On the fidelity of two pure states

Andreas Winter*

November 13, 2000

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.

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 [1], chapter 2. 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

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

*Address: SFB 343, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D–33501 Bielefeld, Germany. Email: `winter@mathematik.uni-bielefeld.de`. Research supported by SFB 343 “Diskrete Strukturen in der Mathematik” of the Deutsche Forschungsgemeinschaft.

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. Note that in the above discussion one of π, τ figures as a state, whereas the other as 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 $1 - \mathrm{Tr}\,\pi\tau$.

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 $\mathfrak{L}(\mathcal{H}) \otimes \mathfrak{L}(\mathcal{H})$. It is even positive — on the separable states! But not on the whole state space: for example consider a pure state vector

$$|\psi\rangle = \alpha|e_0 f_0\rangle + \beta|e_1 f_1\rangle$$

in $\mathcal{H} \otimes \mathcal{H}$, with unit vectors $e_0 \perp e_1$, $f_0 \perp f_1$, and (w.l.o.g.) $\alpha, \beta > 0$ such that $\alpha^2 + \beta^2 = 1$. Then

$$\begin{aligned}|\psi\rangle\langle\psi| = \alpha^2 &|e_0\rangle\langle e_0| \otimes |f_0\rangle\langle f_0| + \beta^2|e_1\rangle\langle e_1| \otimes |f_1\rangle\langle f_1| \\ &+ \alpha\beta|e_0\rangle\langle e_1| \otimes |f_0\rangle\langle f_1| + \alpha\beta|e_1\rangle\langle e_0| \otimes |f_1\rangle\langle f_0|.\end{aligned}$$

Note that

$$\begin{aligned}|\langle e_0|f_0\rangle|^2 &= p = |\langle e_1|f_1\rangle|^2, \\ |\langle e_0|f_1\rangle|^2 &= q = 1 - p = |\langle e_1|f_0\rangle|^2.\end{aligned}$$

Finally, introducing

$$\langle e_0|f_1\rangle = e^{i\gamma}\sqrt{q}, \quad \langle e_1|f_0\rangle = e^{i\delta}\sqrt{q},$$

[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, we can do a tomography [3] of the states, and actually *compute* $\mathrm{Tr}\,\pi\tau$.

we can calculate the 1–component of $F(|\psi\rangle\langle\psi|)$:

$$\begin{aligned}F(|\psi\rangle\langle\psi|)_1 &= \alpha^2 p + \beta^2 p + \alpha\beta\langle e_1|f_0\rangle\langle f_1|e_0\rangle + \alpha\beta\langle e_0|f_1\rangle\langle f_0|e_1\rangle \\ &= p + 2q\alpha\beta\cos(\gamma-\delta),\end{aligned}$$

which may obviously be negative, e.g. $F(|\psi\rangle\langle\psi|)_1 = -1$ for $p=0$, $q=1$, $\alpha=\beta=1/\sqrt{2}$, and $\gamma-\delta=\pi$.

It is interesting to note that we encountered here what is called a *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 trace pairing) is the operator version of this entanglement witness: it has the property

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

but for some entangled states it has negative expected value. One can write out W explicitely:

$$W = \sum_s X_s^* \otimes X_s$$

for any orthonormal basis $X_1, \dots, X_{d^2}$ of $\mathfrak{L}(\mathcal{H})$.

2.2 Two tests

Suppose we are given the PVM $M = (\pi, \mathbb{1}-\pi)\otimes(\tau, \mathbb{1}-\tau)$ on $\mathfrak{L}(\mathcal{H})\otimes\mathfrak{L}(\mathcal{H})$ as a block 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 of M [4] (dual to the tomography of states [3]). This motivates the restriction to a single application of M [5].

Preparing a state ρ on $\mathfrak{L}(\mathcal{H})\otimes\mathfrak{L}(\mathcal{H})$ 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

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

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

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

However, by subsection 2.1 there does not even exist a solution $0 \leq \rho \leq \mathbb{1}$ 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.

In fact, the proof of the following section 3 shows that there is *no* operational quantum extension of the classical identity test at all:

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

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

Thus, we have exhibited a new no–go theorem regarding quantum mechanics, in the line of the no–cloning theorem [6].

3 Universal approximation for two states

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

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

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

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

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

and

$$\pi \otimes \tau \longmapsto U\pi U^* \otimes U\tau U^*, \quad U \in \mathcal{U}(\mathcal{H}).$$

The reasoning is the same as for universal cloning [7] and Bloch vector flipping [8] machines: because of invariance of the fidelity function and triangle inequality, an optimal solution cannot become worse if we average it over the group action using Haar measure.

Since the squared representation of the unitary group has exactly 2 irreducible components, the *symmetric* and the *antisymmetric* subspace, $\mathcal{S}$ and $\mathcal{A}$, respectively, with corresponding projectors $\Pi_{\mathcal{S}}$ and $\Pi_{\mathcal{A}}$, the most general A to consider has the form

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

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

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

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

Now by unitary invariance we may assume that

$$\begin{aligned}
&\pi = |e_1\rangle\langle e_1| \text{ and} \\
&\tau = (u|e_1\rangle + v|e_2\rangle)\,(u\langle e_1| + v\langle e_2|)\,,\ u, v \geq 0,\ u^2 + v^2 = 1.
\end{aligned}$$

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

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

and calculating

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

we end up with

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

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

$$A = \frac{2}{3}\Pi_{\mathcal{S}},$$

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.

Complementing theorem 1 above, note that we *can* obtain *partial* information on the fidelity. For example, the optimal test $T = 2/3\Pi_{\mathcal{S}}$ has the property that

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

4 Summary

We have argued that the fidelity of pure states is the quantum generalization of the classical identity–predicate δ_{xy}, and showed 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 an orthogonal set of states. Finally, we 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.

A 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

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

where we write $\mathbf{A}\left(\pi^{\otimes n} \otimes \tau^{\otimes n}\right)$ 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 $(2n-l, l, 0, \dots, 0)$, $l = 0, \dots, n$. In particular, they all have multiplicity one (for these representation theoretical details we refer the reader to [9]). Denote the subspace projection onto $\mathcal{S}_l$ by S_l.

Since $F(\pi,\tau)^{\otimes m}$ has the constant value $(\mathrm{Tr}\,\pi\tau)^k(1-\mathrm{Tr}\,\pi\tau)^{m-k}$ 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 $\mathrm{Tr}\,\pi\tau = \cos^2\gamma$ and $1-\mathrm{Tr}\,\pi\tau = \sin^2\gamma$, we can define

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

and thus write

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

Observe that with

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

one has $f_k = \mathrm{Tr}\,((\pi^{\otimes n} \otimes \tau^{\otimes n})F_k)$.

By invariance we can write

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

Now, applying invariance once more, we get

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

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

$$\sum_{l=0}^{n} \beta_l \frac{1}{\mathrm{Tr}\,S_l} S_l, \text{ with } \beta_l \geq 0, \quad \sum_{l=0}^{n} \beta_l = 1,$$

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

This makes it seem rather unlikely that we can find

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

let alone minimize this over the α_{kl}.

References

[1] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer, 1993.

[2] B. M. Terhal, "A Family of Indecomposable Positive Linear Maps based on Entangled Quantum States", e–print `quant-ph/9810091`. B. M. Terhal, "Bell Inequalities and the Separability Criterion", Phys. Lett. A, vol. 271, 2000, pp. 319–326.

[3] Consult for example: G. M. D'Ariano, L. Maccone, M. G. A. Paris, "Orthogonality relations in Quantum Tomography", e–print `quant-ph/0005111`.

[4] 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 $\mathrm{Tr}\,\rho A_i$ 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 $\mathrm{Tr}\,A^*B$ for operators.

[5] 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. Suitable combination of the in total 16 outcomes makes occur the outcome [1] with probability $\mathrm{Tr}\,\pi\tau$ (observe that this equals $\mathrm{Tr}\,\pi^\perp\tau^\perp$). [In fact a single application of $(\pi, \mathbb{1} - \pi)$ followed by an application of $(\tau, \mathbb{1} - \tau)$ does the same job.]

This adds another peculiarity to "why two qubits are special" (K. G. H. Vollbrecht, R. F. Werner, e–print `quant-ph/9910064`).

[6] W. K. Wootters, W. H. Zurek, "A single quantum cannot be cloned", Nature, vol. 299, 1982, pp. 802–803.

[7] R. F. Werner, "Optimal cloning of pure states", Phys. Rev. A, vol. 58, no. 3, 1998, pp. 1827–1832, and references therein.

[8] V. Bužek, M. Hillery. R. F. Werner, "Optimal manipulations with qubits: Universal–NOT gate", Phys. Rev. A, vol. 60, no. 4, 1999, pp. R2626–R2629.

[9] D. P. Zhelobenko, *Compact Lie Groups and Their Representations*, vol. 40 of *Translations of Mathematical Monographs*, AMS, Providence, RI, 1973.