On the fidelity of two pure states

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 7095–7101 PII: S0305-4470(01)19132-8

On the fidelity of two pure states

Andreas Winter

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

E-mail: winter@mathematik.uni-bielefeld.de

Received 17 November 2000, in final form 6 April 2001
Published 24 August 2001
Online at stacks.iop.org/JPhysA/34/7095

Abstract

The *fidelity* of two pure states (also known as *transition probability*) is a symmetric function of two operators, and well founded operationally as an event probability in a certain preparation-test pair. Motivated by the idea that the fidelity is the continuous quantum extension of the combinatorial equality function, we enquire whether there exists a *symmetric operational* way of obtaining the fidelity. It is shown that this is impossible. Finally, we discuss the optimal universal approximation by a quantum operation.

PACS numbers: 03.65.Ta, 03.67.−a

1. Introduction

For two pure quantum states $\pi = |\varphi\rangle\langle\varphi|$ and $\tau = |\theta\rangle\langle\theta|$ on the space $\mathcal{H}$, which we assume throughout to be of dimension $d < \infty$, the (pure state) fidelity is

$$F(\pi, \tau) = \operatorname{Tr} \pi\tau = |\langle\varphi|\theta\rangle|^2.$$

Its operational justification is as follows: suppose we test the system for being in state τ, described by the projection valued measure (PVM) $(\tau, \mathbb{1} - \tau)$, then the probability of an affirmative answer, the actual preparation being π, is $F(\pi, \tau)$. It is one of the features of quantum theory that the same probability arises if the system is prepared in state τ, and is tested for π, see the discussion in chapter 2 of [1]. This is reflected in the symmetry of F: $F(\pi, \tau) = F(\tau, \pi)$.

By restricting attention to a set of orthonormal vectors $|x\rangle$, $x \in \mathcal{X}$, one has

$$F(|x\rangle\langle x|, |y\rangle\langle y|) = \delta_{xy} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}.$$

Thus, on $\mathcal{X} \times \mathcal{X}$, F represents the test for equality of two given elements from $\mathcal{X}$. Observe that this characterization is symmetric in the two variables: we can imagine a classical computing machine taking as input x and y from $\mathcal{X}$, which outputs $\delta_{xy} \in \{0, 1\}$.

2. The problem

The question arises whether or not an operational justification for F is possible that is symmetrical too, like the one just given for the identity-predicate (which it generalizes, as we observed). Note that in the above discussion one of π, τ figures as a state, whereas the other is a projection of a test. Hence, two possibilities seem natural: either both have to be given as quantum states, or both as tests. In either case we want to find a procedure to sample the binary distribution $(\mathrm{Tr}\,\pi\tau, 1 - \mathrm{Tr}\,\pi\tau)$ once, i.e. produce the first outcome with probability $\mathrm{Tr}\,\pi\tau$, and the second with probability $1 - \mathrm{Tr}\,\pi\tau$. These two problems will be made precise in the following subsections, and given answers.

2.1. Two states

A would-be fidelity estimator for two unknown states is a map

$$F : \pi \otimes \tau \mapsto (\mathrm{Tr}\,\pi\tau)z_1 + (1 - \mathrm{Tr}\,\pi\tau)z_0$$

where z_0, z_1 are the (orthogonal) idempotent generators of a two-dimensional commutative algebra[1]. As is immediate, this is indeed uniquely extendible to a trace preserving linear map on $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$. It is even positive—on the separable states! But not on the whole state space: for example, consider the pure state vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$$

in $\mathcal{H} \otimes \mathcal{H}$, with orthogonal unit vectors $|0\rangle$ and $|1\rangle$. Then it is easily checked that it must map to $(-1)z_1 + 2z_0$ under F.

In fact, it will turn out useful to write the map F in different terms: since it maps states linearly to binary probability distributions, it is a *test*, which means there is an operator F such that $\mathrm{Tr}\,\pi\tau = \mathrm{Tr}\,((\pi \otimes \tau)F)$, i.e.

$$F(\pi \otimes \tau) = \mathrm{Tr}\,((\pi \otimes \tau)F)z_1 + \mathrm{Tr}\,((\pi \otimes \tau)(\mathbb{1} - F))z_0$$

and it is well known that the unique solution to this equation is the flip operator $F : |\phi\rangle \otimes |\theta\rangle \mapsto |\theta\rangle \otimes |\phi\rangle$, which has eigenvalue 1 on the symmetric subspace, and -1 on the antisymmetric subspace; in particular it is not positive.

It is interesting to note that we encountered here what is called an *entanglement witness* (as introduced by Terhal [2]): a linear map positive on products but negative on certain entangled states which it 'certifies'. The operator $W = F^*(z_1)$ (using the dual map F^* of F with respect to the Hilbert–Schmidt inner product) is the operator version of this entanglement witness: it has the property

$$\mathrm{Tr}\,((\pi \otimes \tau)W) = \mathrm{Tr}\,\pi\tau \geqslant 0$$

but for some entangled states it has negative expected value.

Let us state our finding as a theorem:

Theorem 1. *There is no positive linear map F such that for all π, τ*

$$F(\pi \otimes \tau) = (\mathrm{Tr}\,\pi\tau)z_1 + (1 - \mathrm{Tr}\,\pi\tau)z_0.$$

Equivalently there is no positive operator F such that for all π, τ

$$\mathrm{Tr}\,\pi\tau = \mathrm{Tr}\,((\pi \otimes \tau)F).$$

[1] Note that the restriction to one copy of π, τ each is crucial: if we were allowed to use the preparation device for π, τ indefinitely often, then we could do a tomography (consult, for example, [3]) or any other statistical reconstruction of the states, and actually *compute* $\mathrm{Tr}\,\pi\tau$.

2.2. Two tests

Suppose we are given the PVM $M = (\pi, \mathbb{1} - \pi) \otimes (\tau, \mathbb{1} - \tau)$ on $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ as a black box. What we can do is feed it with an arbitrarily prepared state, and combine the outcomes into two groups. Observe that if we allow multiple uses of the black box we can do a tomography[2] of M (dual to the tomography of states [3]). This motivates the restriction to a single application of M[3]. Note that we make no assumptions other than statistical on the observables, in particular we impose no conditions on the measurement devices regarding their post-measurement states (see again the remark in footnote 3).

Preparing a state ρ on $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ and using it with M, we are supplied with one of four outcomes (11, 10, 01, 00), after which we employ a statistical decision rule: if ij was measured, we vote for 1 with probability $p_{ij} \in [0, 1]$. This is the most general form of the procedure, and we can calculate

$$\begin{aligned}\Pr\{1\} &= p_{11}\operatorname{Tr}(\rho(\pi \otimes \tau)) + p_{10}\operatorname{Tr}(\rho(\pi \otimes (\mathbb{1} - \tau))) + p_{01}\operatorname{Tr}(\rho((\mathbb{1} - \pi) \otimes \tau)) \\ &\quad + p_{00}\operatorname{Tr}(\rho((\mathbb{1} - \pi) \otimes (\mathbb{1} - \tau))) \\ &= (p_{11} - p_{10} - p_{01} + p_{00})\operatorname{Tr}(\rho(\pi \otimes \tau)) + (p_{10} - p_{00})\operatorname{Tr}(\rho(\pi \otimes \mathbb{1})) \\ &\quad + (p_{01} - p_{00})\operatorname{Tr}(\rho(\mathbb{1} \otimes \tau)) + p_{00}\operatorname{Tr}\rho \\ &= (p_{11} - p_{10} - p_{01} + p_{00})\operatorname{Tr}(\rho(\pi \otimes \tau)) + (p_{10} - p_{00})\operatorname{Tr}(\rho_1\pi) \\ &\quad + (p_{01} - p_{00})\operatorname{Tr}(\rho_2\tau) + p_{00}.\end{aligned}$$

This is a polynomial in π and τ with a bilinear, a linear, and a constant part. Hence, for this to be equal to $\operatorname{Tr}\pi\tau$, necessarily

$$p_{10} = p_{01} = p_{00} = 0$$

forcing $p_{11} = 1$ (choose $\pi = \tau$). So, we have to look for a state ρ satisfying

$$\operatorname{Tr}\pi\tau = \operatorname{Tr}(\rho(\pi \otimes \tau)).$$

However, by theorem 1 there does not even exist a solution $0 \leqslant \rho \leqslant \mathbb{1}$ to this equation.

Our result can be understood as another new feature of quantum information as compared to classical information: whereas there is an identity test for classical data, symmetrical in the two inputs, the corresponding natural quantum version, namely the fidelity, is forbidden by the quantum mechanical laws: not only are we unable to access the precise value of it, we cannot even once sample the corresponding Bernoulli variable.

Thus, we have exhibited a new no-go theorem regarding quantum mechanics, in the line of the no-cloning theorem [5].

[2] Since we want to reconstruct a set of selfadjoint operators $(A_1, \ldots, A_n)$ on $\mathcal{H}$, it is sufficient to know the values of $\operatorname{Tr}\rho A_i$ for a spanning set of states ρ. Formally this is the same as state tomography, as it rests on the Hermiticity of the Hilbert–Schmidt inner product $\operatorname{Tr} A^*B$ for operators.

[3] It may be amusing to note that in the case of two qubits (i.e. $\mathcal{H} = \mathbb{C}^2$) and with the promise that M obeys the projection postulate for the post-measurement states, *two* applications of M are sufficient to achieve the goal: M is a complete von Neumann measurement consisting of $\pi \otimes \tau, \pi \otimes \tau^\perp, \pi^\perp \otimes \tau, \pi^\perp \otimes \tau^\perp$.

Apply M once (on an arbitrary initial state), then swap the qubits (which is unitary), and apply M a second time. A suitable combination of the in-total 16 outcomes makes the outcome '1' occur with probability $\operatorname{Tr}\pi\tau$ (observe that this equals $\operatorname{Tr}\pi^\perp\tau^\perp$). More explicitly, denoting the outcomes as (X, Y), where X is the operator from the first measurement, and Y from the second, combine $(\pi \otimes T, \tau \otimes P)$ and $(\pi^\perp \otimes T, \tau^\perp \otimes P)$, where P and T run over the sets $\{\pi, \pi^\perp\}$, $\{\tau, \tau^\perp\}$, respectively, for the outcome '1'. (In fact, a single application of $(\pi, \mathbb{1} - \pi)$ followed by an application of $(\tau, \mathbb{1} - \tau)$ does the same job.)

The reason that we did not consider this to be allowed is that it assumes the projection postulate, apart from obviously working only with qubits.) This adds another peculiarity to 'why two qubits are special' [4].

3. Universal approximation for two states

After failing to find allowed procedures to sample the fidelity distribution $F(\pi, \tau)$, we resort to approximating this ideal behaviour in an optimal way.

To find the optimal approximation to the fidelity estimator, we have to minimize the expression

$$\delta(A) = \max_{\pi,\tau} |\operatorname{Tr}((\pi \otimes \tau)A) - \operatorname{Tr}\pi\tau|$$

with respect to $0 \leqslant A \leqslant \mathbb{1}$. We may assume that the optimal A is invariant under the actions

$$\pi \otimes \tau \longmapsto \tau \otimes \pi$$

and

$$\pi \otimes \tau \longmapsto U\pi U^\dagger \otimes U\tau U^\dagger \qquad U \in \mathcal{U}(\mathcal{H}).$$

The reasoning is the same as for universal cloning [6] and Bloch vector flipping [7] machines: because of the invariance of the fidelity function and triangle inequality, an optimal solution cannot become worse if we average it over the group action using the Haar measure.

Since the squared representation of the unitary group has exactly two irreducible components, the *symmetric* and the *antisymmetric* subspaces, $\mathcal{S}$ and $\mathcal{A}$, respectively, with corresponding projectors $\Pi_\mathcal{S}$ and $\Pi_\mathcal{A}$, the most general A to consider has the form

$$A = \sigma\Pi_\mathcal{S} + \alpha\Pi_\mathcal{A} \qquad 0 \leqslant \sigma, \alpha \leqslant 1.$$

To evaluate $\delta(A)$ choose an orthonormal basis $e_1, \ldots, e_d$ of $\mathcal{H}$. Then

$$\mathcal{S} = \operatorname{span}\left\{ f_i = e_i \otimes e_i,\ f_{ij} = \frac{e_i \otimes e_j + e_j \otimes e_i}{\sqrt{2}} : i < j \right\}$$

and note that the f_i, f_{ij} form an orthonormal basis of $\mathcal{S}$.

Now by unitary invariance we may assume that

$$\pi = |e_1\rangle\langle e_1|$$

and

$$\tau = (u|e_1\rangle + v|e_2\rangle)(u\langle e_1| + v\langle e_2|) \qquad u, v \geqslant 0,\ u^2 + v^2 = 1.$$

Hence, noting $\operatorname{Tr}\pi\tau = u^2$,

$$\begin{aligned}\delta(A) &= \max_{u,v} \left|\sigma \operatorname{Tr}((\pi \otimes \tau)\Pi_\mathcal{S}) + \alpha \operatorname{Tr}((\pi \otimes \tau)\Pi_\mathcal{A}) - u^2\right| \\ &= \max_{u,v} \left|\alpha + (\sigma - \alpha)\operatorname{Tr}((\pi \otimes \tau)\Pi_\mathcal{S}) - u^2\right|\end{aligned}$$

and calculating

$$\begin{aligned}\operatorname{Tr}((\pi \otimes \tau)\Pi_\mathcal{S}) &= \|\Pi_\mathcal{S}|e_1\rangle \otimes (u|e_1\rangle + v|e_2\rangle)\|_2^2 \\ &= |(\langle e_1| \otimes \langle e_1|)(|e_1\rangle \otimes (u|e_1\rangle + v|e_2\rangle))|^2 + |(\langle e_2| \otimes \langle e_2|)(|e_1\rangle \otimes (u|e_1\rangle + v|e_2\rangle))|^2 \\ &\quad + \left|\frac{\langle e_1| \otimes \langle e_2| + \langle e_2| \otimes \langle e_1|}{\sqrt{2}}(|e_1\rangle \otimes (u|e_1\rangle + v|e_2\rangle))\right|^2 \\ &= u^2 + 0 + \frac{v^2}{2} = \frac{1 + u^2}{2}\end{aligned}$$

we end up with

$$\begin{aligned}\delta(A) &= \max_{0\leqslant u^2\leqslant 1}\left|\alpha+(\sigma-\alpha)\frac{1+u^2}{2}-u^2\right| \\ &= \max_{0\leqslant x\leqslant 1}\left|\frac{\sigma+\alpha}{2}+\left(\frac{\sigma-\alpha}{2}-1\right)x\right| \\ &= \max\left\{\frac{\sigma+\alpha}{2},1-\sigma\right\}.\end{aligned}$$

To minimize this we have to choose $\alpha = 0$ and $\sigma = 2/3$. The optimal test is thus

$$A = \tfrac{2}{3}\Pi_{\mathcal{S}}$$

achieving $\delta(A) = \delta_{\min} = 1/3$.

The general case of n copies of the two states, and m samples to be produced, is discussed in the appendix.

In fact, from the proof of optimality we can derive a strengthening of our previous theorem 1: there is *no* operational quantum extension of the classical identity test at all:

Theorem 2. *There is no test T on $\mathcal{H}\otimes\mathcal{H}$ (i.e. $0\leqslant T\leqslant \mathbb{1}\otimes\mathbb{1}$), such that for all states π, τ on $\mathcal{H}$*

$$\begin{aligned}\tau = \pi &\Rightarrow \operatorname{Tr}((\pi\otimes\tau)T) = 1 \\ \tau \perp \pi &\Rightarrow \operatorname{Tr}((\pi\otimes\tau)T) = 0.\end{aligned}$$

Proof is by observing that the maximum defining $\delta(A)$ is achieved for $u^2 = \operatorname{Tr}\pi\tau \in \{0,1\}$ in the above calculation.

Complementing this result, note that we *can*, however, obtain *partial* information on the fidelity. For example, the optimal test $T = \frac{2}{3}\Pi_{\mathcal{S}}$ has the property that

$$\begin{aligned}\operatorname{Tr}((\pi\otimes\tau)T) > \tfrac{1}{2} \qquad &\text{iff} \quad \operatorname{Tr}\pi\tau > \tfrac{1}{2} \\ \operatorname{Tr}((\pi\otimes\tau)T) < \tfrac{1}{2} \qquad &\text{iff} \quad \operatorname{Tr}\pi\tau < \tfrac{1}{2}.\end{aligned}$$

4. Summary

We have argued that the fidelity of pure states is the quantum generalization of the classical identity-predicate δ_{xy}, and have shown that an operational basis for it, similar to the classical way, does not exist. Indeed, there does not exist any quantum operation behaving like δ_{xy} on all orthogonal sets of states. Finally, we have discussed the univeral optimal approximation to the fidelity function, in the simplest case.

Acknowledgements

I would like to thank Cameron Wellard and Serge Massar for conversations on the present paper's subject, during the QCM&C 2000 at Capri. This research was supported by SFB 343 'Diskrete Strukturen in der Mathematik' of the Deutsche Forschungsgemeinschaft.

Appendix. The general case

In this appendix we demonstrate a possible attack on the general case. Unfortunately we find the final optimization problem so hard to solve that we leave the solution open.

Given n copies of each state we want to produce as close an approximation to m samples of $F(\pi,\tau)=(\operatorname{Tr}\pi\tau, 1-\operatorname{Tr}\pi\tau)$ as possible, i.e. a POVM $\mathbf{A}$ indexed by $\{0,1\}^m$ which minimizes

$$\delta(\mathbf{A})=\max_{\pi,\tau}\|\mathbf{A}(\pi^{\otimes n}\otimes\tau^{\otimes n})-F(\pi,\tau)^{\otimes m}\|_1$$

where we write $\mathbf{A}\ (\pi^{\otimes n}\otimes\tau^{\otimes n})$ for the distribution on $\{0,1\}^m$ induced by measuring $\mathbf{A}$ on $\pi^{\otimes n}\otimes\tau^{\otimes n}$. Obviously we can assume that $\mathbf{A}$ is supported on $\mathcal{H}_+^n\otimes\mathcal{H}_+^n$, where $\mathcal{H}_+^n$ is the symmetric subspace in $\mathcal{H}^{\otimes n}$, i.e. the set of all vectors invariant under tensor factor permutation.

By the familiar averaging argument we can assume that all elements of $\mathbf{A}$ are invariant under the action of the unitary group $\mathcal{U}(\mathcal{H})$. This action decomposes $\mathcal{H}_+^n\otimes\mathcal{H}_+^n$ into $n+1$ orthogonal subspaces $\mathcal{S}_l$: the restriction to $\mathcal{S}_l$ is irreducible with highest weight $(2n-l,l,0,\ldots,0), l=0,\ldots,n$. In particular, they all have multiplicity one (for these representation theoretical details we refer the reader to [8]). Denote the subspace projection onto $\mathcal{S}_l$ by S_l.

Since $F(\pi,\tau)^{\otimes m}$ has the constant value $(\operatorname{Tr}\pi\tau)^k(1-\operatorname{Tr}\pi\tau)^{m-k}$ on the sets

$$\mathcal{T}_k=\{x^m\in\{0,1\}^m : x^m \text{ has exactly } k \text{ 0's}\}$$

we may assume that an optimal $\mathbf{A}$ is constant on the $\mathcal{T}_k$ as well. Introducing the angle γ between $|\phi\rangle$ and $|\theta\rangle$, so that $\operatorname{Tr}\pi\tau=\cos^2\gamma$ and $1-\operatorname{Tr}\pi\tau=\sin^2\gamma$, we can define

$$f_k=\mathbf{A}(\pi^{\otimes n}\otimes\tau^{\otimes n})(\mathcal{T}_k)$$
$$p_k=F(\pi,\tau)^{\otimes m}(\mathcal{T}_k)=\binom{m}{k}(\cos^2\gamma)^k(\sin^2\gamma)^{m-k}$$

and thus write

$$\|\mathbf{A}(\pi^{\otimes n}\otimes\tau^{\otimes n})-F(\pi,\tau)^{\otimes m}\|_1=\sum_{k=0}^m|f_k-p_k|.$$

Observe that with

$$F_k=\mathbf{A}^*\big(1_{\mathcal{T}_k}\big)=\sum_{x^m\in\mathcal{T}_k}A_{x^m}$$

one has $f_k=\operatorname{Tr}\left((\pi^{\otimes n}\otimes\tau^{\otimes n})F_k\right)$.

By invariance we can write

$$F_k=\sum_{l=0}^n\alpha_{kl}S_l \quad \text{with } \alpha_{kl}\geqslant 0 \qquad \sum_{k=0}^m\alpha_{kl}=1.$$

Now, applying invariance once more, we get

$$f_k=\operatorname{Tr}\left(\left(\int_{\mathcal{U}(d)}\mathrm{d}\,U(U^{\otimes 2n}\pi^{\otimes n}\otimes\tau^{\otimes n}U^{*\otimes 2n})\right)F_k\right).$$

The integral itself is an invariant state, hence of the form

$$\sum_{l=0}^n\beta_l\frac{1}{\operatorname{Tr}S_l}S_l \quad \text{with } \beta_l\geqslant 0 \qquad \sum_{l=0}^n\beta_l=1$$

and by invariance—third time pays for all—the β_l depend solely on $\operatorname{Tr}\pi\tau$. In fact, it is easily seen that they all are universal homogeneous polynomials in $\cos\gamma$ and $\sin\gamma$ of total degree $2n$.

This makes it seem rather unlikely that we can find

$$\delta(\mathbf{A})=\max_{0\leqslant\gamma\leqslant\pi/2}\sum_{k=0}^m\left|\binom{m}{k}(\cos^2\gamma)^k(\sin^2\gamma)^{m-k}-\sum_{l=0}^n\alpha_{kl}\beta_l(\cos\gamma,\sin\gamma)\right|$$

let alone minimize this over the α_{kl}.

References

[1] Peres A 1993 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer)

[2] Terhal B M 2000 A family of indecomposable positive linear maps based on entangled quantum states *Linear Algebra Appl.* **323** 61–73
(Terhal B M 1998 *Preprint* quant-ph/9810091)
Terhal B M 2000 Bell inequalities and the separability criterion *Phys. Lett.* A **271** 319–26

[3] D'Ariano G M, Maccone L and Paris M G A 2000 Orthogonality relations in quantum tomography *Phys. Lett.* A **276** 25–30

[4] Vollbrecht K G H and Werner R F 2000 Why two qubits are special *J. Math. Phys.* **41** 6772–82
(Vollbrecht K G H and Werner R F 1999 *Preprint* quant-ph/9910064)

[5] Wootters W K and Zurek W H 1982 A single quantum cannot be cloned *Nature* **299** 802–3

[6] Werner R F 1998 Optimal cloning of pure states *Phys. Rev.* A **58** 1827–32 and references therein

[7] Bužek V, Hillery M and Werner R F 1999 Optimal manipulations with qubits: universal-NOT gate *Phys. Rev.* A **60** R2626–9

[8] Zhelobenko D P 1973 *Compact Lie Groups and Their Representations* (*Translations of Mathematical Monographs* 40) (Providence, RI: American Mathematical Society)