

Storage and retrieval of two unknown unitary channels

Michal Sedlák ^{1,2}, Robert Stárek ³, Nikola Horová³, Michal Mičuda ³, Jaromír Fiurášek ³ and Alessandro Bisio ^{4,5,*}¹*RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 84511 Bratislava, Slovakia*²*Faculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic*³*Department of Optics, Faculty of Science, Palacký University, 17. listopadu 1192/12, 77900 Olomouc, Czech Republic*⁴*Dipartimento di Fisica dell'Università di Pavia, via Bassi 6, 27100 Pavia, Italy*⁵*Istituto Nazionale di Fisica Nucleare, Gruppo IV, via Bassi 6, 27100 Pavia, Italy*

(Received 3 September 2025; revised 12 May 2026; accepted 25 June 2026; published 2 September 2026)

We address the task of converting n uses of an unknown unitary transformation into a quantum state (i.e., storage) and later retrieval of the transformation. Specifically, we consider the case where the unknown unitary is selected with equal prior probability from two options. First, we prove that the optimal storage strategy involves the sequential application of the n uses of the unknown unitary, and it produces the optimal state for discrimination between the two possible unitaries. Next, we show that incoherent “measure-and-prepare” retrieval achieves the maximum fidelity between the retrieved operation and the original (qubit) unitary. We then identify the retrieval strategy that maximizes the probability of successfully and perfectly retrieving the unknown transformation. In the regime in which the fidelity between the two possible unitaries is large the probability of success scales as $P_{\text{succ}} = 1 - O(n^{-2})$, which is a quadratic improvement with respect to the case in which the unitaries are drawn from the entire unitary group $U(d)$ with uniform prior probability. Finally, we present an optical experiment for this approach and assess the storage and retrieval quality using quantum tomography of states and processes. The results are discussed in relation to nonoptimal measure-and-prepare strategy, highlighting the advantages of our protocol.

DOI: [10.1103/56sd-xw8b](https://doi.org/10.1103/56sd-xw8b)

I. INTRODUCTION

The nonorthogonality of quantum states is a distinctive feature of quantum theory and the fundamental reason why quantum information cannot be cloned [1–3], gathered without disturbance [4,5], or have its “logical value” inverted [6,7]. A simple system of two quantum states that cannot be perfectly discriminated provides significant insights into many aspects of quantum theory [8]. This remains true when considering transformations as carriers of quantum information. A transformation between possibly different types of quantum systems is, in fact, a more general concept than that of a quantum state, which can be considered as a special case of transformation, whose input system is trivial. As a consequence, there are more ways in which we can process transformations and this leads to a new paradigm for quantum computation and information processing in which transformations are processed by quantum circuits with open slots [9–13] or, more generally, by higher-order maps which may also exhibit indefinite causal order [14–16]. Many tasks within this paradigm were considered like discrimination [17–20], cloning [21,22], information-disturbance tradeoff [23,24],

inversion [25–28], and complex conjugation [29]. With no analogy to states these tasks may differ, for example, by the order in which the transformation to be used and those to be created are available. A single use of a transformation can be modified by a chosen (pre) or (post)transformation or also by using an additional ancillary system interconnecting them.

These new possibilities make a difference, for example, in the discrimination tasks. Two unitary transformations can be nonorthogonal, i.e., not perfectly distinguishable in a single use, but in contrast to states, they can be made perfectly distinguishable if they are used sufficiently many times. On the other hand, similarly to states, completely unknown unitary transformation cannot be cloned. In fact there are not many tasks for just two possible unitary transformations, which would be already studied despite their fundamental interest. The goal of this paper is to partially fill this gap.

In this paper, we focus on the problem of storage and retrieval [30–33] of an unknown unitary transformation that is randomly selected from a set of two unitary transformations, say U_0 and U_1 , which cannot be perfectly discriminated. The scenario is as follows: Currently, we are free to use a black box n times, which either performs unitary U_0 or U_1 . However, later, the black box will no longer be available, and we will be asked to reproduce its action on some unknown input state, say $|\xi\rangle$. If n uses allow to perfectly discriminate between U_0 and U_1 , then the problem trivializes. Therefore, we will from now on assume the nontrivial case in which U_0 and U_1 cannot be perfectly discriminated with n uses. The most general

*Contact author: alessandro.bisio@unipv.it

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

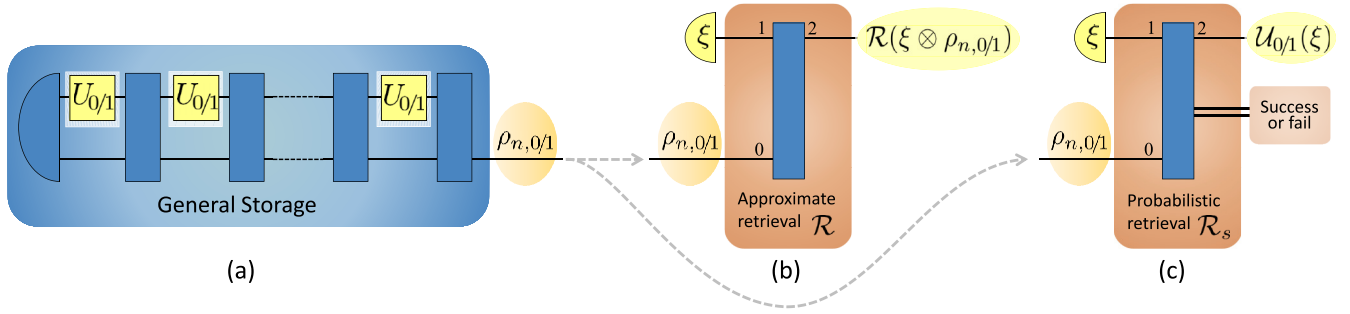


FIG. 1. Illustration of a general storage and retrieval protocol for two d -dimensional unitary transformations U_0, U_1 . These operations are used n times during the storage phase (a). Retrieval can be deterministic (b) and hence yielding only approximation to $\mathcal{U}_{0/1}(\xi)$, or probabilistic (c), yielding perfect application of the stored unitary to any state ξ .

strategy we can use is illustrated in Fig. 1(a) and it consists of two phases: *storage* and *retrieval*. In the storage phase, we run a quantum circuit that makes n calls to the black box and we store the resulting output state $\rho_{n,i}$ in a quantum memory. In the retrieval phase, once the new input $|\xi\rangle\langle\xi|$ becomes available, we feed both $\rho_{n,i}$ and $|\xi\rangle\langle\xi|$ into a retrieving channel $\mathcal{R}$, which should emulate the action of the unknown unitary on the state $|\xi\rangle$. Storage and retrieval can also be interpreted as a machine-learning task [34–37], where the storage phase corresponds to the training of a quantum network. Storage and retrieval tasks for unitary transformations were studied so far only for the whole d -dimensional unitary group [30,31] or its continuous subgroup of qubit phase gates [38]. The storage and retrieval of two unitaries is qualitatively different because of its discreteness, and can be seen as a problem of the same type, but with much stronger prior knowledge about the stored transformation. In fact, the prior knowledge is ultimately high as we are missing just a single classical bit to exactly specify the stored object. Consequently, it is clear that the studied task has the potential to be realized with better performance, while the more widely applicable strategies for storage and retrieval of larger sets of unitaries generally provide only a suboptimal solution to this task.

In the presented results, we address the optimization of storage and retrieval according to two different criteria. In the *approximate deterministic* case, we require that the retrieval works every time (i.e., the retrieval is a quantum channel), and the goal is to maximize the average process fidelity. In the *perfect probabilistic* case, we allow the possibility of retrieval failure (i.e., the retrieval is a two-outcome quantum instrument), but in case of success, the unknown unitary must be retrieved perfectly. Here, the goal is to maximize the probability of success.

First, we show that regardless of the scenario considered in the retrieval phase, the optimal storage phase involves preparing the state that is optimal for the discrimination of the two unitaries [18]. This state can be generated by the sequential application of n uses of the unknown unitary without the need for an ancillary system or entanglement.

Next, we analytically solve the storage and retrieval problem in the approximate deterministic case for qubit unitaries. We demonstrate that the optimal protocol is a “measure-and-prepare” strategy, meaning that quantum memory is not required between the storage and retrieval phases. This feature

was also observed in the case where the unknown unitary is randomly drawn from a group [30].

We then derive the optimal probability of success for the perfect probabilistic case. We observe that in the regime in which the fidelity between the unitaries is big (i.e., far from the regime in which U_0 and U_1 can be perfectly discriminated) the probability of success scales as $P_{\text{succ}} = 1 - O(n^{-2})$. This result can be compared to that in Refs. [31,38], where the unknown unitary was randomly drawn from a group, and the probability of success scaled as $P_{\text{succ}} = 1 - O(n^{-1})$. The more accurate prior knowledge in the present case [two unitaries vs the entire group $U(d)$ with uniform distribution] is responsible for the observed quadratic improvement.

In order to make our results more practically applicable, we derive a short quantum circuit which realizes the optimal perfect probabilistic storage and retrieval of two unknown qubit unitaries. It contains just single-qubit gates, one CNOT gate, and one 3-outcome qubit positive operator-value measure (POVM).

Finally, we also build a quantum linear optical experiment implementing the proposed scheme. The setup uses CNOT gate consisting of partially polarizing beam splitters (PPBS) and utilizing the two-photon interference. The unambiguous measurement part is realized using a Mach-Zehnder-type interferometer made with polarizing beam splitters and wave plates in its arms. We comprehensively characterize the quality of experimentally retrieved quantum operations using quantum process tomography.

The paper is organized as follows. In Sec. II we provide the analytical optimization of the storage and retrieval problem. First we derive the optimal storage protocol and then, in Sec. II A, we solve the optimization of retrieval in the approximate deterministic case and in the perfect probabilistic case. We connect optimal retrieval with the concept of a programmable quantum processor and we interpret the obtained results in this language in Sec. II B. Section II C is devoted to deriving a quantum circuit for the realization of optimal storage and retrieval protocols in the qubit case. In Sec. III we present a proof-of-principle quantum optics experiment that implements the perfect probabilistic version of the optimal storage and retrieval. Section IV contains a discussion of the obtained theoretical and experimental results. Some technical parts of proofs and experiment description are placed in the Appendixes.

II. STORAGE AND RETRIEVAL OF TWO UNITARY CHANNELS

The unknown d -dimensional unitary transformation that we are supposed to use only n times during the storage phase is chosen with equal prior probability from two options, which are denoted as U_0 and U_1 in the rest of the paper. We will denote by $\mathcal{U}_i(\cdot) = U_i \cdot U_i^\dagger$ the action of the given unitary on the density operators.

If U_0 and U_1 are perfectly distinguishable with n uses, then the storage and retrieval problem trivializes. In particular, in the storage phase we can exactly discriminate between U_0 and U_1 . Then, depending on the outcome, in the retrieval phase we prepare the unknown unitary. Therefore, we will assume in the rest of the paper that the unitaries U_0 and U_1 cannot be perfectly distinguishable by their n uses.

If only one use of the unknown transformation is available (i.e., $n = 1$) the storage phase consists of applying the single use of the unitary to a state ρ . On the other hand, if many uses are available (i.e., $n > 1$) the most general storage phase consists of inserting the n uses of the unknown unitary transformation in the open slots of a quantum circuit [see Fig. 1(a)]. At the end of the storage phase, we obtain one of two states $\rho_{n,0}$, $\rho_{n,1}$ depending on the identity of the stored transformation.

A quantum circuit implementing the storage phase is generally a network of quantum transformations that includes channels and measurements. We can always dilate each component of the network to be an isometric transformation [12]. Therefore, we can assume without loss of generality that the states obtained at the end of the storage phase are pure states $\rho_{n,0} = |\psi_{n,0}\rangle\langle\psi_{n,0}|$, $\rho_{n,1} = |\psi_{n,1}\rangle\langle\psi_{n,1}|$. For any pair of states $|a_0\rangle$ and $|a_1\rangle$ with $|\langle a_0|a_1\rangle| \geq |\langle\psi_{n,0}|\psi_{n,1}\rangle|$ one can show that there exists a channel $\mathcal{C}$ such that $\mathcal{C}(|\psi_{n,i}\rangle\langle\psi_{n,i}|) = |a_i\rangle\langle a_i|$. As a consequence the optimal storage phase is the one for which the scalar product $|\langle\psi_{n,0}|\psi_{n,1}\rangle|$ is minimum. However, from the results on the optimal discrimination of unitaries [17,18,39] we know that $|\langle\psi_{n,0}|\psi_{n,1}\rangle| \geq \cos(2n\alpha) > 0$ and 4α corresponds to the length of the smallest arc containing all the eigenvalues of $U_1^\dagger U_0$ on the unit circle. Note that $4n\alpha < \pi$ and therefore $\cos(2n\alpha) > 0$ because we assumed that U_0 and U_1 cannot be perfectly discriminated with n uses.

By suitable fixed preprocessing or postprocessing unitary transformations V and W , we can transform the storage and retrieval protocol for $\{U_0, U_1\}$ into a protocol for $\{WU_0V, WU_1V\}$ with the same performance. The same holds for the discrimination of unitaries. Moreover, the global phase of the unitary operator is irrelevant. Thus, without loss of generality, we can assume that (consider choosing $V = U_0^\dagger$, $W = (U_1U_0^\dagger)^{-1/2}$)

$$U_0 = e^{i\alpha}|0\rangle\langle 0| + \sum_{k=1}^{d-2} e^{i\beta_k}|k\rangle\langle k| + e^{-i\alpha}|d-1\rangle\langle d-1|,$$

$$U_1 = e^{-i\alpha}|0\rangle\langle 0| + \sum_{k=1}^{d-2} e^{-i\beta_k}|k\rangle\langle k| + e^{i\alpha}|d-1\rangle\langle d-1|, \quad (1)$$

where the eigenvalues $e^{i\beta_k}$ are contained in the smallest arc connecting $e^{i\alpha}$ and $e^{-i\alpha}$ and d is the dimension. The optimal

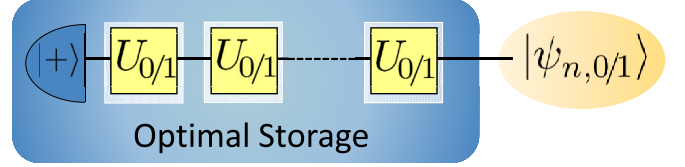


FIG. 2. Optimal quantum circuit for storage of two d -dimensional unitary transformations U_0, U_1 used n times. The circuit is the same for both considered types of retrieval.

storage protocol is thus provided by the sequential strategy derived in Ref. [18] as follows:

Proposition 1 (Optimal storage). Without loss of generality, let us assume that U_0 and U_1 are as in Eq. (1). The optimal storage strategy is then given (see Fig. 2) by applying U_i^n to the input state $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |d-1\rangle)$, i.e., $|\psi_{n,i}\rangle = U_i^n|+\rangle$.

Let us now consider the retrieval phase. The input of this phase is the tensor product between the state $|\psi_{n,i}\rangle$, which encodes information about the unknown unitary U_i , and the state ξ , to which the retrieved transformation should be applied. The result of this action should ideally be $\mathcal{U}_i(\xi)$. We can now approach the problem in at least two different ways.

In the approximate deterministic case, the retrieval is a quantum channel $\mathcal{R}$, which maps bipartite input into single-partite output. Within this approach, the goal is to find the optimal $\mathcal{R}$ such that $\mathcal{R}(|\psi_{n,i}\rangle\langle\psi_{n,i}| \otimes \xi)$ is as close as possible to $\mathcal{U}_i(\xi)$. For the sake of clarity, it is convenient to introduce the following labeling of the Hilbert spaces [see Fig. 1(b)]. We denote with $\mathcal{H}_0$ the system carrying the output of the storage phase ($|\psi_{n,i}\rangle \in \mathcal{H}_0$), with $\mathcal{H}_1$ the system carrying the input ξ , and with $\mathcal{H}_2$ the output system of the retrieval phase, i.e., $\mathcal{R} : \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2)$. In order to assess the quality of the retrieval we choose as a criterion the average of state fidelity over pure input states, namely,

$$F_{\text{avg}} := \int d\phi \sum_{i=0,1} \frac{1}{2} f(\mathcal{R}(\psi_{n,i} \otimes |\phi\rangle\langle\phi|), \mathcal{U}_i(|\phi\rangle\langle\phi|)),$$

$$\psi_{n,i} := |\psi_{n,i}\rangle\langle\psi_{n,i}|, \quad (2)$$

where $d\phi$ is the Haar measure, and $f(\cdot, \cdot)$ is the state fidelity. Equation (2) can be rewritten [40,41] in terms of quantum process fidelity as follows:

$$F_{\text{avg}} = \frac{1}{d+1} + \frac{d}{d+1} F_e, \quad (3)$$

$$F_e \equiv \text{Tr}[RD], \quad (4)$$

$$D := \frac{1}{2d^2} \sum_{i=0,1} |\psi_{n,i}^*\rangle\langle\psi_{n,i}^*| \otimes |U_i\rangle\langle U_i|, \quad (5)$$

$$|\psi_{n,i}^*\rangle := U_i^{*n}|+\rangle, \quad (6)$$

where $R = (\mathcal{I} \otimes \mathcal{R})|I\rangle\langle I| \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1 \otimes \mathcal{H}_2)$ [$|I\rangle\langle I| \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1 \otimes \mathcal{H}_0 \otimes \mathcal{H}_1)$] is the Choi-Jamiołkowski operator [42,43] of the channel $\mathcal{R}$, $|U_i\rangle\langle U_i| \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is the Choi-Jamiołkowski operator of the unitary channel $\mathcal{U}_i$, we used the notation [12] $|A\rangle := \sum_{i,j} A_{i,j}|i\rangle\langle j| = (I \otimes A)|I\rangle$, for an operator $A = \sum_{i,j} A_{i,j}|i\rangle\langle j|$, and A^* denotes the complex conjugate. The isomorphism $A \leftrightarrow |A\rangle$, the Choi-Jamiołkowski isomorphism, the transposition A^T , and the complex

conjugation A^* are defined with respect to some fixed orthonormal basis $\{|i\rangle\}_{i=0}^{d-1}$ which is the same basis in which U_0 and U_1 take the form of Eq. (1). The normalization factor guarantees that $0 \leq F_e \leq 1$. We are then left with the following optimization problem for the approximate deterministic retrieval:

$$F_e = \underset{R}{\text{maximize}} \quad \text{Tr}[RD] \\ \text{subject to} \quad R \geq 0, \quad \text{Tr}_2[R] = I, \quad (7)$$

where Tr_2 is the partial trace on the Hilbert space $\mathcal{H}_2$. Notice that we could loosen the constraint by requiring that R should be a quantum operation ($\text{Tr}_2[R] \leq I$) and not necessarily a quantum channel (i.e., $\text{Tr}_2[R] = I$). It is easy to verify that this change is irrelevant: even if we optimize over the larger set of quantum operations, the optimal solution will still be a quantum channel. Indeed, let R be the optimal quantum operation. Then there exists a quantum channel S such that $R \leq S$ and $\text{Tr}[DR] \leq \text{Tr}[DS]$ because D is a positive operator.

Another approach is perfect probabilistic retrieval. In this case, the retrieval is described by a two-outcome quantum instrument $\mathcal{R} = \{\mathcal{R}_s, \mathcal{R}_f\}$, where the index denotes success (s) or failure (f). If the classical outcome is s , we demand the exact retrieval of the unknown unitary, i.e., $\mathcal{R}_s(\psi_{n,i} \otimes \xi) = \lambda_{\xi,i} \mathcal{U}_i(\xi)$ for any ξ and we define $\lambda_{\xi,i} := \text{Tr}[\mathcal{R}_s(\psi_{n,i} \otimes \xi)]$ (we remind that $\psi_{n,i} := |\psi_{n,i}\rangle\langle\psi_{n,i}|$). The factor $\lambda_{\xi,i}$ is the probability of success of perfect retrieval when the unknown unitary is U_i and the input state is ξ . It is easy to observe that such a probability of success does not depend on ξ , namely, $\lambda_{\xi,i} = \lambda_i$ for any ξ . Indeed, if we had $\lambda_{\xi,i} \neq \lambda_{\xi',i}$ for a pair ξ, ξ' of states, this would imply that $\mathcal{R}_s[\psi_{n,i} \otimes (\xi + \xi')] = \lambda_{\xi,i} \mathcal{U}_i(\xi) + \lambda_{\xi',i} \mathcal{U}_i(\xi') = \mathcal{U}_i(\lambda_{\xi,i} \xi + \lambda_{\xi',i} \xi') \not\propto \mathcal{U}_i(\xi + \xi')$. The perfect retrieval condition can therefore be written as follows:

$$\langle \psi_{n,i}^* | \mathcal{R}_s | \psi_{n,i}^* \rangle = \lambda_i |U_i\rangle\langle U_i|, \quad i = 0, 1 \quad (8)$$

where R_s is the Choi-Jamiolkovski operator of the quantum operation R_s . As a criterion to assess the performance of the retrieval strategy we choose the average probability of success, which in equation reads as

$$P_{\text{succ}} := \frac{1}{2}(\lambda_0 + \lambda_1). \quad (9)$$

It is easy to verify that $\frac{1}{2}(\lambda_0 + \lambda_1) = \text{Tr}[R_s D]$, where D was defined in Eq. (5). Reminding that R_s is a quantum operation if and only if $R_s \geq 0$ and $\text{Tr}_2[R_s] \leq I$, we obtain the following optimization problem for the perfect probabilistic retrieval:

$$P_{\text{succ}} = \underset{R_s}{\text{maximize}} \quad \text{Tr}[R_s D] \\ \text{subject to} \quad R_s \geq 0, \quad \text{Tr}_2[R_s] \leq I, \\ R_s \text{ obeys Eq. (8)}. \quad (10)$$

It is worth noticing that the optimization problems (7) and (10) are almost the same, the only difference being the perfect retrieval condition (8).

A. Optimal retrieval

In this section we will find an analytical solution to the optimization problems of Eqs. (7) and (10) when the unknown unitaries act on a two-dimensional Hilbert space (i.e., U_0 and

U_1 are single-qubit gates). Later, we will present the analytical solution to the optimization problem for perfect probabilistic retrieval for general dimension d . If $d = 2$ operators U_i and states $|\psi_{n,i}\rangle$ simplify as follows:

$$U_0 = e^{i\alpha} |0\rangle\langle 0| + e^{-i\alpha} |1\rangle\langle 1|, \\ U_1 = e^{-i\alpha} |0\rangle\langle 0| + e^{i\alpha} |1\rangle\langle 1|, \quad (11)$$

$$|\psi_{n,0}\rangle = \frac{1}{\sqrt{2}}(e^{i\alpha n} |0\rangle + e^{-i\alpha n} |1\rangle),$$

$$|\psi_{n,1}\rangle = |\psi_{n,0}^*\rangle = \sigma_x |\psi_{n,0}\rangle, \quad \sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|. \quad (12)$$

We remind that $4n\alpha < \pi$ since otherwise the unitaries become perfectly distinguishable.

We now observe that the following commutation relations hold:

$$[D, W(\beta, \gamma, l)] = 0, \quad \forall \beta, \gamma \in \mathbb{R}, \forall l \in \{0, 1\}, \\ W(\beta, \gamma, l) := \sigma^{(l)} \otimes \sigma^{(l)} Z(\beta, \gamma) \otimes \sigma^{(l)} Z^*(\beta, \gamma), \\ Z(\beta, \gamma) := e^{i\beta} |0\rangle\langle 0| + e^{i\gamma} |1\rangle\langle 1|, \\ \sigma^{(0)} := I, \quad \sigma^{(1)} := \sigma_x. \quad (13)$$

We notice that $W(\beta, \gamma, l)$ is a unitary representation of the compact Lie group $G := [\text{U}(1) \times \text{U}(1)] \rtimes \mathbb{Z}$, the semidirect product is defined by the left action $\eta : l \mapsto \eta_l, \eta_0(e^{i\beta}, e^{i\gamma}) = (e^{i\beta}, e^{i\gamma}), \eta_1(e^{i\beta}, e^{i\gamma}) = (e^{i\gamma}, e^{i\beta})$. In order to lighten the notation, we will then write $W(g), g \in G$ in place of $W(\beta, \gamma, l)$. Exploiting this symmetry we can prove the following lemma:

Lemma 1 (Symmetric retrieval is optimal). Without loss of generality, we can assume that the optimal approximate deterministic retrieval R (respectively the optimal perfect probabilistic retrieval R_s) satisfies $[R, W(g)] = 0$ (respectively $[R_s, W(g)] = 0$) for any $g \in G$. Moreover, if $[R_s, W(g)] = 0$ then Eq. (8) holds with

$$\lambda_0 = \lambda_1 = \lambda. \quad (14)$$

Proof. The proof follows the Holevo's averaging argument [44].

Let R be the optimal approximate deterministic retrieval. If R is a quantum operation we observe that also $W^\dagger(g)RW(g)$ is a quantum operation and therefore also the average $\bar{R} = \int_G dg W^\dagger(g)RW(g)$ is a quantum operation. Clearly $[\bar{R}, W(g)] = 0$ for any $g \in G$. Since $[D, W(g)] = 0$ for any $g \in G$ we have that $\text{Tr}[RD] = \text{Tr}[\bar{R}D]$, i.e., $\bar{R}$ is another optimal approximate deterministic retrieval.

The thesis for the perfect probabilistic case follows from the fact that if an optimal R_s obeys the perfect retrieval condition of Eq. (8) with the values λ_0 and λ_1 then the averaged operator $\bar{R}_s$ obeys the perfect retrieval condition of Eq. (8) with $\bar{\lambda}_0 = \bar{\lambda}_1 = \frac{1}{2}(\lambda_0 + \lambda_1)$. ■

Thanks to this result, we can conveniently express the optimal retrieval in a block-diagonal form. First, let us define the following projectors:

$$P := \frac{1}{2}(|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|), \\ P' := \frac{1}{2}(|e'_1\rangle\langle e'_1| + |e'_2\rangle\langle e'_2|), \quad (15)$$

$$|e_1\rangle := |+\rangle|I\rangle, \quad |e_2\rangle := |-\rangle|\sigma_z\rangle, \\ |e'_1\rangle := |+\rangle|\sigma_z\rangle, \quad |e'_2\rangle := |-\rangle|I\rangle. \quad (16)$$

Notice that $P' = (I \otimes \sigma_z)P(I \otimes \sigma_z)$ and $|e'_i\rangle := (I \otimes \sigma_z)|e_i\rangle$. We can now prove the following lemma.

Lemma 2 (Block diagonal retrieval). Without loss of generality, we can assume that the optimal approximate deterministic retrieval R and the optimal perfect probabilistic retrieval R_s satisfy $R = PRP + P'RP'$ and $R_s = PR_sP + P'R_sP'$, respectively. Furthermore, we have

$$\begin{aligned}\langle \psi_{n,i}^* | PR_sP | \psi_{n,i}^* \rangle &= \lambda_A |U_i\rangle \langle U_i|, \\ \langle \psi_{n,i}^* | P'R_sP' | \psi_{n,i}^* \rangle &= \lambda_B |U_i\rangle \langle U_i|\end{aligned}\quad (17)$$

for any $i = 0, 1$.

Proof. One can verify that P and P' are projectors on invariant subspaces of the representation $W(g)$. Since from Lemma 1 we have that $[R, W(g)] = 0$ holds, we have that $R = PRP + P'RP' + QRQ$ where $Q := I - P - P'$ (the same holds for R_s). Since $D = PDP + P'DP'$ the thesis for the approximate deterministic case follows. In the perfect probabilistic case, from inserting $R_s = PR_sP + P'R_sP' + QR_sQ$ into Eq. (8) we conclude that

$$\begin{aligned}\langle \psi_{n,i}^* | PR_sP | \psi_{n,i}^* \rangle &= \lambda_A |U_i\rangle \langle U_i|, \\ \langle \psi_{n,i}^* | P'R_sP' | \psi_{n,i}^* \rangle &= \lambda_B |U_i\rangle \langle U_i|, \\ \langle \psi_{n,i}^* | QR_sQ | \psi_{n,i}^* \rangle &= \lambda_C |U_i\rangle \langle U_i|,\end{aligned}$$

because terms on the left are positive operators, which are supposed to sum up to a rank-one operator $\lambda |U_i\rangle \langle U_i|$. Next, we observe that $Q = I \otimes (|\sigma_x\rangle \langle \sigma_x| + |\sigma_y\rangle \langle \sigma_y|) \equiv I \otimes Q'$, $Q'|U_i\rangle = 0$ and using Eq. (8) we obtain

$$\langle \psi_{n,i}^* | QR_sQ | \psi_{n,i}^* \rangle = Q' \langle \psi_{n,i}^* | R_s | \psi_{n,i}^* \rangle Q' = \lambda_i Q' |U_i\rangle \langle U_i| Q' = 0.$$

Thus, $\langle \psi_{n,i}^* | \otimes \langle \varphi | QR_sQ | \psi_{n,i}^* \rangle \otimes |\varphi\rangle = 0$ for $i = 0, 1$ and $\forall |\varphi\rangle \in \mathcal{H}_{12}$. In other words, support of positive-semidefinite operator QR_sQ has to be orthogonal to the above vectors, which span the whole space. We conclude that the support of QR_sQ is empty, i.e., $QR_sQ = 0$ and the thesis follows. ■

From the block-diagonal structure of Lemma 2, we have a convenient way to rephrase the optimization of the retrieval.

Proposition 2. The optimization of the approximate deterministic retrieval [Eq. (7)] is equivalent to the following state discrimination problem:

$$\begin{aligned}F_e &= \text{maximize}_{A,B} \quad \langle u|A|u\rangle + \langle v|B|v\rangle \\ \text{subject to} \quad & 0 \leq A, B \leq I, \quad A + B \leq I.\end{aligned}\quad (18)$$

The optimization of the perfect probabilistic retrieval (10) is equivalent to the following unambiguous state discrimination problem:

$$\begin{aligned}P_{\text{succ}} &= \text{maximize}_{A,B} \quad \langle u|A|u\rangle + \langle v|B|v\rangle \\ \text{subject to} \quad & 0 \leq A, B \leq I \quad A + B \leq I \\ & \langle v|A|v\rangle = \langle u|B|u\rangle = 0,\end{aligned}\quad (19)$$

where A, B are a 2×2 matrices and we defined

$$\begin{aligned}|u\rangle &:= \begin{pmatrix} c_n c \\ s_n c \end{pmatrix}, \quad |v\rangle := \begin{pmatrix} c_n s \\ -s_n c \end{pmatrix}, \\ s &:= \sin(\alpha), \quad c := \cos(\alpha), \\ s_n &:= \sin(n\alpha), \quad c_n := \cos(n\alpha).\end{aligned}\quad (20)$$

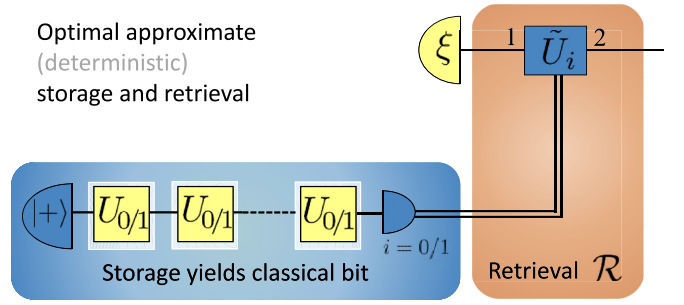


FIG. 3. Optimal approximate deterministic storage and retrieval of two d -dimensional unitary transformations can be realized via optimal discrimination among the two unitaries in the storage phase and conditional preparation of slightly modified unitaries in the retrieval phase.

Proof. The proof is reported in Appendix A. ■

The optimal discrimination of a pair of quantum states has a known analytical solution both in the minimum error and in the unambiguous case. The optimal value of the process fidelity for the approximate deterministic case reads as

$$F_e = \frac{1}{2} + \frac{1}{2} \sqrt{1 - (\sin 2\alpha \cos 2n\alpha)^2}. \quad (21)$$

The optimal strategy can be realized by the measure-and-prepare protocol presented in Fig. 3 (see Appendix B for the details). Since classical information can be cloned, our result can be extended to the case where we must reproduce $m \geq 1$ copies of the unknown unitary with the criterion of maximizing the single-copy fidelity.

The optimal solution to unambiguous state discrimination is known [45,46] and it gives

$$P_{\text{succ}} = \begin{cases} \frac{1 - (\cos 2n\alpha)^2}{2(1 - \cos 2n\alpha \cos 2\alpha)} & \text{if } \alpha \in (0, \chi_n), \\ 1 - \cos 2n\alpha \sin 2\alpha & \text{if } \alpha \in [\chi_n, \frac{\pi}{4n}], \end{cases} \quad (22)$$

where χ_n is the solution of the following trigonometric equation:

$$\cos(2n\alpha) = \frac{1}{\cos(2\alpha) + \sin(2\alpha)}. \quad (23)$$

The range of angles α for which α is below (above) angle χ_n we denote as small- (large-) α regime, respectively. Success probability for various values of n and α is illustrated in Fig. 4. For $\alpha \ll 1$ the fidelity between U_0 and U_1 is large. Therefore, U_0 and U_1 can be perfectly discriminated only with a large number of uses, i.e., $n \gg 1$. Therefore, it makes sense to power expand Eq. (22) around $\alpha = 0$ and study how the probability of success scales with n . We obtain

$$P_{\text{succ}} = 1 - \frac{1}{n^2 + 1} + \frac{n^2 + 2n^4 - 3n^6}{3(n^2 + 1)^2} \alpha^2 + O(\alpha^4). \quad (24)$$

In comparison, the probability of success of the perfect probabilistic storage and retrieval of an unknown unitary transformation which is randomly picked from unitary group $U(d)$ with Haar measure reads as [31,38]

$$P_{\text{succ}} = \begin{cases} 1 - \frac{d^2 - 1}{n + d^2 - 1} & \text{if } U \in U(d), \quad d \geq 2, \\ 1 - \frac{1}{n+1} & \text{if } U = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}. \end{cases} \quad (25)$$

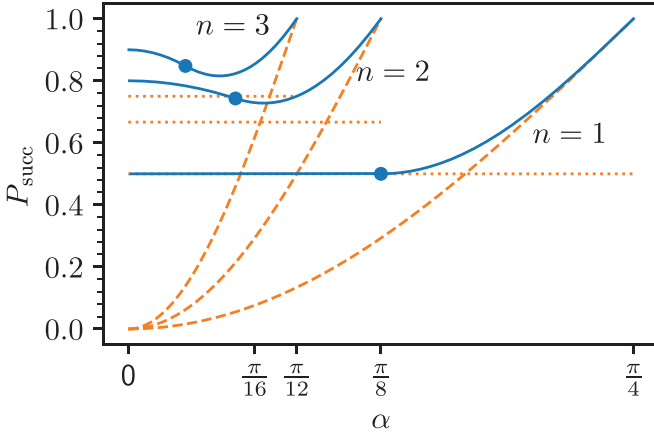


FIG. 4. Optimal success probability of perfect storage and retrieval of two d -dimensional unitary transformations as a function of n the number of black box uses and of the angle α , which equals half of the angular spread of $(U_1)^\dagger U_0$. Dotted and dashed lines represent performance achievable by nonoptimal strategies based on results from Ref. [38] and based on unambiguous discrimination, respectively. The blue circles on the lines mark the transition between small- and large- α regimes of the success probability.

Therefore, a more accurate prior information about the unknown unitary allows us to achieve a quadratic improvement.

Equation (22) provides the optimal probability of success for the perfect probabilistic storage and retrieval of a pair of unitaries also in arbitrary dimension d . Indeed, let U_0 and U_1 be given as in Eq. (1). By restricting the action of U_0 and U_1 on the subspace spanned by $|0\rangle$ and $|d-1\rangle$ we have the two single-qubit unitaries $\tilde{U}_0 := e^{i\alpha}|0\rangle\langle 0| + e^{-i\alpha}|d-1\rangle\langle d-1|$ and $\tilde{U}_1 := e^{-i\alpha}|0\rangle\langle 0| + e^{i\alpha}|d-1\rangle\langle d-1|$. On one hand, the probability of success for the perfect storage and retrieval of the pair $\{\tilde{U}_0, \tilde{U}_1\}$ is given by Eq. (22) and it must be an upper bound to the the probability of success for the perfect storage and retrieval of the pair $\{U_0, U_1\}$. On the other hand, we can provide a retrieval strategy that achieves the bound (see Appendix C for the details).

B. Quantum processor for two unitary transformations

Retrieval phase can be also viewed as a programmable quantum processor since the state resulting from the storage phase programs (encodes) the transformation to be performed on another system.

In the rest of this section we aim to rewrite the obtained analytical formulas in the form more convenient from the perspective of quantum processors. The only relevant parameter for the two possible program states (previously denoted as $|\psi_{n,0}\rangle, |\psi_{n,1}\rangle$) is their mutual angle, which we parametrize through the overlap as $|\langle\psi_{n,0}|\psi_{n,1}\rangle| = \cos \beta$, $\beta \in [0, \pi/2]$, i.e., $\beta = 2n\alpha$, since the optimization of the retrieval leads to the same formulas also if n is considered to be any non-negative number. In this notation, quantum process fidelity, Eq. (21), can be rewritten as

$$F_e = \frac{1}{2} + \frac{1}{2}\sqrt{1 - (\sin 2\alpha \cos \beta)^2}. \quad (26)$$

The corresponding optimal tradeoff between the angle of the program states β , angular spread (4α) of the eigenvalues of

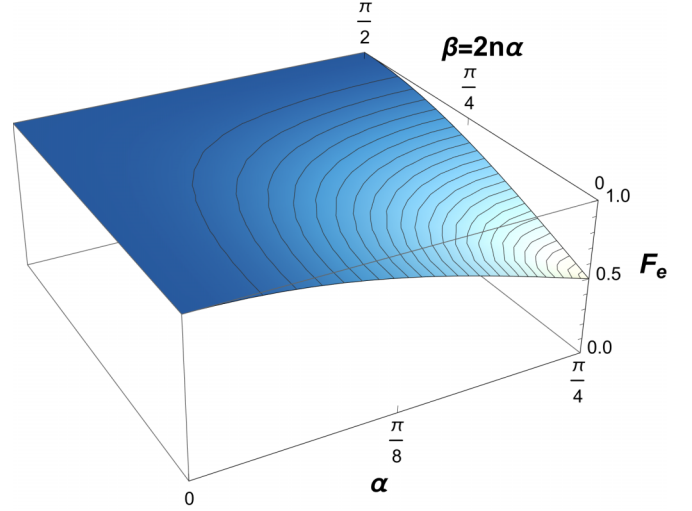


FIG. 5. Maximum achievable quantum process fidelity F_e in the approximate deterministic retrieval as a function of the angle of the program states β and angle α representing $\frac{1}{4}$ of the angular spread of the eigenvalues of the relative unitary $U_1^\dagger U_0$ for unitary transformations U_0, U_1 .

the relative unitary $U_1^\dagger U_0$, and the achievable quantum process fidelity F_e is depicted in Fig. 5.

We can inspect the unambiguous case in the same way. The success probability for unambiguous state discrimination, Eq. (22), can be rewritten in terms of parameters α, β as

$$P_{\text{succ}} = \begin{cases} \frac{1 - (\cos \beta)^2}{2(1 - \cos \beta \cos 2\alpha)} & \text{if } \beta \leq \beta_B(\alpha), \\ 1 - \cos \beta \sin 2\alpha & \text{if } \beta \geq \beta_B(\alpha), \end{cases} \quad (27)$$

$$\beta_B(\alpha) := \arccos \frac{1}{\cos 2\alpha + \sin 2\alpha}. \quad (28)$$

We depict the achievable tradeoff between parameters α, β and success probability in Fig. 6. Let us emphasize that, in contrast to the approximate deterministic case, in perfect probabilistic case Eq. (27) describes the optimal success probability also for any pair of d -dimensional unitary transformations, whose relative unitary $U_1^\dagger U_0$ has angular spread of eigenvalues equal to 4α .

C. Quantum circuit for perfect storage and retrieval of two unitaries

Next, we return to a qubit case ($d = 2$) and we derive a simple quantum circuit realizing the optimal perfect probabilistic retrieval. We will use the following notation to simplify the formulas:

$$\begin{aligned} \tilde{c}_n &:= \cos 2n\alpha, & \tilde{s}_n &:= \sin 2n\alpha, \\ c_n &:= \cos n\alpha, & s_n &:= \sin n\alpha, \\ \tilde{c} &:= \cos 2\alpha, & \tilde{s} &:= \sin 2\alpha, \\ c &:= \cos \alpha, & s &:= \sin \alpha. \end{aligned} \quad (29)$$

Proposition 3. Quantum circuit depicted in Fig. 7 performs optimal perfect probabilistic retrieval of qubit unitaries U_0, U_1 if the qutrit-basis measurement outcomes 0,1 are considered as success and σ_Z correction is performed in case of

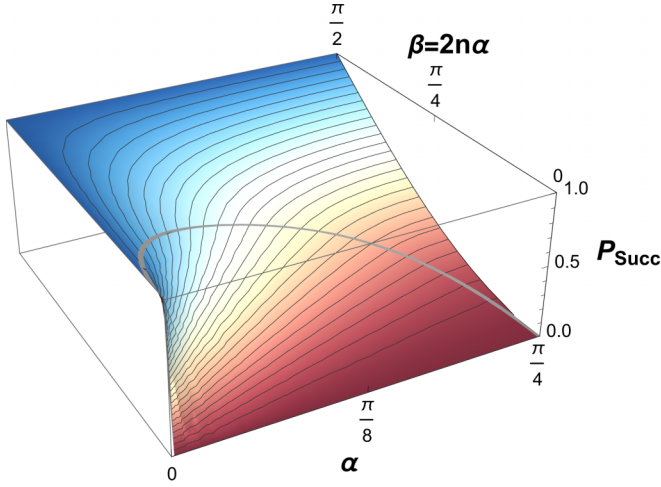


FIG. 6. Maximum achievable success probability P_{succ} in the perfect probabilistic retrieval as a function of the angle of the program states β and angle α representing $\frac{1}{4}$ of the angular spread of the eigenvalues of the relative unitary $U_1^\dagger U_0$ for unitary transformations U_0, U_1 .

outcome 1. Measurement outcome 2 is considered as a failure. In particular, we have

$$\begin{aligned} {}_A\langle 0| \otimes I_2 (M \otimes I) C_X |\psi_{n,k}\rangle \otimes I_1 &= \sqrt{\lambda_A} U_k, \\ {}_A\langle 1| \otimes I_2 (M \otimes \sigma_z) C_X |\psi_{n,k}\rangle \otimes I_1 &= \mp i \sqrt{\lambda_B} U_k, \end{aligned} \quad (30)$$

where

$$\lambda_A = \frac{1 + \tilde{c}_n(\tilde{c} - \tilde{s})}{2}, \quad \lambda_B = \frac{1 - \tilde{c}_n(\tilde{c} + \tilde{s})}{2},$$

and the operator M describes the following qubit-to-qutrit isometry:

$$M := \begin{cases} |a_1\rangle\langle +| + |a_2\rangle\langle -| & \text{if } \alpha \in [\chi_n, \frac{\pi}{4n}], \\ |a'_1\rangle\langle +| + |a'_2\rangle\langle -| & \text{if } \alpha \in (0, \chi_n) \end{cases} \quad (31)$$

with

$$\begin{aligned} |a_1\rangle &:= \sqrt{\lambda_A} \frac{c}{c_n} |0\rangle + \sqrt{\lambda_B} \frac{s}{c_n} |1\rangle + \sqrt{v_+} |2\rangle, \\ |a_2\rangle &:= \sqrt{\lambda_A} \frac{s}{s_n} |0\rangle - \sqrt{\lambda_B} \frac{c}{s_n} |1\rangle - \sqrt{v_-} |2\rangle, \end{aligned}$$

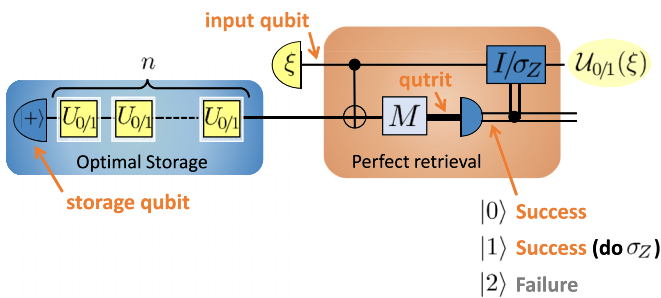


FIG. 7. Quantum circuit for optimal perfect probabilistic storage and retrieval of two single-qubit unitary transformations. Isometry M maps qubit to qutrit and together with the projective measurement of the qutrit forms a three-outcome rank-one POVM.

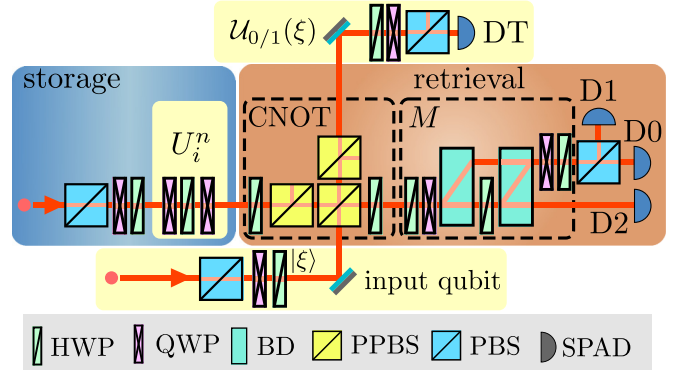


FIG. 8. Experimental implementation of the probabilistic version of storage and retrieval protocol. Experimental setup consisting of half- and quarter-wave plates (HWP, QWP), (partially) polarizing beam splitters [(P)PBS], calcite beam displacers (BD), and single-photon avalanche detectors (SPAD). The encoding of stored operation (U_i^n) is highlighted with a yellow box. CNOT gate and isometry M are highlighted by dashed boxes.

$$\begin{aligned} |a'_1\rangle &:= \sqrt{a}|0\rangle + \sqrt{b}|2\rangle, & |a'_2\rangle &:= \sqrt{b}|0\rangle - \sqrt{a}|2\rangle, \\ v_+ &:= \frac{\tilde{c}_n(1 - \tilde{c}^2 + \tilde{s})}{1 + \tilde{c}_n}, & v_- &:= \frac{\tilde{c}_n(\tilde{c}^2 - 1 + \tilde{s})}{1 - \tilde{c}_n}, \\ a &:= \frac{(1 + \tilde{c})(1 - \tilde{c}_n)}{2(1 - \tilde{c}\tilde{c}_n)}, & b &:= \frac{(1 - \tilde{c})(1 + \tilde{c}_n)}{2(1 - \tilde{c}\tilde{c}_n)}. \end{aligned}$$

It is easy to verify that Eq. (30) holds if we realize that $U_0 = cI + is\sigma_z$, $U_1 = cI - is\sigma_z$ and write the controlled not (CNOT) as $C_X = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes (|+\rangle\langle +| - |-\rangle\langle -|)$. The full proof of Proposition 3 and the derivation of the optimal circuit is placed in Appendix D.

An appealing feature of the proposed circuit is its relatively small complexity: it contains only few elementary elements such as qubits, gates, or their conditioning and, most notably, only one entangling CNOT gate. One might wonder how close to an ideal implementation one can get, especially from the point of view of perfectly implementing the stored unitary transformation. We answer this question in case of quantum optical setups by conducting the experiment reported in the following section.

III. PHOTONIC DEMONSTRATION

We present a proof-of-principle experiment that tests the probabilistic version of the proposed storage and retrieval protocol described in Sec. II C and depicted in Fig. 7. The experiment is based on a linear optical implementation [47]. Correlated photon pairs are generated in a CW-pumped type-II spontaneous parametric down-conversion (SPDC) process at a central wavelength of 810 nm and guided via optical fibers to the experimental setup depicted in Fig. 8.

The qubits are encoded into the polarization states of generated photons. One qubit serves for storage and is initially prepared in state $|+\rangle$ using wave plates. The subsequent stack of wave plates realizes operation U_i^n , introducing the phase shift to the initial storage state. Such a qubit could be, in principle, stored in photonic quantum memory [48,49].

The CNOT gate and isometry M followed by measurement in the qutrit computational basis implement the retrieval. To

retrieve the stored operation and apply it onto the input qubit in a state $|\xi\rangle$, we first entangle both qubits using a probabilistic CNOT gate consisting of partially polarizing beam splitters (PPBS) and utilizing the two-photon interference [50–52]. The storage qubit enters the CNOT gate and acts as a target qubit. State $|\xi\rangle$ of the input qubit serves as a control and it is prepared using another pair of wave plates. The tomographic characterization showed CNOT gate process fidelity of 0.929(1) with the 1.8-kHz rate of detected coincidences.

We realize the POVM measurement on the storage qubit using an optical setup for unambiguous state discrimination (USD). The optical setup for the USD consists of a Mach-Zehnder-type interferometer made with polarizing beam splitters and wave plates in its arms [53]. Such a device attenuates one of the polarization components, making the to-be-discriminated states orthogonal at the expense of success probability. We implement this device using wave plates and an interferometer based on calcite beam displacers [54], offering compact construction and passive phase stability [55]. The corresponding wave plates are set according to the choice of the states that are to be discriminated, which is parametrized by $|\alpha|$. The detailed description of this experimental block, including the relation between parameter $|\alpha|$ and wave-plate angles, is provided in Appendix E. The success of the protocol is heralded by the detection of coincidence DT-D0 (outcome 0) or DT-D1 (outcome 1), while the coincidence DT-D2 (outcome 2) heralds the failure.

In the case of coincidence detection at DT-D1, one has to apply an additional phase flip on the input qubit. Such a feedforward operation could be physically applied using a fast electro-optical phase modulator, fast digital logic, and a sufficiently long delay line for the input qubit [56–63]. However, in the tomographic characterization of the protocol, we can emulate the feedforward using data postprocessing [64]. The final state of the input qubit is analyzed by means of projective measurements implemented by the remaining pair of wave plates, subsequent polarizing beam-splitter cube, and detector DT. The retrieved quantum operations are reconstructed using the maximum-likelihood method [65] from the collected tomographic data.

Let us discuss in detail how we test the protocol. We sweep parameter $|\alpha|$ from 0 to $\frac{\pi}{4n}$, and for each value, we set the discrimination part accordingly. For $|\alpha| > \frac{\pi}{4n}$ the success probability is equal to one. We have tested both stored operations U_i^n , i.e., both positive and negative phase shift applied to the storage qubit.

Having chosen U_i ($i = 0, 1$) in the storage phase we characterize the effective quantum channel for the input qubit using quantum process tomography. We do that by sequentially preparing the input qubit in eigenstates of Pauli operators and measuring the corresponding output states in the X , Y , and Z basis. Here, i -basis measurement means the measurement of expectation value $\langle\sigma_i\rangle$. We record three coincidence tomograms, DT-D0 and DT-D1, corresponding to the protocol success, and DT-D2, corresponding to the failure. We flip the sign of X and Y basis readings in the DT-D1 tomogram to emulate the feedforwarded conditional phase flip corresponding to the conditional σ_0/σ_Z gate from Fig. 7.

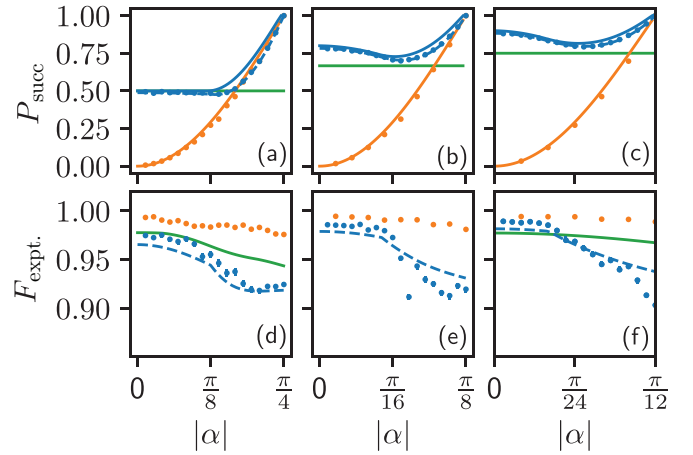


FIG. 9. Success probability (a)–(c) and conditional process fidelity (d)–(f) for $n = 1$ (a), (d), $n = 2$ (b), (e), and $n = 3$ (c), (f) uses of the stored operation. Points show the experimental data, while solid lines represent the ideal theoretical predictions. The $1\text{-}\sigma$ error bars are smaller than the point size. The blue color is related to the presented storage-and-retrieval protocol, and the orange to the measure-and-prepare strategy. The blue dashed lines represent the theoretical predictions based on the reconstructed process matrix of the CNOT gate and additional experimental imperfections. For comparison, green lines in (a), (b), (c), (d), and (f) show the theoretical prediction for the storage and retrieval protocol from [38] that works for an unknown arbitrary phase α and we assumed the same quality of CNOT gate. The same green line for conditional process fidelity for $n = 2$ is below the range reported here since the best corresponding quantum circuit requires eight CNOT gates, in contrast to one or two CNOT gates needed in the other cases.

We can directly evaluate success probability from the tomograms as

$$P_{\text{succ}} = \frac{1}{2} \left(\frac{S_{0+} + S_{1+}}{S_{0+} + S_{1+} + S_{2+}} + \frac{S_{0-} + S_{1-}}{S_{0-} + S_{1-} + S_{2-}} \right), \quad (32)$$

where $S_{j\pm}$ is the sum of all counts in the tomogram obtained from coincidence DT-D j when the sign of the phase was $+$ ($-$). This determination of the retrieval success probability is independent of the success probability of the CNOT gate. By summing tomograms DT-D0 and DT-D1, we obtain an effective tomogram, which we reconstruct to get an estimate of the Choi-Jamiolkowski matrix $C_{\text{expt},i}$ of the quantum channel acting on the input qubit conditioned on the success of the protocol (observation of DT-D0 or DT-D1 coincidence).

We quantify the quality of the retrieval with the average channel fidelity between the experimentally retrieved transformation and the stored unitaries:

$$F_{\text{expt}} = \frac{1}{8} \sum_{i=0,1} \langle\langle U_i | C_{\text{expt},i} | U_i \rangle\rangle. \quad (33)$$

Thus, the probability of success and channel fidelity F_{expt} , plotted using the blue color in Fig. 9, are averages over the choice of U_0 or U_1 (sign of α).

For comparison, we also consider the perfect probabilistic measure-and-prepare strategy in which the storage qubit undergoes USD, and the result is stored in a classical bit if the USD succeeds. Upon retrieval, we simply apply U_0 or U_1 to

the input qubit in state $|\xi\rangle$ based on the value of the saved bit. To test this in the experiment, we first prepared $U_0^n|+\rangle$ at the USD input, noting the relative frequencies of each outcome of the USD, which we denote as $f_0^{(0)}$, $f_1^{(0)}$, and $f_2^{(0)}$ (the outcome 2 is an inconclusive outcome). Theoretically, $f_1^{(0)} = 0$, when we store operation U_0 , but due to experimental imperfections, there is a small fraction of wrong discrimination results. The success probability for stored U_0 is denoted as $P_{\text{succ}}^{(m\&p)}(U_0)$ and is estimated as follows:

$$P_{\text{succ}}^{(m\&p)}(U_0) = \frac{f_0^{(0)} + f_1^{(0)}}{f_0^{(0)} + f_1^{(0)} + f_2^{(0)}}. \quad (34)$$

We aim at evaluating the performance of the whole measure-and-prepare strategy, thus, we also performed process tomography of the direct implementation of channels U_0 and U_1 , which should be applied based on the USD outcome. Obtained Choi-Jamiołkowski matrices are denoted as $C_{\text{dir},0}$ and $C_{\text{dir},1}$.

Then, the Choi operator of the retrieved operation when the inserted unitary is U_0 reads as follows:

$$C_{\text{expt},0}^{(m\&p)} = \frac{f_0^{(0)} C_{\text{dir},0} + f_1^{(0)} C_{\text{dir},1}}{f_0^{(0)} + f_1^{(0)}}. \quad (35)$$

This analysis is also done for storage state $U_1^n|+\rangle$, and finally the obtained fidelities of reconstructed Choi matrices are averaged to obtain $F_{\text{expt}}^{(m\&p)}$ as in Eq. (33). We repeat this procedure for each tested α .

The data in Figs. 9(a)–9(c) show a good agreement between the observed success probability and its theoretical prediction. Error bars in Figs. 9(a)–9(c) are obtained directly by error propagation assuming Poissonian statistics of the detected counts.

As demonstrated, the presented protocol surpasses the measure-and-prepare strategy in terms of success probability. Especially for low α , the success probability given by Eq. (24) is much higher than the one achievable by the measure-and-prepare strategy. However, the fidelity achieved by our protocol in the experiment is lower than the one obtained in the case of the measure-and-prepare strategy, as indicated in Figs. 9(d)–9(f). To determine the error bars in Figs. 9(d)–9(f) we resort to bootstrapping, i.e., we assume the Poissonian statistics of the measured counts and use our knowledge of measurement operators and the reconstructed Choi matrix to generate synthetic data [66]. These data correspond to the repeated execution of the experiment, which we again reconstruct to obtain the fidelity. Because the fidelity distribution is asymmetric close to the boundary value of its definition interval, we use the 0.159 and 0.841 quantiles as lower and upper error bars. This choice is equivalent to one standard deviation in the case of symmetric distributions.

In contrast to the presented protocol, the measure-and-prepare strategy requires only single-qubit operations, which are very robust for polarization qubits, and therefore it exceeds the presented protocol in terms of fidelity.

A. Comparison to storage and retrieval of arbitrary phase gate protocol

Here, we compare the realized protocol to the protocol for storing arbitrary phase gates [38] in the case of $n = 1$ and 3.

In the compared protocol, one first applies the stored phase shift once to the *storage qubit 1*. If we are allowed to use U_i three times, we also apply U_i twice to the *storage qubit 2*, which ends the storage phase. Then in the retrieval phase, we apply the CNOT gate controlled by *input qubit* to the storage qubit 1. If the subsequent measurement of storage qubit 1 in the Z basis yields 0, then the protocol succeeded. If it yields 1 it implies failure for $n = 1$, but for $n = 3$ we apply another CNOT gate controlled by the input qubit to the storage qubit 2 and measure it in the Z basis. Such a protocol would succeed independently on $|\alpha|$ with probability $P_{\text{succ}} = \frac{1}{2}$ for $n = 1$ and $P_{\text{succ}} = \frac{3}{4}$ for $n = 3$, respectively. In Fig. 9, we see that the presented protocol described by Proposition 3 (blue data points) outperforms the protocol [38] (green lines) in terms of success probability.

We utilized the reconstructed Choi matrix of the CNOT gate to theoretically compare both protocols in terms of their robustness to realistic experimental imperfections. We input the imperfect experimental Choi matrix into numerical simulations of the protocol [38] and plotted the achieved fidelity as green solid lines in Figs. 9(d) and 9(f). For $n = 1$ the ideal versions of the compared protocols coincide, i.e., demand the same quantum circuit to be performed, in the small- α regime. From that part of Fig. 9 we see that the experimentally realized protocol has slightly smaller process fidelity due to imperfections in the performed projective measurement, which are not accounted for in the predictions illustrated by the green line. In the large- α regime presented protocol has higher success probability, but due to additional disturbance caused by the imperfect POVM realization lower process fidelity. However, for $n = 3$, the protocol [38] can use up to two applications of the CNOT gates, which on average results in lower fidelity, as indicated in Fig. 9(f). Therefore, the presented protocol is more robust to experimental imperfections and at the same time provides higher success probability in this case.

B. Detailed implementation analysis

Let us discuss in detail the theoretical model used to generate the predictions shown in Fig. 9 and explain the origin of the discrepancies between the ideal fidelities and the experimental realization.

We based the predictions on the Choi matrix of the central CNOT gate obtained by reconstructing the tomographic data. In linear-optical experiments, the entangling gate is typically the dominant source of imperfections. Our motivation for characterizing the CNOT gate was to compare the presented protocol with other similar protocols using entangling gates under comparable conditions, i.e., using the central gate of the same quality. All remaining imperfections were modeled as effective single-qubit unitary errors acting outside the entangling gate.

Before presenting the full model used to obtain the theoretical predictions, we first explain the measured Choi matrix of the CNOT gate using a physical model, in order to identify the dominant sources of imperfections at this stage. While discussing the individual imperfections, we also analyze their impact on the protocol in terms of the retrieved overall process fidelity for $n = 1$. To be specific, we average the fidelity over

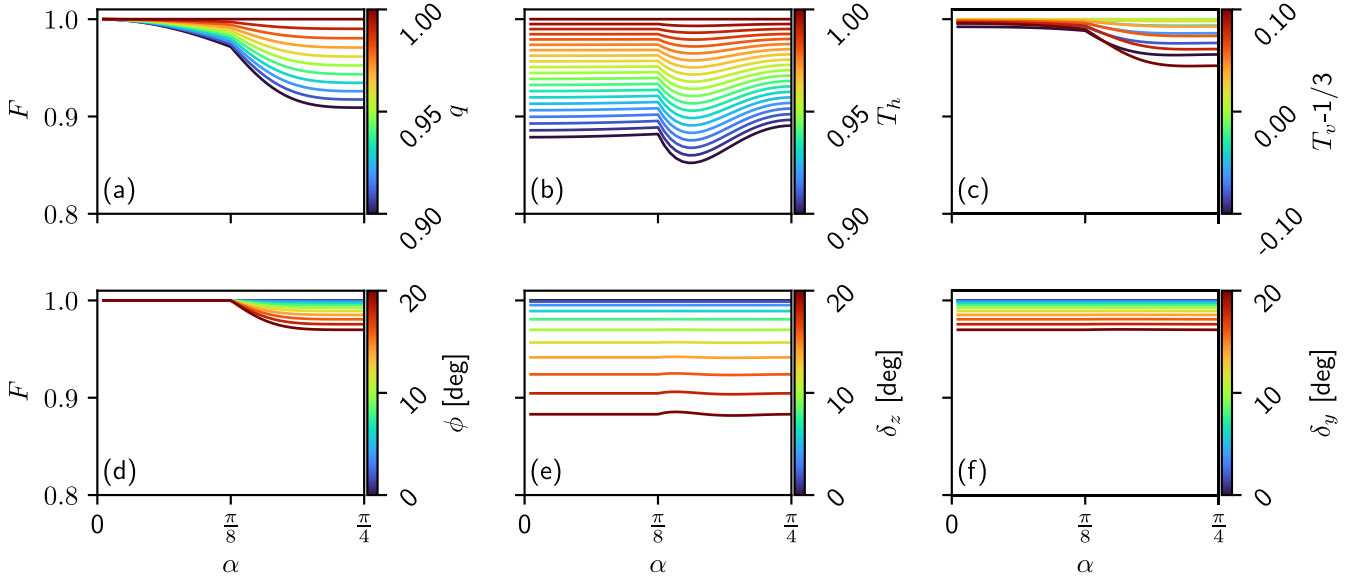


FIG. 10. Analysis of fidelity reduction due to the following individual experimental imperfections: partial distinguishability of the single photons (a), imperfect splitting ratio for horizontally (b) and vertically (c) polarized photons in the central PPBS, phase misalignment in the Mach-Zehnder interferometer (MZI) (d), and z rotation (e) and y rotation (f) of the program qubit. The severity of the imperfection is color coded. The phase shift and qubit rotations are shown in degrees. The graph shows the overall process fidelity of the retrieved operations averaged over both signs of α and both success cases.

the sign of the stored α ,

$$F = \frac{F_+ + F_-}{2},$$

where

$$F_{\pm} = \frac{1}{f_{0\pm} + f_{1\pm}} \sum_{i=0,1} \langle\langle U_{\pm} | C_{i\pm} | U_{\pm} \rangle\rangle f_{i\pm}$$

is the process fidelity of the retrieved operation relative to the ideal stored operation. This definition coincides with the fidelity metric used in Fig. 9.

The used linear-optical logic gate utilizes two-photon interference and, therefore, it is sensitive to the limited indistinguishability of the two interfering photons. This imperfection originates primarily in the SPDC source of photon pairs and is related to the spectral purity of the source and the phase matching in the nonlinear crystal. Additional degradation arises from imperfect spatial and temporal mode matching of the photons at the beam splitter. We describe this imperfection effectively by introducing the probability q of the photons being perfectly indistinguishable, and in the remaining case, with probability $1 - q$, we treat photons as perfectly distinguishable. As described in [67], this imperfection is an effective source of decoherence and leads to incoherent admixing of erroneous terms into the effective Choi matrix, reducing the purity and fidelity of the gate. The partial distinguishability is directly reflected in the visibility of the Hong-Ou-Mandel dip, which was 0.95 in our case. This level of partial distinguishability can readily explain up to 0.03 process infidelity near the threshold α_t , as apparent from the tolerance chart in Fig. 10(a), and is even more pronounced for higher α . With a state-of-the-art SPDC source, such as [68], one could achieve 0.99 visibility, theoretically suppressing the infidelity below the percent level. However, this imperfection

does not explain the loss of fidelity observed near $\alpha = 0$ in Fig. 9.

A critical component of our CNOT gate is the central PPBS, as it facilitates the two-photon interference. We therefore consider deviations of the splitting ratios from their nominal value for both polarizations. The parasitic coupling $T_h < 1$ of horizontally polarized photons causes uneven attenuation of two-qubit populations via destructive two-photon interference and moreover causes the CNOT gate to act partially as a swap gate. An imperfect splitting ratio $T_v \neq \frac{1}{3}$ for vertically polarized photons leads to unequal population amplitudes as well. The unwanted partial swapping has a significant impact on the retrieved fidelity, as shown in Fig. 10(b), and can readily explain the loss of fidelity even for small α . The comparison to Fig. 10(c) reveals that the protocol is more tolerant of deviation of the transmittance T_v . The splitting ratio deviations in the remaining auxiliary PPBSs further contribute to population imbalance, and we have included them into the model as well, but their impact is comparable to that shown in Fig. 10(c).

The performed quantum process tomography captures the entire experimental transformation, including the U_i block. To isolate the CNOT core, we include four single-qubit unitary operations in the model, acting on the individual input and output ports of the CNOT and treated as free fitting parameters. Using a fixed value $q = 0.95$, the fitted model reproduces the reconstructed Choi matrix with a process fidelity of 0.967. The dominant eigenkets of the compared Choi matrices match closely, with 10^{-3} infidelity. Therefore, the main source of mismatch must be of incoherent nature. The fitted transmittances are $T_h = 0.992$, $T_v = 0.597$, $T'_h = 0.603$, $T'_v = 1.0$, $T''_h = 0.584$, $T''_v = 1$, where dashed symbols relate to balancing PPBSs, approximately agree with the independently measured values and the small

discrepancies can be explained by path-dependent losses originating in the calcite interferometer in combination with fiber coupling. The remaining infidelity is likely due to systematic state-preparation and measurement errors, originating primarily from wave-plate retardance deviations, which act nontrivially throughout the tomographic sequence and effectively reduce the reconstructed process purity. Hypothetically increasing q to 0.98 would further improve the agreement between the model and the reconstructed Choi matrix, yielding a fidelity of 0.984.

Having established that the physical model captures the dominant imperfections of the central CNOT gate, we now consider unitary errors occurring outside the entangling operation.

Phase misalignment in the Mach-Zehnder interferometer implementing the discrimination POVM is irrelevant for $\alpha < \alpha_t$ because in this regime the protocol employs two-outcome discrimination and the Mach-Zehnder interferometer is not used, as described in Appendix E. For $\alpha > \alpha_t$, this imperfection reduces the fidelity by altering the relative probabilities of detection events at detectors D0 and D1, thereby introducing errors in state discrimination. This effect partly explains the sharp change in the predicted fidelity trend near α_t observed in Fig. 9.

The phase misalignment may originate from slow drifts of the interferometric phase over the duration of the experiment, resulting in an effective static phase offset. To mitigate these drifts, we split the measurement sequence into several parts and performed phase realignment between them. However, the precision of the phase reset is limited and introduces residual phase errors. By fitting the model, we estimate the effective phase misalignment to be 12.7° . The resulting impact on the retrieved process fidelity is shown in Fig. 10(d).

Finally, we include single-qubit rotations of the storage qubit in the model. We focus primarily on rotations about the z axis, as these have a stronger impact on the protocol due to the fact that the stored unitary operation is encoded in the relative phase. This enhanced sensitivity is evident from a comparison of the z - and y -rotation tolerances shown in Figs. 10(e) and 10(f). In contrast, rotations about the y axis have a comparatively weaker effect on the retrieved process fidelity.

The unwanted z rotation of the storage qubit most likely originates from imperfections in the wave plates used for state preparation and in those implementing the U_i block. Deviations of wave-plate retardances from their nominal values effectively result in additional unitary rotations of the qubit. In the theoretical predictions shown in Fig. 9, we use a z -rotation angle of -1.5° obtained by fitting the theoretical model to the data. Retardance deviations of the wave plates used in state-preparation and tomographic measurements also affect the reconstructed fidelities; however, their action is non-trivial, as the wave-plate settings change for each step of the tomographic sequence.

This type of error would become negligible in integrated photonic chips, which provide interferometric-based single-qubit gates with fidelities reaching 99.9% [69].

We also considered the finite extinction ratios of the polarization components used in the experiment. Since these

typically exceed $1:10^3$, their impact on the retrieved fidelity is negligible and was therefore not included in the model. Similarly, we did not model nonunit interferometric visibility of the Mach-Zehnder interferometer explicitly, as the visibility in our interferometer was 99%. The reduced visibility could be incorporated into the model by averaging over a distribution of interferometric phases. The second-order correlation function characterized the photon-number purity of our SPDC source. Depending on the pump power, it spanned from $g_2(0) = 2 \times 10^{-4}$ to 2×10^{-3} . Based on these values, we decided not to include multiphoton contributions in the model, as the infidelity to be explained is at least an order of magnitude higher.

IV. DISCUSSION

We studied storage and retrieval of unknown unitary transformation, which is chosen from a set of two known possibilities with equal prior probabilities and the retrieved transformation is expected to mimic well the action of the stored unitary on any input state. We proved that the optimal state for storage is the one which is also optimal for the discrimination of the two unitaries and we optimized the retrieval in two scenarios. If the goal is to maximize the average process fidelity for a deterministic strategy, the solution (for qubit unitaries) is of the kind measure-and-prepare (see Fig. 3) and the optimal value of the average fidelity reads as in Eq. (26). It is easy to observe that this result is also optimal if we want to retrieve $m \geq 1$ copies of the unknown unitary with the maximal single-copy fidelity.

We also found the optimal probability of success in the case of a perfect probabilistic retrieval in arbitrary dimension. The result is given by Eq. (22) and we observe that for small values of α the success probability approaches unity as $1/n^2$ with a quadratic improvement over known storage and retrieval strategies, which work for continuous subsets of transformations.

The retrieval phase can be also viewed as a programmable quantum processor since the state resulting from the storage phase programs (encodes) the transformation to be performed on another system. In a suitable basis, the program state encodes a phase shift of $2n\alpha$ and the processor should induce a unitary transformation (or its inverse) with relative phase shifts of at most 2α . Since in our optimization of retrieval operation parameter n can be treated as continuous we can think of this processor as optimally transforming program states with relative phase shift $\pm\beta$ (i.e., states with overlap $\cos \beta$) into two unitary transformations U_0, U_1 , whose relative unitary $U_1^\dagger U_0$ has angular spread of eigenvalues at most 4α . Thus, by performing the optimization of the retrieval we found two classes of optimal processors, which for any given overlap of input states and any choice of two unitary transformations achieve the best possible average process fidelity or best average probability of success, respectively. Since the angular spread is the only relevant parameter, our results easily generalize to the case in which, instead of retrieving U_0 or U_1 , the goal of the protocol is to invert (i.e., retrieving $U_i^\dagger$), conjugate (i.e., retrieving U_i^*), or transpose (i.e., retrieving U_i^T) the unknown unitary. For approximate deterministic case our results hold for qubit unitary transformations, whereas in the perfect

probabilistic case they hold for any pair of finite-dimensional unitary transformations.

Let us discuss implications of our results for perfect probabilistic processors programming N unitary transformations. Having such processor implies that one can use any pair of its program states to program the corresponding two unitary transformations. This sets an upper bound on the success probability for any such chosen pair implied by our results on programming just two unitary transformations.

For qubit unitary transformations we found a quantum circuit performing the optimal perfect probabilistic storage and retrieval using just a single CNOT gate and a suitable unambiguous measurement on its target qubit (see Fig. 7). For quite distinct unitary transformations (large- α regime) two of the measurement outcomes correspond to success, but for one of them unitary (σ_Z gate) correction must be performed on the unmeasured qubit. For almost identical unitary transformations that should be stored (small- α regime) the measurement is projective with just success (no correction) and failure results.

We built a linear-optical experiment implementing the presented optimal circuit. The setup is based on the CNOT gate consisting of partially polarizing beam splitters (PPBS) and utilizing the two-photon interference [50–52,70]. The unambiguous measurement part consists of a Mach-Zehnder-type interferometer made with polarizing beam splitters and wave plates in its arms. Tomographically complete data were recorded and reconstructed using the maximum-likelihood method. The setup was also used to test nonoptimal measure-and-prepare strategy for the perfect retrieval. The success probability of both strategies quite closely follows the theoretically predicted curves. Full tomographic data allowed us also to calculate process fidelity of the recreated transformations. Naturally, measure-and-prepare strategy achieves better fidelities as it uses less multipartite operations, actually only those needed to perform the unambiguous discrimination measurement. Optimal perfect retrieval uses also the CNOT gate, which due to its 93% fidelity of realization is the main cause of the observed infidelities (especially in the small- α regime). Perfect retrieval outperforms the measure-and-prepare strategy especially from the success probability point of view. For $n > 1$ this holds for almost any pair of unitaries (i.e., the whole interval of α parameter). For $n = 1$ and small- α regime the perfect retrieval coincides with probabilistic programmable implementation of phase gates proposed in Ref. [71]. On the other hand, Ref. [38] shows that two applications of this gate implement $n = 3$ storage and retrieval of arbitrary qubit phase gate with probability $\frac{3}{4}$. The reported experimental setup outperforms such strategy both in success probability and in achieved process fidelity if we assume the same quality of CNOT gate as in our experiment.

Just as understanding tasks related to state nonorthogonality offers insight into various quantum phenomena [8], our results shed light on the ultimate limits quantum mechanics imposes on manipulating transformations. Some technical questions are still open, such as search for an efficient quantum circuit for storage and retrieval of qudit unitary transformations in the perfect probabilistic case or optimization of approximate storage and retrieval for qudits.

Our preliminary numerical semidefinite programming investigation of the approximate storage and retrieval for qudit unitary transformations revealed that the optimal process fidelity is no longer solely dependent on the angular spread of the eigenvalues of the relative unitary $U_1^\dagger U_0$, but it has also nontrivial dependence on the angle of the other eigenvalue as well. This illustrates that approximate deterministic version of the storage and retrieval problem for two qudit ($d \geq 3$) unitary transformations has different qualitative and quantitative properties, whose proper investigation goes beyond the scope of this paper.

Our future research could also cover cloning of two unitary transformations and other tasks following the spirit of the above ideas.

ACKNOWLEDGMENTS

A.B. acknowledges financial support from the European Union-Next Generation E.U. through the PNRR MUR Project No. PE0000023-NQSTI. R.S. and M.M. acknowledge the support from the Ministry of the Interior of the Czech Republic, project NU-CRYPT (Project No. VK01030193), and from Ministry of Education, Youth and Sports of the Czech Republic, Grant No. 8C22003 (QD-E-QKD) of the QuantERA II Programme that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101017733. J.F. acknowledges the support from Palacký University, IGA-PrF-2024-008. M.S. was supported by projects APVV-22-0570 (DeQHOST), VEGA 2/0183/21 (DESCOM). M.S. was further supported by funding from QuantERA, an ERA-Net cofund in Quantum Technologies, under the project eDICT.

A.B. and M.S. conceptualized the problem and performed the theoretical calculations. J.F. contributed to the theoretical derivation and supervised the experimental work. R.S., N.H., and M.M. constructed the experimental setup, conducted the measurements, and analyzed the acquired data. All authors contributed to the manuscript's writing.

DATA AVAILABILITY

The data that support the findings of this article are openly available on Zenodo at Ref. [72].

APPENDIX A: PROOF OF PROPOSITION 2

We will split the proof of Proposition 2 into several steps.

Reminding the decomposition $R = PRP + P'RP'$ of Lemma 2, we can write

$$\begin{aligned} PRP &= \sum_{i,j=1}^2 A_{i,j} |e_i\rangle \langle e_j|, \\ P'RP' &= \sum_{i,j=1}^2 B_{i,j} |e'_i\rangle \langle e'_j|. \end{aligned} \quad (\text{A1})$$

Lemma A1. $\text{Tr}_2[R] \leq I$ if and only if $A + B \leq I$, where A and B are the 2×2 matrices with coefficient $A_{i,j}$ and $B_{i,j}$ introduced in Eq. (A1).

Proof. A straightforward computation gives

$$\begin{aligned}\text{Tr}_2[|e_1\rangle\langle e_1|] &= |+\rangle\langle +| \otimes I, \\ \text{Tr}_2[|e_1\rangle\langle e_2|] &= |+\rangle\langle -| \otimes \sigma_z, \\ \text{Tr}_2[|e_2\rangle\langle e_1|] &= |-\rangle\langle +| \otimes \sigma_z, \\ \text{Tr}_2[|e_2\rangle\langle e_2|] &= |-\rangle\langle -| \otimes I, \\ \text{Tr}_2[PRP] &= \tilde{A} \otimes |0\rangle\langle 0| + \sigma_x \tilde{A} \sigma_x \otimes |1\rangle\langle 1|, \\ \text{Tr}_2[P'RP'] &= \tilde{B} \otimes |0\rangle\langle 0| + \sigma_x \tilde{B} \sigma_x \otimes |1\rangle\langle 1|,\end{aligned}$$

where

$$\tilde{X} = X_{1,1}|+\rangle\langle +| + X_{1,2}|+\rangle\langle -| + X_{2,1}|-\rangle\langle +| + X_{2,2}|-\rangle\langle -|.$$

Then we have

$$\text{Tr}_2[R] = (\tilde{A} + \tilde{B}) \otimes |0\rangle\langle 0| + \sigma_x (\tilde{A} + \tilde{B}) \sigma_x \otimes |1\rangle\langle 1|$$

and we have that $\text{Tr}_2[R] \leq I_{01}$ if and only if $\tilde{A} + \tilde{B} \leq I$, which is equivalent with $A + B \leq I$. ■

Lemma A2. Figure of merit from Eq. (7) can be expressed as $\text{Tr}[RD] = \langle u|A|u\rangle + \langle v|B|v\rangle$.

Proof. It is easy to verify that

$$\begin{aligned}|\psi_{n,0}^*\rangle &= c_n|+\rangle - is_n|-\rangle, \quad |U_0\rangle = c|I\rangle + is|\sigma_z\rangle, \\ |\psi_{n,1}^*\rangle &= \sigma_x|\psi_{n,0}^*\rangle, \quad |U_1\rangle = \sigma_x \otimes \sigma_x|U_0\rangle,\end{aligned}\quad (\text{A2})$$

which implies that

$$\begin{aligned}|\psi_{n,0}^*|U_0\rangle &= (c_n c|e_1\rangle + s_n s|e_2\rangle) + i(c_n s|e_1'\rangle - s_n c|e_2'\rangle), \\ |\psi_{n,1}^*|U_1\rangle &= \sigma_x^{\otimes 3}|\psi_{n,0}^*|U_0\rangle.\end{aligned}\quad (\text{A3})$$

By substituting Eq. (A3) into (5) and using (A1) we obtain the thesis. ■

Lemma A3. The perfect retrieval condition of Eq. (17) is equivalent to $\langle v|A|v\rangle = \langle u|B|u\rangle = 0$.

Proof. Via direct computation using Eqs. (16), (A1), and (A2) we can rewrite Eq. (17) as

$$\begin{aligned}A &= \lambda_A |\phi_A\rangle\langle \phi_A|, \quad B = \lambda_B |\phi_B\rangle\langle \phi_B|, \\ |\phi_A\rangle &:= \begin{pmatrix} \frac{c}{s_n} \\ \frac{c_n}{s} \end{pmatrix}, \quad |\phi_B\rangle := \begin{pmatrix} \frac{s}{c_n} \\ -\frac{c}{s_n} \end{pmatrix}.\end{aligned}\quad (\text{A4})$$

We notice that $\langle u|\phi_B\rangle = \langle v|\phi_A\rangle = 0$ and since we are in a two-dimensional space, we have the thesis. ■

Combining Lemma A1 with Lemma A2 yields to Eq. (18). Combining Lemmas A1, A2, and A3 yields to Eq. (19).

APPENDIX B: REALIZATION OF THE OPTIMAL APPROXIMATE DETERMINISTIC RETRIEVAL

The optimal operators A, B from Eq. (18) are one-dimensional orthogonal projectors defining a projective measurement in basis

$$\begin{aligned}A &= |\phi_A\rangle\langle \phi_A|, \quad B = |\phi_B\rangle\langle \phi_B|, \\ |\phi_A\rangle &:= \begin{pmatrix} a \\ b \end{pmatrix}, \quad |\phi_B\rangle := \begin{pmatrix} -b^* \\ a^* \end{pmatrix},\end{aligned}\quad (\text{B1})$$

where $|a|^2 + |b|^2 = 1$ and a, b are suitable real numbers. Note that the figure of merit (18) can be written as $F_e = \eta_u \langle \tilde{u}|A|\tilde{u}\rangle + \eta_v \langle \tilde{v}|B|\tilde{v}\rangle$, where $|\tilde{u}\rangle := \frac{1}{\sqrt{\eta_u}}|u\rangle$, $|\tilde{v}\rangle := \frac{1}{\sqrt{\eta_v}}|v\rangle$ are normalized pure states and $\eta_u + \eta_v = \langle u|u\rangle + \langle v|v\rangle = 1$.

Due to discussion below Eq. (7) we know $A + B = I$. Consequently, the problem is actually a minimum error discrimination problem for pure states $|\tilde{u}\rangle, |\tilde{v}\rangle$ appearing with prior probability η_u, η_v , respectively. Then, the parameters a and b can be determined by the solution of this state-discrimination problem but their particular values are not important for the upcoming discussion. However, to show that realization scheme based on vectors $|\phi_A\rangle, |\phi_B\rangle$ has the same optimal performance it is useful to rewrite figure of merit in terms of parameters a, b , which reads as follows:

$$F = (a c_n c + b s_n s)^2 + (a s_n c + b c_n s)^2. \quad (\text{B2})$$

The matrices A, B completely determine the Choi-Jamiołkowski operator of $\mathcal{R}$, which has rank two. Thus, any dilation of $\mathcal{R}$ must have at least two-dimensional ancilla. One of such minimal dilations is defined by the following unitary transformation:

$$\begin{aligned}G &= (|a_1\rangle\langle +| + |a_2\rangle\langle -|) \otimes |0\rangle\langle 0| \\ &\quad + (|b_1\rangle\langle +| + |b_2\rangle\langle -|) \otimes |1\rangle\langle 1|,\end{aligned}\quad (\text{B3})$$

where

$$\begin{aligned}|a_1\rangle &= a|0\rangle - b^*|1\rangle, \quad |a_2\rangle = b|0\rangle + a^*|1\rangle, \\ |b_1\rangle &= \sigma_z|a_1\rangle, \quad |b_2\rangle = -\sigma_z|a_2\rangle.\end{aligned}\quad (\text{B4})$$

We can write G as

$$G = C_\pi (M \otimes I) \text{CNOT}, \quad (\text{B5})$$

where $\text{CNOT} = I \otimes |0\rangle\langle 0| + \sigma_x \otimes |1\rangle\langle 1|$ is the controlled-NOT gate, $M = |a_1\rangle\langle +| + |a_2\rangle\langle -|$ is qubit unitary transformation, and $C_\pi = I \otimes |0\rangle\langle 0| + \sigma_z \otimes |1\rangle\langle 1|$ is a controlled phase gate. The actual effect of C_π when it is followed by the measurement in the basis $\{|0\rangle_A, |1\rangle_A\}$ of the first ancillary qubit can be equivalently achieved by omitting C_π and taking no action or application of σ_z gate on the second qubit if the measurement outcome was 0 or 1, respectively.

In order to prove optimality of the proposed scheme, we first express the conditionally prepared unitary transformations

$$\begin{aligned}\langle 0| \otimes I G |\psi_{in,0}\rangle \otimes I &= a c_n I + i b s_n \sigma_z =: \sqrt{p} U_{c,00}, \\ \langle 1| \otimes I G |\psi_{in,0}\rangle \otimes I &= i a^* s_n - b^* c_n \sigma_z =: \sqrt{1-p} U_{c,01}, \\ \langle 0| \otimes I G |\psi_{in,1}\rangle \otimes I &= a c_n I - i b s_n \sigma_z =: \sqrt{p} U_{c,10}, \\ \langle 1| \otimes I G |\psi_{in,1}\rangle \otimes I &= -i a^* s_n - b^* c_n \sigma_z =: \sqrt{1-p} U_{c,11},\end{aligned}\quad (\text{B6})$$

where $U_{c,ab}$ are qubit unitary transformations and $p = a^2 c_n^2 + b^2 s_n^2$. As a consequence the Choi operators of the prepared channels in case of the memory states $|\psi_{n,0}\rangle, |\psi_{n,1}\rangle$ are

$$\begin{aligned}X_0 &= p|U_{c,00}\rangle\langle U_{c,00}| + (1-p)|U_{c,01}\rangle\langle U_{c,01}|, \\ X_1 &= p|U_{c,10}\rangle\langle U_{c,10}| + (1-p)|U_{c,11}\rangle\langle U_{c,11}|,\end{aligned}\quad (\text{B7})$$

respectively. Inserting the above expressions into the formula for the average process fidelity of the storage and retrieval we find that

$$\begin{aligned}F_e &= \frac{1}{8}(\langle U_0|X_0|U_0\rangle + \langle U_1|X_1|U_1\rangle) \\ &= (a c_n c + b s_n s)^2 + (a s_n c + b c_n s)^2,\end{aligned}\quad (\text{B8})$$

which coincides with the optimal value of the figure of merit [see Eq. (B2)]. On the other hand, unitary transformation G , defined in Eq. (B3), can be rewritten also in the following way:

$$G = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)(\uparrow| \otimes U_{+A} + \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)(\downarrow| \otimes U_{-A}, \quad (\text{B9})$$

where $U_{\pm A} = aI \pm ib\sigma_z$ are qubit unitary gates and $|\uparrow\rangle = \frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle)$, $|\downarrow\rangle = \frac{1}{\sqrt{2}}(|+\rangle - i|-\rangle)$. Since G is used for a dilation of a channel, it does not matter in which basis we perform the partial trace over the first qubit. Operationally it is equivalent to measuring in some basis [say $\frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$] and ignoring the measurement outcome. With this view we realize that application of G and measurement in basis $\frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$ can be seen as a realization of the following measure-and-prepare strategy for retrieval stage of the protocol. The input state $|\psi_{n,0}\rangle$ or $|\psi_{n,1}\rangle$ is first measured in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ and subsequently unitary transformation U_{+A} or U_{-A} is applied to the main system depending on the obtained outcome. Of course, the measurement could be already performed at the end of the storage phase of the protocol, thus, we found that for maximal process fidelity of deterministic storage and retrieval there exists a measure-and-prepare strategy (see illustration in Fig. 3) which does not require quantum memory and performs optimally.

APPENDIX C: PERFECT PROBABILISTIC RETRIEVAL FOR $d > 2$

In this Appendix we will prove that the value of the probability of success derived in Eq. (22) for the qubit case is also optimal for generic dimension d .

Let us denote with $P_{\text{succ}}(U_0, U_1)$ the optimal probability of success for the perfect probabilistic storage and retrieval of the pair $\{U_0, U_1\}$ of unitaries in dimension d .

Without loss of generality, let U_0 and U_1 be given as in Eq. (1). Let us consider the unitaries

$$\begin{aligned} \tilde{U}_0 &:= e^{i\alpha}|0\rangle\langle 0| + e^{-i\alpha}|d-1\rangle\langle d-1|, \\ \tilde{U}_1 &:= e^{-i\alpha}|0\rangle\langle 0| + e^{i\alpha}|d-1\rangle\langle d-1|. \end{aligned} \quad (\text{C1})$$

Qubit unitaries $\tilde{U}_0$ and $\tilde{U}_1$ are obtained by restricting the action of U_0 and U_1 on the subspace spanned by $|0\rangle$ and $|d-1\rangle$. Only this subspace is used for the optimal storage (see Proposition 1) in both considered situations (U_0 or U_1 vs $\tilde{U}_0$ or $\tilde{U}_1$), thus, the same two states $|\psi_{n,i}\rangle$, $i = 0, 1$, are obtained after the optimal storage. Moreover, the retrieval for U_0/U_1 can be easily modified to perform retrieval for $\tilde{U}_0/\tilde{U}_1$, thus, we conclude $P_{\text{succ}}(U_0, U_1) \leq P_{\text{succ}}(\tilde{U}_0, \tilde{U}_1) = P_{\text{succ}}$, where P_{succ} is given by Eq. (22).

We will now present a retrieval strategy that shows that $P_{\text{succ}}(U_0, U_1)$ achieves the upper bound given by the qubit case. The retrieval strategy is realized by composing an isometry $G: \mathcal{H}_1 \otimes \mathcal{H}_0 \mapsto \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \mathcal{H}_4$ and a projective measurement

$$\Pi_s = I \otimes |0\rangle\langle 0|, \quad \Pi_f = I \otimes |1\rangle\langle 1| \quad (\text{C2})$$

on system $\mathcal{H}_3 \otimes \mathcal{H}_4$ [$\dim(\mathcal{H}_4) = 2$], where $s(f)$ is the outcome corresponding to a successful (failed) retrieval. Let us

consider the following ansatz for the isometry G :

$$G|k\rangle|\psi_{n,i}\rangle = \sqrt{P_{\text{succ}}} e^{i\beta_{k,i}}|k\rangle|\phi_i\rangle|0\rangle + \sqrt{1 - P_{\text{succ}}}|k\rangle|\eta_{i,k}\rangle|1\rangle, \quad (\text{C3})$$

$$\begin{aligned} |\phi_i\rangle, |\eta_{i,k}\rangle &\in \mathcal{H}_3, \quad |0\rangle, |1\rangle \in \mathcal{H}_4, \\ \langle\phi_i|\phi_i\rangle &= \langle\eta_{i,k}|\eta_{i,k}\rangle = 1, \\ \langle\phi_0|\phi_1\rangle &= x, \quad \langle\eta_{0,k}|\eta_{1,k'}\rangle = y_{k,k'}, \\ x, y_{k,k'} &\in \mathbb{C}, \quad |x|, |y_{k,k'}| \leq 1, \end{aligned} \quad (\text{C4})$$

where $|k\rangle$ is the basis of eingestates of U_0 and U_1 , and we defined $\beta_{0,0} = \beta_{d-1,1} = \alpha$, $\beta_{d-1,0} = \beta_{0,1} = -\alpha$, $\beta_{k,0} = \beta_k = -\beta_{k,1}$ in accordance with Eq. (1). Since $\{|k\rangle|\psi_{n,i}\rangle\}$ is a basis of $\mathcal{H}_1 \otimes \mathcal{H}_0$, G is completely specified by Eq. (B3). It is clear that G and the POVM $\{\Pi_s, \Pi_f\}$ provide a perfect probabilistic retrieval of the pair $\{U_0, U_1\}$ with probability of success P_{succ} . We only need to verify that G is an isometry, and that is true if and only if all the scalar products are preserved. Clearly, $|k\rangle|\psi_{n,i}\rangle$ and $|k'\rangle|\psi_{n,j}\rangle$ are sent to orthogonal states for $k \neq k'$. Therefore, G is an isometry if and only if the following set of equations is satisfied:

$$\cos(2n\alpha) = P_{\text{succ}} e^{i(\beta_{k,1} - \beta_{k,0})} x + (1 - P_{\text{succ}}) y_{k,k} \quad \forall k, \quad (\text{C5})$$

$$\text{where } |x|, |y_{k,k}| \leq 1, \quad \text{and } -2\alpha \leq \beta_{k,1} - \beta_{k,0} \leq 2\alpha. \quad (\text{C6})$$

A rather lengthy but straightforward calculation shows that Eq. (C5) is solved by the following choice of parameters which achieve the same probability of success as the qubit case:

$$\begin{aligned} x &= \frac{\tilde{c}_n \tilde{c}}{1 - \tilde{c}_n \tilde{s}}, \quad y_{k,k} = \frac{1 - \tilde{c} \zeta_k}{\tilde{s}}, \\ P_{\text{succ}} &= 1 - \tilde{c}_n \tilde{s} \quad \text{if } \alpha \in [\chi_n, \frac{\pi}{4n}], \end{aligned} \quad (\text{C7})$$

$$\begin{aligned} x &= 1, \quad y_{k,k} = \frac{\tilde{c}_n - P_{\text{succ}} \zeta_k}{1 - P_{\text{succ}}}, \\ P_{\text{succ}} &= \frac{1 - \tilde{c}_n^2}{2(1 - \tilde{c}_n \tilde{c})} \quad \text{if } \alpha \in (0, \chi_n), \end{aligned} \quad (\text{C8})$$

$$\begin{aligned} \zeta_k &:= e^{i(\beta_{k,1} - \beta_{k,0})}, \quad \tilde{c}_n := \cos(2n\alpha), \\ \tilde{c} &:= \cos(2\alpha), \quad \tilde{s} := \sin(2\alpha). \end{aligned} \quad (\text{C9})$$

APPENDIX D: REALIZATION OF THE OPTIMAL PERFECT PROBABILISTIC RETRIEVAL

The goal of this Appendix is to show how the optimal quantum circuit was found and thus prove Proposition 3. We start from Eq. (A4). We notice that $\langle u|\phi_A\rangle = \langle v|\phi_B\rangle = 1$ and consequently from Eq. (19) we get

$$P_{\text{succ}} = \lambda_A + \lambda_B. \quad (\text{D1})$$

We rewrite perfect retrieval conditions from Eq. (19) as

$$\langle \tilde{v}|A|\tilde{v}\rangle = 0, \quad \langle \tilde{u}|B|\tilde{u}\rangle = 0, \quad (\text{D2})$$

where $|\tilde{u}\rangle := \frac{1}{\sqrt{\eta_u}}|u\rangle$, $|\tilde{v}\rangle := \frac{1}{\sqrt{\eta_v}}|v\rangle$ are normalized pure states. On the other hand, the figure of merit from Eq. (19)

reads as

$$P_{\text{succ}} = \eta_u \langle \tilde{u} | A | \tilde{u} \rangle + \eta_v \langle \tilde{v} | B | \tilde{v} \rangle, \quad (\text{D3})$$

where $\eta_u + \eta_v = \langle u | u \rangle + \langle v | v \rangle = 1$. Due to normalization condition $A + B \leq I$ (Lemma A1) operators A, B can be interpreted as elements of a 3-outcome POVM $\{A, B, I - A - B\}$, which performs unambiguous discrimination of pure states $|\tilde{u}\rangle, |\tilde{v}\rangle$ appearing with prior probability η_u, η_v , respectively. The figure of merit (D3) equals the success probability of the discrimination and its optimal solution is known [45]. Depending on the prior probability η_u and the overlap $\mu^2 = |\langle \tilde{u} | \tilde{v} \rangle|^2$ three regimes for the optimal measurement and success probability exist (see, e.g., [46], p. 672). However, for storage and retrieval of two unitaries only two of them will occur since $\eta_u = (1 + \cos 2n\alpha \cos 2\alpha)/2 \geq \frac{1}{2}$. In particular, we obtain Eq. (22) expressed in suitable form for the upcoming steps

$$P_{\text{succ}} = \begin{cases} 1 - \cos 2n\alpha \sin 2\alpha, & \eta_u \leq \frac{1}{1+\mu^2} \text{ (large } \alpha), \\ \frac{1 - (\cos 2n\alpha)^2}{2(1 - \cos 2n\alpha \cos 2\alpha)}, & \eta_u \geq \frac{1}{1+\mu^2} \text{ (small } \alpha), \end{cases} \quad (\text{D4})$$

where $\frac{1}{1+\mu^2} = \frac{1 - (\cos 2n\alpha)^2 (\cos 2\alpha)^2}{1 - (\cos 2n\alpha)^2 \cos 4\alpha}$.

In the large- α regime ($\eta_u \leq \frac{1}{1+\mu^2}$) the optimal values of λ_A, λ_B in Eq. (A4) and the form of the third POVM element read as

$$\lambda_A = \frac{1}{2}[1 + \tilde{c}_n(\tilde{c} - \tilde{s})], \quad \lambda_B = \frac{1}{2}[1 - \tilde{c}_n(\tilde{c} + \tilde{s})], \quad (\text{D5})$$

$$C = I - A - B = |\phi_C\rangle\langle\phi_C|,$$

$$|\phi_C\rangle = \begin{pmatrix} \sqrt{v_+} \\ -\sqrt{v_-} \end{pmatrix}, \quad (\text{D6})$$

which is in accordance with Eq. (31).

On the other hand, for $\eta_u \geq \frac{1}{1+\mu^2}$ the optimal solution turns into a projective measurement

$$\begin{aligned} A &= |\tilde{v}\rangle\langle\tilde{v}|, \quad B = 0, \\ C &= I - A - B = |\tilde{v}^\perp\rangle\langle\tilde{v}^\perp|, \\ |\tilde{v}\rangle &= \begin{pmatrix} \sqrt{a} \\ \sqrt{b} \end{pmatrix}, \quad |\tilde{v}^\perp\rangle = \begin{pmatrix} \sqrt{b} \\ -\sqrt{a} \end{pmatrix}, \end{aligned} \quad (\text{D7})$$

again in accordance with values of a, b defined in Eq. (31).

Here the POVMs are acting on the two-dimensional complex vector space on which matrices A, B act. Elements of matrices A, B define elements of the Choi operator R_s and matrix C serves as a guide for numerous possible completions of quantum operation $\mathcal{R}_s$ into a quantum channel. One can involve isometric dilation of such a channel and a suitable measurement on the ancillary system to realize quantum operation $\mathcal{R}_s$. Thus, in the regime of $\eta_u \leq \frac{1}{1+\mu^2}$ we introduce the following isometry G from $\mathcal{H}_0 \otimes \mathcal{H}_1$ to $\mathcal{H}_A \otimes \mathcal{H}_2$:

$$\begin{aligned} G &= \sqrt{\lambda_A} \left(\frac{c}{c_n} |0\rangle\langle+| \otimes I + \frac{s}{s_n} |0\rangle\langle-| \otimes \sigma_z \right) \\ &+ \sqrt{\lambda_B} \left(\frac{s}{c_n} |1\rangle\langle+| \otimes \sigma_z - \frac{c}{s_n} |1\rangle\langle-| \otimes I \right) \\ &+ \sqrt{v_+} |2\rangle\langle+| \otimes I - \sqrt{v_-} |2\rangle\langle-| \otimes \sigma_z, \end{aligned} \quad (\text{D8})$$

where $\mathcal{H}_A$ is a three-dimensional Hilbert space (qutrit) spanned by orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Using Eqs. (20) and (31) we can verify that $G^\dagger G = I$. Let us calculate the conditional transformations if $|0\rangle$ or $|1\rangle$ are observed on the ancilla system. We obtain

$$\begin{aligned} {}_A\langle 0| \otimes I_2 G |\psi_{n,0}\rangle \otimes I_1 &= \sqrt{\lambda_A}(cI + is\sigma_z) = \sqrt{\lambda_A}U_0, \\ {}_A\langle 0| \otimes I_2 G |\psi_{n,1}\rangle \otimes I_1 &= \sqrt{\lambda_A}(cI - is\sigma_z) = \sqrt{\lambda_A}U_1, \\ {}_A\langle 1| \otimes I_2 G |\psi_{n,0}\rangle \otimes I_1 &= \sqrt{\lambda_B}(s\sigma_z - icI) = -i\sqrt{\lambda_B}U_0, \\ {}_A\langle 1| \otimes I_2 G |\psi_{n,1}\rangle \otimes I_1 &= i\sqrt{\lambda_B}(s\sigma_z + icI) = i\sqrt{\lambda_B}U_1. \end{aligned} \quad (\text{D9})$$

We see that the desired unitary transformations were retrieved exactly since the global phase is irrelevant. The overall success probability reaches its optimal value, since it reads as

$$\begin{aligned} P_{\text{succ}} &= \frac{1}{2} \sum_{i=0,1} \text{Tr}_{A2}[G |\psi_{n,i}\rangle\langle\psi_{n,i}| \otimes \xi G^\dagger W_A \otimes I_2] \\ &= \lambda_A + \lambda_B = 1 - \tilde{c}_n \tilde{s}, \end{aligned} \quad (\text{D10})$$

where ξ is an arbitrary normalized input state to the retrieved transformation, and $W = |0\rangle\langle 0| + |1\rangle\langle 1|$. Similarly, in the regime of $\eta_u \geq \frac{1}{1+\mu^2}$ we define isometry

$$\begin{aligned} H &= a|0\rangle\langle+| \otimes I + b|0\rangle\langle-| \otimes \sigma_z \\ &+ b|2\rangle\langle+| \otimes I - a|2\rangle\langle-| \otimes \sigma_z \end{aligned} \quad (\text{D11})$$

from $\mathcal{H}_0 \otimes \mathcal{H}_1$ to $\mathcal{H}_A \otimes \mathcal{H}_2$, where a, b are defined in Eq. (31). One can check by direct calculation that $H^\dagger H = I$. The main difference with respect to Eq. (D8) is that in this regime one-dimensional subspace along vector $|1\rangle$ is not used, thus, only outcomes in directions of $|0\rangle$ and $|2\rangle$ will appear and they correspond to success and failure, respectively. Indeed we obtain

$$\begin{aligned} {}_A\langle 0| \otimes I_2 H |\psi_{n,0}\rangle \otimes I_1 &= \sqrt{q}(cI + is\sigma_z) = \sqrt{q}U_0, \\ {}_A\langle 0| \otimes I_2 H |\psi_{n,1}\rangle \otimes I_1 &= \sqrt{q}(cI - is\sigma_z) = \sqrt{q}U_1, \\ P_{\text{succ}} &= q = \frac{\tilde{s}_n^2}{2(1 - \tilde{c}\tilde{s}_n)}, \end{aligned} \quad (\text{D12})$$

which agrees with optimal value from Eq. (22).

Our next goal is to find an efficient quantum circuit realization of the optimal protocol. For that purpose we rewrite the isometries G and H in the following form:

$$\begin{aligned} G &= (|a_1\rangle\langle+| + |a_2\rangle\langle-|) \otimes |0\rangle\langle 0| \\ &+ (|b_1\rangle\langle+| + |b_2\rangle\langle-|) \otimes |1\rangle\langle 1|, \\ H &= (|a'_1\rangle\langle+| + |a'_2\rangle\langle-|) \otimes |0\rangle\langle 0| \\ &+ (|b'_1\rangle\langle+| + |b'_2\rangle\langle-|) \otimes |1\rangle\langle 1|, \end{aligned} \quad (\text{D13})$$

where

$$\begin{aligned} F &= |0\rangle\langle 0| - |1\rangle\langle 1| + |2\rangle\langle 2|, \\ |b_1\rangle &= F|a_1\rangle, \quad |b'_1\rangle = F|a'_1\rangle = |a'_1\rangle, \end{aligned} \quad (\text{D14})$$

$$|b_2\rangle = -F|a_2\rangle, \quad |b'_2\rangle = -F|a'_2\rangle = -|a'_2\rangle \quad (\text{D15})$$

and the vectors $|a_1\rangle, |a_2\rangle, |a'_1\rangle, |a'_2\rangle$ are as in Eq. (31).

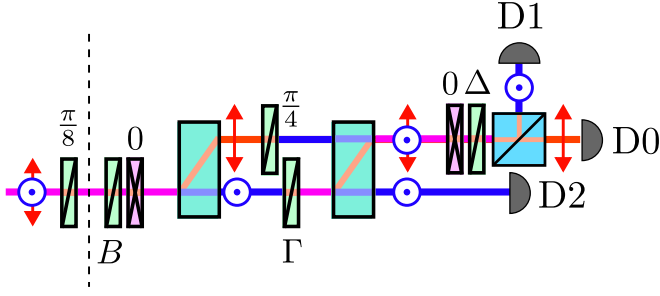


FIG. 11. Optical implementation of the POVM. The circuit implements isometry M followed by measurement in basis $|\uparrow H\rangle = |0_t\rangle$, $|\uparrow V\rangle = |1_t\rangle$, and $|\downarrow V\rangle = |2_t\rangle$. The first wave plate implements Hadamard operation, while the rest of the setup implements operator K . Polarization is denoted by arrow diagrams and colors, blue for vertical polarization and red for horizontal. The general superposition of both polarizations is indicated by the purple color. Half- and quarter-wave plates are depicted using green and purple color, respectively, and the angular orientation of their fast axis with respect to the horizontal direction is described by black captions. POVM measurement is concluded by detecting a photon at one of three single-photon detectors D0-D2.

Now it is obvious we can write G and H as

$$G = C_F M \otimes I C_X, \quad (D16)$$

$$H = M' \otimes I C_X, \quad (D17)$$

where $C_X = I \otimes |0\rangle\langle 0| + \sigma_x \otimes |1\rangle\langle 1|$ is a controlled-NOT (CNOT), $M = |a_1\rangle\langle +| + |a_2\rangle\langle -|$, and $M' = |a'_1\rangle\langle +| + |a'_2\rangle\langle -|$ are qubit-to-qutrit isometries and $C_F = I \otimes |0\rangle\langle 0| + F \otimes |1\rangle\langle 1|$ is a controlled version of unitary F applied to a qutrit. Finally, it is useful to realize what is the actual effect of C_F if it is followed by the measurement in the basis $\{|0\rangle, |1\rangle, |2\rangle\}$ of the qutrit. Subspaces $\{|00\rangle, |01\rangle, |20\rangle, |21\rangle\}$ are left unchanged, while in subspace $\{|10\rangle, |11\rangle\}$ we have $C_F|10\rangle = |10\rangle$, $C_F|11\rangle = -|11\rangle$, which is the same as action of σ_z on the qubit. Thus, instead of implementing C_F gate we might omit it and only in case of measurement outcome $|1\rangle$ on the qutrit we will apply σ_z on the unmeasured qubit to achieve the same resulting state as in the original circuit. In formula,

$${}_A\langle 0| \otimes I_2 G = {}_A\langle 0|M \otimes I C_X, \quad (D18)$$

$${}_A\langle 1| \otimes I_2 G = {}_A\langle 1|M \otimes \sigma_z C_X. \quad (D19)$$

In the regime of $\eta_u \gg \frac{1}{1+\mu^2}$ implementation of C_F gate is not needed and the circuit differs only in the choice of qubit to qutrit isometry. Thus, we proved that the quantum circuit depicted in Fig. 7 performs optimal storage and retrieval protocol for two unitary transformations of a qubit.

APPENDIX E: OPTICAL IMPLEMENTATION OF THE POVM

Here we describe our implementation of the POVM measurement in detail. The corresponding experimental block is depicted in Fig. 11. It consists of half-wave plates (depicted in green) and quarter-wave plates (depicted in purple) which

couple the horizontal $|H\rangle$, depicted in red, and vertical polarizations $|V\rangle$, depicted in blue. These two polarization modes initially occupy the same spatial mode. A calcite beam displacer separates the polarization modes spatially, displacing the horizontally polarized mode laterally by 6 mm, as depicted in the figure. The upper half-wave plate ensures that photons from the upper path always get to the detectors D0 and D1, while the bottom half-wave plate controls the probability of the photon in the bottom path reaching detector D2. The last stack of wave plates controls the splitting ratio between detectors D0 and D1. The black captions in the figure denote the angle between the fast axis of the wave plate and the horizontal direction. The subsequent click of single-photon detector D0, D1, or D2 heralds the detection of the respective POVM elements.

The mapping between input modes $|H\rangle = |0\rangle$, $|V\rangle = |1\rangle$ and output modes $|\uparrow H\rangle = |0_t\rangle$, $|\uparrow V\rangle = |1_t\rangle$, and $|\downarrow V\rangle = |2_t\rangle$ is described by the following relation:

$$C = \tilde{W}(\Delta, \pi) \tilde{W}\left(0, \frac{\pi}{2}\right) D_2 [|\uparrow\rangle\langle\uparrow| \otimes \sigma_x + |\downarrow\rangle\langle\downarrow| \otimes W(\Gamma, \pi)] D_1 W\left(0, \frac{\pi}{2}\right) W(B, \pi), \quad (E1)$$

where D_1 and D_2 describe the action of calcite crystals:

$$D_1 = |\uparrow H\rangle\langle H| + |\downarrow V\rangle\langle V|, \quad (E2)$$

$$D_2 = |\uparrow V\rangle\langle\uparrow V| + |\uparrow H\rangle\langle\downarrow H| + |\downarrow V\rangle\langle\downarrow V|. \quad (E3)$$

The operator

$$W(x, y) = |L_x\rangle\langle L_x| + \exp(-iy)|L_{x+\frac{\pi}{2}}\rangle\langle L_{x+\frac{\pi}{2}}| \quad (E4)$$

describes the action of a rotated wave plate using its eigenstates $|L_x\rangle = \cos(x)|H\rangle + \sin(x)|V\rangle$, x being the orientation of the fast axis and y the wave-plate retardance. The operator $\tilde{W}(x, y)$ describes polarization coupling in the upper spatial mode, i.e.,

$$\tilde{W}(x, y) = \begin{pmatrix} W(x, y) & 0 \\ 0 & 1 \end{pmatrix}. \quad (E5)$$

Note that due to the optimization of the experiment, there effectively is Hadamard gate H at the output of our CNOT gate. Therefore, to find the wave-plate angles B , Γ , and Δ for which the setup implements the isometry M we solve the matrix equation

$$C = M H^\dagger, \quad (E6)$$

where $M = M_l$ for $|\alpha| \leq \alpha_t$ and $M = M_h$ otherwise. The transition point α_t is determined by the following equation:

$$\cos(2n\alpha_t) \cos(2\alpha_t - \pi/4) = \sqrt{2}/2. \quad (E7)$$

The operator $M H^\dagger = K_{l,h}$ reads as

$$M_l H^\dagger = \begin{pmatrix} k_{11} & k_{31} \\ 0 & 0 \\ k_{31} & -k_{11} \end{pmatrix}, \quad (E8)$$

where

$$k_{11} = \sqrt{\frac{(1 + \cos 2\alpha)(1 - \cos 2n\alpha)}{2(1 - \cos 2\alpha \cos 2n\alpha)}}, \quad (E9)$$

$$k_{31} = \sqrt{\frac{(1 - \cos 2\alpha)(1 + \cos 2n\alpha)}{2(1 - \cos 2\alpha \cos 2n\alpha)}}, \quad (\text{E10})$$

and

$$M_h H^\dagger = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \\ k_{31} & k_{32} \end{pmatrix}, \quad (\text{E11})$$

where the elements are

$$k_{11} = \sqrt{\lambda_a} \frac{\cos(\alpha)}{\cos(n\alpha)}, \quad (\text{E12})$$

$$k_{12} = \sqrt{\lambda_a} \frac{\sin(\alpha)}{\sin(n\alpha)}, \quad (\text{E13})$$

$$k_{21} = \sqrt{\lambda_b} \frac{\sin(\alpha)}{\cos(n\alpha)}, \quad (\text{E14})$$

$$k_{22} = -\sqrt{\lambda_b} \frac{\cos(\alpha)}{\sin(n\alpha)}, \quad (\text{E15})$$

$$k_{31} = \sqrt{v_p}, \quad (\text{E16})$$

$$k_{32} = -\sqrt{v_p} \quad (\text{E17})$$

and

$$\lambda_a = [1 + \cos(2n\alpha)(\cos \alpha - \sin \alpha)]^{\frac{1}{2}}, \quad (\text{E18})$$

$$\lambda_b = [1 - \cos(2n\alpha)(\cos \alpha + \sin \alpha)]^{\frac{1}{2}}, \quad (\text{E19})$$

$$v_p = (1 - \cos^2 2\alpha + \sin 2\alpha) \frac{\cos 2n\alpha}{1 + \cos 2n\alpha}, \quad (\text{E20})$$

$$v_n = (-1 + \cos^2 2\alpha + \sin 2\alpha) \frac{\cos 2n\alpha}{1 - \cos 2n\alpha}. \quad (\text{E21})$$

A possible solution is

$$B = \frac{\pi}{4} + \frac{\beta}{4}, \quad (\text{E22})$$

$$\Gamma = \frac{\pi}{4} - \frac{\beta}{2}, \quad (\text{E23})$$

$$\Delta = \frac{\delta}{2}, \quad (\text{E24})$$

where

$$\beta = \arctan\left(\frac{k_{32}}{k_{31}}\right), \quad (\text{E25})$$

$$\gamma = \arcsin\left(\frac{-k_{31}}{\cos(\beta)}\right), \quad (\text{E26})$$

$$\delta = \arcsin\left(-\frac{k_{22} \cos(\gamma) + k_{11}}{\sin \beta \sin^2 \gamma}\right). \quad (\text{E27})$$

To implement isometry M , we apply a Hadamard gate on the input of this setup, i.e., $M = KH$. This Hadamard gate was already effectively present in the setup.

This optical setup could also be reconfigured to perform unambiguous state discrimination between states $[|0\rangle + \exp(\pm 2in\alpha)|1\rangle]/\sqrt{2}$ by setting wave-plate angles to

$$B = \frac{\pi}{4}, \quad (\text{E28})$$

$$\Gamma = \frac{1}{2} \arcsin[\tan(|\alpha|)], \quad (\text{E29})$$

$$\Delta = -\frac{\pi}{8}, \quad (\text{E30})$$

and the first quarter-wave plate to 0 and the second to $\frac{\pi}{4}$. Although the quarter-wave plates are not necessary for these operations, we included them in the experiment because the first two wave plates and the first calcite serve as a variable projector that we use for the tomographic characterization of the gate.

-
- [1] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, *Nature (London)* **299**, 802 (1982).
- [2] V. Bužek and M. Hillery, Quantum copying: Beyond the no-cloning theorem, *Phys. Rev. A* **54**, 1844 (1996).
- [3] R. F. Werner, Optimal cloning of pure states, *Phys. Rev. A* **58**, 1827 (1998).
- [4] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum cryptography without Bell's theorem, *Phys. Rev. Lett.* **68**, 557 (1992).
- [5] P. Busch, "No information without disturbance": Quantum limitations of measurement, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (Springer, Dordrecht, 2009), pp. 229–256.
- [6] H. Bechmann-Pasquinucci and N. Gisin, Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography, *Phys. Rev. A* **59**, 4238 (1999).
- [7] V. Bužek, M. Hillery, and R. F. Werner, Optimal manipulations with qubits: Universal-NOT gate, *Phys. Rev. A* **60**, R2626 (1999).
- [8] C. A. Fuchs, Just two nonorthogonal quantum states, in *Quantum Communication, Computing, and Measurement 2*, edited by P. Kumar, G. M. D'Ariano, and O. Hirota (Springer, Boston, MA, 2002), pp. 11–16.
- [9] P. Selinger, Towards a semantics for higher-order quantum computation, in *Proceedings of the 2nd International Workshop on Quantum Programming Languages*, TUCS General Publication, Vol. 33 (ACM Press, New York, 2004), pp. 127–143.
- [10] G. Gutoski and J. Watrous, Toward a general theory of quantum games, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, 2007), pp. 565–574.
- [11] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [12] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, *Phys. Rev. A* **80**, 022339 (2009).
- [13] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum networks: General theory and applications, *Acta Phys. Slovaca* **61**, 273 (2011).

- [14] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012).
- [15] M. Araújo, P. A. Guérin, and A. Baumeler, Quantum computation with indefinite causal structures, *Phys. Rev. A* **96**, 052315 (2017).
- [16] A. Bisio and P. Perinotti, Theoretical framework for higher-order quantum theory, *Proc. R. Soc. A* **475**, 20180706 (2019).
- [17] A. Acín, Statistical distinguishability between unitary operations, *Phys. Rev. Lett.* **87**, 177901 (2001).
- [18] R. Duan, Y. Feng, and M. Ying, Entanglement is not necessary for perfect discrimination between unitary operations, *Phys. Rev. Lett.* **98**, 100503 (2007).
- [19] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Adaptive versus nonadaptive strategies for quantum channel discrimination, *Phys. Rev. A* **81**, 032339 (2010).
- [20] Q. Zhuang and S. Pirandola, Ultimate limits for multiple quantum channel discrimination, *Phys. Rev. Lett.* **125**, 080505 (2020).
- [21] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Optimal cloning of unitary transformation, *Phys. Rev. Lett.* **101**, 180504 (2008).
- [22] G. Chiribella, Y. Yang, and C. Huang, Universal superreplication of unitary gates, *Phys. Rev. Lett.* **114**, 120504 (2015).
- [23] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, Information-disturbance tradeoff in estimating a unitary transformation, *Phys. Rev. A* **82**, 062305 (2010).
- [24] R. S. Abdul Salam, J. Shamsul Shaari, and S. Mancini, Information disturbance tradeoff in bidirectional QKD, *Phys. Scr.* **99**, 105135 (2024).
- [25] A. Bisio, G. M. D'Ariano, P. Perinotti, and G. Chiribella, Minimal computational-space implementation of multiround quantum protocols, *Phys. Rev. A* **83**, 022325 (2011).
- [26] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations, *Phys. Rev. Lett.* **123**, 210502 (2019).
- [27] S. Yoshida, A. Soeda, and M. Murao, Reversing unknown qubit-unitary operation, deterministically and exactly, *Phys. Rev. Lett.* **131**, 120602 (2023).
- [28] C. Zhu, Y. Mo, Y.-A. Chen, and X. Wang, Reversing unknown quantum processes via virtual combs for channels with limited information, *Phys. Rev. Lett.* **133**, 030801 (2024).
- [29] J. Miyazaki, A. Soeda, and M. Murao, Complex conjugation supermap of unitary quantum maps and its universal implementation protocol, *Phys. Rev. Res.* **1**, 013007 (2019).
- [30] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Optimal quantum learning of a unitary transformation, *Phys. Rev. A* **81**, 032324 (2010).
- [31] M. Sedláč, A. Bisio, and M. Ziman, Optimal probabilistic storage and retrieval of unitary channels, *Phys. Rev. Lett.* **122**, 170502 (2019).
- [32] P. Lewandowska, R. Kukulski, L. Pawela, and Z. Puchała, Storage and retrieval of von Neumann measurements, *Phys. Rev. A* **106**, 052423 (2022).
- [33] F. Grosshans, M. Horodecki, M. Murao, T. Młynik, M. T. Quintino, M. Studziński, and S. Yoshida, Multicopy quantum state teleportation with application to storage and retrieval of quantum programs, *Quantum* **10**, 2105 (2024).
- [34] P. Wittek, *Quantum Machine Learning: What Quantum Computing Means to Data Mining* (Academic Press, Cambridge, MA, 2014).
- [35] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, *Contemp. Phys.* **56**, 172 (2015).
- [36] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [37] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, Challenges and opportunities in quantum machine learning, *Nat. Comput. Sci.* **2**, 567 (2022).
- [38] M. Sedláč and M. Ziman, Probabilistic storage and retrieval of qubit phase gates, *Phys. Rev. A* **102**, 032618 (2020).
- [39] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory effects in quantum channel discrimination, *Phys. Rev. Lett.* **101**, 180501 (2008).
- [40] M. Raginsky, A fidelity measure for quantum channels, *Phys. Lett. A* **290**, 11 (2001).
- [41] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, *Phys. Lett. A* **303**, 249 (2002).
- [42] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**, 285 (1975).
- [43] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Math. Phys.* **3**, 275 (1972).
- [44] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer, New York, 2011), Vol. 1.
- [45] G. Jaeger and A. Shimony, Optimal distinction between two non-orthogonal quantum states, *Phys. Lett. A* **197**, 83 (1995).
- [46] M. Sedláč, Quantum theory of unambiguous measurements, *Acta Phys. Slovaca* **59**, 653 (2009).
- [47] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, *Nature (London)* **409**, 46 (2001).
- [48] Z.-Q. Zhou, W.-B. Lin, M. Yang, C.-F. Li, and G.-C. Guo, Realization of reliable solid-state quantum memory for photonic polarization qubit, *Phys. Rev. Lett.* **108**, 190505 (2012).
- [49] M. Gündoğan, P. M. Ledingham, A. Almasi, M. Cristiani, and H. de Riedmatten, Quantum storage of a photonic polarization qubit in a solid, *Phys. Rev. Lett.* **108**, 190504 (2012).
- [50] N. K. Langford, T. J. Weinhold, R. Prevedel, K. J. Resch, A. Gilchrist, J. L. O'Brien, G. J. Pryde, and A. G. White, Demonstration of a simple entangling optical gate and its use in Bell-state analysis, *Phys. Rev. Lett.* **95**, 210504 (2005).
- [51] N. Kiesel, C. Schmid, U. Weber, R. Ursin, and H. Weinfurter, Linear optics controlled-phase gate made simple, *Phys. Rev. Lett.* **95**, 210505 (2005).
- [52] R. Okamoto, H. F. Hofmann, S. Takeuchi, and K. Sasaki, Demonstration of an optical quantum controlled-NOT gate without path interference, *Phys. Rev. Lett.* **95**, 210506 (2005).
- [53] R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, Experimental demonstration of optimal unambiguous state discrimination, *Phys. Rev. A* **63**, 040305(R) (2001).
- [54] L. Stroh, N. Horová, R. Stárek, I. V. Puthoor, M. Mičuda, M. Dušek, and E. Andersson, Noninteractive XOR quantum oblivious transfer: Optimal protocols and their experimental implementations, *PRX Quantum* **4**, 020320 (2023).
- [55] R. Stárek, M. Miková, I. Straka, M. Dušek, M. Ježek, J. Fiurášek, and M. Mičuda, Experimental realization of swap operation on hyper-encoded qubits, *Opt. Express* **26**, 8443 (2018).

- [56] R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger, High-speed linear optics quantum computing using active feed-forward, *Nature (London)* **445**, 65 (2007).
- [57] G. Vallone, E. Pomarico, F. De Martini, and P. Mataloni, Active one-way quantum computation with two-photon four-qubit cluster states, *Phys. Rev. Lett.* **100**, 160502 (2008).
- [58] X.-S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin, and A. Zeilinger, Quantum teleportation over 143 kilometres using active feed-forward, *Nature (London)* **489**, 269 (2012).
- [59] T.-M. Zhao, H. Zhang, J. Yang, Z.-R. Sang, X. Jiang, X.-H. Bao, and J.-W. Pan, Entangling different-color photons via time-resolved measurement and active feed forward, *Phys. Rev. Lett.* **112**, 103602 (2014).
- [60] J. Sabines-Chesterking, R. Whittaker, S. K. Joshi, P. M. Birchall, P. A. Moreau, A. McMillan, H. V. Cable, J. L. O'Brien, J. G. Rarity, and J. C. F. Matthews, Sub-shot-noise transmission measurement enabled by active feed-forward of heralded single photons, *Phys. Rev. Appl.* **8**, 014016 (2017).
- [61] I. Agresti, D. Poderini, L. Guerini, M. Mancusi, G. Carvacho, L. Aolita, D. Cavalcanti, R. Chaves, and F. Sciarrino, Experimental device-independent certified randomness generation with an instrumental causal structure, *Commun. Phys.* **3**, 110 (2020).
- [62] G. L. Zanin, M. J. Jacquet, M. Spagnolo, P. Schiansky, I. A. Calafell, L. A. Rozema, and P. Walther, Fiber-compatible photonic feed-forward with 99% fidelity, *Opt. Express* **29**, 3425 (2021).
- [63] E. Meyer-Scott, N. Prasannan, I. Dhand, C. Eigner, V. Quiring, S. Barkhofen, B. Brecht, M. B. Plenio, and C. Silberhorn, Scalable generation of multiphoton entangled states by active feed-forward and multiplexing, *Phys. Rev. Lett.* **129**, 150501 (2022).
- [64] N. Horová, R. Stárek, M. Mičuda, M. Kolář, J. Fiurášek, and R. Filip, Deterministic controlled enhancement of local quantum coherence, *Sci. Rep.* **12**, 22455 (2022).
- [65] J. Fiurášek and Z. Hradil, Maximum-likelihood estimation of quantum processes, *Phys. Rev. A* **63**, 020101(R) (2001).
- [66] M. Mičuda, D. Koutný, M. Miková, I. Straka, M. Ježek, and L. Mišta, Experimental demonstration of a fully inseparable quantum state with nonlocalizable entanglement, *Sci. Rep.* **7**, 45045 (2017).
- [67] M. Mičuda, M. Miková, I. Straka, M. Sedlák, M. Dušek, M. Ježek, and J. Fiurášek, Tomographic characterization of a linear optical quantum Toffoli gate, *Phys. Rev. A* **92**, 032312 (2015).
- [68] N. Somaschi, V. Giesz, L. De Santis, J. C. Loredó, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Antón, J. Demory, C. Gómez, I. Sagnes, N. D. Lanzillotti-Kimura, A. Lemaître, A. Auffèves, A. G. White, L. Lanco, and P. Senellart, Near-optimal single-photon sources in the solid state, *Nat. Photon.* **10**, 340 (2016).
- [69] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Universal linear optics, *Science* **349**, 711 (2015).
- [70] M. Mičuda, M. Sedlák, I. Straka, M. Miková, M. Dušek, M. Ježek, and J. Fiurášek, Process-fidelity estimation of a linear optical quantum-controlled-Z gate: A comparative study, *Phys. Rev. A* **89**, 042304 (2014).
- [71] G. Vidal, L. Masanes, and J. I. Cirac, Storing quantum dynamics in quantum states: A stochastic programmable gate, *Phys. Rev. Lett.* **88**, 047905 (2002).
- [72] M. Sedlák, R. Stárek, N. Horová, M. Mičuda, J. Fiurášek, and A. Bisio, Storage and retrieval of two unknown unitary channels [Data set], Zenodo, 2026, <https://doi.org/10.5281/zenodo.21626231>.