Dimension-independent weak value estimation via controlled SWAP operations

Giulio Chiribella,[1, 2, 3, *] Kyrylo Simonov,[4, †] and Xuanqiang Zhao[1, ‡]

[1] *QICI Quantum Information and Computation Initiative, Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong*
[2] *Quantum Group, Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom*
[3] *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada*
[4] *Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria*

(˙Dated: November 8, 2023)

Weak values of quantum observables are a powerful tool for investigating a broad spectrum of quantum phenomena. For this reason, several methods to measure them in the laboratory have been proposed. Some of these methods require weak interactions and postselection, while others are deterministic, but require statistics over a number of experiments growing exponentially with the number of measured particles. Here we propose a deterministic dimension-independent scheme for estimating weak values of arbitrary observables. The scheme, based on coherently controlled SWAP operations, does not require prior knowledge of the initial and final states, nor of the measured observables, and therefore can work with uncharacterized preparation and measurement devices. As a byproduct, our scheme provides an alternative expression for two-time states, that is, states describing quantum systems subject to pre- and post-selections. Using this expression, we show that the controlled-SWAP scheme can be used to estimate weak values for a class of two-time states associated to bipartite quantum states with positive partial transpose.

I. INTRODUCTION

In a seminal 1988 paper [1], Aharonov, Albert, and Vaidman introduced the notion of weak values and showed that they can be experimentally accessed by letting the measured system interact weakly with a pointer in the time interval between a pre-selection and a post-selection. Since then, weak values have proven a powerful tool for analyzing a broad spectrum of quantum phenomena [2–5]. On the foundational side, they provide a lens for experimentally investigating quantum paradoxes [6–8] and Leggett-Garg inequalities [9–12], as well as a quantitative indicator of non-classicality of quantum states [5, 13–15]. On the applied side, they provide probabilistic amplification techniques for quantum metrology [16–23] and for the direct measurement of quantum states [24–27].

Several experimental schemes for measuring weak values have been proposed and demonstrated in the laboratory. Some schemes, based on the original definition of weak values, involve weak measurement interactions and post-selection (see [5] for a review). More recently, there has been a growing interest in methods for estimating weak values without post-selection and weak couplings. An ingenious method was developed by Hoffman [28], who showed that weak values can be estimated by performing standard measurements on the outputs of the universal quantum cloning machine. An alternative method was proposed by Wagner *et al* [29], based on the so-called cyclic shift test [30], which allows one to estimate the trace of the product of multiple density matrices.

A limitation of most of the existing deterministic estimation protocols, however, is that their number of samples generally grows linearly in the system dimension d. For example, the cloning method [28] provides an estimate of the weak value multiplied by a terms of order $1/d$. Hence, obtaining a reliable estimate requires a number of repetitions of the experiment growing at least as d. For a quantum system composed of n particles, this scaling results into an exponential increase of the sample complexity with n. A similar issue arises in the cyclic test approach [29], whose sample complexity is also growing with d in the worst case over all possible observables.

In this paper, we provide a deterministic method for estimating weak values with dimension-independent sample complexity. In our method, two identical quantum systems are prepared in two quantum states associated to the pre- and post-selections appearing in the definition of weak value. Then, the two systems undergo a controlled SWAP operation, which exchanges them or leaves them unchanged depending on the quantum state of a control system. After the controlled-SWAP operation, the two systems undergo a possibly noisy measurement, whose measurement outcomes are used to estimate the weak value by classical post-processing.

The controlled-SWAP scheme studied in this paper also provides insights into the theory of two-time states, a generalized notion of states that describe quantum systems subject to both pre- and post-selections [31–33]. We show that the expectation values associated to two-time states can be obtained from the expectation values associated to ordinary density matrices, by applying a linear fractional transformation that involves a SWAP operation and a partial transpose. Using this result, we show that the controlled-SWAP method can be used to estimate the expectation values of all two-time states associated to density matrices with positive partial transpose (PPT) [34, 35]. Finally, we discuss several adaptations of the controlled-SWAP method that allow to access

* giulio@cs.hku.hk
† kyrylo.simonov@univie.ac.at
‡ xqzhao7@connect.hku.hk

the expectation values for two-time states associated to non-PPT states.

The rest of the paper is structured as follows. In Section II, we briefly review the notion of weak value and put forward a dimension-independent estimation scheme based on controlled-SWAP operations. In Section III, we develop several generalizations of the proposed estimation scheme and introduce a new type of quantities dubbed "double weak values". In Section IV, we provide the reader with an overview of the theory of two-time states. In Section V, we discuss the matrix representation of two-time states. In Section VI, we unravel a fundamental connection between two-time states and the developed estimation scheme. In Section VII, we draw the conclusions and outlook.

II. WEAK VALUE ESTIMATION

Weak values were introduced by Aharonov, Albert and Vaidman, as part of a framework describing pre- and post-selected ensembles [1]. Their original definition referred to the scenario where a quantum system is pre-selected and post-selected in two pure states, described by rays in the system's Hilbert space $\mathcal{H}$: given two (generally unnormalized) vectors $|\psi\rangle \in \mathcal{H}$ and $|\psi'\rangle \in \mathcal{H}$ satisfying the condition $\langle\psi|\psi'\rangle \neq 0$, the weak value of an observable A is

$$W(A\,|\,\psi,\psi') := \frac{\langle\psi'|A|\psi\rangle}{\langle\psi'|\psi\rangle}\,. \tag{1}$$

Here the vector $|\psi\rangle$ ($|\psi'\rangle$) represents the initial (final) state of the system. The observable A is typically taken to be a self-adjoint operator, but more generally could be any linear operator acting on $\mathcal{H}$. Hereafter, the algebra of all linear operators on $\mathcal{H}$ will be denoted by $\mathsf{Lin}(\mathcal{H})$.

The notion of weak value was later generalized to mixed states (see e.g. [36]): for a pair of density matrices ρ and ρ' satisfying the condition $\mathrm{Tr}[\rho\rho'] \neq 0$, the weak value is defined as

$$W(A\,|\,\rho,\rho') = \frac{\mathrm{Tr}[\rho' A\rho]}{\mathrm{Tr}[\rho'\rho]}\,. \tag{2}$$

Eq. (1) can be obtained as a special case of Eq. (2) by setting $\rho = |\psi\rangle\langle\psi|$ and $\rho' = |\psi'\rangle\langle\psi'|$.

An important question is how to measure weak values. In their seminal paper [1], Aharonov, Albert, and Vaidman provided an experimental scheme using weak interactions and postselection. A number of other schemes was subsequently devised by other authors [16–18, 24, 37–49]. The most recent scheme was presented in Ref. [29]. In its simplest version, the scheme is defined for rank-one observables, of the form $A = |\alpha\rangle\langle\alpha|$ for some vector $|\alpha\rangle \in \mathcal{H}$. It consists in two subprotocols: a cyclic test [30] for estimating the trace $\mathrm{Tr}[\rho'|\alpha\rangle\langle\alpha|\rho]]$, and a swap test [50, 51] for estimating the trace $\mathrm{Tr}[\rho'\rho]$. The extension to general self-adjoint observables A is done by using the spectral decomposition $A = \sum_i a_i\,|\alpha_i\rangle\langle\alpha_i|$ and computing the weak value of A as a linear combination of the weak values of the rank-one observables $A_i := |\alpha_i\rangle\langle\alpha_i|$. A limitation of this approach, however, is that in general an observable A can have up to d distinct eigenvalues, and therefore the number of experimental settings needed to estimate the weak value of A grows with the system's dimension in the worst case. For a system of n particles, the number of settings for the estimation of a generic weak value grows exponentially with n.

We now provide a way to estimate weak values of arbitrary observables in a dimension-independent way. Our scheme applies also to infinite-dimensional systems, and does not require the use of ideal projective measurements. Consider a general quantum measurement, described by a positive operator-valued measure (POVM) $(P_j)_{j=1}^N$, satisfying the conditions $P_j \geq 0\,\forall j$ and $\sum_j P_j = I$. We say that this POVM allows estimation of the observable A if A is a linear combination of the POVM operators, namely

$$A = \sum_j x_j\,P_j\,, \tag{3}$$

for suitable coefficients $\{x_j\}_{j=1}^N \in \mathbb{R}^N$ (here and in the following we consider the estimation of self-adjoint observables). An example of POVM that satisfies condition (3) is a noisy measurement of A, corresponding to $P_j = (1-p)\,|\alpha_j\rangle\langle\alpha_j| + p\,I/d$ for $j \in \{1,\dots,d\}$ and $p \in [0,1]$. In this case, Eq. (3) is satisfied with $x_j = (a_j - p\,\mathrm{Tr}[A]/d)/(1-p)$.

The operators P_j give rise to a complex measure

$$q(j|\rho,\rho') := \mathrm{Tr}[P_j\,\rho\,\rho']\,, \tag{4}$$

normalized as $\sum_j q(j|\rho,\rho') = \mathrm{Tr}[\rho\,\rho']$. We call this measure the *weak value (WV) measure.*

When condition Eq. (3) is satisfied, the weak value (2) can be written as the ratio of two expectation values with respect to the WV measure: the expectation value $\mathbb{E}_q(X) := \sum_j x_j\,q(j|\rho,\rho')$ of the random variable X and the expectation value $\mathbb{E}_q(Y) := \sum_j q(j|\rho,\rho')$ of the constant random variable Y with values $y_j = 1\,\forall j$; in formula,

$$W(A\,|\,\rho,\rho') = \frac{\mathbb{E}_q(X)}{\mathbb{E}_q(Y)}\,. \tag{5}$$

In general, the WV measure $q(j|\rho,\rho')$ is not a probability measure, and therefore $\mathbb{E}_q(\cdot)$ are not proper expectation values. Nevertheless, we now show that every expectation value with respect to the WV measure can be evaluated by computing expectations of suitable random variables with respect to a proper probability distribution. To this purpose, we use the following protocol.

Protocol 1 Sampling from WV measure

Inputs. Two copies S_1, S_2 of the system, on which the weak value of (3) is intended to be measured, prepared in the states ρ and ρ', respectively, and an auxiliary qubit C prepared in the state $|+\rangle\langle+|$ with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$.

1. Application of the controlled SWAP gate
$$U = I_{S_1} \otimes I_{S_2} \otimes |0\rangle\langle 0|_C + \mathrm{SWAP}_{S_1 S_2} \otimes |1\rangle\langle 1|_C \qquad (6)$$
to all three systems.

2. Measurement of S_1 with the POVM $(P_j)_j$.

3. Measurement of C with the four-outcome POVM $(R_c)_{c=0}^3$ with
$$\begin{aligned} R_0 &= 1/2|+\rangle\langle+|, &(7)\\ R_1 &= 1/2|-\rangle\langle-|, &(8)\\ R_2 &= 1/2|+i\rangle\langle+i|, &(9)\\ R_3 &= 1/2|-i\rangle\langle-i|, &(10)\end{aligned}$$
where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$. Physically, the POVM $(R_c)_c$ can be realized as a random measurement, by measuring either on the basis $\{|+\rangle, |-\rangle\}$ or on the basis $\{|+i\rangle, |-i\rangle\}$.

Output. The protocol produces a pair of outcomes (j, c), distributed with probability

$$\begin{aligned} p(j,c|\rho,\rho') &:= \mathrm{Tr}\left[(P_j \otimes I \otimes R_c)\, U(\rho \otimes \rho' \otimes |+\rangle\langle+|)U^\dagger\right] \\ &= \frac{1}{8}\Big\{ \mathrm{Tr}[P_j\,\rho] + \mathrm{Tr}[P_j\,\rho'] \\ &\quad + 2\theta(1-c)\,(-1)^c\,\mathrm{Re}\left(\mathrm{Tr}[P_j\rho\rho']\right) \\ &\quad - 2\theta(c-2)\,(-1)^c\,\mathrm{Im}\left(\mathrm{Tr}[P_j\rho\rho']\right)\Big\}, \qquad (11)\end{aligned}$$

where $\theta(t)$ is the Heaviside step function, defined as $\theta(t) = 1$ for $t \geq 0$ and $\theta(t) = 0$ for $t < 0$ (see Appendix A for a derivation of Eq. (11)).

The probability distribution (11) generated by the proposed protocol leads to the following:

Lemma 1. *For every random variable $Z : j \mapsto z_j$, the expectation value of Z with respect to the WV measure $q(j|\rho,\rho')$ in Eq. (4) coincides with the expectation value of the random variable $\widetilde{Z} : (j,c) \mapsto \widetilde{z}_{j,c}$ defined by*

$$\widetilde{z}_{j,c} := 2z_j\,(-1)^c\left[\theta(1-c) - i\theta(c-2)\right] \qquad (12)$$

with respect to the probability distribution $p(j,c|\rho,\rho')$.

Proof. See Appendix B. ■

The above lemma, combined with Eq. (5), implies that the weak value $W(A|\rho,\rho')$ can be estimated by computing the empirical average of the random variables X and Y with respect to the experimental frequencies generated by Protocol 1. The sample complexity of the protocol is provided by the following:

Theorem 1. *With K copies of the input state pair ρ, ρ', the weak value $W(A\,|\,\rho,\rho')$ of an observable $A = \sum_j x_j P_j$ with a POVM $(P_j)_j$ and real coefficients $\{x_j\}_j$ can be estimated up to a small additive error ϵ with a probability no less than $1-\delta$ given*

$$\begin{aligned} K &= \frac{8\ln\left(\frac{6}{\delta}\right)}{\epsilon^2}\left(\frac{x_{\max} + |W(A\,|\,\rho,\rho')|}{\mathrm{Tr}[\rho\rho']}\right)^2 \\ &\quad + O\left(\frac{\ln\frac{1}{\delta}}{\epsilon}\,\frac{x_{\max} + |W(A\,|\,\rho,\rho')|}{(\mathrm{Tr}[\rho\rho'])^2}\right) \qquad (13)\end{aligned}$$

where $x_{\max} := \max_j |x_j|$.

Proof. See Appendix C. ■

When the POVM is the spectral decomposition of the observable A, the maximum value of x_j is simply the norm $\|A\| = \sup_{\||\psi\rangle\|=1} \langle\psi|A|\psi\rangle$. Note that the sample complexity depends on the overlap between the states ρ and ρ', and can become large when the states are nearly orthogonal. However, there is no dependence of the sample complexity on the dimension of the system. As a result, our method provides a reduction of the sample complexity with respect to the method proposed in Ref. [29] in all the cases where the observable A has rank larger than one. Another appealing feature from the experimental point of view is that our method only requires a controlled-SWAP operation, which is generally easier to implement than the controlled shift operations in Ref. [29].

Beside the dimension-independent sample complexity, an interesting feature of the above protocol is that it works with uncharacterized preparation devices: to estimate the weak value $W(A|\,\rho,\rho')$, the experimenter does not need to know in advance the initial state ρ and final state ρ'. Similarly, the experimenter does not need to know in advance the POVM $(P_j)_j$ describing the measurement apparatus: the weak value of any observable of the form $A = \sum_j x_j\,P_j$ can be computed by averaging the random variable x_j with respect to the frequency distribution of the measurement outcome. These features make the above protocol suitable for the tomographic characterization of preparation devices (or measurement devices) using weak values.

III. DOUBLE WEAK VALUES

The controlled-SWAP protocol presented in the previous section lends itself to several generalizations, discussed in this section. These generalizations allow to estimate another type of quantities, which we call "double weak values," and facilitate the connection with the theory of two-time states, discussed in Section VI.

A. Two local measurements

A first generalization of Protocol 1 is to measure both systems S_1 and S_2, instead of measuring only S_1. Measuring S_1 and S_2 with two POVMs $(P_j)_j$ and $(Q_k)_k$, respectively, gives rise to the probability distribution

$$p(j,k,c|\rho,\rho') := \mathrm{Tr}\left[(P_j \otimes Q_k \otimes R_c)\, U(\rho \otimes \rho' \otimes |+\rangle\langle+|)U^\dagger\right] . \tag{14}$$

By sampling over this probability distribution, one can estimate the averages of arbitrary random variables with respect to the complex measure

$$q(j,k\,|\,\rho,\rho') := \mathrm{Tr}[P_j \rho Q_k \rho'] , \tag{15}$$

thereby estimating all quantities of the form

$$\mathrm{Tr}[A\rho B\rho'] , \tag{16}$$

for arbitrary observables A and B in the linear span of $\{P_i\}_i$ and $\{Q_j\}_j$, respectively.

These quantities give rise to a generalization of the notion of weak value:

Definition 1. *The* double weak value *of a pair of observables $A \in \mathsf{Lin}(\mathcal{H})$ and $B \in \mathsf{Lin}(\mathcal{H})$ with respect to the initial state ρ and the final state ρ' is the quantity*

$$W_2(A,B\,|\,\rho,\rho') := \frac{\mathrm{Tr}[A\,\rho\,B\,\rho']}{\mathrm{Tr}[\rho\rho']} . \tag{17}$$

The double weak value generalizes the standard weak value (2), whose expression can be retrieved from Eq. (17) by setting $B = I$; in short, one has

$$W(A\,|\,\rho,\rho') = W_2(A,I\,|\,\rho,\rho') \qquad \forall A,\rho,\rho' \in \mathsf{Lin}(\mathcal{H}) . \tag{18}$$

The double weak value has an interesting physical interpretation. For pure states $\rho = |\psi\rangle\langle\psi| := \psi$ and $\rho' = |\psi'\rangle\langle\psi'| := \psi'$, one has

$$W_2(A,B\,|\,\psi,\psi') = W(A|\psi,\psi')\; W(B|\psi',\psi) , \tag{19}$$

meaning that the double weak value is the product between the weak value of A in the forward time direction, with initial state ψ and final state ψ', and the weak value of B in the backward time direction, with initial state ψ' and final state ψ.

For general mixed states $\rho = \sum_m r_m\, |\psi_m\rangle\langle\psi_m|$ and $\rho' = \sum_n r'_n\, |\psi'_n\rangle\langle\psi'_n|$, the double weak value quantifies the correlations between the weak values of A in the forward time direction and the weak values of B in the backward time direction; in formula:

$$W_2(A,B\,|\,\rho,\rho') = \sum_{m,n} p(m,n)\; W(A|\psi_m,\psi'_n)\; W(B|\psi'_n,\psi_m) , \tag{20}$$

where $p(m,n)$ is the probability distribution

$$p(m,n) := \frac{r_m r'_n\, |\langle\psi_m|\psi'_n\rangle|^2}{\sum_{i,j} r_i r'_j\, |\langle\psi_i|\psi'_j\rangle|^2} . \tag{21}$$

Recalling that the weak values were originally defined in terms of weak measurement processes, the double weak value in Eq. (20) can be interpreted as the correlation between the quantities observed in two weak measurement processes where the roles of the pre- and post-selections are exchanged. These correlations would appear in an exotic scenario where two different agents operate in two opposite time directions [52], with one agent preparing inputs in the past and selecting outputs in the future, and the other agent preparing inputs in the future and selecting outputs in the past. In this setting, the double weak value is the correlation between the values observed by the two agents. It is quite remarkable that such exotic correlations can be experimentally observed through the controlled-SWAP protocol.

B. Local measurements and general bipartite states

A further generalization of Protocol 1 is to replace the two uncorrelated states ρ and ρ' with a single bipartite state $\rho_{\mathrm{in,fin}}$. If systems S_1 and S_2 are measured with POVMs $(P_j)_j$ and $(Q_k)_k$, respectively, the controlled-SWAP protocol produces a triple of outcomes (j,k,c) distributed with probability

$$p(j,k,c|\rho,\rho') := \mathrm{Tr}\left[(P_j \otimes Q_k \otimes R_c)\, U(\rho_{\mathrm{in,fin}} \otimes |+\rangle\langle+|)U^\dagger\right] . \tag{22}$$

By sampling over this probability distribution, one can simulate the averages of arbitrary random variables with respect to the complex measure

$$q(j,k\,|\,\rho_{\mathrm{in,fin}}) := \mathrm{Tr}[(P_j \otimes Q_k)\,\rho_{\mathrm{in,fin}}\,\mathrm{SWAP}] , \tag{23}$$

thereby estimating quantities of the form

$$\mathrm{Tr}[(A \otimes B)\,\rho_{\mathrm{in,fin}}\,\mathrm{SWAP}] \tag{24}$$

for arbitrary observables A and B in the linear span of $\{P_i\}_i$ and $\{Q_j\}_j$, respectively.

As a biproduct, one can also estimate the ratios

$$\frac{\mathrm{Tr}[(A \otimes B)\,\rho_{\mathrm{in,fin}}\,\mathrm{SWAP}]}{\mathrm{Tr}[\rho_{\mathrm{in,fin}}\,\mathrm{SWAP}]} . \tag{25}$$

Note that we do not use the term "(double) weak values" for the ratios in Eq. (25). This omission is intentional: as we will see later in the paper, the proper notion of (double) weak value does not, in general, coincide with the ratios in Eq. (25), but rather with a variant of Eq. (25) where the bipartite state $\rho_{\mathrm{in,fin}}$ is subject to a partial transposition.

C. General bipartite states and joint measurements

A third generalization of the controlled-SWAP protocol consists in performing a joint measurement on the two copies of the system. If the two copies are initially in the bipartite state $\rho_{\rm in,fin}$ and are measured with the joint POVM $(\Pi_j)_j$ after the controlled-SWAP operation, the protocol produces a pair of outcomes (j,c) distributed with probability

$$p(j,c|\rho,\rho') := \mathrm{Tr}\left[(\Pi_j \otimes R_c)\, U(\rho_{\rm in,fin} \otimes |+\rangle\langle+|)U^\dagger\right] . \quad (26)$$

By sampling from this probability distribution, an experimenter can estimate the expectation value of arbitrary random variables with respect to the complex measure

$$q(j|\rho) = \mathrm{Tr}[\Pi_j\, \rho_{\rm in,fin}\, \mathrm{SWAP}]\,, \quad (27)$$

and therefore the value of every quantity of the form

$$\mathrm{Tr}[C\, \rho_{\rm in,fin}\, \mathrm{SWAP}]\,, \quad (28)$$

where $C \in \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H})$ is an arbitrary observable in the linear span of the POVM operators $\{\Pi_j\}$. Again, this allows one to estimate the ratios

$$\frac{\mathrm{Tr}[C\, \rho_{\rm in,fin}\, \mathrm{SWAP}]}{\mathrm{Tr}[\rho_{\rm in,fin}\, \mathrm{SWAP}]} . \quad (29)$$

In the following sections, we will compare the experimentally accessible quantities (28) and (29) with the weak values associated to general two-time states.

IV. TWO-TIME STATES

Weak values can be interpreted as expectation values with respect to a generalized type of quantum states, known as *two-time states* [1, 32, 33]. In the following, we provide a brief review of the notion of two-time state.

The prototype of a two-time state, introduced in the seminal work of Aharonov, Bergmann, and Lebowitz [31], is the linear functional $\lambda : \mathsf{Lin}(\mathcal{H}) \to \mathbb{C}$ defined by the relation

$$\lambda(A) := \langle\psi'|A|\psi\rangle\,, \qquad \forall A \in \mathsf{Lin}(A)\,, \quad (30)$$

where $|\psi\rangle \in \mathcal{H}$ and $|\psi'\rangle \in \mathcal{H}$ are two arbitrary vectors. The complex number $\lambda(A)$ is then interpreted as the *expectation value of the observable A on the two-time state λ*. The functional λ plays the role of an unnormalized state vector in textbook quantum mechanics. In the following we will call λ a *two-time vector*.

The notion of two-time vector was later generalized by Aharonov and Vaidman [32], who considered general linear combinations of the form

$$\lambda(A) := \sum_{i=1}^{N} \langle\psi_i'|\, A\, |\psi_i\rangle\,, \quad (31)$$

where N is a positive integer and $\{|\psi_i\rangle\}_{i=1}^N \subset \mathcal{H}$ and $\{|\psi_i'\rangle\}_{i=1}^N \subset \mathcal{H}$ are arbitrary sets of vectors.

Mathematically, the set of two-time vectors (31) is the set of all linear functionals on the observables of the system.

Proposition 2. *The set of all two-time vectors (31) is the vector space consisting of all linear functionals from $\mathsf{Lin}(\mathcal{H})$ to $\mathbb{C}$.*

Proof. For every linear functional $\lambda : \mathsf{Lin}(\mathcal{H}) \to \mathbb{C}$, there exists one and only one matrix $L \in \mathsf{Lin}(\mathcal{H})$ such that $\lambda(A) = \mathrm{Tr}[L\,A]$. By the singular value decomposition [53], the matrix L can be rewritten as $L = \sum_{i=}^{d} |\psi_i\rangle\langle\psi_i'|$, where $\{|\psi_i\rangle\}$ and $\{|\psi_i'\rangle\}$ are two sets of orthogonal states. Hence, the functional λ has the Aharonov-Vaidman form (31). This proves that every linear functional is a valid two-time vector. The converse is trivial, since every two-time vector is, by definition, a linear functional. ■

The general notion of two-time vector (31) has been extended from pure to mixed states in the work by Silva *et al* [33], who developed a framework for describing pre- and post-selected ensembles of quantum states. We now review this framework, using a slightly different notation that facilitates the connection with the notion of double weak value introduced in this paper.

The transition from pure to mixed two-time states is similar to the transition from state vectors to density matrices.

Definition 2. *The* two-time matrix *corresponding to a two-time vector $\lambda : \mathsf{Lin}(\mathcal{H}) \to \mathbb{C}$ is the linear functional $E_\lambda : \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H}) \to \mathbb{C}$ uniquely defined by the relation*

$$E_\lambda(A \otimes B) := \lambda(A)\, \lambda^\dagger(B)\,, \qquad \forall A, B \in \mathsf{Lin}(\mathcal{H})\,, \quad (32)$$

where $\lambda^\dagger$ is the functional defined by the relation $\lambda^\dagger(A) := \overline{\lambda(A^\dagger)}\,, \forall A \in \mathsf{Lin}(\mathcal{H})$.

Definition 3. *A* (generally unnormalized) two-time density matrix *is linear functional $\omega : \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H}) \to \mathbb{C}$ of the form*

$$\omega = \sum_n E_{\lambda_n}\,, \quad (33)$$

where $(\lambda_n)_n$ are arbitrary linear functionals and $(E_{\lambda_n})_n$ are two-time matrices defined as in Eq. (32).

In the following, two-time density matrices will also be called *two-time states*. Mathematically, the set of all unnormalized two-time states is a convex cone, *i.e.* it contains all convex combinations and all positive multiples of its elements. Silva *et al* [33] showed that the (generally unnormalized) two-time states are in one-to-one correspondence with (generally unnormalized) bipartite density matrices. In the next section we will make this correspondence explicit, providing an expression that connects two-time states to the controlled-SWAP protocol introduced earlier in the paper.

V. MATRIX REPRESENTATION OF TWO-TIME STATES

We now provide an explicit matrix representation of two-time states. To this purpose, we recall that every linear functional $\omega : \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H}) \to \mathbb{C}$ is of the form

$$\omega(C) := \mathrm{Tr}[O\, C] \quad (34)$$

for a suitable matrix $O \in \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H})$. Our goal is to characterize the constraints that the matrix O has to satisfy in order for the functional ω to be a valid two-time state. This goal is achieved by the following theorem:

Theorem 3. *The functional ω in Eq. (34) is an unnormalized two-time state if and only if the matrix O is of the form*

$$O = P^{T_2}\,\mathsf{SWAP}\,, \tag{35}$$

where $P \in \mathsf{Lin}(\mathcal{H}\otimes\mathcal{H})$ is a positive semidefinite matrix and P^{T_2} is the partial transpose of P over the second Hilbert space.

Proof. See Appendix D. ■

Theorem 3 gives an explicit matrix representation of the two-time states: in short, a two-time state must be a functional of the form

$$\omega(C) = \mathrm{Tr}\left[C\,P^{T_2}\,\mathsf{SWAP}\right]\,, \tag{36}$$

for some positive matrix P.

Eq. (36) shows in an explicit way the correspondence between the convex cone of two-time states and the convex cone of positive bipartite matrices. We now analyze the correspondence further, by characterizing the structure of the *normalized* two-time states, defined as follows

Definition 4. *A* normalized two-time state *is a two-time state ω such that $\omega(I) = 1$, where I is the identity operator. The set of normalized two-time states will be denoted by $\mathsf{T}_*(\mathcal{H}\otimes\mathcal{H})$.*

We now provide a characterization of the normalized states, showing that they are in one-to-one correspondence with a convex subset of bipartite density matrices.

Theorem 4. *A linear functional $\omega : \mathsf{Lin}(\mathcal{H}\otimes\mathcal{H}) \to \mathbb{C}$ is a normalized two-time state if and only if it is of the form*

$$\omega(C) = \frac{\mathrm{Tr}[C\,\rho^{T_2}\,\mathsf{SWAP}]}{\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}]} := \omega_\rho(C)\,, \tag{37}$$

where $\rho \in \mathsf{Lin}(\mathcal{H}\otimes\mathcal{H})$ is a normalized density matrix such that $\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}] \neq 0$.

Proof. Suppose that ω is a functional of the form (37). Clearly, ω satisfies the normalization condition $\omega(I) = 1$. Moreover, ω is of the form $\omega(C) = \mathrm{Tr}[C\,P^{T_2}\,\mathsf{SWAP}]$ with $P := \rho/\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}]$. Note that P is positive, because ρ is positive and

$$\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}] = \sum_{i,j}\langle i|\langle i|\rho|j\rangle|j\rangle\,, \tag{38}$$

which implies $\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}] > 0$. Hence, ω is of the form (36), and therefore it is a (normalized) two-time state.

Conversely, suppose that ω is a normalized two-time state. Since ω is a two-time state, it must be of the form (36) for some positive matrix P. Since ω is normalized, one has $\mathrm{Tr}[P^{T_2}\,\mathsf{SWAP}] = \omega(I) = 1$, which implies in particular $P \neq 0$, and, since P is positive, $\mathrm{Tr}[P] \neq 0$. We can then define a normalized density matrix $\rho := P/\mathrm{Tr}[P]$ satisfying the condition $\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}] = \mathrm{Tr}[P^{T_2}\,\mathsf{SWAP}]/\mathrm{Tr}[P] = 1/\mathrm{Tr}[P] \neq 0$. Hence, we have

$$\begin{aligned}\omega(C) &= \mathrm{Tr}[C\,P^{T_2}\,\mathsf{SWAP}]\\ &= \mathrm{Tr}[P]\ \mathrm{Tr}[C\,\rho^{T_2}\,\mathsf{SWAP}]\\ &= \mathrm{Tr}[C\,\rho^{T_2}\,\mathsf{SWAP}]/\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}]\,.\end{aligned} \tag{39}$$

■

Note that the set of bipartite density matrices satisfying the condition $\mathrm{Tr}[\rho^{T_2}\,\mathsf{SWAP}] \neq 0$ can be equivalently characterized as the set of bipartite density matrices that have non-zero overlap with the canonical unnormalized maximally entangled state $|\Phi\rangle := \sum_{i=1}^d |i\rangle\otimes|i\rangle$, as one can see from Eq. (38). We denote this set by

$$\mathsf{D}_*(\mathcal{H}\otimes\mathcal{H}) := \left\{\rho \geq 0\,, \mathrm{Tr}[\rho] = 1\,, \langle\Phi|\rho|\Phi\rangle > 0\right\}. \tag{40}$$

Note also that the correspondence between two-time states and bipartite matrices is a homeomorphism (that is, it is invertible and continuous), as it is given by the linear fractional transformation $\rho \mapsto \omega_\rho$ in Eq. (37). Summarizing, we have the following

Corollary 1. *The sets of two-time states and bipartite density matrices with nonzero overlap with the maximally entangled state are homeomorphic; in formula, $\mathsf{T}_*(\mathcal{H}\otimes\mathcal{H}) \simeq \mathsf{D}_*(\mathcal{H}\otimes\mathcal{H})$.*

Note that the mapping $\rho \mapsto \omega_\rho$ in Eq. (37) is non-linear and it does not preserve convex combinations. Nevertheless, it is linear fractional, and therefore it maps convex combinations into convex combinations, although with generally different weights. Explicitly, a convex combination of density matrices, say $\rho = \sum_i p_i\,\rho_i$, is mapped into a convex combination of two-time states $\omega_\rho = \sum_i q_i\,\omega_{\rho_i}$, with $q_i = p_i\,\langle\Phi|\rho_i|\Phi\rangle/(\sum_j p_j\,\langle\Phi|\rho_j|\Phi\rangle)$. This condition implies that the mapping $\rho \mapsto \omega_\rho$ maps pure bipartite states into extreme points of the set of two-time states, and *vice-versa.*

VI. TWO-TIME STATES AND THE CONTROL-SWAP PROTOCOL

Theorem 4 reveals a fundamental connection between two-time states and the control-SWAP protocol presented earlier in the paper. As we saw in Eq. (29), the control-SWAP protocol allows an experimenter to estimate any quantity of the form

$$\widetilde{\omega}_\rho(C) := \frac{\mathrm{Tr}[C\,\rho\,\mathsf{SWAP}]}{\mathrm{Tr}[\rho\,\mathsf{SWAP}]}\,, \tag{41}$$

for every bipartite observable C and every bipartite quantum state ρ.

The difference between the quantities (41) and the expectation values (37) is the presence of the partial transpose on the second Hilbert space. Physically, one can interpret the partial transpose as the signature of the difference between the spatial correlations accessible with the control-SWAP protocol and the time correlations associated to two-time states.

By comparing Eqs. (41) and (37) we can also have a clear view of the strengths and limitations of the control-SWAP protocol. First, if the density matrix ρ is invariant under partial transpose, an experimentalist who has access to two systems in the joint state ρ can directly estimate the expectation values on the two-time state ω_ρ, by using the control-SWAP protocol.

More generally if the density matrix is positive under partial transpose (PPT) [34, 35], the control-SWAP protocol can provide an estimate of the expectation values on the two time-state ω_ρ, if the experimenter is given access to the quantum state ρ^{T_2}. This is the case of the protocols shown earlier in the paper, where we saw how to estimate the weak values (1) and (2), which correspond to the expectation values of the observable $A \otimes I$ on the two-time states $\omega_{\psi\otimes(\psi')^T}$ and $\omega_{\rho\otimes(\rho')^T}$, respectively. Similarly, the double weak value $\mathrm{Tr}[\rho' A \rho B]/\mathrm{Tr}[\rho'\rho]$ is the expectation of the product observable $A \otimes B$ with respect to the two-time state $\omega_{\rho\otimes(\rho')^T}$.

In constrast, the control-SWAP protocol does not provide, in general, an estimate of the expectation values on a two-time state when the density matrix ρ is not PPT. One way to circumvent this problem would be to approximate the partial transpose operation with a physical process, such as the optimal universal transpose map [54, 55]. However, the estimation protocol resulting from this approach would be dimension-dependent: indeed, optimal universal transpose is the completely positive trace-preserving map given by

$$\mathscr{T}(\rho) = \frac{\rho^{\mathrm{T}} + I}{d+1}\,, \tag{42}$$

and the $1/(d{+}1)$ factor implies that the sample complexity for the estimation of the expectation values $\omega_\rho(C)$ grows linearly with d.

Another approach is to introduce an auxiliary system, and to reduce two-time states associated to non-PPT density matrices to extended two-time states associated to PPT density matrices. This approach can be easily illustrated in the pure state case. In this case, Aharonov and Vaidman [32] showed that every two-time vector $\lambda(A) := \sum_{i=1}^N \langle \psi_i'|\,A\,|\psi_i\rangle$ can be obtained by introducing an auxiliary system of dimension N, and a joint two-time vector $\Lambda(\cdot) := \langle \Psi'|\ \cdot\ |\Psi\rangle$ corresponding to the initial vector $|\Psi\rangle = \sum_i |\psi_i\rangle \otimes |i\rangle \in \mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}$ and final vector $|\Psi'\rangle = \sum_i |\psi_i'\rangle \otimes |i\rangle \in \mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}$, where $\mathcal{H}_{\mathrm{aux}}$ is the Hilbert space of the auxiliary system and $\{|i\rangle\}_{i=1}^N$ is an orthonormal basis for $\mathcal{H}_{\mathrm{aux}}$. With this definition, one has the relation

$$\lambda(A) = \Lambda(A\otimes I_{\mathrm{aux}})\,, \qquad \forall A \in \mathsf{Lin}(\mathcal{H})\,, \tag{43}$$

meaning that the expectation value of the observable A on the two-time vector λ coincides with the expectation value of the observable $A\otimes I_{\mathrm{aux}}$ on the the joint two-time vector Λ. In turn, the two-time vector Λ consists just in a pre-selection and a post-selection to pure states, and therefore it can be reproduced by a product (and therefore PPT) state. Explicitly, we have

$$\frac{\Lambda(A\otimes I_{\mathrm{aux}})}{\Lambda(I\otimes I_{\mathrm{aux}})} = \omega_{\Psi\otimes(\Psi')^T}(A\otimes I_{\mathrm{aux}}\otimes I\otimes I_{\mathrm{aux}})\,, \tag{44}$$

where $\omega_{\Psi\otimes(\Psi')^T}$, defined as in Eq. (37), is a normalized two-time state in $\mathsf{T}_*(\mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}\otimes\mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}})$. Combining Eqs. (43) and (44), we can see that the weak values $\lambda(A)/\lambda(I)$ can be estimated by the control-SWAP protocol, if the experimenter has access to the pure quantum states states proportional to the vectors $|\Psi\rangle$ and $|\overline{\Psi}'\rangle$.

The above argument can be extended from two-time vectors to two-time matrices. For example, the expectation values with respect to the two-time matrix E_λ associated to the two-time vector λ in the previous paragraph can be computed as

$$\frac{E_\lambda(C)}{E_\lambda(I)} = \omega_{\Psi\otimes(\Psi')^T}\Big(W\,(C\otimes I_{\mathrm{aux}}\otimes I_{\mathrm{aux}})\,W^\dagger\Big)\,, \tag{45}$$

where W is the unitary operator that permutes the Hilbert spaces $\mathcal{H}\otimes\mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}\otimes\mathcal{H}_{\mathrm{aux}}$ into $\mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}\otimes\mathcal{H}\otimes\mathcal{H}_{\mathrm{aux}}$. Eq. (45) means that the expectation value on the l.h.s. can be estimated by the controlled-SWAP protocol, if the experimenter has access to the pure quantum states proportional to the vectors $|\Psi\rangle$ and $|\overline{\Psi}'\rangle$.

Overall, the extension approach has the benefit of being dimension-independent, but it requires either the two-time state ω to be known, or the appropriate PPT state of the system and the auxiliary to be provided to the experimenter.

VII. CONCLUSIONS AND OUTLOOK

In this paper we have designed a dimension-independent scheme for estimating weak values of arbitrary observables. The scheme is based on the controlled-SWAP gate, and generates a probability distribution that can be used to sample from a complex measure underlying the weak values of interest. Crucially, the initial and final states of the system as well as the observable itself, whose weak value is estimated, do not have to be *a priori* known to experimenter.

As a byproduct, the structure of the controlled-SWAP scheme provides several insights to theory of two-time states. In particular, we have derived an alternative expression for two-time states, which provides an explicit characterization of the correspondence between two-time states and (a subset of) bipartite density matrices. Using this expression, we characterized the domain of applicability of the control-SWAP protocol, showing that it can be used to estimate the expectation values of all two-time states corresponding to PPT density matrices (including, of course, the product density matrices corresponding to the usual definition of weak values). For two-time states corresponding to non-PPT density matrices, the controlled-SWAP protocol can still be used if the experimenter is given access to an extendend quantum states involving a pair of auxiliary systems.

An interesting open question for future research is whether dimension-independent sampling schemes like ours could be constructed for the estimation of Kirkwood-Dirac (KD) quasiprobability distributions, an important type of complex measures that often arise in quantum information and foundations [14, 56, 57]. KD distributions have a close connection with weak values, and their negativity provides a signature of quantum contextuality in Spekkens' formulation [58]. In turn, KD distributions have found numerous applications beyond quantum foundations, including quantum metrology [59–61], condensed matter physics [56, 62, 63], and thermodynamics [57]. The approach of Wagner *et al* [29] provides a way to estimate the value of the KD distribution at every fixed point. The open question is whether there exist ways to simulate sampling from the KD distribution.

Another interesting direction concerns the theory of two-time states and its relation of the study of causality in quantum theory [64–69]. Recent findings have suggested that anomalous weak values of observables shared between several parties can assist in witnessing the causal relationships between the parties' laboratories [70]. Moreover, two-time states (and multiple-time states in general) themselves can carry information about the underlying causal structure [71] and can be used to witness indefinite causal order of operations [72]. Finally, protocols exploiting controlled causal order of operations in a protocol known as the quantum SWITCH [64, 65] have been shown to be useful for efficient estimation of quantities that can be connected with weak values and KD distributions, such as out-of-time-correlators [73] and incompatibility of quantum observables [74]. Therefore, an interplay between weak values, KD distributions, and indefinite causal order in quantum SWITCH appears as a promising direction for future investigations.

ACKNOWLEDGMENTS

We thank Nicole Yunger Halpern and Billy F. Braasch, Jr. (Quantum Steampunk Laboratory, University of Maryland) for many fruitful discussions and useful comments that helped us improve the manuscript. This work was supported by funding from the Hong Kong Research Grant Council through grant no. 17307520 and through the Senior Research Fellowship Scheme SRFS2021-7S02, and from the John Templeton Foundation through grant 62312, The Quantum Information Structure of Spacetime (qiss.fr). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, "*How the result of a measurement of a component of the spin of a spin-$\frac{1}{2}$ particle can turn out to be* 100," Phys. Rev. Lett. **60**, 1351 (1988).

[2] Y. Aharonov, S. Popescu, and J. Tollaksen, "*A time-symmetric formulation of quantum mechanics*," Phys. Today **63**, 27 (2010).

[3] Y. Shikano, *Theory of "weak value" and quantum mechanical measurements* (InTech Shanghai and Rijeka, 2012).

[4] A. G. Kofman, S. Ashhab, and F. Nori, "*Nonperturbative theory of weak pre-and post-selected measurements*," Phys. Rep. **520**, 43 (2012).

[5] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, "*Colloquium: Understanding quantum weak values: Basics and applications*," Rev. Mod. Phys. **86**, 307–316 (2014).

[6] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, "*Revisiting Hardy's paradox: counterfactual statements, real measurements, entanglement and weak values*," Phys. Lett. A **301**, 130 (2002).

[7] J. S. Lundeen and A. M. Steinberg, "*Experimental joint weak measurement on a photon pair as a probe of Hardy's paradox*," Phys. Rev. Lett. **102**, 020404 (2009).

[8] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, "*Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer*," Science **332**, 1170 (2011).

[9] N. S. Williams and A. N. Jordan, "*Weak Values and the Leggett-Garg Inequality in Solid-State Qubits*," Phys. Rev. Lett. **100**, 026804 (2008).

[10] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, "*Experimental violation of two-party Leggett-Garg inequalities with semiweak measurements*," Phys. Rev. Lett. **106**, 040402 (2011).

[11] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde, "*Violation of the Leggett–Garg inequality with weak measurements of photons*," Proc. Natl. Acad. Sci. **108**, 1256 (2011).

[12] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, "*Partial-measurement backaction and nonclassical weak values in a superconducting circuit*," Phys. Rev. Lett. **111**, 090506 (2013).

[13] M. F. Pusey, "*Anomalous Weak Values Are Proofs of Contextuality*," Phys. Rev. Lett. **113**, 200401 (2014).

[14] J. Dressel, "*Weak values as interference phenomena*," Phys. Rev. A **91**, 032116 (2015).

[15] R. Kunjwal, M. Lostaglio, and M. F. Pusey, "*Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts*," Phys. Rev. A **100**, 042116 (2019).

[16] O. Hosten and P. Kwiat, "*Observation of the Spin Hall Effect of Light via Weak Measurements*," Science **319**, 787 (2008).

[17] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, "*Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification*," Phys. Rev. Lett. **102**, 173601 (2009).

[18] N. Brunner and C. Simon, "*Measuring Small Longitudinal Phase Shifts: Weak Measurements or Standard Interferometry?*" Phys. Rev. Lett. **105**, 010405 (2010).

[19] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, "*Precision frequency measurements with interferometric weak values*," Phys. Rev. A **82**, 063822 (2010).

[20] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, "*Continuous phase amplification with a Sagnac interferometer*," Phys. Rev. A **82**, 011802(R) (2010).

[21] H. F. Hofmann, "*Uncertainty limits for quantum metrology obtained from the statistics of weak measurements*," Phys. Rev. A **83**, 022106 (2011).

[22] H. F. Hofmann, M. E. Goggin, M. P. Almeida, and M. Barbieri, "*Estimation of a quantum interaction parameter using weak measurements: Theory and experiment*," Phys. Rev. A **86**, 040102(R) (2012).

[23] G. Strübi and C. Bruder, "*Measuring ultrasmall time delays of light by joint weak measurements*," Phys. Rev. Lett. **110**, 083605 (2013).

[24] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and

C. Bamber, "*Direct measurement of the quantum wavefunction*," Nature **474**, 188 (2011).

[25] J. S. Lundeen and C. Bamber, "*Procedure for direct measurement of general quantum states using weak measurement*," Phys. Rev. Lett. **108**, 070402 (2012).

[26] J. Z. Salvail, M. Agnew, A. S. Johnson, E. Bolduc, J. Leach, and R. W. Boyd, "*Full characterization of polarization states of light via direct measurement*," Nature Photon. **7**, 316 (2013).

[27] M. Malik, M. Mirhosseini, M. P. J. Lavery, J. Leach, M. J. Padgett, and R. W. Boyd, "*Direct measurement of a 27-dimensional orbital-angular-momentum state vector*," Nat. Commun. **5**, 3115 (2014).

[28] H. F. Hofmann, "*How Weak Values Emerge in Joint Measurements on Cloned Quantum Systems*," Phys. Rev. Lett. **109**, 020408 (2012).

[29] R. Wagner, Z. Schwartzman-Nowik, I. L. Paiva, A. Te'eni, A. Ruiz-Molero, R. Soares Barbosa, E. Cohen, and E. F. Galvão, "*Quantum circuits measuring weak values and Kirkwood-Dirac quasiprobability distributions, with applications*," (2023), arXiv:2302.00705.

[30] M. Oszmaniec, D. J. Brod, and E. F. Galvão, "*Measuring relational information between quantum states, and applications*," (2021), arXiv:2109.10006.

[31] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, "*Time Symmetry in the Quantum Process of Measurement*," Phys. Rev. **134**, B1410 (1964).

[32] Y. Aharonov and L. Vaidman, "*Complete description of a quantum system at a given time*," J. Phys. A: Math. Gen. **24**, 10 (1991).

[33] R. Silva, Y. Guryanova, N. Brunner, N. Linden, A. J. Short, and S. Popescu, "*Pre- and postselected quantum states: Density matrices, tomography, and Kraus operators*," Phys. Rev. A. **89**, 012121 (2014).

[34] A. Peres, "*Separability Criterion for Density Matrices*," Phys. Rev. Lett. **77**, 1413 (1996).

[35] M. Horodecki, P. Horodecki, and R. Horodecki, "*Separability of mixed states: necessary and sufficient conditions*," Phys. Lett. A **223**, 1 (1996).

[36] L. Vaidman, A. Ben-Israel, J. Dziewior, L. Knips, M. Weissl, J. Meinecke, C. Schwemmer, R. Ber, and H. Weinfurter, "*Weak value beyond conditional expectation value of the pointer readings*," Phys. Rev. A **96**, 032114 (2017).

[37] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, "*Realization of a measurement of a "weak value"*," Phys. Rev. Lett. **66**, 1107 (1991).

[38] G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, "*Measurement of quantum weak values of photon polarization*," Phys. Rev. Lett. **94**, 220405 (2005).

[39] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde, "*Violation of the Leggett-Garg inequality with weak measurements of photons*," Proc. Natl. Acad. Sci. USA **108**, 1256 (2011).

[40] C.-F. Li, X.-Y. Xu, J.-S. Tang, J.-S. Xu, and G.-C. Guo, "*Ultrasensitive phase estimation with white light*," Phys. Rev. A **83**, 044102 (2011).

[41] J. Dressel and A. N. Jordan, "*Significance of the imaginary part of the weak value*," Phys. Rev. A **85**, 012107 (2012).

[42] X.-Y. Xu, Y. Kedem, K. Sun, L. Vaidman, C.-F. Li, and G.-C. Guo, "*Phase Estimation with Weak Measurement Using a White Light Source*," Phys. Rev. Lett. **111**, 033604 (2013).

[43] O. S. Magaña-Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, "*Amplification of angular rotations using weak measurements*," Phys. Rev. Lett. **112**, 200401 (2014).

[44] F. Piacentini, A. Avella, M. P. Levi, R. Lussana, F. Villa, A. Tosi, F. Zappa, M. Gramegna, G. Brida, I. P. Degiovanni, and M. Genovese, "*Experiment Investigating the Connection between Weak Values and Contextuality*," Phys. Rev. Lett. **116**, 180401 (2016).

[45] L. Vaidman, "*Weak value controversy*," Phil. Trans. R. Soc. A **375**, 20160395 (2017).

[46] E. Cohen and E. Pollak, "*Determination of weak values of quantum operators using only strong measurements*," Phys. Rev. A **98**, 042112 (2018).

[47] E. Rebufello, F. Piacentini, A. Avella, M. A. de Souza, M. Gramegna, J. Dziewior, E. Cohen, L. Vaidman, I. P. Degiovanni, and M. Genovese, "*Anomalous weak values via a single photon detection*," Light Sci. Appl. **10**, 106 (2021).

[48] P. Adam and L. Diósi, "*Sequential unsharp measurement of photon polarization*," Phys. Rev. A **107**, 063706 (2023).

[49] H. M. Wiseman, A. M. Steinberg, and M. Hallaji, "*Obtaining a single-photon weak value from experiments using a strong (many-photon) coherent state*," AVS Quantum Sci. **5**, 024401 (2023).

[50] A. Barenco, A. Berthiaume, D. Deutsch, A. Ekert, R. Jozsa, and C. Macchiavello, "*Stabilization of Quantum Computations by Symmetrization*," SIAM J. Comput. **26**, 1541 (1997).

[51] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, "*Quantum Fingerprinting*," Phys. Rev. Lett. **87**, 167902 (2001).

[52] G. Chiribella and Z. Liu, "*Quantum operations with indefinite time direction*," Commun. Phys. **5**, 190 (2022).

[53] B. Datta, *Numerical Linear Algebra and Application* (Brooks/Cole Publishing Company, 1995).

[54] P. Horodecki, "*From limits of quantum nonlinear operations to multicopy entanglement witnesses and state spectrum estimation*," (2001), arXiv:quant-ph/0111036.

[55] F. Buscemi, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, "*Optimal realization of the transposition maps*," Phys. Lett. A **314**, 374 (2003).

[56] N. Yunger Halpern, B. Swingle, and J. Dressel, "*Quasiprobability behind the out-of-time-ordered correlator*," Phys. Rev. A **97**, 042105 (2018).

[57] M. Lostaglio, A. Belenchia, A. Levy, S. Hernández-Gómez, N. Fabbri, and S. Gherardini, "*Kirkwood-Dirac quasiprobability approach to quantum fluctuations: Theoretical and experimental perspectives*," Quantum **7**, 1128 (2023).

[58] R. W. Spekkens, "*Contextuality for preparations, transformations, and unsharp measurements*," Phys. Rev. A **71**, 052108 (2005).

[59] D. R. M. Arvidsson-Shukur, N. Yunger Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, "*Quantum advantage in postselected metrology*," Nat. Commun. **11**, 3775 (2020).

[60] J. H. Jenne and D. R. M. Arvidsson-Shukur, "*Quantum Learnability is Arbitrarily Distillable*," (2021), arXiv:2104.09520.

[61] N. Lupu-Gladstein, Y. B. Yilmaz, D. R. M. Arvidsson-Shukur, A. Brodutch, A. O. T. Pang, A. M. Steinberg, and N. Yunger Halpern, "*Negative Quasiprobabilities Enhance Phase Estimation in Quantum-Optics Experiment*," Phys. Rev. Lett. **128**, 220504 (2022).

[62] J. R. González Alonso, N. Yunger Halpern, and J. Dressel, "*Out-of-Time-Ordered-Correlator Quasiprobabilities Robustly Witness Scrambling*," Phys. Rev. Lett. **122**, 040404 (2019).

[63] J. R. González Alonso, N. Shammah, S. Ahmed, F. Nori, and J. Dressel, "*Diagnosing quantum chaos with out-of-time-ordered-correlator quasiprobability in the kicked-top model*," (2022), arXiv:2201.08175.

[64] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "*Beyond quantum computers*," (2009), arXiv:0912.0195.

[65] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "*Quantum computations without definite causal structure*,"

Phys. Rev. A **88**, 022318 (2013).
[66] O. Oreshkov, F. Costa, and Č. Brukner, "*Quantum correlations with no causal order*," Nat. Commun. **3**, 1092 (2012).
[67] Č. Brukner, "*Quantum causality*," Nature Phys. **10**, 259 (2014).
[68] J. Barrett, R. Lorenz, and O. Oreshkov, "*Quantum causal models*," (2019), arXiv:1906.10726.
[69] J. Barrett, R. Lorenz, and O. Oreshkov, "*Cyclic quantum causal models*," Nat. Commun. **12**, 885 (2021).
[70] A. A. Abbott, R. Silva, J. Wechs, N. Brunner, and C. Branciard, "*Anomalous Weak Values Without Post-Selection*," Quantum **3**, 194 (2019).
[71] R. Silva, Y. Guryanova, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu, "*Connecting processes with indefinite causal order and multi-time quantum states*," New J. Phys. **19**, 103022 (2017).
[72] X. Liu, Z. Jia, Y. Qiu, F. Li, and O. Dahlsten, "*Unification of spatiotemporal quantum formalisms: mapping between process and pseudo-density matrices via multiple-time states*," (2023), arXiv:2306.05958.
[73] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, "*Measuring the scrambling of quantum information*," Phys. Rev. A **94**, 040302(R) (2016).
[74] N. Gao, D. Li, A. Mishra, J. Yan, K. Simonov, and G. Chiribella, "*Measuring incompatibility and clustering quantum observables with a quantum switch*," Phys. Rev. Lett. **130**, 170201 (2023).
[75] W. Hoeffding, "*Probability Inequalities for Sums of Bounded Random Variables*," J. Am. Stat. Assoc. **58**, 13–30 (1963).

Appendix A: Derivation of Eq. (11)

At the beginning of the protocol, the two copies of the system (S_1 and S_2) and the control qubit are in the product state

$$\omega = \rho \otimes \rho' \otimes |+\rangle\langle +|. \tag{A1}$$

Application of the controlled SWAP gate U produces the new state

$$\begin{aligned} U\omega U^\dagger \;=\; & \frac{1}{2}\Big\{\rho \otimes \rho' \otimes |0\rangle\langle 0| + \rho' \otimes \rho \otimes |1\rangle\langle 1| \\ & + (\rho \otimes \rho')\mathrm{SWAP} \otimes |0\rangle\langle 1| \\ & + \mathrm{SWAP}(\rho \otimes \rho') \otimes |1\rangle\langle 0|\Big\}. \end{aligned} \tag{A2}$$

A measurement of the first system and auxiliary qubit with the POVM $(P_j)_j$ and $(R_c)_c$, respectively, produces a pair of outcomes (j, c) distributed with probability

$$\begin{aligned} p(j,c|\rho,\rho') \;=\; & \frac{1}{2}\mathrm{Tr}\Big[P_j\rho \otimes \rho' \otimes R_c|0\rangle\langle 0| \\ & + P_j\rho' \otimes \rho \otimes R_c|1\rangle\langle 1| \\ & + (P_j \otimes I)(\rho \otimes \rho')\mathrm{SWAP} \otimes R_c|0\rangle\langle 1|\Big] \\ & + (P_j \otimes I)\mathrm{SWAP}(\rho \otimes \rho') \otimes R_c|1\rangle\langle 0|. \end{aligned}$$

Using the relations

$$\langle 0|R_c|0\rangle = \langle 1|R_c|1\rangle = \frac{1}{4} \qquad \forall c \in \{0,1,2,3\} \tag{A3}$$

and

$$\langle 0|R_c|1\rangle = \frac{(-1)^c}{4}\Big(\theta(1-c) + i\theta(c-2)\Big) \qquad \forall c \in \{0,1,2,3\} \tag{A4}$$

we then obtain

$$\begin{aligned} p(j,c|\rho,\rho') \;=\; & \frac{1}{8}\Big\{\mathrm{Tr}[P_j\rho] + \mathrm{Tr}[P_j\rho'] \\ & + (-1)^c\Big(\theta(1-c) + i\theta(c-2)\Big) \\ & \cdot \mathrm{Tr}\Big[(P_j \otimes I)(\rho \otimes \rho')\mathrm{SWAP}\Big] \\ & + (-1)^c\Big(\theta(1-c) - i\theta(c-2)\Big) \\ & \cdot \mathrm{Tr}\Big[(P_j \otimes I)\mathrm{SWAP}(\rho \otimes \rho')\Big]\Big\}. \end{aligned} \tag{A5}$$

Finally, using the relation $\mathrm{Tr}[(A \otimes B)\,\mathrm{SWAP}] = \mathrm{Tr}[AB]\,, \forall A, B \in \mathsf{Lin}(\mathcal{H})$, we obtain

$$\begin{aligned} p(j,c|\rho,\rho') \;=\; & \frac{1}{8}\Big\{\mathrm{Tr}[P_j\rho] + \mathrm{Tr}[P_j\rho'] \\ & + (-1)^c\theta(1-c)\Big(\mathrm{Tr}[P_j\rho\rho'] + \mathrm{Tr}[\rho P_j\rho']\Big) \\ & + i(-1)^c\theta(c-2)\Big(\mathrm{Tr}[P_j\rho\rho'] - \mathrm{Tr}[\rho P_j\rho']\Big)\Big\} \\ =\; & \frac{1}{8}\Big\{\mathrm{Tr}[P_j\rho] + \mathrm{Tr}[P_j\rho'] \\ & + 2(-1)^c\theta(1-c)\,\mathrm{Re}\,\mathrm{Tr}[P_j\rho\rho'] \\ & - 2(-1)^c\theta(c-2)\,\mathrm{Im}\,\mathrm{Tr}[P_j\rho\rho']\Big\}, \end{aligned} \tag{A6}$$

where the second equality follows from the identity $\mathrm{Tr}[\rho P_j\rho'] = \mathrm{Tr}[\rho'\rho P_j] = \mathrm{Tr}[(P_j\rho\rho')^\dagger] = \overline{\mathrm{Tr}[P_j\rho\rho']}$. This concludes the proof of Eq. (11).

Appendix B: Proof of Lemma 1

The proof of Lemma 1 is a straightforward calculation: we only need to calculate the expectation value of the random variable $\widetilde{Z}$ in Eq. (12) with respect to the probability distribution $p(j,c|\rho,\rho')$. Explicitly, the expectation value is

$$\begin{aligned}
\mathbb{E}_p[\widetilde{Z}] &= \sum_{j,c} \widetilde{z}_{j,c}\, p(j,c|\rho,\rho') \\
&= 2\sum_{j,c} z_j (-1)^c \Big[\theta(1-c) - i\theta(c-2)\Big] p(j,c|\rho,\rho') \\
&= \frac{1}{4}\sum_j z_j \Big(\mathrm{Tr}[P_j\rho] + \mathrm{Tr}[P_j\rho']\Big) \\
&\quad \cdot \sum_c (-1)^c \Big[\theta(1-c) - i\theta(c-2)\Big] \\
&\quad + \frac{1}{2}\sum_c (-1)^{2c}\theta(1-c)\Big[\theta(1-c) - i\theta(c-2)\Big] \\
&\quad \cdot \sum_j z_j \,\mathrm{Re}\,\mathrm{Tr}[P_j\rho\rho'] \\
&\quad - \frac{1}{2}\sum_c (-1)^{2c}\theta(c-2)\Big[\theta(1-c) - i\theta(c-2)\Big] \\
&\quad \cdot \sum_j z_j \,\mathrm{Im}\,\mathrm{Tr}[P_j\rho\rho'] \\
&= \sum_j z_j \Big(\mathrm{Re}\,\mathrm{Tr}[P_j\rho\rho'] + i\,\mathrm{Im}\,\mathrm{Tr}[P_j\rho\rho']\Big) \\
&= \sum_j z_j q(j|\rho,\rho') \\
&= \mathbb{E}_q[Z]\,. \qquad \text{(B1)}
\end{aligned}$$

In summary, the expectation value of the random variable $\widetilde{Z}$ with respect to the probability distribution $p(j,c|\rho,\rho')$ is equal to the expectation value of the random variable Z with respect to the complex distribution $q(j|\rho,\rho')$.

Appendix C: Proof of Theorem 1

The proof is based on two lemmas, provided in the following.

Lemma 2. *Let ρ and ρ' be a pair of states, let $(P_j)_j$ be a POVM, and let $\{x_j\}_j$ be a set of real numbers, with $x_{\max} := \max_j |x_j|$. Let $\nu_{\mathrm{Re}} := \mathrm{Re}(\mathrm{Tr}[\rho' A\rho])$ and $\nu_{\mathrm{Im}} := \mathrm{Im}(\mathrm{Tr}[\rho' A\rho])$ be the real and imaginary parts of the weak value of the observable $A := \sum_j x_j P_j$, respectively. The estimate of ν_{Re} and ν_{Im} obtained from K runs of Protocol 1 has error at most*

$$\epsilon_\nu = 2x_{\max}\sqrt{\frac{1}{K}\ln\frac{2}{\delta}} \qquad \text{(C1)}$$

with probability at least $1-\delta$.

Proof. In accordance with (5), $\mathrm{Tr}[\rho' A\rho] = \mathbb{E}_q[X]$, where q is the WV measure defined in (4), and X is a random variable taking values in $\{x_j\}_j$. Lemma 1 guarantees that the probability distribution (11) generated by Protocol 1 can be equivalently used to estimate $\mathbb{E}_q[X]$. In turn, it can be rewritten as

$$\begin{aligned}
p(j,c|\rho,\rho') &= \frac{1}{2}\Big(\theta(1-c)p_{\mathrm{Re}}(j,c|\rho,\rho') \\
&\quad + \theta(c-2)p_{\mathrm{Im}}(j,c-2|\rho,\rho')\Big), \qquad \text{(C2)}
\end{aligned}$$

with probability distributions

$$\begin{aligned}
p_{\mathrm{Re}}(j,\bar{c}|\rho,\rho') &= \frac{1}{4}\Big\{\mathrm{Tr}[P_j(\rho+\rho')] + 2(-1)^{\bar{c}}\,\mathrm{Re}\,\mathrm{Tr}[\rho' P_j\rho]\Big\} \\
p_{\mathrm{Im}}(j,\bar{c}|\rho,\rho') &= \frac{1}{4}\Big\{\mathrm{Tr}[P_j(\rho+\rho')] - 2(-1)^{\bar{c}}\,\mathrm{Im}\,\mathrm{Tr}[\rho' P_j\rho]\Big\},
\end{aligned}$$

where $\bar{c} \in \{0,1\}$.

Now, recall that the POVM $(R_c)_c$ in Protocol 1 can be obtained by randomly choosing between the projective measurements $(2R_0, 2R_1)$ and $(2R_2, 2R_3)$, which give rise to the probability distributions $p_{\mathrm{Re}}(j,\bar{c}|\rho,\rho')$ and $p_{\mathrm{Im}}(j,\bar{c}|\rho,\rho')$, respectively. Then, a straightforward calculation demonstrates that $\mathbb{E}_{p_{\mathrm{Re}}}[\widetilde{X}] = \nu_{\mathrm{Re}}$ and $\mathbb{E}_{p_{\mathrm{Im}}}[-\widetilde{X}] = \nu_{\mathrm{Im}}$, so that

$$\mathbb{E}_q[X] = \mathbb{E}_{p_{\mathrm{Re}}}[\widetilde{X}] + i\mathbb{E}_{p_{\mathrm{Im}}}[-\widetilde{X}]. \qquad \text{(C3)}$$

In turn, $\widetilde{X}$ is a random variable associated with the corresponding two-outcome projective measurement of auxiliary qubit and taking values in $\{\widetilde{x}_{j,\bar{c}}\}_{j,\bar{c}}$, where $\widetilde{x}_{j,\bar{c}} = (-1)^{\bar{c}}x_j$, and j and $\bar{c}$ are the measurement outcomes. Therefore, for K runs of Protocol 1, we can consider a projective measurement $\{2R_0, 2R_1\}$ of the auxiliary qubit performed in $K/2$ runs and the measurement $\{2R_2, 2R_3\}$ performed in other $K/2$ runs.

In order to estimate ν_{Re}, in the k-th run of the protocol, we associate the measurement $\{2R_0, 2R_1\}$ with a random variable $X_k := \widetilde{X}/(K/2)$ taking values in $\{2(-1)^{\bar{c}}x_j/K\}_{j,\bar{c}}$. Hence, the estimator $X_{\mathrm{Re}} := \sum_{k=1}^{K/2} X_k$ is an unbiased estimator for ν_{Re}. According to the Hoeffding's inequality [75], the probability of estimation with an additive error at least ϵ_ν is upper-bounded as

$$\Pr(|X_{\mathrm{Re}} - \nu_{\mathrm{Re}}| \geq \epsilon_\nu) \leq 2\exp\left(-\frac{2\epsilon_\nu^2}{\sum_k(\max(X_k) - \min(X_k))^2}\right). \qquad \text{(C4)}$$

As each of $K/2$ estimators is bounded as $-2x_{\max}/K \leq X_k \leq 2x_{\max}/K$, where $x_{\max} = \max_j |x_j|$, we obtain

$$\sum_k(\max(X_k) - \min(X_k))^2 = \frac{8x_{\max}^2}{K}. \qquad \text{(C5)}$$

Therefore, the probability (C4) is upper-bounded by

$$\Pr(|X_{\mathrm{Re}} - \nu_{\mathrm{Re}}| \geq \epsilon_\nu) \leq 2\exp\left(-\frac{\epsilon_\nu^2 K}{4x_{\max}^2}\right). \qquad \text{(C6)}$$

Requiring that it does not exceed δ is equivalent to the upper bound

$$2\exp\left(-\epsilon_\nu^2 K/4x_{\max}^2\right) \leq \delta, \qquad \text{(C7)}$$

bounding hence the estimation error as

$$\epsilon_\nu \geq 2x_{\max}\sqrt{\frac{1}{K}\ln\frac{2}{\delta}}. \qquad \text{(C8)}$$

This means that, for $K/2$ runs of Protocol 1, we can make sure that the estimation of ν_{Re} is within an error ϵ_ν with a probability no less than $1-\delta$. This relation between the error and the number of copies also holds for the estimation of the imaginary part ν_{Im}, since, in accordance with (C3), it is associated with the same estimator X_{Re} up to an overall minus sign, which does not change the upper bound (C4). hence the proof. ■

Corollary 2. *Given a pair of states ρ, ρ', for K runs of Protocol 1, the estimate of $\mu := \operatorname{Tr}[\rho'\rho]$ can be guaranteed to have an error at most*

$$\epsilon_\mu = 2\sqrt{\frac{1}{K}\ln\frac{2}{\delta}}, \tag{C9}$$

with probability at least $1-\delta$.

Proof. First, we note that $\mu = \nu_{\text{Re}}$ if $A = I$. Therefore, from the $(K/2)$-round measurement outcomes for estimating ν_{Re}, we can simultaneously construct an unbiased estimator $Y := \sum_{k=1}^{K/2} Y_k$ for μ, where $Y_k := 2(-1)^c/K$. Since μ is a non-negative number, it is not necessary to estimate its imaginary part. Applying the Hoeffding's inequality (C4) and taking into account that $x_{\max} = 1$ for μ, we find that the estimation error is bounded by

$$\epsilon_\mu \le 2\sqrt{\frac{1}{K}\ln\frac{2}{\delta}}. \tag{C10}$$

Therefore, for $K/2$ runs of the Protocol 1, we can make sure that the estimation of μ is within a error ϵ_μ with a probability $1-\delta$. ■

Lemma 3. *Given an observable $A = \sum_j x_j P_j$, where $(P_j)_j$ is a set of POVM elements, and a pair of states ρ, ρ', for K runs of Protocol 1, the estimate of its weak value $W(A\,|\,\rho,\rho')$ can be guaranteed to have an error at most:*

$$\epsilon = \frac{\sqrt{2}(x_{\max} + |W(A\,|\,\rho,\rho')|)}{\mu/\epsilon_\mu - 1}, \tag{C11}$$

where $x_{\max} := \max_j |x_j|$, $\mu = \operatorname{Tr}[\rho'\rho]$, and $\epsilon_\mu = 2\sqrt{\frac{1}{K}\ln\frac{2}{\delta}}$, with probability at least $1-3\delta$.

Proof. First, we we recall that $W(A\,|\,\rho,\rho') = W_{\text{Re}} + iW_{\text{Im}}$, where $W_{\text{Re}} = \nu_{\text{Re}}/\mu$ and $W_{\text{Im}} = \nu_{\text{Im}}/\mu$. Therefore, the corresponding estimation error is given by

$$\epsilon = \sqrt{\epsilon_{W_{\text{Re}}}^2 + \epsilon_{W_{\text{Im}}}^2}, \tag{C12}$$

where $\epsilon_{W_{\text{Re}}}$ and $\epsilon_{W_{\text{Re}}}$ are errors in estimation of the real and imaginary parts of the weak value, respectively. The error in estimation of the real part W_{Re} can be upper-bounded as

$$\begin{aligned}\epsilon_{W_{\text{Re}}} &= \left|\frac{X_{\text{Re}}}{Y} - \frac{\nu_{\text{Re}}}{\mu}\right| \\ &= \left|\frac{X_{\text{Re}}\mu - \nu_{\text{Re}}\mu + \mu\nu_{\text{Re}} - Y\nu_{\text{Re}}}{Y\mu}\right| \\ &\le \frac{|X_{\text{Re}} - \nu_{\text{Re}}|\mu + |\mu - Y||\nu_{\text{Re}}|}{|Y|\mu} \\ &\le \frac{\epsilon_\nu + \epsilon_\mu|W_{\text{Re}}|}{|Y|},\end{aligned} \tag{C13}$$

where the first inequality follows from the triangle inequality $|a+b| \le |a| + |b|$, while the second inequality, in accordance with Lemma 2 and Corollary 2, follows from the upper bounds (C1) and (C9), respectively. As the latter can be given as $\mu - \epsilon_\mu \le Y \le \mu + \epsilon_\mu$, we can assume that the corresponding estimation error is small enough to fulfill the condition $\epsilon_\mu < \mu$, so that $|Y| \ge \mu - \epsilon_\mu \ge 0$ and, thus,

$$\epsilon_{\text{Re}(W)} \le \frac{\epsilon_\nu + \epsilon_\mu|W_{\text{Re}}|}{\mu - \epsilon_\mu}. \tag{C14}$$

Comparing (C1) and (C9), it is easy to note that $\epsilon_\nu = x_{\max}\epsilon_\mu$. Therefore,

$$\epsilon_{W_{\text{Re}}} \le \frac{x_{\max} + |W_{\text{Re}}|}{\mu/\epsilon_\mu - 1}. \tag{C15}$$

Similarly, the error of estimating $\operatorname{Im}(W(A\,|\,\rho,\rho'))$ is bounded from above as

$$\epsilon_{\text{Im}(W)} \le \frac{x_{\max} + |W_{\text{Im}}|}{\mu/\epsilon_\mu - 1}. \tag{C16}$$

Therefore, an upper bound on the total error in estimation of the weak value $W(A|\rho,\rho')$

$$\begin{aligned}\epsilon &\le \frac{\sqrt{(x_{\max} + |W_{\text{Re}}|)^2 + (x_{\max} + |W_{\text{Im}}|)^2}}{\mu/\epsilon_\mu - 1} \\ &= \frac{\sqrt{2x_{\max}^2 + 2x_{\max}\left(|W_{\text{Re}}| + |W_{\text{Im}}|\right) + |W(A|\rho,\rho')|^2}}{\mu/\epsilon_\mu - 1} \\ &\le \frac{\sqrt{2x_{\max}^2 + 4x_{\max}|W(A|\rho,\rho')| + |W(A|\rho,\rho')|^2}}{\mu/\epsilon_\mu - 1} \\ &\le \frac{\sqrt{2}(x_{\max} + |W(A|\rho,\rho')|)}{\mu/\epsilon_\mu - 1},\end{aligned} \tag{C17}$$

is valid if the bounds $|X_{\text{Re}} - \nu_{\text{Re}}| \le \epsilon_\nu$, $|X_{\text{Im}} - \nu_{\text{Im}}| \le \epsilon_\nu$, and $|Y - \mu| \le \epsilon_\mu$ are satisfied. Since each bound is violated with a probability not higher than δ, the violation probability of at least one of them is upper-bounded by 3δ. Hence, the estimation error of $W(A\,|\,\rho,\rho')$ does not exceed the upper bound (C17) with probability at least $1-3\delta$. ■

Proof of Theorem 1. Applying Lemma 3 and taking into account (C9), we can reverse (C11) in order to obtain an expression for the number K of runs of the Protocol 1 in terms of the estimation error ϵ,

$$K = \frac{4\ln\frac{2}{\delta}}{\mu^2}\left(\frac{\sqrt{2}\,(x_{\max} + |W(A\,|\,\rho,\rho')|)}{\epsilon} + 1\right)^2 . \tag{C18}$$

Denoting $\Lambda := \sqrt{2}\,(x_{\max} + |W(A\,|\,\rho,\rho')|)\,/\epsilon$ for the sake of simplicity, we obtain

$$\begin{aligned} K &= \frac{4\ln\frac{2}{\delta}\Lambda^2}{\mu^2}\left(1 + \frac{1}{\Lambda}\right)^2 \\ &= \frac{4\ln\frac{2}{\delta}\Lambda^2}{\mu^2}\left(1 + \frac{2}{\Lambda} + \frac{1}{\Lambda^2}\right) \\ &= \frac{4\ln\frac{2}{\delta}\Lambda^2}{\mu^2} + O\left(\frac{\ln\frac{1}{\delta}\Lambda}{\mu^2}\right), \end{aligned} \tag{C19}$$

where the third equality follows from the assumption that $\Lambda \gg 1$. Finally, relabelling 3δ to δ for the sake of simplicity, we conclude that

$$\begin{aligned} K &= \frac{8\ln\left(\frac{6}{\delta}\right)}{\epsilon^2}\left(\frac{x_{\max} + |W(A\,|\,\rho,\rho')|}{\mathrm{Tr}[\rho\rho']}\right)^2 \\ &\quad + O\left(\frac{\ln\frac{1}{\delta}}{\epsilon}\,\frac{x_{\max} + |W(A\,|\,\rho,\rho')|}{(\mathrm{Tr}[\rho\rho'])^2}\right) \end{aligned} \tag{C20}$$

runs of Protocol 1 are enough to ensure that the estimation of weak value $W(A\,|\,\rho,\rho')$ has an additive error at most ϵ with a probability at least $1-\delta$. Hence the proof.

Appendix D: Proof of Theorem 3

Here we present the proof of Theorem 3. The proof consists of a few lemmas, which are of independent interest as they clarify the mathematical details of the matrix representation of two-time states.

Lemma 4. *Let $\lambda : \mathsf{Lin}(\mathcal{H}) \to \mathbb{C}$ be a two-time vector, written as $\lambda(A) = \mathrm{Tr}[L\,A]$ for a suitable matrix $L \in \mathsf{Lin}(\mathcal{H})$. Then, one has*

$$\lambda^\dagger(A) = \mathrm{Tr}[L^\dagger\,A] \qquad \forall A \in \mathsf{Lin}(\mathcal{H}) \tag{D1}$$

and

$$E_\lambda(C) = \mathrm{Tr}\left[\left(L \otimes L^\dagger\right)\,C\right] \qquad \forall C \in \mathsf{Lin}(\mathcal{H}\otimes\mathcal{H})\,. \tag{D2}$$

Proof. By definition, $\lambda^\dagger(A) = \overline{\lambda(A^\dagger)} = \overline{\mathrm{Tr}[LA^\dagger]} = \mathrm{Tr}[L^\dagger A]$, whence Eq. (D1). Now, one has

$$\begin{aligned} E_\lambda(A\otimes B) &= \mathrm{Tr}\left[L\,A\right]\,\mathrm{Tr}\left[L^\dagger\,B\right] \\ &= \mathrm{Tr}\left[\left(L\otimes L^\dagger\right)\left(A\otimes B\right)\right] \qquad \forall A,B \in \mathsf{Lin}(\mathcal{H})\,. \end{aligned} \tag{D3}$$

Since the product matrices $A\otimes B$ are a spanning set for $\mathsf{Lin}(\mathcal{H}\otimes\mathcal{H})$, Eq. (D2) follows. ■

Lemma 5. *The set of unnormalized two-time density matrices consists of functionals ω of the form $\omega(C) = \mathrm{Tr}[O\,C]$, where*

$$O = \sum_n L_n \otimes L_n^\dagger\,, \tag{D4}$$

and $(L_n)_n \subset \mathsf{Lin}(\mathcal{H})$ are arbitrary matrices.

Proof. Immediate from Eqs. (33) and (D2). ■

Lemma 6. *For every bipartite vector $|\Psi\rangle \in \mathcal{H}\otimes\mathcal{H}$ there exists a matrix $M \in \mathsf{Lin}(\mathcal{H})$ such that*

$$\mathrm{SWAP}\left(|\Psi\rangle\langle\Psi|\right)^{T_2} = M^\dagger \otimes M. \tag{D5}$$

Proof. Let us expand $|\Psi\rangle$ as $|\Psi\rangle = \sum_i |\alpha_i\rangle\otimes|\beta_i\rangle$ for suitable vectors $(|\alpha_i\rangle)_i \subset \mathcal{H}$ and $(|\beta_i\rangle)_i \subset \mathcal{H}$. Then, we have

$$\begin{aligned} \mathrm{SWAP}\;(|\Psi\rangle\langle\Psi|)^{T_2} &= \sum_{i,j} \mathrm{SWAP}\;\left(|\alpha_i\rangle\langle\alpha_j| \otimes |\beta_i\rangle\langle\beta_j|\right)^{T_2} \\ &= \sum_{i,j} \mathrm{SWAP}\,|\alpha_i\rangle\langle\alpha_j| \otimes |\overline{\beta}_j\rangle\langle\overline{\beta}_i| \\ &= \sum_{i,j} |\overline{\beta}_j\rangle\langle\alpha_j| \otimes |\alpha_i\rangle\langle\overline{\beta}_i| \\ &= M^\dagger \otimes M\,, \end{aligned} \tag{D6}$$

with $M := \sum_i |\alpha_i\rangle\langle\overline{\beta}_i|$. ■

Lemma 7. *A matrix $O \in \mathsf{Lin}(\mathcal{H}\otimes\mathcal{H})$ is of the form (D4) if and only if it is of the form (35).*

Proof. Suppose that O is of the form (D4). Let $|\Psi\rangle = \sum_i |\alpha_i\rangle\otimes|\beta_i\rangle$ be an arbitrary vector in $\mathcal{H}\otimes\mathcal{H}$. Then, one has

$$\begin{aligned} \langle\Psi|\;(O\,\mathrm{SWAP})^{T_2}\;|\Psi\rangle &= \mathrm{Tr}\left[(O\,\mathrm{SWAP})^{T_2}\;|\Psi\rangle\langle\Psi|\,\right] \\ &= \mathrm{Tr}\left[O\,\mathrm{SWAP}\;(|\Psi\rangle\langle\Psi|)^{T_2}\,\right] \\ &= \mathrm{Tr}\left[O\,(M^\dagger\otimes M)\right] \\ &= \sum_n \mathrm{Tr}\left[(L_n\otimes L_n^\dagger)\,(M^\dagger\otimes M)\right] \\ &= \sum_n \left|\mathrm{Tr}[L_n^\dagger M]\right|^2 \geq 0\,, \end{aligned} \tag{D7}$$

the third equality following from Lemma 6 and the fourth from Eq. (D4). Since the vector $|\Psi\rangle$ is arbitrary, the matrix $P := (O\,\mathrm{SWAP})^{T_2}$ is positive semidefinite. Moreover, one has $O = P^{T_2}\,\mathrm{SWAP}$. Hence, O is of the form (35).

Conversely, suppose that O has the form (35), namely $O = P^{T_2}\,\mathrm{SWAP}$ for some positive matrix P. Decomposing P as $P = \sum_n |\Psi_n\rangle\langle\Psi_n|$, we obtain

$$\begin{aligned} P^{T_2}\,\mathrm{SWAP} &= \mathrm{SWAP}\,\mathrm{SWAP}\,P^{T_2}\,\mathrm{SWAP} \\ &= \sum_n \mathrm{SWAP}\left(\mathrm{SWAP}\,|\Psi_n\rangle\langle\Psi_n|^{T_2}\right)\,\mathrm{SWAP} \\ &= \sum_n \mathrm{SWAP}\left(M_n^\dagger\otimes M_n\right)\,\mathrm{SWAP} \\ &= \sum_n M_n \otimes M_n^\dagger \end{aligned} \tag{D8}$$

where the third equality follows from Lemma 6. Hence, $P^{T_2}\,\mathrm{SWAP}$ is of the form (D4) with $L_n := M_n$. ■

Proof of Theorem 3. By Lemma 5, the functional ω defined by $\omega(C) = \mathrm{Tr}[O\,C]$ is a two-time state if and only if $O = \sum_n L_n \otimes L_n^\dagger$. Then, Lemma 7 guarantees that there exists a positive semidefinite matrix $P \in \mathsf{Lin}(\mathcal{H} \otimes \mathcal{H})$ such that $O = P^{T_2}\,\mathrm{SWAP}$ if and only if $O = \sum_n L_n \otimes L_n^\dagger$. Hence the proof.