Characterising a universal cloning machine by maximum-likelihood estimation

Massimiliano F. Sacchi
Optics Section, Blackett Laboratory, Imperial College London, London SW7 2BZ, United Kingdom
and Dipartimento di Fisica ‘A. Volta’, Università di Pavia and Unità INFM, via A. Bassi 6, I-27100 Pavia, Italy
(November 21, 2018)

We apply a general method for the estimation of completely positive maps to the 1-to-2 universal covariant cloning machine. The method is based on the maximum-likelihood principle, and makes use of random input states, along with random projective measurements on the output clones. The downhill simplex algorithm is applied for the maximisation of the likelihood functional.

PACS Numbers: 03.65.-w, 03.67.-a, 03.67.Lx

II. MEASURING QUANTUM DEVICES

The maximum-likelihood principle states that the best estimation of unknown parameters is given by the values that are most likely to produce the data one experimenter has observed. Hence, this principle involves the maximisation of a function of the unknown parameters that is given by the theoretical probability of getting the collected data.

Consider a sequence of $K$ independent measurements on the output of a physical device acting on quantum states. Each measurement is described by the element $F_l(x_l)$ of a POVM, where $x_l$ denotes the outcome at the $l$th measurement, and $l = 1, 2, ..., K$. Let us denote by $\rho_l$ the state at the input at the $l$th run. The probability of getting the string of outcomes $\bar{x} = \{x_1, x_2, ..., x_K\}$ is given by

$$ p(\bar{x}) = \Pi_{l=1}^K \text{Tr}[\mathcal{E}(\rho_l)F_l(x_l)] \ .$$

The best estimate of the map $\mathcal{E}$ maximizes the logarithm of Eq. (1)

$$ \mathcal{L}(\mathcal{E}) = \sum_{l=1}^K \log \text{Tr}[\mathcal{E}(\rho_l)F_l(x_l)]$$

over the set of completely positive maps. The likelihood function $\mathcal{L}(\mathcal{E})$ is concave, and in the present case it is defined on the convex set of CP maps. Its maximum is achieved by a single CP map if the data sample is sufficiently large, and the set of measurements is a quorum

The constraints to be imposed in the maximisation problem are the complete positivity and the trace-preserving property of the map $\mathcal{E}$. A trace-preserving CP map is a linear map from operators in Hilbert space $\mathcal{H}$ to operators in $\mathcal{K}$ which can be written in the Kraus form.
\[ \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger , \]  

where
\[ \sum_k A_k^\dagger A_k = \mathbb{1}_\mathcal{H} . \]  

Let \( \dim(\mathcal{H}) = N \) and \( \dim(\mathcal{K}) = M \), and consider an orthonormal basis \( \{ V_i \} \) for the space of linear operators on \( \mathcal{H} \), namely
\[ \text{Tr}[V_i^\dagger V_j] = \delta_{ij} , \]  

and for any operator \( O \)
\[ O = \sum_{i=1}^{N^2} \text{Tr}[V_i^\dagger O] V_i . \]  

Upon defining the operator
\[ S = \sum_{i=1}^{N^2} \mathcal{E}(V_i) \otimes V_i^* , \]  

where \( \ast \) denotes complex conjugation, one can write for linearity
\[ \mathcal{E}(\rho) = \text{Tr}_\mathcal{H}[(\mathbb{1}_\mathcal{K} \otimes \rho^T) S] , \]  

where \( T \) is the transposition. Notice that
\[ \sum_{i=1}^{N^2} V_i \otimes V_i^* = |\Psi\rangle \langle \Psi| , \]  

where \( |\Psi\rangle \) is given by the (unnormalized) maximally entangled state
\[ |\Psi\rangle = \sum_{n=1}^N |n\rangle \otimes |n\rangle . \]  

Hence, one has also
\[ S = \mathcal{E} \otimes \mathbb{1}(|\Psi\rangle \langle \Psi|) . \]  

Eqs. (7) and (8) establish an isomorphism between linear maps from \( \mathcal{H} \) to \( \mathcal{K} \) and linear operators on the tensor-product space \( \mathcal{K} \otimes \mathcal{H} \). Complete positivity and trace-preserving property of \( \mathcal{E} \) imply
\[ S \geq 0 \quad \text{and} \quad \text{Tr}_\mathcal{K}[S] = \mathbb{1}_\mathcal{H} . \]  

For the construction of the likelihood function \( \mathcal{L}(\mathcal{E}) \) the condition \( S \geq 0 \) is crucial. Actually, it allows to write
\[ S = C^\dagger C , \]  

where \( C \) is an upper triangular matrix, with positive diagonal elements. Similarly, one has for the density matrices \( \rho_i^T \) and the POVM’s \( F_i(x_i) \)
\[ \rho_i^T = R_i^\dagger R_i , \quad F_i(x_i) = A_i^\dagger(x_i)A_i(x_i) . \]  

From Eqs. (8), (13) and (14), the likelihood functional in Eq. (2) rewrites
\[ \mathcal{L}(\mathcal{E}) = \mathcal{L}(C) = \sum_{i=1}^K \log \text{Tr}[C^\dagger C(R_i^\dagger R_i) \otimes A_i^\dagger(x_i)A_i(x_i))] \]
\[ = \sum_{i=1}^K \log \sum_{n,m=1}^{N,M} \left| \langle n|C(R_i^\dagger \otimes A_i^\dagger(x_i))|m\rangle \right|^2 , \]  

where \( \{ |n\rangle \} \) denotes an orthonormal basis for \( \mathcal{H} \otimes \mathcal{K} \). On one hand, the parameterisation in Eq. (15) implicitly constrains the complete positivity of the map \( \mathcal{E} \). On the other, the argument of the logarithm is explicitly positive, thus assuring the stability of numerical methods to evaluate \( \mathcal{L}(C) \).

The trace-preserving condition is given in terms of the matrix \( S \) by \( \text{Tr}_\mathcal{K}[S] = \mathbb{1}_\mathcal{H} \). However, the constraint \( \text{Tr}[S] = N \) which follows from \( \text{Tr}_\mathcal{K}[S] = \mathbb{1}_\mathcal{H} \) isolates a closed convex subset of the set of positive matrices. Hence, the maximum of the concave likelihood functional still remains unique under this looser constraint, and one can check \( \text{a posteriori} \) that the condition \( \text{Tr}_\mathcal{K}[S] = \mathbb{1}_\mathcal{K} \) is fulfilled. Using the method of Lagrange multipliers, then one maximises the effective functional
\[ \tilde{\mathcal{L}}(C) = \mathcal{L}(C) - \mu \text{Tr}[C^\dagger C] , \]  

where \( \mathcal{L}(C) \) is given in Eq. (15), and the value of the multiplier \( \mu \) can be obtained as follows. Writing \( S \) in terms of its eigenvectors as \( S = \sum_i s_i^2 |s_i\rangle \langle s_i| \), the maximum likelihood condition \( \partial \mathcal{L}(C)/\partial s_i = 0 \) implies
\[ \sum_{i=1}^K \frac{\text{Tr}[s_i^2 |s_i\rangle \langle s_i|]}{\text{Tr}[s_i^2]} = \mu \frac{\text{Tr}[s_i^2]}{\text{Tr}[s_i^2]} , \]  

Multiplying by \( s_i \) and summing over \( i \) gives \( \mu = K/N \).

**III. CHARACTERISING THE UNIVERSAL CLONING MACHINE**

We consider now the problem of estimating the CP map pertaining to the 1-to-2 universal covariant cloning machine. The map is given by
\[ \mathcal{E}(\rho) = \frac{2}{3} s_2 (\rho \otimes \mathbb{1}) s_2 , \]  

where \( s_2 \) is the projection operator on the symmetric subspace, which is spanned by the set of vectors \( \{ |s_i\rangle \langle s_i|, i = 0 \pm 2 \} \), with \( |s_0\rangle = |00\rangle, |s_1\rangle = 1/\sqrt{2}(|01\rangle + |10\rangle) \), and \( |s_2\rangle = |11\rangle \), where \( \{ |0\rangle, |1\rangle \} \) is a basis for each spin 1/2 system. Using the lexicographic ordering for the basis of the tensor-product Hilbert space, one has
\[
s_2 = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1/2 & 1/2 & 0 \\
0 & 1/2 & 1/2 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

In the following we label with \(A\) and \(B,C\) the Hilbert spaces supporting the input state and the two output copies, respectively. One can apply Eq. (7) to obtain the projector \(s_2\) and \(s_3\) corresponding matrix copies, respectively. One can apply Eq. (7) to obtain the spaces supporting the input state and the two output measurements along random directions \(\vec{r}_i\) and \(\vec{r}_i\) of the \(|\psi\rangle\) in the reconstruction of the density matrix of radiation \(A\) with \(A\) and \(B,C\) of the matrix \(S\) one has

\[
S = \frac{2}{3} \sum_{i=0}^4 \left[ (s_2)^{BC} (s_3)^{BC} \right] \otimes \sigma_i^{A}
\]

The Cholesky decomposition for \(\rho^T_l\) and \(F_l(a_1, b_1)\) writes as in Eqs. (14), with

\[
R_l = \begin{pmatrix}
cos(\theta_l/2) & e^{i\phi_l} \sin(\theta_l/2) \\
0 & 0
\end{pmatrix},
\]

and

\[
\alpha_l = \frac{\alpha_l}{2} + \pi \frac{a_1 - 1}{4},
\]

\[
\gamma_l = \frac{\gamma_l}{2} + \pi \frac{a_1 - 1}{4},
\]

We have now all the ingredients to construct the likelihood function in Eq. (15), which will be a function of the 64 real parameters that specify the triangular matrix \(C\). The problem of the maximisation of \(L(C)\) enters the realm of programming and numerical algebra optimisation, where various techniques are known [2].

In the following we show the results of a simulation obtained by applying the method of downhill simplex [21,22] to find the maximum of the likelihood functional. This method is robust and efficient in case of a relatively small number of parameters. It has been reliably used in the reconstruction of the density matrix of radiation field and spin systems [3], and in the characterisation of quantum communication channels for qubits [3].

The results are shown in Fig. 1. Pure states at the input of the cloning machine have been used, together with projective measurements over the two clones at the output. In both cases we adopted a uniform distribution on the Bloch sphere. The Monte Carlo method has been used to generate \(K = 10000\) data, by using the theoretical probability

\[
p(a_t, b_t) = \frac{2}{3} \text{Tr}[s_1 \rho \otimes s_2 F_l(a_t, b_t)].
\]
FIG. 1. Maximum-likelihood reconstruction of the CP map of the 1-to-2 universal covariant cloning. The picture represents the values of the (real part of the) elements of the matrix \( S \). Random pure states at the input of the cloning machine, and projective measurements along random directions on the two clones at the output have been used, with \( K = 10000 \) couple of measurements. The statistical error in the reconstruction is of the order \( 10^{-2} \). The results compare very well with the theoretical values of Eq. (21).

A lengthy but straightforward calculation gives

\[
p(a_l, b_l) = \frac{1}{4} + \frac{1}{6} (a_l \cos \alpha_l + b_l \cos \gamma_l) \cos \theta_l \\
+ \frac{1}{6} (a_l \sin \alpha_l \cos (\beta_l - \phi_l) + b_l \sin \gamma_l \cos (\delta_l - \phi_l)) \sin \theta_l \\
+ \frac{1}{12} a_l b_l [\cos \alpha_l \cos \gamma_l + \sin \alpha_l \sin \gamma_l \cos (\delta_l - \beta_l)].
\]

(31)

FIG. 2. Average statistical error in the characterisation of the universal covariant cloning machine versus number of data \( K \). The error affects the value of the reconstructed elements of the matrix \( S \) that is univocally related to the CP map of the cloning machine. The dotted line represents the asymptotic dependence on the inverse square root of \( K \), in accordance with the central limit theorem.

In conclusion, we have applied a general method to reconstruct experimentally the completely positive map describing a physical device to the universal covariant cloning machine. The method, based on the maximum likelihood principle, involves the maximisation of a functional, which depends on the results of quantum measurements performed on the clones at the output. The maximisation has to be made over all possible trace-preserving completely positive maps. A suitable parametrisation is allowed by the isomorphism between linear map from Hilbert spaces \( \mathcal{H} \) to \( \mathcal{K} \) and linear operators in \( \mathcal{H} \otimes \mathcal{K} \), along with the Cholesky decomposition of positive matrices. The numerical results we showed here has been obtained by applying the method of the downhill simplex to search the maximum of the likelihood functional. In our example, a good characterisation of the 1-to-2 universal cloning machine has been achieved, with a number of simulated data as low as \( 10^4 \). This is relevant, because some experiments are now feasible, but with low data rate or short stability time. In accordance with the central limit theorem, the statistical error of the characterisation shows the inverse-square-root asymptotic dependence on the number of data.

The method is very general, can be implemented immediately in the lab, and can be adopted in many fields as quantum optics, spins, optical lattices, atoms, ion trap, etc.

ACKNOWLEDGMENTS

The author would like to thank the Leverhulme Trust foundation for partial support. This work has been supported by the Italian Ministero dell’Università e della Ricerca Scientifica e Tecnologica (MURST) under the co-sponsored project 1999 Quantum Information Transmission and Processing: Quantum Teleportation and Error Correction.

[1] W. K. Wootters and W. H. Zurek, Nature 299, 802 (1982); H. P. Yuen, Phys. Lett. A 113, 405 (1986).
[2] V. Bužek and M. Hillery, Phys. Rev. A 54, 1844 (1996); N. Gisin and S. Massar, Phys. Rev. Lett. 79, 2153 (1997).
[3] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello and J. A. Smolin, Phys. Rev. A 57, 2368
(1998); D. Bruß, A. Ekert and C. Macchiavello, Phys. Rev. Lett. 81, 2598 (1998).

[4] R. F. Werner, Phys. Rev. A 58, 1827 (1998).

[5] N. Gisin and S. Massar, Phys. Rev. Lett. 79, 2153 (1997); N. Gisin and B. Huttner, Phys. Lett. A 228, 13 (1997).

[6] G. M. D’Ariano, C. Macchiavello, and M. F. Sacchi, quant-ph/0007063; G. M. D’Ariano, F. De Martini, and M. F. Sacchi, quant-ph/0012025 (to appear on Phys. Rev. Lett.); G. M. D’Ariano and M. F. Sacchi, quant-ph/0009080.

[7] M. F. Sacchi, quant-ph/0009104.

[8] S. L. Braunstein, A. S. Lane, and C. M. Caves, Phys. Rev. Lett. 69, 2153 (1992).

[9] K. Banaszek, G. M. D’Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A 61, 10304(R) (2000).

[10] G. M. D’Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A 62, 023815 (2000); G. M. D’Ariano, M. G. A. Paris, and M. F. Sacchi, quant-ph/0009081.

[11] J. de Pillis, Pacific J. of Math. 23, 129 (1967).

[12] A. Jamiołkowski, Rep. Math. Phys. 3, 275 (1972).

[13] S. Yu, Phys. Rev. A 62, 024302 (2000).

[14] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, quant-ph/0007057.

[15] For the concept of quorum, see G. M. D’Ariano, L. Maccone, and M. G. A. Paris, quant-ph/0006004, Phys. Lett. A 276, 25 (2000).

[16] See, for example, M. A. Nielsen and C. M. Caves, Phys. Rev. A 55, 2547 (1997).

[17] G. M. D’Ariano, P. Lo Presti, and M. F. Sacchi, Phys. Lett. A 272, 32 (2000).

[18] For a positive—but not completely positive—map the condition of positivity of the matrix $S$ is relaxed, by only requiring positivity for tensor product of vectors, namely $\langle \phi \otimes \psi | S | \phi \otimes \psi \rangle \geq 0$ (see Ref. [12]).

[19] Such decomposition is referred to as Cholesky decomposition, and is commonly used in linear programming (see, e.g., Ref. [21]). Moreover, if the matrix $S$ is strictly positive the decomposition is unique.

[20] We mean that $|i\rangle \otimes |j\rangle$ precedes $|k\rangle \otimes |l\rangle$ if and only if either $i < k$ or $i = k$ and $j < l$.

[21] See, for example, P. G. Ciarlet, Introduction to numerical linear algebra and optimisation, Cambridge Univ. Press, Cambridge, 1989.

[22] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in Fortran: The Art of Scientific Computing, Cambridge Univ. Press, Cambridge, 1992.