(k|\phi\psi)$ .

Note that we have  $\sum_k \overline{m}_{\alpha'\beta'}^k m_{\alpha\beta}^k = c^2 \sum_k \overline{A}_{\alpha'\beta'}^k A_{\alpha\beta}^k \ge 0$ , which is a positive operator. By choosing *c* sufficiently small, we can ensure that  $\sum_k M_k \leq 1$ .

Proof of Theorem 2. Consider the rank-one normalized POVM  $M_k = |m_k\rangle \langle m_k|$ ,  $\sum_k M_k = 1$ . We will construct normalized Kraus operators  $A_k$  so that the probabilities of outcomes of measurement  $A_k$ ,  $P_A(k|\phi\psi)$ , are identical to the probabilities of outcomes of the measurement  $M: P_A(k|\phi\psi) =$  $P_{\mathcal{M}}(k|\overline{\phi}\psi)$ . We proceed exactly as in the proof of Theorem 1, part 1 and, in particular, make the choice of Kraus elements Eq. (26) (with the same normalization). Then we have equality in Eq. (28) which shows that the Kraus operators are also normalized.

#### C. Example showing that converse of Theorem 2 does not hold

The following example showing that the converse of Theorem 2 does not hold is based on a version of the unambiguous state discrimination problem.

Denote by  $|\psi_{\pm}\rangle = \alpha |0\rangle \pm \beta |1\rangle$  two nonorthogonal states. Denote the orthogonal states by  $|\psi_{\pm}^{\perp}\rangle = \beta |0\rangle \mp \alpha |1\rangle$ . Similarly, denote by  $|\phi_{\pm}\rangle = \sqrt{1 - \epsilon^2} |0\rangle \pm \epsilon |1\rangle$  two nonorthogonal states. All coefficients  $\alpha > \beta > 0$ ,  $\epsilon > 0$  are real, and all states are normalized:  $\alpha^2 + \beta^2 = 1$ . We will be interested in the case where  $\alpha, \beta$  are fixed and  $\epsilon$  is very small:  $0 \le \epsilon \ll 1$ . Note that, for a counter example, it is in principle sufficient to consider the case when  $\epsilon = 0$ . However, this case is special since the states  $|\phi_{+}\rangle$  are then equal and carry no information. By considering the cases when  $\epsilon > 0$ , we show that counterexamples are rather common.

Consider the problem in which one receives either the states  $|\psi_{+}\rangle |\phi_{+}\rangle$ , or the states  $|\psi_{-}\rangle |\phi_{-}\rangle$ . The measurement can have three outcomes: +, -, 0. Outcome + can only occur if the state was  $|\psi_+\rangle |\phi_+\rangle$  [i.e.,  $P_M(+|-) = 0$ ]. Outcome – can only occur if the state was  $|\psi_{-}\rangle|\phi_{-}\rangle$  [i.e.,  $P_M(-|+) = 0$ ]. Outcome 0 can occur in all cases. The aim is to minimize the probability of occurrence of outcome 0. The theory of USE [29-31] shows that the optimal discrimination probabilities without fixed postselection are  $P_M(+|+) = P_M(-|-) = \sqrt{1 - (\alpha^2 - \beta^2)^2 (1 - 2\epsilon^2)^2} =$  $\sqrt{1 - (\alpha^2 - \beta^2)^2} + O(\epsilon^2)$  where we expand to first order in  $\epsilon^2$ , and  $P_M(0|+) = P_M(0|-) = 1 - \sqrt{1 - (\alpha^2 - \beta^2)^2} + O(\epsilon^2).$ 

Now consider the related problem where some of the information is flowing from the future. The aim is to distinguish between the two ensembles  $\langle \phi_+ || \psi_+ \rangle$  and  $\langle \phi_- || \psi_- \rangle$ . The measurement, given by Kraus operators, can have three outcomes: +, -, 0. Outcome + can only occur if the ensemble is  $\langle \phi_+ || \psi_+ \rangle$  [i.e.,  $P_A(+|-) = 0$ ]. Outcome – can only occur if the ensemble is  $\langle \phi_- || \psi_- \rangle$  [i.e.,  $P_A(-|+) = 0$ ]. Outcome 0

can occur in all cases. The aim is to minimize the probability of occurrence of outcome 0. To this end we consider the Kraus operators  $A_+ = |0\rangle \langle \psi_-^{\perp}| / \sqrt{2\alpha^2}$ ,  $A_- = |0\rangle \langle \psi_+^{\perp}| / \sqrt{2\alpha^2}$ , and  $A_0 = |1\rangle \langle 0|\sqrt{1-\beta^2/\alpha^2}$ . One checks that  $A_+^{\dagger}A_+ + A_-^{\dagger}A_- +$  $A_0^{\dagger}A_0 = 1$ . One easily computes that  $P_A(+|+) = P_A(-|-) =$  $1 - O(\epsilon^2)$  and  $P_A(0|+) = P_A(0|-) = O(\epsilon^2)$ .

Thus, in this example, in the absence of fixed postselection, the outcome 0 occurs with probability  $P_M(0|+) = P_M(0|-) =$ O(1) when all the information comes from the past and occurs with probability  $P_A(0|+) = P_A(0|-) = O(\epsilon^2)$  when some of the information flows from the future. The gain is dramatic. The origin of the gain is that the states  $|\phi_{+}\rangle$  contain very little information, since  $\epsilon$  is small, but when the state  $\langle \phi_{\pm} |$  is postselected, it can be used to strongly decrease the probability of occurrence of the unwanted outcome 0.

# VII. COVARIANT MEASUREMENTS **ON SPIN-1/2 PARTICLES**

## A. Stating the problem

We illustrate the above formalism by the case of covariant measurements on spin-1/2 particles. Suppose that the parameter to be estimated is a direction uniformly distributed on the sphere:  $\theta \equiv \Omega \in S_2$ . This direction is encoded in the preand postselected state of spin-1/2 particles. The spins are polarized in the direction  $\Omega$  or the opposite direction  $-\Omega$ . The task is to estimate the direction  $\Omega$ . To each outcome k of the measurement one thus associates a guessed direction  $\tilde{\Omega}(k)$ . The quality of the estimate is gauged with the fidelity  $F = \cos^2 \Phi/2$  where  $\Phi$  is the angle between the true direction  $\Omega$  and the guessed direction  $\tilde{\Omega}$ .

When there is no postselection the solution of this stateestimation problem is well known; see [5-11]. We summarize some of these results. Throughout this section we denote by N the total number of spins.

(1) When the initial state consists of N parallel spins  $|\uparrow_{\Omega}^{\otimes N}\rangle$ ,

the optimal fidelity is (N + 1)/(N + 2). (2) When the initial state  $|\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2}\rangle$  consists of N/2 spins in direction  $\Omega$  and N/2 spins in direction  $-\Omega$  (here N is even), the optimal fidelity is 0.7887 for N = 2, 0.8848 for N = 4, and 0.9235 for N = 6.

(3) There is an optimal encoding of the direction  $\Omega$  into states of the form  $R_{\Omega}|\psi\rangle$ , where  $R_{\Omega}$  is the rotation that maps direction +z onto the direction  $\tilde{\Omega}$ . In the case of N spins, the optimal fidelity for the optimal choice of  $\psi$  is 0.7887 for N = 2, 0.8873 for N = 4, and 0.9306 for N = 6.

The standard approach to these estimation problems is to use covariant measurements. By covariant measurements we mean that there exists a POVM element  $M_{\tilde{\Omega}}$  for each possible guessed direction  $\tilde{\Omega} \in S_2$ . These POVM elements are related to each other by  $M_{\tilde{\Omega}} = R_{\tilde{\Omega}} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger}$  where  $R_{\tilde{\Omega}}$  is the rotation that maps direction +z onto direction  $\tilde{\Omega}$  and  $M_{\tilde{z}}$  is the POVM element associated to the guessed direction +z.

Here we consider the problem of estimating the unknown preselected state  $|\uparrow_{\Omega}^{\otimes n}\downarrow_{\Omega}^{\otimes m}\rangle$  in the presence of a fixed postselected state, or the unknown pre- and postselected ensemble  $\langle \uparrow_{\Omega}^{\otimes k} \downarrow_{\Omega}^{\otimes l} || \uparrow_{\Omega}^{\otimes (n-l)} \downarrow_{\Omega}^{\otimes (m-k)} \rangle$  in the presence of a fixed postselected state.

Covariant measurements can also be used in the case of measurements on pre- and postselected ensembles. In the usual approach to state estimation used in [5-11] one can show that covariant measurements perform at least as well as any other measurements. We have not been able to show this in the present case because of the more complicated form of the fidelities. However, covariant measurements are an interesting category to consider because they allow for detailed calculations. Here we will restrict ourselves to covariant measurements could perform better for the problems considered here.

We therefore consider subnormalized POVM elements that are related through  $M_{\bar{\Omega}} = R_{\bar{\Omega}} M_{\bar{z}} R_{\bar{\Omega}}^{\dagger}$ , or subnormalized Kraus operators that are related through  $A_{\bar{\Omega}} = R_{\bar{\Omega}} A_{\bar{z}} R_{\bar{\Omega}}^{\dagger}$ , where  $R_{\bar{\Omega}}$ is the rotation that maps direction +z onto direction  $\tilde{\Omega}$ .

# B. Covariant measurements and equivalence between information flowing from past and future

We note that spin-1/2 states pointing in opposite directions are related through convex conjugation and the action of a fixed unitary:  $|\downarrow_{\Omega}\rangle = i\sigma_y|\uparrow_{\Omega}\rangle$ . Therefore, theorems 1 and 2 apply. We also expect theorems 1 and 2 to apply if we restrict ourselves to covariant measurements. We now show that this is indeed the case.

*Theorem 3.* The relations and equivalences between estimation of preselected ensembles and pre- and postselected ensembles expressed in theorems 1 and 2 also hold if one considers covariant measurements (as defined above) on the ensembles  $|\uparrow_{\Omega}^{\otimes n}\downarrow_{\Omega}^{\otimes m}\rangle$  and  $\langle\uparrow_{\Omega}^{\otimes k}\downarrow_{\Omega}^{\otimes l}||\uparrow_{\Omega}^{\otimes (n-l)}\downarrow_{\Omega}^{\otimes (m-k)}\rangle$ , for any k,l, with n,m fixed.

*Proof of Theorem 3*. The proof follows easily from the proofs of theorems 1 and 2.

Note that, without changing the state-estimation problem, we can consider the equivalent ensembles  $|\uparrow_{\Omega}^{\otimes n} \overline{\uparrow_{\Omega}^{\otimes m}}\rangle$  and  $\langle\uparrow_{\Omega}^{\otimes k} \overline{\uparrow_{\Omega}^{\otimes l}}||\uparrow_{\Omega}^{\otimes (n-l)} \overline{\uparrow_{\Omega}^{\otimes (m-k)}}\rangle$  since they differ from the original ensemble only by fixed unitaries.

A covariant rank-one POVM element on the above state has the form  $M_{\tilde{\Omega}} = |m_{\tilde{\Omega}}\rangle \langle m_{\tilde{\Omega}}|$  with

$$|m_{\tilde{\Omega}}\rangle = (U_{\tilde{\Omega}})^{\otimes n} (\overline{U_{\tilde{\Omega}}})^{\otimes m} |m_{\tilde{z}}\rangle, \qquad (31)$$

with  $U_{\tilde{\Omega}}$  being the 2 × 2 matrix that takes a spin 1/2 pointing in the z direction to the  $\tilde{\Omega}$  direction. Similarly, a covariant Kraus operator acting on the above state has the form

$$A_{\tilde{\Omega}} = (U_{\tilde{\Omega}})^{\otimes k} (\overline{U_{\tilde{\Omega}}})^{\otimes l} A_{\tilde{z}} (U_{\tilde{\Omega}}^{\dagger})^{\otimes (n-l)} (\overline{U_{\tilde{\Omega}}}^{\dagger})^{\otimes (m-k)}.$$
(32)

The key to the proofs of theorems 1 and 2 are the mappings Eqs. (26) and (29) between rank-one POVM elements and Kraus operators. It is easy to see by direct substitution that these mappings conserve the covariant character of the measurements. That is, if we take a covariant rank-one POVM element of the form of Eq. (31) and insert it in Eq. (26), we obtain a covariant Kraus operator of the form of Eq. (32). Similarly, if we take a covariant Kraus operator of the form of Eq. (32) and insert it in Eq. (29), we obtain a covariant rank-one POVM element of the form of Eq. (31). Therefore Theorem 3 holds.

#### C. Preselected parallel spins and fixed postselected state

We now discuss two examples involving preselected ensembles of spin-1/2 particles with fixed postselection. In the first example we obtained an analytical result for an arbitrary number N of spins, while for the example of Sec. VII D we had to resort to a symbolic math program and only obtained (numerical) results for  $N \le 6$  spins. We discuss the calculations for the first example in detail and treat the second example more succinctly.

In this subsection we consider the case where the spins are preselected in the state  $|\uparrow_{\Omega}^{\otimes N}\rangle$  and there is a fixed postselected state  $\langle 0|$ . The fidelity can be expressed as

$$F_{||}^{\rm pre} = \frac{1}{4\pi} \int d\Omega \frac{\int d\tilde{\Omega} \langle \uparrow_{\Omega}^{\otimes N} | R_{\tilde{\Omega}} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger} | \uparrow_{\Omega}^{\otimes N} \rangle \cos^2 \Phi/2}{\int d\tilde{\Omega} \langle \uparrow_{\Omega}^{\otimes N} | R_{\tilde{\Omega}} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger} | \uparrow_{\Omega}^{\otimes N} \rangle},$$
(33)

where  $M_{\tilde{z}}$  is the POVM acting on the spins when the guessed direction is +z, normalized according to

$$\int d\tilde{\Omega} R_{\tilde{\Omega}} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger} \leqslant \mathbb{1}.$$
(34)

We note that we can rewrite  $|\uparrow_{\Omega}^{\otimes N}\rangle\langle\uparrow_{\Omega}^{\otimes N}| = R_{\Omega}|\uparrow_{z}^{\otimes N}\rangle\langle\uparrow_{z}^{\otimes N}|R_{\Omega}^{\dagger}|$  to obtain

$$F_{||}^{\text{pre}} = \frac{1}{4\pi} \int d\Omega \frac{\int d\tilde{\Omega} \langle \uparrow_{z}^{\otimes N} | R_{\Omega}^{\dagger} R_{\tilde{\Omega}} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger} R_{\Omega} | \uparrow_{z}^{\otimes N} \rangle \cos^{2} \Phi/2}{\int d\tilde{\Omega} \langle \uparrow_{z}^{\otimes N} | R_{\Omega}^{\dagger} R_{\Omega} M_{\tilde{z}} R_{\tilde{\Omega}}^{\dagger} R_{\Omega} | \uparrow_{z}^{\otimes N} \rangle}$$
(35)

Note also that the integrals over  $\Omega$  and  $\overline{\Omega}$  can be replaced by integrals over the whole SU(2) group using the uniform Haar measure (since any rotation can be decomposed into a rotation around *z*, a rotation that brings *z* to  $\Omega$ , and a rotation around  $\Omega$ ) to obtain

$$F_{\parallel}^{\rm pre} = \int dU \frac{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | U^{\dagger} \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} U | \uparrow_z^{\otimes N} \rangle \cos^2 \Phi/2}{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | U^{\dagger} \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} U | \uparrow_z^{\otimes N} \rangle}$$
  
$$= \int dU \frac{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} | \uparrow_z^{\otimes N} \rangle \cos^2 \Phi/2}{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} | \uparrow_z^{\otimes N} \rangle}$$
  
$$= \frac{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} | \uparrow_z^{\otimes N} \rangle \cos^2 \Phi/2}{\int d\tilde{U} \langle \uparrow_z^{\otimes N} | \tilde{U} M_{\tilde{z}} \tilde{U}^{\dagger} | \uparrow_z^{\otimes N} \rangle}, \qquad (36)$$

where in the second line we have absorbed the rotation U into the rotation  $\tilde{U}$ , and where in the last line we recall that  $\Phi$  is the angle between the *z* axis and the direction onto which the *z* axis is rotated by rotation  $\tilde{U}$ . Note how the use of covariant measurements has enabled an important simplification: in going from Eq. (33) to Eq. (36) we have removed one integral. Equation (36) can be reexpressed as

$$F_{||}^{\rm pre} = \frac{{\rm Tr}CM_{\tilde{z}}}{{\rm Tr}DM_{\tilde{z}}},\tag{37}$$

where

$$C = \int d\tilde{U}\tilde{U}^{\dagger} |\uparrow_{z}^{\otimes N}\rangle \langle\uparrow_{z}^{\otimes N} |\tilde{U}\cos^{2}\Phi/2$$
(38)

and

$$D = \int d\tilde{U}\tilde{U}^{\dagger} |\uparrow_{z}^{\otimes N}\rangle \langle\uparrow_{z}^{\otimes N}|\tilde{U}.$$
 (39)

Now recall that, without loss of generality, the POVM elements can be taken to be rank-one  $M_{\tilde{z}} = |m_{\tilde{z}}\rangle \langle m_{\tilde{z}}|$ . Upon varying with respect to the components of  $|m_{\tilde{z}}\rangle$ , one obtains the equations

$$C|m_{\tilde{z}}\rangle = \lambda D|m_{\tilde{z}}\rangle$$

with

$$\lambda = \frac{\mathrm{Tr}CM_{\tilde{z}}}{\mathrm{Tr}DM_{\tilde{z}}}.$$

Hence the maximum fidelity  $F_{||}^{\text{pre}}$  is given by the largest solution  $\lambda$  of det $(C - \lambda D) = 0$  [compare with Eq. (37)].

It remains to compute the matrices *C* and *D*. To this end we note that the vector  $|\uparrow_z^{\otimes N}\rangle$  has total angular momentum  $S = [\frac{N}{2}(\frac{N}{2}+1)]^{1/2}$  and that, under rotation, the total angular momentum does not change. We can thus restrict our analysis to the space of total angular momentum  $S = [\frac{N}{2}(\frac{N}{2}+1)]^{1/2}$ whose dimension is N + 1. A convenient basis of this space are the eigenvectors of  $S_z$  which we denote  $|m\rangle$ ,  $m = -N/2, \ldots, N/2$ . If U is the rotation that takes direction +z to direction  $\theta$ ,  $\varphi$ , then

$$U^{\dagger}|\uparrow_{z}^{\otimes N}\rangle = \sum_{m=-N/2}^{N/2} \cos^{N/2+m}\left(\frac{\theta}{2}\right) \sin^{N/2-m}\left(\frac{\theta}{2}\right) \times e^{-i(N/2-m)\varphi} \sqrt{\binom{N}{N/2-m}} |m\rangle$$
(40)

and  $\cos \frac{\Phi}{2} = \cos \frac{\theta}{2}$ . Inserting these expressions into Eqs. (38) and (39) and integrating over  $\varphi$  and then  $\theta$  with the uniform measure over the sphere yields that the matrices  $C = \frac{m+N/2+1}{(N+1)(N+2)}\delta_{mm'}$  and  $D = \frac{1}{N+1}\delta_{mm'}$  are both diagonal in this basis. The maximum fidelity [i.e., the largest solution of  $\det(C - \lambda D) = 0$ ] is therefore

$$\max F_{||}^{\rm pre} = \frac{N+1}{N+2}.$$
 (41)

Thus, if the direction  $\Omega$  is encoded into *N* parallel spins, then the presence of a fixed postselected state does not help one in estimating the direction  $\Omega$ , at least if we restrict ourselves to covariant measurements.

#### D. Preselected antiparallel spins and fixed postselected state

Let now consider the case where the spins are preselected to be antiparallel; that is, to be in the state  $|\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2}\rangle$  (for *N* even) and there is a fixed postselected state  $\langle 0|$ . In this case, the fidelity reads

$$F_{\text{anti}||}^{\text{pre}} = \frac{1}{4\pi} \int d\Omega \frac{\int dR \langle \uparrow_{\Omega}^{\otimes N/2} \downarrow_{\Omega}^{\otimes N/2} | RM_{\tilde{z}}R^{\dagger} | \uparrow_{\Omega}^{\otimes N/2} \downarrow_{\Omega}^{\otimes N/2} \rangle \cos^{2} \Phi/2}{\int dR \langle \uparrow_{\Omega}^{\otimes N/2} \downarrow_{\Omega}^{\otimes N/2} | RM_{\tilde{z}}R^{\dagger} | \uparrow_{\Omega}^{\otimes N/2} \downarrow_{\Omega}^{\otimes N/2} \rangle}.$$
(42)

Using exactly the same reasoning as above one can bring this to the form

$$F_{\text{anti}||}^{\text{pre}} = \frac{\text{Tr}C'M_{\tilde{z}}}{\text{Tr}D'M_{\tilde{z}}},\tag{43}$$

where

$$C' = \int d\tilde{U}\tilde{U}^{\dagger} |\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2} \rangle \langle\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2} |\tilde{U}\cos^{2}\Phi/2$$

and

$$D' = \int d\tilde{U}\tilde{U}^{\dagger} |\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2}\rangle \langle\uparrow_{\Omega}^{\otimes N/2}\downarrow_{\Omega}^{\otimes N/2}|\tilde{U}.$$

The maximum fidelity is given by the largest solution  $\lambda$  of det $(C' - \lambda D') = 0$ . In this case the computation of the matrices C' and D' is more complicated. Using a symbolic mathematics program, we could compute these matrices for N = 2,4,6, yielding for the optimal fidelities  $F_{\text{anti}||}^{\text{pre}} = 0.7887$  for N = 2, 0.8873 for N = 4, and 0.9306 for N = 6.

Thus we see that in the case of covariant measurements on antiparallel spins, the presence of a fixed postselected ancilla leads to a small improvement in the fidelity (we can go from case 2 above to the optimal fidelities case 3 above). At present we do not understand why sometimes there is an improvement and sometimes not.

#### VIII. CONCLUSION

In summary we have raised the question of state estimation in pre- and postselected ensembles and setup a general formalism for this problem. In the examples we studied we found two main processes that play a role:

(1) The measurer uses the future to dump into it the results he does not want. No attempt at all is made to use information coming from the future.

(2) The measurer tries to use the information from the future and no attempt at all is made to use the future as a dump.

In general, a measurement procedure may combine these two ideas.

Our first general result, Theorem 1, shows that, when the future can be used to dump unwanted results, then information coming from the future and the complex conjugate information coming from the past are equivalent. This was illustrated by the examples involving covariant measurements on spin-1/2 particles discussed in Sec. VII. Our second general result, Theorem 2, shows that, when the future cannot be used to dump unwanted results, then information coming from the future is always at least as informative as the complex conjugate information coming from the past.

Obviously this is only a first study of estimating pre- and postselected ensembles. Our results and examples show that sometimes the presence of a fixed postselection or the presence of information flowing from the future can dramatically improve the precision with which states can be estimated, but that in other cases the improvement is small, or even nonexistent. (For instance, compare the dramatic gain in [28] with the absence of gain in the example of Sec. VIIC for two very related state-estimation problems.) Future investigations will tell us when information coming from the future can be more informative than the complex conjugate information coming from the past, when using the future as a dump (i.e., having a fixed postselected state  $\langle 0|$ ) helps and when it does not, etc.

Finally, let us comment on the conceptual implications of pre- and postselection. The dynamics of physical systems are invariant under time reversal. But the "measurement postulate" of quantum mechanics breaks this invariance. The theory of pre- and postselection is an attempt to correct this and to have a theory of microphysics that is genuinely invariant under time reversal. But as [28] and the present work show, this approach has dramatic consequences. The hierarchy of computational complexity and much of the structure of quantum information break down. For instance, since two states which are arbitrarily close together can be distinguished with certainty, an analog of Holevo's theorem will not hold. Defining a unit of quantum information in the pre- and postselected setting (analog to the usual qubit) is thus bound to be far more complicated and involve significant conceptual steps.

We do not know what the solution to this conundrum is. Is it possible to formulate a genuinely time invariant and satisfactory theory of microphysics? If so, how deep a reformulation of physics will it require?

## ACKNOWLEDGMENTS

We acknowledge financial support by EU projects Qubit Applications (QAP contract 015848) and Quantum Computer Science (QCS contract 255961), and by the InterUniversity Attraction Pole—Belgium Science Policy—project P6/10 Photonics@be.

- [1] Y. Aharonov, P. G. Bergmann, and J. Lebowitz, Phys. Rev. 134, B1410 (1964).
- [2] Y. Aharonov and L. Vaidman, Lect. Notes Phys. 734, 399 (2007).
- [3] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Phys. Rev. A 79, 052110 (2009).
- [4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
- [5] A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory (North-Holland, Amsterdam, 1982).
- [6] S. Massar and S. Popescu, Phys. Rev. Lett. 74, 1259 (1995).
- [7] N. Gisin and S. Popescu, Phys. Rev. Lett. 83, 432 (1999).
- [8] S. Massar, Phys. Rev. A 62, 040101 (2000).
- [9] E. Bagan, M. Baig, A. Brey, R. Munoz-Tapia, and R. Tarrach, Phys. Rev. A 63, 052309 (2001).
- [10] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. A 64, 022305 (2001).
- [11] E. Bagan, M. Baig, and R. Munoz-Tapia, e-print arXiv:quant-ph/0106155.
- [12] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
- [13] J. S. Lundeen and A. M. Steinberg, Phys. Rev. Lett. **102**, 020404 (2009); K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, New J. Phys. **11**, 033011 (2009).
- [14] L. Hardy, Phys. Rev. Lett. 68, 2981 (1992).
- [15] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, Phys. Lett. A **301**, 130 (2002).
- [16] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, Nature (London) 474, 188 (2011).

- [17] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, Science 332, 1170 (2011).
- [18] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, Phys. Rev. Lett. 92, 043601 (2004).
- [19] N. Brunner, V. Scarani, M. Wegmuller, M. Legre, and N. Gisin, Phys. Rev. Lett. 93, 203902 (2004).
- [20] N. Brunner, A. Acin, D. Collins, N. Gisin, and V. Scarani, Phys. Rev. Lett. 91, 180402 (2003).
- [21] H. M. Wiseman, Phys. Rev. A 65, 032111 (2002).
- [22] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. 66, 1107 (1991); A. D. Parks, D. W. Cullin, and D. C. Stoudt, Proc. R. Soc. A 454, 2997 (1998); 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); R. Mir *et al.*, New J. Phys. 9, 287 (2007); M. Goggin *et al.*, Proc. Natl. Acad. Sci. USA 108, 1256 (2011).
- [23] Y. Aharonov and L. Vaidman, Phys. Rev. A 41, 11 (1990).
- [24] O. Hosten and P. Kwiat, Science 319, 787 (2008).
- [25] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. **102**, 173601 (2009); D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A **80**, 041803 (2009).
- [26] O. Zilberberg, A. Romito, and Y. Gefen, Phys. Rev. Lett. 106, 080405 (2011).
- [27] N. Brunner and C. Simon, Phys. Rev. Lett. 105, 010405 (2010).
- [28] S. Aaronson, Proc. R. Soc. A 461, 3473 (2005).
- [29] D. Dieks, Phys. Lett. A 126, 303 (1988).
- [30] I. D. Ivanovic, Phys. Lett. A 123, 257 (1987).
- [31] A. Peres, Phys. Lett. A **128**, 19 (1988).