Optimality of programmable quantum measurements

D. Pérez-García
Max Planck Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, Garching, D-85748, Germany
and Departamento de Matemática Aplicada, Universidad Rey Juan Carlos, C/Tulipan s/n, 28933 Móstoles (Madrid), Spain
(Received 9 February 2006; published 24 May 2006)

We prove that for a programmable measurement device that approximates *every* POVM with an error $\leqslant\delta$, the dimension of the program space has to grow at least polynomially with $\frac{1}{\delta}$. In the case of qubits we can improve the general result by showing a linear growth. This proves the optimality of the programmable measurement devices recently designed in G. M. D'Ariano and P. Perinotti, Phys. Rev. Lett. **94**, 090401 (2005).

DOI: 10.1103/PhysRevA.73.052315 PACS number(s): 03.67.−a

I. INTRODUCTION

One of the most important features of present day computers is their universality. That is, the same computer can achieve many different tasks by changing the program that runs in it. The analog quantum concept—the search for universal quantum devices—has attracted a lot of attention along the short history of quantum information and quantum computation. The idea behind this is the same as in the classical situation: the economy of resources. Given the difficulty of creating a quantum device, it would be desirable to create it as universal (or multipurpose) as possible. Most of the work in this direction has been concentrated in the two basic quantum operations: unitaries (or more generally channels) [1–6] and measurements [7–10].

In this paper we will restrict ourselves to the latter. That is, we will study measurement apparatus that can be programmed to achieve any generalized measurement (POVM). The program will be an ancillary quantum state that can be changed depending on the POVM one wants to get. The possible applications of these kind of devices are considerably large: measurement based quantum computation, eavesdropping of quantum encrypted information, and in general every quantum protocol in which one wants to change the measurements on the fly. However, exact universally programming of measurements is impossible [7,8] as a consequence of the no-go theorem for programmability of unitary transformations [1]. Hence, one has to restrict oneself to schemes that approximate any measurement with a fixed error δ.

The first universal programmable quantum device was designed in [8], but it needed a dimension m of the ancilla that grows exponentially versus the inverse of the error δ^{-1}. This was dramatically improved in [10], where only a polynomial (and linear in the case of qubits) growth of m is required. In this paper we will show that the results in [10] are optimal, by showing the following:

(1) The polynomial growth cannot be improved, that is, the dimension of the ancillary system has to scale at least like $(\frac{1}{\delta})^{(d-1)/2}$, where d is the dimension of the original system.

(2) The linear growth cannot be improved in the case of qubits, that is, we have at least the scaling $\frac{1}{\delta}$.

It is quite surprising that the estimate for the general case does not give the optimal exponent in the case of qubits. The reason comes from the techniques used in the paper. Since they come from the theory of convex bodies, we need a *real* vector space. This is straightforward in the case of qubits, where one has the Bloch sphere representation, but a much more artificial trick has to be used in the general case.

II. PRELIMINARIES

Let us recall that in quantum mechanics, the statistics of a generic measurement apparatus (with discrete sampling space) is described by a positive operator valued measure (POVM), that is, a set of positive operators (one for each possible outcome) $P^j\geqslant 0$ on the Hilbert space of the system such that $\Sigma_j P^j=\mathbb{1}$. The statistics of the outcomes j for an input state ρ are given by the Born rule

$$p(j|\rho)=\mathrm{tr}(\rho P^j).$$

How to design then a programmable POVM? The idea is to build a device that acts on the original system (we will call its dimension d) and on some ancillary or program system (with dimension m) that can be tuned just by changing the state ("quantum software") in the ancilla. Clearly, the most general programmable measurement device would be a fixed unitary U in both the system and the ancilla, followed by a POVM $(F^j)_j$ in the joint system. That is, if the ancilla state is σ, we will get the statistics

$$p(j|\rho)=\mathrm{tr}(U(\rho\otimes\sigma)U^\dagger F^j)\quad\forall\ j,\rho.$$

Including the unitary U in the POVM $(F^j)^j$ restricts our study to programmable POVMs $(G^j)_{j=1}^d$ of the form

$$G^j=\mathrm{tr}_a[(\mathbb{1}\otimes\sigma)F^j], \tag{1}$$

where $(F^j)_{j=1}^d$ is a POVM in the joint system.

How to *measure* the distance between the original POVM $\mathbb{P}=(P^j)_{j=1}^d$ and the programmed one $\mathbb{G}=(G^j)_{j=1}^d$? Maybe the most natural measure is given by the usual distance between the probability distributions of the outcomes; in our case

$$d(\mathbb{P},\mathbb{G})=\max_\rho\sum_{j=1}^d|\mathrm{tr}[\rho(P^j-G^j)]|.$$

Clearly $\max_j\|P^j-G^j\|_\infty\leqslant d(\mathbb{P},\mathbb{G})\leqslant d\max_j\|P^j-G^j\|_\infty$, and since $\frac{1}{d}\|\cdot\|_1\leqslant\|\cdot\|_\infty\leqslant\|\cdot\|_1$, and d is a fixed constant (we are

just interested in the asymptotics in m), we will be able to reason also with the trace norm $\|\cdot\|_1$.

So we want to construct a POVM $(F^j)_{j=1}^d$ in the joint system such that, for any POVM $(P^j)_{j=1}^d$ in the original system, there exists an ancillary state σ such that $d(\mathbb{G},\mathbb{P})\leq\delta$, where $\mathbb{G}=(G^j)_j$ is given by (1). It is argued in [10] that it is enough to approximate POVMs that are given by one-dimensional orthogonal projectors $P^j=|\phi^j\rangle\langle\phi^j|$. We will call such a POVM $(F^j)_{j=1}^d$ a δ-universal programmable measurement (δ-UPM).

III. RESULTS

The first result of this paper is to relate the dimension of the ancillary system with the existence of disjoint nets of balls in some Hilbert space (we will prove it at the end of the paper).

Theorem 1. For a δ-UPM, the dimension m of the ancillary system verifies $m\geq\frac{1}{d}A$, where A is the cardinality of a net of normalized pure states $(|\phi_\alpha\rangle)_{\alpha=1}^A$ in a d-dimensional Hilbert space such that

$$D(|\phi_\alpha\rangle,|\phi_\beta\rangle)>\sqrt{8d\delta} \tag{2}$$

for any $\alpha\neq\beta$, where

$$D(|\phi\rangle,|\varphi\rangle)=\tfrac{1}{2}\||\phi\rangle\langle\phi|-|\varphi\rangle\langle\varphi|\|_1=\sqrt{1-|\langle\phi|\varphi\rangle|^2}. \tag{3}$$

Then, to establish large lower bounds for the dimension of the ancilla m, it is enough to obtain large lower bounds for A.

A. The case of qubits

Let us start with the case of qubits. The key estimate will be given by the following lemma.

Lemma 2. Let us consider a small ϵ $\left(0<\epsilon<\frac{1}{10}\right)$. Then, in the unit sphere S of a n-dimensional real Hilbert space $\mathcal{H}$ $(n\geq2)$, one can take $\frac{1}{(10\epsilon)^{n-1}}$ elements x_j with the property that $\|x_j-x_i\|>\epsilon$ if $j\neq i$.

Proof. Let us consider a maximal subset $(x_j)_{j=1}^J$ such that $\|x_j-x_i\|\geq2\epsilon>\epsilon$ if $j\neq i$. By maximality, S can be covered by balls of radius 2ϵ centered in the x_j's. Moreover, if $1\geq\|x\|\geq1-\epsilon$, we have that there exists a j such that $\|\frac{x}{\|x\|}-x_j\|\leq2\epsilon$ and so

$$\|x-x_j\|\leq\left\|x-\frac{x}{\|x\|}+\frac{x}{\|x\|}-x_j\right\|\leq\left\|x-\frac{x}{\|x\|}\right\|+2\epsilon\leq3\epsilon.$$

This means that, if B is the unit ball of $\mathcal{H}$, the set $C=\cup_{j=1}^J\{x_j+3\epsilon B\}$ covers the shell $R=\{x\in\mathcal{H}:1-\epsilon\leq\|x\|\leq1\}$. Therefore

$$J3^n\epsilon^n\,\mathrm{vol}(B)\geq\mathrm{vol}(C)\geq\mathrm{vol}(R)=[1-(1-\epsilon)^n]\mathrm{vol}(B),$$

and so

$$J\geq\frac{1}{3^n\epsilon^n}(1-(1-\epsilon)^n)(*)$$

$$\geq\frac{1}{3^n}\left(\frac{1-\epsilon}{\epsilon}\right)^{n-1}\geq\frac{1}{(10\epsilon)^{n-1}}.(**)$$

To see $(*)$ one has to show that $1-(1-\epsilon)^n\geq\epsilon(1-\epsilon)^{n-1}$, which is clearly true since

$$1\geq(1-\epsilon)^{n-1}=\epsilon(1-\epsilon)^{n-1}+(1-\epsilon)^n.$$

To see $(**)$ it is enough to notice that $\frac{1}{3^n}\geq\frac{1}{9^{n-1}}$ (since $n\geq2$) and that $(1-\epsilon)\geq\frac{9}{10}$ (since $0<\epsilon<\frac{1}{10}$). ■

Now we consider the Bloch sphere. The distance D of (3) corresponds to $\frac{1}{2}$ the usual (Hilbert) distance in the Bloch sphere ([11] page 404). Therefore, using Lemma 2 (now $n=3$, $\epsilon=2\sqrt{8d\delta}$, and $d=2$), we can take a net of $\frac{1}{6400}\frac{1}{\delta}$ pure states with property (2). This immediately implies by Theorem 1 (now $d=2$) that the dimension m that we need in the ancilla to get a δ-UPM has to verify

$$m\geq\frac{1}{12800}\frac{1}{\delta}.$$

That is, the linear growth obtained in [10] is optimal.

Remark 3. Since the main aim under study is the growth of m with $\frac{1}{\delta}$, we are quite careless with the constants and then, as one can easily see, the constant $\frac{1}{12800}$ is far from optimal.

B. The general case

Let us now turn to the general case. Since we do not have now such a good real representation as the Bloch sphere, we will play the trick of restricting to some *real part* $\mathcal{H}_\mathbb{R}$ of our d-dimensional Hilbert space $\mathcal{H}$. That is, we fix one orthonormal basis $|i\rangle$ and we consider the elements that are of the form $\Sigma_{i=1}^d\lambda_i|i\rangle$ with the λ_i's real. This is a d-dimensional real Hilbert space $\mathcal{H}_\mathbb{R}$ with the inherited norm [given by $(\Sigma_i|\lambda_i|^2)^{1/2}$].

Now we also have to *identify* vectors up to global phases. For this reason we consider in the unit sphere of $\mathcal{H}_\mathbb{R}$ a maximal set $X=(x_j)_{j=1}^J$ of J elements with the following two properties:

(P1) $\|x_j-x_i\|\geq2\epsilon$ if $j\neq i$,

(P2) if $x\in X$, then $-x\in X$.

We have that the unit sphere $S_\mathbb{R}$ of $\mathcal{H}_\mathbb{R}$ can be covered by balls of radius 2ϵ centered in the x_j's. Let us see it.

If it is not the case, there exists an $x\in S_\mathbb{R}$ such that $\|x-x_j\|>2\epsilon$ for every j. By (P2), also $\|-x-x_j\|>2\epsilon$ for every j, and this implies that $X\cup\{x,-x\}$ also fulfills (P1) and (P2), which contradicts the maximality of X.

Then, we can reason as in Lemma 2. For any $x\in\mathcal{H}_\mathbb{R}$, with $1\geq\|x\|\geq1-\epsilon$, there exists a j such that $\|\frac{x}{\|x\|}-x_j\|\leq2\epsilon$ and so $\|x-x_j\|\leq3\epsilon$. This means that, if $B_\mathbb{R}$ is the unit ball of $\mathcal{H}_\mathbb{R}$, the set $C=\cup_{j=1}^J\{x_j+3\epsilon B_\mathbb{R}\}$ covers the shell $R=\{x\in\mathcal{H}_\mathbb{R}:1-\epsilon\leq\|x\|\leq1\}$. Therefore

$$J3^d\epsilon^d\,\mathrm{vol}(B_\mathbb{R})\geq\mathrm{vol}(C)\geq\mathrm{vol}(R)=[1-(1-\epsilon)^n]\mathrm{vol}(B_\mathbb{R})$$

and hence $J\geq\frac{1}{(10\epsilon)^{d-1}}$.

Now we choose either x or $-x$ for every $x\in X$ and obtain another sequence $(|\phi_j\rangle)_{j=1}^{J/2}$ of $\frac{J}{2}$ elements, that one can see in the original Hilbert space $\mathcal{H}$, for which, if $j\neq i$,

$$\| |\phi_j\rangle - |\phi_i\rangle \| \geq 2\epsilon \quad \text{by (P1)},$$

$$\| |\phi_j\rangle + |\phi_i\rangle \| \geq 2\epsilon \quad \text{by (P1) applied to } -|\phi_i\rangle.$$

Therefore,

$$\begin{aligned} D(|\phi_j\rangle,|\phi_i\rangle) &= \sqrt{1-|\langle\phi_j|\phi_i\rangle|^2} \geq \sqrt{1-|\langle\phi_j|\phi_i\rangle|}(*) \\ &= \min\{\sqrt{1-\langle\phi_j|\phi_i\rangle},\sqrt{1+\langle\phi_j|\phi_i\rangle}\}(**) \\ &= \frac{1}{\sqrt{2}}\min\{\| |\phi_j\rangle - |\phi_i\rangle\|, \| |\phi_j\rangle \\ &\quad + |\phi_i\rangle\|\} \geq \sqrt{2}\epsilon > \epsilon. \end{aligned}$$

$(*)$ is simply because we are inside a *real* Hilbert space $\mathcal{H}_{\mathbb{R}}$. $(**)$ comes from the equality $\| |\phi_j\rangle \pm |\phi_i\rangle\|^2 = 2(1 \pm \langle\phi_j|\phi_i\rangle)$, which is true also because we are in a real Hilbert space.

So now, taking $\epsilon = \sqrt{8d\delta}$, one can apply Theorem 1 to get that the dimension of the ancilla m needed to have a δ-UPM has to be $m \geq \frac{1}{d}\frac{J}{2}$, since we have a net of $\frac{J}{2}$ pure states verifying property (2) in Theorem 1. But $J \geq \frac{1}{(10\epsilon)^{d-1}}$, and this gives

$$m \geq k(d)\left(\frac{1}{\delta}\right)^{(d-1)/2},$$

for $\frac{1}{k(d)} = 2d(10\sqrt{8d})^{d-1}$ (which is again far form optimal).

So the best growth for m is polynomial in $\frac{1}{\delta}$. This implies that the control unitary (which has polynomial growth [10]) is essentially optimal among the programmable quantum measurements.

C. The proof of the theorem

Let us finish the paper with the proof of Theorem 1. We will need the following lemma.

Lemma 4. If (for every $1 \leq i \leq I$) $0 \leq \lambda_i \leq 1$, $|\psi_i\rangle$ and $|\phi\rangle$ are normalized and

$$\left\| \sum_{i=1}^{I} \lambda_i |\psi_i\rangle\langle\psi_i| - |\phi\rangle\langle\phi| \right\|_1 \leq \epsilon, \tag{4}$$

then there exists an i_0 such that $D(|\phi\rangle,|\psi_{i_0}\rangle) \leq \sqrt{2\epsilon}$.

Proof. Taking trace in (4) we have that $\lambda = \Sigma_i \lambda_i$ verifies $|\lambda - 1| \leq \epsilon$. By defining $\tilde{\lambda}_i = \frac{\lambda_i}{\lambda}$ we get that $\Sigma_i \tilde{\lambda}_i = 1$ and still

$$\begin{aligned} &\left\| \sum_i \tilde{\lambda}_i |\psi_i\rangle\langle\psi_i| - |\phi\rangle\langle\phi| \right\|_1 \\ &= \left\| \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| - |\phi\rangle\langle\phi| + \sum_i \lambda_i \left(\frac{1}{\lambda} - 1\right)|\psi_i\rangle\langle\psi_i| \right\|_1 \\ &\leq \epsilon + \sum_i \lambda_i \left| \frac{1}{\lambda} - 1 \right| = \epsilon + |1-\lambda| \leq 2\epsilon. \end{aligned}$$

Now

$$\begin{aligned} 1 - \sum_i \tilde{\lambda}_i |\langle\phi|\psi_i\rangle|^2 &= \left| \langle\phi| \left(\sum_i \tilde{\lambda}_i |\psi_i\rangle\langle\psi_i| - |\phi\rangle\langle\phi| \right) |\phi\rangle \right| \\ &\leq \left\| \sum_i \tilde{\lambda}_i |\psi_i\rangle\langle\psi_i| - |\phi\rangle\langle\phi| \right\|_\infty \leq 2\epsilon, \end{aligned}$$

which means that there exists an i_0 with $|\langle\phi|\psi_{i_0}\rangle|^2 \geq 1 - 2\epsilon$; and hence $D(|\phi\rangle,|\psi_{i_0}\rangle) \leq \sqrt{2\epsilon}$. ■

Now we want to choose a programmable POVM $(F^j)_{j=1}^d$ that approximates any observable $(|\phi^j\rangle\langle\phi^j|)_{j=1}^d$ with an error $\leq \delta$. This means that

$$d((G^j)_j,(|\phi^j\rangle\langle\phi^j|)_j) = \max_\rho \sum_j |\mathrm{tr}[\rho(G^j - |\phi^j\rangle\langle\phi^j|)]| \leq \delta,$$

where $G^j = \mathrm{tr}_a[(\mathbb{1} \otimes |\varphi\rangle\langle\varphi|)F^j]$ for some $|\varphi\rangle$ (by convexity it is enough to consider pure states in the ancilla).

This implies, in particular

$$\|G^1 - |\phi\rangle\langle\phi|\|_\infty \leq \max_\rho |\mathrm{tr}[\rho(G^1 - |\phi\rangle\langle\phi|)]| \leq \delta$$

for any arbitrary pure state $|\phi\rangle$ (where the ancilla $|\varphi\rangle$ that defines G^1 can depend on $|\phi\rangle$). Using that $\frac{1}{d}\|\cdot\|_1 \leq \|\cdot\|_\infty \leq \|\cdot\|_1$ we obtain $\|G^1 - |\phi\rangle\langle\phi|\|_1 \leq d\delta$ for every $|\phi\rangle$. That is, if we take a sequence $(|\phi_\alpha\rangle)_{\alpha=1}^A$ of pure states with property (2) (as in the statement of Theorem 1), we have that for every α there exists a $|\varphi_\alpha\rangle$ such that $\|\rho_\alpha - |\phi_\alpha\rangle\langle\phi_\alpha|\|_1 \leq d\delta$, where $\rho_\alpha = \mathrm{tr}_a[(\mathbb{1} \otimes |\varphi_\alpha\rangle\langle\varphi_\alpha|)F^1]$.

Now we take the spectral decomposition $F^1 = \Sigma_{i=1}^I \lambda_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$ ($0 \leq \lambda_i \leq 1$ and $|\tilde{\psi}_i\rangle$ normalized). Theorem 1 will be then proven if we can show that

$$A \leq I.$$

To see it let us fix an α. Calling $|\psi_i^\alpha\rangle = \langle\varphi_\alpha|\tilde{\psi}_i\rangle$ we have that $\rho_\alpha = \Sigma_i \lambda_i |\psi_i^\alpha\rangle\langle\psi_i^\alpha|$. By Lemma 4 (with $\epsilon = d\delta$) we have that there exists an i_α such that

$$D(|\phi_\alpha\rangle,|\psi_{i_\alpha}^\alpha\rangle) \leq \sqrt{2d\delta}. \tag{5}$$

Now, for $\alpha \neq \beta$

$$\begin{aligned} D(|\tilde{\psi}_{i_\alpha}\rangle,|\tilde{\psi}_{i_\beta}\rangle) &\geq D(|\phi_\alpha\rangle|\varphi_\alpha\rangle,|\phi_\beta\rangle|\varphi_\beta\rangle) - D(|\tilde{\psi}_{i_\alpha}\rangle,|\phi_\alpha\rangle|\varphi_\alpha\rangle) \\ &\quad - D(|\tilde{\psi}_{i_\beta}\rangle,|\phi_\beta\rangle|\varphi_\beta\rangle) > 0. \end{aligned}$$

To see it, it is enough to notice that, by (5), both $D(|\tilde{\psi}_{i_\alpha}\rangle,|\phi_\alpha\rangle|\varphi_\alpha\rangle)$ and $D(|\tilde{\psi}_{i_\beta}\rangle,|\phi_\beta\rangle|\varphi_\beta\rangle)$ are bounded by $\sqrt{2d\delta}$; and by (2), $D(|\phi_\alpha\rangle|\varphi_\alpha\rangle,|\phi_\beta\rangle|\varphi_\beta\rangle) > \sqrt{8d\delta}$.

Therefore, if $\alpha \neq \beta$, we have that $|\tilde{\psi}_{i_\alpha}\rangle \neq |\tilde{\psi}_{i_\beta}\rangle$, which means that $i_\alpha \neq i_\beta$; and hence $A \leq I$. QED.

IV. CONCLUSION

In conclusion, we have proven that the universal programmable measurements designed in [10] are optimal in the sense of the resources (dimension of the program space) needed to build them. This opens the door of a key question: how to physically implement them?

Another question that arises from this paper, apart from improving the constants (see Remark 3), is to fill in the gap between the lower and the upper bounds [10] found for the exponents in the general case: $\frac{d-1}{2} \leq \text{exponent} \leq d(d-1)$. Notice that in the case of qubits the optimal exponent 1 does not coincide with any of the general bounds.

As for the techniques, this paper shows once more the close connection between quantum information and the rich mathematical theory of convex bodies (other recent applications can be found for instance in [12–17] and the references therein). Our belief is that this connection will give much more in the near future.

ACKNOWLEDGMENTS

The author would like to thank M. M. Wolf, J. I. Cirac, and especially P. Perinotti and G. M. D'Ariano for valuable discussions concerning this paper. This work has been partially supported by Spanish Grant No. MTM-2005-0082.

[1] M. A. Nielsen and I. L. Chuang, Phys. Rev. Lett. **79**, 321 (1997).
[2] A. Y. Vlasov, e-print quant-ph/0103119.
[3] G. Vidal and J. I. Cirac, e-print quant-ph/0012067.
[4] M. Hillery, V. Buzek, and M. Ziman, Phys. Rev. A **65**, 022301 (2002).
[5] M. Hillery, M. Ziman, and V. Buzek, Phys. Rev. A **69**, 042311 (2004).
[6] M. Hillery, M. Ziman, and V. Buzek, e-print quant-ph/0510161.
[7] M. Dusek and V. Buzek, Phys. Rev. A **66**, 022112 (2002).
[8] J. Fiurasek, M. Dusek, and R. Filip, Phys. Rev. Lett. **89**, 190401 (2002).
[9] J. Fiurasek and M. Dusek, Phys. Rev. A **69**, 032302 (2004).
[10] G. M. D'Ariano and P. Perinotti, Phys. Rev. Lett. **94**, 090401 (2005).
[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[12] P. Hayden, D. W. Lueng, and A. Winter, e-print quant-ph/0407049.
[13] S. D. Bartlett, P. Hayden, and R. W. Spekkens, Phys. Rev. A **72**, 052329 (2005).
[14] P. Hayden, D. Leung, and G. Smith, Phys. Rev. A **71**, 062339 (2005).
[15] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371–391 (2004).
[16] S. J. Szarek, Phys. Rev. A **72**, 032304 (2005).
[17] S. J. Szarek, I. Bengtsson, and K. Zyczkowski, e-print quant-ph/0509008.