Optimal quantum-programmable projective measurement with linear optics

Ulysse Chabaud,[1,*] Eleni Diamanti,[1] Damian Markham,[1] Elham Kashefi,[1,2] and Antoine Joux[3,†]
[1]*Laboratoire d'Informatique de Paris 6, CNRS, Sorbonne Université, 4 Place Jussieu, 75005 Paris, France*
[2]*School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, Scotland*
[3]*Chaire de Cryptologie de la Fondation SU, Sorbonne Université, Institut de Mathématiques de Jussieu – Paris Rive Gauche, CNRS, INRIA, Université Paris Diderot, Campus Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France*

(Received 17 September 2018; published 14 December 2018)

We present a scheme for a universal device which can be programed by quantum states to approximate a chosen projective measurement to a given precision. Our scheme can be viewed as an extension of the swap test to the instance where one state is supplied many times. As such, it has many potential applications given the variety of quantum information tasks which make use of the swap test. In particular, we show that our scheme is optimal for state discrimination under the one-sided error requirement, and optimally approximates any projective measurement. Furthermore, we propose a practical implementation of our scheme with passive linear optics, which involves a simple interferometer composed only of balanced beam splitters.

DOI: 10.1103/PhysRevA.98.062318

I. INTRODUCTION

In a typical experiment performing a quantum measurement, the choice of measurement is encoded in macroscopic, classical information in the experimental setup. For example, it can be encoded into the reflectivity of a beam splitter, the phase in the branch of an interferometer, or the spacial direction of a Stern-Gerlach device. Often these choices are made beforehand and fixed. In some cases they can be programed in a single setup (for example, using thermo-optic phase shifters [1]). In all these cases, however, the choice of measurement basis is effectively programed classically.

In this work we consider the case where the choice of measurement is instead controlled by a quantum state. There are several reasons why one may consider a quantum state to control the choice of measurement. This state may be an output of a quantum computer or a communication protocol, for example, which is not known beforehand and only accessible as a quantum state. For example, in the cryptographic setting, nonorthogonal states can be used to remotely program a measurement which allows one to test the behavior of a remote party. This is the essence behind the delegated blind verified quantum computation in [2]. At a fundamental level, quantum-programmable measurements separate as much as possible the choice of measurement basis and the bulk of the physical measurement apparatus, which could be interesting in probing foundational questions, for example, in tests of contextuality where information about which measurements are being carried out leads to loopholes [3–5].

A related and, in a sense, more general problem is that of a programmable quantum computer, where a quantum program state is used to encode a unitary to be run on a generic quantum computing device (gate array), first proposed by Nielsen and Chuang [6]. There it was shown that to do so deterministically requires orthogonal program states for every different unitary. To use the continuous parameters available in quantum states to encode more computations, the best one can do is probabilistic. In principle these techniques can be used to program quantum measurements. Indeed, since the original proposal there have been several alternative schemes, extensions, and applications, including programmable quantum state discriminators and measurements [7–11]. These results, however, are either too general to consider the type of efficiency we show here, or specialized to tasks which are different from our simple setting (for example, state discrimination [11]).

We cast our problem as follows, illustrated in Fig. 1. One has $M-1$ program registers, each prepared in the state $|\psi\rangle$ corresponding to the choice of measurement basis, and a single input register prepared in some state $|\phi\rangle$. Our aim is to output a classical bit corresponding to a projective measurement, where 0 represents the outcome $|\psi\rangle$ and 1 represents its complement. In an ideal measurement the result 0 would occur with probability $|\langle\phi|\psi\rangle|^2$. However, this is impossible for finite M. This follows from standard arguments based on the linearity of quantum mechanics, in analogy to the necessity of orthogonal program states for the computation mentioned above. (See, for example, Ref. [6] for the case of programmable universal quantum computation, which easily extends to our case.) We can thus ever only approximate perfect measurements. In our case we parametrize this approximation by ϵ, requiring that the result 0 is returned with probability ϵ close to $|\langle\phi|\psi\rangle|^2$ (see Sec. III for a formal definition).

We present a scheme which achieves this optimally in terms of how ϵ scales with M, under the condition that if the input is $|\psi\rangle$, the measurement always returns 0. This so-called one-sided error requirement [12] makes sense for various potential applications where it is important not to be wrong for this answer. One such example is the link between our scheme and the swap test [13].

*ulysse.chabaud@gmail.com
†antoine.joux@m4x.org

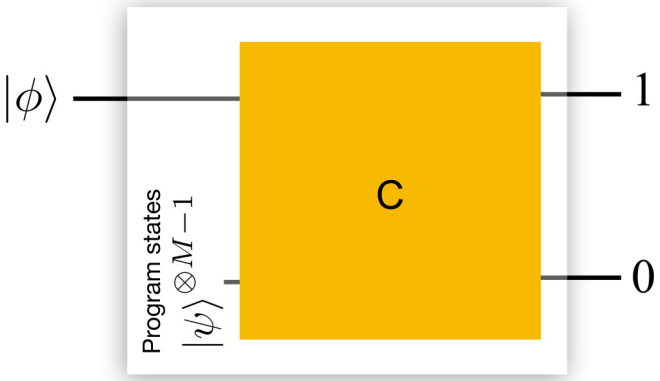

FIG. 1. Programmable projective measurement. Given an input $|\phi\rangle$ and $M-1$ program registers $|\psi\rangle^{\otimes M-1}$, and allowing for possible ancillas (not pictured here), we apply some circuit C, independent of $|\psi\rangle$, and output a binary result where 0 is associated to projecting onto $|\psi\rangle$ and 1 to its complement.

In the swap test, two unknown quantum states are compared using a controlled-swap operation. This test is especially relevant for the task of state discrimination. The general task of assessing if a set of M arbitrary states are identical has been addressed in [14,15]. To solve this in generality requires controlled permutations for all possible permutations and therefore scales exponentially in circuit size. If one restricts oneself to the case where one has $M/2$ copies of one state and $M/2$ copies of the other, one can apply the construction in [15] to get an optimal result. However, this scaling is not much better than simply doing the original swap test $M/2$ times, yet the corresponding test is much more difficult.

From this point of view, the interesting cases of two-states comparison is if one has an asymmetric number of one compared state compared to the other. In the most extreme case one would have just one copy of one state and $M-1$ copies of the other, which is exactly the case we consider for our programmable projective measurement, viewing the program state as the one we have many copies of. In particular, the $M=2$ case reduces to the swap test.

Moreover, the swap test has been shown equivalent to the linear optical Hong-Ou-Mandel effect [16]. Generalizing this equivalence, we present a practical solution to our problem with linear optics using the Hadamard interferometer [17,18].

The next sections are organized as follows. In Sec. II we introduce the circuits for the swap test and its generalization, the swap test of order M. We show in Sec. III that these circuits can be used for programmable projective measurement and prove their optimality. We then present in Sec. IV a simple linear optical interferometer to implement our scheme. For completeness, we introduce in Sec. V a general family of interferometers which reproduce the appropriate statistics. We conclude with an interpretation of our results and discuss various applications in Sec. VI.

II. SWAP CIRCUIT OF ORDER M

The swap test [13] provides an efficient probabilistic tool to compare two unknown quantum states. It takes as input two

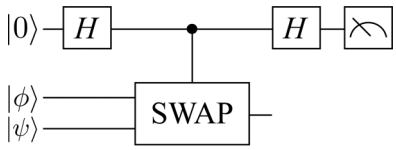

FIG. 2. Circuit representation of a swap test. The ancilla qubit is measured in the computational basis.

quantum states $|\phi\rangle$ and $|\psi\rangle$ that are not entangled and outputs 0 with probability $\frac{1}{2}+\frac{1}{2}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{1}{2}-\frac{1}{2}|\langle\phi|\psi\rangle|^2$, where $\langle\phi|\psi\rangle$ is the overlap between the states $|\phi\rangle$ and $|\psi\rangle$. When the measurement outcome is 0 (resp. 1), we conclude that the states were identical (resp. different), up to a global phase.

A circuit implementing the swap test is represented in Fig. 2, where an ancilla is first prepared in the $|+\rangle$ state by a Hadamard gate,

$$H=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \tag{1}$$

which controls a swap between the two systems being tested.

The swap test meets the so-called one-sided error requirement [12], i.e., if the input states are identical, the test will always declare them as identical. On the other hand, if the input states are different, the test can obtain a wrong conclusion and declare the states identical. The probability that this happens is strictly less than 1; hence by repeating the test various times, the probability that the sequence of tests never answers 1 can be brought down arbitrarily close to zero, exponentially fast. However, the swap test is destructive in the sense that the output states of a previous test cannot be reused for a new test because they become maximally entangled during the test [16]. This means that in order to boost the correctness of the test in this manner, multiple copies of both states must be available.

Let $M\geqslant 2$. We introduce the following generalization of the swap test, in the context where one has access to various copies of a reference state $|\psi\rangle$ but to only a single copy of the other tested state $|\phi\rangle$:

Definition 1. The swap test of order M is a binary test that takes as input a state $|\phi\rangle$ and $M-1$ copies of a state $|\psi\rangle$, and outputs 0 with probability $\frac{1}{M}+\frac{M-1}{M}|\langle\phi|\psi\rangle|^2$ and 1 with probability $(\frac{M-1}{M})(1-|\langle\phi|\psi\rangle|^2)$. If the outcome 0 (resp. 1) is obtained, the test concludes that the states $|\phi\rangle$ and $|\psi\rangle$ were identical (resp. different).

Such a test clearly satisfies the one-sided error requirement. In the following, we are restricted to the swap test of order M when M is a power of 2, writing $n=\ln M$. We introduce the swap circuit of order M (Fig. 3) that acts on M input qubits by applying n consecutive layers of products of swap gates controlled by n ancilla qubits. These ancilla qubits are first initialized in the $|+\rangle$ state using Hadamard gates. Then, they are used as control qubits for the gates $S_0,\ldots,S_{n-1}$, which can be applied in any order, where for all $k\in\{0,\ldots,n-1\}$

$$S_k=\bigotimes_{\substack{i\in[0,2^k-1],\\ j\in[0,2^{n-k-1}-1]}}\mathrm{SWAP}[j2^{k+1}+i,\, j2^{k+1}+i+2^k], \tag{2}$$

with SWAP $[i,j]$ being the unitary operation that swaps the ith and jth qubits for $i,j\in\{0,\ldots,M-1\}$. These controlled

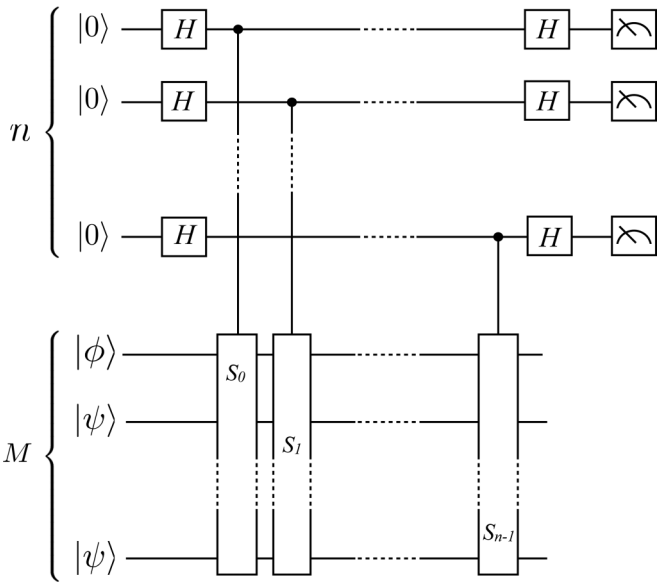

FIG. 3. Swap circuit of order M. The unitaries S_k are tensor products of swap gates described in the main text (2). The $n = \ln M$ ancilla qubits are measured in the computational basis at the end of the computation. The probability of obtaining 0 for all measurement outcomes is $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$.

gates are applied to the input states $|\phi\rangle, |\psi\rangle, \ldots, |\psi\rangle$ (one copy of a state $|\phi\rangle$ and $M-1$ copies of a state $|\psi\rangle$). Finally, a Hadamard gate is applied to each ancilla, which is then measured in the computational basis. By a simple induction, we obtain that the probability of obtaining the outcome 0 for all ancilla qubits is the squared norm of the following state:

$$\frac{1}{M}(|\phi\psi\ldots\psi\rangle + |\psi\phi\ldots\psi\rangle + \cdots + |\psi\ldots\psi\phi\rangle), \tag{3}$$

which depends only on the overlap between the states $|\phi\rangle$ and $|\psi\rangle$. More precisely,

$$\Pr(0,\ldots,0) = \frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2. \tag{4}$$

The swap circuit of order M thus implements the swap test of order M. Indeed, if the outcome $(0,\ldots,0)$ is obtained, the test outputs 0 and we conclude that the states were identical, while for any other outcome the test outputs 1 and we conclude that the states were different. Note that in the case where $M = 2$, the scheme reduces to the original swap test.

Because the $M-1$ last input states are identical, swapping them acts as the identity. This can be used to simplify the swap circuit of order M by replacing the $n = \ln M$ layers of swap gates in Eq. (2) by the following n layers $S'_0, \ldots, S'_{n-1}$, which have to be applied in this order:

$$S'_k = \bigotimes_{l=0}^{2^k-1} \mathrm{SWAP}[l, l+2^k]. \tag{5}$$

This reduces the total number of swap gates from $\frac{M\ln M}{2}$ to $M-1$ without changing the number of ancilla qubits. This

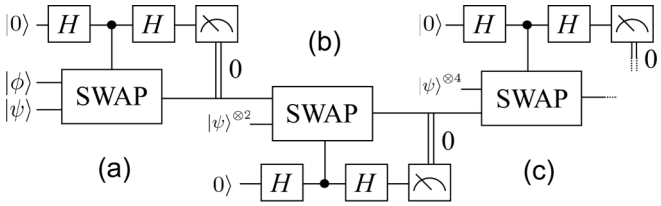

FIG. 4. The simplified swap circuit of order M consisting of $n = \ln M$ consecutive swap tests. (a) The first swap test compares the input states $|\phi\rangle$ and $|\psi\rangle$. (b) If this test is not able to tell apart the input states, i.e., if its outcome is 0, then the second swap test compares the bipartite output state of the first test with the state $|\psi\rangle^{\otimes 2}$. (c) If this test outcome is again 0, then the third swap test compares the quadripartite output state of the second test with the state $|\psi\rangle^{\otimes 4}$, and so on. If the n outcomes are 0, the test concludes that the states $|\phi\rangle$ and $|\psi\rangle$ were identical.

circuit has a simple structure of $n = \ln M$ consecutive swap tests (Fig. 4). For $k \in \{0,\ldots,n-1\}$, conditioned on all the previous outputs being 0, the kth swap test compares the output state of the previous test and the state $|\psi\rangle^{\otimes 2^k}$. Here, the swap test of two multipartite quantum states consists of applying a swap test to each of their corresponding subsystems. However, this multipartite swap test uses only a single ancilla qubit controlling the product of swap gates, as in Eq. (5), instead of an ancilla qubit for each pair of subsystems.

We now prove the optimality of the swap test of order M under the one-sided error requirement, i.e., we show that it achieves the lowest error probability in comparing states $|\phi\rangle$ and $|\psi\rangle$ given $M-1$ copies of $|\psi\rangle$ and one copy of $|\phi\rangle$ such that the one-sided error requirement is satisfied.

For this purpose, we first derive a more general result. In Ref. [15], the authors consider the problem of testing if M quantum states are identical or not (the so-called *identity test*), with the promise that all the states are pairwise identical or orthogonal. In particular, they show that the optimal value for the error probability of any identity test with these assumptions satisfying the one-sided error requirement is $\frac{1}{M}$. We extend this result to the case where the states to be compared are no longer assumed pairwise identical or orthogonal:

Theorem 1. Under the one-sided error requirement, any identity test of M unknown quantum states $|\psi_0\rangle, \ldots, |\psi_{M-1}\rangle$ has an error probability at least

$$\frac{1}{M!}\sum_{\sigma\in\mathcal{S}_M}\prod_{k=0}^{M-1}\langle\psi_k|\psi_{\sigma(k)}\rangle, \tag{6}$$

where $\mathcal{S}_M$ is the symmetric group over $\{0,\ldots,M-1\}$.

Proof. An identity test satisfying the one-sided error requirement can only be wrong when declaring identical states (outputting 0) that were not identical. Hence, to prove Theorem 1, it suffices to lower bound the probability of outputting 0 for any identity test. This is done by showing that the optimal identity test consists of a projection onto the symmetric subspace of the input states Hilbert space. We give a detailed proof in Appendix A. ■

Applying Theorem 1 with $|\psi_0\ldots\psi_{M-1}\rangle = |\phi\psi\ldots\psi\rangle$ implies that the value $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$ is a lower bound for the error probability of any identity test of M states

$|\phi\rangle, |\psi\rangle, \ldots, |\psi\rangle$ (one copy of a state $|\phi\rangle$ and $M-1$ copies of a state $|\psi\rangle$). With Definition 1 we directly obtain the following result:

Corollary 1. The swap test of order M has optimal error probability $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$ under the one-sided error requirement.

The swap circuit of order M is thus optimal for quantum state identity testing with an input $|\phi\rangle, |\psi\rangle, \ldots, |\psi\rangle$, under the one-sided error requirement, since it implements the swap test of order M. In the next section, we show that the swap circuit of order M can be used to implement a programmable projective measurement.

III. CIRCUIT FOR PROGRAMMABLE PROJECTIVE MEASUREMENT

Given that a projective measurement with respect to a state $|\psi\rangle$ is a process that takes as input a state $|\phi\rangle$ and outputs 0 with probability $|\langle\phi|\psi\rangle|^2$ and 1 with probability $1-|\langle\phi|\psi\rangle|^2$, we introduce the natural notion of projective measurement with finite error:

Definition 2. Given a quantum state $|\psi\rangle$ and $\epsilon > 0$, a projective measurement with error ϵ with respect to the reference state $|\psi\rangle$ is a process that takes as input a quantum state $|\phi\rangle$ and outputs 0 with probability $P(0)$ and 1 with probability $P(1)$, such that $|P(0) - (|\langle\phi|\psi\rangle|^2)| \leqslant \epsilon$ and $|P(1) - (1 - |\langle\phi|\psi\rangle|^2)| \leqslant \epsilon$.

Note that the two conditions in the previous definition are equivalent, since $P(0) + P(1) = 1$. It will thus suffice to consider, e.g., the first condition. In this context, under the one-sided error requirement, a projective measurement with any error ϵ always outputs 0 if the input state is equal to the reference state.

Theorem 2. A swap circuit of order M can be used to perform a projective measurement with error $\frac{1}{M}$ under the one-sided error requirement. Moreover, it is optimal in the sense that it uses the minimum number of copies of the reference state for achieving such an error.

Proof. For the swap circuit of order M, we have $\Pr(0, \ldots, 0) = \frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$, so we can consider the whole circuit except the state $|\phi\rangle$ as a black box in Fig. 3, and postprocess the measurement outcomes D as follows: if $D = (0, \ldots, 0)$, output 0, and output 1 otherwise (Fig. 5). The setup now takes a single state $|\phi\rangle$ in input and outputs 0 with probability $P(0) = \frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$, and 1 with probability $P(1) = 1 - P(0)$. We have $|P(0) - (|\langle\phi|\psi\rangle|^2)| \leqslant \frac{1}{M}$ and when $|\phi\rangle = |\psi\rangle$, we have $P(0) = 1 = |\langle\phi|\psi\rangle|^2$, and hence this device performs a projective measurement with error $\frac{1}{M}$ and meets the one-sided error requirement.

We now prove the optimality of this device in terms of resources, i.e., we show that any device implementing a projective measurement with error $\frac{1}{M}$ and meeting the one-sided error requirement cannot use less than $M-1$ copies of the reference state.

We consider a device that implements a projective measurement with error ϵ, with respect to a reference state $|\psi\rangle$, using N copies of this reference state. This device takes as input a quantum state $|\phi\rangle$ and outputs 0 with probability $P_\phi(0)$ and 1 with probability $P_\phi(1) = 1 - P_\phi(0)$. By Definition 2, the probability of outputting 0 satisfies $|P_\phi(0) - (|\langle\phi|\psi\rangle|^2)| \leqslant \epsilon$.

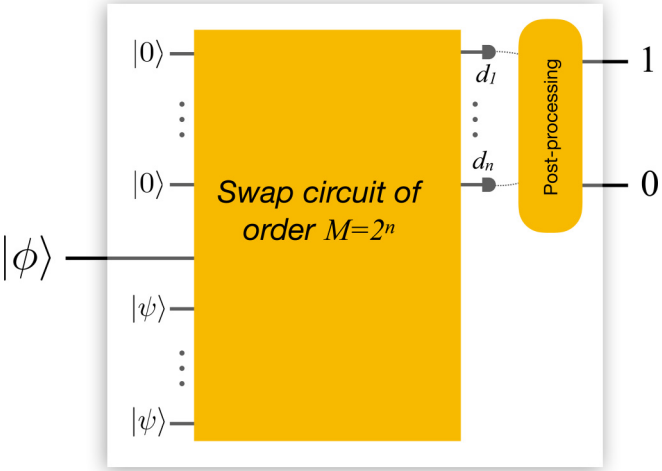

FIG. 5. The swap circuit of order M used as a programmable projective measurement device. It takes as input a state $|\phi\rangle$, and the internal measurement outcomes are postprocessed such that the device outputs 0 with probability $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{M-1}{M}(1 - |\langle\phi|\psi\rangle|^2)$. The programmable resource is the state $|\psi\rangle$, and the process uses $M-1$ copies of this state as well as $n = \ln M$ ancillas.

When the input state $|\phi\rangle$ is orthogonal to the reference state $|\psi\rangle$, the probability $P_{\phi,\perp}(0)$ of outputting 0 thus satisfies

$$P_{\phi,\perp}(0) \leqslant \epsilon. \tag{7}$$

On the other hand, we can use this device to perform an identity test of $N+1$ states $|\phi\rangle, |\psi\rangle, \ldots, |\psi\rangle$ (one copy of the state $|\phi\rangle$ and N copies of the state $|\psi\rangle$): if the output 0 (resp. 1) is obtained, we conclude that the states were identical (resp. different). This device meets the one-sided error requirement, so by Theorem 1 it has error probability at least $\frac{1}{N+1} + \frac{N}{N+1}|\langle\phi|\psi\rangle|^2$. This error probability corresponds to the probability of outputting 0 when the input states are different. In particular, when the input state $|\phi\rangle$ is orthogonal to the reference state $|\psi\rangle$, the probability $P_{\phi,\perp}(0)$ of outputting 0 thus satisfies

$$P_{\phi,\perp}(0) \geqslant \frac{1}{N+1}. \tag{8}$$

Combining both inequalities (7, 8) we obtain $\frac{1}{N+1} \leqslant \epsilon$ or equivalently, $N \geqslant \frac{1}{\epsilon} - 1$. For $\epsilon = \frac{1}{M}$, this amounts to $N \geqslant M-1$, which completes the proof. ■

Theorem 2 implies that given a large enough swap circuit and the ability to produce many copies of a state $|\psi\rangle$, one can projectively measure any state with respect to the state $|\psi\rangle$ up to arbitrary small error. This error scales as the inverse of the number of copies. The circuit can thus be used as a programmable projective measurement device, where the programmable resource is the reference state $|\psi\rangle$ whose number of copies can be adjusted to control the precision of the measurement (Fig. 5).

The implementation of the swap circuit of order M is, however, challenging due to the presence of many controlled-swap gates. In order to lower the implementation requirements, we study in the next section the Hadamard interferometer

and show that its statistics can be efficiently postprocessed to reproduce those of a swap circuit of order M, without the need for ancillas. This comes at the cost that the device no longer has a quantum output, which does not matter for most applications. In particular, we show that the Hadamard interferometer provides a simple linear optical platform for implementing the programmable projective measurement that we have described.

IV. INTERFEROMETER FOR PROGRAMMABLE PROJECTIVE MEASUREMENT

In what follows, we consider optical unitary interferometers of size M which take as input one single photon in a quantum state $|\phi\rangle$ and $M-1$ indistinguishable single photons in a state $|\psi\rangle$, one in each spatial mode. (The spatial modes of the interferometers are indexed from 0 to $M-1$.) These states should be thought of as encoded in additional degrees of freedom of the photons (e.g., polarization, time bins). The output modes are measured using photon number resolving detection.

There exist complex amplitudes α and β and a state $|\psi^\perp\rangle$ with $\langle\psi|\psi^\perp\rangle = 0$ such that

$$|\phi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle, \tag{9}$$

where $\alpha = \langle\psi|\phi\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$. We have the following homomorphism property for single-photon states:

$$|1_\phi\rangle = |1_{\alpha\psi+\beta\psi^\perp}\rangle = \alpha|1_\psi\rangle + \beta|1_{\psi^\perp}\rangle, \tag{10}$$

where for any state $|\chi\rangle$, $|1_\chi\rangle$ is the state of a single photon encoding the state $|\chi\rangle$. It thus suffices to compute the output statistics separately when $|\phi\rangle = |\psi\rangle$ (*indistinguishable case*) and when $|\phi\rangle = |\psi^\perp\rangle$ (*distinguishable case*) to obtain the output statistics in the general case by linearity. The probability of detecting the photon number pattern $D = (d_0, \ldots, d_{M-1})$, or equivalently, that the kth detector detects d_k photons for all $k \in \{0, \ldots, M-1\}$, is then

$$\begin{aligned}\Pr(D) &= |\alpha|^2 \Pr_i(D) + |\beta|^2 \Pr_d(D) \\ &= \Pr_d(D) + |\langle\phi|\psi\rangle|^2 [\Pr_i(D) - \Pr_d(D)],\end{aligned} \tag{11}$$

where $\Pr_i(D)$ is the probability in the indistinguishable case and $\Pr_d(D)$ is the probability in the distinguishable case. The single-photon encoding maps identity of quantum states to distinguishability of single photons. Note that for any measurement outcome $D = (d_0, \ldots, d_{M-1})$, we have $d_0 + \cdots + d_{M-1} = M$, since an interferometer is a passive device that does not change the total number of photons. For any interferometer of size M, we prove in Appendix B the following inequality:

$$\Pr_d(D) \geqslant \frac{\Pr_i(D)}{M}, \tag{12}$$

for any detection pattern D. Combining this inequality with Eq. (11) yields

$$\Pr(D) \geqslant \left(\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2\right)\Pr_i(D). \tag{13}$$

This last expression is valid for any interferometer and can be used to retrieve, in the context of linear optics, the error probability bound for state identity testing under the one-sided error requirement obtained in Corollary 1. Indeed, assume that E is a detection event, which could be a disjoint union of detection events, used for an identity test: If E is obtained we conclude that the states were identical (or equivalently, that the photons were indistinguishable); otherwise we assume that the states were different (or equivalently, that the first photon was distinguishable from the others). The one-sided error requirement can thus be written as $\Pr_i(E) = 1$: indistinguishable photons always pass the test. For different input states $|\phi\rangle$ and $|\psi\rangle$, the error probability of the corresponding test is then given by $\Pr(E)$, which by Eq. (13) is lower bounded by $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$.

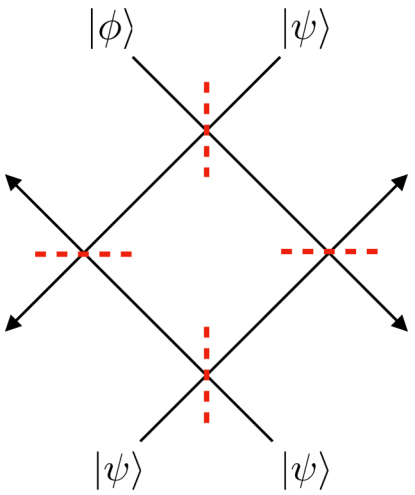

FIG. 6. Hadamard interferometer with four input modes. The dashed red lines represent balanced beam splitters. The input states are one single photon in state $|\phi\rangle$ and three single photons in state $|\psi\rangle$, one in each mode.

We now study a particular unitary interferometer when the size M is a power of 2, namely, the Hadamard interferometer [17,18], and show that it provides a practical and simple implementation of the swap test of order M. For $M = 4$ spatial modes (Fig. 6), this interferometer is described by the Hadamard-Walsh transform of order 2:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} H & H \\ H & -H \end{pmatrix}, \tag{14}$$

where H is a Hadamard matrix, see Eq. (1).

In the general case, the Hadamard interferometer of order M is described by the Hadamard-Walsh transform of order $n = \ln M$, which is defined by induction:

$$H_{k+1} = \frac{1}{\sqrt{2}}\begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix}, \tag{15}$$

with $H_0 = 1$ and $H_1 = H$. We can now state our main result linking the Hadamard interferometer and the swap test of order M.

Theorem 3. The output statistics of the Hadamard interferometer of order M can be classically postprocessed in time $O(M \ln M)$ to reproduce those of the swap test of order M.

Proof. We give hereafter an overview of the proof and refer to Appendix C for further details.

Due to the structure of the Hadamard-Walsh transform, we are able to show that there exists a collection of detection patterns which saturate the bound in Eq. (13) and characterize this collection. We introduce the $M \times M$ matrix,

$$S = (s_{ij})_{0 \leqslant i,j \leqslant M-1} = \sqrt{M} H_n, \tag{16}$$

thus omitting the normalization factor. The matrix S only has $+1$ and -1 entries. We show that its rows, together with the elementwise multiplication, form a group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$. We define for all measurement outcomes $D = (d_0, \ldots, d_{M-1})$ the function

$$\pi(D) = \sum_{i=0}^{M-1} \prod_{j=0}^{M-1} (s_{ij})^{d_j} \tag{17}$$

and exploit the aforementioned group structure to obtain the following equivalences:

$$\begin{aligned} \pi(D) \neq 0 &\Leftrightarrow \pi(D) = M \\ &\Leftrightarrow \Pr_i(D) \neq 0 \\ &\Leftrightarrow \Pr_d(D) = \frac{\Pr_i(D)}{M}. \end{aligned} \tag{18}$$

With the first two lines, the condition $\pi(D) = 0$ is directly equivalent to having a detection event D that can only be witnessed in the distinguishable case. In other words, the detection patterns D such that $\pi(D) = 0$ can only occur if $\langle\phi|\psi\rangle \neq 0$. On the other hand, with the third equivalence, the detection patterns D such that $\pi(D) \neq 0$ are those that saturate the bound obtained in Eq. (13). The Hadamard interferometer can thus be used to compare the states $|\phi\rangle$ and $|\psi\rangle$: If the outcome D obtained satisfies $\pi(D) = M$, we conclude that the states were identical; otherwise $\pi(D) = 0$ and we conclude that the states were different. We show, in particular, that the interferometer described by the unitary matrix H_n satisfies

$$\Pr[\pi(D) = M] = \frac{1}{M} + \frac{M-1}{M} |\langle\phi|\psi\rangle|^2 \tag{19}$$

and

$$\Pr[\pi(D) = 0] = 1 - \Pr[\pi(D) = M], \tag{20}$$

for any detection pattern D. Hence the identity test using the Hadamard interferometer of order M is a swap test of order M. The measurement outcomes D have to be postprocessed by computing $\pi(D)$. Using the group structure of the matrix S, we show that this can be done in time $O(M \ln M)$. ■

Note that the group structure invoked in the proof is preserved under permutations, so Theorem 3 also applies to the unitary interferometers described by permutations of the Hadamard-Walsh transform.

The conclusion to be drawn from Theorem 3 is that as long as a state $|\psi\rangle$ can be encoded using single photons, then one can perform a swap test of order M with respect to the state $|\psi\rangle$ using the Hadamard interferometer of order M and an efficient classical postprocessing of the measurement outcomes. The postprocessing consists of the following parity test: Given the measurement outcome $D = (d_0, \ldots, d_{M-1})$, where $d_0 + \cdots + d_{M-1} = M$, construct the matrix S_D from the matrix $S = \sqrt{M} H_n$ by keeping the kth column only if d_k is odd. If the rows $(1, 2, 4, \ldots, 2^{n-1})$ of S_D all have an even number of -1, output 0; output 1 otherwise. This means that the postprocessing requires only the parity of the photon number in each output mode.

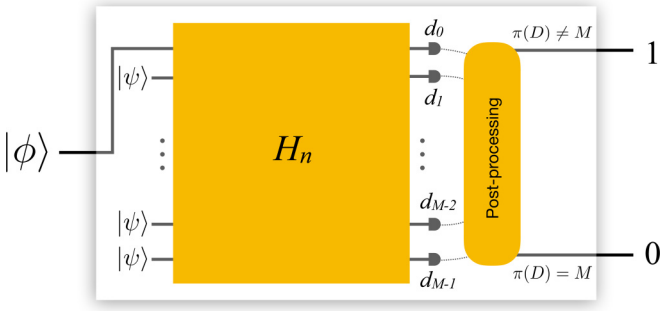

FIG. 7. The Hadamard interferometer of order M used as a programmable projective measurement device. A single photon in the state $|\phi\rangle$ goes through a linear interferometer along with $M-1$ indistinguishable single photons in the state $|\psi\rangle$. The parity of the number of photons in each output mode is measured and efficiently postprocessed, such that the device outputs 0 with probability $\frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2$ and 1 with probability $\frac{M-1}{M}(1 - |\langle\phi|\psi\rangle|^2)$.

In particular, the photon number resolving detectors can be replaced by detecting the parity of the number of photons in each output mode. Detecting this parity can, for example, be achieved with microwave technology [19–21]. Also, only $M-1$ detectors are necessary, since the parity of the number of photons in the remaining mode can be deduced from the parities of the other modes, given that the total number of photons is M.

Using the argument developed in the proof of Theorem 2, by considering the $M-1$ photons and the interferometer as a black box (Fig. 7) whose outcomes are postprocessed as described above, we also deduce the following result from Theorem 3:

Corollary 2. The Hadamard interferometer of order M can be used to perform a projective measurement with error $\frac{1}{M}$ using a classical postprocessing of its measurement outcomes that takes time $O(M \ln M)$.

Interestingly, the unitary interferometers described by the Hadamard-Walsh transform and its permutations are not the only unitary interferometers which can reproduce the statistics of a swap test with efficient postprocessing, and indeed we present a generalization in Sec. V. However, it is the simplicity of the Hadamard interferometer in terms of experimental implementation that motivates our interest towards this interferometer. In particular, this interferometer can be simply implemented with a few balanced beam splitters. A result by Reck *et al.* [22] states that any $M \times M$ unitary interferometer can be implemented using phase shifters and at most $\frac{M(M-1)}{2}$ beam splitters, possibly unbalanced. For the Hadamard interferometer, only $\frac{M \ln M}{2}$ balanced beam splitters are needed and no phase shifters. The proof of this statement is based on a simple induction detailed in Appendix D.

V. GROUP GENERALIZATION FOR ANY VALUE OF THE SIZE PARAMETER M

The Hadamard interferometer requires the size parameter M to be a power of 2. This requirement can be relaxed,

possibly raising the experimental requirements at the same time. Indeed, for any value of M, one can associate to any Abelian group of order M an interferometer of size M which has the desired statistics. This is the object of the following result that uses the invariant factor decomposition of an Abelian group:

Theorem 4. Let G be an Abelian group of order M. Then there exists $N \in \mathbb{N}^*$ and $a_1, \ldots, a_N \in \mathbb{N}^*$, where $a_i | a_{i+1}$ for $i \in \{1, \ldots, N-1\}$ and $a_1 \ldots a_N = M$, such that the interferometer described by the $M \times M$ unitary matrix

$$U_G = \frac{1}{\sqrt{M}} F_{a_1} \otimes \cdots \otimes F_{a_N}, \tag{21}$$

where $F_a = (e^{\frac{2i\pi}{a} kl})_{0 \leqslant k,l \leqslant a-1}$, is the quantum Fourier transform (QFT) of order a for all $a \in \mathbb{N}^*$, can perform a $\frac{1}{M}$-approximate projective measurement with a postprocessing of its measurement outcomes that takes time at most $M \times N$. The rows of $F_G = \sqrt{M} U_G$ together with the elementwise multiplication form a group isomorphic to G.

Proof. We use the notations of the theorem. The invariant factor decomposition of G gives

$$G \simeq (\mathbb{Z}/a_1\mathbb{Z}) \otimes \cdots \otimes (\mathbb{Z}/a_N\mathbb{Z}), \tag{22}$$

where $N \in \mathbb{N}^*$ and $a_1, \ldots, a_N \in \mathbb{N}^*$ are unique, satisfying $a_i | a_{i+1}$ for $i \in \{1, \ldots, N-1\}$ and $a_1 \ldots a_N = M$. Given that the rows of F_a together with the elementwise multiplication form a group isomorphic to $(\mathbb{Z}/a\mathbb{Z})$ for all $a \in \mathbb{N}^*$, the rows of $F_G = (f_{ij})_{0 \leqslant i,j \leqslant M-1} = \sqrt{M} U_G$ together with the elementwise multiplication form a group isomorphic to G.

Since the group structure was the only argument invoked in the proof of Theorem 3, the same conclusion can be drawn here, by following the same argument:

$$\Pr[\pi(D) = M] = \frac{1}{M} + \frac{M-1}{M} |\langle \phi | \psi \rangle|^2, \tag{23}$$

where

$$\pi(D) = \sum_{i=0}^{M-1} \prod_{j=0}^{M-1} (f_{ij})^{d_j}. \tag{24}$$

The group G is finitely generated by N elements, so N rows of F_G are sufficient to generate all its rows by elementwise multiplication. The condition $\pi(D) = M$ can thus be checked in time at most $M \times N$. ∎

In particular, for $G \simeq (\mathbb{Z}/M\mathbb{Z})$, the corresponding interferometer is described by the (normalized) QFT of order M, while for $G \simeq (\mathbb{Z}/2\mathbb{Z})^n$, we retrieve Theorem 3 and the Hadamard interferometer.

VI. CONCLUSION AND DISCUSSION

We have presented an optimal scheme for a programmable projective measurement device and a linear optical implementation, the Hadamard interferometer, which is straightforward and efficient. This could, for example, be used to design a photonic circuit which would act as a universal projective measurement device for a broad range of potential applications, from quantum information and cryptography to tests of contextuality.

The Hadamard interferometer is easily implementable, but this comes at the cost that we are detecting all modes, i.e., that there is no quantum output, unlike for the swap circuit of order M. However, for most applications, it is only the classical output statistics of the circuit that matters, as is the case, e.g., for quantum state identity testing.

Our scheme can also be interpreted as an optimal swap test when one has a single copy of one state and $M-1$ of the other. Given the breadth of applications of the swap test for entanglement testing [23–25], communications [13,26,27], quantum machine learning [28,29], etc., one can anticipate our result will have applications also in these domains.

We have chosen to phrase the problem in terms of $M-1$ copies of the state $|\psi\rangle$. In principle, we could have chosen any other encoding of the quantum input into $M-1$ registers. The reason for our choice is twofold. First, it is part of the envisaged problem setting—we imagine a device producing states encoding our measurement, for example, these could be the output of a computation. Second, we do so in order to separate as much as possible the resource of $M-1$ program systems and the process of translating them into a measurement. In particular, if one had any other encoding, for example, into some entangled states, this encoding process could be incorporated into the circuit representing the generic measurement apparatus. In this sense the most quantum information that can be contained about the state $|\psi\rangle$ in $M-1$ systems is $M-1$ copies of the state $|\psi\rangle$; anything more can be done afterwards. See, for example, [30] for a similar discussion in the case of programmable quantum computation of U(1) rotations.

This result also gives rise to a natural interpretation of the notion of projective measurement in quantum mechanics, as a comparison between one state and several copies of another state using an interferometer. In the macroscopic limit, when many copies of a reference eigenstate are available, we retrieve a macroscopic classically programmable quantum measurement setup.

For completeness, it could be interesting to characterize the full class of interferometers that are optimal for state identity testing under the one-sided error requirement, as we only gave a broad class of such interferometers using a group construction. We conjecture that the Hadamard interferometer will remain the simplest to implement among this class of optimal schemes. It would also be interesting to consider the influence of real experimental conditions, as our scheme assumes that the input states are pure. The one-sided error requirement is also a challenge experimentally, as any interferometer would suffer from the effects of imperfection and noise. We leave these analyses open for future work.

ACKNOWLEDGMENTS

We kindly acknowledge F. Grosshans and A. Olivo for interesting and inspiring discussions. This work has been supported in part by the European Union's H2020 Programme under Grant Agreement No. ERC-669891, by the European Research Council Starting Grant QUSCO, and by the ANR COMB project.

APPENDIX A: PROOF OF OPTIMALITY

An identity test on a Hilbert space $\mathcal{H}$ is a binary test which can be written as a positive-operator-valued measure $\{\Pi_0, \Pi_1\}$, with $\Pi_0 + \Pi_1 = I$. Such a test takes as input a pure tensor product state $|\psi_0 \dots \psi_{M-1}\rangle \in \mathcal{H}^{\otimes M}$ and outputs 0 with probability

$$P(0) = \mathrm{Tr}[\Pi_0 |\psi_0 \dots \psi_{M-1}\rangle\langle\psi_0 \dots \psi_{M-1}|], \tag{A1}$$

and 1 with probability

$$P(1) = 1 - P(0) = \mathrm{Tr}[\Pi_1 |\psi_0 \dots \psi_{M-1}\rangle\langle\psi_0 \dots \psi_{M-1}|]. \tag{A2}$$

If the output 0 is obtained, we conclude that we had $|\psi_0\rangle = \dots = |\psi_{M-1}\rangle$, whereas if the output 1 is obtained, we conclude that the states were not all identical. The one-sided error requirement can thus be written as

$$\forall |\psi\rangle,\ \mathrm{Tr}[\Pi_1 |\psi\rangle\langle\psi|^{\otimes M}] = 0. \tag{A3}$$

Following [31], the symmetric subspace of $\mathcal{H}^{\otimes M}$ can be characterized as

$$S = \mathrm{span}\{|\psi\rangle^{\otimes M} : |\psi\rangle \in \mathcal{H}\}, \tag{A4}$$

and the orthogonal projector onto this space can be written as

$$P_S = \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} P_\sigma, \tag{A5}$$

where for all $\sigma \in \mathcal{S}_M$ and all $|\psi_0 \dots \psi_{M-1}\rangle \in \mathcal{H}^{\otimes M}$ we have $P_\sigma |\psi_0 \dots \psi_{M-1}\rangle = |\psi_{\sigma(0)} \dots \psi_{\sigma(M-1)}\rangle$. Given the characterization of the symmetric subspace, the one-sided error requirement in Eq. (A3) implies that the supports of P_S and Π_1 are disjoint. The support of P_S is thus included in the support of Π_0, given that $\Pi_0 + \Pi_1 = I$, and this implies in turn that $\Pi_0 \geqslant P_S$ by positivity of Π_0.

The error probability of the identity test under the one-sided error requirement is given by the probability of outputting the result 0 while the states were not all identical:

$$\begin{aligned} P(0) &= \mathrm{Tr}[\Pi_0 |\psi_0 \dots \psi_{M-1}\rangle\langle\psi_0 \dots \psi_{M-1}|] \\ &\geqslant \mathrm{Tr}[P_S |\psi_0 \dots \psi_{M-1}\rangle\langle\psi_0 \dots \psi_{M-1}|] \\ &\geqslant \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} \mathrm{Tr}[P_\sigma |\psi_0 \dots \psi_{M-1}\rangle\langle\psi_0 \dots \psi_{M-1}|] \\ &\geqslant \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} \mathrm{Tr}[|\psi_{\sigma(0)} \dots \psi_{\sigma(M-1)}\rangle\langle\psi_0 \dots \psi_{M-1}|] \\ &\geqslant \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} \prod_{k=0}^{M-1} \langle\psi_k|\psi_{\sigma(k)}\rangle, \end{aligned} \tag{A6}$$

where in the third line we used the expression of the orthogonal projector P_S onto the symmetric subspace.

APPENDIX B: STATISTICS OF AN INTERFEROMETER

Recall that we consider optical unitary interferometers of size M which take as input one single photon in a quantum state $|\phi\rangle$ and $M-1$ indistinguishable single photons in a state $|\psi\rangle$, one in each spatial mode, indexed from 0 to $M-1$. The output modes are measured using photon number detection. A measurement outcome thus has the form $D = (d_0, \dots, d_{M-1})$, with $d_0 + \dots + d_{M-1} = M$.

The permanent of an $M \times M$ matrix $T = (t_{ij})_{0 \leqslant i,j \leqslant M-1}$ is defined by

$$\mathrm{Per}(T) = \sum_{\sigma \in \mathcal{S}_M} \prod_{k=0}^{M-1} t_{k\sigma(k)}, \tag{B1}$$

where $\mathcal{S}_M$ is the symmetric group over $\{0, \dots, M-1\}$. We now compute $\mathrm{Pr}_i(D)$ and $\mathrm{Pr}_d(D)$ for all detection patterns D.

In the indistinguishable case, M indistinguishable photons, one in each mode, are sent through a linear optical network described by an $M \times M$ unitary matrix $U = (u_{ij})_{0 \leqslant i,j \leqslant M-1}$. The probability of a detection event D can be computed (see, e.g., [32]) as

$$\mathrm{Pr}_i(D) = \frac{|\mathrm{Per}(U_D)|^2}{D!}, \tag{B2}$$

where $D! = d_0! \dots d_{M-1}!$, and where U_D is the matrix obtained from U by repeating d_k times the kth column for $k \in \{0, \dots, M-1\}$.

In the distinguishable case, $M-1$ indistinguishable photons are sent in modes $1, \dots, M-1$ through a linear optical network described by an $M \times M$ unitary matrix $U = (u_{ij})_{0 \leqslant i,j \leqslant M-1}$, along with one additional photon in the zeroth mode in an orthogonal state. Since it is fully distinguishable from the others, the additional photon behaves independently; hence the probability of detecting the photon number pattern D for one distinguishable photon and $M-1$ indistinguishable photons in input is

$$\mathrm{Pr}_d(D) = \sum_{\substack{k=0 \\ d_k \neq 0}}^{M-1} \mathrm{Pr}_i(D - 1_k)\mathrm{Pr}_i(1_k). \tag{B3}$$

This last expression formalizes the fact that the $M-1$ indistinguishable photons give a detection pattern $D - 1_k$ which, completed by the additional distinguishable photon in the kth output mode, forms the pattern D. Developing this expression with Eq. (B2) yields

$$\mathrm{Pr}_d(D) = \frac{1}{D!} \sum_{\substack{k=0 \\ d_k \neq 0}}^{M-1} d_k |u_{0k} \mathrm{Per}(U_{0,D-1_k})|^2, \tag{B4}$$

where $U_{0,D-1_k}$ is the matrix obtained from U by removing the zeroth row, then by repeating d_l times the lth column for $l \neq k$ and by repeating $d_k - 1$ times the kth column.

In order to obtain more readable expressions, we define for all $k \in \{0, \dots, M-1\}$ and for any detection pattern D,

$$p_k(D) = \begin{cases} \frac{u_{0k}\mathrm{Per}(U_{0,D-1_k})}{\sqrt{D!}} \text{ if } d_k \neq 0, \\ 0 \text{ otherwise.} \end{cases} \tag{B5}$$

Using the Laplace expansion of the permanent, the previous equations (B2) and (B4) are rewritten

$$\mathrm{Pr}_i(D) = \left| \sum_{k=0}^{M-1} d_k p_k(D) \right|^2 \tag{B6}$$

and

$$\mathrm{Pr}_d(D) = \sum_{k=0}^{M-1} d_k |p_k(D)|^2. \tag{B7}$$

Since $\sum_{k=0}^{M-1} d_k = M$, we obtain, using the Cauchy-Schwarz inequality with the complex vectors $\{\sqrt{d_k}\}_{0\leqslant k\leqslant M-1}$ and $\{\sqrt{d_k}p_k(D)\}_{0\leqslant k\leqslant M-1}$,

$$\mathrm{Pr}_d(D) \geqslant \frac{\mathrm{Pr}_i(D)}{M}, \tag{B8}$$

for any detection pattern D.

APPENDIX C: PROOF OF THEOREM 3

Let us define

$$S = (s_{ij})_{0\leqslant i,j\leqslant M-1} = \sqrt{M}H_n, \tag{C1}$$

thus omitting the normalization factor. We have

$$S = \underbrace{\sqrt{2}H \otimes \cdots \otimes \sqrt{2}H}_{n\ times}, \tag{C2}$$

where H is a Hadamard matrix. The rows of $\sqrt{2}H$, together with the elementwise multiplication, form a group isomorphic to $\mathbb{Z}/2\mathbb{Z}$, and thus the rows of S together with the elementwise multiplication form a group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$. As a consequence, multiplying elementwise all the rows of S by its ith row for a given i amounts to permuting the rows of S. Let $D = (d_0, \ldots, d_{M-1})$ and $k \in \{0, \ldots, M-1\}$ such that $d_k \neq 0$. Let also S_{D-1_k} be the matrix obtained from S by repeating d_l times the lth column for $l \neq k$ and $d_k - 1$ the kth column. For all $i \in \{0, \ldots, M-1\}$, one can obtain the matrix $S_{0,D-1_k}$ (with the zeroth row removed) from the matrix $S_{i,D-1_k}$ (with the ith row removed) by multiplying elementwise all rows by the ith row and permuting the rows. Since the permanent is invariant by row permutation, we obtain, for all $i \in \{0, \ldots, M-1\}$ and all $k \in \{0, \ldots, M-1\}$ such that $d_k \neq 0$,

$$\mathrm{Per}(S_{i,D-1_k}) = \epsilon_{ik}(D)\mathrm{Per}(S_{0,D-1_k}), \tag{C3}$$

where $\epsilon_{ik}(D) = s_{ik}\prod_{j=0}^{M-1}(s_{ij})^{d_j}$. Finally, we use the Laplace row expansion formula for the permanent of S_D to obtain, for all $D = (d_0, \ldots, d_{M-1})$ and all $k \in \{0, \ldots, M-1\}$ such that $d_k \neq 0$,

$$\begin{aligned}\mathrm{Per}(S_D) &= \sum_{i=0}^{M-1} s_{ik}\mathrm{Per}(S_{i,D-1_k}) \\ &= \left(\sum_{i=0}^{M-1} s_{ik}\epsilon_{ik}(D)\right)\mathrm{Per}(S_{0,D-1_k}) \\ &= \left(\sum_{i=0}^{M-1}\prod_{j=0}^{M-1}(s_{ij})^{d_j}\right)\mathrm{Per}(S_{0,D-1_k}) \\ &= \pi(D)\mathrm{Per}(S_{0,D-1_k}),\end{aligned} \tag{C4}$$

where we used Eq. (C3) in the second line. With the general expressions of $\mathrm{Pr}_i(D)$ (B2) and $\mathrm{Pr}_d(D)$ (B4), this equation implies

$$M\mathrm{Pr}_i(D) = \pi(D)^2\mathrm{Pr}_d(D). \tag{C5}$$

With the Laplace column expansion formula for the permanent of S_D and the last line of Eq. (C4), we also obtain

$$M^2\mathrm{Pr}_i(D) = \pi(D)^2\mathrm{Pr}_i(D). \tag{C6}$$

In particular, combining Eqs. (C5) and (C6),

$$M^2\pi(D)^2\mathrm{Pr}_d(D) = \pi(D)^4\mathrm{Pr}_d(D). \tag{C7}$$

Now $\mathrm{Pr}_d(D)$ is nonzero for all D, since by Eq. (B4) it is a sum of moduli squared of permanents of $(2^n-1)\times(2^n-1)$ matrices, which in turn cannot vanish by a result of [33]. Hence the previous equation is rewritten

$$M\pi(D) = \pi(D)^2. \tag{C8}$$

As a consequence, $\pi(D) = M$ or $\pi(D) = 0$ for all D. Combining Eqs. (C5) and (C8) we obtain

$$\begin{aligned}\pi(D) \neq 0 &\Leftrightarrow \pi(D) = M \\ &\Leftrightarrow \mathrm{Pr}_i(D) \neq 0 \\ &\Leftrightarrow \mathrm{Pr}_d(D) = \frac{\mathrm{Pr}_i(D)}{M},\end{aligned} \tag{C9}$$

and thus

$$\begin{aligned}\mathrm{Pr}_i[\pi(D) = M] &= \sum_{\pi(D)=M} \mathrm{Pr}_i(D) \\ &= \sum_{\mathrm{Pr}_i(D)\neq 0} \mathrm{Pr}_i(D) \\ &= 1.\end{aligned} \tag{C10}$$

We also obtain

$$\begin{aligned}\mathrm{Pr}_d[\pi(D) = M] &= \sum_{\pi(D)=M} \mathrm{Pr}_d(D) \\ &= \frac{1}{M}\sum_{\pi(D)=M} \mathrm{Pr}_i(D) \\ &= \frac{1}{M}.\end{aligned} \tag{C11}$$

We finally conclude by combining Eqs. (C10) and (C11), and Eq. (11):

$$\begin{aligned}\mathrm{Pr}[\pi(D) = M] &= \sum_{\pi(D)=M} \mathrm{Pr}(D) \\ &= \frac{1}{M} + \frac{M-1}{M}|\langle\phi|\psi\rangle|^2.\end{aligned} \tag{C12}$$

The postprocessing mentioned in the main text, i.e., computing $\pi(D)$, can be done efficiently in time $O(M \ln M)$ for any detection pattern $D = (d_0, \ldots, d_{M-1})$. Indeed, let S_D be the $M \times M$ matrix obtained from S by repeating d_k times the kth column for $k \in \{0, \ldots, M-1\}$. The expression $\pi(D)$ in Eq. (17) is the sum of the product of the elements of each row of S_D. Since the entries of the matrix S are only $+1$ and -1, $\pi(D) = M$ if and only if the number of -1 on the rows of S_D is even for all rows. The condition $\pi(D) = M$ can thus be written as a system of M linear equations modulo 2. Since $(\mathbb{Z}/2\mathbb{Z})^n$ is finitely generated by n elements, the

M rows of S_D can be generated with at most n rows using elementwise multiplication, for any measurement outcome D. Hence, computing the parity of the number of -1 on each row of S_D, which is equivalent to testing $\pi(D) = M$, can be done by computing at most $n = \ln M$ parity equations, with a number of terms in each equation which is at most M.

A simple induction shows that a possible choice for the rows whose parity has to be tested is the rows with index 2^k for $k \in \{0, \ldots, n-1\}$ (the rows of the matrix being indexed from 0 to $M-1$).

APPENDIX D: THE HADAMARD INTERFEROMETER CAN BE IMPLEMENTED WITH A FEW BALANCED BEAM SPLITTERS

Let I_k be the $k \times k$ identity matrix for all k. The size M is a power of 2, with $n = \ln M$. We prove by induction over n that there exist $P_0(n), \ldots, P_{n-1}(n)$ permutation matrices of order $M/2$, such that

$$H_n = \prod_{k=0}^{n-1} P_k(n)(I_{M/2} \otimes H)P_k(n)^T. \tag{D1}$$

Since multiplying matrices is equivalent to setting up experimental devices in sequence, and given that H is the matrix describing a balanced beam splitter, Eq. (D1) implies the result we want to prove.

For $n = 1$, we have $M = 2$, and Eq. (D1) is true with $P_0(1) = I_1$. For brevity, we define for all k,

$$H^{(k)} = I_k \otimes H. \tag{D2}$$

Assuming that Eq. (D1) is true for n, we use the recursive definition of the Hadamard-Walsh transform,

$$H_{n+1} = H \otimes H_n, \tag{D3}$$

along with properties of the tensor product of matrices in order to obtain

$$\begin{aligned} H_{n+1} &= (H_n \otimes I_2)H^{(M)} = Q(I_2 \otimes H_n)Q^T H^{(M)} \\ &= Q\left[I_2 \otimes \prod_{k=0}^{n-1} P_k(n)H^{(M/2)}P_k(n)^T\right]Q^T H^{(M)} \\ &= Q\left[\prod_{k=0}^{n-1}(I_2 \otimes P_k(n))H^{(M)}(I_2 \otimes P_k(n)^T)\right]Q^T H^{(M)} \\ &= \prod_{k=0}^{n-1}[Q(I_2 \otimes P_k(n))]H^{(M)}[Q(I_2 \otimes P_k(n))]^T H^{(M)}, \end{aligned} \tag{D4}$$

where Q is a permutation matrix of order M and where in the third line we have used Eq. (D1). Setting $P_k(n+1) = Q(I_2 \otimes P_k(n))$ for $k \in \{0, \ldots, n-1\}$ and $P_n(n+1) = I_M$ proves Eq. (D1) for $n+1$, since these matrices are permutation matrices of order M. This completes the induction and the proof of the result.

[1] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh *et al.*, Science **349**, 711 (2015).

[2] J. F. Fitzsimons and E. Kashefi, Phys. Rev. A **96**, 012303 (2017).

[3] D. A. Meyer, Phys. Rev. Lett. **83**, 3751 (1999).

[4] R. Clifton and A. Kent, Proc. R. Soc. London, Ser. A, **456**, 2101 (2000).

[5] A. Winter, J. Phys. A: Math. Theor. **47**, 424031 (2014).

[6] M. A. Nielsen and I. L. Chuang, Phys. Rev. Lett. **79**, 321 (1997).

[7] G. Vidal and J. I. Cirac, arXiv:0012067.

[8] M. Dušek and V. Bužek, Phys. Rev. A **66**, 022112 (2002).

[9] M. Roško, V. Bužek, P. R. Chouha, and M. Hillery, Phys. Rev. A **68**, 062302 (2003).

[10] M. Ziman and V. Bužek, Phys. Rev. A **72**, 022343 (2005).

[11] J. A. Bergou, V. Bužek, E. Feldman, U. Herzog, and M. Hillery, Phys. Rev. A **73**, 062334 (2006).

[12] H. Buhrman and L. Fortnow, in *Annual Symposium on Theoretical Aspects of Computer Science* (Springer, New York, 1999), pp. 100–109.

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

[14] A. Chefles, E. Andersson, and I. Jex, J. Phys. A: Math. Gen. **37**, 7315 (2004).

[15] M. Kada, H. Nishimura, and T. Yamakami, J. Phys. A: Math. Theor. **41**, 395309 (2008).

[16] J. C. Garcia-Escartin and P. Chamorro-Posada, Phys. Rev. A **87**, 052330 (2013).

[17] A. Crespi, Phys. Rev. A **91**, 013811 (2015).

[18] A. Crespi, R. Osellame, R. Ramponi, M. Bentivegna, F. Flamini, N. Spagnolo, N. Viggianiello, L. Innocenti, P. Mataloni, and F. Sciarrino, Nat. Commun. **7**, 10469 (2016).

[19] S. Haroche, M. Brune, and J.-M. Raimond, J. Mod. Opt. **54**, 2101 (2007).

[20] B. Vlastakis, G. Kirchmair, Z. Leghtas, S. E. Nigg, L. Frunzio, S. M. Girvin, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Science **342**, 607 (2013).

[21] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff *et al.*, Nature (London) **511**, 444 (2014).

[22] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[23] F. Mintert, M. Kuś, and A. Buchleitner, Phys. Rev. Lett. **95**, 260502 (2005).

[24] S. Walborn, P. S. Ribeiro, L. Davidovich, F. Mintert, and A. Buchleitner, Nature (London) **440**, 1022 (2006).

[25] A. W. Harrow and A. Montanaro, J. Assoc. Comput. Mach. **60**, 3 (2013).

[26] J. N. de Beaudrap, Phys. Rev. A **69**, 022307 (2004).

[27] N. Kumar, E. Diamanti, and I. Kerenidis, Phys. Rev. A **95**, 032337 (2017).
[28] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[29] S. Lloyd, M. Mohseni, and P. Rebentrost, Nat. Phys. **10**, 631 (2014).
[30] A. Brazier, V. Bužek, and P. L. Knight, Phys. Rev. A **71**, 032306 (2005).
[31] A. W. Harrow, arXiv:1308.6595.
[32] S. Aaronson and A. Arkhipov, Theory Comput. **9**, 143 (2013).
[33] R. Simion and F. W. Schmidt, Discrete Mathematics **46**, 107 (1983).