

## Optimal Cloning of Unitary Transformation

Giulio Chiribella,<sup>\*</sup> Giacomo Mauro D'Ariano,<sup>+</sup> and Paolo Perinotti<sup>‡</sup>

*QUIT Group, Dipartimento di Fisica "A. Volta" and INFN Sezione di Pavia, via Bassi 6, 27100 Pavia, Italy<sup>§</sup>*

(Received 7 April 2008; published 30 October 2008)

After proving a general no-cloning theorem for black boxes, we derive the optimal universal cloning of unitary transformations, from one to two copies. The optimal cloner is realized by quantum channels with memory, and greatly outperforms the optimal measure-and-reprepare cloning strategy. Applications are outlined, including two-way quantum cryptographic protocols.

DOI: 10.1103/PhysRevLett.101.180504

PACS numbers: 03.67.Ac, 03.65.Ta, 03.67.Dd

The no-cloning theorem [1] is one of the cornerstones of quantum information, at the basis of the security of quantum cryptography [2], and challenging various protocols, from optimal estimation to error correction. Despite the long-dated attention to cloning of quantum states, cloning of quantum transformations is still a completely unexplored topic. Cloning a transformation  $\mathcal{T}$  means exploiting a single use of  $\mathcal{T}$  inside a quantum circuit, which thus performs the transformation  $\mathcal{T} \otimes \mathcal{T}$  on bipartite states. This elementary copying task is particularly relevant to the recent trend in quantum information, with the role of the information carrier played more and more by transformations instead of states, e.g., in gate discrimination [3,4], programming [5], teleportation [6–8], and tomography [9,10], along with multiround games [11] and cryptographic protocols like bit commitment with anonymous-state encoding [12,13]. Here cloning is the great absent of the list, partly due to the intrinsic difficulty in treating manipulations of transformations instead of states. Such a difficulty has been overcome by the method of Ref. [14], which allows one to optimize tasks where the input and output are transformations.

Cloning quantum transformations can be used for copying quantum software with a limited number of uses, and in other informational contexts. An interesting application is in the security analysis of multiround cryptographic protocols with encoding on secret transformations. Consider, for example, the following alternative version of the BB84 cryptographic protocol [15], where Alice uses two orthogonal bases of unitary transformations instead of states,  $B_1 = \{\sigma_\mu\}$  and  $B_2 = \{U\sigma_\mu\}$ , where  $\mu = 0, 1, 2, 3$ ,  $\sigma_0$  is the identity,  $\sigma_{1,2,3}$  are the three Pauli matrices, and  $U = (I + i \sum_{k=1}^3 \sigma_k)/2$  is a rotation of  $2\pi/3$  around the axis  $\mathbf{n} = (1, 1, 1)/\sqrt{3}$ . The protocol works as follows: Bob prepares an arbitrary maximally entangled state  $|B\rangle$  of two qubits and sends half of it to Alice, who applies one of the unitaries and sends the output to Bob, as in dense coding [16]. Afterwards, Bob measures the two qubits either on the Bell basis  $\{(\sigma_\mu \otimes I)|B\rangle\}$  or on the rotated basis  $\{(U\sigma_\mu \otimes I)|B\rangle\}$ , which are mutually unbiased for any maximally entangled state  $|B\rangle$ . After publicly announcing their choice of bases, and discarding cases where the bases were different, Alice and Bob use the values of index  $\mu$  to

establish a secret key. In this protocol, a naïve eavesdropping strategy would be for Eve trying to estimate the unitaries, by swapping the qubit sent by Bob with a qubit prepared by her—e.g., in a known maximally entangled state  $|E\rangle$ , half of which she keeps for herself—and then intercepting the qubit sent back by Alice. To prevent this attack Alice can randomly ask Bob to send his half of the entangled state, and to later reveal  $|B\rangle$ , so that she can check whether the received state was actually from Bob. However, Eve can perform coherent attacks that are much more efficient than naïve estimation. Among coherent attacks, the first and most natural to investigate is quantum cloning.

Differently from pure states, cloning of transformations is impossible even classically. This is a consequence of a general no-cloning theorem, holding not only for states, but also for transformations and any other kind of black boxes (e.g., measuring devices). Denoting by  $p$  the minimum of the worst case error probability in discriminating between two black boxes  $\mathcal{O}_1$  and  $\mathcal{O}_2$ , we have the following theorem, containing as a special case the no-cloning of quantum states [1]: two black boxes cannot be perfectly cloned by a single use unless  $p = 0$  (perfect discrimination) or  $p = 1/2$  (random guess, no discrimination at all). The proof is simple: If perfect cloning is possible, we can get three copies, perform three times the minimum error discrimination, and use majority voting to decide the most likely between  $\mathcal{O}_1$  and  $\mathcal{O}_2$  with worst case error probability  $p' = p^2(3 - 2p)$ . Since  $p$  is the minimum error probability, it must be  $p \leq p'$ , whose acceptable solutions are only  $p = 0$  and  $p = 1/2$  [17]. Application of the theorem to many black boxes  $\{\mathcal{O}_i\}_{i=1,\dots,k}$  yields the following cloning-discrimination equivalence: if for any  $i, j$  the error probability  $p_{ij}$  is not  $1/2$ , then perfect cloning is possible iff perfect discrimination is possible [18]. As a consequence, classical transformations, e.g., permutations of a classical register, cannot be cloned by a single use (there is no way to discriminate arbitrary permutations of the letters  $\{a, b, c\}$  by evaluation on a single letter). Likewise, quantum transformations, e.g., unitary gates, cannot be cloned by a single use (there is no way to discriminate arbitrary gates by a single use [3]).

The existence of a no-cloning theorem immediately raises the question about the performances of optimal cloners. In addition to possible cryptographic applications, the problem has a fundamental interest in itself, as the relation (if any) between optimal cloning of transformations and cloning of states is not *a priori* obvious.

In this Letter we derive the optimal universal cloner, which produces two approximate copies of a completely unknown unitary gate in dimension  $d < \infty$ , showing that entanglement with a quantum memory allows one to outperform any classical cloning strategy. For qubits the global channel fidelity of the clones is  $F_{\text{clon}} = 46.65\%$ , significantly larger than the fidelity of the optimal measure-and-prepare scheme  $F_{\text{est}} = 31\%$ , and of the random guess (using the given unitary on the first system, and performing a random unitary on the second)  $F_{\text{ran}} = 25\%$ . Surprisingly, cloning of unitary gates has no relation with cloning of maximally entangled states, in spite of the two sets being commonly considered as equivalent. Not only cloning maximally entangled states is always a suboptimal step for cloning unitary gates, but also any other scheme involving application of the unknown gate to a maximally entangled state (or any other fixed state) is necessarily suboptimal. As it will be shown, this also highlights a fundamental difference between the two tasks of cloning and learning quantum transformations.

The derivation of the optimal cloner exploits the recent toolbox of quantum circuit architecture theory [14], which allows optimization of quantum networks for any possible manipulation of quantum channels, including cloning and estimation. In this framework any channel  $\mathcal{C}$  from  $\mathcal{S}(\mathcal{H}_{\text{in}})$  to  $\mathcal{S}(\mathcal{H}_{\text{out}})$  [ $\mathcal{S}(\mathcal{H})$  denoting states on  $\mathcal{H}$ ] is described by its Choi operator  $C = \mathcal{C} \otimes I(|I\rangle\langle I|)$ , where  $|I\rangle = \sum_{i=1}^d |i\rangle|i\rangle$  is an unnormalized maximally entangled vector in  $\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{in}}$ . For a unitary channel  $\mathcal{U}(\rho) = U\rho U^\dagger$ , the Choi operator is  $|U\rangle\langle U|$ , with  $|U\rangle = (U \otimes I)|I\rangle$ . A quantum network for  $N$ -to- $M$  cloning is a network with  $N$  open slots in which the  $N$  input copies are inserted, and is also described by a suitable Choi operator  $R^{(N)}$  ( $N = 1$  for one-to-two cloning, see Fig. 1). If the Hilbert spaces of the inputs are labeled with even numbers from 0 to  $2N$ , and the output spaces with odd numbers from 1 to  $2N + 1$ , the Choi operator  $R^{(N)}$  is a non-negative operator on the tensor product  $\bigotimes_{k=0}^{2N+1} \mathcal{H}_k$  satisfying the recursive normalization condition

$$\text{Tr}_{2N+1}[R^{(N)}] = I_{2N} \otimes R^{(N-1)}, \quad (1)$$

where  $\text{Tr}_{2N+1}$  denotes the partial trace over the Hilbert space  $\mathcal{H}_{2N+1}$  of the  $N + 1$ th output system, and  $R^{(N-1)}$  is the Choi operator of a network with  $N - 1$  open slots, which in turn satisfies Eq. (1) with  $N$  replaced by  $N - 1$ . A network with  $N = 0$  open slots is a quantum channel from  $\mathcal{S}(\mathcal{H}_0)$  to  $\mathcal{S}(\mathcal{H}_1)$ , and has the normalization  $\text{Tr}_1[R^{(0)}] = I_0$ . Inserting  $N$  channels  $C_1, \dots, C_N$  in the  $N$  slots of a network, we obtain a new channel  $\mathcal{C}'$  from  $\mathcal{S}(\mathcal{H}_0)$  to  $\mathcal{S}(\mathcal{H}_{2N+1})$ , with Choi operator given by [14]

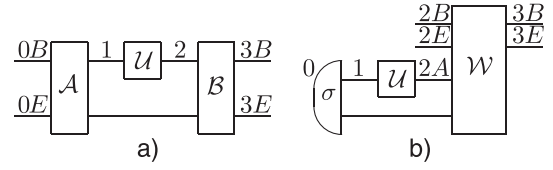


FIG. 1. (a) One-to-two cloning of unitaries. Two input systems are first processed by Eve with channel  $\mathcal{A}$ , which entangles system 1 and a quantum memory  $\mathcal{M}$ . While the memory  $\mathcal{M}$  is kept by Eve, system 1 is sent to Alice, who applies the secret gate  $\mathcal{U}$ , and sends back output 2. Then, Eve applies the channel  $\mathcal{B}$ , producing two output systems, so that the overall transformation from inputs to outputs optimally emulates  $\mathcal{U}^{\otimes 2}$ . (b) One-to-two quantum learning of unitary. In a training phase, the example  $\mathcal{U}$  is applied locally on the bipartite state  $\sigma$ , and stored in the state  $\sigma_U = (U \otimes I)\sigma(U^\dagger \otimes I)$ . Then, two input systems interact with  $\sigma_U$ , undergoing a transformation that optimally emulates  $\mathcal{U}^{\otimes 2}$ .

$$C' = \text{Tr}_{1,2,\dots,2N}[(I_0 \otimes C_1^* \otimes \dots \otimes C_N^* \otimes I_{2N+1})R^{(N)}]. \quad (2)$$

In one-to-two cloning ( $N = 1$ ) the first input  $\mathcal{H}_0$  and the last output  $\mathcal{H}_3$  must have a bipartite structure,  $\mathcal{H}_0 = \mathcal{H}_{0B} \otimes \mathcal{H}_{0E}$  and  $\mathcal{H}_3 = \mathcal{H}_{3B} \otimes \mathcal{H}_{3E}$ , since the ultimate aim of the network is to mimic the bipartite channel  $\mathcal{U}_B \otimes \mathcal{U}_E$  on Bob's and Eve's systems. Then, the normalization of the Choi operator in Eq. (1) gives

$$\text{Tr}_3[R^{(1)}] = I_2 \otimes R^{(0)}, \quad \text{Tr}_1[R^{(0)}] = I_0. \quad (3)$$

Inserting the gate  $\mathcal{U}$  in the network, we obtain the bipartite channel  $\mathcal{C}'_U$ , which according to Eq. (2) is given by  $C'_U = \text{Tr}_{1,2}[(I_0 \otimes |U\rangle\langle U|_{1,2} \otimes I_3)R^{(1)}]$ .

We derive now the cloning network for which the channel  $\mathcal{C}'_U$  most closely resembles  $\mathcal{U}_B \otimes \mathcal{U}_E$ . As a figure of merit we use the global channel fidelity, uniformly averaged over the unknown unitaries

$$F = \int dU \frac{1}{d^4} \text{Tr}[C_U |U\rangle\langle U|^{\otimes 2}] \\ = \frac{1}{d^4} \int dU \langle U | \langle U | \langle U^* | R^{(1)} | U \rangle | U \rangle | U^* \rangle. \quad (4)$$

Note that  $F = 1$  if and only if  $C_U = \mathcal{U}^{\otimes 2}$  for any  $U$ , corresponding to perfect cloning. Exploiting symmetry then provides a radical simplification of the problem:

**Lemma 1.** The optimal cloning network maximizing the channel fidelity (4) can be assumed without loss of generality to be covariant, i.e., with a Choi operator  $R^{(1)}$  satisfying the commutation relation

$$[R^{(1)}, V_0^{\otimes 2} \otimes V_1^* \otimes W_2^* \otimes W_3^{\otimes 2}] = 0 \quad \forall V, W \in \text{SU}(d). \quad (5)$$

**Proof.** Let  $R^{(1)}$  be optimal. Then take its average  $\overline{R^{(1)}} = \int dV \int dW \mathcal{V}_0^{\otimes 2} \otimes \mathcal{V}_1^* \otimes \mathcal{W}_2^* \otimes \mathcal{W}_3^{\otimes 2} R^{(1)}$ , where  $\mathcal{V}$ ,  $\mathcal{V}^*$ ,  $\mathcal{W}^*$ ,  $\mathcal{W}$  are the unitary channels corresponding to

$V, V^*, W^*, W$ . It is immediate to see that  $\overline{R^{(1)}}$  satisfies Eqs. (5) and (3), and has the same fidelity as  $R^{(1)}$ . ■

The representation  $V^{\otimes 2} \otimes V^*$  in Eq. (5) can be decomposed into irreducible blocks as follows. First, one has  $V^{\otimes 2} = V_+ \oplus V_-$ , where  $V_\pm$  is the irreducible block acting in symmetric (antisymmetric) subspace  $\mathcal{H}_\pm \subset \mathcal{H}^{\otimes 2}$ , of dimension  $d_\pm = d(d \pm 1)/2$ . Then, one can further decompose  $V_+ \otimes V^* = V_{\alpha,+} \oplus V_{\beta,+}$ , where  $V_{\alpha,+} (V_{\beta,+})$  is the irreducible block acting on the subspace  $\mathcal{H}_{\alpha,+} (\mathcal{H}_{\beta,+}) \subset \mathcal{H}_+ \otimes \mathcal{H}$ , of dimension  $d_\alpha = d [d_\beta = d(d_+ - 1)]$ . Similarly,  $V_- \otimes V^* = V_{\alpha,-} \oplus V_{\gamma,-}$ , corresponding to irreducible subspaces  $\mathcal{H}_{\alpha,-}, \mathcal{H}_{\gamma,-} \subset \mathcal{H}_- \otimes \mathcal{H}$  of dimensions  $d_\alpha = d, d_\gamma = d(d_- - 1)$ , respectively. Note that the subspaces  $\mathcal{H}_{\alpha,+}$  and  $\mathcal{H}_{\alpha,-}$  carry equivalent representations, and that for qubits the block  $\mathcal{H}_{\gamma,-}$  does not show up. By Schur lemmas, the Choi operator  $R^{(1)}$  in Eq. (5) must be of the form  $R^{(1)} = \sum_{\mu, \nu \in \mathbf{S}} \sum_{i,j,k,l=\pm} r_{ik,jl}^{\mu\nu} T_{ij}^\mu \otimes T_{kl}^\nu$  where  $\mathbf{S} = \{\alpha, \beta, \gamma\}$ ,  $r_{ik,jl}^{\mu\nu}$  is a non-negative matrix for any  $\mu, \nu$ , and  $T_{ij}^\mu = \sum_{n=1}^{d_\mu} |\mu, i, n\rangle \langle \mu, j, n|$  is the isomorphism between the two equivalent subspaces  $\mathcal{H}_{\mu,i}$  and  $\mathcal{H}_{\mu,j}$  ( $T_{+-}^\beta = T_{-+}^\beta = T_{+-}^\gamma = T_{-+}^\gamma = 0$ ). Exploiting this fact, we obtain for the fidelity the following expression:

$$F = \frac{1}{d^4} \sum_{\mu \in \mathbf{S}} \sum_{i,j=\pm} d_\mu r_{ii,jj}^{\mu\mu} \quad (6)$$

while the normalization constraint of Eq. (3) becomes

$$\sum_{\mu, \nu} d_\mu d_\nu t_i^{\mu\nu} = d_i d, \quad t_i^{\mu\nu} := \sum_k r_{ik,ik}^{\mu\nu}, \quad i = \pm. \quad (7)$$

We are now ready to derive the optimal cloner:

**Theorem 1.** The fidelity of the optimal universal cloning of unitary transformations is  $F_{\text{clon}} = (d + \sqrt{d^2 - 1})/d^3$ . The value  $F_{\text{clon}}$  is achieved by a network as in Fig. 1(a) with preprocessing channel  $\mathcal{A}$  from  $\mathcal{S}(\mathcal{H}_{0B} \otimes \mathcal{H}_{0E})$  to  $\mathcal{S}(\mathcal{H}_{0A} \otimes \mathcal{M})$ ,  $\mathcal{M}$  being a memory qubit, given by

$$\mathcal{A}(\rho) = \sum_{i,j=\pm} \text{Tr}_{0E}[P_i \rho P_j] \otimes |i\rangle \langle j| \quad (8)$$

( $P_\pm$  orthogonal projector on  $\mathcal{H}_\pm$ , and  $\{|+\rangle, |-\rangle\}$  orthonormal basis for  $\mathcal{M}$ ), and postprocessing channel  $\mathcal{B}$  from  $\mathcal{S}(\mathcal{H}_{0B} \otimes \mathcal{M})$  to  $\mathcal{S}(\mathcal{H}_{3B} \otimes \mathcal{H}_{3E})$ , given by

$$\mathcal{B}(\sigma) = \sum_{i,j=\pm} \frac{d}{\sqrt{d_i d_j}} P_i [|\langle i|\sigma|j\rangle \otimes I_{3E}] P_j. \quad (9)$$

Accordingly, the approximate cloning of  $U$  is a channel  $\mathcal{C}'_U$  from  $\mathcal{S}(\mathcal{H}_{0B} \otimes \mathcal{H}_{0E})$  to  $\mathcal{S}(\mathcal{H}_{3B} \otimes \mathcal{H}_{3E})$ :

$$\begin{aligned} \mathcal{C}'_U(\rho) &= \mathcal{B} \circ (\mathcal{U} \otimes I_{\mathcal{M}}) \circ \mathcal{A}(\rho) \\ &= \sum_{i,j=\pm} \frac{d}{\sqrt{d_i d_j}} P_i [U \text{Tr}_{0E}[P_i \rho P_j] U^\dagger \otimes I] P_j. \end{aligned} \quad (10)$$

**Proof.** For the fidelity we have the following bound:

$$\begin{aligned} F &\leq \frac{1}{d^4} \left( \sum_{i=\pm} \sqrt{\sum_{\mu \in \mathbf{S}} d_\mu r_{ii,ii}^{\mu\mu}} \right)^2 \leq \frac{1}{d^4} \left( \sum_{i=\pm} \sqrt{\sum_{\mu \in \mathbf{S}} d_\mu t_i^{\mu\mu}} \right)^2 \\ &\leq \frac{1}{d^4} \left( \sum_{i=\pm} \max_{\mu \in \mathbf{S}} \left\{ \sqrt{\frac{d_i d}{d_\mu}} \right\} \right)^2. \end{aligned}$$

The first inequality comes from Schwartz inequality applied to the non-negative matrix  $a_{i,j} = \sum_{\mu} d_\mu r_{ii,ii}^{\mu\mu}$ , the second from the definition of  $t_i^{\mu\mu}$  [Eq. (7)], and the third from constraint (7). Since the maximum in the bound is achieved for minimum  $d_\mu$ , i.e., for  $\mu = \alpha$ , we have  $F \leq 1/d^4 (\sqrt{d_+} + \sqrt{d_-})^2 \equiv F_{\text{clon}}$ . To conclude achievability, we directly compute the fidelity between  $\mathcal{C}'_U$  [Eq. (10)] and  $\mathcal{U}^{\otimes 2}$ , which yields  $F(\mathcal{C}'_U, \mathcal{U}^{\otimes 2}) = F_{\text{clon}} \forall U$ . ■

Let us now clarify the meaning of the pre- and post-processing channels  $\mathcal{A}$  and  $\mathcal{B}$  in the optimal network. First, channel  $\mathcal{A}$  can be extended to a unitary interaction between the input systems  $\mathcal{H}_{0B}, \mathcal{H}_{0E}$  and the memory  $\mathcal{M}$ :  $\mathcal{A}(\rho) = \text{Tr}_{0E}[V(\rho \otimes |0\rangle \langle 0|) V^\dagger]$ , where  $|0\rangle = (|+\rangle + |-\rangle)/\sqrt{2} \in \mathcal{M}$ , and  $V$  is the controlled-swap  $V = I \otimes |+\rangle \langle +| + S \otimes |-\rangle \langle -|$ ,  $S|\phi\rangle|\psi\rangle = |\psi\rangle|\phi\rangle$ . Such an extension has a very intuitive meaning in terms of quantum parallelism: for bipartite input  $|\Psi\rangle_{BE}$  the single-system unitary  $U$  is made to work on both  $B$  and  $E$  by applying it to the superposition  $|\Psi\rangle_{BE} + S|\Psi\rangle_{BE}$  and discarding  $E$ . Less intuitive, and much more intriguing, is the meaning of channel  $\mathcal{B}$ . It is an extension of optimal universal cloning of pure states [19]: if system  $\mathcal{H}_{0B}$  and the ancilla  $\mathcal{M}$  are prepared in the state  $|\psi\rangle|+\rangle$ , then we obtain  $\mathcal{B}(|\psi\rangle \langle \psi| \otimes |+\rangle \langle +|) = d/d_+ [P_+ (|\psi\rangle \langle \psi| \otimes I) P_+]$ , which are indeed two optimal clones of  $|\psi\rangle$ . This means that realizing the optimal cloning of unitaries is a harder task than realizing the optimal cloning of states: an eavesdropper that is able to optimally clone unitaries must also be able to optimally clone pure states. This suggests that cryptographic protocols based on gates (such as the two-way protocol in the introduction) might be harder to attack than protocols based on states.

The performances of the optimal cloner crucially depend on entanglement with the quantum memory  $\mathcal{M}$ : Suppose that after channel  $\mathcal{A}$  the ancillary qubit  $\mathcal{M}$  decoheres on the basis  $\{|+\rangle, |-\rangle\}$ . Then, the approximate cloning of  $U$  is no longer given by Eq. (10), but rather by its decohered version  $\tilde{\mathcal{C}}'_U(\rho) = \sum_{i=\pm} d/d_i [P_i (U \text{Tr}_{0E}[P_i \rho P_i] \otimes I) P_i]$ . Direct calculation of the fidelity in this case gives  $F_{\text{deco}} = 1/d^2$ , which is exactly the same fidelity  $F_{\text{ran}}$  that one would achieve by applying  $U$  on the first system and by performing a randomly chosen unitary on the second. For large  $d$ , the optimal fidelity achieved with the quantum memory is essentially twice this value. Another classical cloning strategy would be to optimally estimate the unknown unitary, getting an estimate  $\hat{U}$ , and then performing  $\hat{U}^{\otimes 2}$  on the input systems  $\mathcal{H}_{0B}$  and  $\mathcal{H}_{0E}$ . Using the optimal estimation strategy of Ref. [20] we can readily evaluate the



fidelity of estimation to be  $F_{\text{est}} = 6/d^4$  for  $d > 2$ ,  $F_{\text{est}} = 5/16$  for  $d = 2$ . Note that, as far as it concerns the global fidelity, for  $d > 2$  estimation is by far worse than the crude decohered strategy described above.

We answer now a natural question: since there is a canonical isomorphism between unitaries  $U$  and maximally entangled states  $|U\rangle = (U \otimes I)|I\rangle$ , one might wonder whether the optimal cloning of  $U$  can be achieved via cloning of the state  $|U\rangle$ . Surprisingly at first sight, the answer is negative. To prove this fact, we put ourselves in a slightly more general scenario: we apply the unknown gate  $U$  to an arbitrary bipartite state  $\sigma \in \mathcal{S}(\mathcal{P})$  (not necessarily maximally entangled), and use the state  $\sigma_U := (U \otimes I)\sigma(U^\dagger \otimes I)$  to program a transformation  $\mathcal{L}_U$  on the two systems  $\mathcal{H}_{0B}$ ,  $\mathcal{H}_{0E}$ , given by  $\mathcal{L}_U(\rho) = \text{Tr}_{\mathcal{P}}[W(\rho \otimes \sigma_U)W^\dagger]$ , where  $W$  is a suitable interaction. Again, the goal is to maximize the fidelity between  $\mathcal{L}_U$  and  $\mathcal{U}_B \otimes \mathcal{U}_E$ . This is an elementary instance of quantum learning, in which a training set of examples— $N$  uses of the unknown  $\mathcal{U}$ —is provided in a first stage ( $N = 1$  here), and, after the training has been concluded, the learning machine is asked to optimally emulate  $\mathcal{U}^{\otimes M}$  ( $M = 2$  here). Using the same method illustrated for cloning, we can find the optimal one-to-two learning network with Choi operator  $L^{(1)}$ , for which the normalization (3) now reads

$$\text{Tr}_3[L^{(1)}] = I_2 \otimes \rho_1, \quad \text{Tr}[\rho_1] = 1, \quad (11)$$

where  $I_2$  acts on the tripartite space  $\mathcal{H}_2 = \mathcal{H}_{2B} \otimes \mathcal{H}_{2E} \otimes \mathcal{H}_{2A}$  and the space  $\mathcal{H}_0$  is one-dimensional (see Fig. 1). Using the symmetry argument of Lemma 1, we can restrict the optimization to Choi operators satisfying  $[L^{(1)}, V_1^* \otimes V_{2B,2E}^{\otimes 2} \otimes W_{2A}^* \otimes W_{3B,3E}^{\otimes 2}] = 0$ , i.e., of the form  $L^{(2)} = \sum_{\mu, \nu \in \mathcal{S}} \sum_{i,j,k,l=\pm} l_{ik,jl}^{\mu\nu} T_{ij}^\mu \otimes T_{kl}^\nu$ . The normalization of Eq. (11) then becomes  $\sum_{\nu,m} d_\nu l_{im,jm}^{\mu\nu} = \delta_{ij}$  for any  $\mu \in \mathcal{S}$ , and  $i, j = \pm$ . Maximizing the fidelity under this constraint we then obtain the maximum value  $F_{\text{learn}} = 6/d^4$  for  $d > 2$  and  $F_{\text{learn}} = 5/d^4$  for  $d = 2$ , exactly the same value of optimal estimation. Therefore, any scheme based on the application of  $U$  on a fixed input state will be extremely poor compared with the optimal cloner. This highlights the fundamental difference between quantum learning and cloning: in learning one has to first apply the unknown gate  $U$  to a fixed state  $\sigma$ , which implies an irreversible degradation of its computational power. Note that the difference between cloning and learning is a specific trait of quantum channels, while for states there is no difference between the two tasks. Moreover, one can regard the quantum learning of unitary transformations as a special case of gate programming [5], e.g., for one-to-two learning the program is of the form  $\sigma_U = (U \otimes I)\sigma(U^\dagger \otimes I)$  and the target is  $U \otimes U$ .

In conclusion, in this Letter we proved a general no-cloning theorem for black boxes, and derived the optimal universal cloning of unitary transformations from one-to-two copies. The optimal cloner is realized via a quantum

channel with memory, and greatly outperforms the optimal measure-and-prepare strategy. Exploring the deep relations among cloning, learning, and programming of quantum transformations is a natural development of our work and an interesting avenue for future research.

We thank S. De Zordo for preliminary calculations in his master thesis. G. C. is grateful to M. Murao, M. Hayashi, and A. Winter for useful and enjoyable discussions. This work is supported by the EC through the networks SECOCQ and CORNER.

\*chiribella@fisicavolta.unipv.it

+dariano@unipv.it

\*perinotti@fisicavolta.unipv.it

§http://www.qubit.it

- [1] W. K. Wootters and W. H. Zurek, *Nature (London)* **299**, 802 (1982); D. Dieks, *Phys. Lett. A* **92**, 271 (1982).
- [2] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, *Rev. Mod. Phys.* **77**, 1225 (2005).
- [3] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, *Phys. Rev. Lett.* **87**, 270404 (2001).
- [4] A. Acín, *Phys. Rev. Lett.* **87**, 177901 (2001).
- [5] M. A. Nielsen and I. L. Chuang, *Phys. Rev. Lett.* **79**, 321 (1997).
- [6] S. F. Huelga, J. A. Vaccaro, A. Chefles, and M. B. Plenio, *Phys. Rev. A* **63**, 042303 (2001).
- [7] J. Eisert, K. Jacobs, P. Papadopoulos, and M. B. Plenio, *Phys. Rev. A* **62**, 052317 (2000).
- [8] Y.-F. Huang, X.-F. Ren, Y.-S. Zhang, L.-M. Duan, and G.-C. Guo, *Phys. Rev. Lett.* **93**, 240501 (2004).
- [9] G. M. D'Ariano and P. Lo Presti, *Phys. Rev. Lett.* **86**, 4195 (2001).
- [10] D. Leung, *J. Math. Phys. (N.Y.)* **44**, 528 (2003).
- [11] G. Gutoski and J. Watrous, in *Proceedings of the 39th Annual ACM Symposium on Theory of Computation* (ACM, New York, 2007), p. 565.
- [12] H. P. Yuen, arXiv:quant-ph/0207089.
- [13] G. M. D'Ariano, D. Kretschmann, D. M. Schlingemann, and R. F. Werner, *Phys. Rev. A* **76**, 032328 (2007).
- [14] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [15] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.
- [16] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [17] Note that  $p$  cannot exceed  $\frac{1}{2}$ : with a random guess one can always reduce the error probability to  $\frac{1}{2}$ .
- [18] If perfect discrimination is possible, discrimination plus reparation gives perfect cloning. Vice versa, if perfect cloning is possible, the no-cloning theorem implies  $p_{ij} = 0$  for any  $i \neq j$ . Then we need only  $k - 1$  perfect clones to achieve perfect discrimination via  $k - 1$  pairwise tests.
- [19] R. F. Werner, *Phys. Rev. A* **58**, 1827 (1998).
- [20] G. Chiribella, G. M. D'Ariano, and M. F. Sacchi, *Phys. Rev. A* **72**, 042338 (2005).