SUB– AND SUPER–FIDELITY AS BOUNDS FOR QUANTUM FIDELITY

JAROSŁAW ADAM MISZCZAK   ZBIGNIEW PUCHAŁA

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences,
Bałtycka 5, 44-100 Gliwice, Poland

PAWEŁ HORODECKI

Faculty of Applied Physics and Mathematics, Gdańsk University of Technology,
Narutowicza 11/12, 80-952 Gdańsk, Poland
and
National Quantum Information Centre of Gdańsk,
Andersa 27, 81-824 Sopot, Poland

ARMIN UHLMANN

Institute of Theoretical Physics, University of Leipzig,
Vor dem Hospitaltore 1, D-04103 Leipzig, Germany

KAROL ŻYCZKOWSKI

Instytut Fizyki im. Smoluchowskiego, Uniwersytet Jagielloński,
Reymonta 4, 30-059 Kraków, Poland
and
Centrum Fizyki Teoretycznej, Polska Akademia Nauk,
Aleja Lotników 32/44, 02-668 Warszawa, Poland

We derive several bounds on fidelity between quantum states. In particular we show that fidelity is bounded from above by a simple to compute quantity we call super–fidelity. It is analogous to another quantity called sub–fidelity. For any two states of a two–dimensional quantum system ($N = 2$) all three quantities coincide. We demonstrate that sub– and super–fidelity are concave functions. We also show that super–fidelity is super–multiplicative while sub–fidelity is sub–multiplicative and design feasible schemes to measure these quantities in an experiment. Super–fidelity can be used to define a distance between quantum states. With respect to this metric the set of quantum states forms a part of a $N^2 - 1$ dimensional hypersphere.

1   Introduction

By processing quantum information we wish to transform a quantum state in a controlled way. Taking into account inevitable interaction with an environment and possible imperfection of real dynamics it is then crucial to characterize quantitatively, to what extend a given quantum state gets close to its target. For this purpose one often uses *fidelity* [1], here denoted by F. That quantity has also been called *transition probability* [2]: Operationally it is the maximal success probability of changing a state to another one by a measurement in a larger quantum

system. If both quantum states are pure, fidelity is the squared overlap between them.

In the general case fidelity between any two mixed states is the function of the trace norm of the product of their square roots. Thus analytical evaluation of fidelity, or its direct experimental measurement becomes a cumbersome task. Hence there is a need for other quantities, which bound fidelity and are easier to compute and measure.

The aim of this work is to present some bounds for fidelity and to develop experimental schemes to estimate it for an arbitrary pair of mixed quantum states. In particular we find an upper bound for fidelity by a simple quantity which is the function of purity of both states and the trace of their product. Since it possesses some nice algebraic properties we believe it may become useful in future research and propose to call it *super–fidelity*. In a sense it is a quantity complementary to the one forming the lower bound proved in [3], and we tend to call *sub–fidelity*. For any two one–qubit states all three quantities coincide. Fidelity is well known to be multiplicative with respect to the tensor product. In this work we prove that super–fidelity is concave and super–multiplicative, while sub–fidelity is concave and sub–multiplicative.

Fidelity can be used to define the Bures distance between quantum states and the Bures angle. As shown by Uhlmann in [4] the Bures geometry of the set of one–qubit states ($N = 2$), is equivalent to a three-dimensional hemisphere $\frac{1}{2}S^3$. The set of density operators, Ω_N, becomes the space of non-constant curvature by the Bures metric for $N \geq 3$, [5].

We construct distance and angle analogous to the Bures distance out of super–fidelity in a similar way. With respect to this metric the set Ω_N forms a fragment of a $N^2 - 1$ dimensional hypersphere with the maximally mixed state $\rho_* := \mathbb{I}/N$ at the pole. A linear function of super–fidelity was earlier used by Chen *et al.* [6] to analyze the set of mixed quantum states and demonstrate its hyperbolic interpretation.

This paper is organized as follows. In Section II the definition and basic properties of fidelity are reviewed. Sections III and IV are devoted to bounds on fidelity. In Section V we define sup– and super–fidelity and investigate their properties. Experimental schemes designed to measure these quantities are presented in section VI. In Section VII we analyze the geometry of the set of quantum states induced by the distance derived by super–fidelity. Concluding remarks are followed by appendices, in which we prove necessary lemmas and present the collection of useful algebraic facts.

2   Fidelity between quantum states

Consider an N– dimensional Hilbert space $\mathcal{H}_N$. A linear operator $\rho : \mathcal{H}_N \to \mathcal{H}_N$ represents quantum state if it is Hermitian, semipositive, $\rho = \rho^\dagger \geq 0$, and normalized, $\mathrm{tr}\rho = 1$. Let Ω_N denote the set of all mixed quantum states of size N.

Fidelity between quantum states ρ_1 and ρ_2 is defined as [2, 1],

$$F(\rho_1, \rho_2) = (\mathrm{tr}|\sqrt{\rho_1}\sqrt{\rho_2}|)^2 = ||\rho_1^{1/2}\rho_2^{1/2}||_1^2, \tag{1}$$

where $||\cdot||_1$ is Schatten 1-norm (trace norm),

$$||A||_1 = \mathrm{tr}|A| := \mathrm{tr}\sqrt{AA^\dagger}. \tag{2}$$

Alternatively, the trace norm of an operator can be expressed as the sum of its singular values, $||A||_1 = \sum_{i=1}^n \sigma_i(A)$. Here $\sigma_i(A)$ is equal to the square root of the corresponding eigenvalue of the positive matrix $AA^\dagger$ – see e.g. [7].

There are different uses of the name *fidelity*. In [1] the older notion *transition probability* has been renamed *fidelity* by Jozsa. In [8] $\sqrt{F}$ has been called *fidelity*, while [9] uses Jozsa's notion, and to the latter convention we shall stick in calling fidelity the expression in Eq. (1).

For pure states the definition (1) is reduced to the transition probability. If one state is pure, $\rho_1 = |\psi\rangle\langle\psi|$, then $F(\rho_1, \rho_2) = \langle\psi|\rho_2|\psi\rangle$. Hence for any two pure states their fidelity is equal to their squared overlap, $F(\psi, \phi) = |\langle\psi|\phi\rangle|^2 =: \kappa$.

Fidelity enjoys several important properties [2, 10, 11, 12, 1], which can also be proved on state spaces of unital C*-algebras. Some of them are:

i) **Bounds:** $0 \leq F(\rho_1, \rho_2) \leq 1$. Furthermore $F(\rho_1, \rho_2) = 1$ iff $\rho_1 = \rho_2$, while $F(\rho_1, \rho_2) = 0$ iff $\text{supp}(\rho_1) \perp \text{supp}(\rho_2)$.

ii) **Symmetry:** $F(\rho_1, \rho_2) = F(\rho_2, \rho_1)$.

iii) **Unitary invariance:** $F(\rho_1, \rho_2) = F(U\rho_1 U^\dagger, U\rho_2 U^\dagger)$, for any unitary operator U.

iv) **Concavity:** $F(\rho, a\rho_1 + (1-a)\rho_2) \geq aF(\rho, \rho_1) + (1-a)F(\rho, \rho_2)$, for $a \in [0, 1]$.

v) **Multiplicativity:** $F(\rho_1 \otimes \rho_2, \rho_3 \otimes \rho_4) = F(\rho_1, \rho_3)F(\rho_2, \rho_4)$.

vi) **Joint concavity:** $\sqrt{F}(a\rho_1 + (1-a)\rho_2, a\rho_1' + (1-a)\rho_2') \geq a\sqrt{F}(\rho_1, \rho_1') + (1-a)\sqrt{F}(\rho_2, \rho_2')$, for $a \in [0, 1]$.

For further analysis of fidelity properties it is instructive to work with eigenvalues of a matrix $\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}$. Let us denote them by λ_i, $i = 1, \ldots, N$. This matrix is positive so its eigenvalues and singular values coincide. Unless otherwise stated, we tacitly assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$. The root fidelity reads

$$\sqrt{F(\rho_1, \rho_2)} = \text{tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} = \sum_{i=1}^{N} \lambda_i. \tag{3}$$

Squaring this equation one obtains a compact expression for fidelity,

$$F(\rho_1, \rho_2) = \left(\sum_{i=1}^{N} \lambda_i\right)^2 = \text{tr}\rho_1\rho_2 + 2\sum_{i<j} \lambda_i\lambda_j, \tag{4}$$

where we have taken into account that $\text{tr}\rho_1\rho_2 = \text{tr}\sqrt{\rho_1}\rho_2\sqrt{\rho_1} = \sum_{i=1}^{N} \lambda_i^2$. The matrix $\sqrt{\rho_1}\rho_2\sqrt{\rho_1}$ is similar to $\rho_1\rho_2$ and they share the same set of N eigenvalues.

3   Bounds for fidelity

We shall need some further algebraic definitions. For any matrix X of size N with a set of eigenvalues $\{\lambda_1, \ldots, \lambda_N\}$ we define elementary symmetric functions $s_m(X)$ as the elementary symmetric function of its eigenvalues [17, Def. 1.2.9]. For instance, the second and third elementary symmetric functions read

$$s_2(X) = \sum_{i<j} \lambda_i\lambda_j, \tag{5}$$

$$s_3(X) = \sum_{i<j<k} \lambda_i\lambda_j\lambda_k. \tag{6}$$

For any matrix of rank r the highest non-vanishing symmetric function reads $s_r(X) = \prod_{i=1}^{r} \lambda_i$. In the generic case $r = N$ we have $s_N(X) = \det(X)$.

In this section we shall list several bounds for fidelity, some of which are well known in the literature. Let us start by stating a simple result,

$$F(\rho_1, \rho_2) \leq \mathrm{tr}\rho_1 \mathrm{tr}\rho_2, \tag{7}$$

which follows directly from Fact 1 (see Appendix A) if we set $\nu = 1/2$. This fact implies the property $F(\rho_1, \rho_2) \leq 1$.

Expression (4) implies the following lower bound

$$\mathrm{tr}\rho_1\rho_2 \leq F(\rho_1, \rho_2) \leq N\mathrm{tr}|\rho_1\rho_2|. \tag{8}$$

To get the upper bound we use Fact 2 (see Appendix A) and set $\nu = 1/2$ to obtain $||\sqrt{\rho_1}\sqrt{\rho_2}||_1^2 \leq N||\rho_1\rho_2||_1$.

Let us now denote the spectra of the states ρ_1 and ρ_2, by vectors $\vec{p}$ and $\vec{q}$, respectively. The fidelity between them is then bounded by the classical fidelity between diagonal density matrices [13]

$$F(p^{\uparrow}, q^{\downarrow}) \leq F(\rho_1, \rho_2) \leq F(p^{\uparrow}, q^{\uparrow}), \tag{9}$$

where the arrows up (down) indicate that the eigenvalues are put in the nondecreasing (nonincreasing) order.

The lower bound in (8) can be improved, since the following result is true [3]

$$F(\rho_1, \rho_2) \geq \mathrm{tr}\rho_1\rho_2 + \sqrt{2}\sqrt{(\mathrm{tr}\rho_1\rho_2)^2 - \mathrm{tr}\rho_1\rho_2\rho_1\rho_2}. \tag{10}$$

The above inequality is saturated for any pair of one–qubit states. Furthermore, the above inequality is an equality if the rank of $\rho_1\rho_2$ does not exceed two. On the other hand, the inequality is strict if that rank is larger than two — see Appendix E. For completeness we present the simple proof of inequality (10) in Appendix B.

Another lower bound is obtained if the rank of $\rho_1\rho_2$ is exactly r. If s_r denotes the rth elementary symmetric function then

$$F(\rho_1, \rho_2) \geq \mathrm{tr}\rho_1\rho_2 + r(r-1)\sqrt[r]{s_r(\rho_1\rho_2)}. \tag{11}$$

This bound is proved in Appendix C. If both states are generic, i.e. if they are of the maximal rank the above formula reads

$$F(\rho_1, \rho_2) \geq \mathrm{tr}\rho_1\rho_2 + N(N-1)\sqrt[N]{\det\rho_1\det\rho_2}. \tag{12}$$

The key result of this paper consist in the following upper bound, in a sense complementary to (10).

Theorem 1 *For any density matrices ρ_1 and ρ_2 we have*

$$F(\rho_1, \rho_2) \leq \mathrm{tr}\rho_1\rho_2 + \sqrt{(1 - \mathrm{tr}\rho_1^2)(1 - \mathrm{tr}\rho_2^2)}. \tag{13}$$

Before presenting the proof in the subsequent section let us first note that the bound is saturated if at least one of the states is pure. Furthermore, an equality holds for any two mixed states of size $N = 2$. To show this property observe that in this case the sum in (4) consists of a single term $2\lambda_1\lambda_2 = 2\sqrt{\det(\rho_1\rho_2)} = \sqrt{2\det(\rho_1)}\sqrt{2\det(\rho_2)}$. Since for any one-qubit state one has $2\det(\rho) = 1 - \mathrm{tr}\rho^2$ an equality in (13) follows. This fact was already known to Hübner [14]. In a similar way we treat the more general case of $N = 3$ in Appendix E, for which some other equations for fidelity are derived.

4 Proof of the main upper bound

The notion of the second symmetric function (5) allows us to write the expression

$$[(\mathrm{tr}X)^2 - (\mathrm{tr}X^2)] = 2s_2(X). \tag{14}$$

Note that if X has nonnegative eigenvalues then $s_2(X) \geq 0$.

Using (14) we can rewrite fidelity

$$F(\rho_1, \rho_2) = \mathrm{tr}\rho_1\rho_2 + 2s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right), \tag{15}$$

and

$$\sqrt{(1 - \mathrm{tr}\rho_1^2)(1 - \mathrm{tr}\rho_2^2)} = 2\sqrt{s_2(\rho_1)s_2(\rho_2)}. \tag{16}$$

Thus the Theorem 1 can be equally expressed as an inequality

$$s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) \leq \sqrt{s_2(\rho_1)s_2(\rho_2)}. \tag{17}$$

The proof of (17) is decomposed into two Lemmas, the proof of which can be found in Appendix D.

Lemma 1 *For given density matrices ρ_1, ρ_2 with eigenvalues $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ respectively*

$$s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) \leq s_2\left(\sqrt{\mathrm{diag}(p)\mathrm{diag}(q)}\right), \tag{18}$$

where $\mathrm{diag}(p)$ and $\mathrm{diag}(q)$ denote diagonal matrices with entries on diagonal $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ respectively.

Lemma 2 *With notation as in Lemma 1, we have*

$$s_2\left(\sqrt{\mathrm{diag}(p)\mathrm{diag}(q)}\right) \leq \sqrt{s_2(\mathrm{diag}(p))s_2(\mathrm{diag}(q))} = \sqrt{s_2(\rho_1)s_2(\rho_2)}. \tag{19}$$

Proof of Theorem 1. For given density matrices ρ_1 and ρ_2 with eigenvalues $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ respectively. We denote diagonal matrices with entries on diagonal $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ as $\mathrm{diag}(p)$ and $\mathrm{diag}(q)$ respectively.

$$
\begin{aligned}
F(\rho_1, \rho_2) &= \mathrm{tr}\rho_1\rho_2 + 2s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) \\
&\leq \mathrm{tr}\rho_1\rho_2 + 2s_2\left(\sqrt{\mathrm{diag}(p)\mathrm{diag}(q)}\right) \\
&\leq \mathrm{tr}\rho_1\rho_2 + 2\sqrt{s_2(\rho_1)s_2(\rho_2)} = \mathrm{tr}\rho_1\rho_2 + \sqrt{(1 - \mathrm{tr}\rho_1^2)(1 - \mathrm{tr}\rho_2^2)}.
\end{aligned}
$$

Making use of Lemma 1 for the first inequality and of Lemma 2 for the second one we arrive at the inequality (13). □

5   Sub– and super–fidelity and their properties

5.1   Definition and basic facts

We shall start this section with a general **definition**. For any two hermitian operators A and B let us define two quantities

$$
\begin{aligned}
E(A,B) &= \mathrm{tr}AB + \sqrt{2[(\mathrm{tr}AB)^2 - \mathrm{tr}ABAB]}, &\quad (20)\\
G(A,B) &= \mathrm{tr}AB + \sqrt{(\mathrm{tr}A)^2 - \mathrm{tr}A^2}\sqrt{(\mathrm{tr}B)^2 - \mathrm{tr}B^2}. &\quad (21)
\end{aligned}
$$

For any two density operators their traces are equal to unity, so $E(\rho_1, \rho_2)$ and $G(\rho_1, \rho_2)$ have lower bound (10) and upper bound (13), respectively. Thus both universal bounds for fidelity can be rewritten as

$$
E(\rho_1, \rho_2) \le F(\rho_1, \rho_2) \le G(\rho_1, \rho_2). \quad (22)
$$

Note that both bounds require the evaluation of three traces only, so they are easier to compute than the original fidelity. As shown in Section 3 for $N = 2$ all three quantities are equal, so we propose to call $E(\rho_1, \rho_2)$ and $G(\rho_1, \rho_2)$ as *sub–* and *super–fidelity*. These names are additionally motivated by the following appealing properties:

i') **Bounds:** $0 \le E(\rho_1, \rho_2) \le 1$ and $0 \le G(\rho_1, \rho_2) \le 1$.

ii') **Symmetry:** $E(\rho_1, \rho_2) = E(\rho_2, \rho_1)$ and $G(\rho_1, \rho_2) = G(\rho_2, \rho_1)$.

iii') **Unitary invariance:** $E(\rho_1, \rho_2) = E(U\rho_1 U^\dagger, U\rho_2 U^\dagger)$ and $G(\rho_1, \rho_2) = G(U\rho_1 U^\dagger, U\rho_2 U^\dagger)$, for any unitary operator U.

iv') **Concavity:**

> **Proposition 1** *Sub– and super–fidelity are concave, that is for $A, B, C, D \in \Omega_N$ and $\alpha \in [0, 1]$ we have*
>
> $$
> \begin{aligned}
> E(A, \alpha B + (1-\alpha)C) &\ge \alpha E(A, B) + (1-\alpha)E(A, C), &\quad (23)\\
> G(A, \alpha B + (1-\alpha)C) &\ge \alpha G(A, B) + (1-\alpha)G(A, C). &\quad (24)
> \end{aligned}
> $$

v') **Properties of the tensor product:**

> **Proposition 2** *Super–fidelity is super–multiplicative, that is for $A, B, C, D \in \Omega_N$*
>
> $$
> G(A \otimes B, C \otimes D) \ge G(A, C)G(B, D), \quad (25)
> $$
>
> while
>
> **Proposition 3** *Sub–fidelity is sub–multiplicative, that is for $A, B, C, D \in \Omega_N$*
>
> $$
> E(A \otimes B, C \otimes D) \le E(A, C)E(B, D). \quad (26)
> $$

Properties **i')**, **ii')** and **iii')** follow from the properties of $\mathrm{tr}AB$ and definitions (20) and (21). In this section we prove properties **iv')** and **v')**.

Proof of Proposition 1. The definitions (20) and (21) can be rewritten in terms of the aforementioned elementary symmetric functions (5) using relation (14),

$$
\begin{aligned}
E(A,B) &= \mathrm{tr}AB + 2\sqrt{s_2(AB)}, & (27) \\
G(A,B) &= \mathrm{tr}AB + 2\sqrt{s_2(A)s_2(B)}. & (28)
\end{aligned}
$$

In general r^{th} root of s_r is concave on the cone of positive operators [41]. This implies concavity of G directly. To get concavity of E we can replace matrix AB by the similar matrix $A^{1/2}BA^{1/2}$ which is positive. Using the concavity of $\sqrt{s_2(A^{1/2}BA^{1/2})}$ we obtain the result. □.

Proof of Proposition 2. First we note that super–fidelity is not multiplicative. As an example we can take

$$
A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \ C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \ D = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \tag{29}
$$

in which case we have

$$
\frac{1}{2} = G(A \otimes B, C \otimes D) > G(A,C)G(B,D) = 0. \tag{30}
$$

To prove the proposition we write

$$
G(A \otimes B, C \otimes D) = \mathrm{tr}AC\,\mathrm{tr}BD + \sqrt{(1 - \mathrm{tr}A^2\mathrm{tr}B^2)(1 - \mathrm{tr}C^2\mathrm{tr}D^2)}, \tag{31}
$$

and

$$
G(A,C)G(B,D) = \left(\mathrm{tr}AC + \sqrt{(1 - \mathrm{tr}A^2)(1 - \mathrm{tr}C^2)} \right) \left(\mathrm{tr}BD + \sqrt{(1 - \mathrm{tr}B^2)(1 - \mathrm{tr}D^2)} \right). \tag{32}
$$

Denoting $\mathrm{tr}A^2 = \alpha, \mathrm{tr}B^2 = \beta, \mathrm{tr}C^2 = \gamma$ and $\mathrm{tr}D^2 = \delta$ we have to show that

$$
\sqrt{(1 - \alpha\beta)(1 - \gamma\delta)} \geq \mathrm{tr}AC\sqrt{(1 - \beta)(1 - \delta)} + \mathrm{tr}BD\sqrt{(1 - \alpha)(1 - \gamma)} + \sqrt{(1 - \alpha)(1 - \gamma)(1 - \beta)(1 - \delta)}.
$$

Now from Fact 6 with $a = 2$ (see Appendix B) one has

$$
\mathrm{tr}AC \leq \sqrt{\mathrm{tr}A^2\mathrm{tr}C^2} = \sqrt{\alpha\gamma} \tag{33}
$$

and

$$
\mathrm{tr}BD \leq \sqrt{\mathrm{tr}B^2\mathrm{tr}D^2} = \sqrt{\beta\delta}. \tag{34}
$$

Thus it is enough to show that

$$
\sqrt{(1 - \alpha\beta)(1 - \gamma\delta)} \geq \sqrt{\alpha\gamma}\sqrt{(1 - \beta)(1 - \delta)} + \sqrt{\beta\delta}\sqrt{(1 - \alpha)(1 - \gamma)} + \sqrt{(1 - \alpha)(1 - \gamma)(1 - \beta)(1 - \delta)}. \tag{35}
$$

We define two vectors

$$
X = \begin{pmatrix} \sqrt{\alpha}\sqrt{1 - \beta} \\ \sqrt{\beta}\sqrt{1 - \alpha} \\ \sqrt{1 - \alpha}\sqrt{1 - \beta} \end{pmatrix} \ \text{and} \ Y = \begin{pmatrix} \sqrt{\gamma}\sqrt{1 - \delta} \\ \sqrt{\delta}\sqrt{1 - \gamma} \\ \sqrt{1 - \gamma}\sqrt{1 - \delta} \end{pmatrix}. \tag{36}
$$

Note that

$$\langle X|Y\rangle = \sqrt{\alpha\gamma}\sqrt{(1-\beta)(1-\delta)} + \sqrt{\beta\delta}\sqrt{(1-\alpha)(1-\gamma)} + \sqrt{(1-\alpha)(1-\gamma)(1-\beta)(1-\delta)} \quad (37)$$

and

$$\langle X|X\rangle = (1-\alpha\beta) \text{ and } \langle Y|Y\rangle = (1-\gamma\delta). \quad (38)$$

Now by combining (38) with (37) and using Cauchy–Schwarz inequality

$$\sqrt{\langle X|X\rangle\langle Y|Y\rangle} \geq \langle X|Y\rangle, \quad (39)$$

we obtain (35).   □.

Proof of Proposition 3.   To show sub–multiplicativity of sub–fidelity we write the definition (20) for a tensor product,

$$
\begin{aligned}
E(A \otimes B, C \otimes D) &= \operatorname{tr}[(A \otimes B)(C \otimes D)] \\
&\quad + \sqrt{2\left[\left(\operatorname{tr}(A \otimes B)(C \otimes D)\right)^2 - \operatorname{tr}(A \otimes B)(C \otimes D)(A \otimes B)(C \otimes D)\right]} \\
&= \operatorname{tr}AC\operatorname{tr}BD + \sqrt{2[(\operatorname{tr}AC\operatorname{tr}BD)^2 - \operatorname{tr}ACAC\operatorname{tr}BDBD]}.
\end{aligned}
$$

The product of two sub–fidelities reads

$$
\begin{aligned}
E(A,C)E(B,D) &= (\operatorname{tr}AC + \sqrt{2[(\operatorname{tr}AC)^2 - \operatorname{tr}ACAC]})(\operatorname{tr}BD + \sqrt{2[(\operatorname{tr}BD)^2 - \operatorname{tr}BDBD]}) \\
&= \operatorname{tr}AC\operatorname{tr}BD + \operatorname{tr}AC\sqrt{2[(\operatorname{tr}BD)^2 - \operatorname{tr}BDBD]} + \operatorname{tr}BD\sqrt{2[(\operatorname{tr}AC)^2 - \operatorname{tr}ACAC]} \\
&\quad + \sqrt{2[(\operatorname{tr}AC)^2 - \operatorname{tr}ACAC]}\sqrt{2[(\operatorname{tr}BD)^2 - \operatorname{tr}BDBD]}.
\end{aligned}
$$

For short we denote

$$
\begin{aligned}
\alpha &= \operatorname{tr}AC, \quad a = \operatorname{tr}ACAC, \\
\beta &= \operatorname{tr}BD, \quad b = \operatorname{tr}BDBD.
\end{aligned}
$$

We have $\alpha^2 \geq a$ and $\beta^2 \geq b$. By rewriting above expressions in the new notation we obtain

$$E(A \otimes B, C \otimes D) = \alpha\beta + \sqrt{2[\alpha^2\beta^2 - ab]}$$

and

$$E(A,C)E(B,D) = \alpha\beta + \alpha\sqrt{2[\beta^2 - b]} + \beta\sqrt{2[\alpha^2 - a]} + \sqrt{2[\alpha^2 - a]}\sqrt{2[\beta^2 - b]}.$$

Now we write

$$
\begin{aligned}
\sqrt{2[\alpha^2\beta^2 - ab]} &= \sqrt{2[\alpha^2\beta^2 - ab + ab - ab - \alpha^2 b + \alpha^2 b - \beta^2 a + \beta^2 a]} \\
&= \sqrt{2[(\alpha^2 - a)(\beta^2 - b) + b(\alpha^2 - a) + a(\beta^2 - b)]}.
\end{aligned}
$$

Making use of subadditivity of square root we obtain

$$\sqrt{2[\alpha^2\beta^2 - ab]} \leq \sqrt{2(\alpha^2 - a)(\beta^2 - b)} + \sqrt{2b(\alpha^2 - a)} + \sqrt{2a(\beta^2 - b)}.$$

Because $2 < 4$, $a \leq \alpha^2$ and $b \leq \beta^2$ we get

$$\sqrt{2[\alpha^2\beta^2 - ab]} \leq \sqrt{4(\alpha^2 - a)(\beta^2 - b)} + \beta\sqrt{2(\alpha^2 - a)} + \alpha\sqrt{2(\beta^2 - b)}.$$

And as a result we obtain the desired inequality

$$E(A \otimes B, C \otimes D) \quad \leq \quad E(A, C)E(B, D).$$

□.

For any pair of Hermitian operators X_1 and X_2 let us now define a quadratic *Lorentz form*

$$(X_1, X_2)_L := [(\text{tr}X_1)(\text{tr}X_2) - \text{tr}X_1X_2]. \tag{40}$$

To find out the motivation standing behind this name let us expand a Hermitian operator X in an operator basis, $X = \sum_{j=0}^{N^2-1} a_j H_j$. We assume that the basis is orthogonal, $\text{tr}H_j H_k = \delta_{jk}$, the first operator is proportional to identity, $H_0 = \mathbb{I}/\sqrt{N}$, and all other operators H_j are traceless. Then the form (40) gives

$$(X, X)_L = (\text{tr}X)^2 - \text{tr}X^2 = \frac{(N-1)}{N}a_0^2 - \sum_{j=1}^{N^2-1} a_j^2, \tag{41}$$

which is of Minkowski–Lorentz type. By the help of this notion, super–fidelity can be written as

$$G(A, B) = \text{tr}AB + [(A, A)_L(B, B)_L]^{1/2}, \tag{42}$$

while sub–fidelity reads

$$E(A, B) = \text{tr}AB + [2(AB, AB)_L]^{1/2}. \tag{43}$$

The *forward cone* with respect to the form (40) is given by operators X satisfying

$$(X, X)_L \geq 0 \text{ and } \text{tr}X \geq 0. \tag{44}$$

Since the density matrices are normalized, $\text{tr}\rho = 1$, the form $(\rho, \rho)_L$ is non–negative.
For a Lorentz form any two forward directed Hermitian matrices A and B satisfy

$$(A, A)_L(B, B)_L \leq [(A, B)_L]^2. \tag{45}$$

Substituting this bound into expression (42) we arrive at an upper bound for super–fidelity

$$G(A, B) \leq \text{tr}AB + (\text{tr}A)(\text{tr}B) - (\text{tr}AB) = (\text{tr}A)(\text{tr}B). \tag{46}$$

For the case of normalized density matrices, $\text{tr}\rho = 1$ we get $G(\rho_1, \rho_2) \leq 1$.
Using the bound (11) for density operators we introduce a third quantity

$$E'(A, B) = \text{tr}AB + r(r-1)\sqrt[r]{s_r(AB)}, \tag{47}$$

where r is the rank of matrix AB. Note that for $r = 2$ this formula is reduced to an expression (27) for sub–fidelity, hence in this case $E' = E$.

Since $(s_r(X))^{1/r}$ is concave for density operators we infer that the quantity $E'(A, B)$, defined in equation (47), is separately concave in A and in B.

5.2 Examples and classical analogues

To observe sub– and super–fidelity in action consider a family of mixed states

$$\rho_a = a|\psi\rangle\langle\psi| + (1-a)\mathbb{I}/N, \tag{48}$$

which interpolates between arbitrary pure state $|\psi\rangle$ and the maximally mixed state. It is straightforward to compute the fidelity between the state ρ_a and the maximally mixed state $\rho_* := \mathbb{I}/N$,

$$F(\rho_a, \rho_*) = \frac{1}{N^2}\left(\sqrt{(N-1)a+1} + (N-1)\sqrt{1-a}\right)^2, \tag{49}$$

as well as other bounds

$$E(\rho_a, \rho_*) = \frac{1}{N} + \sqrt{2}\frac{1}{N}\sqrt{1 - \frac{1}{N}}\sqrt{1-a^2}, \tag{50}$$

$$E'(\rho_a, \rho_*) = \frac{1}{N} + \left(1 - \frac{1}{N}\right)\sqrt[N]{((N-1)a+1)(1-a)^{N-1}}, \tag{51}$$

$$G(\rho_a, \rho_*) = \frac{1}{N} + \left(1 - \frac{1}{N}\right)\sqrt{1-a^2}. \tag{52}$$

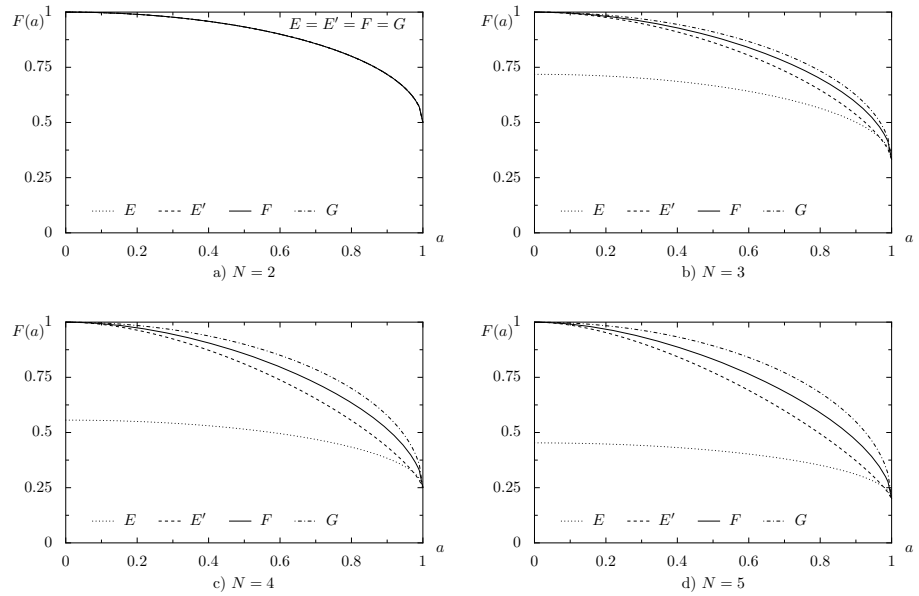These results are plotted in Fig. 1 for $N = 2, 3, 4, 5$.

Fig. 1. The comparison of sub–fidelity E, bound E', fidelity F (solid line) and super–fidelity G. Each plot shows these quantities calculated for the maximally mixed state and a state (48) depending on the parameter a. For a one–qubit system, case a) $N = 2$, one has $E = E' = F = G$. Note the difference between these quantities shown for $N = 3, 4, 5$. In this case $E > E'$ for a close to unity.

For $N = 2$ all these quantities coincide, and the quality of the approximation goes down with the system size N, as expected. Looking at the graph one could imagine that relation

$E \leq E'$ is fulfilled. However, such an equality does not hold as we found a counter example: the pair of states analyzed in the figure with parameter a very close to unity.

One may work out several other examples, for which sub– and super–fidelity are easy to find. Explicit formulas are simple in the case of two commuting density matrices ρ_p and ρ_q with spectra given by vectors $\vec{p}$ and $\vec{q}$, respectively. In such a classical case these quantities read

$$E(\rho_p, \rho_q) = \sum_{i=1}^{N} p_i q_i + \sqrt{2 \left[\left(\sum_{i=1}^{N} p_i q_i \right)^2 - \sum_{i=1}^{N} p_i^2 q_i^2 \right]}, \tag{53}$$

$$F(\rho_p, \rho_q) = \left(\sum_{i=1}^{N} \sqrt{p_i q_i} \right)^2, \tag{54}$$

$$G(\rho_p, \rho_q) = \sum_{i=1}^{N} p_i q_i + \sqrt{\left(1 - \sum_{i=1}^{N} p_i^2 \right) \left(1 - \sum_{i=1}^{N} q_i^2 \right)}. \tag{55}$$

5.3   The difference $G - F$

In view of the inequality (22) it is natural to ask how big the difference $G - F$ might be. Since both quantities coincide if one of the states is pure, let us analyze the case of two mixed states living in orthogonal subspaces.

More precisely, let us fix an even dimensionality of the Hilbert space $N = 2M$, and define two diagonal states, each supported in M dimensional space, $\rho_1 = \frac{2}{N}\text{diag}(1, \ldots, 1, 0, \ldots, 0)$ and $\rho_2 = \frac{2}{N}\text{diag}(0, \ldots, 0, 1, \ldots, 1)$. Since they are supported by orthogonal subspaces their fidelity vanishes, $F(\rho_1, \rho_2) = 0$. On the other hand the definition (21) gives their super–fidelity

$$G(\rho_1, \rho_2) = \frac{N - 2}{N}, \tag{56}$$

equal in this case to the difference $G - F$. As expected for $N = 2$ we get $G = F = 0$. However, for N large enough the difference $G - F$ may become arbitrarily close to unity.

Thus working with super–fidelity G in place of fidelity F one needs to remember that this approximation works fine for small systems or where at least one of the states is pure enough.

6   On measurement methods

6.1   Associated physical observables

Here we shall shortly discuss possibilities of measurement of both sub– and super–fidelities in physical experiments. The approach below follows the techniques used in state spectrum estimation [20] and nonlinear entanglement detection and/or estimation which has been developed significantly last years (see [21, 22] and references therein). Those approaches exploited the properties of SWAP operator and other permutation unitary operations to get the properties of single state rather than the relation of different states. There were little exceptions: one was a quantum network measurement of an overlap of the two states [20]. Here we shall follow the latter idea since we want to estimate the *distance* of two different quantum states. In particular we shall see that it is possible to measure these quantities with help of not more than two collective observables. This fact may be helpful in experimental comparison of

two different stationary sources of quantum states. Quite remarkably, as we shall see below, with help of similar techniques, super–fidelity can be represented by only three experimental probabilities which makes it very friendly from an experimental point of view.

We start by providing a simple example. First one can see that to calculate sub– and super–fidelity it is necessary to calculate the values of the terms of the form $\mathrm{tr} AB$. Let $A, B \in M_2(\mathbb{C})$. In this case

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \tag{57}$$

and

$$\mathrm{tr} AB = \mathrm{tr} \left[\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \right] = a_{11}b_{11} + a_{21}b_{12} + a_{12}b_{21} + a_{22}b_{22}. \tag{58}$$

On the other hand this value can be calculated using SWAP gate as

$$
\begin{aligned}
\mathrm{tr}\left[\mathrm{SWAP}(A \otimes B)\right] &= \mathrm{tr} \left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} \right] \tag{59} \\
&= \mathrm{tr} \left[\begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} \right] = \mathrm{tr} AB. \tag{60}
\end{aligned}
$$

To address the question of measurability of the quantities (20), (21) let us first recall the corresponding permutation operators which we shall need subsequently. The first one will be just SWAP operator (example of which is the SWAP gate presented above) $V_{12} : \mathcal{H}_N \otimes \mathcal{H}_N \to \mathcal{H}_N \otimes \mathcal{H}_N$ which is defined by the action

$$V_{12}|\phi_1\rangle \otimes |\psi_2\rangle = |\psi_2\rangle \otimes |\phi_1\rangle. \tag{61}$$

This is a Hermitian operator and as a such it represents an observable. It has a simple eigendecomposition in the form

$$V_{12} = P_{12}^{(+)} - P_{12}^{(-)}, \tag{62}$$

where projections $P_{12}^{(\pm)}$ onto symmetric and antisymmetric subspaces of $\mathcal{H}_N \otimes \mathcal{H}_N$ are

$$P_{12}^{\pm} = \frac{1}{2}(\mathbb{I}_{12} \pm V_{12}). \tag{63}$$

Below we shall omit the indices and use the notation V and $P^{(\pm)}$, if it does not lead to confusion. An important property usually exploited in case of entanglement detection is that the formula (60) holds for the SWAP operator V of any dimension [23]. Apart form that operation we will also need a family of unitary permutation matrices $V_{1234}^{\pi} : \mathcal{H}_N^{\otimes 4} \to \mathcal{H}_N^{\otimes 4}$,

$$V_{1234}^{\pi}|\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle \otimes |\psi_4\rangle = |\psi_{\pi(1)}\rangle \otimes |\psi_{\pi(2)}\rangle \otimes |\psi_{\pi(3)}\rangle \otimes |\psi_{\pi(4)}\rangle, \tag{64}$$

where π represents any chosen permutation of the indices $(1, 2, 3, 4)$. For simplicity we shall drop the indices using the notation V^{π}.

Let us define the set $\mathcal{S}$ of all eight permutations that do not map the sequence $(1, 2, 3, 4)$ into a one having odd or even elements one after another. For instance, the permutations defined by the ranges (2341) or (3214) belong to $\mathcal{S}$, while (2314) or (1423) do not. For a fixed set $\mathcal{S}' \subset \mathcal{S}$ and some $\pi_0 \in \mathcal{S}$ we define the following observable:

$$W^{\mathcal{S}', \pi_0} = \frac{1}{2|\mathcal{S}'|} \left(\sum_{\pi \in \mathcal{S}'} V^\pi V^{\pi_0} V^\pi + \sum_{\pi \in \mathcal{S}'} V^{\pi^{-1}} V^{\pi_0^{-1}} V^{\pi^{-1}} \right). \qquad (65)$$

A special case is the observable $W^{\{\pi_0\}, \pi_0} = (V^{\pi_0} + V^{\pi_0^{-1}})/2$ with π_0 being just some cyclic permutation (cf. [21] and references therein). The choice of the permutation π_0 and/or the subset $\mathcal{S}'$ may be motivated by a specific physical situation.

In the case of single qubit sources ($N = 2$) the observables (65) have highly degenerated spectra and the corresponding eigenvectors have very symmetric forms. In particular the observable $W^{\{\pi_0\}, \pi_0}$ has spectrum $\{1, -1, 0\}$ which means that its mean value requires probabilities of only two outcomes of incomplete von Neumann measurement. The observable has the spectral decomposition $W^{\{\pi_0\}, \pi_0} = Q^{(+)} - Q^{(-)}$ where support of the projector $Q^{(+)}$ is spanned by eigenvectors $\{|\phi_1\rangle = |0000\rangle, |\phi_2\rangle = |1111\rangle, |\phi_3\rangle = (|0111\rangle + |1011\rangle + |1101\rangle + |1110\rangle)/2, |\phi_4\rangle = (|0011\rangle + |0110\rangle + |1001\rangle + |1100\rangle)/2, |\phi_5\rangle = (|0101\rangle + |1010\rangle)/\sqrt{2}, |\phi_6\rangle = \sigma_x^{\otimes 4} |\psi_4\rangle\}$ while the support of the second projector $Q^{(-)}$ (orthogonal to $Q^{(+)}$) corresponds to $\{I \otimes \sigma_z^{\otimes 2} \otimes I |\phi_3\rangle, I^{\otimes 2} \otimes \sigma_z^{\otimes 2} |\phi_4\rangle, I^{\otimes 3} \otimes \sigma_z |\phi_5\rangle, I \otimes \sigma_z^{\otimes 2} \otimes I |\phi_6\rangle\}$.

To illustrate how to measure the quantities E and G suppose now we can perform collective measurements on two and four copies of both quantum states. We plan measurements that allow two or four copies of analyzed states to interact. Then (cf. [20, 21, 22] and references therein) the sub– and super–fidelities can be represented in terms of averages of following observables,

$$E(\rho_1, \rho_2) = \mathrm{tr} V \rho_1 \otimes \rho_2 + \sqrt{2[(\mathrm{tr} V \rho_1 \otimes \rho_2)^2 - \mathrm{tr} W^{\mathcal{S}, \pi_0} \rho_1 \otimes \rho_2 \otimes \rho_1 \otimes \rho_2]}, \qquad (66)$$

$$G(\rho_1, \rho_2) = \mathrm{tr} V \rho_1 \otimes \rho_2 + \sqrt{1 - \mathrm{tr} V \rho_1 \otimes \rho_1} \sqrt{1 - \mathrm{tr} V \rho_2 \otimes \rho_2}. \qquad (67)$$

There are two simple but important observations to be made. The sub–fidelity E can be measured with help of two setups: (i) the one measuring the observable V and (ii) the second one measuring observable $W^{\mathcal{S}, \pi_0}$. Each setup requires one source: setup (i) needs the source that creates, say, pairs $\rho_1 \otimes \rho_2$, while setup (ii) requires a source producing quadruples of the form, say, $\rho_1 \otimes \rho_2 \otimes \rho_1 \otimes \rho_2$.

Our scheme will work also for a worse source that produces one of the pairs (quadruples) $\{\rho_1 \otimes \rho_2, \rho_2 \otimes \rho_1\}$ ($\{\rho_1 \otimes \rho_2 \otimes \rho_1 \otimes \rho_2, \rho_2 \otimes \rho_1 \otimes \rho_2 \otimes \rho_1\}$) at random according to an unknown biased probability distribution, which will not affect the results of the corresponding estimate for sub–fidelity.

The second observation is that the super–fidelity G can be measured with help of *single* setup, namely the one that measures observable V, but requires its application to three types of sources i.e. the ones creating pairs $\rho_1 \otimes \rho_1$, $\rho_2 \otimes \rho_2$, and, say, $\rho_1 \otimes \rho_2$. Again, the last source may produce at random one of the pairs $\{\rho_1 \otimes \rho_2, \rho_2 \otimes \rho_1\}$ and this will not affect the estimate for super–fidelity.

It is very interesting to study the form of super–fidelity in terms of directly measurable quantities, *i.e. probabilities*, since it has a simple optical implementation. Let us introduce

the probabilities of the projection onto the antisymmetric subspace of $\mathcal{H}_N \otimes \mathcal{H}_N$:

$$p_{ij}^{(-)} = \mathrm{tr} P^{(-)} \rho_i \otimes \rho_j, \quad i, j = 1, 2. \tag{68}$$

Then super–fidelity has a particularly nice form,

$$G(\rho_1, \rho_2) = 1 - 2 \left(p_{12}^{(-)} - \sqrt{p_{11}^{(-)} p_{22}^{(-)}} \right) , \tag{69}$$

which is crucial for further discussion. Note that the super–fidelity can be represented in terms of only *three* probabilities that can be measured in a single set-up. One can perform a simple consistency test by checking, whether the combination of experimental probabilities satisfy (up to error bars) the condition $p_{12}^{(-)} - \sqrt{p_{11}^{(-)} p_{22}^{(-)}} \leq 0.5$ – otherwise one had an unphysical result, since super–fidelity can not be negative. Note that the probability $p_{11}^{(-)}$ has been already measured experimentally for two copies of composite systems in context of entanglement detection [25] or estimation [26] under some assumptions about the nature of the sources. In subsection below we shall refer to the scheme analogous to the one utilized in Ref. [25].

It is interesting to note that if the state ϱ is of d-dimensional type, then reproduction of sub– and super–fidelity *via* quantum tomography requires $2d^2 - 2$ independent quantities to be estimated since each of the two states is described by $d^2 - 1$ real parameters. On the other hand, to find the quantities E and G in the way described above one requires only two or three independent real quantities (probabilities) to be estimated independly on how large the dimension d is. The price to be payed is, of course, that one must perform collective experiments. Preparation of reliabe setups of such experiments might be a good test for quantum engeneering.

6.2   *Measuring super–fidelity of states representing photons polarizations*

Consider now physical setup that would compare two states of polarization of single photon in terms of super–fidelity G. In this case the density matrix is defined on Hilbert space isomorphic to C^2 where the horizontal (vertical) polarization, usually denoted by $|H\rangle$ ($|V\rangle$) corresponds to the standard basis element $|0\rangle$ ($|1\rangle$). Suppose one has memoryless sources of two types S_i ($i = 1, 2$) sending photons in polarization states ρ_1, ρ_2 respectively.

The experimental setup is elementary. We have sources S_i, S_j, where we put either $i = j = 1, 2$ (sources of the same type) or, say $i = 1$, $j = 2$ (different sources) then we have a beamsplitter (in equal distance to the source) and two detectors behind it (see Fig. 2). If two photons form sources S_i, S_j meet on the beamsplitter and the two detectors click, we have so–called anticoalescence event, which happens with probability $p_{ij}^{(-)}$ [25]. Otherwise we deal with a coalescence result which occurs with probability $p_{ij}^{(+)} = 1 - p_{ij}^{(-)}$. Putting all three probabilities of anti-coalescence into formula (69) we reproduce the expression for super–fidelity.

This seems to be the most easy experiment with two sources to perform. Such an experiment can be realized for two sources of photons engineered with help of controlled decoherence (in a way similar to Ref. [27]) corresponding to two different mixed states of a qubit.

The above scheme immediately extends to the case of states ρ_1, ρ_2 are defined on $N = 2^n$-dimensional Hilbert space representing polarization degrees of freedom of n photons. In this
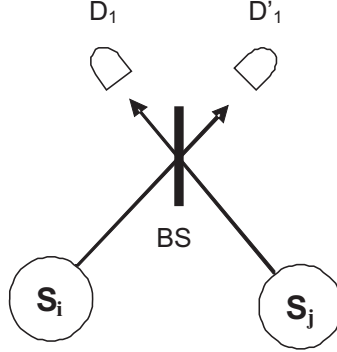
Fig. 2. Elementary scheme with one beam-splitter BS. The sources S_i, S_j are chosen to be, in turn, ot the same ($i = j = 1, 2$) and different ($i = 1$, $j = 2$) type. Single click in either of the detectors D, D' corresponds to projection into symmetric two-qubit subspace of photon polarization, while two clicks represent projection onto one-dimensional antisymmetric (singlet) subspace.

case the total Hilbert space is $\mathcal{H}_N = (\mathbb{C}^2)^{\otimes n}$ and the scheme reads as in Fig. 3 (compare [28, 22]). If the probability $p_{ij}^{(s_k),k}$ with $s_k = -1$ ($s_k = +1$) corresponds to anticoalescence (coalescence) on k-th beamsplitter, i.e. it represents the probability of two clicks (one click) in the pair of detectors D_k, D'_k, then the total probabilities:

$$p_{ij}^{(-)} = \sum_{s_1, s_2 \ldots, s_n:\ s_1 s_2 \ldots s_n = -1} p_{ij}^{(s_1),1} p_{ij}^{(s_2),2} \ldots p_{ij}^{(s_n),n}, \quad i, j = 1, 2 \tag{70}$$

are these we put into (69). In the formula above we count all the cases when an odd number of anti-coalescence events occurs, provided that there is no photon losses during the experiment.
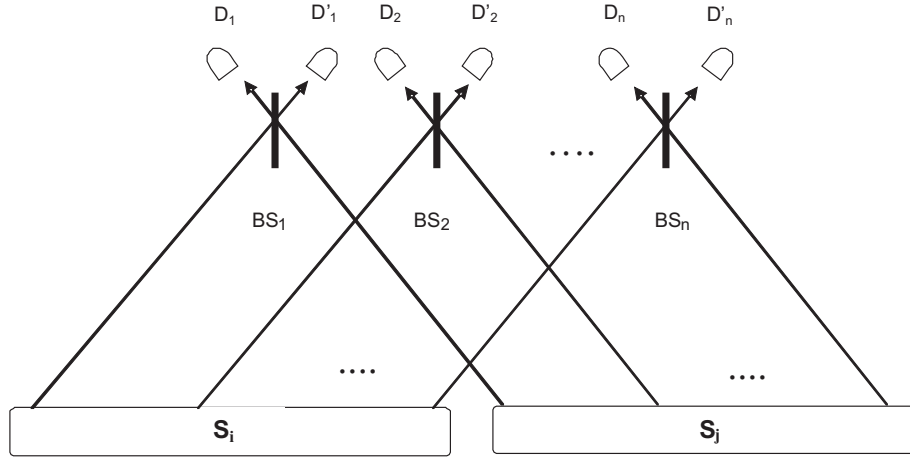

Fig. 3. The scheme for measurement of super–fidelity of the states of n-photon polarizations. According to (70) only the events with double clicks in odd number of detector pairs contribute to each of the three probabilities in the formula (69).

For $n = 2$ this type of experiment has already been performed with two two-photon sources producing entangled states [25]. However, the sources were considered to provide

the same state on average rather than two different ones. A similar reasoning was used in another recent experiment, in which photon polarization and momentum degrees of freedom were used to estimate the concurrence where additional strong assumption about purity of each copies were also used [26]. In general, measurements schemes of quantities like purity, concurrence, sub– and super–fidelity in the collective framework like the one presented here requires the assumption that the sources producing states are stationary and memoryless. Quite remarkably this is the same assumption one makes in quantum tomography. As discussed in [29, 30] there may be difficulties with satisfying it in real experimental scenarios for instance due to classical correlations between the consecutive copies of the system. In other words the condition of having the global state in ϱ^N form (which mathematically corresponds to quantum de Finnetti condition [31]) may not be obeyed. This point requires more analysis, and leads, in general to nontrivial issues. It seems however that stationarity and memoryless character of the source can usually be satisfied approximately. Then the present measurement would serve (similarly like quantum tomography does, though may be in a way more sensitive to source correlations) as an approximative, coarse-grained-like characteristic of the states sources under reasonable physical assumptions.

6.3   Quantum networks

There is yet another method of detection of quantities that may be considered here. This is a method based on quantum networks. It is known that a unitary operation U acting on state σ but *controlled* by a qubit in the superposition state $|+\rangle \equiv \frac{1}{2}(|0\rangle + |1\rangle)$ reproduces the value $Re(\mathrm{tr}U\sigma)$ directly as a mean value of the Pauli matrix $\langle \sigma_x \rangle$ measured on the controlled qubit [24, 20]. This fact allows us to measure certain nonlinear functions of the state. To get $\mathrm{tr}\rho^k$ one takes k copies of the state, $\sigma = \rho^{\otimes k}$ and takes for U the operator of cyclic permutation, in full analogy to V^π used in previous subsection. To measure the overlap of two matrices, ρ_1, ρ_2, one takes $\sigma = \rho_1 \otimes \rho_2$ and uses the SWAP operator, $U = V$.

The corresponding network is already provided explicitly in Ref. [20], so we shall not write it down here. Such a network allows one to measure all three quantities needed to reproduce the super–fidelity G. Indeed, the network produces *directly* (as mean values of σ_z on controlled qubit) all three mean values: $\mathrm{tr}\rho_i\rho_j$, $i = 1, 2$ provided that states forming the input of the controlled part of the network are $\rho_i \otimes \rho_j$. Some alternative constructions of programmable networks designed to measure super–fidelity are also possible.

Similarly, sub–fidelity E can also be estimated with a network-based experimental scheme. Following reasoning from Ref. [21] one constructs the following programmable quantum network – see Fig. 4. Depending on the program state $|\Psi_{12}\rangle$ as a mean value of σ_z of the measured controlling qubit one gets

(i) $\mathrm{tr}\rho_1\rho_2$ if $|\Psi_{12}\rangle = |0\rangle|0\rangle$,

(ii) $\mathrm{tr}\rho_1\rho_2\rho_1\rho_2$ if $|\Psi_{12}\rangle = |1\rangle|0\rangle$,

(iii) $\frac{1}{2}\left(\mathrm{tr}\rho_1\rho_2\rho_1\rho_2 - (\mathrm{tr}\rho_1\rho_2)^2\right)$ if $|\Psi_{12}\rangle = (|0\rangle|1\rangle + |1\rangle|0\rangle)/\sqrt{2}$ *i.e.* if it is in Bell state.

The last quantity up to the factor $(-\frac{1}{4})$ is just the quantity that occurs under the square root in the formula for E. In general to estimate the sub–fidelity E we may "run" the first "program" and either the second or the third one. Alternatively, we may run all three
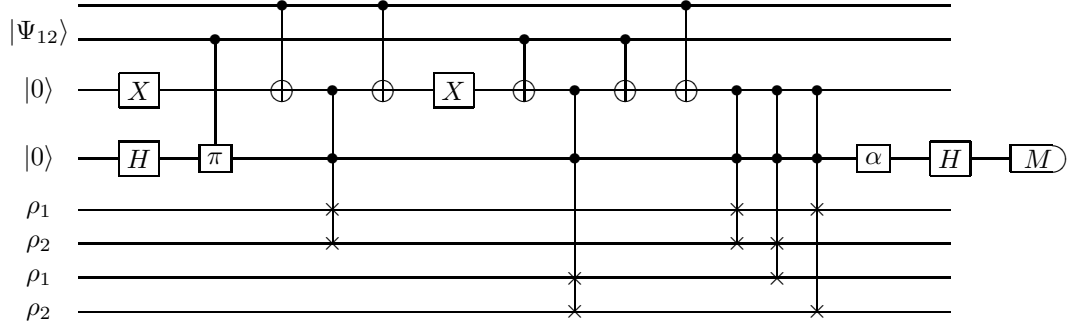
Fig. 4.   Example of programmable network allowing to measure in particular the quantities $\mathrm{tr}(\rho_1\rho_2\rho_1\rho_2)$ and $\frac{1}{2}\mathrm{tr}(\rho_1\rho_2\rho_1\rho_2) - \mathrm{tr}(\rho_1\rho_2)^2$. The state $|\Psi_{12}\rangle$ represents the program (see the main text). Symbol M corresponds to the measurement of Pauli matrix σ_z. The last phase gate usually is chosen to be $\alpha = 0$ unless the interferometric picture with visibility is needed (cf. Ref. [20]). See [8] for the description of quantum gates used in this circuit.

programs and use the data to verify the accuracy of the experiment by comparing the two partially independent estimates of E obtained in that way. It is easy tu see, that the same network can be used to estimate super–fidelity G if one puts as an input $\rho_i \otimes \rho_i \otimes \rho_j \otimes \rho_j$, $j = 1, 2$.

7   Distances and geometry of the space of states

7.1   Hilbert-Schmidt distance and flat geometry

The geometry of the space of quantum states depends on the metric used [32, 33, 34, 9]. The set Ω_N of mixed states of size N reveals the Euclidean (flat) geometry if it is analyzed with respect to the *Hilbert-Schmidt distance*,

$$D_{HS}(\rho_1, \rho_2) = \sqrt{\mathrm{tr}[(\rho_1 - \rho_2)^2]}. \tag{71}$$

To demonstrate this property let us first concentrate on the simplest case, $N = 2$. Making use of the notion of a coherence vector $\vec{\tau}$ any state of a qubit can be written in the *Bloch representation*

$$\rho = \frac{\mathbb{I}}{N} + \vec{\tau} \cdot \vec{\lambda}. \tag{72}$$

Here $\vec{\lambda}$ denotes the vector of three rescaled traceless Pauli matrices $\{\sigma^x, \sigma^y, \sigma^z\}/\sqrt{2}$, which are orthogonal in the sense of the Hilbert-Schmidt scalar product, $\langle\lambda^k|\lambda^m\rangle = \mathrm{tr}(\lambda^k)^\dagger\lambda^m = \delta_{km}$. Together with $\lambda^0 = \mathbb{I}/\sqrt{2}$ they form an orthonormal basis in the space of complex density matrices of size two. Due to Hermiticity of ρ the three-dimensional Bloch vector $\vec{\tau}$ is real. Positivity condition implies $|\vec{\tau}| \leq 1/\sqrt{2} = R_2$ with equality for pure states, which form the Bloch sphere of radius R_2. Representation (72) implies that for any state of a qubit $\mathrm{tr}\rho^2 = 1/2 + |\tau|^2$.

Consider two arbitrary density matrices and express their difference $\rho_1 - \rho_2$ in the Bloch form. The entries of this difference consist of the differences between components of both

Bloch vectors $\vec{\tau}_1$ and $\vec{\tau}_2$. Therefore Hilbert-Schmidt distance induces the flat (Euclidean) geometry of Ω_2,

$$D_{HS}\big(\rho_{\vec{\tau}_1}, \rho_{\vec{\tau}_2}\big) = D_E(\vec{\tau}_1, \vec{\tau}_2), \tag{73}$$

where D_E is the Euclidean distance between both Bloch vectors in $\mathbb{R}^3$.

It is worth to add that expression (73) holds for an arbitrary N. In this case $\vec{\tau}$ is a real vector with $N^2 - 1$ components, while the vector $\vec{\lambda} = \{\lambda^k\}_{k=1}^{N^2-1}$ in (72) denotes the set of $N^2 - 1$ traceless generators of the group $SU(N)$. Positivity of ρ implies that the length of the Bloch vector is limited by

$$|\vec{\tau}| \leq D_{HS}(\mathbb{I}/N, |\psi\rangle\langle\psi|) = \sqrt{\frac{N-1}{N}} =: R_N. \tag{74}$$

For $N = 2$ the condition $|\vec{\tau}| \leq R_2$ is sufficient to imply that the corresponding matrix is positive and represents a state, while for $N \geq 3$ it is only a necessary condition [9]. This is related to the fact that with respect to the flat, H–S geometry the set Ω_2 forms a full 3-ball, while for larger N the set Ω_N forms a convex subset of the (N^2-1)-dimensional ball of radius R_N centered at $\rho_* = \mathbb{I}/N$.

7.2   *Bures distance and the geometry it induces*

The notion of fidelity, introduced in (1), can be used to define the *Bures distance* [35, 2]

$$D_F(\rho_1, \rho_2) = \sqrt{2 - 2\sqrt{F(\rho_1, \rho_2)}}. \tag{75}$$

or the *Bures length* [36] (later called *angle* in [8]),

$$D'_F(\rho_1, \rho_2) := \arccos\sqrt{F(\rho_1, \rho_2)} = \frac{1}{2}\arccos\Big(2F(\rho_1, \rho_2) - 1\Big). \tag{76}$$

For any pair of pure states the Bures length coincides with their Fubini–Study distance, $D'_F\big(\rho_\psi, \rho_\phi\big) = d_{FS}\big(|\psi\rangle, |\phi\rangle\big) = \arccos|\langle\psi|\phi\rangle|$.

The Bures metric is distinguished by its rather special properties: it is a Riemannian, monotone metric [37], Fisher adjusted metric [33], closely related to the statistical distance [32].

It is not difficult to describe the geometry of the set of mixed states of a single qubit induced by the Bures metric. Consider a mixed state $\rho \in \Omega_2$ and its transformation proposed in [4]

$$\rho(x, y, z) \rightarrow \Big(x, y, z, t = \sqrt{\frac{1}{2} - x^2 - y^2 - z^2}\Big). \tag{77}$$

It blows up the Bloch ball $\mathbf{B}^3$ of radius $R_2 = 1/\sqrt{2}$ into a hyper-hemisphere $\frac{1}{2}S^3$ of the same radius. The original variables (x, y, z) denote the parameters of the state in the Bloch vector representation. The auxiliary variable reads $t = \sqrt{1/2 - |\tau|^2}$ in terms of the Bloch vector, so that $t^2 + \mathrm{tr}\rho^2 = 1$. The maximally mixed state $\rho_* = (0, 0, 0)$, is mapped into a hyper-pole. It is equally distant from all pure states located at the hyper-equator S^2, which form the boundary of Ω_2.

Any state ρ is uniquely represented by an 'extended Bloch vector', $\vec{v} = (x, y, z, t)$ of length R_2. The auxiliary variable reads $t = \sqrt{1/2 - |\tau|^2}$ in terms of the Bloch vector, so that

$t^2 + \text{tr}\rho^2 = 1$. Consider two states ρ_1 and ρ_2, described by two vectors $\vec{v}_1$ and $\vec{v}_2 \in \mathbb{R}^4$, which form the angle ϑ. Since for any one-qubit states the bound (13) becomes an equality, we see that fidelity between them reads $F(\rho_1, \rho_2) = \text{tr}\rho_1\rho_2 + \sqrt{t_1^2 t_2^2} = 1/2 + \vec{\tau}_1 \cdot \vec{\tau}_2 + t_1 t_2$. This can be rewritten with the use of extended vectors $\vec{v}_i$ and the angle between them, $F = 1/2 + \vec{v}_1 \cdot \vec{v}_2 = 1/2 + R_2^2 \cos \vartheta$. Since $R_2^2 = 1/2$ we find

$$\vartheta = \arccos(2F - 1) = 2D'_F(\rho_1, \rho_2), \tag{78}$$

which shows that the Bures length (76) between any two mixed states is proportional to the Riemannian distance between the corresponding points at the Uhlmann hemisphere.

Making use of the fidelity F one can also define other distances in the space of quantum states. For instance, Gilchrist et al. [38] have shown that the *root infidelity*

$$C(\rho_1, \rho_2) = \sqrt{1 - F(\rho_1, \rho_2)} \tag{79}$$

satisfies the triangle inequality and thus introduces a metric. This very quantity can be used to bound the trace distance $D_{\text{tr}}(\rho_1, \rho_2) = \frac{1}{2}\text{Tr}|\rho_1 - \rho_2| \leq C(\rho_1, \rho_2)$ from above [39, 9]. Note that the Bures distance, Bures length and root infidelity are functions of the same quantity, so they generate the same topology.

7.3 Modified Bures length

In analogy to (75) and (76) one may ask whether

$$D_G(\rho_1, \rho_2) = \sqrt{2 - 2\sqrt{G(\rho_1, \rho_2)}}. \tag{80}$$

and

$$D'_G(\rho_1, \rho_2) := \arccos\sqrt{G(\rho_1, \rho_2)} \tag{81}$$

define distances. This is obvious for $N = 2$ because $F = G$. The situation changes for $N \geq 3$, for which the D_F and D_G do differ and only $F \leq G$ is valid.

We do not know, whether D_G and D'_G are distances. However, it can be proved that a direct analogue of the root infidelity (79)

$$C'(\rho_1, \rho_2) = \sqrt{1 - G(\rho_1, \rho_2)} \tag{82}$$

is a genuine distance. The same is true for the *modified Bures length*,

$$D'_M(\rho_1, \rho_2) = \arccos G(\rho_1, \rho_2). \tag{83}$$

Proof.   Let us call $\mathcal{L}$ the direct sum of the real linear space of all Hermitian operators and the 1-dimensional space of real numbers. Its elements are $\{H, x\}$, H Hermitian, x a real number. $\mathcal{L}$ becomes Euclidean (i.e. a real Hilbert space) by defining the scalar product

$$(\{H_1, x_1\}, \{H_2, x_2\}) = \text{tr}H_1 H_2 + x_1 x_2. \tag{84}$$

Let us denote by $B(\mathcal{L})$ the unit ball of $\mathcal{L}$ and by $S(\mathcal{L})$ the unit sphere. Our proof rests on the embedding of the Hermitian operators

$$B_N = \{H| \text{ tr}H = 1, \text{tr}H^2 \leq 1\} \tag{85}$$

into $S(\mathcal{L})$ by

$$H \to \xi_H := \left\{ H, \sqrt{1 - \operatorname{tr}H^2} \right\}. \tag{86}$$

Clearly, $(\xi_H, \xi_H) = 1$, and from (84) we get

$$(\xi_H, \xi_{H'}) = G(H, H'). \tag{87}$$

Now it is obvious that $\sqrt{2 - 2G}$ is the Euclidean distance between ξ_H and ξ'_H, provided H and H' belong to B_N. Because the density operators form a subset of B_N, (82) is a distance.

From (87) we get $G(H, H') = \cos \alpha$, where α is the angle from which ξ_H and ξ'_H are seen from the center of the ball $B(\mathcal{L})$. Thus, $\arccos G = \alpha$ and, in particular, (83) is a distance. $\square$.

Let us now return to the two conditions of (85). They are equivalent with

$$B_N = \left\{ H \mid \operatorname{tr}H = 1, \operatorname{tr}(H - \frac{1}{N}\mathbb{I})^2 \leq \frac{N-1}{N} \right\} \tag{88}$$

and they describe the smallest ball containing the state space. B_N is an affine translate by $1/N$ of the generalized Bloch-ball, [8]. B_N is centered at $A = N^{-1}\mathbb{I}$ and is of radius $\sqrt{(N-1)/N}$.

Above we have embedded B_N by the map (86) into the sphere $S(\mathcal{L})$. Just this gives the opportunity to apply Mielnik's definition [40] for a transition probability (he also called it *affine ratio*) of two extremal states of a compact convex set. In our case the compact convex set is $B(\mathcal{L})$ and its extremal part is $S(\mathcal{L})$. At the case at hand, Mielnik's procedure starts with first choosing an extremal point $\xi \in S(\mathcal{L})$ and selecting all affine functions l satisfying $l(\xi) = 1$ and $0 \leq l \leq 1$ on $B(\mathcal{L})$. Any such function can be written

$$l(\eta) = \frac{a + (\xi, \eta) + 1}{a + 2}, \tag{89}$$

with $a \geq 0$ and $\eta \in \mathcal{L}$ arbitrarily. Now we have to vary over all these affine functions,

$$p_M(\eta, \xi) := \min_l l(\eta) = \min_a \frac{a + (\xi, \eta)}{2 + a}, \tag{90}$$

to get Mielnik's transition probability

$$p_M(\xi, \eta) = \frac{1 + (\xi, \eta)}{2}. \tag{91}$$

Returning to $H, H' \in B_N$, we can write

$$p_M(H, H') := p_M(\xi_H, \xi_{H'}) = \frac{1 + G(H, H')}{2} \tag{92}$$

and, in becoming even more special by choosing two density operators for H and H' in the equation above, we arrive at

$$D_M(\rho_1, \rho_2) = 2\sqrt{1 - p_M(\rho_1, \rho_2)} = 2\sin\frac{\alpha}{2}, \tag{93}$$

(using $2\cos^2 = 1 + \cos$) and also at

$$D'_M(\rho_1, \rho_2) = 2\arccos\sqrt{p_M(\rho_1, \rho_2)}. \tag{94}$$

8   Concluding remarks

In this paper we analyzed various bounds for quantum fidelity. Two quantities, we propose to call sub– and super–fidelity, posses particularly nice properties. On one hand these quantities form universal lower and upper bounds for the fidelity. Moreover, with respect to the tensor product they display sub– and super–multiplicativity.

On the other hand, quantities E and G are much easier to calculate than the original fidelity F. To compute any of these bounds it is enough to evaluate three traces only. Thus one can expect, the quantities introduced in this paper might become useful for various tasks of the theory of quantum information processing. Furthermore, under a realistic assumption that several copies of both states are available, it is possible to design a scheme to measure experimentally sub– and super–fidelity between arbitrary mixed states. For instance, the measurement of super–fidelity is possible if one has three copies of each state. In this paper we have worked out concrete schemes of such experiments concerning the super–fidelity between any two mixed states representing the polarization of photons.

Acknowledgements

Note added

After this paper was submitted we learned about a related work by Mendonca et al. [43] in which the super–fidelity was independently introduced and was called an 'alternative fidelity' measure. In this valuable work the authors provide an alternative proof of super–multiplicativity of G, discuss its relation to the trace distance and analyze the distance G induces into the space of mixed quantum states, and prove that G is *jointly* concave in its two arguments.

References

1. R. Jozsa, *J. Mod. Opt.* **41**, 2315 (1994).
2. A. Uhlmann, *Rep. Math. Phys.* 9 273 (1976).
3. A. Uhlmann, Conference "Probability and Operator Algebras with Applications in Mathematical Physics", Erdö Research Center, 2000 (unpublished).
4. A. Uhlmann, *The metric of Bures and the geometric phase,* in *Groups and related Topics*, ed. R. Gierelak *et al.*, Kluver, Dodrecht, 1992.
5. J. Dittmann, The scalar curvature of the Bures metric on the space of density matrices, *J. Geom. Phys.* **31**, 16-24 (1999).
6. J.-L. Chen, L. Fu, A.A. Ungar, and X.-G. Zhao, Alternative fidelity measure between two states of an N–state quantum system, *Phys. Rev. A* **65**, 054304 (2002).
7. R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, 1997.

8. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

9. I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, Cambridge, 2006.

10. H. Araki and G. Raggio: A remark on transition probability, *Lett. Math. Phys.* **6**, 237–240 (1982).

11. P. M. Alberti, A note on the transition probability over C*-algebras, *Lett. Math. Phys.* **7**, 25–32 (1983).

12. P. M. Alberti and A. Uhlmann, Stochastic Linear Maps and Transition Probability, *Lett. Math. Phys* **7**, 107–112 (1983).

13. D. Markham, J. A. Miszczak, Z. Puchała, K. Życzkowski, Quantum state discrimination: a geometric approach, *Phys. Rev. A* **77**, 042111 (2008).

14. M. Hübner, Explicit computation of the Bures distance for density matrices, *Phys. Lett. A* **163**, 239 (1992).

15. M. Buliga, *Majorisation with applications to the calculus of variations* arXiv:math/0105044.

16. I. D. Coope, On Matrix Trace Inequalities and Related Topics for Products of Hermitian Matrices *J. Math. Analysis and Applications* **188**, 3 (1994), pp. 999-1001.

17. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

18. R. A. Horn, C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 2006.

19. H. Lütkenpohl, *Handbook of Matrices*, John Wiley & Sons, 1996.

20. A. K. Ekert, C. M. Alves, D. K. Oi, M. Horodecki, P. Horodecki and L. C. Kwek , *Phys. Rev. Lett.* **88**, 217901 (2002).

21. R. Augusiak, M. Demianowicz and P. Horodecki, *Phys. Rev. A* **77**, 030301(R) (2008).

22. R. Augusiak, J. Stasińska and P. Horodecki, *Phys. Rev. A* **77**, 012333 (2008).

23. R. F. Werner, *Phys. Rev. A* **40**, 4277 (1989).

24. E. Knill and R. Laflamme, *Phys. Rev. Lett.* **81**, 5672 (1998).

25. F. A. Bovino, G. Castagnoli, A. Ekert, P. Horodecki, C. Moura Alves and A. V. Sergienko, *Phys. Rev. Lett* **95**, 240407 (2005).

26. S. P. Walborn, P. H. Souto Ribeiro, L. Davidovich, F. Mintert and A. Buchleitner, *Nature* **440**, 1022 (2006).

27. N. A. Peters, J. B. Altepeter, D. A. Branning, E. R. Jeffrey, T.-C. Wei and P. G. Kwiat, *Phys. Rev. Lett.* **92**, 133601 (2004).

28. C. M. Alves and D. Jaksch, *Phys. Rev. Lett.* **93**, 110501 (2004).

29. S. J. van Enk, *Can measuring entanglement be easy?*, arXiv:quant-ph/0606017.

30. S. J. van Enk, N. Lutkenhaus, H. J. Kimble, *Phys. Rev. A* **75**, 052318 (2007).

31. C. M. Caves *et al.*, *J. Math. Phys.* **43**, 4537 (2002).

32. S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, *Phys. Rev. Lett.* **72**, 3439 (1994).

33. D. Petz and C. Sudár, Geometries of quantum states, *J. Math. Phys.* **37**, 2662 (1996).

34. K. Życzkowski and W. Słomczyński, Monge metric on the sphere and geometry of quantum states, *J. Phys.* **A 34**, 6689 (2001)

35. D.J.C. Bures, An extension of Kakutani theorem on infinite product measures to the tensor of semifinite W*-algebras, *Trans. Am. Math. Soc.* **135**, 199 (1969).

36. A. Uhlmann, Geometric phases and related structures, *Rep. Math. Phys.* **36**, 461 (1995).

37. D. Petz, Monotone metrics on matrix spaces, *Linnear Algebra Appl.* **244**, 81 (1996).

38. A. Gilchrist, N. K. Langford, M. A. Nielsen, Distance measures to compare real and ideal quantum processes *Phys. Rev. A* **71** 062310 (2005).

39. C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, *IEEE Trans. Inf. Theor.* **45**, 1216-1227 (1999).

40. B. Mielnik, Theory of filters, *Commun. Math. Phys.* **15** 15 (1974).

41. M. Marcus and L. Lopes, Inequalities for symmetric functions and Hermitian matrices, *Canad. J. Math.* **8**, pp. 524-531 (1956).

42. P. Biler and A. Witkowski, *Problems in Mathematical Analysis*, CRC Press, 1990.

43. P. E. M. F. Mendonca, R. d. J. Napolitano, M. A. Marchiolli, C. J. Foster, and Y.-C. Liang, *An alternative fidelity measure for quantum states*, arXiv:0806.1150

Appendix A   Algebraic facts

In this appendix we collect useful algebraic facts, which are used in the main body of the paper.

Fact 1 (From corollary IX.5.3 in [7]) *For any positive matrices A and B and every unitarily invariant norm $||| \cdot |||$ we have*

$$|||A^\nu B^{1-\nu}||| \leq |||A|||^\nu |||B|||^{1-\nu}, \tag{A.1}$$

where $\nu \in [0,1]$.

Fact 2 (From corollary IX.5.4 in [7]) *For any positive matrices A and B and every unitarily invariant norm $||| \cdot |||$ we have*

$$|||A^\nu B^\nu||| \leq |||\mathbb{I}|||^{1-\nu} |||AB|||^\nu, \tag{A.2}$$

where $\nu \in [0,1]$.

Next two facts can be found in [16].

Fact 3 *Matrix AB is similar to matrices $\sqrt{A}B\sqrt{A}$ and $\sqrt{B}A\sqrt{B}$.*

Fact 4 *For positive matrices A and B matrix AB has positive eigenvalues.*

Fact 5 *If $p_1 + p_2 + \cdots + p_n = 1$ and $p_i \geq 0$ then*

$$1 - p_1^2 - p_2^2 - \cdots - p_n^2 = \sum_{i \neq j} p_i p_j. \tag{A.3}$$

Proposition 4 *Let g be defined as*

$$g(x) = \sum_{i \neq j} \sqrt{x_i}\sqrt{x_j}. \tag{A.4}$$

For $x, y \in \mathbb{R}_+^n$ such that

$$\prod_{i=1}^k x_i \leq \prod_{i=1}^k y_i, \text{ for } k = 1, \ldots, n, \tag{A.5}$$

with equality for $k = n$, we have

$$g(x) \leq g(y). \tag{A.6}$$

Proof.   We introduce notation

$$g_i(\cdot) = \frac{\partial g}{\partial x_i}(\cdot). \tag{A.7}$$

Direct computation shows that function g satisfies

$$u_1 g_1(u) \geq u_2 g_2(u) \geq \cdots \geq u_n g_n(u), \tag{A.8}$$

for $u \in \mathbb{R}^n$ such that $u_1 \geq u_2 \geq \cdots \geq u_n \geq 0$. We denote $\alpha_i = \log(x_i)$ and $\beta_i = \log(y_i)$. Note that (A.5) can be rewritten as

$$\sum_{i=1}^{k} \alpha_i \leq \sum_{i=1}^{k} \beta_i \text{ for } k = 1, \ldots, n, \tag{A.9}$$

with equality for $k = n$. We define new function

$$h(v) = g(e^{v_1}, e^{v_2}, \ldots, e^{v_n}). \tag{A.10}$$

For a given vector u such that $u_1 \geq u_2 \geq \cdots \geq u_n \geq 0$, and $v_i = \log(u_i)$ we have

$$v_1 \geq v_2 \geq \cdots \geq v_n. \tag{A.11}$$

Using (A.8) we can write

$$e^{v_1} g_1(e^{v_1}, e^{v_2}, \ldots, e^{v_n}) \geq \cdots \geq e^{v_1} g_n(e^{v_1}, e^{v_2}, \ldots, e^{v_n}). \tag{A.12}$$

Now from above and (A.10) we have

$$h_1(v) \geq h_2(v) \geq \cdots \geq h_n(v). \tag{A.13}$$

Note now that function h satisfies condition from [15, Theorem 3.6] and thus it is *Schur-convex*, so

$$h(\alpha) \leq h(\beta). \tag{A.14}$$

Using (A.10) we can write

$$g(x) \leq g(y). \tag{A.15}$$

Thus the proof is complete. □.

Fact 6 (Hölder's inequality [19]) *For $a > 1, b = a/(a-1)$ and positive semidefinite A and B we have*

$$\text{tr}(AB) \leq (\text{tr}A^a)^{1/a}(\text{tr}B^b)^{1/b}. \tag{A.16}$$

Fact 7 *For density matrices A and B we have*

$$1 - \sqrt{\text{tr}A^2}\sqrt{\text{tr}B^2} \geq \sqrt{1 - \text{tr}A^2}\sqrt{1 - \text{tr}B^2}. \tag{A.17}$$

Proof. This inequality can be rewritten in equivalent form

$$1 - 2\sqrt{\text{tr}A^2\text{tr}B^2} + \text{tr}A^2\text{tr}B^2 \geq 1 - \text{tr}A^2 - \text{tr}B^2 + \text{tr}A^2\text{tr}B^2, \tag{A.18}$$

which is equivalent to

$$\sqrt{\text{tr}A^2\text{tr}B^2} \leq \frac{\text{tr}A^2 + \text{tr}B^2}{2}. \tag{A.19}$$

This completes the proof since for any positive numbers the arithmetic mean is always greater than or equal to the geometric mean. □.

Fact 8 (Maclaurin inequality [42, p. 5]) *For a given matrix A of rank r and with r positive eigenvalues we have*

$$\sqrt[k]{\frac{s_k(A)}{\binom{r}{k}}} \geq \sqrt[k+1]{\frac{s_{k+1}(A)}{\binom{r}{k+1}}} \tag{A.20}$$

for $1 \leq k < r$.

Appendix B   Proof of the lower bound (10)

To prove that sub–fidelity E is not larger than fidelity F, let us take a look at equations (15) and (27) in which both quantities are expressed in terms of the second symmetric function. We can rewrite the function s_2, which forms fidelity,

$$s_2(\sqrt{A^{1/2}BA^{1/2}}) = \sum_{i<j} \lambda_i(\sqrt{A^{1/2}BA^{1/2}})\lambda_j(\sqrt{A^{1/2}BA^{1/2}}) \tag{B.1}$$

$$= \sum_{i<j} \sqrt{\lambda_i(A^{1/2}BA^{1/2})}\sqrt{\lambda_j(A^{1/2}BA^{1/2})} \tag{B.2}$$

$$= \sum_{i<j} \sqrt{\lambda_i(AB)}\sqrt{\lambda_j(AB)}. \tag{B.3}$$

The last equality is the consequence of similarity of matrices $A^{1/2}BA^{1/2}$ and AB. Making use of subadditivity of square root we obtain

$$s_2(\sqrt{A^{1/2}BA^{1/2}}) \geq \sqrt{\sum_{i<j} \lambda_i(AB)\lambda_j(AB)} \tag{B.4}$$

$$= \sqrt{s_2(AB)}. \tag{B.5}$$

As a consequence we get

$$F(A,B) = \mathrm{tr}AB + 2s_2(\sqrt{A^{1/2}BA^{1/2}}) \geq \mathrm{tr}AB + 2\sqrt{s_2(AB)} = E(A,B). \tag{B.6}$$

Appendix C   Proof of the lower bound (11)

To prove inequality (11) we use Fact 8 (Maclaurin inequality) and obtain

$$\left(\frac{s_2\left(\sqrt{A^{1/2}BA^{1/2}}\right)}{\binom{r}{2}}\right)^{1/2} \geq \left(\frac{s_r\left(\sqrt{A^{1/2}BA^{1/2}}\right)}{\binom{r}{r}}\right)^{1/r}. \tag{C.1}$$

Using Fact 3 we get

$$s_2\left(\sqrt{A^{1/2}BA^{1/2}}\right) \geq \binom{r}{2}\left(s_r\left(\sqrt{A^{1/2}BA^{1/2}}\right)\right)^{2/r} = \binom{r}{2}\left(\prod_{i=1}^{r} \lambda_i\left(\sqrt{A^{1/2}BA^{1/2}}\right)\right)^{2/r}$$

$$= \binom{r}{2}\left(\prod_{i=1}^{r} \sqrt{\lambda_i\left(A^{1/2}BA^{1/2}\right)}\right)^{2/r} = \binom{r}{2}\left(\prod_{i=1}^{r} \lambda_i\left(AB\right)\right)^{1/r}$$

$$= \binom{r}{2}\sqrt[r]{s_r(AB)}.$$

Now using (15) we write

$$F(A,B) = \text{tr}AB + 2s_2\left(\sqrt{A^{1/2}BA^{1/2}}\right) \geq \text{tr}AB + r(r-1)\sqrt[r]{s_r(AB)}. \tag{C.2}$$

Appendix D   Proofs of Lemmas

Proof of Lemma 1.   Observe that the matrix $\rho_1^{1/2}\rho_2\rho_1^{1/2}$ is similar to $\rho_1\rho_2$ and thus

$$
\begin{aligned}
2s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) &= \sum_{i\neq j} \lambda_i\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right)\lambda_j\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) \\
&= \sum_{i\neq j}\sqrt{\lambda_i\left(\rho_1^{1/2}\rho_2\rho_1^{1/2}\right)\lambda_j\left(\rho_1^{1/2}\rho_2\rho_1^{1/2}\right)} \\
&= \sum_{i\neq j}\sqrt{\lambda_i\left(\rho_1\rho_2\right)\lambda_j\left(\rho_1\rho_2\right)},
\end{aligned}
$$

where $\lambda_i(A)$ denotes i^{th} eigenvalue of a matrix A.

Let us define a function $g : R^n \to R$ which acts on a vector $\vec{x}$ of non-negative numbers

$$g(\vec{x}) := \sum_{i\neq j}\sqrt{x_i x_j}. \tag{D.1}$$

It allows one to rewrite

$$2s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) = g\left(\vec{\lambda}(\rho_1\rho_2)\right), \tag{D.2}$$

where $\vec{\lambda}(A)$ denotes the vector of eigenvalues of A.

From [18, Theorem 3.3.2 and 3.3.4] we obtain

$$\prod_{i=1}^{k}\lambda_i(\rho_1\rho_2) \leq \prod_{i=1}^{k}\lambda_i(\rho_1)\lambda_i(\rho_2) \text{ for } k = 1, \ldots, n, \tag{D.3}$$

with equality for $k = n$. Making use of Proposition 4 from Appendix A with $x_i = \lambda_i(\rho_1\rho_2)$ and $y_i = \lambda_i(\rho_1)\lambda_i(\rho_2)$ we obtain

$$g\left(\vec{\lambda}(\rho_1\rho_2)\right) \leq g\left(\vec{\lambda}(\rho_1) \circ \vec{\lambda}(\rho_2)\right), \tag{D.4}$$

where $\circ$ denotes Hadamard product, [18, Definition 7.5.1]. Now making use of (D.2) we get

$$s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right) \leq s_2\left(\sqrt{\text{diag}(\vec{\lambda}(\rho_1))\text{diag}(\vec{\lambda}(\rho_2))}\right). \tag{D.5}$$

And thus the proof is complete.   □.

Proof of Lemma 2.   For given density matrices ρ_1, ρ_2 with eigenvalues $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ respectively. We denote diagonal matrices with entries on diagonal $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ as $\text{diag}(p), \text{diag}(q)$ respectively.

Rewriting the second elementary function s_2 we obtain

$$2s_2\left(\sqrt{\text{diag}(p)\text{diag}(q)}\right) = \sum_{i\neq j}\sqrt{p_i q_i}\sqrt{p_j q_j}. \tag{D.6}$$

On the other hand

$$2\sqrt{s_2(\rho_1)s_2(\rho_2)} = \sqrt{\left(1 - \sum p_i^2\right)\left(1 - \sum q_i^2\right)}. \tag{D.7}$$

Let us define vectors $x, y \in \mathbb{R}^{n^2}$

$$x_{i,j} = \sqrt{p_i p_j}(1 - \delta_{i,j}), \ y_{i,j} = \sqrt{q_i q_j}(1 - \delta_{i,j}), \tag{D.8}$$

where $x_{i,j} = x_{n(i-1)+j}$. Using Cauchy–Schwarz inequality

$$|\langle x|y\rangle| \leq \sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle}, \tag{D.9}$$

we get

$$\sum_{i\neq j} \sqrt{p_i q_i}\sqrt{p_j q_j} \leq \sqrt{\left(1 - \sum p_i^2\right)\left(1 - \sum q_i^2\right)}. \tag{D.10}$$

This completes the proof. □.

Appendix E The case $N = 3$

In this section we are going to study the fidelity of two states ρ_1 and ρ_2 in the case where the rank r of their product $\rho_1\rho_2$ is not greater than 3. As in Section 2 we will denote eigenvalues of $\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}$ by λ_i, so eigenvalues of $\rho_1^{1/2}\rho_2\rho_1^{1/2}$ are given by $\{\lambda_i^2\}$ and by similarity we have that the eigenvalues of $\rho_1\rho_2$ are also given by $\{\lambda_i^2\}$. Since $r \leq 3$, not more than three eigenvalues of $\rho_1\rho_2$ are positive, so the third symmetric function (6) reads $s_3(\rho_1\rho_2) = (\lambda_1\lambda_2\lambda_3)^2$. This is so for any two states of a qutrit, so for $N = 3$ one has $s_3(\rho_1\rho_2) = \det(\rho_1\rho_2)$.

Consider now the expression for fidelity (4) which can be rewritten with the use of the second symmetric function,

$$F(\rho_1, \rho_2) = \mathrm{tr}\rho_1\rho_2 + 2s_2\left(\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right). \tag{E.1}$$

The square of the symmetric function presented in the above equation, can be written as $\left(\sum_{i<j}\lambda_i\lambda_j\right)^2 = \sum_{i<j}\lambda_i^2\lambda_j^2 + R$. The reminder R, defined implicitly by this equation, is equal to zero if $r \leq 2$ and the sum consists of a single term only. It is difficult to handle R generally. But if $r = 3$ one has

$$R = \lambda_1\lambda_2(\lambda_2\lambda_3 + \lambda_3\lambda_1) + ... \ = 2(\lambda_1\lambda_2\lambda_3)(\lambda_1 + \lambda_2 + \lambda_3). \tag{E.2}$$

However, in the particular case $r \leq 3$ discussed here used to (3) one has $\lambda_1 + \lambda_2 + \lambda_3 = \sqrt{F}$ while $\lambda_1\lambda_2\lambda_3 = \sqrt{s_3(\rho_1\rho_2)}$. Combining this with (4) we get the equation for fidelity satisfied for $r \leq 3$

$$F = \mathrm{tr}\rho_1\rho_2 + 2\sqrt{s_2(\rho_1\rho_2) + 2\sqrt{F}\sqrt{s_3(\rho_1\rho_2)}}. \tag{E.3}$$

In the case $r \leq 2$ the third function s_3 vanishes, so this equation leads to an expression, $F = \mathrm{tr}\rho_1\rho_2 + 2\sqrt{s_2(\rho_1\rho_2)} = \mathrm{tr}\rho_1\rho_2 + 2\lambda_1\lambda_2$, already discussed in Section 2.

Another relation for fidelity is due to the fact that an assumption $r \leq 3$ implies that

$$\sum_{j<k} \lambda_j \lambda_k = (\lambda_1 \lambda_2 \lambda_3)(\lambda_1^{-1} + \lambda_2^{-1} + \lambda_3^{-1}). \tag{E.4}$$

Therefore in this case one has

$$s_2\left(\sqrt{A_1^{1/2} B A^{1/2}}\right) = \sqrt{\det(AB) F(1/A, 1/B)}. \tag{E.5}$$

Lifting for a moment the assumption that the arguments of fidelity have to be normalized, we arrive therefore at another equation for fidelity satisfied for $r \leq 3$,

$$F(A, B) = \mathrm{tr}\, AB + 2\sqrt{\det(AB) F(1/A, 1/B)}. \tag{E.6}$$