REVIEW

Quantum fidelity measures for mixed states

IOP Publishing Reports on Progress in Physics
Rep. Prog. Phys. **82** (2019) 076001 (29pp) https://doi.org/10.1088/1361-6633/ab1ca4

Review

Quantum fidelity measures for mixed states

Yeong-Cherng Liang[1,6], Yu-Hao Yeh[1], Paulo E M F Mendonça[2,3], Run Yan Teh[4], Margaret D Reid[4,5] and Peter D Drummond[4,5,7]

[1] Department of Physics, National Cheng Kung University, Tainan 701, Taiwan
[2] Academia da Força Aérea, Caixa Postal 970, 13643-970, Pirassununga, SP, Brazil
[3] Melbourne Graduate School of Education, University of Melbourne, Melbourne, VIC 3010, Australia
[4] Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne, VIC 3122, Australia
[5] Institute of Theoretical Atomic, Molecular and Optical Physics (ITAMP), Harvard University, Cambridge, MA, United States of America
[6] Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan

E-mail: ycliang@mail.ncku.edu.tw, pmendonca@gmail.com, rteh@swin.edu.au, mdreid@swin.edu.au and pdrummond@swin.edu.au

Received 17 October 2018, revised 8 February 2019
Accepted for publication 25 April 2019
Published 24 June 2019

Corresponding Editor Professor Dr Matthias Troyer

Abstract

Applications of quantum technology often require fidelities to quantify performance. These provide a fundamental yardstick for the comparison of two quantum states. While this is straightforward in the case of pure states, it is much more subtle for the more general case of mixed quantum states often found in practice. A large number of different proposals exist. In this review, we summarize the required properties of a quantum fidelity measure, and compare them, to determine which properties each of the different measures has. We show that there are large classes of measures that satisfy all the required properties of a fidelity measure, just as there are many norms of Hilbert space operators, and many measures of entropy. We compare these fidelities, with detailed proofs of their properties. We also summarize briefly the applications of these measures in teleportation, quantum memories and quantum computers, quantum communications, and quantum phase-space simulations.

Keywords: fidelity, quantum information, quantum technology, entanglement, quantum memory, quantum logic

(Some figures may appear in colour only in the online journal)

Contents

[7] Author to whom any correspondence should be addressed.

1. Introduction

Fidelity is a central concept to quantum information. It provides a mathematical prescription for the quantification of the *degree of similarity* of a pair of quantum states. In practice, there are many situations where such a comparison is useful. For example, since any experimental preparation of a quantum state is limited by imperfections and noise, one is generally interested to find how close the state actually produced is to the state whose production was intended. This is a common issue in quantum communications and quantum computing, where one is interested in either generating or sending precisely defined quantum states in the face of noise and other sources of error.

Another common application arises in the context of entanglement quantification [1]: the closer a given quantum state is to the set of separable states, the less entangled it is, and vice-versa (see, however, [2]). Measuring and computing fidelities between quantum states is at the heart of various quantum information tasks. In recent years, fidelity measure has also been applied extensively to study quantum phase transitions. For a review on this subject, see [3, 4] and references therein.

For pure states, fidelity is well-defined. Yet pure states, by their nature, are exactly what one does *not* expect in a noisy, real-world environment. Moreover, in large quantum systems, one needs to measure exponentially many parameters in order fully determine the quantum states, thus making the task infeasible in practice. Instead, one could hope to obtain some partial information about the produced quantum states, for example, by performing tomography on some (random) subsets of the multipartite quantum states. Importantly, a generic multipartite pure state is highly entangled across any bipartition [5], hence its reduced states on subsets are typically mixed. Thus, more realistically, one must expect to deal with impure or mixed target states that are obtained by examining a subsystem obtained by tracing over a larger environment.

In all real-world experiments, there is also noise or coupling to an environment. This is an essential part of the quantum-classical transition, since it is the coupling to an environment that allows measurements. From the point of view of a quantum technologist or engineer, the environment causes decoherence, and this is the main challenge in many quantum technology applications. Just as relevant is the fact that one often wishes to analyze the performance of one component—for example, a quantum gate—embedded in a larger device, so that the environment is an essential part of the system of interest, as explained above. Thus, one may argue that the fidelity for mixed states is generic, and therefore the most practical type of fidelity. Pure state fidelity represents an idealized case only, which is typically non-scalable. For mixed states, however, fidelity has no clearly unique definition, and a number of different approaches exist.

Here the question is really one of distance: given two *density matrices*, how close are they to each other in the appropriate Hilbert space? This is an important issue, for example, if one wishes to understand how accurately a given approximate calculation replicates that of some target density matrix. The concept of distance in any vector space is never a uniquely defined concept without further considerations being applied.

There are other situations where mixed state fidelity might seem to be important, but frequently they are more subtle than would appear at first sight. For example, while it is known [6] that one cannot clone quantum states, there is great interest in the idea of optimal cloning [7, 8], in which a state is copied as well as is allowed by quantum theory. This requires one to quantify the fidelity of the clone or copy, in order to decide if a given cloning strategy is really optimal. Similarly, while quantum memories and quantum teleportation are permitted by quantum mechanics, the real world of the laboratory leads to inevitable noise and errors. Again, it is a mixed state fidelity that is important, since the original state that is teleported or stored was typically not a pure state originally. These applications may involve multiple states as well, which requires not just fidelities, but averages over them.

In this review, we analyze these issues by considering measures of fidelity that apply to mixed as well as pure states. In section 2, we review the properties of a fidelity measure as defined by Jozsa [9], which apply to the mixed state case, and illustrate that these are satisfied by a large number of proposed

definitions of fidelity measure. This leads us to compare the different properties of these measures, which is the main purpose of this review. In sections 3 and 4 we summarize the respective mathematical properties, and make comparisons of the values the different fidelity measures can have. We choose for specific comparison finite-dimensional 'qudit' states with varying degrees of purity, and also make comparisons based on randomly generated density operators. Finally in section 5, we review the application of fidelity measures to quantum information protocols such as teleportation [10], quantum memories [11] and quantum gates [12]. Here, the fidelity measure can indicate a level of security for a quantum state transfer, or indicate the effectiveness of a logic operation. We present a useful tool—phase space fidelity—for evaluating the theoretical prediction of fidelity where systems and theories are complex.

As applications are based on experimental measurements, we conclude the review with a short discussion of the different types of experimental fidelity measurements that have been reported in the literature. These range from measurements on atomic states such as in ion traps, to photonic qubit states that might imply entanglement based on post-selection. The issue of defining the appropriate Hilbert space for the fidelity measurement is also discussed. In the summary, we present our main conclusion, which is that for some situations different fidelity measures to the commonly used Uhlmann–Jozsa measure may be advantageous. The appendix gives detailed proofs of novel results where the proofs are too lengthy for the main part of the review.

2. Fidelity measures for mixed states

2.1. Measured and relevant Hilbert spaces

Given the wide variety of quantum technologies involved, it is not surprising that there are many ways to measure fidelity. These may depend on the applications envisaged, or simply on what is measurable in an experiment. Yet these are quite different issues. What is feasible in an experiment may not be the fidelity measurement that is needed. To understand this point, we must introduce the concept of measured and relevant Hilbert space, which is fundamental to understanding mixed state fidelity.

To motivate this analysis, we note that in certain types of relatively well-structured quantum states, such as the Greenberger–Horne–Zeilinger [13] states, large-dimensional fidelity measurements have been reported, with 10–14 qubit Hilbert spaces [14–16] being treated in ion-trap, photonic and superconducting quantum circuit environments. Such measurements are especially important in allowing the detection of decoherence and super-decoherence, where the decoherence rates increase with system size [17, 18].

These experiments are impressive demonstrations, but scaling them to even larger sizes is likely to become increasingly difficult owing to their exponential complexity, leading to exponentially many measurements being required in a general tomographic measurement. For this reason, we expect measurements of mixed state fidelity to become more common in large Hilbert spaces, as explained in the introduction.

The first difficulty in carrying out a physical fidelity measurement is to identify the relevant Hilbert space. All quantum systems are coupled to other modes of the universe, but the state of Betelgeuse is usually only relevant if one wishes to communicate over a 640 light year distance. Thus, a useful fidelity measurement only measures the relevant fidelity. Suppose the global quantum state is divided into the relevant and irrelevant parts, with an orthogonal basis $|\psi_i\rangle_{\rm rel}\,|\phi_i\rangle_{\rm irr}$:

$$|\Psi\rangle = \sum_{ij} C_{ij}\,|\psi_i\rangle_{\rm rel}\,|\phi_j\rangle_{\rm irr}\,. \tag{2.1}$$

Next, the relevant density matrix is obtained by the usual partial trace procedure, so that

$$\rho = \sum_{ijk} C_{ij}C^*_{kj}\,|\psi_i\rangle_{\rm rel}\,\langle\psi_k|_{\rm rel}. \tag{2.2}$$

The irrelevant portion, $|\phi\rangle_{\rm irr}$, of the quantum state may change in time. Yet given a large enough separation, and for localized interactions, this will generally not alter the relevant density matrix, owing to the trace over irrelevant parts. However, if the product does not factorize initially, so the relevant and irrelevant parts are entangled, then one does always have a mixed state. This is the generic situation, and one should not assume that a state is pure, without tomographic measurements that verify purity.

While subdivision is necessary, *where should the dividing line be*? For example, the internal states of all the different ions in an ion trap, and their motional states, may all be relevant to an ion-trap quantum computer. However, the state of motion of the vacuum chamber will not be. Thus, the mixed state fidelity is a *relative* measure. It depends on the relevant Hilbert spaces. One will get different results depending on how large the relevant Hilbert space is, which is application-dependent.

Yet the relevant space in many cases may still be much larger than the one in which the fidelity is *measured*. For example, one might only be able to measure the internal states of one or two ions [19]. The total relevant Hilbert space can be larger than this. This means that there is some lost of information which may be important to the application. One is still entitled to claim to have measured a fidelity, but it is clearly not the only relevant fidelity.

This means that the definition and understanding of a mixed state fidelity is very central to quantum technologies. The original idea of fidelity [20], as first introduced to the quantum information community, is not immediately applicable to an arbitrary pair of density matrices. However, it is implicit in this definition that for a pair of pure states $\rho = |\phi\rangle\langle\phi|$ and $\sigma = |\psi\rangle\langle\psi|$, their fidelity should be defined by the transition probability between the two states, i.e.

$$\mathcal{F}(|\phi\rangle\langle\phi|\,,|\psi\rangle\langle\psi|) = |\langle\psi|\phi\rangle|^2\,. \tag{2.3}$$

As pointed out subsequently [9], this is indeed a natural candidate for a fidelity measure since it corresponds to the closeness of states in the usual geometry of Hilbert space.

When one of the quantum states, say, ρ is mixed, there also exists a generalization of equation (2.3) in terms of the transition probability between the two states, namely,

$$\mathcal{F}(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\,\rho\,|\psi\rangle\,. \tag{2.4}$$

Note that this was also implicitly defined in the original fidelity measure [20]. In [21], this expression has also been referred as Schumacher's fidelity.

2.2. Desirable properties of mixed state fidelities

As one might expect, not every function of two density matrices provides a physically reasonable generalization of equation (2.4) to a pair of mixed states.

For example, at first glance it may seem that

$$\mathcal{F}(\rho, \sigma) = \mathrm{tr}(\rho\,\sigma), \tag{2.5}$$

serves as a useful generalization of equation (2.4) to an arbitrary pair of mixed states. However, the Hilbert–Schmidt inner product leads to an unsatisfactory generalization of fidelity [9]. For example, let us denote by $\mathbf{1}_d$ the identity operator acting on a d-dimensional Hilbert space. Then, adopting equation (2.5) as the mixed state fidelity would imply that all pairs of density matrices for a two-state or qubit system of the form $(\mathbf{1}_2/2, |\phi\rangle\langle\phi|)$, are just as similar as the identical pair $(\mathbf{1}_2/2, \mathbf{1}_2/2)$.

This problem is soluble through a suitable normalization, as we show below, but it illustrates the need for a suitable definition of a fidelity measure. In order to avoid such difficulties, Jozsa proposed the following list of fidelity axioms [9], which should be satisfied by any sensible generalization of equation (2.4) to a pair of mixed states:

(J1a) $\mathcal{F}(\rho, \sigma) \in [0, 1]$
(J1b) $\mathcal{F}(\rho, \sigma) = 1$ *if and only if* $\rho = \sigma$
(J2) $\mathcal{F}(\rho, \sigma) = \mathcal{F}(\sigma, \rho)$
(J3) $\mathcal{F}(\rho, \sigma) = \mathrm{tr}(\rho\,\sigma)$ if either ρ or σ is a pure state
(J4) $\mathcal{F}(U\rho\,U^\dagger, U\sigma U^\dagger) = \mathcal{F}(\rho, \sigma)$ for all unitary U

Henceforth, we shall refer to this set of conditions as Jozsa's axioms.

Apart from these, it is convenient to append to this list the requirement that any fidelity measure should vanish when applied to quantum states of orthogonal support, i.e.

(J1c) $\mathcal{F}(\rho, \sigma) = 0$ *if and only if* $\rho\,\sigma = 0$

Throughout, the requirements (J1)–(J4) will be taken as the most basic requirements to be satisfied by any generalization of fidelity measure for a pair of mixed states.

One may be interested in further subtleties in determining fidelity, beyond the Jozsa axioms, depending on the application of the measure. For example, in the case of a mixed system, the idea of a state-by-state fidelity could be important. One may wish to investigate a cloning, communication or quantum memory experiment. Suppose, in the experiment, the input state is an unknown qudit state of dimension d, and has maximum entropy. In other words, one has $\rho = \mathbf{1}_d/d$. According to Jozsa's criterion, the highest fidelity output state σ would be another, identical, maximal entropy state. Yet this would not be a useful criterion on its own—it is a necessary, but not sufficient measure.

Here, one might wish to have the maximum fidelity for every expected input state, pure or mixed, with an appropriate weight given by their relative probability. This requires an understanding of the correlations between the input and output states, given a communication alphabet $\rho_A, \rho_B, \ldots$, which is related to the conditional information measures found in communication theory [22]. Such more general issues cannot be investigated by just using simple fidelity measures of the Jozsa type, and more sophisticated process fidelity measures [23] are needed, which we discuss in detail later. However, even in this case, a fidelity measure can be useful provided it is applied relative to every input density matrix in the relevant communication alphabet, and then averaged with its probability of appearance. This is called the average fidelity:

$$\langle\mathcal{F}(\rho, \sigma)\rangle = \sum P_j \mathcal{F}(\rho_j, \sigma_j), \tag{2.6}$$

where the pair ρ_j, σ_j occur with probability P_j. However, the average fidelity may involve averages over mixed state fidelities. Hence this concept is a relative one, and depends on the precise definition of fidelity used.

2.3. Uhlmann–Jozsa fidelity

The most widely-employed generalization of Schumacher's fidelity that has been proposed in the literature is the *Uhlmann–Jozsa (U–J) fidelity* $\mathcal{F}_1$ [9, 24], as the maximal *transition probability* between the purification of a pair of density matrices ρ and σ. Our choice of notation for this fidelity will become evident in section 2.5:

$$\mathcal{F}_1(\rho, \sigma) := \max_{|\psi\rangle, |\varphi\rangle} |\langle\psi|\varphi\rangle|^2 = \left(\mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2. \tag{2.7}$$

In order to understand this definition, we note that here, $|\psi\rangle$ is what is called a purification of ρ. A purification is a state in a notional extension of the Hilbert space $\mathcal{H}$ of ρ to an enlarged space $\mathcal{H}' = \mathcal{H} \otimes \mathcal{H}_2$, such that $|\psi\rangle$ is a member of this larger space. The limitation on $|\psi\rangle$ is that when the projector, $|\psi\rangle\langle\psi|$, is traced over the auxiliary Hilbert space $\mathcal{H}_2$, it reduces to ρ, i.e.

$$\rho = \mathrm{tr}_{\mathcal{H}_2}\left[|\psi\rangle\langle\psi|\right].$$

This fidelity measure is often referred simply as *the fidelity*. Nevertheless, the reader should beware that some authors (e.g. those of [3, 12, 25, 26]) have referred, instead, to the square root of $\mathcal{F}_1(\rho, \sigma)$ as the fidelity. We show below that this fidelity measure, $\mathcal{F}_1(\rho, \sigma)$, is one of a large class of similar norm-based measures called $\mathcal{F}_p(\rho, \sigma)$.

Given the wide use of this definition of $\mathcal{F}_1(\rho, \sigma)$, one might naturally wonder if there is any need to search further, especially as this definition carries with it a number of desirable properties. There are also some difficulties with this approach, however, and we list them here:

- The U–J fidelity requires one to calculate or measure traces of square roots of matrices. This is not trivial in cases of large or infinite density matrices.
- The conceptual basis of the U–J fidelity is that both the density matrices being compared are derived from an

identical enlarged space $\mathcal{H}'$. This is not always true in many applications of fidelity.

- Since the U–J fidelity is a maximum over purifications, the measured U–J fidelity on a subspace is always greater or equal to the true relevant fidelity for two pure states. This may introduce a bias in estimating pure-state fidelities given a measurement over a reduced Hilbert space.

This leads to an obvious mathematical question:

- *Does the Jozsa set of requirements lead uniquely to the U–J fidelity, or do other alternatives exist?*

The purpose of this review is to answer this question. We show that, indeed, other alternatives do exist that satisfy the Jozsa axioms. A similar type of situation exists for the quantum entropy, where it is known that there are many entropy-like measures. Some are more suitable for given applications than others (see, e.g. [27–29] and references therein).

2.4. Alternative fidelities

One of the first proposals of a fidelity measure alternative to $\mathcal{F}_1$ was that provided by Chen *et al* in [30]:

$$\mathcal{F}_C(\rho,\sigma) := \frac{1-r}{2} + \frac{1+r}{2}\mathcal{F}_N(\rho,\sigma), \tag{2.8}$$

where $r = \frac{1}{d-1}$, d is the dimension of the state space, and

$$\mathcal{F}_N(\rho,\sigma) := \mathrm{tr}(\rho\,\sigma) + \sqrt{1-\mathrm{tr}(\rho^2)}\sqrt{1-\mathrm{tr}(\sigma^2)}. \tag{2.9}$$

For two-dimensional quantum states, $\mathcal{F}_1|_{d=2} = \mathcal{F}_C|_{d=2} = \mathcal{F}_N|_{d=2}$ [21]. Moreover, $\mathcal{F}_C$ admits a hyperbolic geometric interpretation [31] in terms of the generalized Bloch vectors [32–34].

In 2008, $\mathcal{F}_N$ itself was proposed as an alternative fidelity measure in [21]; the same quantity was also independently introduced in [35] by the name of super-fidelity, as it provides an *upper bound* on $\mathcal{F}_1$.

At about the same time, the square of the quantum affinity $A(\rho,\sigma)$ was proposed in [36] (see also [37]) as a fidelity measure, by the name of *A-fidelity*:

$$\mathcal{F}_A(\rho,\sigma) := \left[\mathrm{tr}\left(\sqrt{\rho}\sqrt{\sigma}\right)\right]^2. \tag{2.10}$$

It is worth noting that, in contrast to $\mathcal{F}_N$, the A-fidelity $\mathcal{F}_A$ provides a *lower bound* on $\mathcal{F}_1$.

The super-fidelity $\mathcal{F}_N$ clearly does not satisfy the axiom (J1c). In response to this [38], one may introduce the quantity

$$\mathcal{F}_{GM}(\rho,\sigma) := \frac{\mathrm{tr}(\rho\,\sigma)}{\sqrt{\mathrm{tr}(\rho^2)\,\mathrm{tr}(\sigma^2)}}, \tag{2.11}$$

which is, instead, incompatible with axiom (J3). $\mathcal{F}_{GM}$ can be seen as the Hilbert–Schmidt inner product between ρ and σ normalized by the geometric mean (GM) of their purities $\mathrm{tr}(\rho^2)$ and $\mathrm{tr}(\sigma^2)$.

Lastly, let us point out another quantity of special interest. The non-logarithmic variety of the quantum Chernoff bound, $\mathcal{F}_Q$, is defined by Audenaert *et al* [39] as:

Table 1. Compatibility of existing fidelity measures against Jozsa's axioms.

	J1a	J1b	J1c	J2	J3	J4
$\mathcal{F}_1$	√	√	√	√	√	√
$\mathcal{F}_Q$	√	√	√	√	√	√
$\mathcal{F}_N$	√	√	×	√	√	√
$\mathcal{F}_C$	√	√	×	√	×	√
$\mathcal{F}_A$	√	√	√	√	×	√
$\mathcal{F}_{GM}$	√	√	√	√	×	√

$$\mathcal{F}_Q(\rho,\sigma) := \min_{0\leqslant s\leqslant 1} \mathrm{tr}(\rho^s\,\sigma^{1-s}). \tag{2.12}$$

This quantity was not originally proposed as a fidelity measure. Instead, it is related to the (asymptotic) probability of error incurred in discriminating between quantum states ρ and σ when one has access to arbitrarily many copies of them. Nonetheless, we will include $\mathcal{F}_Q$ in our subsequent discussion as it does have many desirable properties of a fidelity measure.

In fact, amongst all the generalized fidelities formulas proposed so far, only $\mathcal{F}_Q$ and $\mathcal{F}_1$ fully comply with Jozsa's axioms (see table 1). These will be our main focus in the review out of these previously known fidelities, although we point out that there is also an infinite class of norm-based fidelities that comply with Jozsa's axioms as well as these two. We note however that $\mathcal{F}_Q$ is also computationally challenging. Not only does it involve fractional powers, but one must optimize over a continuous set of candidate measures, each involving different fractional powers.

2.5. Norm-based fidelities

In addition, there are many norm-based fidelity measures that satisfy these axioms. Consider the operator $A = \sqrt{\rho}\sqrt{\sigma}$ whose properties are closely related to the overlap of two density matrices. There is an infinite class of unitarily invariant norms of linear operators, called the Schatten–von-Neumann norms (or more commonly Schatten norms), which can be used to measure the size of such operators. These are defined, for $p \geqslant 1$, as [40]:

$$\|A\|_p \equiv \left(\mathrm{tr}\left[\left(AA^\dagger\right)^{p/2}\right]\right)^{1/p}. \tag{2.13}$$

These norms all satisfy Hölder and triangle inequalities, so that $\|A_1A_2\|_p \leqslant \|A_1\|_{2p}\,\|A_2\|_{2p}$. This particular inequality can be deduced, for example, from corollary IV.2.6 of [40]. They are not yet suitable as fidelity measures, as they must be normalized appropriately to satisfy the fidelity axioms. Hence, keeping this in mind, we define a p-fidelity as:

$$\mathcal{F}_p(\rho,\sigma) := \frac{\left\|\sqrt{\rho}\sqrt{\sigma}\right\|_p^2}{\max\left[\|\sigma\|_p^2, \|\rho\|_p^2\right]}. \tag{2.14}$$

The proof that the axioms are satisfied is given in appendix A. We note that, for the special case of $p = 1$, $\mathcal{F}_1(\rho,\sigma)$ is exactly the Uhlmann–Jozsa fidelity. This follows since for any

Hermitian density matrix σ, $\|\sigma\|_1^2 = (\mathrm{tr}\sigma)^2 = 1$, which is the same for all density matrices. Hence, the normalizing term in this case is

$$\max\left[\|\sigma\|_1^2, \|\rho\|_1^2\right] = \max\left[(\mathrm{tr}\sigma)^2, (\mathrm{tr}\rho)^2\right] = 1. \tag{2.15}$$

In this review, we focus on $\mathcal{F}_1(\rho,\sigma)$ and $\mathcal{F}_2(\rho,\sigma)$, which have especially desirable properties. In particular, $\mathcal{F}_2(\rho,\sigma)$, which is defined as:

$$\mathcal{F}_2(\rho,\sigma) := \frac{\mathrm{tr}(\rho\,\sigma)}{\max\left[\mathrm{tr}(\rho^2), \mathrm{tr}(\sigma^2)\right]}, \tag{2.16}$$

uses the Hilbert–Schmidt operator measure $\|A\|_2$, which is often simpler to calculate than $\|A\|_1$. In fact, all of the even order p-fidelities can be easily evaluated, as they reduce to the form:

$$\mathcal{F}_{2p}(\rho,\sigma) := \frac{\left\{\mathrm{tr}\left[(\rho\,\sigma)^p\right]\right\}^{1/p}}{\max\left\{\left[\mathrm{tr}(\rho^{2p})\right]^{1/p}, \left[\mathrm{tr}(\sigma^{2p})\right]^{1/p}\right\}}. \tag{2.17}$$

We finally note that a very similar type of circumstance occurs for entropy, which is also sometimes used to calculate distances between two density matrices. The traditional von Neumann entropy measure involves a logarithm, and is often difficult to compute or measure. This can be generalized to the Rényi entropy [41], which is:

$$S_p(\rho) = \frac{p}{1-p}\ln\|\rho\|_p. \tag{2.18}$$

The Rényi entropy reduces to the usual von Neumann entropy, $S(\rho)$ in the limit of $p \to 1$, just as the generalized fidelity defined in (2.14) reduces to the Uhlmann–Jozsa fidelity, in the same limit. Both generalizations have advantages in simplifying computations [42]. Experimentally, it is worth noting that a measurement of $S_2(\rho)$ without resorting to collective measurements has recently been achieved in trapped ions [43] (see also [44] and its references.)

2.6. Hilbert–Schmidt fidelities

Although the measure $\mathcal{F}_{\mathrm{GM}}$ does not comply with all of Jozsa's axioms, its functional form suggests alternatives that are also worth investigating, using the Hilbert–Schmidt norm. An example is the norm-based fidelity $\mathcal{F}_2(\rho,\sigma)$ in the previous subsection, which is a special case of $\mathcal{F}_p(\rho,\sigma)$.

To this end, note that for an arbitrary symmetric, non-vanishing function f that takes the purity of ρ and σ as arguments, one can introduce the functional of $f[\mathrm{tr}(\rho^2), \mathrm{tr}(\sigma^2)]$

$$\mathcal{F}_f(\rho,\sigma) = \frac{\mathrm{tr}(\rho\,\sigma)}{f[\mathrm{tr}(\rho^2), \mathrm{tr}(\sigma^2)]}, \tag{2.19}$$

which are easily seen to satisfy a number of Jozsa's axioms. Specifically, the symmetric property of f and the cyclic property of trace guarantees that the axiom (J2) is satisfied by $\mathcal{F}_f$, whereas the non-vanishing nature of f guarantees that (J1c) is fulfilled. In addition, the fact that f only takes the purity of ρ and σ as arguments ensures that $\mathcal{F}_f$ complies with (J4).

Table 2. Compatibility of $\mathcal{F}_2$ and the candidate fidelity measures defined in (2.20) against Jozsa's axioms.

	J1a	J1b	J1c	J2	J3	J4
$\mathcal{F}_2$	√	√	√	√	√	√
$\mathcal{F}_{\mathrm{AM}}$	√	√	√	√	×	√
$\mathcal{F}_{\mathrm{HM}}$	×	×	√	√	×	√
$\mathcal{F}_{\mathrm{min}}$	×	×	√	√	×	√

The advantage of measures like this is that they only involve the use of operator expectation values, in the sense that $\mathrm{tr}(\rho\,\sigma)$ is the expectation of σ given the state ρ, or vice-versa. Such measures tend to be readily expressed and accessible using standard quantum mechanical techniques applicable to infinite-dimensional Hilbert spaces. By contrast, measures involving nested square roots of operators as found with $\mathcal{F}_1(\rho,\sigma)$ are not so readily calculated using standard quantum techniques in large Hilbert spaces, which is important when one is treating bosonic cases. Similar issues arise in the large Hilbert spaces that occur in many-body theory [42].

Two classes of functions naturally fit into the above requirements, namely, means (arithmetic, geometric or harmonic) and extrema (minimum or maximum) of the purities of ρ and σ. Other than the geometric mean and the maximum—which give, respectively $\mathcal{F}_{\mathrm{GM}}$ and $\mathcal{F}_2$—the other functions give, explicitly,

$$\begin{aligned}
\mathcal{F}_{\mathrm{AM}}(\rho,\sigma) &:= \frac{2\mathrm{tr}(\rho\,\sigma)}{\mathrm{tr}(\rho^2)+\mathrm{tr}(\sigma^2)},\\
\mathcal{F}_{\mathrm{HM}}(\rho,\sigma) &:= \frac{\mathrm{tr}(\rho\,\sigma)\left[\mathrm{tr}(\rho^2)+\mathrm{tr}(\sigma^2)\right]}{2\mathrm{tr}(\rho^2)\mathrm{tr}(\sigma^2)},\\
\mathcal{F}_{\mathrm{min}}(\rho,\sigma) &:= \frac{\mathrm{tr}(\rho\,\sigma)}{\min\left[\mathrm{tr}(\rho^2), \mathrm{tr}(\sigma^2)\right]},
\end{aligned} \tag{2.20}$$

where AM and HM stand for arithmetic and harmonic mean, respectively. The compatibility of these new measures against Jozsa's axioms is summarized in table 2.

The (in)consistency of equation (2.20) with axiom (J1c), (J2), (J3) and (J4) can be verified easily either by inspection or by the construction of counter-examples. Likewise, the normalization of $\mathcal{F}_2$ follows easily from Cauchy–Schwarz inequality whereas the *incompatibility* of $\mathcal{F}_{\mathrm{min}}$ and $\mathcal{F}_{\mathrm{HM}}$ with (J1a) can be verified easily, for example, by considering the following pair of 3×3 density matrices:

$$\rho = \Pi_0, \quad \sigma = \frac{3}{4}\Pi_0 + \frac{1}{8}(\Pi_1 + \Pi_2), \tag{2.21}$$

where, for convenience, we denote

$$\Pi_i := |i\rangle\langle i|, \tag{2.22}$$

as the rank-1 projector corresponding to the ith computational basis state, labelled $|0\rangle, |1\rangle, \ldots$. Explicitly, one finds that

$$\mathcal{F}_{\mathrm{HM}}(\rho,\sigma) = \frac{153}{152}, \tag{2.23}$$

and

$$\mathcal{F}_{\min}(\rho,\sigma)=\frac{24}{19}, \tag{2.24}$$

both greater than 1, thus being incompatible with Jozsa axiom (J1a). As for the normalization of $\mathcal{F}_{\rm AM}$ (and $\mathcal{F}_{\rm GM}$), its proofs can be found in appendix A.

In what follows, we will investigate the compatibility of the fidelity measures listed in tables 1 and 2 against other desirable properties that have been considered. We will, however, dismiss $\mathcal{F}_{\rm HM}$ and $\mathcal{F}_{\min}$ from our discussion as they do not even meet the basic requirement of normalization. Apart from these we will mainly focus on $\mathcal{F}_2$. This meets all the required fidelity axioms, and will be termed the Hilbert–Schmidt fidelity.

The great advantage of $\mathcal{F}_2$ in computational terms is that as well as complying with the extended version of Jozsa's axioms, it is also relatively straightforward to compute and to measure. It only involves expectation values of Hermitian operators, which are computable and measurable with a variety of standard techniques in quantum mechanics (see, e.g., [43, 45]).

3. Auxiliary fidelity properties

Let us now look into other auxiliary properties that have been discussed in the literature. We shall focus predominantly in the three measures that satisfy all the Jozsa axioms, namely, $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_Q$. However, for completeness, we also provide a summary of our understanding of the various properties of those candidate fidelity measures that satisfy (J1a), (J1b), (J2) and (J4).

3.1. Concavity properties

A fidelity measure $\mathcal{F}(\rho,\sigma)$ is said to be separately concave if it is a concave function of any of its argument. More precisely, $\mathcal{F}(\rho,\sigma)$ is concave in its first argument if for arbitrary density matrices σ, ρ_i and arbitrary $p_i\geqslant 0$ such that $\sum_i p_i=1$,

$$\mathcal{F}\left(\sum_i p_i\rho_i,\sigma\right)\geqslant\sum_i p_i\mathcal{F}(\rho_i,\sigma). \tag{3.1}$$

By the symmetry of $\mathcal{F}$ (Jozsa's axiom (J2)), i.e. $\mathcal{F}(\rho,\sigma)=\mathcal{F}(\sigma,\rho)$, a fidelity measure that is concave in its first argument is also concave in its second argument.

A stronger concavity property is also commonly discussed in the literature. Specifically, $\mathcal{F}(\rho,\sigma)$ is said to be jointly concave in both of its arguments if:

$$\mathcal{F}\left(\sum_i p_i\rho_i,\sum_j p_j\sigma_j\right)\geqslant\sum_i p_i\mathcal{F}(\rho_i,\sigma_i). \tag{3.2}$$

This is a stronger concavity property in the sense that if equation (3.2) holds, so must equation (3.1). This can be seen by setting $\sigma_j=\sigma$ for all j in (3.2). Conversely, if $\mathcal{F}(\rho,\sigma)$ is not separately concave, it also cannot be jointly concave. Essentially, concavity property of a fidelity measure tells us how the average state-by-state fidelity compares with the fidelity between the two resulting ensembles of density matrices, a point which we will come back to in section 5.2. A summary of the concavity properties of the various fidelity measures considered can be found in table 3.

Table 3. Summary of the concavity properties of various fidelity measures. The first column gives a list of the various measures, while the second and third column give the compatibility of each measure against the concavity property. An asterisk [a] means that the square root of the measure satisfies the property. Throughout, we use the symbol [b] to indicate that a particular property was—to our knowledge—not discussed in the literature previously. A question mark ? indicates that no counterexample has been found.

	Separate	Joint
$\mathcal{F}_1$	$\surd$ [24]	$\times$ [a]
$\mathcal{F}_2$	$\times$ [b]	$\times$
$\mathcal{F}_{\rm Q}$	$\surd$	$\surd$ [39]
$\mathcal{F}_{\rm N}$	$\surd$	$\surd$ [21]
$\mathcal{F}_{\rm C}$	?	?
$\mathcal{F}_{\rm GM}$	$\times$ [38]	$\times$
$\mathcal{F}_{\rm AM}$	$\times$ [b]	$\times$
$\mathcal{F}_A$	? [a] [25]	$\times$ [ba] [25]

To see that equation (3.1) does not hold for $\mathcal{F}_2$ (as well as $\mathcal{F}_{\rm GM}$ and $\mathcal{F}_{\rm AM}$), it suffices to set $p_1=p_2=\frac{1}{2}$ and consider the qubit or 2×2 density matrices $\rho_1=\frac{1}{10}(\Pi_0+9\Pi_1)$, $\rho_2=\frac{1}{5}(\Pi_0+4\Pi_1)$ and $\sigma=\frac{1}{5}(3\Pi_0+2\Pi_1)$.

On the other hand, to see that $\mathcal{F}_1$ and $\mathcal{F}_{\rm A}$ are not jointly concave, it suffices to consider $p_1=\frac{49}{100}$, $p_2=\frac{1}{2}$, $p_3=\frac{1}{100}$ together with the qutrit or 3×3 density matrices $\rho_1=\Pi_2$, $\sigma_1=\Pi_0$, $\rho_2=\sigma_2=\Pi_1$, $\rho_3=\frac{1}{5}(3\Pi_1+2\Pi_2)$, $\sigma_3=\frac{1}{5}(2\Pi_0+3\Pi_1)$ in (3.2).

3.2. Multiplicativity under tensor products

A fidelity measure $\mathcal{F}(\rho,\sigma)$ is said to be multiplicative if for all density matrices ρ_i, σ_i and for all integer $n\geqslant 2$,

$$\mathcal{F}\left(\bigotimes_{i=1}^n\rho_i,\bigotimes_{j=1}^n\sigma_j\right)=\prod_{i=1}^n\mathcal{F}(\rho_i,\sigma_i); \tag{3.3}$$

likewise $\mathcal{F}(\rho,\sigma)$ is said to *super-multiplicative* if

$$\mathcal{F}\left(\bigotimes_{i=1}^n\rho_i,\bigotimes_{j=1}^n\sigma_j\right)\geqslant\prod_{i=1}^n\mathcal{F}(\rho_i,\sigma_i). \tag{3.4}$$

Clearly, if $\mathcal{F}(\rho,\sigma)$ is (super)multiplicative for $n=2$, it is also (super)multiplicative in the general scenario.

Two special instances of multiplicativity under tensor products are worth mentioning. The first of which concerns the comparison of two quantum states when one has access to n copies of each state. In this case, a (super)multiplicative measure $\mathcal{F}(\rho,\sigma)$ is also (super)multiplicative in its tensor powers, i.e.

$$\mathcal{F}(\rho^{\otimes n},\sigma^{\otimes n})=[\mathcal{F}(\rho,\sigma)]^n. \tag{3.5}$$

The other special instance concerns the scenario when ρ and σ are each appended with an uncorrelated state τ. In this case, multiplicativity demands

Table 4. Summary of the multiplicativity of the various fidelity measures. The first column gives the list of candidate measures $\mathcal{F}$. From the second to the fourth column, we have, respectively, the multiplicativity of the various measures $\mathcal{F}$ under the addition of an uncorrelated ancillary state, under tensor powers and under the general situation of (3.3).

	Ancilla	Tensor powers	General
$\mathcal{F}_1$	$\surd$ [9]	$\surd$ [9]	$\surd$ [9]
$\mathcal{F}_2$	$\surd^{b}$	$\surd^{b}$	Super^{b}
$\mathcal{F}_Q$	$\surd$ [26]	$\surd^{b}$	Super^{b}
$\mathcal{F}_N$	Super	Super	Super [21]
$\mathcal{F}_C$	Super	Super	Super^{b}
$\mathcal{F}_{GM}$	$\surd$ [38]	$\surd$ [38]	$\surd$ [38]
$\mathcal{F}_{AM}$	$\surd^{b}$	$\times^{b}$	$\times^{b}$
$\mathcal{F}_A$	$\surd$ [25]	$\surd$ [25]	$\surd$ [25]

Table 5. Summary of the behavior of fidelity measures under quantum operations. The first column gives the list of candidate measures $\mathcal{F}$. From the second to the fourth column, we have, respectively, the non-decreasing monotonicity of the measures $\mathcal{F}$ under partial trace operation, under projective measurements (see text for details) and under general quantum operations. That is, we mark an entry with a tick $\surd$ if equation (3.7) holds for the corresponding CPTP map.

	Partial trace	Projection	General
$\mathcal{F}_1$	$\surd$	$\surd$	$\surd$ [12]
$\mathcal{F}_2$	$\times^{b}$	$\times^{b}$	$\times$
$\mathcal{F}_Q$	$\surd$	$\surd$	$\surd$ [26]
$\mathcal{F}_N$	$\times$ [21]	?	$\times$
$\mathcal{F}_C$	$\times^{b}$	?	$\times$
$\mathcal{F}_{GM}$	$\times^{b}$	$\times^{b}$	$\times$
$\mathcal{F}_{AM}$	$\times^{b}$	$\times^{b}$	$\times$
$\mathcal{F}_A$	$\surd$ [25]	$\surd$ [25]	$\surd$ [37]

$$\mathcal{F}(\rho\otimes\tau,\sigma\otimes\tau)=\mathcal{F}(\rho,\sigma)\mathcal{F}(\tau,\tau)=\mathcal{F}(\rho,\sigma). \tag{3.6}$$

Intuitively, if $\mathcal{F}(\rho,\sigma)$ is also a measure of the overlap between ρ and σ, one would expect that multiplicativity is satisfied, at least, in these two special instances. In table 4, we summarize the multiplicativity properties of the various fidelity measures. For a proof of the (super)multiplicativity of $\mathcal{F}_2$, $\mathcal{F}_Q$ (and $\mathcal{F}_C$), and a counterexample showing that $\mathcal{F}_{AM}$ is in general not multiplicative nor supermultiplicative, we refer the reader to A.5.

3.3. Monotonicity under quantum operations

The physical operation of appending a given quantum state ρ by a fixed quantum state τ discussed above is an example of what is known as a completely positive trace preserving (CPTP) map. If we denote a general CPTP map by $\mathcal{E}:\rho\to\mathcal{E}(\rho)$, it is often of interest to determine, for all density matrices ρ and σ, if the inequality:

$$\mathcal{F}(\mathcal{E}(\rho),\mathcal{E}(\sigma))\geqslant\mathcal{F}(\rho,\sigma) \tag{3.7}$$

is satisfied for a given candidate fidelity measure $\mathcal{F}$. In particular, a fidelity measure that satisfies this inequality is said to be non-contractive (or equivalently, monotonically non-decreasing) under quantum operations.

In this regard, it is worth noting that a measure $\mathcal{F}$ that (1) complies with the requirement of unitary invariance (J4), (2) is invariant under the addition of an uncorrelated ancillary state and is either (3a) non-contractive under partial trace operation or is (3b) jointly concave is also non-contractive under general quantum operations. The sufficiency of conditions (1), (2) and (3a) follow directly from the Stinespring representation (see, e.g. [46]) of CPTP maps while that of (1), (2) and (3b) also make use of a specific representation of the partial trace operation as a convex mixture of unitary transformations, see, e.g. equation (33) of [47]. For example, since $\mathcal{F}_A$ is monotonic [25] under partial trace operation (likewise for $\mathcal{F}_1$ and $\mathcal{F}_Q$), the above sufficiency condition allows us to conclude that $\mathcal{F}_A$ is also monotonic under general quantum operations.

Apart from the partial trace operation and the extension of a quantum state by a fixed ancillary state, the measurement of a quantum state in some fixed basis followed by forgetting the measurement outcome is another class of CPTP maps that one frequently encounters in the context of quantum information. In particular, if each measurement operator (the Kraus operator) is a rank-1 projector, the post-measurement state would be the corresponding eigenstate. In this case, an evaluation of the fidelity between the different outputs of the CPTP map corresponds to an evaluation of the fidelity between the corresponding classical probability distributions.

The monotonicity of the various candidate fidelity measures for the few different CPTP maps discussed above is summarized in table 5. That $\mathcal{F}_2$ may be contractive under partial trace operation can be seen by considering the two-qubit density matrices $\rho=\Pi_1\otimes\rho_B$, $\sigma=\frac{1}{2}\mathbf{1}_2\otimes\sigma_B$ where

$$\rho_B=\begin{bmatrix}0.3 & 0.3\\ 0.3 & 0.7\end{bmatrix},\quad \sigma_B=\begin{bmatrix}0.06 & 0.2\\ 0.2 & 0.94\end{bmatrix}$$

and the partial trace of ρ and σ over subsystem B. As for the monotonicity of $\mathcal{F}_2$, $\mathcal{F}_{GM}$ and $\mathcal{F}_{AM}$ under projective measurements, one may verify that these measures may be indeed contractive by considering the qubit density matrices

$$\begin{aligned}\rho&=\begin{bmatrix}0.35 & -0.25-0.2\,\mathrm{i}\\ -0.25+0.2\,\mathrm{i} & 0.65\end{bmatrix},\\ \sigma&=\begin{bmatrix}0.82 & -0.2-0.24\,\mathrm{i}\\ -0.2+0.24\,\mathrm{i} & 0.18\end{bmatrix},\end{aligned} \tag{3.8}$$

and a rank-1 projective measurement in the computational basis. For the monotonicity of $\mathcal{F}_C$, $\mathcal{F}_{GM}$ and $\mathcal{F}_{AM}$ under the partial trace operation, we refer the reader to A.5 for counterexamples.

The monotonicity under partial trace obtained in $\mathcal{F}_1$ and $\mathcal{F}_Q$ means that a mixed state fidelity measured according to $\mathcal{F}_1$ and $\mathcal{F}_Q$ is greater than or equal to the corresponding fidelity when evaluated on a larger Hilbert space. This implies that there is a bias in using these measured fidelities on a smaller Hilbert space as an estimator for the fidelity on an enlarged relevant Hilbert space, when using $\mathcal{F}_1$ and $\mathcal{F}_Q$.

This potentially undesirable property is not shared by $\mathcal{F}_2$. However, the general question of which fidelity is the best

Table 6. Metric properties for some functionals of the fidelity measures $\mathcal{F} = \mathcal{F}(\rho,\sigma)$, as discussed in section 2.

	$\arccos[\sqrt{\mathcal{F}}]$	$\sqrt{1-\sqrt{\mathcal{F}}}$	$\sqrt{1-\mathcal{F}}$
$\mathcal{F}_{\mathrm{I}}$	$\surd$ [23]	$\surd$ [23]	$\surd$ [23]
$\mathcal{F}_2$	$\times^{\mathrm{b}}$	$\times^{\mathrm{b}}$	$\surd^{\mathrm{b}}$
$\mathcal{F}_{\mathrm{Q}}$	$\times^{\mathrm{b}}$	$\times^{\mathrm{b}}$	$\times^{\mathrm{b}}$
$\mathcal{F}_{\mathrm{N}}$	$\times$ [21]	$\times$ [21]	$\surd$ [21]
$\mathcal{F}_{\mathrm{C}}$	?	?	$\surd^{\mathrm{b}}$
$\mathcal{F}_{\mathrm{GM}}$	$\times^{\mathrm{b}}$	$\times^{\mathrm{b}}$	$\surd^{\mathrm{b}}$
$\mathcal{F}_{\mathrm{AM}}$	$\times^{\mathrm{b}}$	$\times^{\mathrm{b}}$	?
$\mathcal{F}_{A}$	?	$\surd$ [37]	$\surd^{\mathrm{b}}$

unbiased estimator under partial trace operations from a randomized extension of the measured Hilbert space appears to be an open problem.

3.4. Metrics

Intuitively, one expects that if $\mathcal{F}(\rho,\sigma)$ is a measure of the degree of similarity or overlap between ρ and σ, a proper distance measure, i.e. a metric can be constructed via some functionals of $\mathcal{F}(\rho,\sigma)$ which vanishes for $\rho=\sigma$.

In this section, we review what is known about the metric properties of three functionals of $\mathcal{F}(\rho,\sigma)$, namely, $\arccos[\sqrt{\mathcal{F}(\rho,\sigma)}]$, $\sqrt{1-\sqrt{\mathcal{F}(\rho,\sigma)}}$ and $\sqrt{1-\mathcal{F}(\rho,\sigma)}$ for the various fidelity measures discussed in the previous section. Following the literature, one may want to refer to these functionals, respectively, as the *modified* Bures angle [12], the *modified* Bures distance [48] and the *modified* sine distance [49]. The results are summarized in table 6 while a proof of the respective metric properties can be found in appendix B.

Somewhat surprisingly, despite the fact that $\mathcal{F}_{\mathrm{Q}}$ shares many nice properties with $\mathcal{F}_{\mathrm{I}}$, none of these functionals derived from $\mathcal{F}_{\mathrm{Q}}$ actually behave like a metric for the space of density matrices. This can be verified, for example, by noticing a violation of the triangle inequality for all these functionals of $\mathcal{F}_{\mathrm{Q}}$ under the choice

$$\begin{aligned}\rho &= \frac{1}{10}(3\Pi_0 + 7\Pi_1),\\ \sigma &= \frac{1}{100}(\Pi_0 + 99\Pi_1),\\ \tau &= \frac{1}{5}(\Pi_0 + 4\Pi_1).\end{aligned} \tag{3.9}$$

As for a violation of the triangle inequality by the other functionals presented in the table, it is sufficient to consider the following qutrit density matrices:

$$\begin{aligned}\rho &= \frac{1}{5}(\Pi_1 + 4\Pi_2),\\ \sigma &= \Pi_0,\\ \tau &= \frac{1}{5}\Pi_0 + \frac{1}{20}\Pi_1 + \frac{3}{4}\Pi_2.\end{aligned} \tag{3.10}$$

Notice that other than the functionals considered above, it was also shown in [36] that $\max_\tau |\mathcal{F}_{\mathrm{N}}(\rho,\tau) - \mathcal{F}_{\mathrm{N}}(\sigma,\tau)|$, a functional constructed from the super-fidelity $\mathcal{F}_{\mathrm{N}}$, is also a metric. In fact, with very similar arguments, the same authors showed in [50] that the same functional with $\mathcal{F}_{\mathrm{I}}$ instead of $\mathcal{F}_{\mathrm{N}}$ is also a metric.

4. Comparisons, bounds, and relations between measures

To understand the differences between the fidelity measures, one must compare them quantitatively. In some cases there are rigorous bounds that relate the fidelities, while in other cases comparisons are made graphically. As a rough guide, the average behavior of full-rank, random density matrices indicates that when making comparisons with $p > 1$, our numerics for small p and d suggest that:

$$\mathcal{F}_p(\rho,\sigma) \lesssim \mathcal{F}_{\mathrm{I}}(\rho,\sigma) \lesssim \mathcal{F}_{\mathrm{Q}}(\rho,\sigma) \tag{4.1}$$

often holds. However, this is not a hard and fast rule. In qubit cases these are rather strict bounds (at least for $p=2$). In larger Hilbert spaces, however, these inequalities are only approximately true, with exceptions that strongly depend on the rank of the density matrices being compared.

These different cases are explained below.

4.1. Bounds

First, we provide a summary of inequalities relating some of these fidelity measures, or some bounds on them. Although there are also known inequalities (see, e.g., [51, 52]) bounding the Uhlmann-Jozsa fidelity between two density matrices ρ,σ and the trace distance between them, we focus here on the bounds relating the various fidelity measures. To begin with, it was established in [53] (see also [37] and [25]) that

$$\mathcal{F}_{\mathrm{I}}(\rho,\sigma) \leqslant \sqrt{\mathcal{F}_{\mathrm{A}}(\rho,\sigma)} \leqslant \sqrt{\mathcal{F}_{\mathrm{I}}(\rho,\sigma)}. \tag{4.2}$$

Later, in [26] (see also [39]), these inequalities were rediscovered and extended to

$$\mathcal{F}_{\mathrm{I}}(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{Q}}(\rho,\sigma) \leqslant \sqrt{\mathcal{F}_{\mathrm{A}}(\rho,\sigma)} \leqslant \sqrt{\mathcal{F}_{\mathrm{I}}(\rho,\sigma)}, \tag{4.3}$$

where the second of these inequalities follows directly from the definition of $\mathcal{F}_{\mathrm{A}}$ and $\mathcal{F}_{\mathrm{Q}}$ given, respectively, in equations (2.10) and (2.12). At about the same time, $\mathcal{F}_{\mathrm{N}}$ was also shown [35] to be an upper bound on $\mathcal{F}_{\mathrm{I}}$, i.e.

$$\mathcal{F}_{\mathrm{I}}(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{N}}(\rho,\sigma). \tag{4.4}$$

Here, we add to this list by showing that $\mathcal{F}_2(\rho,\sigma)$ actually provides a lower bound to $\mathcal{F}_{\mathrm{N}}(\rho,\sigma)$.

Theorem 1. *For arbitrary Hermitian matrices ρ and σ such that* $\mathrm{tr}(\rho^2) \leqslant 1$ *and* $\mathrm{tr}(\sigma^2) \leqslant 1$

$$\mathcal{F}_2(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{N}}(\rho,\sigma). \tag{4.5}$$

Proof. First, let us rewrite an arbitrary pair of d-dimensional density matrices ρ,σ using an orthonormal basis of Hermitian matrices $\vec{\Upsilon} = (\Upsilon_0, \Upsilon_1, \ldots, \Upsilon_{d^2-1})$,

$$\rho = \vec{u}\cdot\vec{\Upsilon} \quad \text{and} \quad \sigma = \vec{v}\cdot\vec{\Upsilon}, \tag{4.6}$$

where $\vec{u}, \vec{v} \in \mathbb{R}^{d^2}$ are the expansion coefficients of ρ and σ in the basis $\vec{\Upsilon}$. We may now rewrite $\mathcal{F}_2$ and $\mathcal{F}_{\mathrm{N}}$ as

$$\mathcal{F}_2(\rho,\sigma) = \frac{\vec{u}\cdot\vec{v}}{\max(u^2,v^2)},$$
$$\mathcal{F}_{\mathrm{N}}(\rho,\sigma) = \vec{u}\cdot\vec{v} + \sqrt{1-u^2}\sqrt{1-v^2}, \tag{4.7}$$

where $u = \|\vec{u}\|_2$ and $v = \|\vec{v}\|_2$.

It is straightforward to check that inequality (4.5) holds true if and only if the following inequality is true:

$$\frac{\vec{u}\cdot\vec{v}}{\max(u^2,v^2)} \leqslant \sqrt{\frac{1-\min(u^2,v^2)}{1-\max(u^2,v^2)}}, \tag{4.8}$$

which obviously holds if $u = v$. For definiteness, let $u > v$ and we have

$$\frac{\vec{u}\cdot\vec{v}}{\sqrt{1-v^2}} \leqslant \frac{u^2}{\sqrt{1-u^2}}. \tag{4.9}$$

This is easily seen to hold since, for $u > v$, the numerator of the l.h.s. is dominated by the numerator of the r.h.s., whereas the denominator of the l.h.s. dominates the denominator of the r.h.s.. The case where $v > u$ is completely analogous. □

Corollary 1. *For arbitrary qubit density matrices ρ and σ,*

$$\mathcal{F}_2(\rho,\sigma) \leqslant \mathcal{F}_1(\rho,\sigma). \tag{4.10}$$

Proof. This follows trivially from theorem 1 and the fact that for qubit density matrices $\mathcal{F}_1(\rho,\sigma) = \mathcal{F}_{\mathrm{N}}(\rho,\sigma)$, see [21]. □

From the inequality of arithmetic and geometric means, as well as the definitions given in equations (2.11), (2.20) and (2.16), it is easy to see that the following bounds hold:

$$\mathcal{F}_2(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{AM}}(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{GM}}(\rho,\sigma). \tag{4.11}$$

Besides, some straightforward calculation starting from the definitions given in equations (2.8) and (2.9) leads to

$$\mathcal{F}_{\mathrm{N}}(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{C}}(\rho,\sigma). \tag{4.12}$$

Some remarks are now in order. Given that $\mathcal{F}_{\mathrm{N}}(\rho,\sigma) = \mathcal{F}_1(\rho,\sigma)$ for qubit states and the fact that $\mathcal{F}_{\mathrm{Q}}(\rho,\sigma)$ and $\mathcal{F}_{\mathrm{N}}(\rho,\sigma)$ both provide an upper bound on $\mathcal{F}_1(\rho,\sigma)$, one may wonder:

(1) Could $\mathcal{F}_{\mathrm{N}}(\rho,\sigma)$ provide a lower bound on $\mathcal{F}_{\mathrm{Q}}(\rho,\sigma)$, or the other way around?
(2) Since $\mathcal{F}_{\mathrm{A}}(\rho,\sigma)$ and the sub-fidelity [35] defined as: $\mathrm{tr}(\rho,\sigma) + \sqrt{2\{[\mathrm{tr}(\rho\,\sigma)]^2 - \mathrm{tr}(\rho\,\sigma\,\rho\,\sigma)\}}$ both provide a lower bound on $\mathcal{F}_1(\rho,\sigma)$, could it be that one of these quantities also lower bounds the other?
(3) Could it be that equation (4.10) also holds for higher-dimensional density matrices?

Here, let us note that counterexamples to *all* of the above conjectures can be easily found by considering pairs of qutrit density matrices. However, we leave open the possibility of bounding these quantities using nonlinear (including polynomial) functionals of the other fidelity measures.

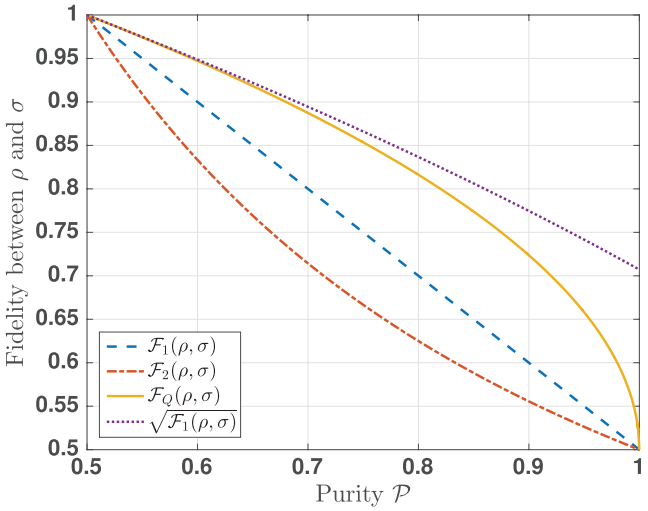

Figure 1. The fidelity between the single-qubit (mixed) states given in equation (4.13) as a function of purity for the Uhlmann–Jozsa fidelity $\mathcal{F}_1$, equation (2.7), the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16) and the non-logarithmic variety of the quantum Chernoff bound $\mathcal{F}_{\mathrm{Q}}$, equation (2.12). For comparison, see equation (4.3), we have also included in the plot the square root of $\mathcal{F}_1$ (sometimes referred to as *the fidelity*, see [12]) as a function of the purity $\mathcal{P}$.

4.2. Comparisons: interpolated qubit states

How large are the differences between these measures? In order to understand this, we first turn to a simple but concrete example. To this end, consider two families of qubit density matrices that interpolate between maximally mixed states—that are necessarily identical—and pure states that are distinct, with an interpolation parameter of $r \in [0,1]$:

$$\rho(r) = \frac{1}{2}\left(\mathbf{1}_2 + r\sigma_x\right), \quad \sigma(r) = \frac{1}{2}\left(\mathbf{1}_2 + r\sigma_z\right). \tag{4.13}$$

The purity of a density operator ρ is given by $\mathcal{P} = \mathrm{tr}\left(\rho^2\right)$. For the two quantum states in this comparison, their purities are set to be equal and are parametrized by r, i.e.

$$\mathcal{P} = \frac{1}{2}\left(1 + r^2\right), \tag{4.14}$$

where $r = 0$ corresponds to a maximally mixed state while $r = 1$ corresponds to a pure state. Note that although these density matrices might look at first as though they are a sum of two mixed states, they are not. In the limit of $r = 1$, each becomes a distinct pure state, with a non-vanishing inner product.

In figure 1, we show the result of the fidelities $\mathcal{F}_1(\rho,\sigma)$ (see equation (2.7)), $\mathcal{F}_2(\rho,\sigma)$ (see equation (2.16)) and $\mathcal{F}_{\mathrm{Q}}(\rho,\sigma)$ (see equation (2.12)) as a function of the purity $\mathcal{P}$, equation (4.14), of the states given in equation (4.13). It is worth noting that for these states, the inequalities of equations (4.4), (4.11) and (4.12) are all saturated, thereby giving $\mathcal{F}_1(\rho,\sigma) = \mathcal{F}_{\mathrm{N}}(\rho,\sigma) = \mathcal{F}_{\mathrm{C}}(\rho,\sigma)$ and

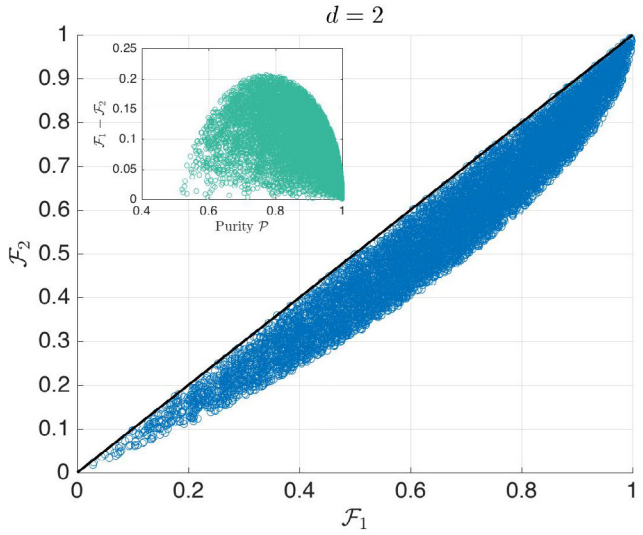

Figure 2. Scatter plot showing the Uhlmann–Jozsa fidelity $\mathcal{F}_1$, equation (2.7), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of single-qubit (mixed) density matrices. The black line satisfying $\mathcal{F}_1 = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_1 - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

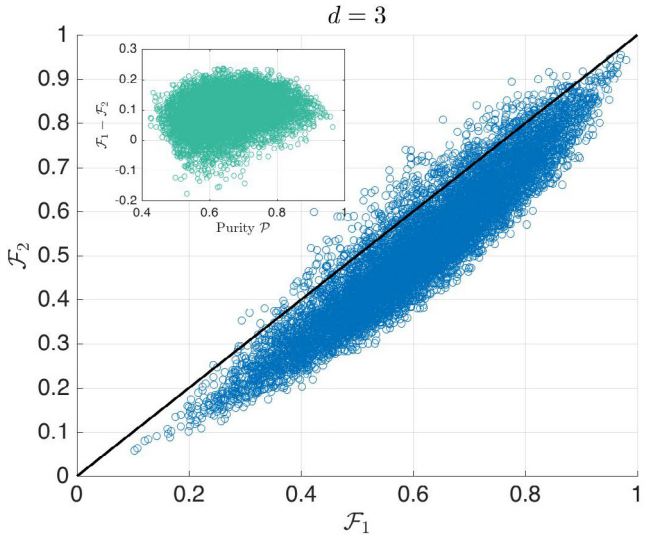

Figure 3. Scatter plot showing the Uhlmann–Jozsa fidelity $\mathcal{F}_1$, equation (2.7), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of single-qutrit (mixed) density matrices. The black line satisfying $\mathcal{F}_1 = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_1 - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

$\mathcal{F}_2(\rho,\sigma) = \mathcal{F}_{\rm AM}(\rho,\sigma) = \mathcal{F}_{\rm GM}(\rho,\sigma)$. Moreover, for these states, it can also be verified that $\sqrt{\mathcal{F}_{\rm A}(\rho,\sigma)} = \mathcal{F}_{\rm Q}(\rho,\sigma)$, thereby saturating the second inequality in equation (4.3).

4.3. Comparisons: random density matrices

As a second type of comparison, we generate two random density matrices and compare them, in a Hilbert space of arbitrary dimension. In general, these are generated by taking a random Gaussian matrix $\boldsymbol{g}$, whose elements g_{ij} are complex random numbers of unit variance. Starting from a complex Gaussian matrix, called a Ginibre ensemble [54], a random, positive-semidefinite density matrix with unit trace is generated [55] by letting:

$$\rho = \frac{gg^\dagger}{\mathrm{tr}\,(gg^\dagger)}. \tag{4.15}$$

Some investigations of random matrix fidelity have been carried out previously using the $\mathcal{F}_1$ fidelity [56], while here we focus on the comparative behavior of the different fidelity measures.

In figure 2, we compare plots with two qubit random matrices, and give scatter plots of both $\mathcal{F}_1 - \mathcal{F}_2$ against the maximum purity of the pair, as well as $\mathcal{F}_2$ against $\mathcal{F}_1$. We note that for this case, as expected from equation (4.10), $\mathcal{F}_1 \geqslant \mathcal{F}_2$ for $d = 2$. In addition, there is a lower bound on the state purity, since $\mathcal{P} \geqslant 0.5$ for qubit density matrices.

In figure 3, we compare plots with two qutrit random matrices, again with scatter plots of both $\mathcal{F}_1 - \mathcal{F}_2$ against the maximum purity of the pair, together with the two fidelities plotted against each other. For this case, with $d = 3$, both $\mathcal{F}_1 > \mathcal{F}_2$, and $\mathcal{F}_1 < \mathcal{F}_2$ are possible. As a concrete example of the latter case, one only has to consider a comparison of two diagonal qutrit density matrices, where p is a real coefficient such that $0 < p < 1$:

$$\rho = (1-p)\,\Pi_1 + p\Pi_2, \quad \sigma = (1-p)\,\Pi_0 + p\Pi_2. \tag{4.16}$$

This has the property that $\mathcal{F}_1 < \mathcal{F}_2$ for any value of p such that $0 \neq p \neq 1$, since $\mathcal{F}_2$ is divided by the maximum purity, and this is less than unity. However, although extremely simple, this is also atypical. In almost all cases, the two density matrices being compared are not simultaneously diagonal. For the average case, we can see instead that $\langle \mathcal{F}_1 \rangle > \langle \mathcal{F}_2 \rangle$. This corresponds to the intuitive expectation that, since $\mathcal{F}_1$ is a maximum over purifications, it will generally have a bias towards high values.

Finally in the case of two random matrices of larger dimension, we compare two 10×10 qudit matrices, with the other details as previously. For this case, just as with $d = 3$, both $\mathcal{F}_1 > \mathcal{F}_2$, and $\mathcal{F}_1 < \mathcal{F}_2$ are possible, although the fraction of cases with $\mathcal{F}_1 > \mathcal{F}_2$ has increased substantially as shown in figure 4. Again, for the average case, $\langle \mathcal{F}_1 \rangle > \langle \mathcal{F}_2 \rangle$. As expected, there is a lower bound on the state purity, since $\mathcal{P} \geqslant 1/d$ in a $d-$dimensional density matrix. Here one has $d = 10$, so $\mathcal{P} \geqslant 0.1$ in a 10 dimensional density matrix. The average fidelity is greatly reduced in this case, since the larger Hilbert space dimension reduces the probability that two random matrices will be similar.

The corresponding plots comparing $\mathcal{F}_2$ and $\mathcal{F}_{\rm Q}$ can be found, respectively, in figures 5–7.

5. Applications

5.1. Fidelity in quantum physics

Most current applications of quantum fidelity take place in the context of quantum technology. Since fidelity is a

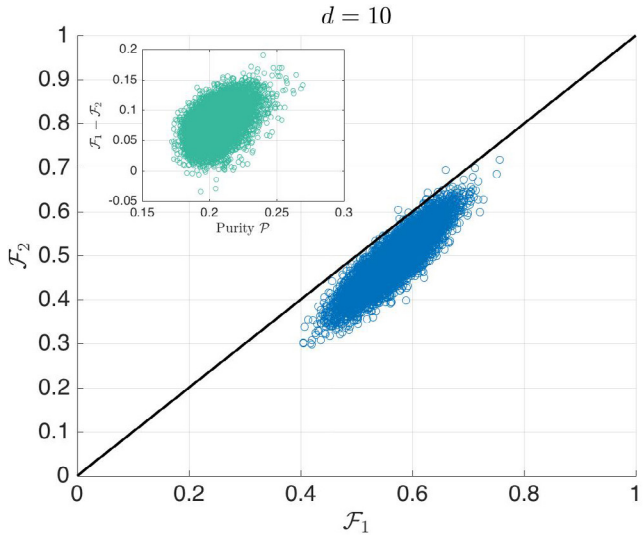

Figure 4. Scatter plot showing the Uhlmann–Jozsa fidelity $\mathcal{F}_1$, equation (2.7), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of 10×10 density matrices. The black line satisfying $\mathcal{F}_1 = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_1 - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

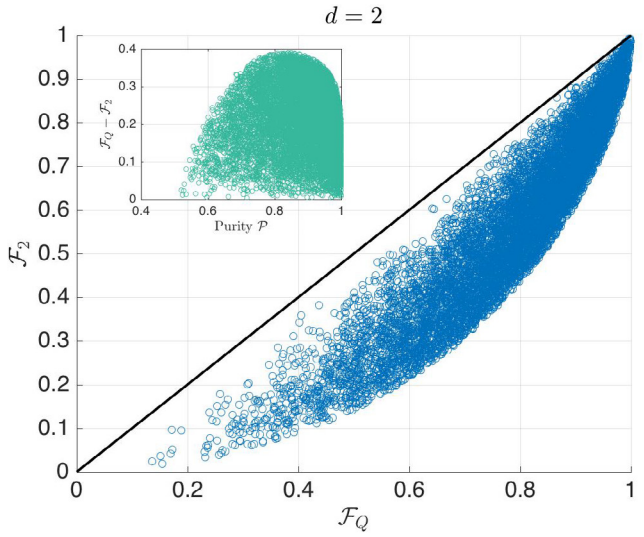

Figure 5. Scatter plot showing the non-logarithmic variety of the quantum Chernoff bound $\mathcal{F}_Q$, equation (2.12), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of single-qubit (mixed) density matrices. The black line satisfying $\mathcal{F}_Q = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_Q - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

relative measure, the most appropriate fidelity depends on the proposed application. These technologies usually have a well-defined purpose. For example, one may use a quantum technology like squeezing [57, 58] or Einstein–Poldosky–Rosen (EPR) correlation [59]/steering [60–62] to measure gravitational waves [63, 64] in a more sensitive way. As well as precision metrology, other applications include enhanced communications [22], cryptography [65], quantum computing [66], many-body quantum simulators [67, 68], quantum data processing and storage [69], and the growing area of quantum thermodynamics [70].

Each of these fields has their own criteria for success. However, the components of the technology ultimately depend on the realization of certain quantum states and their processing. Hence, knowledge of fidelity can help to measure how close one is to achieving the required quantum states that are utilized at a given stage in a quantum process [12], see, e.g. [71–73] and references therein. This concept is also applicable more widely, in fundamental physics problems like quantum phase transitions.

A fidelity measure, as with some other (operational) figures of merit (see, e.g. [74–76]), is sometimes [73] used to prove that a device is quantum in nature, rather than operating at a simply classical level. More generally, one can analyze quantum operations in terms of process fidelity and average fidelity, as explained below.

5.2. Average fidelity in quantum processes with pure input states

Jozsa's axioms do not lead to a unique function of two density matrices that generalizes equation (2.4) to a pair of mixed states. In this section, we would like to examine a desirable feature of a fidelity measure that arises naturally in a quantum communication scenario [22], although this concept also arises in other types of quantum processes. This also allows us to relate a fidelity measure as a measure of the degree of similarity between two quantum states to the notion of fidelity originally introduced by Schumacher [20]. We focus especially on the three candidate fidelities that satisfy the axioms, and have a physical interpretation: $\mathcal{F}_1$, $\mathcal{F}_Q$ and $\mathcal{F}_2$.

We note that the uniform average fidelity over all possible pure input states is often termed $\mathcal{F}_{ave}$. Like the more complex process fidelity, this can be used as a means of analyzing performance of quantum logic gates [77, 78]. Here we allow $\mathcal{F}_{ave}$ to have non-uniform initial probabilities, and we note that, for pure states, the fidelities being averaged are not dependent on the *choice* of fidelity measure, since these are unique (by Jozsa's axiom (J3)) if one of the density matrices being compared is a pure state.

Suppose we attempt to copy, store, perform idealized logic operations or transmit a number of pure states ρ_j, each occurring with probability p_j. Imagine further that the pure states ρ_j, are the desired output of the logic operations (communication protocol). Evidently, this combination of states can also be described by the mixed state $\rho = \sum_j p_j \rho_j$. Now, let us further imagine that during the physical processes of interest, there is a probability of error ϵ, in which case a random error state ρ_0 is generated. For simplicity, we shall assume (unless otherwise stated) that ρ_0 is orthogonal to ρ_j, for all $j \neq 0$.

A physical scenario that matches the above description consists of transmitting photon number states through some quantum channel, but with probability ϵ, the Fock state $\rho_j = |j\rangle\langle j|$ gets transformed to the vacuum state $|0\rangle\langle 0|$, or some other undesirable Fock state that corresponds to an error.

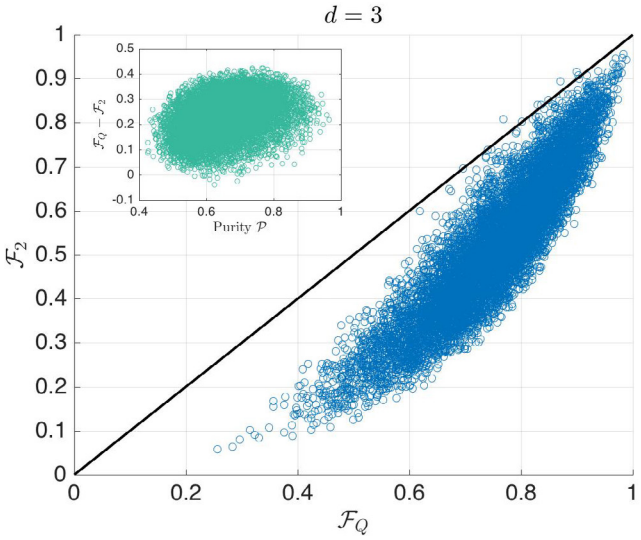

Figure 6. Scatter plot showing the non-logarithmic variety of the quantum Chernoff bound $\mathcal{F}_{\mathrm{Q}}$, equation (2.12), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of single-qutrit (mixed) density matrices. The black line satisfying $\mathcal{F}_{\mathrm{Q}} = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_{\mathrm{Q}} - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

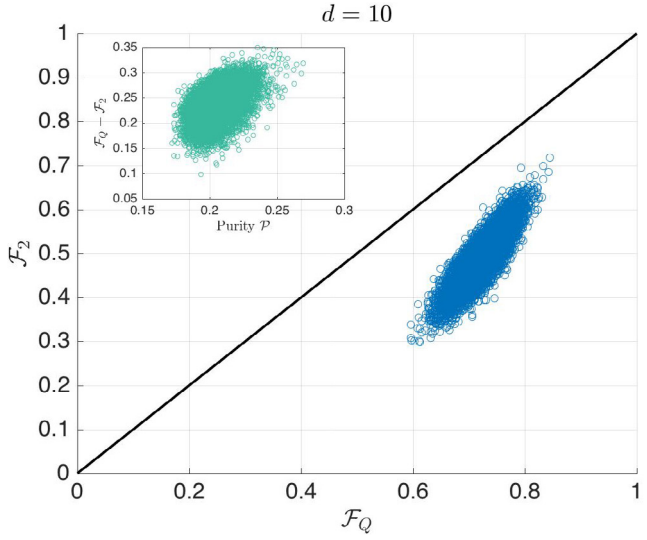

Figure 7. Scatter plot showing the non-logarithmic variety of the quantum Chernoff bound $\mathcal{F}_{\mathrm{Q}}$, equation (2.12), and the Hilbert–Schmidt fidelity $\mathcal{F}_2$, equation (2.16), for 10^4 random pairs of single-qutrit (mixed) density matrices. The black line satisfying $\mathcal{F}_{\mathrm{Q}} = \mathcal{F}_2$ is a guide for the eye. The inset shows a scatter plot of the difference $\mathcal{F}_{\mathrm{Q}} - \mathcal{F}_2$ versus the *maximum* purity $\mathcal{P}$ of these pairs of density matrices.

However, we do not assume here that the inputs are necessarily orthogonal, as there are applications in cryptography [65] where non-orthogonality serves as an important part of the communication protocol.

Let us denote the output state for the transmission of ρ_j as $\sigma_j = \epsilon\rho_0 + (1-\epsilon)\rho_j$. The overall combination of the output states can also be described by the following mixed state:

$$\sigma = \sum_j p_j\sigma_j = \epsilon\rho_0 + (1-\epsilon)\rho. \tag{5.1}$$

The mixed state fidelity can now be examined from two points of view: the *average state-by-state fidelity*, and also the *fidelity as a mixed state*. For a sensible generalization of Schumacher's fidelity to a pair of mixed states, it seems reasonable to require that the average state-by-state fidelity corresponds exactly to the mixed state fidelity, at least, under some circumstances.

To this end, let us first define the average state-by-state fidelity as a function of the probabilities and states:

$$\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\sigma}) = \sum_j p_j \mathcal{F}(\rho_j, \sigma_j). \tag{5.2}$$

Since the inputs are individually pure states, the fidelity is just the state overlap, and one then obtains:

$$\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\sigma}) = \sum_j p_j \mathrm{tr}(\rho_j\sigma_j) = 1 - \epsilon. \tag{5.3}$$

In the case of $\mathcal{F}_2$, the average fidelity matches the mixed fidelity, so that

$$\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\sigma}) = \mathcal{F}(\rho, \sigma) \tag{5.4}$$

whenever the average signal state ρ has equal to or larger purity than that of the average output state σ, i.e. $\mathrm{tr}(\sigma^2) \leqslant \mathrm{tr}(\rho^2)$. This is a common situation, although not universal.

If all (input) states including the error state are orthogonal to each other, and hence diagonal in the same basis, it is easy to verify that the Uhlmann–Jozsa fidelity $\mathcal{F}_{\mathrm{I}}$, the non-logarithmic variety of the quantum Chernoff bound, $\mathcal{F}_{\mathrm{Q}}$, as well as the A-fidelity $\mathcal{F}_{\mathrm{A}}$ comply with the desired requirement of equation (5.4), without any additional conditions. For a proof, see appendix A.

It is also not difficult to check that if ρ is a pure state, $\mathcal{F}_{\mathrm{N}}$ also satisfies the desired requirement; although it is a rare situation that a pure state would be used for communications in this way, owing to the extremely small information content. As for $\mathcal{F}_{\mathrm{C}}$, $\mathcal{F}_{\mathrm{AM}}$ and $\mathcal{F}_{\mathrm{GM}}$, there does not appear to be a generic situation where equation (5.4) will hold true for all error probabilities $0 < \epsilon < 1$.

Let us analyze briefly the case where average output state has a greater purity than the signal state. At first this seems unlikely, but in fact it can occur. If the error state ρ_0 is a pure state, and this occurs with probability one, then the output is always a pure state. This could be a vacuum state, owing to an extremely serious error condition: the channel being completely absorbing. Under these conditions, one has that $\mathcal{F}_{\mathrm{ave}} = 0$, which is expected: every input of the alphabet results in an incorrect output. This is true of any fidelity measure that satisfies our generalized Jozsa axiom list, due to the orthogonality condition, and hence it is true of $\mathcal{F}_{\mathrm{I}}$, $\mathcal{F}_{\mathrm{Q}}$ and $\mathcal{F}_2$.

Finally, we note that the most general error state ρ_0 depends on the input state ρ_j, and is neither pure nor orthogonal to all other inputs. Under these circumstances, $\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\sigma})$ is still

well-defined but it may not correspond to any of the mixed fidelities defined here.

5.3. Average fidelity in quantum processes with mixed input states

A quantum process with pure input states is an ideal scenario that is unlikely to occur in reality. As well as the outputs, even the inputs are most likely to be mixed states, as they are typically multi-mode and will be subject to timing jitter, losses and spontaneous emission noise, to name just a few possible error conditions.

In this case, we may have a scenario like the one given above, except that all the states are mixed. We again assume for simplicity that errors in the output channel are *orthogonal* to all inputs, as before. The average state-by-state fidelity is now a function of the probabilities, states and the fidelity measure:

$$\mathcal{F}_{\text{ave}}\left(\boldsymbol{p},\boldsymbol{\rho},\boldsymbol{\sigma}\right)=\sum_{j}p_j\mathcal{F}(\rho_j,\sigma_j). \tag{5.5}$$

In general, for $\mathcal{F}_{\text{I}}$, $\mathcal{F}_{\text{Q}}$, and $\mathcal{F}_{\text{A}}$, the situation becomes complex, as there is no guarantee that all the states involved can be diagonalized in the same basis. If they do, we will again end up with equation (5.4). This makes an evaluation of the corresponding fidelities generally difficult, since, for example, the products of matrix square roots that occur in $\mathcal{F}_{\text{I}}$ are extremely nontrivial in general.

As for $\mathcal{F}_2$, a reduction holds if the individual input states ρ_j and the average state ρ has a degraded purity owing to errors (i.e. $\text{tr}(\rho_j^2)\geqslant\text{tr}(\sigma_j^2)$ and $\text{tr}(\rho^2)\geqslant\text{tr}(\sigma^2)$), although, as noted above, this may not be true for an extremely lossy channel. With this simplification the fidelity for each possible input is the same:

$$\mathcal{F}_2(\rho_j,\sigma_j)=\frac{\text{tr}(\rho_j\sigma_j)}{\text{tr}(\rho_j^2)}=1-\epsilon. \tag{5.6}$$

In such cases, we have the same result for the average fidelity as for pure state inputs, namely:

$$\mathcal{F}_{\text{ave}}\left(\boldsymbol{p},\boldsymbol{\rho},\boldsymbol{\sigma}\right)=\sum_{j}p_j\mathcal{F}_2(\rho_j\sigma_j)=1-\epsilon=\mathcal{F}_2(\rho,\sigma). \tag{5.7}$$

As with the pure state case, we emphasize that the average fidelity is therefore only the same as the mixed state fidelity in rather special circumstances, even when using the simpler $\mathcal{F}_2$ measure of fidelity. In general, the assumption that the error state is orthogonal to all inputs is not always valid, and as a result they measure different properties of the quantum process. Note also that the requirement of equation (5.4) amounts to demanding that the joint concavity inequality of equation (3.2) is saturated in this communication scenario.

5.4. Teleportation and cloning fidelity

As a measure of the degree of similarity between two quantum states, fidelity occurs naturally in assessing the quality of a quantum communication [20] channel. A well-known example of such a channel is the so-called teleportation channel [10, 79] where the *unknown* quantum state of a physical system is transferred from a sender to a receiver via the help of a shared quantum resource between the two ends. If one could transport the system itself directly to the receiver intact, this task is trivial. However, since channels are usually far from ideal—they may be lossy, for example, especially over long distances—the quantum states are readily degraded on transmission. Moreover, an *unknown* quantum information cannot be cloned [6], or else measured and regenerated, without degradation. Thus, whenever a high-fidelity output state is required at the end-location (for example, a qubit that is to be part of a quantum secure network), a direct transmission of the quantum signal would generally not be the preferred option. Instead, teleportation is likely to serve as an essential component of future quantum communication networks [80].

A protocol for the *ideal* teleportation of qubit states was first proposed by Bennett *et al* [10] and extensively discussed in a very general setting by Horodecki *et al* [79]. Suppose Alice has in her possession an arbitrary qudit state $|\psi_i\rangle$ that she wants to teleport to Bob. The protocol involves Alice and Bob sharing an EPR resource, or more precisely a maximally entangled two-qudit state.

Alice makes a joint measurement (in the basis of maximally entangled states) on her half of the entangled qudit and the state to be teleported. She transmits via a classical communication channel those measurement results to Bob, who reconstructs the original state as $|\psi_f\rangle$ by applying a local unitary correction to his half of the maximally entangled two-qudit state.

Fidelity between the initial state and the final state is the standard figure of merit used to measure the effectiveness of the teleportation transfer [81]. Specifically, the quality of the teleportation channel Λ_ρ is often quantified via the *average* fidelity [79]:

$$\mathcal{F}_{\text{tele}}(\Lambda_\rho)=\int \text{d}\psi_i\,\langle\psi_i|\,\Lambda_\rho(|\psi_i\rangle\langle\psi_i|)|\psi_i\rangle \tag{5.8}$$

where ρ is the density matrix of the shared resource and the integral is performed over the Haar measure (i.e. a uniform distribution over all possible pure input states of a *fixed* Hilbert space dimension). Not surprisingly, $\mathcal{F}_{\text{tele}}(\Lambda_\rho)$ depends on the degree of entanglement of ρ. For instance, if ρ is maximally entangled, one has an ideal teleportation channel and thus $\mathcal{F}_{\text{tele}}=1$. In general, for a given resource state ρ of local dimension d, the maximal average fidelity achievable was found to be [79]

$$\mathcal{F}_{\text{tele}}^{\text{max}}(\Lambda_\rho)=\frac{\mathcal{F}_{\text{max}}d+1}{d+1} \tag{5.9}$$

where $\mathcal{F}_{\text{max}}$ is the singlet fraction (or more appropriately, the fully-entangled fraction) of ρ, which is the largest possible overlap (i.e. Schumacher's fidelity) between ρ and a maximally entangled two-qudit state:

$$\mathcal{F}_{\text{max}}=\max_{U_A,U_B}\mathcal{F}(\rho,U_A\otimes U_B\left|\Phi_d^+\right\rangle\!\left\langle\Phi_d^+\right|U_A^\dagger\otimes U_B^\dagger). \tag{5.10}$$

Here U_A and U_B are, respectively, local unitary operator acting on the Hilbert space of Alice's and Bob's qudit.

5.4.1. Teleportation fidelity and bounds. The problem of establishing the quality of a practical teleportation experimental process then becomes the problem of a fidelity measurement. To this end, it is insightful to compare the fidelity bound of equation (5.9) against that arising from employing an optimal cloning machine [7, 82, 83]. For simplicity, we shall restrict our discussion to the $d = 2$, i.e. the qubit case. For general fidelity bounds of cloning, we refer the reader to [8]. In the case of an arbitrary qubit, it is known that a classical strategy of teleportation, whereby the state is measured and then regenerated, will incur extra noise (see, e.g. [79, 81, 84, 85]). This limits the fidelity for any classical measure-and-regenerate protocol, to $\mathcal{F}_{\text{tele}}^{\text{class}} \leqslant \frac{2}{3}$ [79, 81, 82]. It is also known that the fidelity $\mathcal{F} = \frac{5}{6}$ is the maximum for any symmetric $1 \to 2$ cloning process, i.e. $\mathcal{F} > \frac{5}{6}$ ensures that there can be no 'copies' $|\psi_c\rangle^{\otimes 2}$ taken of the state $|\psi_i\rangle$ that has a fidelity $\mathcal{F} = |\langle\psi_i|\psi_c\rangle|^2$ greater than $\frac{5}{6}$ [7, 86, 87]. The *ideal* teleportation protocol clearly exceeds this bound because the state $|\psi_i\rangle$ at Alice's location is destroyed by her measurements. The final teleported state at Bob's location is therefore not a copy of the state $|\psi_i\rangle$, but a unique secure transfer of it. In view of these limits, a fidelity $\mathcal{F}_{\text{tele}} > \frac{2}{3}$ is the benchmark figure of merit used to justify quantum teleportation in qubit teleportation experiments [88, 89]. The no-cloning teleportation is achieved when $\mathcal{F}_{\text{tele}} > \frac{5}{6}$. For qubit-teleportation experiments carried out photonically, the fidelity estimates were evaluated for a post-selected subensemble, selected conditionally on all photons being detected at Bob's location [88] (see also [90]). The post-selection was required due to poor detection efficiencies.

5.4.2. Continuous variable teleportation. A protocol for continuous variable teleportation was developed by Vaidmann [91] as well as Braunstein and Kimble [92]. Here, the EPR resource was a continuous-variable EPR entangled state [61, 93]. Defining 'quantum teleportation' as taking place where there can be no classical measure-and-regenerate strategy that can replicate the teleportation fidelity $\mathcal{F}_{\text{tele}}$, it was shown that $\mathcal{F}_{\text{tele}} > 1/2$ is the benchmark fidelity bound to demonstrate the quantum teleportation of an unknown coherent state [84]. This fidelity has been achieved experimentally for optical states [94–97]. The fidelity at which there can be no replica of the state at a location different to Bob's corresponds to the no-cloning fidelity [98]. For coherent states, the fidelity $\mathcal{F} = 2/3$ is the maximum for any cloning process [99], and the fidelity $\mathcal{F}_{\text{tele}} > 2/3$ is hence the no-cloning benchmark for the teleportation of a coherent state. The no-cloning teleportation limit has been achieved for coherent states [97]. Although reporting lower teleportation fidelities, the continuous variable experiments did not rely on *any* post-selection of data.

5.4.3. Connection with other desired properties. Fidelity is the figure of merit used to quantify the success and usefulness of the teleportation protocol. For example, no-cloning teleportation allows bounds to be placed on the quality of any unwanted copies of the teleported state. As a consequence, fidelity is also used to determine the constraints on the nature of the entangled resource, so that certain conditions are met. For example, by using the fidelity bounds, it was proved [100] that any entangled two-qubit (mixed) state is useful for quantum teleportation if arbitrary local operations assisted by classical communications (including arbitrary local filtering operations) are allowed prior to the teleportation experiment. Similarly, a connection has been shown, between no-cloning teleportation and the requirement of a steerable resource [101], between the fully-entangled fraction and a steerable [60] resource, as well as a Bell-nonlocal [102] resource, see, e.g. [103] and references therein.

In the context of the papers examined for this review, the teleportation fidelity is most frequently defined with respect to a pure state that Alice wants to transport. Bob's teleported state will generally be mixed. However, one can see that more generally the fidelity for the teleportation of a mixed state needs also to be considered, given that the state prepared at Alice's location would not usually be pure.

5.5. Fidelity in phase space

In the following, we show how the fidelity can be computed with phase space methods [104]. The symmetrically ordered Wigner function representation [105] was first applied to dynamical problems by Moyal [106]. Although it is is generally non-positive, it is common to use tomography to measure the Wigner function [107], to represent a density matrix. Other schemes using a classical-like phase-space for bosons include the anti-normally ordered, positive Q-function distribution [108], and the normally-ordered P-function distribution [109], which is non-positive and singular in some cases.

These have been generalized to positive distributions on non-classical phase spaces [110–112], which are normally-ordered, non-singular, positive representations. These have been employed in many different fields such as quantum optics [113–115], Bose–Einstein condensates [116–118] and quantum opto-mechanical systems [119–121], as well as spin [122–124] and Fermi [125] systems.

In these methods, a density operator $\hat{\rho}$ is generically represented as

$$\hat{\rho} = \int P(\vec{\alpha})\, \hat{\Lambda}(\vec{\alpha})\, \mathrm{d}\vec{\alpha}, \tag{5.11}$$

where $\hat{\Lambda}(\vec{\alpha})$ is a projection operator that forms the basis in the description of a density operator and $P(\vec{\alpha})$ is the quasi-probability density that corresponds to that operator basis. In the simplest cases, $\vec{\alpha} = \boldsymbol{\alpha} = (\alpha_1, ..., \alpha_M)$ is a real or complex vector in the relevant M-mode phase-space for the first

phase-space representations defined using a classical phase-space. This can have increased dimensionality in more recent mappings.

Using equation (5.11), we can immediately see how the $\mathcal{F}_2$ fidelity can be computed. In particular, we show explicitly how the quantity $\mathrm{tr}\,(\rho\,\sigma)$ is obtained:

$$\begin{aligned}\mathrm{tr}(\rho\,\sigma) &= \mathrm{tr}\left[\int\int P_\rho\left(\vec{\alpha}\right) P_\sigma\left(\vec{\beta}\right)\hat{\Lambda}\left(\vec{\alpha}\right)\hat{\Lambda}\left(\vec{\beta}\right)\mathrm{d}\vec{\alpha}\,\mathrm{d}\vec{\beta}\right]\\ &= \int\int P_\rho\left(\vec{\alpha}\right) P_\sigma\left(\vec{\beta}\right) D(\vec{\alpha},\vec{\beta})\,\mathrm{d}\vec{\alpha}\,\mathrm{d}\vec{\beta}. \end{aligned} \tag{5.12}$$

Depending on the particular phase space representation, $D(\vec{\alpha},\vec{\beta}) \equiv \mathrm{tr}\left[\hat{\Lambda}\left(\vec{\alpha}\right)\hat{\Lambda}\left(\vec{\beta}\right)\right]$ in equation (5.12) takes different forms. An expression for this quantity was calculated by Cahill and Glauber [126], for classical phase-space methods, and it is generally only well-behaved for the Wigner and *P*-function methods. The Glauber *P*-function, although often singular, was proposed as an approximate method in fidelity tomography using this approach [127].

Here we focus on two of the most useful representations in a typical phase space numerical simulation: the Wigner and positive *P*-representations. The first of these is a rather classical-like mapping, although generally not positive-definite, while the second is always probabilistic, although defined on a phase-space that doubles the classical dimensionality. In the Wigner representation, $\vec{\alpha} \equiv \boldsymbol{\alpha}$, and

$$D(\boldsymbol{\alpha},\boldsymbol{\beta}) = \pi^M \delta^M\left(\boldsymbol{\alpha}-\boldsymbol{\beta}\right), \tag{5.13}$$

where M is the dimension of the vector $\boldsymbol{\alpha}$ and the Mth dimensional Dirac delta function is $\delta^M\left(\boldsymbol{\alpha}-\boldsymbol{\beta}\right) = \delta\left(\alpha_1-\beta_1\right)...\delta\left(\alpha_M-\beta_M\right)$. This leads to [126]

$$\mathrm{tr}(\rho\,\sigma) = \pi^M \int P_\rho\left(\boldsymbol{\alpha}\right) P_\sigma\left(\boldsymbol{\alpha}\right)\,\mathrm{d}\boldsymbol{\alpha}. \tag{5.14}$$

In the positive *P* representation, a single mode is characterized by two complex numbers, so $\vec{\alpha} = (\alpha,\alpha^+)$. The notation α^+ indicates that this variable represents a conjugate operator, and is stochastically complex conjugate to α in the mean. This doubles the dimension of the relevant classical phase space. There is an intuitive interpretation that it allows one to map superpositions directly into a phase-space representation. The density operator in positive *P* representation is then given by equation (5.11), but now the operator bases have the form:

$$\hat{\Lambda}\left(\vec{\alpha}\right) = \frac{|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}^+|}{\langle\boldsymbol{\alpha}^{+*}|\boldsymbol{\alpha}\rangle}, \tag{5.15}$$

and the product trace is:

$$D(\vec{\alpha},\vec{\beta}) = \frac{\langle\boldsymbol{\beta}^+|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}^+|\boldsymbol{\beta}\rangle}{\langle\boldsymbol{\alpha}^{+*}|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}^{+*}|\boldsymbol{\beta}\rangle}. \tag{5.16}$$

The quantity $\mathrm{tr}(\rho\,\sigma)$ is more complicated in this case but still follows the structure of equation (5.12). Fidelity measures of the form given in equation (2.16) also involve the purity, $\mathrm{tr}\left(\rho^2\right)$. This can be calculated similarly.

One great advantage of phase space methods is that quasi-probability densities allow numerical simulation to be carried out. Typically, samples are drawn from these probability densities, which are then evolved dynamically. Finally, observables of interest are computed by the Monte Carlo method, which is usually the only practical technique for very large Hilbert spaces. Likewise, fidelity measures can be computed numerically in a typical Monte Carlo scheme. In particular, we consider $\mathcal{F}_2$, which, as we will discuss, is the most tractable form of fidelity.

Let ρ be the initial density operator of the system and we want to compute the fidelity of a quantum state at a later time, which is characterized by the density operator σ, with respect to ρ. Suppose that the initial state ρ and its corresponding phase space distribution are known in a numerical simulation.

For simplicity, suppose that ρ is a pure state, which is a common but not essential assumption, and implies that $\mathrm{tr}\left(\rho^2\right) = 1$. Even for cases where $\mathrm{tr}\left(\rho^2\right) < 1$, the final state after a time evolution will usually be no purer than the initial state. There are exceptions to this rule, since a dissipative time-evolution can evolve a mixed state of many particles to a pure vacuum state, but we first consider the case of non-increasing purity here for definiteness.

In other words, $\max\left[\mathrm{tr}\left(\rho^2\right),\mathrm{tr}\left(\sigma^2\right)\right] = \mathrm{tr}\left(\rho^2\right)$. This is convenient as the exact (quasi)-probability distribution for σ is not known and only a set of samples of this distribution is available, which leads to the sampled fidelity we discuss next.

Next, consider the sampled fidelity in the Wigner representation. The quantity $\mathrm{tr}\,(\rho\,\sigma)$ in equation (5.14) in the Monte Carlo scheme is given by:

$$\begin{aligned}\mathrm{tr}\,(\rho\,\sigma) &= \pi^M \int P_\rho\left(\boldsymbol{\alpha}\right) P_\sigma\left(\boldsymbol{\alpha}\right)\,\mathrm{d}\boldsymbol{\alpha}\\ &\approx \pi^M \frac{1}{N_{\mathrm{samples}}}\sum_{i=1}^{N_{\mathrm{samples}}} P_\rho\left(\boldsymbol{\alpha}_i\right), \end{aligned} \tag{5.17}$$

where N_{samples} is the sample size of the probability distribution P_σ. We note that there can be issues with the fact that the same random variable occurs in both the distributions, leading to practical problems if both the distributions are sampled. This can be avoided if one of the Wigner functions is known analytically.

The same quantity can be computed in the positive *P* representation. It is then possible to use two independent sets of random variables, so that both the distributions can be obtained from random sampling:

$$\mathrm{tr}\,(\rho\,\sigma) \approx \frac{1}{N_{\mathrm{samples}}^2}\sum_{i,j}^{N_{\mathrm{samples}}} D(\vec{\alpha}_i,\vec{\beta}_j). \tag{5.18}$$

Here, the factor N_{samples}^2 comes from the product of $P_\rho\left(\vec{\alpha}\right) P_\sigma\left(\vec{\beta}\right)$ in equation (5.12) under the usual assumption of equally weighted samples. This shows that it is possible to compute $\mathcal{F}_2$ fidelities from a phase-space simulation. This is useful when trying to predict performance of a quantum technology or memory in an application involving storage of an exotic quantum state. We emphasize that if one of the calculated states is a pure state, then all fidelity measures give the same result.

Admittedly, this quantity is more complicated than equation (5.17) in the Wigner representation. In addition, the sampling error can be very large in some cases, as discussed by Rosales-Zarate and Drummond [128]. When this occurs, representations such as the generalized Gaussian representations [125, 129–131] can be employed, and clearly the purities can be estimated in a similar way if the initial state is not pure.

Overall, $\mathcal{F}_2$ appears to be the most suitable fidelity measure in a dynamical simulation or measurement using phase-space techniques, where only the initial state with its probability distribution is known. It is the only measure using easily computable Hilbert–Schmidt norms that satisfies all of the Jozsa axioms.

5.6. Techniques of fidelity measurement

As pointed out in section 2.1, fidelity is a relative measure. The results of fidelity measurements on different Hilbert spaces are not the same. The $\mathcal{F}_1$ fidelity may improve if the measured Hilbert space has a lower dimensionality, simply because it is defined as a maximum over all possible purifications. Hence, the $\mathcal{F}_2$ fidelity has the advantage that it is generally less biased towards high values, as shown in the examples of the previous section, and is always true for qubits.

The adage of being cautious in comparing apples to oranges should be remembered. In general terms we will distinguish six different approaches that are described below, as applied to typical physical implementations [12, 132]. The real utility of a given fidelity measure is how well it matches the requirements of a given application. Analyzing quantum logic gates and memories is one of the most widespread and useful applications of fidelity, and hence we will give examples of these applications.

These general considerations about physical implementation apply to all of the various applications listed in this section. The examples referenced here are necessarily incomplete, as this is not a full review of experimental implementations. Nevertheless, some typical experimental measurements are referenced in each of the following application examples.

5.6.1. Atomic tomography fidelity. Atomic or ionic fidelity measurements involve a finite, stationary, closed quantum system, where each state can be accessed and projected [19, 133]. These measurements are usually relatively simple. To obtain the entire density matrix for a calculation of mixed state fidelity involves a tomographic measurement. Thus, for example, in qubit tomography one must measure both the diagonal elements, which are level occupations, as well as off-diagonal elements that are obtained through Rabi rotations that transform off-diagonal elements into level occupations for measurement. This approach is easiest to implement when the Hilbert space has only two or three levels. Typical examples of this technique involve trapped ions, whose level occupations are measured using laser pulses and fluorescence photo-detection.

It is not always clear in such measurements how the translational state is measured, or if it is even part of the relevant Hilbert space, which is necessary in order to understand how the fidelity is defined. The problem is that the full quantum state of an isolated ion or atom always has both internal and center-of-mass degrees of freedom, so that a pure state is:

$$|\Psi\rangle = \sum_{ij} C_{ij} |\psi_i\rangle_{\rm int} |\phi_j\rangle_{\rm CM} . \tag{5.19}$$

Here $|\psi\rangle_{\rm int}$ is the internal state defined by the level structure, while $|\phi\rangle_{\rm CM}$ defines the center-of-mass degree of freedom. The actual density matrix even for a single ion or atom therefore involves different spatial modes for the center-of-mass, such that:

$$\rho = \sum_{ijkl} \rho_{ijkl} |\psi_i\rangle_{\rm int} |\phi_j\rangle_{\rm CM} \langle\phi_k|_{\rm CM} \langle\psi_l|_{\rm int} . \tag{5.20}$$

Next, we can consider two possible situations:

Full tomography: Suppose that the target state is $|\Psi_0\rangle = |\psi_0\rangle_{\rm int} |\phi_0\rangle_{\rm CM}$, so that the center-of-mass position is part of the relevant Hilbert space. Under these conditions, only the measured states with $k = j = 0$ are in the same *overall* quantum state as the target state. This may prevent complete visibility in an interference measurement in which the center-of-mass position is relevant. If this is the case, one should consider the center-of-mass position as part of the relevant Hilbert space. Hence one must consider rather carefully if it is necessary to investigate the translational state fidelity as well in this type of application. This has been investigated in ion-trap quantum computer gate fidelity measurements [134].

Partial tomography: The center-of-mass part of the Hilbert space may not matter if the internal degrees of freedom are decoupled sufficiently from the spatial degrees of freedom, so that only the internal degrees of freedom are relevant over the time-scales that are of interest. In these cases the density matrix can be written as:

$$\rho = \rho_{\rm int} \otimes \rho_{\rm CM} . \tag{5.21}$$

Provided this factorization is maintained throughout the experiment, it may well be enough to only measure the internal degrees of freedom. However, any spin-orbit or similar effective force that couples the internal and translational degrees of freedom will cause entanglement. This will sometimes mean a reduction in fidelity, since the entangled state can become mixed after tracing out the spatial degrees. Relatively high fidelities have been measured with this approach. Depending on the system, this can be viewed as occurring because coupling to phonons is weak [135] or because experiments occur on faster time-scales than the atomic motion [68].

5.6.2. Photonic fidelity. Photonic measurements are typical of quantum memories [11, 136], communications or cryptography when photons are used as the information carrier. In the case of a quantum memory, a quantum state is first encoded in a well-defined spatiotemporal mode(s), then dynamically coupled into the memory subsystem, stored for a chosen

period, and coupled out into a second well-defined spatiotemporal field mode(s) where it can be measured [137]. We note that temporal mode structure is an essential part of defining a quantum state.

The actual quantum state in these cases is defined as an outer product of photonic states in each possible mode $|n_k\rangle_k$, where n_k is the photon number in each mode, so that:

$$|\Psi\rangle = \sum_{\boldsymbol{n}} C_{\boldsymbol{n}} |n_1\rangle_1 \otimes \ldots |n_K\rangle_K . \tag{5.22}$$

Here each mode has an associated mode function $\boldsymbol{u}_k$, which is typically localized in space-time, since technology applications are carried out in finite regions of space, over finite time-intervals. We implicitly assume a finite total number of modes K, although there is no physical upper bound except possibly that from quantum gravity.

To obtain the output density matrix, the most rigorous approach is to use pulsed homodyne detection to isolate the mode(s) used, with a variable phase delay or other methods to measure the off-diagonal elements. This gives a projected quadrature expectation value of the relevant single mode operator. By tomographic reconstruction, one can obtain the Wigner function [107]. We show in section 5.5 that this phase-space technique directly gives the quantum fidelity as an $\mathcal{F}_2$ measure. Obtaining any other fidelity measure generally requires the reconstructed density matrix. See, however [35, 138], where the authors discuss a direct measurement of the superfidelity, $\mathcal{F}_{\mathrm{N}}$, together with a lower bound called the sub-fidelity for photonic states encoded in the polarization degree of freedom. Since $\mathcal{F}_{\mathrm{N}} = \mathcal{F}_1$ for qubit states [21], the method of [35] actually amounts to a direct measurement of $\mathcal{F}_1$ for these qubit states without resorting to quantum state tomography.

5.6.3. Conditional fidelity. In some types of photonic fidelity measurement the state may not be found at all in some of the measurements. This is typically the case in photodetection experiments with low photon number, where a qubit can be encoded into two spatial or polarization modes, as $|\psi\rangle = \frac{1}{\sqrt{|a|^2+|b|^2}}(a\,|0\rangle_1\,|1\rangle_2 + b\,|1\rangle_1\,|0\rangle_2)$. Problems arise when no photon is detected at all, either because the photodetector was inefficient, or because the photon was lost during the transmission, thereby making the input a vacuum state, which is in a larger Hilbert space.

As a result, reported measurements are sometimes defined by simply conditioning all results on the presence of a detected photon(s). Unless the target state is itself defined to be the conditioned state, this conditional fidelity is best regarded as an upper bound for the true mixed state fidelity, which includes these loss effects. The potential difficulty with conditional fidelity measured in this way is that it essentially involves an assumption of fair sampling. In other words, photons may be lost through detector inefficiency, but they may also be lost in any number of other ways.

While detector loss can be regarded as simply a measurement issue, unrelated to the state itself, there is also a possibility that the state was already degraded before it reached the detector, and hence the true fidelity is lower than estimated by the conditional measurement. Yet in many applications, like quantum logic gates and computing, one may have to repeat the same quantum memory process many times in succession. In these cases any inefficiency that occurs prior to detection grows exponentially with the number of gates, and can become an important issue. In other quantum information processes however, it might be argued that this effect is not relevant. In quoting fidelity, it is thus important to match the target state with the final intended application.

5.6.4. Inferred fidelity. In a similar way to efficiency problems that lead to conditional fidelity measures, the spatiotemporal mode may change from measurement to measurement in photonic experiments. This leads to a mixed state. As a result, the increased number of modes present can enlarge the Hilbert space in a way that is not detectable through measurements of photodetection events without using interferometry or local oscillators. This approach is sometimes combined with a conditional measurement.

An example of this approach is a recently reported quantum memory for orbital angular momentum qubits [139]. This experiment has many robust and useful features, using spatial mode projection to ensure that the correct transverse mode is matched from input to output. However, the report does not describe how longitudinal or temporal mode structure was determined, leaving this issue as an open question at this stage.

A photon-counting approach cannot usually detect the full mode structure. For example, suppose one has a wide range of longitudinal modes that can be occupied, having distinct frequencies and/or temporal mode structures, and each occurring with a probability $P_k \ll 1$, so that:

$$\rho = \sum_k P_k\,|0\rangle_1 \ldots |1\rangle_k \ldots \langle 0|_1 \ldots \langle 1|_k \ldots . \tag{5.23}$$

This is a mixed state in which a single photon could be in any longitudinal mode with a given probability P_k.

Let us now compare this with a desired pure state σ, for example:

$$\sigma = |1\rangle_1 \ldots |0\rangle_k \ldots \langle 1|_1 \ldots \langle 0|_k \ldots .. \tag{5.24}$$

It is clear that, for any definition of fidelity, $\mathcal{F} = \mathrm{tr}\,(\sigma\rho) = P_1 \ll 1$. One could attempt to measure this fidelity with a photon-counting measurement, combined with the *assumption* that there is only one longitudinal mode present. If all measurements give exactly one count, then this measurement, combined with the single-mode assumption would lead to an inferred state fidelity of $\mathcal{F}_{\mathrm{inf}} = 1$. This does not match the true fidelity in this example.

Fidelity measurements like these generally make the assumption that the mode structure that is measured matches the desired mode structure, even when it is not measured directly. As a result, the inferred fidelity may be higher than the true fidelity, and should be considered as an upper bound. This may cause problems if one must carry out a binary quantum logic operation with input signals derived from two different sources such that the modes should be matched in time and/or frequency. Under these conditions, it is the true

fidelity, including the effects of losses and modal infidelity that is important. These questions have been investigated in experiments that carry out full tomographic measurements to reconstruct single-photon Fock states using homodyne measurement techniques [140].

5.6.5. Cloned fidelity. A fifth type of fidelity measurement is obtained as a variant of a quantum game in which a number of copies of a quantum state may be recorded or stored [141]. From subsequent measurements, it is possible to infer, for example, using maximum likelihood measurements, what the original state was. The inferred state can be compared with the original using fidelities. Unfortunately the no-cloning theorem tells us that multiple copies of any single quantum state cannot be obtained reliably. Hence, while one can infer a fidelity from multiple copies of a state, the entire process that includes first generating multiple copies of a quantum state will always involve an initial reduction in fidelity. This is important in some types of application.

5.6.6. Logic and process fidelity. Finally, we turn to a different type of fidelity used to analyze quantum processes rather than states or density matrices. Quantum processes are also known [46] as quantum channels, or mathematically as completely positive maps or super-operators. For the case of a quantum process, one may wish to analyze the fidelity of an actual quantum operation to an intended quantum operation. This may include any quantum technology from logic gates to memories, or indeed any input-output process. An operation is defined in the general sense of any quantum map $\mathcal{E}(\rho)$, from an input density matrix ρ_{in} to an output density matrix ρ_{out}. Their fidelity is discussed by Gilchrist *et al* [23].

Just as any density matrix has a matrix representation in the Hilbert space of state vectors of dimension d, quantum channels have a matrix representation in terms of a basis set of d^2 quantum operators A_j, where $\text{tr}\left(A_j^\dagger A_k\right) = \delta_{jk}$. Using this basis, any quantum operation can be written as:

$$\rho_{\text{out}} = \mathcal{E}(\rho_{\text{in}}) = \sum_{mn} P_{mn} A_m \rho_{\text{in}} A_n^\dagger. \tag{5.25}$$

Here P_{mn} are the elements of the so-called process matrix P, which provides a convenient way of representing the operation $\mathcal{E}$.

At first, this seems rather different to density matrices, as discussed throughout this review. Yet it is easy to show via the Choi–Jamiolkowski isomorphism [142, 143] that one can define a new quantum 'density matrix' [23] on the enlarged Hilbert space of dimension d^2, such that $\rho^{\mathcal{E}} = P/d$. Hence any fidelity or distance measure for quantum states can also be applied to processes, by the simple technique of dimension squaring. We will not investigate this in detail here except to remark that all of the different fidelity measures used for density matrices can be applied directly to quantum processes. In fact, even though $\mathcal{F}_{\text{GM}}$ does not comply with all of Jozsa's axioms, its applicability to non-positive-semidefinite operators has made it more appealing to some authors (see, e.g. [144, 145] and references therein) than $\mathcal{F}_{\text{I}}$ in quantifying the performance of certain quantum gates. However, in general, for process fidelity it is common to impose, instead, additional requirements for the fidelity in addition to the axioms used here.

A typical example is the measurement of quantum process fidelity in a CNOT gate [146]. In this early photonic measurement, the counting fidelity was measured using conditional techniques. Thus, as explained above, these results should be regarded as an upper bound to the actual quantum fidelity, once mode-mismatch errors and losses are included. Other, more recent, process fidelity measurements with quantum logic gates have been carried out with ion traps [77, 147], liquid nuclear magnetic resonance [148], solid-state silicon [149] and superconducting Josephson qubits [150].

6. Summary

We have reviewed the requirements for a mixed state fidelity measure [9], and analyzed a number of candidate measures of fidelity for their compliance with these requirements, as well as other considerations. While there are many candidates, most of them do not fully comply with the Jozsa axioms, although some of these alternatives do have useful properties.

Despite the above observation, there do exist an infinite number of compliant fidelities. Among them, three well-defined measures that fully satisfy the Jozsa axioms for fidelity measures are of particular interest, due to their physical interpretation and measurement properties: these are the Uhlmann–Jozsa fidelity $\mathcal{F}_{\text{I}}$, the non-logarithmic variety of the quantum Chernoff bound $\mathcal{F}_{\text{Q}}$ and the Hilbert–Schmidt fidelity, $\mathcal{F}_2$. It is worth noting that both $\mathcal{F}_{\text{I}}$ and $\mathcal{F}_2$ are particular cases of an infinite family of Jozsa-compliant fidelity measures $\mathcal{F}_p$, each associated with a Schatten-von-Neumann p-norm.

In this review, we have focused on two specific cases of these norm based fidelity measures $\mathcal{F}_p$, as well as the quantum Chernoff bound $\mathcal{F}_{\text{Q}}$. Analyzing the properties of this family of candidate measures for other integer values of $p > 2$ is clearly something that may be of independent interest. On the other hand, despite much effort, the validity of a few desired properties of various candidate fidelity measures remains unknown (see tables 3, 5 and 6 for details). For each of these conjectured properties, at least 2000 optimizations with different initial starting points have been carried out for each Hilbert space dimension $d = 2, 3, \ldots, 10$. Given the fact that intensive numerical searches have been carried out for counterexamples to these properties for small Hilbert space dimensions, we are inclined to conjecture that these properties are indeed valid.

An intriguing result is that of the $\mathcal{F}_p$ fidelities investigated, the $\mathcal{F}_{\text{I}}$ fidelity gives the largest average values when random density matrices are compared. While this relationship is only universal for the qubit case—otherwise there are occasional exceptions—it is found on average for higher dimensional Hilbert spaces as well. This is clearly related to the fact that the $\mathcal{F}_{\text{I}}$ fidelity is defined as a maximum fidelity over purifications. These purifications represent an unmeasured portion

of Hilbert space. Hence, measuring $\mathcal{F}_1$ on a subspace could introduce a bias compared to a more complete measurement on a larger relevant space, if there are additional errors in the unmeasured part of the relevant Hilbert space.

To conclude, while the Uhlmann–Jozsa measure is the most widely known fidelity measure, there are other alternatives which have properties that can make them preferable under some circumstances. They may be either simpler to compute or more relevant to certain applications. For example, the Hilbert–Schmidt fidelity measure $\mathcal{F}_2$ is well-defined even for unnormalized density matrices, and appears less biased towards high values. Finding out the full implications of this and other mathematical properties of the candidate measures, however, is too broad a research topic to be considered within the present review.

Note added. After the completion of this work, we became aware of the work of [151], which has provided also a fidelity measure that can be cast in the form of (2.19) and which complies with all of Jozsa's axioms.

Acknowledgments

YCL, PEMFM, and PDD contributed equally towards this work. This work is supported by the Ministry of Science and Technology, Taiwan (Grants No. 104-2112-M-006-021-MY3 and 107-2112-M-006-005-MY2) and the Center for Quantum Technology, Hsinchu, Taiwan. PDD and MDR thank the Australian Research Council and the hospitality of the Institute for Atomic and Molecular Physics (ITAMP) at Harvard University, supported by the NSF. YCL acknowledges useful discussions with N Gisin and N Sangouard. PDD thanks B Sparkes for useful discussions.

Appendix A. Detailed proofs

In this appendix, detailed proofs are presented for the fidelity results in the earlier sections, where they are not given already.

A.1. Norm based fidelity properties

Theorem A.1. *All norm-based fidelities, $\mathcal{F}_p$, obey the Jozsa axioms for $p \geqslant 1$.*

Proof.

(J1a) $\mathcal{F}_p(\rho,\sigma) \in [0,1]$. The minimum bound is trivial, since norms are positive semi-definite. That the maximum bound holds is obtained from the Hölder inequality, since: $\left\|\sqrt{\rho}\sqrt{\sigma}\right\|_p^2 \leqslant \left\|\sqrt{\rho}\right\|_{2p}^2 \left\|\sqrt{\sigma}\right\|_{2p}^2 = \|\rho\|_p \|\sigma\|_p \leqslant \max\left[\|\sigma\|_p^2, \|\rho\|_p^2\right]$.

(J1b) $\mathcal{F}_p(\rho,\sigma) = 1$ *if and only if* $\rho = \sigma$. Clearly, $\mathcal{F}_p(\rho,\rho) = 1$ for identical operators, since $\left\|\sqrt{\rho}\sqrt{\rho}\right\|_p = \|\rho\|_p^2 = \max\left[\|\rho\|_p^2, \|\rho\|_p^2\right]$. To prove the converse, note from the above proof of J1a that the maximum bound is attained if and only if the Hölder inequality becomes an equality. Taking into account of the normalization of density matrices, we thus see that the maximum bound of $\mathcal{F}_p(\rho,\sigma) = 1$ is attained if and only if $\rho = \sigma$.

(J1c) $\mathcal{F}_p(\rho,\sigma) = 0$ *if and only if* $\rho\,\sigma = 0$. This follows since $\rho\,\sigma = 0 \Longleftrightarrow \sqrt{\rho}\sqrt{\sigma} = 0 \Longleftrightarrow \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_p = 0$.

(J2) $\mathcal{F}_p(\rho,\sigma) = \mathcal{F}_p(\sigma,\rho)$ is clearly true from the symmetry of the matrix norm under transposition.

(J3) $\mathcal{F}_p(\rho,\sigma) = \mathrm{tr}(\rho\,\sigma)$ if either ρ or σ is a pure state. From the definition, we have $\|\sqrt{\rho}\sqrt{\sigma}\|_p^2 = \left[\mathrm{tr}\left(\sqrt{\rho}\,\sigma\,\sqrt{\rho}\right)^{\frac{p}{2}}\right]^{\frac{2}{p}}$. Let ρ be a pure state, then $\rho = \sqrt{\rho} = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$, the expression thus simplifies to $\left[(\langle\psi|\,\sigma\,|\psi\rangle)^{\frac{p}{2}}\mathrm{tr}\left(\rho^{\frac{p}{2}}\right)\right]^{\frac{2}{p}}$, which reduces to $\langle\psi|\,\sigma\,|\psi\rangle = \mathrm{tr}(\rho\,\sigma)$ since $\mathrm{tr}\left(\rho^{\frac{p}{2}}\right) = \mathrm{tr}\rho = 1$. Similarly, the normalizing factor is unity, since $\|\rho\|_p^2 = 1 \geqslant \|\sigma\|_p^2$, and the argument holds if ρ, σ are interchanged.

(J4) $\mathcal{F}_p(U\rho\,U^\dagger, U\sigma U^\dagger) = \mathcal{F}_p(\rho,\sigma)$. This follows from the unitary invariance of the matrix norm. □

A.2. Normalization

We now show that both $\mathcal{F}_{\mathrm{AM}}$ and $\mathcal{F}_{\mathrm{GM}}$ obey axiom (J1a) and (J1b).

Theorem A.2. $\mathcal{F}_{\mathrm{AM}}(\rho,\sigma), \mathcal{F}_{\mathrm{GM}}(\rho,\sigma) \in [0,1]$ *with the upper bound attained if and only if* $\rho = \sigma$.

Proof. The non-negativity of $\mathcal{F}_{\mathrm{AM}}$ and $\mathcal{F}_{\mathrm{GM}}$ is obvious. Next, we prove that these quantities are upper bounded by 1. To this end, we recall that the geometric mean between two numbers is always upper bounded by its arithmetic mean, hence, $\mathcal{F}_{\mathrm{AM}}(\rho,\sigma) \leqslant \mathcal{F}_{\mathrm{GM}}(\rho,\sigma) \leqslant 1$, where the second inequality follows easily from the Cauchy–Schwarz inequality.

Moreover, the Cauchy–Schwarz inequality is saturated if and only if its entries are scalar multiples of each other. Since our entries have unity trace (density matrices), saturation can only occur if $\rho = \sigma$. It is easy to see by inspection that both $\mathcal{F}_{\mathrm{AM}}(\rho,\rho)$ and $\mathcal{F}_{\mathrm{GM}}(\rho,\rho)$ indeed equal to unity. Thus $\mathcal{F}_{\mathrm{AM}}(\rho,\sigma) = \mathcal{F}_{\mathrm{GM}}(\rho,\sigma) = 1$ if and only if $\rho = \sigma$. □

A.3. Multiplicativity

Theorem A.3. *The measure $\mathcal{F}_2$ is generally super-multiplicative, but is multiplicative when appended by an uncorrelated ancillary state, or when considering tensor powers of the same states.*

Proof. Let us define $\rho = \rho_1 \otimes \rho_2$ and $\sigma = \sigma_1 \otimes \sigma_2$, then start by noting that

$$\mathcal{F}_2(\rho,\sigma) = \frac{\mathrm{tr}(\rho_1\sigma_1)\mathrm{tr}(\rho_2\sigma_2)}{\max\left[\mathrm{tr}(\rho_1^2)\mathrm{tr}(\rho_2^2), \mathrm{tr}(\sigma_1^2)\mathrm{tr}(\sigma_2^2)\right]} \tag{A.1}$$

and

$$\mathcal{F}_2(\rho_1,\sigma_1)\mathcal{F}_2(\rho_2,\sigma_2) = \frac{\prod_{i=1}^2 \mathrm{tr}(\rho_i\sigma_i)}{\prod_{i=1}^2 \max\left[\mathrm{tr}(\rho_i^2),\mathrm{tr}(\sigma_i^2)\right]}. \tag{A.2}$$

Clearly, saturation of inequality (3.4) is obtained if the denominators of the two equations above coincide, i.e.

$$\max\left[\prod_{i=1}^2 \mathrm{tr}(\rho_i^2), \prod_{j=1}^2 \mathrm{tr}(\sigma_j^2)\right] = \prod_{i=1}^2 \max\left[\mathrm{tr}(\rho_i^2),\mathrm{tr}(\sigma_i^2)\right]. \tag{A.3}$$

It is easy to check that this equation holds when (at least) one of the following is observed

- $\mathrm{tr}(\rho_i^2) = \mathrm{tr}(\sigma_i^2)$ for some $i = 1, 2$,
- $\mathrm{tr}(\rho_i^2) > \mathrm{tr}(\sigma_i^2)$ for both $i = 1, 2$,
- $\mathrm{tr}(\rho_i^2) < \mathrm{tr}(\sigma_i^2)$ for both $i = 1, 2$.

Note that the first condition is satisfied for the scenario when each quantum state is appended, respectively, by a quantum state with the same purity (e.g. when they are both appended by same ancillary state τ). On the other hand, the second/third condition is satisfied for the scenario of tensor powers, i.e. $\rho_1 = \rho_2$ etc, see equation (3.5). When none of the above conditions is satisfied, it is easy to see that we must have the r.h.s. of equation (A.3) larger than its l.h.s., and hence super-multiplicativity. For example, if $\mathrm{tr}(\rho_1^2) > \mathrm{tr}(\sigma_1^2)$ and $\mathrm{tr}(\rho_2^2) < \mathrm{tr}(\sigma_2^2)$, then the r.h.s. of (A.3) becomes $\mathrm{tr}(\rho_1^2)\mathrm{tr}(\sigma_2^2)$ which has to be larger than both $\mathrm{tr}(\rho_1^2)\mathrm{tr}(\rho_2^2)$ or $\mathrm{tr}(\sigma_1^2)\mathrm{tr}(\sigma_2^2)$ by assumption. The proof for the case when $\mathrm{tr}(\rho_1^2) < \mathrm{tr}(\sigma_1^2)$ and $\mathrm{tr}(\rho_2^2) > \mathrm{tr}(\sigma_2^2)$ proceeds analogously. □

Theorem A.4. *The measure $\mathcal{F}_\mathrm{C}$ is generally supermultiplicative.*

Proof. Let $d_1 := \dim\rho_1 = \dim\sigma_1$, $d_2 := \dim\rho_2 = \dim\sigma_2$ and, in terms of these, define $r := (d_1-1)^{-1}$, $s := (d_2-1)^{-1}$ and $t := (d_1d_2-1)^{-1}$ (or, equivalently, $t = rs/(1+r+s)$), in such a way that $r, s \in (0, 1]$. Then, the statements of the supermultiplicativity of $\mathcal{F}_\mathrm{C}$ (to be proved) and of $\mathcal{F}_\mathrm{N}$ (proved in [21]) can be expressed, respectively, as

$$\begin{aligned} &2(1-t)+2(1+t)\mathcal{F}_\mathrm{N}(\rho_1\otimes\rho_2,\sigma_1\otimes\sigma_2) \\ &\geqslant [(1-r)+(1+r)x]\,[(1-s)+(1+s)y], \end{aligned} \tag{A.4}$$

$$\begin{aligned} &2(1-t)+2(1+t)\mathcal{F}_\mathrm{N}(\rho_1\otimes\rho_2,\sigma_1\otimes\sigma_2) \\ &\geqslant 2(1-t)+2(1+t)xy, \end{aligned} \tag{A.5}$$

where, for brevity, we have defined $x := \mathcal{F}_\mathrm{N}(\rho_1,\sigma_1) \in [0,1]$ and $y := \mathcal{F}_\mathrm{N}(\rho_2,\sigma_2) \in [0,1]$. In what follows, the validity of inequality inequality (A.4) is established by showing that the r.h.s. of (A.5) dominates the r.h.s. of inequality (A.4); an inequality that can be written as

$$2(1-t)-(1-r)(1-s) \geqslant f_{r,s}(x,y) \tag{A.6}$$

where we have defined the functions

$$\begin{aligned} f_{r,s}(x,y) :=\ &(1+r)(1-s)x+(1-r)(1+s)y \\ &-[2(1+t)-(1+r)(1+s)]\,xy. \end{aligned} \tag{A.7}$$

To see that inequality (A.6) holds, it is our interest to find out the maximum of the functions $f_{r,s}(x,y)$. The extreme point of $f_{r,s}(x,y)$ is given by respectively setting the partial derivative of x and y to zero. It can be verified that, unless $r, s \in \{0, 1\}$, the extreme points of $f_{r,s}(x,y)$ lie outside the domain $x, y \in [0, 1]$. In the former cases, the extreme points lie on the boundaries of the domain. To identify them, note that the functions $f_{r,s}(x,y)$ are increasing both in x and y for all values of parameters $r, s \in [0, 1]$. This follows, for example, from the observation that their partial derivatives with respect to x and y are always linear and take non-negative values in the extremes of the domain $x, y \in [0, 1]$. Indeed,

$$\begin{aligned} \left.\frac{\partial f_{r,s}(x,y)}{\partial x}\right|_{y=0} &= (1+r)(1-s) \geqslant 0 \\ \left.\frac{\partial f_{r,s}(x,y)}{\partial x}\right|_{y=1} &= \frac{2r(1+r)}{1+r+s} > 0 \\ \left.\frac{\partial f_{r,s}(x,y)}{\partial y}\right|_{x=0} &= (1-r)(1+s) \geqslant 0 \\ \left.\frac{\partial f_{r,s}(x,y)}{\partial y}\right|_{x=1} &= \frac{2s(1+s)}{1+r+s} > 0. \end{aligned} \tag{A.8}$$

Thanks to that, it suffices to verify the validity of inequality (A.6) for $x = 1$ and $y = 1$, where $f_{r,s}(x,y)$ is maximal. In this case, however, a straightforward simplification process shows that the inequality is satisfied with saturation. □

Theorem A.5. *The measure $\mathcal{F}_\mathrm{Q}$ is generally super-multiplicative under tensor products, but is multiplicative when appended by an uncorrelated ancillary state, or when considering tensor powers of the same states.*

Proof. For given density matrices ρ_1, ρ_2, σ_1 and σ_2, we see that

$$\begin{aligned} \mathcal{F}_\mathrm{Q}(\rho_1\otimes\rho_2,\sigma_1\otimes\sigma_2) &= \min_s\ \mathrm{tr}\left[(\rho_1\otimes\rho_s)^s\,(\sigma_1\otimes\sigma_2)^{1-s}\right] \\ &= \min_s\ \mathrm{tr}(\rho_1^s\,\sigma_1^{1-s})\mathrm{tr}(\rho_2^s\,\sigma_2^{1-s}) \\ &= \min_s\ f_1(s)f_2(s) \\ &= f_1(s^*)f_2(s^*) \end{aligned} \tag{A.9}$$

where $f_i(s) := \mathrm{tr}(\rho_i^s\,\sigma_i^{1-s})$, s^* is a minimizer of the function $f_1(s)f_2(s)$, and it is worth reminding that each $f_i(s)$ is a convex function of s [39].

On the other hand, it also follows from the definition of $\mathcal{F}_\mathrm{Q}$ that

$$\begin{aligned} \mathcal{F}_\mathrm{Q}(\rho_1,\sigma_1)\mathcal{F}_\mathrm{Q}(\rho_2,\sigma_2) &= \min_{s_1} f_1(s_1)\ \min_{s_2} f_2(s_2) \\ &= f_1(s_1^*)f_2(s_2^*), \end{aligned} \tag{A.10}$$

where $s_i^* \in \mathcal{S}_i$ is a minimizer of $f_i(s_i)$, and $\mathcal{S}_i \subseteq [0,1]$ is the set of minimizers of $f_i(s_i)$. Note that the convexity of f_i guarantees that $\mathcal{S}_i$ is a convex interval in $[0,1]$.

Evidently, for general density matrices ρ_1, ρ_2, σ_1 and σ_2, the set of minimizers for $f_1(s)$ and $f_2(s)$ differ, i.e. $\mathcal{S}_1 \neq \mathcal{S}_2$. Without loss of generality, let us assume that $\min \mathcal{S}_1 \leqslant \min \mathcal{S}_2$. There are now two cases to consider, if $\max \mathcal{S}_1 \geqslant \min \mathcal{S}_2$, we must also have $\mathcal{S}_2 \cap \mathcal{S}_1 \neq \{\}$, and hence the minimizer s^* for $f_1(s)f_2(s)$ must also minimize f_1 and f_2. In this case, we have,

$$\mathcal{F}_{\mathrm{Q}}(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) = \mathcal{F}_{\mathrm{Q}}(\rho_1, \sigma_1)\mathcal{F}_{\mathrm{Q}}(\rho_2, \sigma_2), \tag{A.11}$$

meaning that multiplicativity holds true for this set of density matrices.

In the event that $\max \mathcal{S}_1 < \min \mathcal{S}_2$, we have $\mathcal{S}_2 \cap \mathcal{S}_1 = \{\}$, which implies that any minimizer s^* for $f_1(s)f_2(s)$ must be such that $s^* \notin \mathcal{S}_1$ and/or $s^* \notin \mathcal{S}_2$. As a result, we must have $f_i(s^*) \geqslant f_i(s_i^*)$ for all i, and thus

$$\begin{aligned}\mathcal{F}_{\mathrm{Q}}(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) &= f_1(s^*)f_2(s^*), \\ &\geqslant f_1(s_1^*)f_2(s_2^*), \\ &= \mathcal{F}_{\mathrm{Q}}(\rho_1, \sigma_1)\mathcal{F}_{\mathrm{Q}}(\rho_2, \sigma_2),\end{aligned} \tag{A.12}$$

which demonstrates the super-multiplicativity of $\mathcal{F}_{\mathrm{Q}}$.

For tensor powers of the same state, see equation (3.5), the fact that $\mathrm{tr}(A \otimes B) = (\mathrm{tr}\, A)\,(\mathrm{tr}\, B)$ and the above arguments make it evident that the minimizer for a single copy is also the minimizer for an arbitrary number of copies. Thus, $\mathcal{F}_{\mathrm{Q}}$ is multiplicative under tensor powers. Similarly, for an uncorrelated ancillary state τ, the definition of $\mathcal{F}_{\mathrm{Q}}$, and axiom (J1b) immediately imply that the measure is multiplicative when each quantum state is appended by τ, i.e. $\mathcal{F}_{\mathrm{Q}}(\rho \otimes \tau, \sigma \otimes \tau) = \mathcal{F}_{\mathrm{Q}}(\rho, \sigma)$. □

Theorem A.6. *The measure $\mathcal{F}_{\mathrm{AM}}$ is multiplicative under uncorrelated ancilla state, that is*

$$\mathcal{F}_{\mathrm{AM}}(\rho \otimes \tau, \sigma \otimes \tau) = \mathcal{F}_{\mathrm{AM}}(\rho, \sigma). \tag{A.13}$$

Proof. The proof follows trivially from the application of the tensor product identities

$$(A \otimes B)(C \otimes D) = AC \otimes BD \tag{A.14}$$

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B). \tag{A.15}$$

Explicitly,

$$\begin{aligned}\mathcal{F}_{\mathrm{AM}}(\rho \otimes \tau, \sigma \otimes \tau) &= \frac{2\mathrm{tr}(\rho\,\sigma \otimes \tau^2)}{\mathrm{tr}(\rho^2 \otimes \tau^2) + \mathrm{tr}(\sigma^2 \otimes \tau^2)} \\ &= \frac{2\mathrm{tr}(\rho\,\sigma)\mathrm{tr}(\tau^2)}{[\mathrm{tr}(\rho^2) + \mathrm{tr}(\sigma^2)]\mathrm{tr}(\tau^2)} \\ &= \frac{2\mathrm{tr}(\rho\,\sigma)}{[\mathrm{tr}(\rho^2) + \mathrm{tr}(\sigma^2)]} \\ &= \mathcal{F}_{\mathrm{AM}}(\rho, \sigma),\end{aligned} \tag{A.16}$$

where equation (A.14) was used to establish the second equality and equation (A.15) was used to establish the third equality. □

A.4. Proofs of average fidelity properties

Here, we give the proofs showing when the respective fidelity measure satisfies equation (5.4). Throughout, as mentioned in sections 5.2 and 5.3, we will assume that the error state ρ_0 is orthogonal to *all* of the signal state, i.e.

$$\rho_j \rho_0 = \rho_0 \rho_j = 0 \quad \forall j \neq 0. \tag{A.17}$$

If, in addition to equation (A.17), all signal states are orthogonal to each other, i.e.

$$\rho_j \rho_k = 0 \quad \forall j \neq k, \tag{A.18}$$

then all ρ_j commute pairwise, and hence diagonalizable in the same basis. In this case, we will see that equation (5.4) holds for $\mathcal{F}_{\mathrm{I}}$, $\mathcal{F}_{\mathrm{Q}}$ and $\mathcal{F}_{\mathrm{A}}$ independent of the purity of the signal state ρ_j.

Proof. Recall from the main text that in the present discussion, the average state-by-state fidelity is

$$\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\sigma}) = \sum_{i \neq 0} p_i \mathcal{F}(\rho_i, \sigma_i). \tag{A.19}$$

Note, however, that equation (A.18) implies that for all $i \neq 0$,

$$\begin{aligned}\mathcal{F}_{\mathrm{I}}(\rho_i, \sigma_i) &= \left(\mathrm{tr}\sqrt{\sqrt{\rho_i}\sigma_i\sqrt{\rho_i}}\right)^2, \\ &= \left(\mathrm{tr}\sqrt{\sqrt{\rho_i}\left[\epsilon\rho_0 + (1-\epsilon)\rho_i\right]\sqrt{\rho_i}}\right)^2, \\ &= \left(\sqrt{1-\epsilon}\,\mathrm{tr}\sqrt{\sqrt{\rho_i}\rho_i\sqrt{\rho_i}}\right)^2, \\ &= (1-\epsilon)\,(\mathrm{tr}\rho_i)^2 = 1-\epsilon\end{aligned} \tag{A.20}$$

while

$$\begin{aligned}\mathcal{F}_{\mathrm{I}}(\rho, \sigma) &= \left(\mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2, \\ &= \left(\mathrm{tr}\sqrt{\sqrt{\rho}\left[\epsilon\rho_0 + (1-\epsilon)\rho\right]\sqrt{\rho}}\right)^2, \\ &= \left(\sqrt{1-\epsilon}\,\mathrm{tr}\sqrt{\sqrt{\rho}\,\rho\,\sqrt{\rho}}\right)^2, \\ &= (1-\epsilon)\left(\mathrm{tr}\sqrt{\rho^2}\right)^2 = 1-\epsilon.\end{aligned}$$

Substituting equation (A.20) into equation (A.19) and comparing the resulting expression with the above equation then verifies equation (5.4) for $\mathcal{F}_{\mathrm{I}}$.

Next, we shall prove the equivalence for $\mathcal{F}_{\mathrm{Q}}$. As with $\mathcal{F}_{\mathrm{I}}$, equation (A.18) implies that for all $i \neq 0$,

$$\begin{aligned}\mathcal{F}_{\mathrm{Q}}(\rho_i,\sigma_i) &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}(\rho_i^s\,\sigma_i^{1-s}),\\ &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}(\rho_i^s\,[\epsilon\rho_0+(1-\epsilon)\rho_i]^{1-s}),\\ &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}(\rho_i^s\,[\epsilon^{1-s}\rho_0^{1-s}+(1-\epsilon)^{1-s}\rho_i^{1-s}]),\\ &= \min_{0\leqslant s\leqslant 1}(1-\epsilon)^{1-s}\mathrm{tr}(\rho_i)=1-\epsilon, \end{aligned}\tag{A.21}$$

where the last equality follows from the fact that $0\leqslant 1-\epsilon\leqslant 1$, and thus $(1-\epsilon)^{1-s}\geqslant 1-\epsilon$ for all $0\leqslant s\leqslant 1$.

In a similar manner, the simultaneous diagonalizability of ρ and ρ_0 gives

$$\begin{aligned}\mathcal{F}_{\mathrm{Q}}(\rho,\sigma) &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}(\rho^s\,\sigma^{1-s}),\\ &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}\left\{\rho^s\,[\epsilon\rho_0+(1-\epsilon)\rho]^{1-s}\right\},\\ &= \min_{0\leqslant s\leqslant 1}\mathrm{tr}\left\{\rho^s\,[\epsilon^{1-s}\rho_0^{1-s}+(1-\epsilon)^{1-s}\rho^{1-s}]\right\},\\ &= \min_{0\leqslant s\leqslant 1}(1-\epsilon)^{1-s}\mathrm{tr}(\rho)=1-\epsilon. \end{aligned}\tag{A.22}$$

Substituting equation (A.21) into equation (A.19) and comparing with the last equation immediately leads to the verification of equation (5.4) for $\mathcal{F}_{\mathrm{Q}}$ whenever equation (A.18) holds.

To prove the equivalence for $\mathcal{F}_{\mathrm{A}}$, we note from equation (A.18) that for all $i\neq 0$,

$$\begin{aligned}\mathcal{F}_{\mathrm{A}}(\rho_i,\sigma_i) &= \left[\mathrm{tr}\left(\sqrt{\rho_i}\sqrt{\sigma_i}\right)\right]^2,\\ &= \left[\mathrm{tr}\left(\sqrt{\rho_i}\sqrt{\epsilon\rho_0+(1-\epsilon)\rho_i}\right)\right]^2,\\ &= \left\{\mathrm{tr}\left[\sqrt{\rho_i}\left(\sqrt{\epsilon}\sqrt{\rho_0}+\sqrt{1-\epsilon}\sqrt{\rho_i}\right)\right]\right\}^2,\\ &= \left[\sqrt{1-\epsilon}\,\mathrm{tr}\left(\rho_i\right)\right]^2=1-\epsilon. \end{aligned}\tag{A.23}$$

Similarly, equation (A.18) implies that

$$\begin{aligned}\mathcal{F}_{\mathrm{A}}(\rho,\sigma) &= \left[\mathrm{tr}\left(\sqrt{\rho}\sqrt{\sigma}\right)\right]^2,\\ &= \left[\mathrm{tr}\left(\sqrt{\rho}\sqrt{\epsilon\rho_0+(1-\epsilon)\rho}\right)\right]^2,\\ &= \left\{\mathrm{tr}\left[\sqrt{\rho}\left(\sqrt{\epsilon}\sqrt{\rho_0}+\sqrt{1-\epsilon}\sqrt{\rho}\right)\right]\right\}^2,\\ &= \left[\sqrt{1-\epsilon}\,\mathrm{tr}\left(\rho\right)\right]^2=1-\epsilon. \end{aligned}\tag{A.24}$$

Hence, by substituting equation (A.23) into equation (A.19) and comparing with the last equation immediately leads to the verification of equation (5.4) for $\mathcal{F}_{\mathrm{A}}$ whenever equation (A.18) holds. □

Notice that if instead of equation (A.18), we have the promise that

$$\mathrm{tr}(\rho_j^2)\geqslant\mathrm{tr}(\sigma_j^2),\quad \mathrm{tr}(\rho^2)\geqslant\mathrm{tr}(\sigma^2),\tag{A.25}$$

then

$$\begin{aligned}\mathcal{F}_2(\rho,\sigma) &= \frac{\mathrm{tr}(\rho\,\sigma)}{\max\left[\mathrm{tr}(\rho^2),\mathrm{tr}(\sigma^2)\right]},\\ &= \frac{\mathrm{tr}\left\{\rho\,[\epsilon\rho_0+(1-\epsilon)\rho]\right\}}{\mathrm{tr}(\rho^2)},\\ &= 1-\epsilon. \end{aligned}\tag{A.26}$$

Equation (5.7) then follows by combining equations (5.6), (A.19) and the last equation above.

Finally, note that if instead we have the premise that $\mathrm{tr}(\rho^2)=1$, then $\rho=\rho_j$, $\sigma=\sigma_j$, and there is only term in the sum of equation (A.19). Consequently, we must also have $\mathcal{F}_{\mathrm{N}}(\rho,\sigma)=\mathcal{F}_{\mathrm{N}}(\rho_j,\sigma_j)=\mathcal{F}_{\mathrm{ave}}(\boldsymbol{p},\boldsymbol{\rho},\boldsymbol{\sigma})$ in this case.

A.5. Counterexamples

We provide here some counterexamples that have been left out in the main text for showing certain desired properties of various fidelity measures.

To verify that $\mathcal{F}_{\mathrm{AM}}$ is generally not (super)multiplicative, it suffices to consider $\rho=\frac{1}{5}(\Pi_0+4\Pi_1)$ and $\sigma=\frac{1}{2}\mathbf{1}_2$. An explicit calculation gives $\mathcal{F}_{\mathrm{AM}}(\rho\otimes\rho,\sigma\otimes\sigma)\approx 0.702$ while $[\mathcal{F}_{\mathrm{AM}}(\rho,\sigma)]^2\approx 0.718>\mathcal{F}_{\mathrm{AM}}(\rho\otimes\rho,\sigma\otimes\sigma)$, thus showing a violation of the desired (super)multiplicative property.

To see that $\mathcal{F}_{\mathrm{C}}$, $\mathcal{F}_{\mathrm{GM}}$ and $\mathcal{F}_{\mathrm{AM}}$ can be contractive under the partial trace operation, it suffices to consider the following pair of two-qubit density matrices (written in the product basis):

$$\rho=\Pi_0\otimes\Pi_0,\quad \sigma=\frac{1}{8}\begin{pmatrix}3&0&0&\sqrt{3}\\0&4&0&0\\0&0&0&0\\\sqrt{3}&0&0&1\end{pmatrix},\tag{A.27}$$

and consider the partial trace of ρ and σ over the first qubit subsystem.

Appendix B. Metric properties

In this appendix, detailed proofs are given for the metric properties of fidelities.

B.1. General definitions

For a given fidelity measure $\mathcal{F}(\rho,\sigma)$, let us define the following functionals of $\mathcal{F}$:

$$\begin{aligned}A[\mathcal{F}(\rho,\sigma)] &:= \arccos[\sqrt{\mathcal{F}(\rho,\sigma)}],\\ B[\mathcal{F}(\rho,\sigma)] &:= \sqrt{2-2\sqrt{\mathcal{F}(\rho,\sigma)}},\\ C[\mathcal{F}(\rho,\sigma)] &:= \sqrt{1-\mathcal{F}(\rho,\sigma)}. \end{aligned}\tag{B.1}$$

In what follows, we will provide the proofs of the metric properties of these functionals of $\mathcal{F}$ for the various fidelity measures discussed in section 3.3. By a metric, we mean a mapping $\mathfrak{D}$ on a set S such that for every $a,b,c\in S$, the mapping $\mathfrak{D}:S\times S\to\mathbb{R}$ satisfies the following properties:

(M1) $\mathfrak{D}(a,b) \geqslant 0$ (Nonnegativity),
(M2) $\mathfrak{D}(a,b) = 0$ iff $a = b$ (Identity of Indiscernible),
(M3) $\mathfrak{D}(a,b) = \mathfrak{D}(b,a)$ (Symmetry),
(M4) $\mathfrak{D}(a,c) \leqslant \mathfrak{D}(a,b) + \mathfrak{D}(b,c)$ (Triangle Inequality).

Our main tool is a simplified version of Schoenberg's theorem [152], reproduced as follows.

Theorem B.1 (Schoenberg). *Let $\mathcal{X}$ be a nonempty set and $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ a function such that $K(x,y) = K(y,x)$ and $K(x,y) \geqslant 0$ for all $x, y \in \mathcal{X}$, with saturation iff $x = y$. If the implication*

$$\sum_{i=1}^{n} c_i = 0 \Rightarrow \sum_{i,j=1}^{n} K(x_i, x_j) c_i c_j \leqslant 0 \tag{B.2}$$

holds for all $n \geqslant 2$, $\{x_1, \ldots, x_n\} \subseteq \mathcal{X}$ and $\{c_1, \ldots, c_n\} \subseteq \mathbb{R}$, then $\sqrt{K}$ is a metric.

The theorem has previously been used in [21] to show that $C[\mathcal{F}_{\mathrm{N}}]$ is a metric for the space of density matrices. However, the alternative proof for the metric properties of $B[\mathcal{F}_1(\rho,\sigma)]$ and $C[\mathcal{F}_1(\rho,\sigma)]$ given in [21] is flawed due to an erroneous application of the above theorem.

B.2. $\mathcal{F}_2$ metric properties

Here, we will show that $C[\mathcal{F}_2(\rho,\sigma)]$ is a metric for the space of density matrices. Because $\mathcal{F}_2$ complies with Jozsa's axioms (J1) and (J2), it is immediate that $C[\mathcal{F}_2]$ is non-negative (M1), fulfills the indiscernible identity (M2) and is symmetric (M3). Hence, in order to establish $C[\mathcal{F}_2]$ as a genuine metric, we only have to prove that it satisfies the triangle inequality

$$C[\mathcal{F}_2(\rho,\sigma)] \leqslant C[\mathcal{F}_2(\rho,\tau)] + C[\mathcal{F}_2(\tau,\sigma)] \tag{B.3}$$

for *arbitrary* d-dimensional density matrices ρ, σ and τ. To this end, we shall first prove the following lemma:

Lemma B.1. *For any $\vartheta, \varphi \in [0, 2\pi]$ and $p, q \in [0,1]$*

$$\begin{aligned}\sqrt{1 - pq\cos(\vartheta + \varphi)} &\leqslant \sqrt{1 - p\cos\vartheta} + \sqrt{1 - q\cos\varphi},\\ \sqrt{1 - pq\cos(\vartheta - \varphi)} &\geqslant \sqrt{1 - p\cos\vartheta} - \sqrt{1 - q\cos\varphi}\end{aligned} \tag{B.4}$$

Proof. We start by showing that both inequalities hold in the special case of $p = q = 1$. To see this, note that

$$\begin{aligned}\sqrt{1 - \cos(\vartheta + \varphi)} &= \sqrt{2}\left|\sin\frac{\vartheta}{2}\cos\frac{\varphi}{2} + \sin\frac{\varphi}{2}\cos\frac{\vartheta}{2}\right|\\ &\leqslant \sqrt{2}\left|\sin\frac{\vartheta}{2}\cos\frac{\varphi}{2}\right| + \sqrt{2}\left|\sin\frac{\varphi}{2}\cos\frac{\vartheta}{2}\right|\\ &\leqslant \sqrt{2}\left|\sin\frac{\vartheta}{2}\right| + \sqrt{2}\left|\sin\frac{\varphi}{2}\right|\\ &= \sqrt{1 - \cos\vartheta} + \sqrt{1 - \cos\varphi}\end{aligned} \tag{B.5}$$

where we have used the trigonometric identity $\sqrt{1 - \cos x} = \sqrt{2}|\sin\frac{x}{2}|$ in the first and last lines, the inequality $|x + y| \leqslant |x| + |y|$ in the second line, and the fact that cosines are upper bounded by 1 in the third line.

To see that the second inequality of (B.4) also holds for $p = q = 1$, we rewrite it in the equivalent form

$$\left|\sin\left(\frac{\vartheta - \varphi}{2}\right)\right| \geqslant \sin\left(\frac{\vartheta}{2}\right) - \sin\left(\frac{\varphi}{2}\right), \tag{B.6}$$

which is clearly valid if $\sin\frac{\vartheta}{2} \leqslant \sin\frac{\varphi}{2}$. If instead $\sin\frac{\vartheta}{2} > \sin\frac{\varphi}{2}$, then both sides of inequality (B.6) are non-negative and can thus be squared to yield an equivalent inequality that can be simplified to

$$4\sin\left(\frac{\vartheta}{2}\right)\sin\left(\frac{\varphi}{2}\right)\sin^2\left(\frac{\vartheta - \varphi}{4}\right) \geqslant 0. \tag{B.7}$$

Since this clearly holds for $\vartheta, \varphi \in [0, 2\pi]$, inequality (B.4) for $p = q = 1$ must also hold.

In order to show Lemma B.1 for general $p, q \in [0,1]$, let us *define* the angles $\Theta, \Phi \in [0, \pi]$ as follows

$$\begin{aligned}\cos\Theta &= p\cos\vartheta, \quad \sin\Theta = \sqrt{1 - p^2\cos^2\vartheta},\\ \cos\Phi &= q\cos\varphi, \quad \sin\Phi = \sqrt{1 - q^2\cos^2\varphi}.\end{aligned} \tag{B.8}$$

This gives

$$\begin{aligned}&\pm\cos(\Theta \pm \Phi)\\ &= \pm\left(pq\cos\vartheta\cos\varphi \mp \sqrt{1 - p^2\cos^2\vartheta}\sqrt{1 - q^2\cos^2\varphi}\right)\\ &\leqslant \pm\left(pq\cos\vartheta\cos\varphi \mp pq\sqrt{1 - \cos^2\vartheta}\sqrt{1 - \cos^2\varphi}\right)\\ &= \pm pq(\cos\vartheta\cos\varphi \mp |\sin\vartheta\sin\varphi|)\\ &\leqslant \pm pq\cos(\vartheta \pm \varphi).\end{aligned} \tag{B.9}$$

Inequalities (B.4) can now be obtained from inequality (B.9) as follows:

$$\begin{aligned}\pm\sqrt{1 - pq\cos(\vartheta \pm \varphi)} &\leqslant \pm\sqrt{1 - \cos(\Theta \pm \Phi)}\\ &\leqslant \pm\left(\sqrt{1 - \cos\Theta} \pm \sqrt{1 - \cos\Phi}\right)\\ &= \pm\left(\sqrt{1 - p\cos\vartheta} \pm \sqrt{1 - q\cos\varphi}\right),\end{aligned}$$

where the second inequality follows from the already verified inequality of (B.4) in the case of $p = q = 1$. □

Theorem B.2. *$C[\mathcal{F}_2(\rho,\sigma)]$ is a metric for the space of density matrices*

Proof. Using equations (4.6) and (4.7), the triangle inequality of (B.3) can be rewritten as

$$\sqrt{1 - \frac{\vec{r}\cdot\vec{s}}{\mu(\vec{r},\vec{s})}} \leqslant \sqrt{1 - \frac{\vec{r}\cdot\vec{t}}{\mu(\vec{r},\vec{t})}} + \sqrt{1 - \frac{\vec{t}\cdot\vec{s}}{\mu(\vec{t},\vec{s})}}, \tag{B.10}$$

where $\mu(\vec{r},\vec{s}) \equiv \max(r^2, s^2)$ and the entries of the vectors $\vec{r}, \vec{s}, \vec{t} \in \mathbb{R}^{d^2}$ are the expansion coefficients of ρ, σ, and τ, respectively, in some basis of Hermitian matrices.

Without loss of generality, we henceforth assume that ρ and τ are, respectively, the density matrices of maximal and minimal purities (i.e. $r \geqslant s \geqslant t$). Hence, (B.3) unfolds into three inequalities to be proven, which correspond to $C[\mathcal{F}_2(\rho,\sigma)] \leqslant C[\mathcal{F}_2(\rho,\tau)] + C[\mathcal{F}_2(\tau,\sigma)]$, $C[\mathcal{F}_2(\sigma,\tau)] \leqslant C[\mathcal{F}_2(\rho,\sigma)] + C[\mathcal{F}_2(\rho,\tau)]$ and $C[\mathcal{F}_2(\rho,\tau)] \leqslant C[\mathcal{F}_2(\rho,\sigma)] + C[\mathcal{F}_2(\tau,\sigma)]$ respectively. In terms of the expansion coefficients vectors $\vec{r}$, $\vec{s}$, and $\vec{t}$, these be written as:

$$\begin{aligned}
\sqrt{1-\frac{s}{r}\cos\theta_{rs}} &\leqslant \sqrt{1-\frac{t}{r}\cos\theta_{rt}} + \sqrt{1-\frac{t}{s}\cos\theta_{ts}},\\
\sqrt{1-\frac{t}{s}\cos\theta_{ts}} &\leqslant \sqrt{1-\frac{t}{r}\cos\theta_{rt}} + \sqrt{1-\frac{s}{r}\cos\theta_{rs}},\\
\sqrt{1-\frac{t}{r}\cos\theta_{rt}} &\leqslant \sqrt{1-\frac{s}{r}\cos\theta_{rs}} + \sqrt{1-\frac{t}{s}\cos\theta_{ts}},
\end{aligned} \tag{B.11}$$

where the angles θ_{rs}, θ_{rt}, and $\theta_{ts} \in [0,\pi]$ have been *defined* (in accordance with the Cauchy–Schwarz inequality) as follows:

$$\cos\theta_{rs} = \frac{\vec{r}\cdot\vec{s}}{rs}, \quad \cos\theta_{rt} = \frac{\vec{r}\cdot\vec{t}}{rt}, \quad \cos\theta_{ts} = \frac{\vec{t}\cdot\vec{s}}{ts}. \tag{B.12}$$

Besides, since

$$\cos(\theta_{rs}+\theta_{ts}) \leqslant \cos\theta_{rt} \leqslant \cos(\theta_{rs}-\theta_{ts}), \tag{B.13}$$

(a demonstration of which will be given at the end of this proof), the replacement of $\cos\theta_{rt}$ with either $\cos(\theta_{rs}-\theta_{ts})$ or $\cos(\theta_{rs}+\theta_{ts})$ in inequality (B.11), yields the following *stronger* set of inequalities to be proven, where we define $\Delta^{\pm} = \theta_{rs} \pm \theta_{ts}$:

$$\begin{aligned}
\sqrt{1-\frac{t}{r}\cos(\Delta^-)} &\geqslant \sqrt{1-\frac{s}{r}\cos\theta_{rs}} - \sqrt{1-\frac{t}{s}\cos\theta_{ts}}\\
\sqrt{1-\frac{t}{r}\cos(\Delta^-)} &\geqslant \sqrt{1-\frac{t}{s}\cos\theta_{ts}} - \sqrt{1-\frac{s}{r}\cos\theta_{rs}}\\
\sqrt{1-\frac{t}{r}\cos(\Delta^+)} &\leqslant \sqrt{1-\frac{s}{r}\cos\theta_{rs}} + \sqrt{1-\frac{t}{s}\cos\theta_{ts}}.
\end{aligned} \tag{B.14}$$

Note that $\frac{t}{r} = \frac{s}{r}\cdot\frac{t}{s}$ and $\frac{t}{r}, \frac{s}{r}, \frac{t}{s} \in (0,1]$ by our assumption that $r \geqslant s \geqslant t > 0$. Thus, through appropriate identifications of these ratios with p, q and applying lemma B.1, inequalities (B.14), and hence the desired triangle inequalities (B.10) follow.

We complete this proof with a demonstration of inequalities (B.13). Consider, first, the following pair of vectors, each of which being orthogonal to $\vec{s}$

$$\vec{u} = \vec{r} - \frac{r}{s}\cos\theta_{rs}\vec{s} \quad \text{and} \quad \vec{v} = \vec{t} - \frac{t}{s}\cos\theta_{ts}\vec{s}. \tag{B.15}$$

Using these and the orthogonality of $\vec{u}, \vec{v}$ to $\vec{s}$, we may expand the scalar product between $\vec{r}$ and $\vec{t}$ as

$$\vec{r}\cdot\vec{t} = rt\cos\theta_{rt} = \vec{u}\cdot\vec{v} + rt\cos\theta_{rs}\cos\theta_{ts}. \tag{B.16}$$

Then, using $\vec{u}\cdot\vec{v} = uv\cos\theta_{uv}$ with $u = r\sin\theta_{rs}$, $v = t\sin\theta_{ts}$, and $\theta_{uv} \in [0,\pi]$, we arrive at

$$\cos\theta_{rt} = \sin\theta_{rs}\sin\theta_{ts}\cos\theta_{uv} + \cos\theta_{rs}\cos\theta_{ts}. \tag{B.17}$$

From here, the trivial inequalities $\cos\theta_{uv} \geqslant -1$ and $\cos\theta_{uv} \leqslant 1$ imply, respectively, the upper and lower bounds on $\cos\theta_{rt}$ in (B.13). □

B.3. $\mathcal{F}_{\mathrm{C}}$ metric properties

Theorem B.3. *$C[\mathcal{F}_{\mathrm{C}}(\rho,\sigma)]$ is a metric for the space of density matrices of fixed Hilbert space dimension.*

Proof. For density matrices ρ_i acting on d-dimensional complex Hilbert space $\mathbb{C}^d$, note that $r = \frac{1}{d-1}$ is a constant. Thus, for any $c_i's \in \mathbb{R}$ and which are such that $\sum_i c_i = 0$, we have

$$\begin{aligned}
&\sum_{i,j} c_i c_j \left[1 - \mathcal{F}_{\mathrm{C}}(\rho_i,\rho_j)\right]\\
&= -\sum_{i,j} c_i c_j \left[\frac{1-r}{2} + \frac{1+r}{2}\mathcal{F}_{\mathrm{N}}(\rho_i,\rho_j)\right]\\
&= -\frac{1+r}{2}\sum_{i,j} c_i c_j \mathcal{F}_{\mathrm{N}}(\rho_i,\rho_j).
\end{aligned} \tag{B.18}$$

From the proof of the metric property of $C[\mathcal{F}_{\mathrm{N}}(\rho,\sigma)]$ given in [21], we know that the last expression above must be non-positive. Hence, $\sqrt{1-\mathcal{F}_{\mathrm{C}}(\rho,\sigma)}$ is a metric for the space of density matrices of fixed dimension. □

B.4. $\mathcal{F}_{\mathrm{GM}}$ metric properties

Theorem B.4. *$C[\mathcal{F}_{\mathrm{GM}}(\rho,\sigma)]$ is a metric for the space of density matrices.*

Proof. For any $c_i's \in \mathbb{R}$ and which are such that $\sum_i c_i = 0$, note that

$$\begin{aligned}
\sum_{i,j} c_i c_j \left[1 - \mathcal{F}_{\mathrm{GM}}(\rho_i,\rho_j)\right] &= -\sum_{i,j} c_i c_j \frac{\mathrm{tr}(\rho_i\rho_j)}{\sqrt{\mathrm{tr}(\rho_i^2)\mathrm{tr}(\rho_j^2)}}\\
&= -\mathrm{tr}\left[\left(\sum_i \frac{c_i\rho_i}{\sqrt{\mathrm{tr}(\rho_i^2)}}\right)^2\right] \leqslant 0.
\end{aligned} \tag{B.19}$$

Hence, $C[\mathcal{F}_{\mathrm{GM}}(\rho,\sigma)] = \sqrt{1-\mathcal{F}_{\mathrm{GM}}(\rho,\sigma)}$ is a metric for the space of density matrices. □

B.5. $\mathcal{F}_{\mathrm{A}}$ metric properties

The metric properties of $B[\mathcal{F}_{\mathrm{A}}(\rho,\sigma)]$ were first mentioned in [37], while those of $A[\mathcal{F}_{\mathrm{A}}(\rho,\sigma)]$ and $C[\mathcal{F}_{\mathrm{A}}(\rho,\sigma)]$ were numerically investigated in [36], suggesting that they may indeed be metrics for the space of density matrices. In section C.3, [21], it was briefly mentioned that both $B[\mathcal{F}_{\mathrm{A}}(\rho,\sigma)]$ and

$C[\mathcal{F}_{\rm A}(\rho,\sigma)]$ can be proved to be a metric using Schoenberg's theorem. Here, we shall provide an explicit proof of these facts. An alternative proof for $B[\mathcal{F}_{\rm A}(\rho,\sigma)]$ can also be found in [153].

Proof. For any $c_i's \in \mathbb{R}$ and which are such that $\sum_i c_i = 0$, note that

$$\sum_{i,j} c_i c_j \left[1 - \sqrt{\mathcal{F}_{\rm A}(\rho_i,\rho_j)}\right] = -\sum_{i,j} c_i c_j \mathrm{tr}\left(\sqrt{\rho_i}\sqrt{\rho_j}\right)$$
$$= -\mathrm{tr}\left[\left|\sum_i c_i \sqrt{\rho_i}\right|^2\right] \leqslant 0. \tag{B.20}$$

Likewise, we can show that:

$$\sum_{i,j} c_i c_j \left[1 - \mathcal{F}_{\rm A}(\rho_i,\rho_j)\right] = -\sum_{i,j} c_i c_j \left[\mathrm{tr}\left(\sqrt{\rho_i}\sqrt{\rho_j}\right)\right]^2$$
$$= -\sum_{i,j} c_i c_j \mathrm{tr}\left[\left(\sqrt{\rho_i}\otimes\sqrt{\rho_i}\right)\left(\sqrt{\rho_j}\otimes\sqrt{\rho_j}\right)\right]$$
$$= -\mathrm{tr}\left[\left|\sum_i c_i \sqrt{\rho_i}\otimes\sqrt{\rho_i}\right|^2\right] \leqslant 0. \tag{B.21}$$

This concludes our proof for the metric properties of $B[\mathcal{F}_{\rm A}(\rho,\sigma)]$ and $C[\mathcal{F}_{\rm A}(\rho,\sigma)]$. □

ORCID iDs

Yeong-Cherng Liang https://orcid.org/0000-0002-2899-5842
Yu-Hao Yeh https://orcid.org/0000-0003-0523-8758
Paulo E M F Mendonça https://orcid.org/0000-0003-0975-6915
Run Yan Teh https://orcid.org/0000-0001-8328-0711
Margaret D Reid https://orcid.org/0000-0002-4101-0921
Peter D Drummond https://orcid.org/0000-0003-4763-8549

References

[1] Gühne O and Tóth G 2009 Entanglement detection *Phys. Rep.* **474** 1–75
[2] Rosset D, Ferretti-Schöbitz R, Bancal J-D, Gisin N and Liang Y-C 2012 Imperfect measurement settings: implications for quantum state tomography and entanglement witnesses *Phys. Rev.* A **86** 062325
[3] Gu S-J 2010 Fidelity approach to quantum phase transitions *Int. J. Mod. Phys.* B **24** 4371–458
[4] Braun D, Adesso G, Benatti F, Floreanini R, Marzolino U, Mitchell M W and Pirandola S 2018 Quantum-enhanced measurements without entanglement *Rev. Mod. Phys.* **90** 35006
[5] Hayden P, Leung D W and Winter A 2006 Aspects of generic entanglement *Commun. Math. Phys.* **265** 95–117
[6] Wootters W K and Zurek W H 1982 A single quantum cannot be cloned *Nature* **299** 802–3
[7] Buzek V and Hillery M 1996 Quantum copying: beyond the no-cloning theorem *Phys. Rev.* A **54** 1844–52
[8] Scarani V, Iblisdir S, Gisin N and Acín A 2005 Quantum cloning *Rev. Mod. Phys.* **77** 1225–56
[9] Jozsa R 1994 Fidelity for mixed quantum states *J. Mod. Opt.* **41** 2315–23
[10] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels *Phys. Rev. Lett.* **70** 1895
[11] Lvovsky A I, Sanders B C and Tittel W 2009 Optical quantum memory *Nat. Photon.* **3** 706
[12] Nielsen M A and Chuang I L 2000 *Quantum Computation and Information* (Cambridge: Cambridge University Press)
[13] Greenberger D M, Horne M A and Zeilinger A 1989 *Bell's Theorem, Quantum Theory, and Conceptions of the Universe* (Dordrecht: Kluwer) pp 69–72
[14] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish M W A, Hänsel W, Hennrich M and Blatt R 2011 14-qubit entanglement: creation and coherence *Phys. Rev. Lett.* **106** 130506
[15] Song C *et al* 2017 10-qubit entanglement and parallel logic operations with a superconducting circuit *Phys. Rev. Lett.* **119** 180511
[16] Chen L-K *et al* 2017 Observation of ten-photon entanglement using thin BiB_3O_6 crystals *Optica* **4** 77–83
[17] Reid M D, Opanchuk B, Rosales-Zárate L and Drummond P D 2014 Quantum probabilistic sampling of multipartite 60-qubit Bell-inequality violations *Phys. Rev.* A **90** 012111
[18] Galve F, Mandarino A, Paris M G A, Benedetti C and Zambrini R 2017 Microscopic description for the emergence of collective dissipation in extended quantum systems *Sci. Rep.* **7** 42050
[19] Leibfried D *et al* 2003 Experimental demonstration of a robust, high-fidelity geometric two ion-qubit phase gate *Nature* **422** 412
[20] Schumacher B 1995 Quantum coding *Phys. Rev.* A **51** 2738
[21] Mendonca P E M F, Napolitano R D J, Marchiolli M A, Foster C J and Liang Y-C 2008 Alternative fidelity measure between quantum states *Phys. Rev.* A **78** 052330
[22] Caves C M and Drummond P D 1994 Quantum limits on bosonic communication rates *Rev. Mod. Phys.* **66** 481
[23] Gilchrist A, Langford N K and Nielsen M A 2005 Distance measures to compare real and ideal quantum processes *Phys. Rev.* A **71** 062310
[24] Uhlmann A 1976 The 'transition probability' in the state space of a $*$-algebra *Rep. Math. Phys.* **9** 273–9
[25] Luo S and Zhang Q 2004 Informational distance on quantum-state space *Phys. Rev.* A **69** 032106
[26] Audenaert K M R, Nussbaum M, Szkoła A and Verstraete F 2008 Asymptotic error rates in quantum hypothesis testing *Commun. Math. Phys.* **279** 251–83
[27] Hu X and Ye Z 2006 Generalized quantum entropy *J. Math. Phys.* **47** 023502
[28] Müller-Lennert M, Dupuis F, Szehr O, Fehr S and Tomamichel M 2013 On quantum Rényi entropies: a new generalization and some properties *J. Math. Phys.* **54** 122203
[29] Dupuis F, Krämer L, Faist P, Renes J M and Renner R 2013 Generalized entropies *17th Int. Congress on Mathematical Physics* (Singapore: World Scientific) pp 134–53
[30] Chen J-L, Fu L, Ungar A A and Zhao X 2002 Alternative fidelity measure between two states of an n-state quantum system *Phys. Rev.* A **65** 054304
[31] Chen J-L, Fu L, Ungar A A and Zhao X-G 2002 Geometric observation for bures fidelity between two states of a qubit *Phys. Rev.* A **65** 024303
[32] Chen J-L and Ungar A A 2002 The bloch gyrovector *Found. Phys.* **32** 531–65
[33] Kimura G 2003 The bloch vector for n-level systems *Phys. Lett.* A **314** 339–49
[34] Byrd M S and Khaneja N 2003 Characterization of the positivity of the density matrix in terms of the coherence vector representation *Phys. Rev.* A **68** 062322

[35] Miszczak J A, Puchała Z, Horodecki P, Uhlmann A and Zyczkowski K 2009 Sub-and super-fidelity as bounds for quantum fidelity *Quantum Inf. Comput.* **9** 103–30

[36] Ma Z, Zhang F-L and Chen J-L 2008 Geometric interpretation for the a fidelity and its relation with the Bures fidelity *Phys. Rev.* A **78** 064305

[37] Raggio G A 1984 Generalized transition probabilities and applications *Quantum Probability and Applications to the Quantum Theory of Irreversible Processes* (New York: Springer) pp 327–35

[38] Wang X, Yu C-S and Yi X X 2008 An alternative quantum fidelity for mixed states of qudits *Phys. Lett.* A **373** 58–60

[39] Audenaert K M R, Calsamiglia J, Munoz-Tapia R, Bagan E, Masanes A L L and Verstraete F 2007 Discriminating states: the quantum Chernoff bound *Phys. Rev. Lett.* **98** 160501

[40] Bhatia R 1997 *Matrix Analysis* vol 169 (New York: Springer)

[41] Rényi A 1961 On measures of entropy and information *Technical Report* Hungarian Academy of Science Budapest Hungary

[42] Hastings M B, González I, Kallin A B and Melko R G 2010 Measuring Renyi entanglement entropy in quantum monte carlo simulations *Phys. Rev. Lett.* **104** 157201

[43] Brydges T, Elben A, Jurcevic P, Vermersch B, Maier C, Lanyon P B P, Blatt R and Roos C F 2018 Probing entanglement entropy via randomized measurements *Science* **364** 260

[44] Elben A, Vermersch B, Dalmonte M, Cirac J I and Zoller P 2018 Rényi entropies from random quenches in atomic hubbard and spin models *Phys. Rev. Lett.* **120** 050406

[45] Elben A, Vermersch B, Roos C F and Zoller P 2019 Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states *Phys. Rev.* A **99** 052323

[46] Preskill J 2015 *Lecture Notes for Physics 229: Quantum Information and Computation* (CreateSpace Independent Publishing Platform)

[47] Carlen E A and Lieb E H 2008 A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity *Lett. Math. Phys.* **83** 107–26

[48] Hübner M 1992 Explicit computation of the Bures distance for density matrices *Phys. Lett.* A **163** 239–42

[49] Rastegin A E 2006 Sine distance for quantum states (arXiv:quant-ph/0602112)

[50] Ma Z, Zhang F-L and Chen J-L 2009 Fidelity induced distance measures for quantum states *Phys. Lett.* A **373** 3407–9

[51] Fuchs C A and van de Graaf J 1999 Cryptographic distinguishability measures for quantum-mechanical states *IEEE Trans. Inf. Theory* **45** 1216

[52] Zhang L, Bu K and Wu J 2015 A lower bound on the fidelity between two states in terms of their trace-distance and max-relative entropy *Linear Multilinear* A **64** 801–6

[53] Raggio G A 1982 Comparison of Uhlmann's transition probability with the one induced by the natural positive cone of von Neumann algebras in standard form *Lett. Math. Phys.* **6** 233–6

[54] Ginibre J 1965 Statistical ensembles of complex, quaternion, and real matrices *J. Math. Phys.* **6** 440–9

[55] Życzkowski K, Penson K A, Nechita I and Collins B 2011 Generating random density matrices *J. Math. Phys.* **52** 062201

[56] Życzkowski K and Sommers H-J 2005 Average fidelity between random quantum states *Phys. Rev.* A **71** 032313

[57] Caves C M 1981 Quantum-mechanical noise in an interferometer *Phys. Rev.* D **23** 1693

[58] Walls D F 1983 Squeezed states of light *Nature* **306** 141–6

[59] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 777–80

[60] Wiseman H M, Jones S J and Doherty A C 2007 Steering, entanglement, nonlocality, and the Einstein–Podolsky–Rosen paradox *Phys. Rev. Lett.* **98** 140402

[61] Reid M D, Drummond P D, Bowen W P, Cavalcanti E G, Lam P K, Bachor U H A and Leuchs G 2009 Colloquium: the Einstein–Podolsky–Rosen paradox: from concepts to applications *Rev. Mod. Phys.* **81** 1727

[62] Ma Y, Miao H, Pang B H, Evans M, Zhao C, Harms J, Schnabel R and Chen Y 2017 Proposal for gravitational-wave detection beyond the standard quantum limit through EPR entanglement *Nat. Phys.* **13** 776

[63] Abbott B P *et al* 2016 Observation of gravitational waves from a binary black hole merger *Phys. Rev. Lett.* **116** 061102

[64] Einstein A 1918 Über gravitationswellen *Sitzungsber. K. Preuss. Akad. Wiss.* **1** 154

[65] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 Quantum cryptography *Rev. Mod. Phys.* **74** 145

[66] Steane A 1998 Quantum computing *Rep. Prog. Phys.* **61** 117

[67] Fialko O, Opanchuk B, Sidorov A, Drummond P D and Brand J 2015 Fate of the false vacuum: towards realization with ultra-cold atoms *Europhys. Lett.* **110** 56001

[68] Bernien H *et al* 2017 Probing many-body dynamics on a 51-atom quantum simulator *Nature* **551** 579

[69] Schumacher B and Nielsen M A 1996 Quantum data processing and error correction *Phys. Rev.* A **54** 2629

[70] Millen J and Xuereb A 2016 Perspective on quantum thermodynamics *New. J. Phys.* **18** 011002

[71] Yang T H, Vértesi T, Bancal J-D, Scarani V and Navascués M 2014 Robust and versatile black-box certification of quantum devices *Phys. Rev. Lett.* **113** 040401

[72] Kaniewski J 2016 Analytic and nearly optimal self-testing bounds for the Clauser–Horne–Shimony–Holt and Mermin inequalities *Phys. Rev. Lett.* **117** 070402

[73] Sekatski P, Bancal J-D, Wagner S and Sangouard N 2018 Certifying the building blocks of quantum computers from bell's theorem *Phys. Rev. Lett.* **121** 180505

[74] Chen S-L, Budroni C, Liang Y-C and Chen Y-N 2016 Natural framework for device-independent quantification of quantum steerability, measurement incompatibility, and self-testing *Phys. Rev. Lett.* **116** 240401

[75] Cavalcanti D and Skrzypczyk P 2016 Quantitative relations between measurement incompatibility, quantum steering, and nonlocality *Phys. Rev.* A **93** 052112

[76] Rosset D, Buscemi F and Liang Y-C 2018 Resource theory of quantum memories and their faithful verification with minimal assumptions *Phys. Rev.* X **8** 021033

[77] Knill E, Leibfried D, Reichle R, Britton J, Blakestad R B, Jost C J D, Ozeri R, Seidelin S and Wineland D J 2008 Randomized benchmarking of quantum gates *Phys. Rev.* A **77** 012307

[78] Harty T P, Allcock D T C, Ballance C J, Guidoni L, Janacek H A, Linke N M, Stacey D N and Lucas D M 2014 High-fidelity preparation, gates, memory, and readout of a trapped-ion quantum bit *Phys. Rev. Lett.* **113** 220501

[79] Horodecki M, Horodecki P and Horodecki R 1999 General teleportation channel, singlet fraction, and quasidistillation *Phys. Rev.* A **60** 1888–98

[80] Kimble H J 2008 The quantum internet *Nature* **453** 1023

[81] Popescu S 1994 Bell's inequalities versus teleportation: what is nonlocality? *Phys. Rev. Lett.* **72** 797–9

[82] Gisin N and Massar S 1997 Optimal quantum cloning machines *Phys. Rev. Lett.* **79** 2153

[83] Bruss D, Ekert A and Macchiavello C 1998 Optimal universal quantum cloning and state estimation *Phys. Rev. Lett.* **81** 2598–601

[84] Hammerer K, Wolf M M, Polzik E S and Cirac J I 2005 Quantum benchmark for storage and transmission of coherent states *Phys. Rev. Lett.* **94** 150503

[85] Massar S and Popescu S 2005 Optimal extraction of information from finite quantum ensembles *Asymptotic Theory of Quantum Statistical Inference: Selected Papers* (Singapore: World Scientific) pp 356–64

[86] Bužek V and Hillery M 1998 Universal optimal cloning of arbitrary quantum states: from qubits to quantum registers *Phys. Rev. Lett.* **81** 5003

[87] Bruß D, DiVincenzo D P, Ekert A, Fuchs C A, Macchiavello C and Smolin J A 1998 Optimal universal and state-dependent quantum cloning *Phys. Rev.* A **57** 2368

[88] Bouwmeester D, Pan J W, Mattle K, Eibl M, Weinfurter H and Zeilinger A 1997 Experimental quantum teleportation *Nature* **390** 575

[89] Boschi D, Branca S, De Martini F, Hardy L and Popescu S 1998 Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein–Podolsky–Rosen channels *Phys. Rev. Lett.* **80** 1121

[90] Wang X-L, Cai X-D, Su Z-E, Chen M-C, Wu D, Li L, Liu N-L, Lu C-Y and Pan J-W 2015 Quantum teleportation of multiple degrees of freedom of a single photon *Nature* **518** 516

[91] Vaidman L 1994 Teleportation of quantum states *Phys. Rev.* A **49** 1473

[92] Braunstein S L and Kimble H J 1998 Teleportation of continuous quantum variables *Quantum Information with Continuous Variables* (New York: Springer) pp 67–75

[93] Reid M D 1989 Demonstration of the Einstein–Podolsky–Rosen paradox using nondegenerate parametric amplification *Phys. Rev.* A **40** 913

[94] Furusawa A, Sørensen J L, Braunstein S L, Fuchs C A, Kimble H J and Polzik E S 1998 Unconditional quantum teleportation *Science* **282** 706–9

[95] Bowen W P, Treps N, Buchler B C, Schnabel R, Ralph T C, Bachor T H-A and Lam P K 2003 Experimental investigation of continuous-variable quantum teleportation *Phys. Rev.* A **67** 032302

[96] Zhang T C, Goh K W, Chou C W, Lodahl P and Kimble H J 2003 Quantum teleportation of light beams *Phys. Rev.* A **67** 033802

[97] Takei N, Yonezawa H, Aoki T and Furusawa A 2005 High-fidelity teleportation beyond the no-cloning limit and entanglement swapping for continuous variables *Phys. Rev. Lett.* **94** 220502

[98] Grosshans F and Grangier P 2001 Quantum cloning and teleportation criteria for continuous quantum variables *Phys. Rev.* A **64** 010301

[99] Cerf N J, Ipe A and Rottenberg X 2000 Cloning of continuous quantum variables *Phys. Rev. Lett.* **85** 1754

[100] Verstraete F and Verschelde H 2003 Optimal teleportation with a mixed state of two qubits *Phys. Rev. Lett.* **90** 097901

[101] He Q, Rosales-Zárate L, Adesso G and Reid M D 2015 Secure continuous variable teleportation and Einstein–Podolsky–Rosen steering *Phys. Rev. Lett.* **115** 180502

[102] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 Bell nonlocality *Rev. Mod. Phys.* **86** 419–78

[103] Hsieh C-Y, Liang Y-C and Lee R-K 2016 Quantum steerability: characterization, quantification, superactivation, and unbounded amplification *Phys. Rev.* A **94** 062120

[104] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 Distribution functions in physics: fundamentals *Phys. Rep.* **106** 121–67

[105] Wigner E 1932 On the quantum correction for thermodynamic equilibrium *Phys. Rev.* **40** 749–59

[106] Moyal J E 1949 Quantum mechanics as a statistical theory *Math. Proc. Camb. Phil. Soc.* **45** 99–124

[107] Lvovsky A I and Raymer M G 2009 Continuous-variable optical quantum-state tomography *Rev. Mod. Phys.* **81** 299

[108] Husimi K 1940 Some formal properties of the density matrix *Proc. Phys. Math. Soc. Japan.* 22 264–314

[109] Glauber R J 1963 Coherent and incoherent states of the radiation field *Phys. Rev.* **131** 2766–88

[110] Drummond P D and Gardiner C W 1980 Generalised P-representations in quantum optics *J. Phys. A: Math. Gen.* **13** 2353

[111] Gilchrist A, Gardiner C W and Drummond P D 1997 Positive P representation: application and validity *Phys. Rev.* A **55** 3014–32

[112] Deuar P and Drummond P D 2002 Gauge *P* representations for quantum-dynamical problems: removal of boundary terms *Phys. Rev.* A **66** 033812

[113] Drummond P D, McNeil K J and Walls D F 1981 Non-equilibrium transitions in sub/second harmonic generation *Optica Acta* **28** 211–25

[114] Drummond P D, Gardiner C W and Walls D F 1981 Quasiprobability methods for nonlinear chemical and optical systems *Phys. Rev.* A **24** 914–26

[115] Drummond P D and Hardman A D 1993 Simulation of quantum effects in Raman-active waveguides *Europhys. Lett.* **21** 279–84

[116] Kheruntsyan K V, Olsen M K and Drummond P D 2005 Einstein–Podolsky–Rosen correlations via dissociation of a molecular Bose–Einstein condensate *Phys. Rev. Lett.* **95** 150405

[117] Opanchuk B, He Q Y, Reid M D and Drummond P D 2012 Dynamical preparation of Einstein–Podolsky–Rosen entanglement in two-well Bose–Einstein condensates *Phys. Rev.* A **86** 023625

[118] Opanchuk B and Drummond P D 2013 Functional Wigner representation of quantum dynamics of Bose–Einstein condensate *J. Math. Phys.* **54** 042107

[119] Kiesewetter S, He Q Y, Drummond P D and Reid M D 2014 Scalable quantum simulation of pulsed entanglement and Einstein–Podolsky–Rosen steering in optomechanics *Phys. Rev.* A **90** 043805

[120] Kiesewetter S, Teh R Y, Drummond P D and Reid M D 2017 Pulsed entanglement of two optomechanical oscillators and furry's hypothesis *Phys. Rev. Lett.* **119** 023601

[121] Teh R Y, Kiesewetter S, Reid M D and Drummond P D 2017 Simulation of an optomechanical quantum memory in the nonlinear regime *Phys. Rev.* A **96** 013854

[122] Arecchi F T, Courtens E, Gilmore R and Thomas H 1972 Atomic coherent states in quantum optics *Phys. Rev.* A **6** 2211–37

[123] Agarwal G S 1981 Relation between atomic coherent-state representation, state multipoles, and generalized phase-space distributions *Phys. Rev.* A **24** 2889–96

[124] Barry D W and Drummond P D 2008 Qubit phase space: SU(n) coherent-state *P* representations *Phys. Rev.* A **78** 052108

[125] Corney J F and Drummond P D 2006 Gaussian phase-space representations for fermions *Phys. Rev.* B **73** 125112

[126] Cahill K E and Glauber R J 1969 Density operators and quasiprobability distributions *Phys. Rev.* A **177** 1882–902

[127] Lobino M, Korystov D, Kupchak C, Figueroa E, Sanders B C and Lvovsky A I 2008 Complete characterization of quantum-optical processes *Science* **322** 563–6

[128] Rosales-Zárate L E C and Drummond P D 2011 Linear entropy in quantum phase space *Phys. Rev.* A **84** 042114

[129] Corney J F and Drummond P D 2003 Gaussian quantum operator representation for bosons *Phys. Rev.* A **68** 063822

[130] Corney J F and Drummond P D 2005 Gaussian operator bases for correlated fermions *J. Phys. A: Math. Gen.* **39** 269
[131] Joseph R R, Rosales-Zárate L E C and Drummond P D 2018 Phase space methods for majorana fermions *J. Phys. A: Math. Theor.* **51** 245302
[132] DiVincenzo D P 2000 The physical implementation of quantum computation *Fortschr. Phys.* **48** 771–83
[133] Longdell J J and Sellars M J 2004 Experimental demonstration of quantum-state tomography and qubit-qubit interactions for rare-earth-metal-ion-based solid-state qubits *Phys. Rev.* A **69** 032307
[134] Ospelkaus C, Warring U, Colombe Y, Brown K R, Amini J M, Leibfried D and Wineland D J 2011 Microwave quantum logic gates for trapped ions *Nature* **476** 181
[135] Fuchs G D, Burkard G, Klimov P V and Awschalom D D 2011 A quantum memory intrinsic to single nitrogen-vacancy centres in diamond *Nat. Phys.* **7** 789–93
[136] Chanelière T, Hétet G and Sangouard N 2018 Chapter two—quantum optical memory protocols in atomic ensembles *Advances in Atomic, Molecular, and Optical Physics* vol 67 (New York: Academic) pp 77–150
[137] He Q Y, Reid M D, Giacobino E, Cviklinski J and Drummond P D 2009 Dynamical oscillator-cavity model for quantum memories *Phys. Rev.* A **79** 022310
[138] Bartkiewicz K, Lemr K and Miranowicz A 2013 Direct method for measuring of purity, superfidelity, and subfidelity of photonic two-qubit mixed states *Phys. Rev.* A **88** 052104
[139] Nicolas A, Veissier L, Giner L, Giacobino E, Maxein D and Laurat J 2014 A quantum memory for orbital angular momentum photonic qubits *Nat. Photon.* **8** 234
[140] Lvovsky A I, Hansen H, Aichele T, Benson O, Mlynek J and Schiller S 2001 Quantum state reconstruction of the single-photon Fock state *Phys. Rev. Lett.* **87** 050402
[141] Massar S and Popescu S 1995 Optimal extraction of information from finite quantum ensembles *Phys. Rev. Lett.* **74** 1259–63
[142] Choi M D 1975 Completely positive linear maps on complex matrices *Linear Algebr. Appl.* **10** 285
[143] Jamiołkowski A 1974 An effective method of investigation of positive maps on the set of positive definite operators *Rep. Math. Phys.* **5** 415–24
[144] Feng G, Xu G and Long G 2013 Experimental realization of nonadiabatic holonomic quantum computation *Phys. Rev. Lett.* **110** 190501
[145] Zhang J, Souza A M, Brandao F D and Suter D 2014 Protected quantum computing: interleaving gate operations with dynamical decoupling sequences *Phys. Rev. Lett.* **112** 050502
[146] O'Brien J L, Pryde G J, Gilchrist A, James D F V, Langford N K, Ralph T C and White A G 2004 Quantum process tomography of a controlled-not gate *Phys. Rev. Lett.* **93** 080502
[147] Benhelm J, Kirchmair G, Roos C F and Blatt R 2008 Towards fault-tolerant quantum computing with trapped ions *Nat. Phys.* **4** 463
[148] Ryan C A, Laforest M and Laflamme R 2009 Randomized benchmarking of single-and multi-qubit control in liquid-state NMR quantum information processing *New. J. Phys.* **11** 013034
[149] Veldhorst M *et al* 2015 A two-qubit logic gate in silicon *Nature* **526** 410
[150] Lucero E, Hofheinz M, Ansmann M, Bialczak R C, Katz N, Neeley M, O'Connell A D, Wang H, Cleland A N and Martinis J M 2008 High-fidelity gates in a single Josephson qubit *Phys. Rev. Lett.* **100** 247001
[151] Ektesabi A, Behzadi N and Faizi E 2017 Improved bound for quantum-speed-limit time in open quantum systems by introducing an alternative fidelity *Phys. Rev.* A **95** 022115
[152] Schoenberg I J 1938 Metric spaces and positive definite functions *Trans. Am. Math. Soc.* **44** 522–36
[153] Ma Z-H and Chen J-L 2011 Metrics of quantum states *J. Phys. A: Math. Theor.* **44** 195303