Discrimination of mixed quantum states:
Reversible maps and unambiguous strategies

Inaugural-Dissertation

zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Matthias Kleinmann
aus Ichenhausen

Juni 2008

Aus dem Institut für Theoretische Physik, Lehrstuhl III

der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent:	Prof. Dr. D. Bruß
1. Koreferent:	Prof. Dr. R. Egger
2. Koreferent:	Prof. Dr. N. Lütkenhaus

Tag der mündlichen Prüfung: 30. Juni 2008

Abstract

The discrimination of two mixed quantum states is a fundamental task in quantum state estimation and quantum information theory. In quantum state discrimination a quantum system is assumed to be in one of two possible – in general mixed – non-orthogonal quantum states. The discrimination then consists of a measurement strategy that allows to decide in which state the system was before the measurement. In *unambiguous* state discrimination the aim is to make this decision without errors, but it is allowed to give an inconclusive answer. Especially interesting are measurement strategies that minimize the probability of an inconclusive answer.

A starting point for the analysis of this optimization problem was a result by Eldar *et al.* [Phys. Rev. A **69**, 062318 (2004)], which provides non-operational necessary and sufficient conditions for a given measurement strategy to be optimal. These conditions are reconsidered and simplified in such a way that they become operational. The simplified conditions are the basis for further central results: It is shown that the optimal measurement strategy is unique, a statement that is e.g. of importance for the complexity analysis of optimal measurement devices. The optimal measurement strategy is derived for the case, where one of the possible input states has at most rank two, which was an open problem for many years. Furthermore, using the optimality criterion it is shown that there always exists a threshold probability for each state, such that below this probability it is optimal to exclude this state from the discrimination strategy.

If the two states subject to discrimination can be brought to a diagonal structure with (2×2)-dimensional blocks, then the unambiguous discrimination of these states can be reduced to the unambiguous discrimination of pure states. A criterion is presented that allows to identify the presence of such a structure for two self-adjoint operators. This criterion consists of the evaluation of three commutators and allows an explicit construction of the (2×2)-dimensional blocks.

As an important application of unambiguous state discrimination, unambiguous state comparison, i.e., the question whether two states are identical or not, is generalized and optimal measurements for this problem are constructed.

If for a certain family of states, a physical device maps the input state to an output state, such that a second device can be built that yields back the original input state, such a map is called reversible on this family. With respect to state discrimination, such reversible maps are particularly interesting, if the output states are pure. A complete characterization of all families that allow such a reversible and purifying map is provided. If the states are mapped to pure states, but the map itself is not reversible, upper and lower bounds are analyzed for the "deviation from perfect faithfulness", a quantity which measures the deviation from a reversible mapping.

Zusammenfassung

Die zweifelsfreie Unterscheidung zweier gemischter Quantenzustände gehört zu den fundamentalen Fragestellungen im Bereich der Quantenzustandsabschätzung und der Quanteninformationstheorie. Die Aufgabe bei der Unterscheidung zweier nicht orthogonaler gemischter Quantenzustände besteht darin, durch ein geeignetes Messverfahren festzustellen, in welchem von zwei möglichen – im allgemeinen gemischten – Zuständen sich ein Quantensystem befindet. Das Wesentliche bei der *zweifelsfreien* Zustandsunterscheidung ist, dass bei diesem Verfahren niemals eine falsche Identifikation stattfindet, aber dafür einer dritte, "unschlüssige" Antwort zugelassen ist. Typischerweise sind hierbei Verfahren von besonderem Interesse, welche mit möglichst geringer Wahrscheinlichkeit eine unschlüssige Antwort liefern.

Als Ausgangspunkt für die Analyse dieses Optimierungsproblems dient ein Ergebnis von Eldar u.a. [Phys. Rev. A **69**, 062318 (2004)], welches hinreichende und notwendige (jedoch schwer auswertbare) Bedingungen für die Optimalität eines gegebenen Messverfahrens aufstellt. Diese Bedingungen werden nun aufgegriffen und soweit vereinfacht, dass sie direkt auswertbar sind. Diese vereinfachten Bedingungen bilden die Grundlage für weitere zentrale Ergebnisse: Es wird gezeigt, dass das optimal Messverfahren für die zweifelsfreie Zustandsunterscheidung eindeutig ist; diese Aussage hat z.B. Auswirkungen auf die Komplexitätsanalyse von optimalen Messverfahren. Durch die vereinfachten Optimalitätsbedingungen wird es auch möglich, die optimale Messung für den Fall zu konstruieren, dass einer der Zustände Rang 2 hat; ein Problem, welches viele Jahre lang ungelöst war. Des weiteren wird gezeigt, dass es einen Grenzwert für die *a-priori*-Wahrscheinlichkeit eines jeden Zustands gibt, unterhalb dessen dieser Zustand bei einem optimalen Messverfahren nicht berücksichtigt wird.

Falls die beiden zu unterscheidenden Zustände in eine Form mit (2×2)-dimensionalen Blöcken gebracht werden können, so kann die zweifelsfreie Unterscheidung dieser Zustände auf die Unterscheidung von reinen Zuständen zurückgeführt werden. Hier wird nun ein konstruktives Kriterium bestehend aus drei Kommutatoren aufgestellt, welches es ermöglicht die Existenz solcher Blöcke aufzuspüren und in diesem Falle die Blöcke auch explizit zu konstruieren.

Eine wichtige Anwendung der zweifelsfreien Zustandsunterscheidung ist der zweifelsfreie Zustandsvergleich (d.h. die Frage, ob zwei Zustände gleich sind). Diese Anwendung wird verallgemeinert und optimale Messverfahren werden hierfür entwickelt.

Wenn es für eine Familie von Zuständen möglich ist, diese Zustände mittels einer physikalischen Apparatur so auf andere Zustände abzubilden, dass dieser Vorgang durch eine weitere Apparatur wieder rückgängig gemacht werden kann, so heißt die erstere Abbildung *reversibel*. In Hinblick auf die Zustandsunterscheidung ist eine solche Abbildung von besonderem Interesse, falls die resultierenden Zustände rein sind. In der vorliegende Arbeit werden alle Familien von Zuständen, welche eine solche reversible Abbildung erlauben, vollständig charakterisiert. Es werden obere und untere Schranken an die Abweichung von der Abbildungstreue für Abbildungen aufgestellt, welche zwar auf reine Zustände abbilden, jedoch nicht vollständig reversibel sind.

Arguments against the epistemic interpretation of quantum states: Shall we adopt, then, the epistemic interpretation? Not so fast. The epistemic interpretation is also objectionable. The objections have to do with the apparent need for a knower. For example: What if the knower is a physicist who had a martini before trying to "know"? What if a person who knows just a little physics learns of the result? What if he had a martini? Somehow we feel that such questions are irrelevant.

Shimon Malin in "What are Quantum States?", Quantum Information Processing, Vol. 5, pp. 233-237 (2006)

List of included publications

[A] Kleinmann, M., Kampermann, H., and Bruß, D. (2005). Generalization of quantum-state comparison. *Phys. Rev. A*, 72(3):032308.

[B] Kleinmann, M., Kampermann, H., Meyer, T., and Bruß, D. (2006). Physical purification of quantum states. *Phys. Rev. A*, 73(6):062309.

[C] Kleinmann, M., Kampermann, H., Meyer, T., and Bruß, D. (2007). Purifying and reversible physical processes. *Appl. Phys. B*, 86(3):371–375.

[D] Kleinmann, M., Kampermann, H., Raynal, P., and Bruß, D. (2007). Commutator relations reveal solvable structures in unambiguous state discrimination. *J. Phys. A: Math. Theor.*, 40(36):F871–F878.

[E] Kleinmann, M., Kampermann, H., and Bruß, D. (2008) Structural approach to unambiguous discrimination of two mixed states. `arxiv:0803.1083 [quant-ph]`, *submitted.*

Contents

1 Introduction

The description of nature by means of quantum mechanics is one of the foundations of modern physics. While the framework of quantum mechanics produces very reliable predictions for the behavior of physical systems, the concepts of quantum mechanics do not possess a direct meaning in our everyday experience. Maybe at the heart of this discrepancy is the concept of a quantum state. With the advent of quantum information theory, the properties of quantum states themselves became a central topic. The states are considered as a concept detached from a specific physical embedding and so new questions arise, e.g.[1], *What correlations can be extracted from a composite quantum state when measuring different degrees of freedom?*, *If two remote parties share a quantum state, can they use this state for certain communication primitives?*, *Which output states can be achieved by a physical process, given a characterization of possible input states?*, or *How can one decide whether a system is in state "1" or in state "2"?*

The first question lead to the result, that correlations between quantum systems can be stronger than it would be possible with a local classical description [2]. One of the most prominent results in quantum information theory is an answer to the second question: quantum mechanics in particular allows secret communication between two parties over a public channel [3, 9]. The last two questions are typical for the field of state estimation and state discrimination. The main focus in this thesis will be a major line of research in state discrimination, the unambiguous discrimination of two mixed states.

A fundamental property of quantum mechanics is, that two quantum states cannot be distinguished perfectly by any physical device, unless they are orthogonal. However, imperfect discrimination is possible. The most straightforward approach is to build a device which will distinguish the states imperfectly, such that the device announces the wrong state as rarely as possible. This discrimination strategy is called *minimum error discrimination*. But quantum mechanics also allows a different strategy. It is possible to build a device that will never give a wrong identification, but the device will fail with a non-vanishing probability and in this case no identification at all will be provided. This idea was introduced by Ivanovic [13] and Dieks [8] and is called *unambiguous state discrimination*. The main open problem in unambiguous state discrimination is to maximize the probability of successful identification.

This thesis provides a systematic analysis of this optimization problem for the case of two mixed states. The main results are new classes of optimal measurements and an affirmative answer to the question whether the optimal measurement is unique,

[1] Obviously, this list is biased.

cf.[2] Pub. [E]. As an important application of unambiguous state discrimination, the unambiguous comparison of quantum states is analyzed and generalized in Pub. [A]. During the analysis of unambiguous state discrimination, results where achieved which are relevant also in a more general context: In Pub. [D] a commutator criterion is derived, which allows to detect a common two-dimensional structure in a pair of self-adjoint operators. In Pub. [B] and Pub. [C] reversible physical maps are analyzed, that map families of mixed states to pure states.

2 Quantum states and quantum operations

2.1 Mixed quantum states and purifications

In a description of an ideal, isolated quantum system, the state of the complete system can be considered to be a pure quantum state. However, this description does not correspond to the most general situation, in which the system may have been interacting with other quantum systems. Then, in general, the system under consideration is in a mixed state, and is conveniently described by a density operator on a Hilbert space[3], i.e., by a positive semi-definite operator of unit trace.

The set of density operators is convex, i.e., if σ and σ' are density operators, then for any $0 \leq p \leq 1$ also $\rho = p\,\sigma + (1-p)\,\sigma'$ is a density operator. In such a situation, due to the linearity of quantum mechanics and as a consequence of Born's rule, the system can be considered to be with probability p in state σ and with probability $1-p$ in state σ'. If a state cannot be written as a convex combination of two different density operators, then the state is a *pure* state. (It is also common to use normalized vectors in order to represent pure states. However, in this thesis the main focus is on mixed states and hence the description of a pure state in terms of a density operator is preferred.) Since the complete set of quantum states is the convex hull of all pure states, a mixed state can always be seen as a classical mixture of pure states.

When a quantum system naturally decomposes into two subsystems A and B (i.e., into two independent degrees of freedom), then the complete system is described by a density operator ρ on a tensor product of Hilbert spaces $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$. The state of subsystem A is described by a density operator ρ_A on $\mathscr{H}_A$, where ρ_A is given by the *partial trace* of ρ over $\mathscr{H}_B$, $\rho_A = \operatorname{tr}_B \rho$. The partial trace is defined as $\operatorname{tr}_B \rho = \sum_i \langle i|\rho|i\rangle$, where $\{|i\rangle\}$ is an orthonormal basis of $\mathscr{H}_B$. Even if ρ is a pure state, ρ_A in general is not pure. Conversely, any mixed state can be seen to be the partial trace of a pure state: If

[2] Pub. [A] - Pub. [E] denote publications which are included in this thesis.

[3] Throughout this thesis, the dimension of the underlying Hilbert space is always assumed to be finite.

ρ is a density operator on $\mathscr{H}$, then there exists a pure state ρ_{pur} on an extended Hilbert space $\mathscr{H} \otimes \mathscr{H}_{\text{aux}}$, such that $\rho = \operatorname{tr}_{\text{aux}} \rho_{\text{pur}}$. In this case ρ_{pur} is called a *purification* of ρ.

2.2 Projective measurements

The measurement process is a description of the action of a "macroscopic" measurement device on a quantum system. This process forms the interface between a physical system at quantum level and the observed "classical" events at the measurement apparatus.

An ideal measurement apparatus may have outcomes[4] $1, \ldots, M$, where for each single measurement exactly one outcome "μ" will occur. The characterization of such an apparatus at quantum level is given by a projection valued measure $(\Pi_1, \ldots, \Pi_M)$, where Π_μ corresponds to the measurement outcome "μ". A projection valued measure (as used in this context) is a resolution of unity into a sum of orthogonal projectors, i.e., $\sum_\mu \Pi_\mu = \mathbb{1}$ and $\Pi_\mu \Pi_\mu = \Pi_\mu = \Pi_\mu^\dagger$.

The probability of the measurement outcome "μ" for a system in state ρ is given by Born's rule, $P(\mu|\rho) = \operatorname{tr}(\Pi_\mu \rho)$. If the system was described by the state ρ, then after the measurement it is described by the state $\sum_\mu \Pi_\mu \rho \Pi_\mu$. (In this description, the knowledge of the measurement outcome is ignored.)

2.3 Completely positive maps and generalized measurements

A typical question in quantum state estimation and quantum information theory is, whether a certain mapping $\rho_i \mapsto \rho'_i$ for a family of input states $(\rho_1, \ldots, \rho_N)$ and a family of output states $(\rho'_1, \ldots, \rho'_N)$ can – at least in principle – be realized by a physical apparatus. (In general the density operators ρ_i and ρ'_i act on different Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$, respectively.)

Such a mapping can be realized if and only if the mapping $\rho_i \mapsto \rho'_i$ can be extended to a map Λ on the set of self-adjoint operators, such that Λ is linear, completely positive, and trace preserving. A linear map Λ is positive, if it maps positive semi-definite operators to positive semi-definite operators, i.e., if $\sigma \geq 0$ includes $\Lambda[\sigma] \geq 0$. A map Λ is completely positive, if for any extension $\mathscr{H} \otimes \widetilde{\mathscr{H}}$ of $\mathscr{H}$, the map $(\Lambda \otimes \text{id})$ again is positive. A map Λ is trace preserving if $\operatorname{tr} \sigma = \operatorname{tr} \Lambda[\sigma]$.

Analogously, there exists a complete characterization of the most general measurement apparatus. For a measurement apparatus with outcomes $\mu = 1, \ldots, M$, such a *generalized measurement* is given by a positive operator valued measure (POVM). A POVM $(E_1, \ldots, E_M)$ is a decomposition of unity into a sum of positive semi-definite operators, i.e., $\sum_\mu E_\mu = \mathbb{1}$ and $E_\mu \geq 0$. Any physically realizable measurement can be

[4] In this theses, only measurements with a finite number of outcomes are considered.

described by a POVM, where for a general state ρ, the probability of the outcome "μ" is given by $P(\mu|\rho) = \mathrm{tr}(E_\mu \rho)$.

3 Quantum state estimation

In many fields of modern physics, quantum mechanics has proved to be a satisfactory tool in order to make predictions on physical systems. An admissible question is, to what extent the concepts of quantum mechanics themselves exactly correspond to the actual situation in nature. The aim in *quantum state estimation* is to explore a physical system in such a way that a density operator can be announced, which reliable describes the state of the system. The field of quantum state estimation can be roughly split into three different areas, namely state tomography, parameter estimation, and state discrimination.

In *state tomography* one is equipped with a certain (relatively high) number of identically prepared quantum states. The tomography consists of an unprejudiced strategy in order to perform measurements on these states and to evaluate the measurement data, such that a density operator can be announced, which describes the input state approximately. – Typically, state tomography is used in order to verify that a certain quantum state was successfully prepared. But e.g. in quantum key distribution ("quantum cryptography"), it can be used in order to exclude the presence of an eavesdropper [7, 18].

A specialization of state tomography is *parameter estimation*. Here typically the input state is known to be from a continuous family (ρ_x) which is parametrized by $x \in U \subset \mathbb{R}^n$. Using again a certain number of identically prepared quantum states, which are all in the same unknown state $\rho_y \in (\rho_x)$, measurements are performed and the data is evaluated in such a way that with respect to a quality function $Q_y \colon U \to \mathbb{R}_0^+$, the best estimate of x is achieved.

State discrimination deals only with a single copy of the input state. The physical system subject to discrimination is prepared in one of the states $1, \ldots, N$, described by density operators $\rho_1, \ldots, \rho_N$. The preparation is randomly, where the probability of occurrence for the state "μ" is given by $p_\mu > 0$ with $\sum_\mu p_\mu = 1$. The task in state discrimination is to find a measurement which can discriminate between these states, i.e., to find a POVM $(E_1, \ldots, E_M)$ with $M \geq N$, such that the measurement result "μ" corresponds to an identification of state "μ". There exist two main branches of state discrimination:

In *minimum error discrimination*, there are exactly as many measurement outcomes as possible input states, i.e., $M = N$. The measurement maximizes the probability of correct identification (and hence minimizes the error), $P_{\text{succ}}^{\text{MED}} = \sum_\mu p_\mu \operatorname{tr}(E_\mu \rho_\mu)$. For two mixed states, the optimal measurement was found by Helstrøm [11].

In *unambiguous state discrimination*, no wrong identification is allowed, i.e., $\operatorname{tr}(E_\mu \rho_\nu) = 0$ for all $\nu \neq \mu \leq N$. Unless all states are orthogonal, no perfect discrimination is possible and hence there must be an additional measurement outcome ($M = N + 1$), which corresponds to an inconclusive result. This additional measurement outcome is usually labeled by "?". The task is to find an optimal measurement, which again maximizes the success probability $P_{\text{succ}}^{\text{USD}} = \sum_\mu p_\mu \operatorname{tr}(E_\mu \rho_\mu)$.

3.1 Distance measures on the set of density operators

When focusing on quantum states, a quite natural question is, *How similar are two density operators ρ_1 and ρ_2?* This question is closely related to quantum state discrimination and surprisingly, there exists no satisfactory answer to this question.

For a distance on the set of density operators, there are several reasonable requirements. In mathematics there are three axioms which should apply for a distance, namely positivity, symmetry, and the triangular inequality: If $d(\rho_1, \rho_2)$ denotes a distance between two density operators ρ_1 and ρ_2, then the axioms are (i)(a) $d(\rho_1, \rho_2) \in \mathbb{R}_0^+$ where (b) $d(\rho_1, \rho_2) = 0$ if and only if $\rho_1 = \rho_2$, (ii) $d(\rho_1, \rho_2) = d(\rho_2, \rho_1)$, and (iii) $d(\rho_1, \rho_2) \leq d(\rho_1, \rho_3) + d(\rho_3, \rho_2)$ for any density operator ρ_3.

In physical terms, a distance potentially should represent a measure of "similarity" or "distinguishability". In these terms, in particular the axioms (i)(b) and (iii) may become less important. But from a physical point of view, an additional condition may become desirable: No physical process should increase the distance of ρ_1 and ρ_2. Assuming, that the distance measure is defined for density operators acting on a Hilbert space of arbitrary dimension, this condition can be split into three different conditions (iv) $d(\rho_1, \rho_2) = d(\rho_1 \otimes \sigma, \rho_2 \otimes \sigma)$ for all density operators σ, (v) $d(\rho_1, \rho_2) = d(U\rho_1 U^\dagger, U\rho_2 U^\dagger)$ for any unitary operator U, and (vi) $d(\operatorname{tr}_X \rho_1, \operatorname{tr}_X \rho_2) \leq d(\rho_1, \rho_2)$ for any subsystem X.

Two important distance measures are the *trace distance* $d_1(\rho_1, \rho_2) = \frac{1}{2} \operatorname{tr} |\rho_1 - \rho_2|$ and the *Bures distance* $d_B(\rho_1, \rho_2)^2 = 2 - 2 \operatorname{tr} |\sqrt{\rho_1}\sqrt{\rho_2}|$, where $|A| = \sqrt{A^\dagger A}$. Both distances fulfill all axioms (i)-(vi).

One of the most prominent appearances of the trace distance is in the success probability of minimum error discrimination. For two states ρ_1 and ρ_2 having the same *a priori* probability $\frac{1}{2}$, the success probability in minimum error discrimination is given by $P_{\text{succ}}^{\text{MED}} = \frac{1}{2}(1 + d_1(\rho_1, \rho_2))$. In unambiguous state discrimination, on the other hand, the Bures distance appears indirectly: An important upper bound on the success probability is given by $P_{\text{succ}}^{\text{USD}} \leq d_B(\rho_1, \rho_2)^2/2$.

3.2 Optimal unambiguous state discrimination of two mixed states

The main topic under investigation in this thesis is unambiguous state discrimination. A measurement $(E_1, E_2, E_?)$, which can unambiguously discriminate between two states ρ_1 and ρ_2 must satisfy $\mathrm{tr}(E_1\rho_2) = 0$ and $\mathrm{tr}(E_2\rho_1) = 0$. Let $\ker A = \{|\phi\rangle \in \mathscr{H} \mid A|\phi\rangle = 0\}$ denote the kernel of an operator A and write $\mathrm{supp}\, A = (\ker A)^\perp$ (the orthocomplement of $\ker A$) for the support of A. Then the condition $\mathrm{tr}(E_1\rho_2) = 0$ is equivalent to the condition $\mathrm{supp}\, E_1 \subseteq \ker \rho_2$ and analogously $\mathrm{tr}(E_2\rho_1) = 0$ is equivalent to $\mathrm{supp}\, E_2 \subseteq \ker \rho_1$. These properties in principle allow to construct a non-trivial measurement for any two states, whenever $\ker \rho_1 \supsetneq \{0\}$ or $\ker \rho_2 \supsetneq \{0\}$ holds.

The aim in optimal unambiguous state discrimination is to find a measurement that maximizes the success probability $P_{\text{succ}} = p_1 \,\mathrm{tr}\, E_1\rho_1 + p_2 \,\mathrm{tr}\, E_2\rho_2$, where p_1 is the *a priori* probability of ρ_1 and p_2 is the *a priori* probability of ρ_2. For the case where ρ_1 and ρ_2 are both pure, the optimal measurement was provided by Jaeger and Shimony [14].

Surprisingly, most of the optimal solutions for the mixed state case so far utilize the result for the pure state case. This is possible due to the following reasons. As shown by Raynal *et al.* [20], the part of the Hilbert space which is given by $\mathrm{supp}\, \rho_1 \cap \mathrm{supp}\, \rho_2$ cannot be used at all for unambiguous discrimination. On the other hand [20] a prefect discrimination can always be performed in the subspaces $\mathrm{supp}\, \rho_1 \cap \ker \rho_2$ and $\ker \rho_1 \cap \mathrm{supp}\, \rho_2$, without decreasing the success probability. Hence it is possible to reduce the optimization problem in unambiguous discrimination to a new problem ρ_1' and ρ_2' with according *a priori* probabilities p_1' and p_2', respectively. For this reduced problem one can achieve that $\mathrm{supp}\, \rho_1' \cap \mathrm{supp}\, \rho_2' = \{0\}$, $\mathrm{supp}\, \rho_1' \cap \ker \rho_2' = \{0\}$, and $\ker \rho_1' \cap \mathrm{supp}\, \rho_2' = \{0\}$. After this reduction the rank of both density operators are equal, $\mathrm{rank}\, \rho_1' = \mathrm{rank}\, \rho_2'$, where $\mathrm{rank}\, \rho_1' \leq \mathrm{rank}\, \rho_1$ and $\mathrm{rank}\, \rho_2' \leq \mathrm{rank}\, \rho_2$. Hence in particular the unambiguous discrimination of a pure state and a mixed state always reduces to the discrimination of two pure states [5, 6].

A second way to utilize the pure state discrimination is a blockwise application. Assume that ρ_1 and ρ_2 are composed from mutually orthogonal pairs of pure states, $\rho_1 = \sum_k \eta_1^k \sigma_1^k$ and $\rho_2 = \sum_k \eta_2^k \sigma_2^k$, where $\mathrm{tr}[(\sigma_1^k + \sigma_2^k)(\sigma_1^\ell + \sigma_2^\ell)] = 0$ for all $k \neq \ell$, and $\eta_\mu^k > 0$ with $\sum_k \eta_1^k = 1$ and $\sum_k \eta_2^k = 1$. Since the pairs are mutually orthogonal, it is possible to apply the result by Jaeger and Shimony in each block independently [4, 19]. A very similar idea leads to a lower bound on the success probability. For this lower bound, first a von-Neumann measurement $\rho_i \mapsto \sum_k \Pi_k \rho_i \Pi_k$ is applied to the states, such that after this measurement, the states are composed from mutually orthogonal pairs of pure states, as above. An explicit construction of such a strategy (which is, in general, not optimal) was provided by Rudolph *et al.* [21].

Also the upper bound on the success probability, mentioned at the end of the previous subsection, is based on the optimal measurement for two pure states [21]. Let ρ_1^{pur} and ρ_2^{pur} denote purifications of ρ_1 and ρ_2, respectively, such that $\operatorname{tr}_{\mathrm{aux}} \rho_1^{\mathrm{pur}} = \rho_1$ and $\operatorname{tr}_{\mathrm{aux}} \rho_2^{\mathrm{pur}} = \rho_2$. Then the optimal success probability for discriminating the purifications cannot be lower than the one for the original states, since the mapping $\rho_\mu^{\mathrm{pur}} \mapsto \rho_\mu$ can be performed by a physical device. Hence the success probability for discriminating the pure states is an upper bound on the success probability of the original problem.

The analysis of the attainability of this bound leads to the "fidelity form measurement" [12, 19]. The bound can be reached, if and only if certain positivity conditions are satisfied [19]. The name "fidelity form measurement" originates in the fact, that the measurement operators E_1 and E_2 contain the operators $\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$ and $\sqrt{\sqrt{\rho_2}\rho_1\sqrt{\rho_2}}$. The trace of these operators is equal to the Uhlmann fidelity $\operatorname{tr}|\sqrt{\rho_1}\sqrt{\rho_2}|$.

—

The optimization problem in unambiguous state discrimination is a typical example of a convex optimization problem, i.e., an optimization problem of a convex function over a convex set. This in principle allows an efficient numerical solution. However, in many examples this turns out to be less reasonable due to the specific structure of unambiguous state discrimination. More important, a numerical solution will not allow a deeper insight into the structure of unambiguous state discrimination. Important questions here concern a classification and the possible uniqueness of the optimal measurement, or the complexity of a measurement with respect to a possible experimental implementation. But the analytical methods developed for general convex optimization allow to give insight into the structure of the optimization problem in unambiguous state discrimination. An important result which is based on such an analysis are the optimality conditions by Eldar *et al.* [10]. These conditions roughly state that a given measurement in unambiguous state discrimination is optimal if and only if a positive semi-definite operator Z can be found, such that a certain set of equations and positivity conditions is satisfied. However, as the operator Z is unknown, these results cannot be used in a constructive way and only a few, very specialized applications were known for these optimality conditions.

4 Summary of results

4.1 Optimal unambiguous state discrimination

As outlined in Sec. 3.2, the only generic solution known for the optimal unambiguous discrimination of two mixed states was for the case of a pure state and an arbitrary mixed state [5, 6]. One of the central results in Pub. [E] is the construction of the optimal measurement for a pair of states, where both states have rank two ("solution in four dimensions"). Due to the reduction theorems[5] by Raynal *et al.* [20], this result extends to the case of a state of rank two and an arbitrary mixed state.

A first step towards this solution in four dimensions is a general classification of the optimal measurement, according to the rank of E_1 and E_2, cf. Pub. [E] (Sec. 3.2): If r denotes the rank of the states ρ_1 and ρ_2 (after the reductions have been performed), then the number of different measurement *types* is given by $\frac{1}{2}(r+1)(r+2)$, where a measurement type is characterized by the rank of E_1 and E_2. Since the discrimination task is symmetric with respect to the interchange of ρ_1 and ρ_2, one only has to consider $\lfloor(\frac{r}{2}+1)^2\rfloor$ different measurement *classes*, where $\lfloor x\rfloor$ denotes the floor function. It follows that in the case of two pure states, only two measurement classes exist, while in the case of two rank two states, already four classes have to be considered.

The two measurement classes which occur in the unambiguous discrimination of two pure states can be generalized to arbitrary states. The first class is the case where $\operatorname{rank} E_1 = \operatorname{rank} E_2 = r$. According to Pub. [E] (Sec. 5.2), the optimal measurement in this case is always given by the "fidelity form measurement" [12, 19]. The second measurement class, which can be generalized from the pure state case are "single state detection measurements", where either $E_1 = 0$ or $E_2 = 0$ holds. A single state detection measurement with $E_1 = 0$ is optimal if and only if[6] $\rho_1(p_2\rho_2 - p_1\rho_1)\rho_1 \geq 0$, cf. Pub. [E] (Proposition 12). In particular for any ρ_1 and ρ_2 with $\operatorname{supp}\rho_1 \cap \operatorname{supp}\rho_2 = \{0\}$, there always exist *a priori* probabilities p_1 and p_2, such that the single state detection measurement with $E_1 = 0$ is optimal, typically for $\frac{1}{p_1} \gg 1$, cf. Pub. [E] (Eq. (33)). An analogous analysis holds for $E_2 = 0$.

The basis for the analysis of the single state detection measurement and also the basis for the remaining steps in order to find the solution in four dimensions is a simplification of the optimality conditions by Eldar *et al.*[10], cf. Pub. [E] (Sec. 4). It is possible to eliminate the unknown operator Z from the original conditions, so that the optimality conditions can be rephrased as follows, cf. Pub. [E] (Corollary 9): *A measurement* $(E_1, E_2, E_?)$ *is optimal if and only if*

$$(\Lambda_1 - \Lambda_2)E_?(p_2\rho_2 - p_1\rho_1)E_?(\Lambda_1 + \Lambda_2) \geq 0$$
$$(\Lambda_1 - \Lambda_2)E_?(p_2\rho_2 - p_1\rho_1)E_?(\mathbb{1} - E_?) = 0,$$

where Λ_1 *is the projector onto* $\ker\rho_2$ *and* Λ_1 *the projector onto* $\ker\rho_1$*.* This result has several consequences. First, in many situations from the symmetry of the problem it is possible to guess an optimal measurement. This criterion allows to verify the optimality

[5] These reduction theorems are simplified in Pub. [E] (Sec. 3.1).

[6] Here it is always assumed that $\operatorname{supp}\rho_1 + \operatorname{supp}\rho_2 = \mathscr{H}$ holds.

of this measurement, a result that otherwise may be difficult to achieve. At second it is possible to try to find solutions for $E_?$ from the equations above. In particular the solution in four dimensions is based on this idea.

A further consequence is more indirect. From the second condition in the optimality criterion, it is possible to show that given $\operatorname{supp} E_?$ of an optimal measurement, one can reconstruct $E_?$ itself, cf. Pub. [E] (Lemma 10 and Eq. (9)). Given $E_?$ on the other hand, it is always possible to uniquely reconstruct also E_1 and E_2, cf. Pub. [E] (Proposition 3.) Since the rank of $E_?$ is given by $\operatorname{rank} E_? = \operatorname{rank}(\rho_1\rho_2)$, cf. Pub. [E] (Theorem 4), it follows from the convexity of the set of all optimal measurements, that the optimal measurement is unique, cf. Pub. [E] (Proposition 11).

The uniqueness of the optimal measurement itself also has important consequences. One of it concerns the complexity analysis of optimal measurements, e.g. whether a measurement can be implemented locally by remote parties. Another one is that the uniqueness helps to construct new solutions: if it is possible to show, that if a certain measurement would be optimal, a second measurement would exist with the same success probability, then this is a contradiction to the uniqueness statement and hence neither of the measurements can be optimal. This type of argument is important in order to derive the solution in four dimensions.

4.2 Unambiguous state comparison

State comparison was originally introduced by Barnett *et al.*[1] and is is an elementary task in quantum information processing. Consider a system composed from two identical subsystems, where each subsystem is independently prepared in either of the pure states π_1 or π_2, i.e., the composed system is described by either of the four density operators $\pi_1 \otimes \pi_1$, $\pi_1 \otimes \pi_2$, $\pi_2 \otimes \pi_1$, or $\pi_2 \otimes \pi_2$. The task is to find the optimal measurement, that can unambiguously discriminate whether both systems are in the same state or in a different state, i.e., to distinguish between the set $\{\pi_1 \otimes \pi_1, \pi_2 \otimes \pi_2\}$ and the set $\{\pi_1 \otimes \pi_2, \pi_2 \otimes \pi_1\}$.

In Pub. [A] a natural generalization of this task was defined: *Given C quantum states, each of them taken from a family of N states $(\pi_1, \dots, \pi_N)$ that occur with corresponding* a priori *probabilities $(q_1, \dots, q_N)$. Unambiguous state comparison "C out of N" is performed by doing a measurement, which allows with probability P_{succ} to decide without doubt whether all C states are equal or whether at least one of them is different. The best probability of success is reached in optimal state comparison.*

This problem can be mapped to the (optimal) unambiguous discrimination of two mixed states, namely with $\gamma_a = \sum_i (q_i\pi_i)^{\otimes C}$ and $\gamma_b = (\sum_i q_i\pi_i)^{\otimes C} - \gamma_a$ to the unambiguous discrimination of $\rho_a = \gamma_a/p_a$ and $\rho_b = \gamma_b/p_b$, appearing with *a priori* probabilities given by $p_a = \operatorname{tr}\gamma_b$ and $p_b = \operatorname{tr}\gamma_b$.

For the case "2 out of 2" with two pure states π_1 and π_2, this is solved in Pub. [A]. The optimal success probability turns out not to be reachable by any local measurement strategy, i.e., by a strategy where both systems subject to comparison are only connected by a classical channel. In order to show this, it is proven that the optimal local strategy is given by the naïve approach, to unambiguously distinguish π_1 and π_2 in each system and than compare the measurement results. (Actually it is shown, that no separable measurement can be better than this strategy. This results is slightly stronger, since not all separable measurements can be realized by a local strategy.)

The solution of the case "2 out of 2" is based on the fact, that this comparison task can be reduced to the discrimination of two pure states. However, as shown in Pub. [A], the case "2 out of N" for $N > 2$ corresponds to the unambiguous discrimination of two mixed states, both having rank N. For the case where all pure states $(\pi_1, \ldots, \pi_N)$ occur with the same *a priori* probability $1/N$ and all states have the same mutual distance, $d_1(\pi_i, \pi_j) \equiv \sin\vartheta$ for all $i \neq j$, state comparison "2 out of N" can be simplified to the N-fold unambiguous discrimination of two pure states, as shown in Pub. [D].

The comparison of a string of C states, where each state is taken from an alphabet of two pure states (π_1, π_2) with a probability distribution (p_1, p_2) can be mapped to the unambiguous discrimination of a mixed state of rank 2 and a state of rank $2^C - 2$. Hence it can be solved by the methods developed in Pub. [E], cf. Ref. [16].

4.3 Common block diagonal structures

In the discussion of unambiguous state discrimination in Sec. 3.2 it was mentioned that the optimization problem reduces to the pure state case, whenever the two states can be decomposed into mutually orthogonal pairs of pure states. However, given two arbitrary states, it is not obvious whether such a decomposition exists. The main result in Pub. [D] is an operable criterion in order to detect and construct such decompositions.

More generally, a *common block diagonal structure* (CBS) of two operators A and B is a projection valued measure (Π_i), such that $[A, \Pi_i] = 0$ and $[B, \Pi_i] = 0$ for all i. In this case the map $X \mapsto \sum_i \Pi_i X \Pi_i$ does not change A and B. Denoting by n the maximal rank of all Π_i, such a common block diagonal structure is called *at most n-dimensional.* The pairs of states, which possess an at most *two*-dimensional CBS, form a superclass of the states that can be decomposed into mutually orthogonal pairs of pure states, cf. Pub. [D] (Lemma 1). However, also for the case of an at most *two*-dimensional CBS, the optimal measurement for unambiguous discrimination reduces to the pure state case, cf. Ref. [19] and Pub. [D].

The main theorem in Pub. [D] is a criterion to identify an at most two-dimensional

CBS: *Two self-adjoint operators A and B on $\mathscr{H}$ with* $\ker A + \ker B = \mathscr{H}$ *have an at most two-dimensional common block diagonal structure, if and only if* $[A, ABA] = 0$, $[A, AB^2A] = 0$, *and* $[B, BA^2B] = 0$. This result it significantly simpler than a similar result due to Laffey [17], which in contrast holds for any pair of positive semi-definite operators. In addition the proof of the theorem in Pub. [D] is constructive (while the one in Ref. [17] is not), so that an explicit construction of the CBS is possible.

For the application to the unambiguous discrimination of two mixed states, the condition $\ker \rho_1 + \ker \rho_2 = \mathscr{H}$ is no restriction, since it is equivalent to $\operatorname{supp} \rho_1 \cap \operatorname{supp} \rho_2 = \{0\}$ and this situation can always be achieved by virtue of the first reduction theorem [20]. Indeed, this reduction and also the second reduction [20] can only create, but not destroy any CBS, cf. Pub. [D].

4.4 Purifying and reversible maps

The success probability for the unambiguous discrimination of two mixed states is upper bounded by the "fidelity bound" introduced by Rudolph *et al.* [21]. The basic idea for this bound is that by definition, a purification ρ_{pur} on $\mathscr{H} \otimes \mathscr{H}_{\text{aux}}$ of a density operator ρ on $\mathscr{H}$ can physically be mapped to the original state ρ by taking the partial trace over the auxiliary system, $\rho = \operatorname{tr}_{\text{aux}} \rho_{\text{pur}}$. After a physical map, the success probability of a generalized measurement cannot be higher than before the map. Hence an upper bound on the success probability for the unambiguous discrimination of (ρ_1, ρ_2) is given by the optimal success probability for unambiguously discriminating the purification of both states.

A physical reason that this bound cannot always be attained, lies in the fact that the partial trace on a family of pure states in general is not reversible. A physical process is reversible on a family of states, if the process can be undone by virtue of a second physical process. That is, a completely positive and trace preserving map Λ is reversible on a family of states (ρ_i), if there exists a completely positive and trace preserving map Λ', such that $(\Lambda' \circ \Lambda)[\rho_i] = \rho_i$ for all i.

In Pub. [B] and Pub. [C] a complete characterization of all families of states is given, that allow the existence of a reversible map to a family of pure states. Such a family roughly consists of a family of pure states (π_x) tensored with a common mixed state σ, i.e., $(\pi_x \otimes \sigma)$. For a family with only two members, an operational criterion was provided in Pub. [B], which allows to verify whether a reversible and purifying map exists. A pair (ρ_1, ρ_2) can be reversibly mapped to a pair of pure states, if and only if $d_1(\rho_1, \rho_2) = \mathcal{D}(\rho_1, \rho_2)$. Here $d_1(\rho_1, \rho_2) = \frac{1}{2} \operatorname{tr} |\rho_1 - \rho_2|$ is the trace distance and $\mathcal{D}(\rho_1, \rho_2)$ denotes the worst case distinguishability as defined in Pub. [B]: For a mixed state ρ, denote by $\mathcal{Q}_\rho$ the set of pure states π, such that $\rho - \alpha\pi \geq 0$ for some $\alpha > 0$. Then

$\mathcal{D}(\rho_1, \rho_2) = \inf d_1(\pi_1, \pi_2)$, where the infimum is taken over all $\pi_1 \in \mathcal{Q}_{\rho_1}$ and $\pi_2 \in \mathcal{Q}_{\rho_2}$. (Note, that there is an efficient way to calculate the worst case distinguishability, cf. Pub. [B].) This quantity is not a proper distance measure, since it does not satisfy the axioms (i)(a) and (iii) discussed in Sec. 3.1. But it has a meaning in terms of the "distinguishability" of states, cf. Pub. [C].

Since pure states are much easier to deal with than mixed states, it might also be desirable to consider physical processes, that map a pair of mixed states to a pair of pure states, but to allow, that these maps might not be perfectly reversible. For such a map Λ, the inequality $d_1(\Lambda[\rho_1], \Lambda[\rho_2]) \leq \mathcal{D}(\rho_1, \rho_2)$ holds and equality can always be achieved, cf. Pub. [B] and Pub. [C]. As figure of merit for maps, that are not perfectly reversible, in Pub. [B], the notion of "deviation from perfect faithfulness" was introduced. If a map Λ with pure output is perfectly reversible, than one can always achieve, that the output of the map is a purification of the input state, i.e., $\rho_i = \mathrm{tr}_{\mathrm{aux}}(\Lambda[\rho_i])$. The deviation from perfect faithfulness then is defined as $\sum_i p_i d_1(\rho_i, \mathrm{tr}_{\mathrm{aux}}(\Lambda[\rho_i]))$, where p_i is the *a priori* probability of ρ_i. In Pub. [B] upper and lower bounds on this deviation were derived for the case of two input states.

5 Outlook

In the unambiguous discrimination of mixed states, there are still many open questions. Although the optimal measurement for the case of a mixed state of rank two and an arbitrary mixed state was found, the structure of this solution is rather complicated. It would be interesting to deepen the analysis of this solution and provide more examples. On the other hand the next step – the optimal measurement for a mixed state of rank three and an arbitrary mixed state – is expected to be even more complicated. From the six measurement classes, that will occur in this solution, only the analysis of the single state detection measurement and the analysis of the fidelity form measurement is complete. Another generalization, for which so far almost no results are known, is the unambiguous discrimination of more than two states. However, due to the methods developed in Pub. [E], certain results that are known from the two state case can be generalized to the many state case [15].

In the analysis of block diagonal structures, the focus was to find structures of dimension two. One of the main motivations for the analysis of two-dimensional structures was the fact, that the general unambiguous discrimination was only known in two dimensions. Due to the solution in four dimensions, it now in particular becomes interesting to also identify blocks of dimensions four. With respect to the many-state case (and also for general considerations) a possible generalization to more than two operators should be considered.

Also in the analysis of reversible and incomplete maps, several question remain open. If two states cannot be mapped to a pure state in a reversible way, only upper and lower bounds are known for the best mapping to pure states. The optimization problem in this case – namely to minimize the deviation from perfect faithfulness – might give a deeper insight into the structure of such purifying maps. On the other hand also the analysis of maps, that are reversible, but do not perfectly map to pure states, would be an interesting physical question.

6 List of main results

- The optimal measurement for the unambiguous discrimination of two mixed states is unique.

- An analytic expression for the optimal measurement for the unambiguous discrimination of two mixed states was derived for the case where one of the states has rank two.

- For any two mixed states with non-overlapping support, there exist a certain threshold probability for each of the states, such that below this *a priori* probability the optimal measurement for unambiguous discrimination will never detect this state ("single state detection"). Given two states, this threshold probability can be calculated explicitly.

- Given a measurement for the unambiguous discrimination of two mixed states, this measurement is optimal if and only if certain conditions are satisfied. These conditions are operational.

- The optimal measurement for the unambiguous comparison of two states, taken from a family of two pure states, cannot be implemented by any local measurement strategy.

- For self-adjoint operators A and B with non-overlapping support, the commutators $[A, ABA]$, $[A, AB^2A]$, and $[B, BA^2B]$ vanish simultaneously if and only if A and B have an at most two-dimensional common block diagonal structure. The reductions in unambiguous state discrimination do not destroy any common block diagonal structure.

- There exist (even continuous) families of mixed states, that can reversibly be mapped to pure states by a physical device. These families are *essentially pure*. For a pair of states which is not essentially pure, it can be possible to map the states to pure states, such that this map is close to be reversible.

7 Bibliography

[1] Barnett, S. M., Chefles, A., and Jex, I. (2003). Comparison of two unknown pure quantum states. *Phys. Lett. A*, 307(4):189–195.

[2] Bell, J. S. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200.

[3] Bennett, C. H. and Brassard, G. (1984). Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalora, India*, pages 175–179, New York. IEEE.

[4] Bergou, J. A., Feldman, E., and Hillery, M. (2006). Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination. *Phys. Rev. A*, 73(3):032107.

[5] Bergou, J. A., Herzog, U., and Hillery, M. (2003). Quantum filtering and discrimination between sets of boolean functions. *Phys. Rev. Lett.*, 90(25):257901.

[6] Bergou, J. A., Herzog, U., and Hillery, M. (2005). Optimal unambiguous filtering of a quantum state: An instance in mixed state discrimination. *Phys. Rev. A*, 71(4):042314.

[7] Bruß, D. (1998). Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81(14):3018–3021.

[8] Dieks, D. (1988). Overlap and distinguisability of quantum states. *Phys. Lett. A*, 126:303–306.

[9] Ekert, A. (1991). Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661.

[10] Eldar, Y. C., Stojnic, M., and Hassibi, B. (2004). Optimal quantum detectors for unambiguous detection of mixed states. *Phys. Rev. A*, 69:062318.

[11] Helstrøm, C. W. (1976). *Quantum Detection and Estimation Theory.* Acad. Press, New York.

[12] Herzog, U. and Bergou, J. A. (2005). Optimum unambiguous discrimination of two mixed quantum states. *Phys. Rev. A*, 71(5):050301(R).

[13] Ivanovic, I. (1987). How to differentiate between non-orthogonal states. *Phys. Lett. A*, 123:257–259.

[14] Jaeger, G. and Shimony, A. (1995). Optimal distinction between two non-orthogonal quantum states. *Phys. Lett. A*, 197:83–87.

[15] Kleinmann, M., Kampermann, H., and Bruß, D. Optimal unambiguous discrimination of many states. *in preparation.*

[16] Kleinmann, M., Kampermann, H., and Bruß, D. Unambiguous state discrimination: optimal solution and case study. *in preparation.*

[17] Laffey, T. J. (1977). Simultaneous quasidiagonalization of complex matrices. *Lin. Alg. Appl.*, 16(3):189–201.

[18] Meyer, T., Kampermann, H., Kleinmann, M., and Bruß, D. (2006). Finite key analysis for symmetric attacks in quantum key distribution. *Phys. Rev. A*, 74(4):042340.

[19] Raynal, P. and Lütkenhaus, N. (2005). Optimal unambiguous state discrimination of two density matrices: Lower bound and class of exact solutions. *Phys. Rev. A*, 72:022342, 049909(E).

[20] Raynal, P., Lütkenhaus, N., and van Enk, S. J. (2003). Reduction theorems for optimal unambiguous state discrimination of density matrices. *Phys. Rev. A*, 68(2):022308.

[21] Rudolph, T., Spekkens, R. W., and Turner, P. S. (2003). Unambiguous discrimination of mixed states. *Phys. Rev. A*, 68(1):010301(R).

Publication A

Generalization of quantum-state comparison
Physical Review A, Vol. **72**, p. 032308 (2005) *(7 pages)*
M. Kleinmann, H. Kampermann, and D. Bruß

Generalization of quantum-state comparison

M. Kleinmann,* H. Kampermann, and D. Bruß

Institut für Theoretische Physik III, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

(Received 2 March 2005; published 6 September 2005)

We investigate the unambiguous comparison of quantum states in a scenario that is more general than the one that was originally suggested by Barnett *et al.* First, we find the optimal solution for the comparison of two states taken from a set of two pure states with arbitrary *a priori* probabilities. We show that the optimal coherent measurement is always superior to the optimal incoherent measurement. Second, we develop a strategy for the comparison of two states from a set of N pure states, and find an optimal solution for some parameter range when $N=3$. In both cases we use the reduction method for the corresponding problem of mixed-state discrimination, as introduced by Raynal *et al.*, which reduces the problem to the discrimination of two pure states only for $N=2$. Finally, we provide a necessary and sufficient condition for unambiguous comparison of mixed states to be possible.

DOI: 10.1103/PhysRevA.72.032308 PACS number(s): 03.67.−a, 03.65.Ta

I. INTRODUCTION

The laws of quantum mechanics do not allow the perfect discrimination of two nonorthogonal quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$. Consequently, given a set of nonorthogonal states $\{|\psi_1\rangle,|\psi_2\rangle\}$, it is also impossible to find out with probability one whether two quantum states, drawn from this set, are identical (namely, the total state is either $|\psi_1\psi_1\rangle$ or $|\psi_2\psi_2\rangle$) or different (i.e., the total state is either $|\psi_1\psi_2\rangle$ or $|\psi_2\psi_1\rangle$). What is the optimal probability of success, when no errors are allowed? This problem has been introduced by Barnett, Chefles, and Jex [1] and is called unambiguous quantum state comparison. It has been solved for the case that the *a priori* probabilities for the two ensemble states are equal [1]. The task of determining whether C given states taken from a set of N pure states with equal *a priori* probabilities are identical or not has been investigated in Refs. [2,3].

In this paper, we consider the most general case of unambiguous state comparison, also admitting mixed states. We provide sufficient and necessary conditions for which this task can succeed. Furthermore, the comparison of two states drawn from a set of N pure states with arbitrary *a priori* probabilities is investigated, and an optimal solution is found for the case $N=2$, as well as for a range of parameters in the case $N=3$, using the reduction techniques for mixed-state discrimination developed in Ref. [4]. This method is also applied for general N. We address the question of how much can be gained in the optimal coherent strategy (i.e., with global measurements on the two given states), as compared to the best incoherent strategy (i.e., consecutive measurements).

Our paper is organized as follows: in Sec. II, we define the most general state comparison problem, and explain the connection to mixed-state discrimination. In Sec. III, we find the optimal solution for comparing two states, drawn from a set of two states. In Sec. IV, we develop the formalism for the comparison of two out of N states, and apply it to $N=3$. In Sec. V we derive sufficient and necessary conditions for the general task of mixed-state comparison to be successful, before concluding in Sec. VI.

II. GENERAL STATE COMPARISON

Let us define the task of state comparison in the most general way: Given C quantum states of arbitrary dimension, each of them taken from a set of N possible (in general mixed) quantum states $\{\pi_1,\ldots,\pi_N\}$ that occur with corresponding *a priori* probabilities $\{q_1,\ldots,q_N\}$. Unambiguous state comparison "C out of N" is performed by doing a measurement, which allows with probability P to decide without doubt whether all C states are equal, or whether at least one of them is different. The best possible probability of success P_{opt} is reached in optimal state comparison.

A measurement is most generally described as a positive operator-valued measurement (POVM), i.e., a decomposition of the identity operator into a set of n positive operators [5],

$$F_1,\ldots,F_n \geqslant 0, \quad \text{satisfying} \sum_i F_i = \mathbb{1}. \tag{1}$$

The probability for a system in a state ϱ_k to yield the outcome corresponding to F_i is given by $p_k\,\mathrm{tr}(F_i\varrho_k)$, where p_k is the *a priori* probability for the system being in state ϱ_k. For the task of unambiguous state comparison, we need at least two measurements F_a and F_b, having vanishing probabilities in the case where the total state is composed of different or equal states, respectively. This means that for all $(p_k,\varrho_k) \in\{(q_{i_1}\cdots q_{i_C},\pi_{i_1}\otimes\cdots\otimes\pi_{i_C})|i_1,\ldots,i_C\in\{1,\ldots,N\}\}$ we demand

$$p_k\,\mathrm{tr}(F_a\varrho_k) > 0 \Leftrightarrow \exists\ m\colon \varrho_k = \pi_m^{\otimes C}, \tag{2a}$$

$$p_k\,\mathrm{tr}(F_b\varrho_k) > 0 \Leftrightarrow \nexists m\colon \varrho_k = \pi_m^{\otimes C}. \tag{2b}$$

However, measurements which satisfy this defining property will in general not sum up to the identity, thus admitting the inconclusive measurement $F_?=\mathbb{1}-F_a-F_b$, which has to be a positive operator. In order to find an *optimal* solution to the

*Electronic address: kleinmann@thphy.uni-duesseldorf.de problem, one has to minimize the probability for the inconclusive answer $\Sigma_k p_k \operatorname{tr}(F_? \varrho_k)$, or equivalently maximize the rate of success given by

$$P = \sum_k p_k \operatorname{tr}[(F_a + F_b)\varrho_k]. \tag{3}$$

The problem of finding the optimal measurement for state comparison can be addressed by considering the optimal solution of a related problem, namely unambiguous state discrimination. Here, two states ρ_a and ρ_b have to be distinguished without error, but admitting an inconclusive answer. In order to see the connection between the two tasks, consider the mixed states

$$\rho_a = \frac{1}{\eta_a}\sum_i (q_i \pi_i)^{\otimes C}, \tag{4a}$$

$$\rho_b = \frac{1}{\eta_b}\left(\sum_i q_i \pi_i\right)^{\otimes C} - \frac{\eta_a}{\eta_b}\rho_a, \tag{4b}$$

with *a priori* probabilities

$$\eta_a = \sum_i q_i^C \text{ and } \eta_b = 1 - \eta_a. \tag{4c}$$

Now, a POVM, which satisfies Eq. (2) also has

$$F_a \rho_b = 0 \text{ and } F_b \rho_a = 0, \tag{5}$$

and furthermore the probability of success (3) which has to be optimized can be rewritten as

$$P = \eta_a \operatorname{tr}(F_a \rho_a) + \eta_b \operatorname{tr}(F_b \rho_b). \tag{6}$$

These equations are characteristic for unambiguous state discrimination. Thus an optimal solution to the problem of unambiguous discrimination (UD) of ρ_a and ρ_b, which in addition satisfies Eq. (2), is also the optimal solution to the related problem of unambiguous state comparison. The task of optimal UD of mixed states has been studied in the literature [4,6–9].

III. STATE COMPARISON "TWO OUT OF TWO"

We first consider explicitly the most simple case of state comparison, namely, "two out of two" with the states subject to comparison being pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, both of which are vectors in a Hilbert space of any dimension. The two states may appear with arbitrary (but nonvanishing) *a priori* probabilities q_1 and q_2. The trivial cases where both states are colinear or orthogonal are not considered. Without loss of generality the phase between the two states can be chosen to be real, so that their overlap is determined by their relative angle ϑ,

$$\cos\vartheta := \langle\psi_1|\psi_2\rangle \in \,]0,1[\, . \tag{7}$$

We consider the related UD problem of the corresponding mixed states, which are according to Eqs. (4a)–(4c) given by

$$\rho_a = \frac{1}{\eta_a}(q_1^2|\psi_1\psi_1\rangle\langle\psi_1\psi_1| + q_2^2|\psi_2\psi_2\rangle\langle\psi_2\psi_2|), \tag{8a}$$

$$\rho_b = \frac{1}{2}(|\psi_1\psi_2\rangle\langle\psi_1\psi_2| + |\psi_2\psi_1\rangle\langle\psi_2\psi_1|), \tag{8b}$$

appearing with *a priori* probabilities

$$\eta_a = q_1^2 + q_2^2 \text{ and } \eta_b = 2q_1q_2. \tag{8c}$$

Note that $\eta_a \geqslant \eta_b$ always holds. In what follows, we construct an optimal solution of this related UD problem and then show that the POVM of this solution satisfies Eq. (2), thus providing an optimal solution of the unambiguous state comparison task.

A. Reduction to the nontrivial subspace

It has been shown by Raynal, Lütkenhaus, and van Enk [4] that the optimal UD of mixed states can be reduced to a subspace of the Hilbert space in such a way that the relevant density matrices, acting on the reduced space, have equal rank and their kernels form nonorthogonal subspaces, the intersection of which is zero. This is achieved in two reduction steps: In the *first reduction step*, the Hilbert space is reduced to its nontrivial part, removing that part of the Hilbert space, where no UD is possible at all. We will denote this reduced space as $\mathcal{H}$. It is given by the particular space, where

$$\mathcal{S}_{\rho_a} \cap \mathcal{S}_{\rho_b} = 0 \text{ and } \mathcal{K}_{\rho_a} \cap \mathcal{K}_{\rho_b} = 0 \tag{9}$$

holds. Here $\mathcal{K}_\rho$ is the kernel of ρ and $\mathcal{S}_\rho$ is its support, defined as the orthocomplement to the kernel [12]. Thus $\mathcal{H}$ contains only the direct sum of the support of ρ_a and ρ_b, i.e., $\mathcal{H} = \mathcal{S}_{\rho_a} \oplus \mathcal{S}_{\rho_b}$.

For our system, we have

$$\mathcal{S}_{\rho_a} = \operatorname{span}(|\psi_1\psi_1\rangle, |\psi_2\psi_2\rangle), \tag{10a}$$

$$\mathcal{S}_{\rho_b} = \operatorname{span}(|\psi_1\psi_2\rangle, |\psi_2\psi_1\rangle), \tag{10b}$$

which already satisfy $\mathcal{S}_{\rho_a} \cap \mathcal{S}_{\rho_b} = \{0\}$ due to the linear independence of $|\psi_1\rangle$ and $|\psi_2\rangle$. For the further calculation it is convenient to rewrite both supports in an appropriate basis of $\mathcal{H}$. Therefore consider complementary normalized vectors $|\bar\psi_1\rangle, |\bar\psi_2\rangle \in \operatorname{span}(|\psi_1\rangle, |\psi_2\rangle)$, which are in the same plane as $|\psi_1\rangle$ and $|\psi_2\rangle$, but orthogonal to the corresponding vector, i.e., $|\bar\psi_1\rangle \perp |\psi_1\rangle$ and $|\bar\psi_2\rangle \perp |\psi_2\rangle$. Then, an orthonormal basis of $\mathcal{H}$ is given by

$$|e_{1,2}\rangle = \frac{1}{\sqrt{2n_\pm}}(|\psi_1\psi_1\rangle \pm |\psi_2\psi_2\rangle), \tag{11a}$$

$$|e_{3,4}\rangle = \frac{1}{\sqrt{2n_\pm}}(|\bar\psi_1\bar\psi_2\rangle \pm |\bar\psi_2\bar\psi_1\rangle), \tag{11b}$$

with $n_\pm = \sqrt{1 \pm \cos^2\vartheta}$. In Eq. (11a), the + (−) sign refers to the index 1 (2) and in Eq. (11b) to 3 (4), respectively.

By this choice, one immediately has $\mathcal{K}_{\rho_a} = \operatorname{span}(|e_3\rangle, |e_4\rangle)$ and $|e_2\rangle \in \mathcal{K}_{\rho_b}$. Let us denote by $P_+ = |e_1\rangle\langle e_1| + |e_3\rangle\langle e_3|(P_- = |e_2\rangle\langle e_2| + |e_4\rangle\langle e_4|)$ the projector onto that subspace, which is symmetric (antisymmetric) under exchanging $|\psi_1\rangle$ and $|\psi_2\rangle$. Then, due to $|\psi_1\psi_2\rangle \in \mathcal{S}_{\rho_b}$,

$$|\gamma\rangle := \frac{\sqrt{2}}{n_+}P_+|\psi_1\psi_2\rangle = \frac{\sqrt{2}}{n_+}P_+|\psi_2\psi_1\rangle \in \mathcal{S}_{\rho_b} \tag{12}$$

must hold, where $|\gamma\rangle$ is normalized and has the components

$$|\langle e_1|\gamma\rangle| = \frac{2\cos\vartheta}{n_+^2} \text{ and } |\langle e_3|\gamma\rangle| = \frac{\sin^2\vartheta}{n_+^2}. \tag{13}$$

Since $P_- + P_+ = \mathbb{1}_{\mathcal{H}}$, the second spanning vector of $\mathcal{S}_{\rho_b}$ has to be $P_-|\psi_1\psi_2\rangle = -P_-|\psi_2\psi_1\rangle$. This vector, however, cannot have any component in direction of $|e_2\rangle \in \mathcal{K}_{\rho_b}$ and therefore has to be parallel to $|e_4\rangle$. Thus we finally write the nontrivial Hilbert space $\mathcal{H}$ as

$$\mathcal{H} = \mathcal{S}_{\rho_a} \oplus \mathcal{S}_{\rho_b} \equiv \text{span}(|e_1\rangle, |e_2\rangle) \oplus \text{span}(|\gamma\rangle, |e_4\rangle). \tag{14}$$

Due to the particular choice of basis, we further find $\mathcal{K}_{\rho_b} = \text{span}(|\gamma^\perp\rangle, |e_2\rangle)$, where $|\gamma^\perp\rangle$ is a normalized vector satisfying $|\gamma^\perp\rangle \perp |\gamma\rangle$ and $P_-|\gamma^\perp\rangle = P_-|\gamma\rangle \equiv 0$.

B. Optimal solution

In the *second reduction step* shown in Ref. [4], one reduces the space by those parts, which allow perfect UD. These parts are given by

$$\mathcal{K}_a^\cap := \mathcal{K}_{\rho_a} \cap \mathcal{S}_{\rho_b} \text{ and } \mathcal{K}_b^\cap := \mathcal{K}_{\rho_b} \cap \mathcal{S}_{\rho_a}. \tag{15}$$

The Hilbert space $\mathcal{H}$ can then be decomposed into

$$\mathcal{H} = \mathcal{H}' \oplus \mathcal{K}_a^\cap \oplus \mathcal{K}_b^\cap, \tag{16}$$

where $\mathcal{H}'$ is conveniently chosen to be the orthocomplement of $\mathcal{K}_a^\cap \oplus \mathcal{K}_b^\cap$. Denoting by $P_{\mathcal{H}'}$ the projector onto $\mathcal{H}'$, and further writing ζ_a, ζ_b for appropriate normalization constants, the density matrices

$$\rho_a' = \frac{1}{\zeta_a}P_{\mathcal{H}'}\rho_a P_{\mathcal{H}'} \text{ and } \rho_b' = \frac{1}{\zeta_b}P_{\mathcal{H}'}\rho_b P_{\mathcal{H}'} \tag{17}$$

are states acting on $\mathcal{H}'$ and having *a priori* probabilities

$$\eta_a' = \frac{\eta_a\zeta_a}{\zeta} \text{ and } \eta_b' = 1 - \eta_a', \tag{18}$$

where $\zeta := \zeta_a\eta_a + \zeta_b\eta_b$. Suppose that P' is the optimal rate of success for this reduced problem. Then the optimal rate of success of the complete problem was shown [4] to be

$$P_{\text{opt}} = 1 - (1 - P')\zeta. \tag{19}$$

In our basis, we immediately find

$$\mathcal{K}_a^\cap = \text{span}(|e_3\rangle, |e_4\rangle) \cap \mathcal{S}_{\rho_b} = \text{span}(|e_4\rangle), \tag{20a}$$

$$\mathcal{K}_b^\cap = \text{span}(|\gamma^\perp\rangle, |e_2\rangle) \cap \mathcal{S}_{\rho_a} = \text{span}(|e_2\rangle), \tag{20b}$$

since $|\gamma\rangle \nparallel |e_3\rangle$ and $|\gamma^\perp\rangle \nparallel |e_1\rangle$ holds.

Now the optimization problem can be reduced to $\mathcal{H}' = \text{span}(|e_1\rangle, |e_3\rangle)$. Since the remaining problem is two dimensional, it can be considered as the well-known discrimination of pure states. Indeed, the problem reduces to the UD of

$$\rho_a' = \frac{1}{\zeta_a}P_+\rho_a P_+ = |e_1\rangle\langle e_1|, \tag{21a}$$

$$\rho_b' = \frac{1}{\zeta_b}P_+\rho_b P_+ = |\gamma\rangle\langle\gamma|. \tag{21b}$$

Calculating the normalization factors $\zeta_a = \text{tr}(P_+\rho_a)$ and $\zeta_b = \text{tr}(P_+\rho_b)$, one obtains $\zeta_a = \zeta_b = \zeta = \frac{1}{2}n_+^2$ and thus the *a priori* probabilities of the reduced problem remain unchanged, $\eta_a' \equiv \eta_a$ and $\eta_b' \equiv \eta_b$. Jaeger and Shimony have derived [10] the optimal UD of two pure states with an unbalanced probability distribution. Using their result for the discrimination between $|e_1\rangle$ and $|\gamma\rangle$, the optimal rate of success for UD of ρ_a and ρ_b calculates to

$$P_{\text{opt}} = \begin{cases} 1 - 2\sqrt{\eta_a\eta_b}\cos\vartheta & \text{if } (\mathcal{C}_1) \\ \dfrac{n_-^2}{n_+^2}\left(1 - \dfrac{\eta_b}{2}\sin^2\vartheta\right) & \text{else,} \end{cases} \tag{22}$$

where $(\mathcal{C}_1)$ is the condition

$$\cos\vartheta < \sqrt{\frac{\eta_a}{\eta_b}}\left(1 - \sqrt{\frac{\eta_a - \eta_b}{\eta_a}}\right). \tag{22'}$$

Further, the optimal POVM of the reduced problem is given by

$$F_a' = \alpha|\gamma^\perp\rangle\langle\gamma^\perp| \text{ and } F_b' = \beta|e_3\rangle\langle e_3|. \tag{23}$$

In the region, where (22′) holds,

$$\alpha = \frac{1 - \sqrt{\frac{\eta_b}{\eta_a}}|\langle e_1|\gamma\rangle|}{|\langle e_3|\gamma\rangle|^2}, \tag{24a}$$

$$\beta = \frac{1 - \sqrt{\frac{\eta_a}{\eta_b}}|\langle e_1|\gamma\rangle|}{|\langle e_3|\gamma\rangle|^2}, \tag{24b}$$

and $\alpha = 1$, $\beta = 0$ elsewhere. The optimal measurement of the full problem is then given by

$$F_a = F_a' + P_{\mathcal{K}_b^\cap} \text{ and } F_b = F_b' + P_{\mathcal{K}_a^\cap}, \tag{25}$$

where $P_{\mathcal{K}_b^\cap} \equiv |e_2\rangle\langle e_2|$ and $P_{\mathcal{K}_a^\cap} \equiv |e_4\rangle\langle e_4|$. The fact that the projectors $|e_2\rangle\langle e_2|$ and $|e_4\rangle\langle e_4|$ have to be part of the optimal POVMs F_a and F_b, respectively, was already obvious from the structure of the kernels and supports, since $|e_2\rangle$ and $|e_4\rangle$ are orthogonal and part of either $\mathcal{S}_{\rho_a}$ or $\mathcal{S}_{\rho_b}$.

Now one easily verifies that condition (2) holds for this measurement, by noting that $|\langle\psi_1\psi_1|e_2\rangle|^2 = |\langle\psi_2\psi_2|e_2\rangle|^2 > 0$ and $|\langle\psi_1\psi_2|e_4\rangle|^2 = |\langle\psi_2\psi_1|e_4\rangle|^2 > 0$. Thus we have found the optimal solution for unambiguous two-dimensional state comparison. Furthermore, as we discuss in the following, this solution is *always* better then a separable measurement on both states, which becomes manifest by the fact that F_a and F_b are not separable, i.e., the partial transpose fails to be positive semidefinite.

C. Discussion

In the literature, an optimal solution for the problem of state comparison has only been found for the case of equal

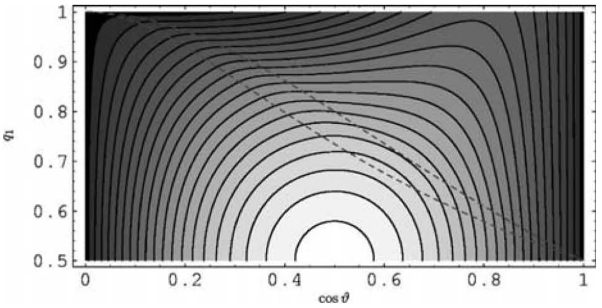

FIG. 1. Contour plot of the gain $P_{\rm opt}-P_{\rm sep}$, where higher gain corresponds to brighter shade. White stands for a gain value of 0.25, black for a value of 0.0125, and each contour line corresponds to a step of 0.0125. The dashed lines divide the set of parameters into regions where both (22′) and (27′) hold (lower left), neither of both condition holds (top right) and (27′) holds, but (22′) does not (remaining small stripe).

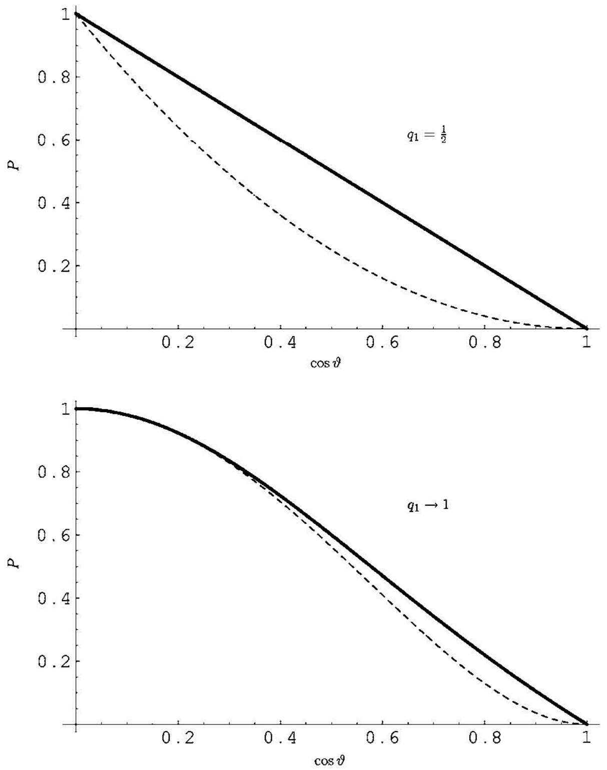

FIG. 2. Rate of success for state comparison "two out of two" with $q_1=\frac{1}{2}$ (upper graph) and $q_1 \to 1$ (lower graph). The solid line is the optimal result, and the dashed line corresponds to the best separable measurement.

probabilities. Barnett, Chefles, and Jex [1] showed that in this case the optimal rate of success is given by $P=1-\cos\vartheta$, which is our result for $q_1=q_2=\frac{1}{2}$. This particular result was also obtained by Rudolph, Spekkens, and Turner [7], by providing a general upper and lower bound for the rate of success of an UD of mixed states. Their upper bound matches our result only in situations where (22′) holds. On the other hand, their lower bound turns out to match our optimal result for all parameters and thus our calculation has proven that their lower bound is indeed optimal for the UD of ρ_a and ρ_b.

Let us compare our result with the naïve incoherent strategy, where both states are measured consecutively. The straightforward approach of the optimal POVM $\{\widetilde{F}_1,\widetilde{F}_2,\widetilde{F}_?\}$ for unambiguous discrimination between $|\psi_1\rangle$ and $|\psi_2\rangle$ leads to

$$F_a^{\rm sep} = \widetilde{F}_1 \otimes \widetilde{F}_1 + \widetilde{F}_2 \otimes \widetilde{F}_2, \tag{26a}$$

$$F_b^{\rm sep} = \widetilde{F}_1 \otimes \widetilde{F}_2 + \widetilde{F}_2 \otimes \widetilde{F}_1. \tag{26b}$$

This naïve method is indeed the *optimal* separable measurement, as shown in the Appendix. It has a rate of success given by the square of the success probability for unambiguous discrimination of $|\psi_1\rangle$ and $|\psi_2\rangle$, i.e., [10],

$$P_{\rm sep} = \begin{cases} (1-2\sqrt{q_1 q_2}\cos\vartheta)^2 & \text{if } (C_2) \\ q_{\rm max}^2 \sin^4\vartheta & \text{else,} \end{cases} \tag{27}$$

where $q_{\rm max}$ is the maximum of q_1 and q_2, and (C_2) is the condition

$$\cos\vartheta < \sqrt{\frac{1-q_{\rm max}}{q_{\rm max}}}. \tag{27′}$$

In Fig. 1 we show the gain $P_{\rm opt}-P_{\rm sep}$, which of course is always positive or zero. This gain has its absolute maximum of $\frac{1}{4}$ at $q_1=\frac{1}{2}$ and $\vartheta=\pi/3$. While for fixed angles the maximum gain is always at $q_1=\frac{1}{2}$, one finds for fixed *a priori* probabilities that at some regions there are two maxima. The maximum in low values of cos ϑ appears, where (27′) holds without having (22′) satisfied. Also note that the gain function is asymmetric in cos ϑ, while it is symmetric in q_1. In Fig. 2, the gain of the coherent vs the incoherent strategy is illustrated for the parameters $q_1=\frac{1}{2}$ and $q_1 \to 1$.

IV. STATE COMPARISON "TWO OUT OF *N*"

Next, we investigate the problem of unambiguous state comparison "two out of *N*" for pure states. As shown by Chefles *et al.* [2] for equal probabilities and in Sec. V for arbitrary probabilities, this can only work if all *N* states are linearly independent, thus spanning an *N*-dimensional Hilbert space. Again this unambiguous state comparison is related to the UD of

$$\rho_a = \frac{1}{\eta_a}\sum_i^N q_i^2 |\psi_i\psi_i\rangle\langle\psi_i\psi_i|, \tag{28a}$$

$$\rho_b = \frac{1}{\eta_b}\sum_{i\neq j}^N q_i q_j |\psi_i\psi_j\rangle\langle\psi_i\psi_j|, \tag{28b}$$

having *a priori* probabilities

$$\eta_a = \sum q_i^2, \quad \eta_b = \sum_{i\neq j} q_i q_j. \tag{28c}$$

We immediately obtain

$$\mathcal{S}_{\rho_a} = \bigoplus_{i} \mathrm{span}(|\psi_i\psi_i\rangle), \tag{29a}$$

$$\mathcal{S}_{\rho_b} = \bigoplus_{i\neq j} \mathrm{span}(|\psi_i\psi_j\rangle) = \bigoplus_{i>j} \mathrm{span}(|\psi_i\psi_j\rangle \pm |\psi_j\psi_i\rangle). \tag{29b}$$

Due to linear independence $\mathcal{S}_{\rho_a}\cap\mathcal{S}_{\rho_b}=\{0\}$ holds and thus the first reduction step yields $\mathcal{H}=\mathcal{S}_{\rho_a}\oplus\mathcal{S}_{\rho_b}$. Note that the dimension of $\mathcal{S}_{\rho_a}$ is now in general much smaller than the one of $\mathcal{S}_{\rho_b}$, because $\dim\mathcal{S}_{\rho_a}=N$ while $\dim\mathcal{S}_{\rho_b}=N^2-N$. In what follows we show in a constructive way that the N-dimensional state comparison in general is related to such an UD of mixed states, which cannot be reduced to UD of pure states.

The second reduction step can be performed as follows. The antisymmetric subspace $\mathcal{H}^-=\oplus_{i>j}\mathrm{span}(|\psi_i\psi_j\rangle-|\psi_j\psi_i\rangle)$ is part of $\mathcal{K}_a^\cap\equiv\mathcal{K}_{\rho_a}\cap\mathcal{S}_{\rho_b}$, since

$$\mathcal{S}_{\rho_a}\perp\mathcal{H}^- \text{ and } \mathcal{S}_{\rho_b}\supset\mathcal{H}^-. \tag{30}$$

Further, $\mathcal{S}_{\rho_a}$ is part of the symmetric subspace $\mathcal{H}^+=\oplus_{i\geq j}\mathrm{span}(|\psi_i\psi_j\rangle+|\psi_j\psi_i\rangle)$ and thus, due to $\mathcal{H}^-\perp\mathcal{H}^+$, we have the orthogonal decomposition

$$\mathcal{K}_a^\cap=\mathcal{K}_a^{\cap^-}\oplus\mathcal{K}_b^{\cap^+}, \tag{31}$$

with $\mathcal{K}_a^{\cap-}:=\mathcal{H}^-$ and $\mathcal{K}_a^{\cap+}:=\mathcal{H}^+\cap\mathcal{K}_{\rho_a}$. In order to obtain $\mathcal{K}_a^{\cap+}$, let $C_{ij}:=\langle\psi_i|\psi_j\rangle$ be the Hermitian overlap matrix and A_{ij} be a lower triangular coefficient matrix. Then $\mathcal{K}_a^{\cap+}$ is given by all vectors $\Sigma_{i>j}A_{ij}(|\psi_i\psi_j\rangle+|\psi_j\psi_i\rangle)$, which satisfy

$$\forall\ k\ \langle\psi_k\psi_k|\sum_{i>j}A_{ij}(|\psi_i\psi_j\rangle+|\psi_j\psi_i\rangle)=0,$$

$$\Leftrightarrow\forall\ k\sum_{i>j}C_{ki}A_{ij}C_{kj}=0,$$

$$\Leftrightarrow\forall\ k\,[CAC^T]_{kk}=0. \tag{32}$$

This set of *linear* equations may eliminate up to N out of $N(N-1)/2$ coefficients A_{ij}, thus

$$\frac{N(N-1)}{2}\geq\dim(\mathcal{K}_a^{\cap^+})\geq\max\left\{\frac{N(N-3)}{2},0\right\}. \tag{33}$$

The space $\mathcal{K}_b^\cap\equiv\mathcal{K}_{\rho_b}\cap\mathcal{S}_{\rho_a}$ on the other hand is given by all vectors out of $\mathcal{S}_{\rho_a}$, which are orthogonal to $|\psi_i\psi_j\rangle+|\psi_j\psi_i\rangle$ for all $i>j$. With a diagonal coefficient matrix B this yields

$$\forall\ i>j\quad[CBC^T]_{ij}=0. \tag{34}$$

Thus, we have

$$N\geq\dim\mathcal{K}_b^\cap\geq\max\left\{\frac{N(3-N)}{2},0\right\}. \tag{35}$$

Since the dimension of the reduced Hilbert space is given as $\dim\mathcal{H}'=\dim\mathcal{H}-(\dim\mathcal{K}_a^{\cap-}+\dim\mathcal{K}_a^{\cap+})-\dim\mathcal{K}_b^\cap$, we finally arrive at the main result of this section,

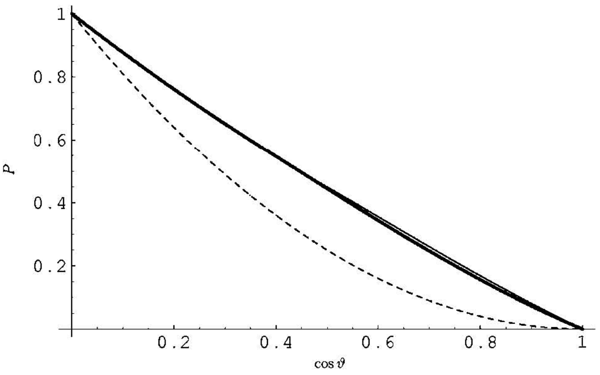

FIG. 3. Bounds for the probability of success for state comparison "two out of three," with equal *a priori* probabilities and relative angles. The solid lines are an upper [9] and a lower bound [7], while the dashed line corresponds to the separable measurement.

$$0\leq\dim\mathcal{H}'\leq\begin{cases}2 & \text{if } N=2\\ 2N & \text{if } N>2.\end{cases} \tag{36}$$

The case $N=2$, considered in Sec. III, turns out to play a special role, since here always $\dim\mathcal{K}_b^\cap>0$ holds, cf. Eq. (35). We point out that these bounds are tight. This can be directly verified by considering a system of states with equal overlap, i.e., a system with

$$\cos\vartheta:=\langle\psi_i|\psi_j\rangle\in[0,1[\quad\forall\ i\neq j. \tag{37}$$

Then for the trivial case (i.e., $\vartheta=\pi/2$) $\dim\mathcal{H}'=0$ holds, while the upper bound is reached whenever $\vartheta<\pi/2$. Thus state comparison for two out of three states may already lead to a nontrivial UD problem, as illustrated in the following.

A. Example: "Two out of three"

As an example of a case, where state comparison does not reduce to UD of pure states, $N=3$ is considered. We specialize to the case where the states $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\psi_3\rangle$ subject to comparison satisfy Eq. (37) with $0<\vartheta<\pi/2$ and assume all *a priori* probabilities to be equal, $q_1=q_2=q_3=\frac{1}{3}$.

The previous discussion of the related UD problem showed that this related problem can be reduced to a Hilbert space $\mathcal{H}'$ of dimension $\dim\mathcal{H}'=\dim\mathcal{S}_{\rho_a}+\dim\mathcal{S}_{\rho_b}=3+3$. Since $N=3$ this has the consequence that $\mathcal{K}_a^{\cap+}=\mathcal{K}_b^\cap=\{0\}$. Thus $\mathcal{H}'$ exactly consists of the symmetric subspace of $\mathcal{H}\equiv\mathcal{S}_{\rho_a}\oplus\mathcal{S}_{\rho_b}$, i.e., $\mathcal{H}'=\mathcal{H}^+$. However, for the remaining UD problem, no general optimal solution is known and we thus calculate the tightest upper and lower bounds for the rate of success known so far, i.e., the lower bound provided by Rudolph *et al.* [7] and the upper bound shown by Raynal and Lütkenhaus [9]. These bounds together with the rate of success for the separable measurement are shown in Fig. 3. Again, the incoherent measurement is always worse than the measurement used to construct the lower bound. In addition one finds that for

$$\cos\vartheta\leq\frac{\sqrt{2}-\sqrt{\sqrt{2}}}{2-\sqrt{2}} \tag{38}$$

(i.e., $\vartheta/\pi\gtrsim0.375$) the lower and upper bound coincide, revealing the *optimal* solution of UD of ρ_a and ρ_b in that region to be

$$P_{\text{opt}} = 1 - \frac{\sqrt{8}}{9}(4 \cos \vartheta - \cos^2 \vartheta). \tag{39}$$

One can also show that in this region the optimal measurement satisfies the defining property (2) and thus also solves the problem of optimal state comparison.

V. MIXED-STATE COMPARISON

In this section we investigate in what situations a measurement can exist, which satisfies the defining property (2). We have the following:

Proposition 1. Unambiguous state comparison "C out of N" for a set of mixed states $\{\pi_1, \ldots, \pi_N\}$ with arbitrary but nonvanishing *a priori* probabilities can be realized if and only if $\forall\, i$,

$$\mathcal{S}_{\pi_i} \not\subset \sum_{k \neq i} \mathcal{S}_{\pi_k}. \tag{40}$$

Proof. For the *if* part it is enough to show that there is a POVM, given by $\{\widetilde{F}_1, \ldots, \widetilde{F}_N, \widetilde{F}_?\}$, such that

$$\text{tr}(\widetilde{F}_i \pi_j) > 0 \Leftrightarrow i = j. \tag{41}$$

In order to construct such a POVM, denote by P_i the projector onto the orthocomplement of $\Sigma_{k\neq i}\mathcal{S}_{\pi_k}$. Then from Eq. (40) it follows that there is at least one vector $|\varphi\rangle \in \mathcal{S}_{\pi_i}$, such that $|\phi_i\rangle := P_i|\varphi\rangle$ satisfies $\langle\phi_i|\phi_i\rangle = 1$. These vectors $|\phi_i\rangle$ by construction satisfy $\langle\phi_i|\pi_i|\phi_i\rangle > 0$ for each i, while $\langle\phi_i|\pi_j|\phi_i\rangle = 0$ for all $j \neq i$. The choice $\widetilde{F}_i = (1/N)|\phi_i\rangle\langle\phi_i|$ satisfies (41) and further has $\widetilde{F}_? = \mathbb{1} - \Sigma_i \widetilde{F}_i \geq 0$. Indeed, for any $|\psi\rangle$ out of the complete Hilbert space,

$$\langle\psi|\widetilde{F}_?|\psi\rangle = \langle\psi|\psi\rangle - \frac{1}{N}\sum_i |\langle\phi_i|\psi\rangle|^2 \geq 0 \tag{42}$$

holds by virtue of the Cauchy-Schwarz inequality.

For the *only if* part we use that any unambiguous state comparison measurement solves (not necessarily in an optimal way) the related unambiguous state discrimination problem. However, assuming that for some i

$$\mathcal{S}_{\pi_i} \subset \sum_{k \neq i} \mathcal{S}_{\pi_k}, \tag{43}$$

we show, that no UD measurement can satisfy $\text{tr}(F_a \pi_i^{\otimes C}) > 0$, thus being a contradiction to Eq. (2a).

In order to show this contradiction, note that for positive operators A and B,

$$\mathcal{S}_{A+B} = \mathcal{S}_A + \mathcal{S}_B, \tag{44a}$$

$$\mathcal{S}_{A\otimes B} = \mathcal{S}_A \otimes \mathcal{S}_B. \tag{44b}$$

Further we use a lemma, shown by Raynal, Lütkenhaus, and van Enk in Ref. [4], which states that $\text{tr}(AB) = 0$, if and only if $\mathcal{S}_A \perp \mathcal{S}_B$. Now, assuming Eq. (43), it follows that

$$\mathcal{S}_{\pi_i^{\otimes C}} = \mathcal{S}_{\pi_i}^{\otimes C} \subset \sum_{k\neq i} \mathcal{S}_{\pi_k} \otimes \mathcal{S}_{\pi_i}^{\otimes(C-1)} \subset \mathcal{S}_{\rho_b}. \tag{45}$$

However, by the Lemma of Ref. [4], the requirement $\text{tr}(F_a\rho_b) = 0$ [cf. Eq. (5)] is equivalent to $\mathcal{S}_{F_a} \perp \mathcal{S}_{\rho_b}$. This implies $\mathcal{S}_{F_a} \perp \mathcal{S}_{\pi_i^{\otimes C}}$ or equivalently $\text{tr}(F_a \pi_i^{\otimes C}) = 0$ and completes the proof. ■

For the comparison of qubits this proposition implies that unambiguous comparison "C out of N" can only be realized for $N=2$ and *pure states*. For unambiguous state comparison "C out of N" of pure states in any dimension, Proposition 1 reduces to the result of Chefles *et al.* [2]. They found that state comparison can only be realized for linearly independent states. Another direct consequence from Proposition 1 is the fact that density matrices which contain a proportion of the identity (e.g., by being sent through a depolarizing channel, or by adding white noise in an experiment) can never be compared unambiguously.

VI. CONCLUSIONS

We have addressed the question of unambiguous state comparison with general *a priori* probabilities. Our method consists of reducing the corresponding problem of unambiguous mixed-state discrimination to a nontrivial subspace [4]. We analytically solve the case for comparing two states drawn from a set of two states, finding the optimal POVMs and the optimal rate of success. There is a considerable gain of the optimal coherent strategy over the best incoherent strategy. While this case reduces to the discrimination between two pure states, the comparison of two states drawn from a set of three states is shown to lead to a nontrivial mixed-state discrimination task. So far, the optimal solution is only found for certain parameter ranges.

The more general task of comparing two states from a set of N states is exceedingly difficult. No general solution to this problem exists. Here, we have presented an upper bound for the dimension of the reduced Hilbert space. This bound is shown to be reached for states with equal overlap. We have also provided a necessary and sufficient condition for unambiguous comparison of mixed states to be possible.

Note added: Recently, we learned about related work by Herzog and Bergou [11], who found the same expression as Eq. (22) for optimal unambiguous state comparison of two states drawn from a set of two states.

ACKNOWLEDGMENTS

We would like to thank N. Lütkenhaus, T. Meyer, and P. Raynal for enlightening discussions. This work was partially supported by the EC programme SECOQC.

APPENDIX: OPTIMAL SEPARABLE MEASUREMENT "TWO OUT OF TWO"

This appendix is dedicated to show that with the naïve measurement given in Eq. (26), indeed the optimal separable solution was found. That is, the optimal *separable* unambiguous state comparison measurement for two states drawn from a set of two pure states $\{|\psi_1\rangle, |\psi_2\rangle\}$ is solved in an optimal way by performing optimal unambiguous state discrimination in each subsystem.

A general element of a separable POVM $\{F_x\}$ is of the form

$$F_x = \sum_{i,j} c_{x,ij} F_{x,i}^{(1)} \otimes F_{x,j}^{(2)}, \tag{A1}$$

where the non-negative coefficients $c_{x,ij}$ account for the relative contribution of each of the terms containing the positive local POVM elements $F_{x,i}^{(k)}$.

First we show that in our case no measurement outcome of either subsystem can be used to adapt the measurement of the other. Consider without loss of generality that a measurement first takes place in subsystem 1 and yields with probability $p_{x,i}^{(1)}$ the outcome (x,i). This measurement is applied to the global state $\rho := \eta_a \rho_a + \eta_b \rho_b = (q_1 |\psi_1\rangle\langle\psi_1| + q_2 |\psi_2\rangle\langle\psi_2|)^{\otimes 2}$ and yields in subsystem 2

$$\mathrm{tr}_1[(F_{x,i}^{(1)} \otimes \mathbb{1})\rho] = p_{x,i}^{(1)}(q_1|\psi_1\rangle\langle\psi_1| + q_2|\psi_2\rangle\langle\psi_2|), \tag{A2}$$

which is, up to the factor $p_{x,i}^{(1)}$, independent of the outcome (x,i). Thus the local measurements can be optimized in each subsystem separately, and one is free to choose the same (optimal) measurement in both systems due to the symmetry of ρ_a and ρ_b. Therefore we can drop the upper label (k) on the local measurement elements in the following.

Furthermore, one is forced to choose these measurements to be UD measurements. Indeed, $\mathrm{tr}(F_a\rho_b) = \mathrm{tr}(F_b\rho_a) = 0$, only if for each $x \in \{a,b\}$ and for all l, either $\mathrm{tr}(F_{x,l}|\psi_1\rangle\langle\psi_1|) = 0$ or $\mathrm{tr}(F_{x,l}|\psi_2\rangle\langle\psi_2|) = 0$. We prove this statement by contradiction: Suppose that at least one term $(c_{a,ij}F_{a,i} \otimes F_{a,j})$ of F_a contains at least one local POVM element $F_{a,m}$ (where $m \in \{i,j\}$), having a nonvanishing expectation value for both states, i.e.,

$$\langle\psi_1|F_{a,m}|\psi_1\rangle > 0 \text{ and } \langle\psi_2|F_{a,m}|\psi_2\rangle > 0. \tag{A3}$$

It follows that

$$\mathrm{tr}[(c_{x,ij}F_{a,i} \otimes F_{a,j})\rho_a] > 0 \tag{A4}$$

and

$$\mathrm{tr}[(c_{x,ij}F_{a,i} \otimes F_{a,j})\rho_b] > 0, \tag{A5}$$

which is in which is in contradiction to $\mathrm{tr}(F_a\rho_b) = 0$. An analogous argument holds for F_b.

Without losing any information, an UD measurement can always be reduced to have the measurement elements $\{F_1, F_2, F_?\}$, with $\langle\psi_2|F_1|\psi_2\rangle = \langle\psi_1|F_2|\psi_1\rangle = 0$. In order to make this a valid choice for the local measurements of unambiguous state comparison, in addition the conditions (2) have to be satisfied, i.e.,

$$\alpha := \langle\psi_1|F_1|\psi_1\rangle > 0 \text{ and } \beta := \langle\psi_2|F_2|\psi_2\rangle > 0. \tag{A6}$$

From the consideration above, we find that F_a and F_b are of the form

$$F_a = F_1 \otimes F_1 + F_2 \otimes F_2, \tag{A7}$$

$$F_b = F_1 \otimes F_2 + F_2 \otimes F_1. \tag{A8}$$

The optimal separable state comparison corresponds to $F_1 = \tilde{F}_1$ and $F_2 = \tilde{F}_2$ as defined in Eq. (26). Thus we have shown that in this case the optimal separable unambiguous state comparison strategy is indeed given by consecutive optimal UD measurements.

Let us mention that for the optimal UD measurement the conditions $\alpha > 0$ and $\beta > 0$ do not always hold: in those situations, where condition (27′) is not satisfied, $\alpha = 0$ or $\beta = 0$. But changing α and β (under the constraint $\mathbb{1} - F_1 - F_2 \geqslant 0$) infinitesimally, affects the probability of success only infinitesimally. In this limit, we consider the optimal unambiguous state discrimination measurement as a valid choice for F_1 and F_2.

We conjecture that also in the more general scenario of unambiguous state comparison of "C out of N" states, the best separable measurement is given by performing unambiguous state discrimination in each subsystem. However, the proof by contradiction given above for "two out of two" cannot be generalized in a straightforward way for the operator F_b. We leave the generalization as an open question for future work.

[1] S. M. Barnett, A. Chefles, and I. Jex, Phys. Lett. A **307**, 189 (2003).

[2] A. Chefles, E. Andersson, and I. Jex, J. Phys. A **37**, 7315 (2004).

[3] I. Jex, E. Andersson, and A. Chefles, J. Mod. Opt. **51**, 505 (2004).

[4] P. Raynal, N. Lütkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003).

[5] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1995).

[6] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).

[7] T. Rudolph, R. W. Spekkens, and P. S. Turner, Phys. Rev. A **68**, 010301(R) (2003).

[8] C. Zhang, Y. Feng, and M. Ying, quant-ph/0410073.

[9] P. Raynal and N. Lütkenhaus, quant-ph/0502165

[10] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).

[11] U. Herzog and J. Bergou, Phys. Rev. A **71**, 050301(R) (2005).

[12] Since a density matrix is Hermitian, in particular the support of a density matrix is identical to its range.

Publication B

Physical purification of quantum states
Physical Review A, Vol. **73**, p. 062309 (2006) *(5 pages)*
M. Kleinmann, H. Kampermann, T. Meyer, and D. Bruß

PHYSICAL REVIEW A **73**, 062309 (2006)

Physical purification of quantum states

M. Kleinmann,* H. Kampermann, T. Meyer, and D. Bruß
Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

(Received 16 September 2005; revised manuscript received 10 March 2006; published 8 June 2006)

We introduce the concept of a physical process that purifies a mixed quantum state, taken from a set of states, and investigate the conditions under which such a purification map exists. Here, a purification of a mixed quantum state is a pure state in a higher-dimensional Hilbert space, the reduced density matrix of which is identical to the original state. We characterize all sets of mixed quantum states, for which perfect purification is possible. Surprisingly, some sets of two noncommuting states are among them. Furthermore, we investigate the possibility of performing an imperfect purification.

DOI: 10.1103/PhysRevA.73.062309 PACS number(s): 03.67.−a, 03.65.−w

I. INTRODUCTION

A fundamental entity in quantum mechanics and quantum information is a mixed quantum state. A mixed quantum state can be either understood as a statistical mixture of pure quantum states, or as being part of a higher-dimensional, pure state—a *purification* of the mixed state. Formally, given the decomposition $\rho=\Sigma_i p_i|\chi_i\rangle\langle\chi_i|$, where $p_i\geqslant 0$ and $\Sigma_i p_i=1$, an example for a purification of ρ is given by $|\psi\rangle=\Sigma_i\sqrt{p_i}|\chi_i\rangle|a_i\rangle$, with the auxiliary states $|a_i\rangle$ being mutually orthogonal. This abstract point of view was, so far, the main impetus for discussing purifications of a single, *known* quantum state [1,2].

In this paper, we consider the purification of an *unknown* quantum state. More precisely, we introduce the fundamental question whether there exists a *physical process* (i.e., a completely positive map) that takes any state of a given set to one of its purifications. (We remind the reader for clarity that there exists a different notion of "purification" in the literature, referring to the process of performing operations on several identical copies of a given state, such that the purity of some of them is increased; a typical application is entanglement distillation.) Our aim is to characterize all sets of states for which a purifying map exists. The existence of such a process implies a nontrivial physical equivalence between certain sets of mixed and pure quantum states.

Let us introduce our concepts and outline the structure of this paper. As already pointed out above, a purification of a mixed state has to satisfy two characteristic properties: first, it has to be pure, and second, tracing out the auxiliary system has to yield back the original state. We call the second property *faithfulness* and name a process a *perfect purifier* for a mixed state, when the output achieves both properties. It is straightforward to prove that the linearity of quantum mechanics does not allow the existence of a perfect purifier for a completely unknown quantum state, i.e., a state taken from the set of *all* states. However, will dropping the condition of faithfulness *or* the one of purity allow nontrivial purification processes for an unknown quantum state? It will be shown in Theorem 1 that this is not the case. Consequently, in Sec. III we will restrict the set of possible input states, and investigate the properties of purifying maps acting on the most simple nontrivial set, namely a set of only two mixed states. While keeping the condition of purity, we will find that the deviation from perfect faithfulness depends on a purely geometric quantity of the two inputs. This result will allow us to derive lower and upper bounds on the achievable faithfulness. Since these bounds do not exclude perfect faithfulness for certain pairs of states, we then, in Sec. IV proceed to investigate the existence of a perfect purifier in general. Theorem 2 completely characterizes all sets of states that can be purified perfectly. Finally, we will provide an operational test for a given pair of states that allows to check whether a physical purification is possible.

II. THE GENERAL PURIFICATION TASK

In the following we will denote by $\mathcal{M}$ a given set of mixed states, represented by density operators that act on a finite-dimensional Hilbert space $\mathcal{H}$. The elements $\rho_i\in\mathcal{M}$ are allowed to have unbalanced *a priori* probabilities $\eta_i>0$, satisfying $\Sigma_i\eta_i=1$. We consider deterministic physical processes represented by completely positive and trace preserving [14] linear maps Λ that take any density operator acting on $\mathcal{H}$ to a density operator acting on $\mathcal{H}\otimes\mathcal{H}_{\text{aux}}$, where $\mathcal{H}_{\text{aux}}$ denotes an auxiliary space of unspecified dimension. We refer to such a physical process as a *perfect purifier* if for each $\rho_i\in\mathcal{M}$, the output $\Lambda[\rho_i]$ is pure as well as faithful, i.e., $\text{tr}_{\text{aux}}\Lambda[\rho_i]=\rho_i$. If these conditions are not met, we will measure the average output purity by $p=\Sigma_i\eta_i\,\text{tr}\,\Lambda[\rho_i]^2$ and the average faithfulness by $f=1-\Sigma_i\eta_i\|\rho_i-\text{tr}_{\text{aux}}\,\Lambda[\rho_i]\|$. Here, $\|\rho-\sigma\|=\frac{1}{2}\text{tr}|\rho-\sigma|$ denotes the trace distance, where $|A|=\sqrt{A^\dagger A}$. The trace distance is a good measure for the distinguishability of two states as it vanishes for identical states and is equal to one for orthogonal states. In particular, the success probability for the minimum error discrimination procedure [3,4] of two states having equal *a priori* probability depends linearly on the trace distance of the states. We call any deterministic process a *purifier* of $\mathcal{M}$, if it does not decrease the average purity of $\mathcal{M}$.

For the universal case where the set $\mathcal{M}$ contains all possible density operators acting on a given Hilbert space, neither relaxing the condition of purity nor relaxing the condition of faithfulness allows nontrivial purifiers.

Theorem 1. (i) Any universal purifier with perfect output

*Electronic address: kleinmann@thphy.uni-duesseldorf.de purity is a constant map. (ii) A universal purifier with perfect faithfulness does not increase the purity of any state.

Proof. We prove (i) by contradiction. Suppose there exists a purifier Λ such that $\Lambda[\rho]$ is pure for any state ρ, and with the property that at least for two states ρ_1 and ρ_2, $\Lambda[\rho_1] \neq \Lambda[\rho_2]$ holds. But for the state $\rho_3 = (\rho_1 + \rho_2)/2$, the purity of $\Lambda[\rho_3] = (\Lambda[\rho_1] + \Lambda[\rho_2])/2$ requires $\Lambda[\rho_1] = \Lambda[\rho_2]$.

Proof of statement (ii): perfect faithfulness of a universal purifier requires that *any* pure state $|\phi\rangle\langle\phi|$ is mapped onto the state $|\phi\rangle\langle\phi| \otimes \sigma_\phi$ for some state σ_ϕ acting on $\mathcal{H}_{\mathrm{aux}}$. For any state ρ we find with the spectral decomposition $\rho = \Sigma p_i |\lambda_i\rangle\langle\lambda_i|$ that due to linearity $\mathrm{tr}\,\Lambda[\rho]^2 = \mathrm{tr}(\Sigma_i p_i |\lambda_i\rangle\langle\lambda_i| \otimes \sigma_{\lambda_i})^2 = \Sigma_i p_i^2 \mathrm{tr}\,\sigma_{\lambda_i}^2 \leqslant \Sigma_i p_i^2 = \mathrm{tr}\,\rho^2$, i.e., no state can become purer by the action of Λ. □

Let us mention that there is some similarity of the arguments given in the proof above with the no-cloning theorem [5–7]. In both scenarios, linearity of quantum mechanics forbids the existence of some physical process, when the input set contains *all* states. Even when the set of input states is restricted to two pure states, perfect quantum cloning is impossible, as follows from unitarity. It was furthermore shown that broadcasting (a natural generalization of quantum cloning to mixed input states) is possible for a set of two mixed states, if and only if the states commute [8]. The same criterion does *not* apply for purification maps: a pair of orthogonal or identical states can, of course, be purified perfectly—but in any other case of commuting states we will show that perfect purification is impossible. Yet for some noncommuting states, a perfect purification process exists.

III. TWO-STATE PURIFIERS WITH PURE OUTPUT

In this section we will focus on the case of two input states and perfect output purity, i.e., a deterministic process which takes any state from the set $\mathcal{M} = \{\rho, \rho'\}$ to a pure state. A characteristic quantity for purification will turn out to be the *worst-case distinguishability* $\mathcal{D}(\rho, \rho')$, which denotes the trace distance of the two closest states that may appear physically in the ensembles of ρ and ρ', i.e.,

$$\mathcal{D}(\rho, \rho') = \min_{|\chi\rangle, |\chi'\rangle} \| |\chi\rangle\langle\chi| - |\chi'\rangle\langle\chi'| \|, \tag{1}$$

where $|\chi\rangle$ and $|\chi'\rangle$ are normalized vectors in the range of ρ and ρ', respectively. (We point out that this quantity can be calculated by taking the sine of the smallest canonical angle [9] between the range of ρ and the range of ρ'.) The notion of distinguishability here refers to the success probability of a minimum error discrimination, as explained above.

Although at first sight the worst-case distinguishability resembles a distance, mathematically speaking it is none: The triangular inequality does not hold, and $\mathcal{D}(\rho, \rho') = 0$ is true for some $\rho \neq \rho'$. Note that any two states with overlapping ranges have, in fact, a vanishing worst-case distinguishability. On the other hand, $\mathcal{D}(\rho, \rho') = 1$ is equivalent to ρ and ρ' being orthogonal, i.e., $\|\rho - \rho'\| = 1$. Thus commuting states are either orthogonal or have a vanishing worst-case distinguishability.

A. Characterization of two-state purifiers

We are now in the position to study the general consequences of perfect output purity. Suppose that Λ is a purifier of ρ and ρ' with perfect output purity. As a defining property of any normalized vector $|\chi\rangle$ in the range of ρ one can write $\rho = \alpha |\chi\rangle\langle\chi| + \beta \tilde{\rho}$ with positive numbers α and β, and positive semidefinite $\tilde{\rho}$. Using the same convexity argument as in the proof of Theorem 1 (i), it follows that $\Lambda[|\chi\rangle\langle\chi|] = \Lambda[\rho]$. An analogous argument holds for all vectors $|\chi'\rangle$ in the range of ρ'. Thus we have $\|\Lambda[\rho] - \Lambda[\rho']\| = \|\Lambda[|\chi\rangle\langle\chi|] - \Lambda[|\chi'\rangle\langle\chi'|]\| \leqslant \| |\chi\rangle\langle\chi| - |\chi'\rangle\langle\chi'| \|$, where in the inequality we used that a deterministic physical process Λ cannot increase the trace distance between two states [10]. By choosing for $|\chi\rangle\langle\chi|$ and $|\chi'\rangle\langle\chi'|$ the states with minimal distance [cf. definition in Eq. (1)], we have shown that for maps Λ where $\Lambda[\rho]$ as well as $\Lambda[\rho']$ are pure,

$$\mathcal{D}(\rho, \rho') \geqslant \|\Lambda[\rho] - \Lambda[\rho']\| \tag{2}$$

must hold.

It is important that there always exists a map which reaches equality in Eq. (2). In order to see this, one constructs a canonical basis [9] of the ranges of both states, i.e., an orthonormal basis $\{|\chi_i\rangle\}$ of the range of ρ and $\{|\chi_i'\rangle\}$ of the range of ρ', such that in addition $\langle\chi_i|\chi_j'\rangle = 0$ holds for all $i \neq j$. One can show that there always exists a map, which decreases the distance of two pure states by an arbitrary value. Such a map is now applied in each of the orthogonal subspaces spanned by $\{|\chi_i\rangle, |\chi_i'\rangle\}$, such that the distance $\| |\chi_i\rangle\langle\chi_i| - |\chi_i'\rangle\langle\chi_i'| \|$ decreases to be $\mathcal{D}(\rho, \rho')$. The composed map has the property, that if applied to ρ and ρ', an orthonormal eigenbasis for both output states exists, such that all nonorthogonal eigenvectors (one of the output of ρ and one of ρ') have a distance $\mathcal{D}(\rho, \rho')$. Now a map can readily be found, which maps the output states to pure states having a distance $\mathcal{D}(\rho, \rho')$. The fact that one can always reach the equality in Eq. (2) completes the characterization of the output of a general process, which maps two input states ρ and ρ' to two pure states.

B. Bounds on two-state purifiers

As an application of the result in Sec. III A we now estimate the faithfulness of a purifier with perfect output in the case of two input states. For this purpose we assume that the state ρ (ρ') occurs with *a priori* probability η (η'), where $\eta' \geqslant \eta$ without loss of generality. We denote the deviation from perfect faithfulness by δ, i.e.,

$$\delta = \eta \|\rho - \mathrm{tr}_{\mathrm{aux}} \Lambda[\rho]\| + \eta' \|\rho' - \mathrm{tr}_{\mathrm{aux}} \Lambda[\rho']\|. \tag{3}$$

Using the triangular inequality for the trace distance,

$$\|\rho - \rho'\| \leqslant \|\rho - \mathrm{tr}_{\mathrm{aux}} \Lambda[\rho]\| + \|\mathrm{tr}_{\mathrm{aux}} \Lambda[\rho] - \mathrm{tr}_{\mathrm{aux}} \Lambda[\rho']\| + \|\mathrm{tr}_{\mathrm{aux}} \Lambda[\rho'] - \rho'\|$$

holds, and we obtain due to Eq. (2) the lower bound

$$\delta \geqslant \eta [\|\rho - \rho'\| - \mathcal{D}(\rho, \rho')]. \tag{4}$$

A straightforward upper bound on δ for the optimal process (i.e., minimal δ) can be obtained by considering a con-

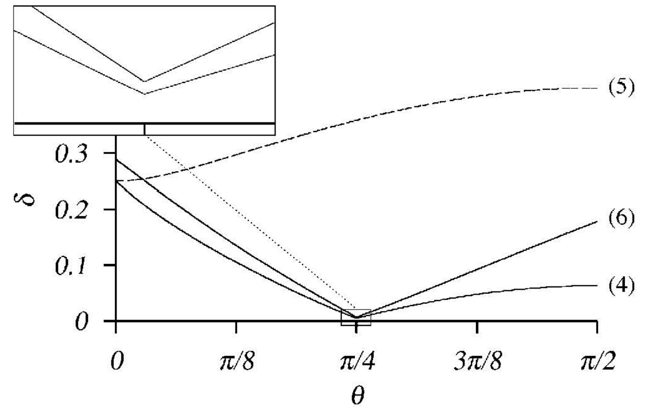

FIG. 1. Example for lower and upper bounds on the optimal deviation from perfect faithfulness δ of a two-state purifier with pure output. See main text for explanation.

stant purifier that produces a perfect purification of ρ'. This leads to the first upper bound

$$\delta_{\text{opt}} \leq \eta\|\rho-\rho'\|. \qquad (5)$$

A more sophisticated upper bound on δ is given by using the map which reaches the equality in Eq. (2). One chooses the output of ρ' to be a purification of ρ' and the output of ρ to be a pure state, which is as close as possible—according to Eq. (2)—to a purification of ρ. Since the maximal overlap of all purifications for two states ρ and ρ' is given by the Uhlmann fidelity $F(\rho,\rho')=\mathrm{tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}$ [11,12], we find with $\sin\alpha=\mathcal{D}(\rho,\rho')$ and $\cos\beta=F(\rho,\rho')$ the second upper bound

$$\delta_{\text{opt}} \leq \eta\sin(\beta-\alpha). \qquad (6)$$

Let us give an explicit example for these bounds. We consider the states $\rho=\frac{1}{2}(|0\rangle\langle 0|+|1\rangle\langle 1|)\otimes|0\rangle\langle 0|$ and $\rho'=\frac{2}{3}|0\rangle\langle 0|\otimes|+\rangle\langle +|+\frac{1}{3}|1\rangle\langle 1|\otimes|\theta\rangle\langle\theta|$, which appear with equal *a priori* probability, where $|\theta\rangle=\cos\theta|0\rangle+\sin\theta|1\rangle$ and $|+\rangle=(|0\rangle+|1\rangle)/\sqrt{2}$. In Fig. 1 the bounds for the optimal deviation from faithfulness δ are shown: the lower bound as given in Eq. (4), the first (dashed line) and second upper bound, cf. Eq. (5) and (6). At $\theta=0$ the ranges of both states share the vector $|1\rangle\otimes|0\rangle$ and thus the worst-case distinguishability vanishes and the optimal faithfulness is given by the upper bound in Eq. (5). The second upper bound and the lower bound almost coincide at $\theta=\pi/4$ with $0.0050<\delta<0.0072$. Note that the upper bounds cross each other, i.e., depending on the input state, either the first or the second upper bound is tighter.

An interesting question in this context is the following: given two quantum states, does a better distinguishability (in the sense of minimum error discrimination) imply a better faithfulness? The surprising answer is no: in the example given above, the trace distance of the two states monotonically increases from $\theta=0$ to $\theta=\pi/2$, while the deviation from faithfulness has its minimum at $\theta=\pi/4$. The examples illustrates that the worst-case distinguishability is indeed an important quantity for purifying processes. This is remarkable, as the worst-case distinguishability is purely determined by the geometric features of the states, whereas the statistical weights in the ensembles do not play any role. Note that a related, but not purely geometric quantity $F_1^{\dagger}(\rho,\rho')$ was introduced in [13].

IV. SETS THAT CAN BE PURIFIED PERFECTLY

Finally, our focus turns to the general analysis of perfect purifiers. The existence of a perfect purifier for a set $\mathcal{M}$ has far-reaching implications, as it is possible to convert all states in $\mathcal{M}$ to pure states in a reversible way. An investigation of the property of reversibility indeed turns out to be the key for understanding perfect purification: Suppose that we have a purifier Λ of a set $\mathcal{M}$ with perfect output purity (but not necessarily perfect faithfulness), and some completely positive and trace preserving map Λ', such that for any $\rho_i \in \mathcal{M}$ this map is the reverse map of Λ, i.e., $\Lambda'[\Lambda[\rho_i]]=\rho_i$. The action of any completely positive and trace preserving map can always be formulated as appending a (pure) ancilla state, performing a unitary rotation and finally tracing out an appropriate subsystem. We write Λ' in this manner and apply everything, apart from tracing out, to the output of Λ. For this composed map we write the shorthand notation $\tilde{\Lambda}$. The output of $\tilde{\Lambda}$ is still pure for any state in $\mathcal{M}$ and the remaining step of the map Λ', namely the trace over the subsystem, yields back the original state, thus $\tilde{\Lambda}$ is a *perfect* purifier of $\mathcal{M}$.

In order to further approach the characterization of sets that can be purified perfectly, we call a set of states *essentially pure*, if every state from the set can be globally rotated into a tensor product of a pure state and a common mixed contribution, or in more technical terms: A set of states $\mathcal{M}$ is called essentially pure, if one can find states ω_{aux} and σ_{B}, a unitary transformation U, and a set of *pure* states $\mathcal{P}_{\text{A}}$, such that for all $\rho_i \in \mathcal{M}$ there is a corresponding pure state $|\phi_i\rangle\langle\phi_i| \in \mathcal{P}_{\text{A}}$ with

$$\rho_i \otimes \omega_{\text{aux}} = U(|\phi_i\rangle\langle\phi_i| \otimes \sigma_{\text{B}})U^{\dagger}. \qquad (7)$$

Note that the tensor product symbol on the two sides of this equation, in general, denotes *different* splits of the composite system: on the left-hand side one sees the composition of the original system and an auxiliary system, while on the right-hand side the composition refers to some system A and some system B. Essentially pure sets can be purified perfectly: A process which appends ω_{aux} to ρ_i, performs $U^{\dagger}$ and traces out system B produces a pure state for any state in $\mathcal{M}$. On the other hand a process, which appends σ_{B} to $|\phi_i\rangle\langle\phi_i|$, performs U and traces out the auxiliary system, undoes the action of the purifying map. Thus, a perfect purifier of $\mathcal{M}$ exists. Of course, a union of essentially pure sets, where any two states taken from different sets are orthogonal, can also be purified perfectly. We call such a union an *orthogonal union* of essentially pure sets.

Theorem 2. For a set of states $\mathcal{M}$, the following statements are equivalent: (i) A perfect purifier of $\mathcal{M}$ exists. (ii) There exists a completely positive and trace preserving map, which maps any state in $\mathcal{M}$ to a pure state and does not change the trace distance of any two states in $\mathcal{M}$. (iii) $\mathcal{M}$ is an orthogonal union of essentially pure sets.

Proof. Our motivation for the definition of orthogonal unions of essentially pure sets was indeed that this property implies the existence of a perfect purifier. Thus, we have already shown that (iii) implies (i). Furthermore, from the fact that no process can increase the trace distance, together with the existence of a reversible map, (ii) is a direct consequence of (i). Thus it only remains to show that (ii) implies (iii): If (ii) holds for an $\mathcal{M}$ that is a union of mutually orthogonal subsets, there exist maps that satisfy (ii) for each subset. Therefore, we can assume without loss of generality that one cannot split the set $\mathcal{M}$ into orthogonal parts. With $|a\rangle\langle a|$ being a pure auxiliary state and $U^\dagger$ a unitary transformation, we can write the action of Λ as $\rho \mapsto \mathrm{tr}_B U^\dagger(\rho \otimes |a\rangle\langle a|)U$, where B denotes an appropriate subsystem. Since the output of Λ for a state $\rho_i \in \mathcal{M}$ is a pure state (represented by a projector Φ_i), we have $U^\dagger(\rho_i \otimes |a\rangle\langle a|)U = \Phi_i \otimes \sigma_i$, with σ_i a state in subsystem B. The final step is now to show that $\sigma_i = \sigma_j$ holds. For any two states $\rho_i, \rho_j \in \mathcal{M}$, due to the assumption (ii),

$$\|\Phi_i - \Phi_j\| = \|\rho_i - \rho_j\| = \|\Phi_i \otimes \sigma_i - \Phi_j \otimes \sigma_j\| \qquad (8)$$

holds. A minimum error discrimination [3,4] in subsystem B on the right-hand side can be written as $\sigma_i \to q_i|0\rangle\langle 0| + (1 - q_i)|1\rangle\langle 1|$ and $\sigma_j \to (1-q_j)|0\rangle\langle 0| + q_j|1\rangle\langle 1|$, where $(q_i + q_j)/2 = (1 + \|\sigma_i - \sigma_j\|)/2$ is the success probability for the optimal discrimination measurement. We find

$$\|\Phi_i \otimes \sigma_i - \Phi_j \otimes \sigma_j\|^2 \geqslant [\|q_i\Phi_i - (1-q_j)\Phi_j\| + \|(1-q_i)\Phi_i - q_j\Phi_j\|]^2 \geqslant \|\Phi_i - \Phi_j\|^2 + \|\sigma_i - \sigma_j\|^2 \mathrm{tr}(\Phi_i\Phi_j), \qquad (9)$$

where in the first step we used, that the discrimination procedure cannot increase the trace distance. The second inequality follows from a lengthy but straightforward calculation. From a comparison with Eq. (8) either $\sigma_i = \sigma_j$ or $\mathrm{tr}(\Phi_i\Phi_j) = 0$ (or both) must hold. The latter case implies ρ_i to be orthogonal to ρ_j, i.e., if $\sigma_i \neq \sigma_j$ for two states, then one can split $\mathcal{M}$ into two orthogonal sets, in contrast to our assumption. □

This Theorem completely characterizes all sets of states that can be purified perfectly, cf. also Eq. (7). It is surprising that one can even purify a set of *continuous* states, meaning that the set may contain infinitesimally close neighbors. It is also worth mentioning that all states in an essentially pure set share the same spectrum and pairwise have a completely degenerate set of canonical angles [9]. What is the lowest dimension, in which perfect purification is possible for nonorthogonal mixed states? This cannot happen unless the dimension of the Hilbert space is at least four: In two and three dimensions, only pure states can have identical spectra without having an overlapping range.

Although essentially pure sets can be characterized in a explicit manner and have a lot of straightforward features, there is no obvious method to verify whether a given set is of the structure as specified in Eq. (7). However, for the case, where $\mathcal{M}$ consists of only two states, there exists a computable test: From the lower bound on δ derived in Eq. (4) it follows that $\|\rho - \rho'\| = \mathcal{D}(\rho, \rho')$ is a necessary condition for the existence of a perfect two-state purifier. It is also a sufficient condition: For any two states ρ and ρ' there is a map Λ such that $\|\Lambda[\rho] - \Lambda[\rho']\| = \mathcal{D}(\rho, \rho')$, thus if $\|\rho - \rho'\| = \mathcal{D}(\rho, \rho')$, this map satisfies part (ii) of Theorem 2, i.e., ρ and ρ' can be purified perfectly. Note, that it is also straightforward to prove that the upper bound on δ_{opt} in Eq. (6) vanishes if and only if there is a perfect purifier of ρ and ρ'.

V. CONCLUSIONS

In summary, we have introduced the concept of purification as a physical map, and studied its properties: without any prior knowledge of the input state a perfect purifier cannot exist. Relaxing one of the two characteristic properties of a purifier, purity and faithfulness, does not lead to a nontrivial universal process either. We have investigated the case when the input set contains only two states and found a characterization of the output of any map, which takes both states to a pure state. Using this tool, we derived bounds on the deviation from perfect faithfulness (i.e., the distance of the partial trace of the output state and the original state). We also completely characterized all sets of states that can be purified perfectly. Roughly speaking, any such set can be globally rotated into a set of pure states, tensored with a common mixed contribution. Surprisingly, we found that some sets of noncommuting states can be purified, in contrast to the situation of broadcasting. For the case of sets with only two states, we provided an operational test to check whether perfect purification is possible.

In this paper we have presented some of the basic properties of purifying completely positive maps. Several questions remain open. One direction of future work is to consider the maximal possible purity of a purifier in the case of perfect faithfulness. Furthermore, the analysis of purifiers for sets with more than two states will be the subject of further research.

ACKNOWLEDGMENTS

We acknowledge discussions with Norbert Lütkenhaus, Armin Uhlmann, and Reinhard Werner. This work was partially supported by the EC programmes SECOQC and SCALA.

[1] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).

[2] A. Bassi and G. Ghirardi, Phys. Lett. A **309**, 24 (2003).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[4] U. Herzog and J. A. Bergou, Phys. Rev. A **70**, 022302 (2004).

[5] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).

[6] D. Dieks, Phys. Lett. **92**, 271 (1982).

[7] H. P. Yuen, Phys. Lett. A **113**, 405 (1986).

[8] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, Phys. Rev. Lett. **76**, 2818 (1996).

[9] G. W. Stewart and J.-g. Sun, *Matrix Pertubation Theory* (Academic Press, San Diego, 1990).

[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[11] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).

[12] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[13] A. Uhlmann, Rep. Math. Phys. **46**, 319 (2000).

[14] The output of a probabilistic process Λ_p with success rate $\mathrm{tr}\,\Lambda_\mathrm{p}[\rho]$ is (for our purposes) physically equivalent to a deterministic process $\Lambda_\mathrm{d}:\rho\mapsto\Lambda_\mathrm{p}[\rho]+(1-\mathrm{tr}\,\Lambda_\mathrm{p}[\rho])|\varphi\rangle\langle\varphi|$, with $|\varphi\rangle\langle\varphi|$ being orthogonal to all $\Lambda_\mathrm{p}[\rho]$.

Publication C

Purifying and reversible physical processes
Applied Physics B, Vol. **86**, p. 371 (2007) *(5 pages, invited article)*
M. Kleinmann, H. Kampermann, T. Meyer, and D. Bruß

Appl. Phys. B 86, 371–375 (2007)
DOI: 10.1007/s00340-006-2515-4

Applied Physics B
Lasers and Optics

Purifying and reversible physical processes

M. KLEINMANN✉
H. KAMPERMANN
T. MEYER
D. BRUß

Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Received: 7 August 2006
Published online: 12 December 2006 • © Springer-Verlag 2006

ABSTRACT Starting from the observation that reversible processes cannot increase the purity of any input state, we study deterministic physical processes, which map a set of states to a set of pure states. Such a process must map any state to the same pure output, if purity is demanded for the input set of all states. But otherwise, when the input set is restricted, it is possible to find non-trivial purifying processes. For the most restricted case of only two input states, we completely characterise the output of any such map. We furthermore consider maps, which combine the property of purity and reversibility on a set of states, and we derive necessary and sufficient conditions on sets, which permit such processes.

PACS 03.67.-a; 03.65.-w

1 Introduction

The notion of a pure quantum state plays a special role in quantum information theory. Many problems, such as separability or the existence of a particular quantum protocol can easily be solved if one restricts the problem to pure quantum states only. However, mixed states endow quantum systems with many properties (such as bound entanglement), that cannot be found for systems described by pure states. In our contribution we investigate physical processes which transform a given set of mixed states to a set of pure states. If such a process exists, then it may for example be possible to infer from the properties of the pure output states some properties of the input states. Such a kind of conclusion is particularly powerful, if the purifying map can be chosen to be reversible, since then one can consider the set of pure output states and the set of input states as physically equivalent.

We consider deterministic physical processes, i.e., processes which map any possible input quantum state with the probability of one to a corresponding output quantum state. The states of the input quantum system are represented by density operators $\varrho_{\text{in}} \in \mathcal{B}_+(\mathcal{H}_{\text{in}})$, i.e., positive semidefinite operators with trace one acting on the finite-dimensional complex vector space $\mathcal{H}_{\text{in}}$. Analogously $\varrho_{\text{out}} \in \mathcal{B}_+(\mathcal{H}_{\text{out}})$ shall represent the set of states of the output quantum system. Any deterministic physical process can be written as a completely positive and trace preserving linear (CPTP) map $\Lambda : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$, where $\mathcal{B}(\mathcal{H}_{\text{in, out}})$ denotes the space of linear operators on $\mathcal{H}_{\text{in, out}}$. In this language, the trace preserving condition reflects the fact that we restrict our considerations to deterministic processes. In Sect. 3 we will argue that this restriction is indeed necessary to have a proper definition of a purifying map.

A reversible process is a physical process, where the action of the process on any physical state can be undone by another physical process, i.e., a CPTP map Λ is reversible if one can find an inverse map Λ' which is also CPTP and satisfies $(\Lambda' \circ \Lambda)[\varrho_{in}] = \varrho_{in}$ for any density operator ϱ_{in}. The most common example are unitary processes $E_U : \mathcal{X} \mapsto U\mathcal{X}U^\dagger$, where U is a unitary transformation, $UU^\dagger = 1$. Here obviously the inverse map is given by $(E_U)' = E_{U^\dagger}$. Another class of reversible processes that is important for our purposes is described by

$$E_\sigma : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{aux}}) : \mathcal{X} \mapsto \mathcal{X} \otimes \sigma\,, \tag{1}$$

where $\sigma \in \mathcal{B}_+(\mathcal{H}_{\text{aux}})$ is some arbitrary density operator. The inverse map for this process is the partial trace over the auxiliary system, $(E_\sigma)' = \text{tr}_{\text{aux}}$. A remarkable property of this inverse map is, that it does not depend on σ and hence cannot be reversible. Note that neither the action of E_U nor the one of E_σ increases the purity $\text{tr}(\varrho^2)$ of any density operator ϱ. Indeed a process, which is reversible on the set of all states cannot increase the purity of even a single state: Let us first consider reversible maps, for which the reverse map is the partial trace (e.g. E_σ). For such a reversible map Λ, the output of any pure state $\Pi_{|\varphi\rangle} \equiv |\varphi\rangle\langle\varphi|$ must be $\Lambda[\Pi_{|\varphi\rangle}] = \Pi_{|\varphi\rangle} \otimes \sigma_\varphi$ for some state σ_φ. For any state ϱ_{in} we find with the spectral decomposition $\varrho_{\text{in}} = \sum_i p_i \Pi_{|\lambda_i\rangle}$ that due to linearity, $\text{tr}(\Lambda[\varrho_{\text{in}}]^2) = \sum p_i^2 \text{tr}(\sigma_{\lambda_i}^2) \leq \text{tr}(\varrho_{\text{in}}^2)$, i.e., no state can become purer by the action of Λ. Now consider a general reversible map Λ. For the reverse process Λ', by virtue of Stinespring's dilation theorem [1, 2] one can write the most general form of a CPTP map,

$$\Lambda' : \mathcal{X} \mapsto \text{tr}_{\text{aux}} U(\mathcal{X} \otimes \Pi_{|\text{anc}\rangle})U^\dagger\,. \tag{2}$$

From this we define

$$\begin{aligned}\Gamma_{\Lambda',\Lambda;|\text{anc}\rangle} :& \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{aux}}) \\ :& \mathcal{X} \mapsto U(\Lambda[\mathcal{X}] \otimes \Pi_{|\text{anc}\rangle})U^\dagger\,. \end{aligned}\tag{3}$$

✉ Fax: +49 221 81 11337, E-mail: kleinmann@thphy.uni-duesseldorf.de

The inverse map of $\Gamma_{\Lambda',\Lambda}$ obviously is $(\Gamma_{\Lambda',\Lambda})' = \mathrm{tr}_{\mathrm{aux}}$ and by construction, $\mathrm{tr}(\Gamma_{\Lambda',\Lambda}[\varrho_{\mathrm{in}}]^2) = \mathrm{tr}(\Lambda[\varrho_{\mathrm{in}}]^2)$ holds for all ϱ_{in}. Using the previous result, we find $\mathrm{tr}(\Gamma_{\Lambda',\Lambda}[\varrho_{\mathrm{in}}]^2) = \mathrm{tr}(\Lambda[\varrho_{\mathrm{in}}]^2) \leq \mathrm{tr}(\varrho_{\mathrm{in}}^2)$, i.e., no state can become purer by the action of a reversible process.

Since a process that is reversible on all states cannot improve the purity of any state, one would guess that a process which maps all states to a pure state cannot be reversible for any state. Such a process is called a purifying process, i.e., a CPTP map Λ is purifying, if $\mathrm{tr}(\Lambda[\varrho_{\mathrm{in}}]^2) = 1$ holds for any input state ϱ_{in}. The action of such a map, indeed, has to map any state to the same pure output state: suppose $\Lambda[\varrho_1] \neq \Lambda[\varrho_2]$. Then for $\varrho_3 = (\varrho_1 + \varrho_2)/2$ we find $\Lambda[\varrho_3] = (\Lambda[\varrho_1] + \Lambda[\varrho_2])/2$, which only can be pure, if $\Lambda[\varrho_1] = \Lambda[\varrho_2]$ in contradiction to our assumption. Thus a purifying process must destroy any information of the input state and cannot be reversible at all.

So, the properties of reversibility and output purity are completely incompatible for a physical process, if one demands these properties to hold on all possible input states. Our approach now is to require these properties only on a certain subset of states $\mathcal{M} \subset \mathcal{B}_+(\mathcal{H}_{\mathrm{in}})$. In Sect. 2 we analyze the properties of maps, which map at least two mixed states to pure states. The result of this investigation will completely characterize any such map. As an application of this result we will provide a lower bound on the trace distance of any two product states $\varrho_1 \otimes \sigma_1$ and $\varrho_2 \otimes \sigma_2$. In a brief excursion in Sect. 3 we will show that, if we allow probabilistic processes, non-trivial examples of reversible and purifying processes can easily be constructed. But we will also show that the definition of a probabilistic process to some extent contradicts the properties of a purifying process. In Sect. 4 we will then characterize any set of states, for which a deterministic reversible and purifying map exists and discuss in some detail the structure of such sets. Finally, we conclude in Sect. 5.

2 Purifying processes of two states

In the previous analysis we ruled out the possibility of a non-trivial process, which takes all states $\mathcal{B}_+(\mathcal{H}_{\mathrm{in}})$ to a corresponding pure state in $\mathcal{B}_+(\mathcal{H}_{\mathrm{out}})$. So the question arises, to what extent this also holds if one demands a pure output only for a subset of states $\mathcal{M} \subset \mathcal{B}_+(\mathcal{H}_{\mathrm{in}})$. More technically, for a CPTP map Λ, let us write $\mathrm{pur}(\Lambda) = \{\varrho_{\mathrm{in}} \in \mathcal{B}_+(\mathcal{H}_{\mathrm{in}}) \mid \mathrm{tr}(\Lambda[\varrho_{\mathrm{in}}]^2) = 1\}$ for the set of states which gets purified by the action of Λ. For a purifying process of $\mathcal{M}$ we demand $\mathcal{M} \subset \mathrm{pur}(\Lambda)$. In this section we will only deal with the most simple non-trivial case where only two states ϱ_1 and ϱ_2 are to be mapped onto a pure state, i.e., $\mathcal{M} = \{\varrho_1, \varrho_2\} \subset \mathrm{pur}(\Lambda)$. Let us consider the case where we already have two purifying maps Λ_A and Λ_B acting on ϱ_1 and ϱ_2, and without loss of generality assume

$$d(\Lambda_A[\varrho_1], \Lambda_A[\varrho_2]) \geq d(\Lambda_B[\varrho_1], \Lambda_B[\varrho_2])\,. \tag{4}$$

(Here, $\mathrm{d}(\varrho, \sigma) = \frac{1}{2}\mathrm{tr}|\varrho - \sigma|$ with $|\mathcal{X}| = \sqrt{\mathcal{X}\mathcal{X}^\dagger}$ denotes the trace distance of ϱ and σ.) Then there exists a CPTP map Ω, such that up to a global unitary transformation, $\Lambda_B[\varrho_i] = (\Omega \circ \Lambda_A)[\varrho_i]$ for $i = 1, 2$: for two pure states $\Pi_{|\psi_1\rangle}$ and $\Pi_{|\psi_2\rangle}$ one can reduce the angle defined by $\sin\vartheta = d(\Pi_{|\psi_1\rangle}\Pi_{|\psi_2\rangle})$ to an arbitrary angle $\varphi < \vartheta$ via the CPTP map

$$\Omega_{\varphi;|\psi_1\rangle,|\psi_2\rangle}\colon\ \mathcal{X} \mapsto \sum_{\alpha=1}^{3} A_\alpha \mathcal{X} A_\alpha^\dagger, \tag{5a}$$

with A_α being the Kraus operators [3]

$$A_1 = |\psi_1\rangle\langle\psi_1| + a|\psi_1^\perp\rangle\langle\psi_1^\perp|, \tag{5b}$$

$$A_2 = (\sqrt{1-b^2}\,|\psi_1\rangle + \sqrt{b^2-a^2}\,|\psi_1^\perp\rangle)\,\langle\psi_1^\perp|\,, \tag{5c}$$

$$A_3 = 1 - |\psi_1\rangle\langle\psi_1| - |\psi_1^\perp\rangle\langle\psi_1^\perp|, \tag{5d}$$

where $a = \tan\varphi\cot\vartheta$, $b = \sin\varphi/\sin\vartheta$, and $|\psi_1^\perp\rangle \in \mathrm{span}\{|\psi_1\rangle, |\psi_2\rangle\}$ is a normalised vector orthogonal to $|\psi_1\rangle$. Now let $\Pi_{|\psi_i\rangle} = \Lambda_A[\varrho_i]$ and choose $\sin\varphi = d(\Lambda_B[\varrho_1], \Lambda_B[\varrho_2])$. Then, up to a global unitary transformation, $\Lambda_B[\varrho_i] = (\Omega_{\varphi;|\psi_1\rangle,|\psi_2\rangle} \circ \Lambda_A)[\varrho_i]$ holds. Since we can mimic the action of Λ_B on ϱ_1 and ϱ_2 by using the map Λ_A, we would always prefer Λ_A in favor of Λ_B. Thus among all purifying processes of two states we are most interested in those which maximize the trace distance of the corresponding output states.

This trace distance of the output of a purifying map Λ is upper bounded by a geometric quantity depending on ϱ_1 and ϱ_2, namely by the worst case distinguishability $\mathcal{D}(\varrho_1\varrho_2)$ [4],

$$d(\Lambda[\varrho_1], \Lambda[\varrho_2]) \leq \mathcal{D}(\varrho_1, \varrho_2)\,. \tag{6}$$

We now want to give a physical interpretation of this inequality. In quantum mechanics, an ensemble of pure states $\Pi_{|\varphi_j\rangle}$ with probabilities $p_j > 0$ (where $\sum p_j = 1$) is described by the mixed state $\varrho = \sum_j p_j \Pi_{|\varphi_j\rangle}$. In general, many different ensembles lead to the same density operator ϱ, and it is a prediction of quantum mechanics that it is impossible to physically distinguish between such different ensembles. Having said that, for a given mixed state ϱ, a pure state $\Pi_{|\varphi\rangle}$ may physically appear if and only if $\Pi_{|\varphi\rangle}$ can be part of an ensemble that is represented by ϱ, i.e., if and only if a positive number p exists, such that $\varrho - p\Pi_{|\varphi\rangle}$ is positive semidefinite. Let us denote the collection of all such pure states $\Pi_{|\varphi\rangle}$ by

$$\begin{aligned}\mathcal{Q}_\varrho &= \{\Pi_{|\varphi\rangle} \mid \exists p > 0\colon\ \varrho - p\Pi_{|\varphi\rangle} \geq 0\}\\ &\equiv \{\Pi_{|\varphi\rangle} \mid |\varphi\rangle \in \mathrm{supp}\,\varrho\},\end{aligned} \tag{7}$$

where $\mathrm{supp}\,\varrho$ is the support of ϱ, i.e., the orthocomplement of the kernel of ϱ. The worst-case distinguishability is now defined as

$$\begin{aligned}\mathcal{D}(\varrho_1\varrho_2) &= \inf_{\Pi_{|\varphi_i\rangle} \in \mathcal{Q}_{\varrho_i}} \mathrm{d}\left(\Pi_{|\varphi_1\rangle}, \Pi_{|\varphi_2\rangle}\right)\\ &\equiv \min_k(\sin\vartheta_k),\end{aligned} \tag{8}$$

where ϑ_k denote the Jordan angles [5] between $\mathrm{supp}\,\varrho_1$ and $\mathrm{supp}\,\varrho_2$.

Let us continue the physical motivation of (6). The maximal success probability for distinguishing two mixed states via a measurement ("minimum error discrimination") is given by [6, 7]

$$P_{\mathrm{MED}}(\varrho_1, \varrho_2) = (1 + d(\varrho_1, \varrho_2))/2\,, \tag{9}$$

where we assumed that both states have equal a priori probabilities. Hence P_{MED} is the average success probability for distinguishing the ensemble of pure states denoted by ϱ_1 and ϱ_2. In a physical experiment, each single measurement is performed on a pure state out of the ensembles, i.e., the task of the discrimination measurement is to distinguish between a state in $\mathcal{Q}_{\varrho_1}$ and a state in $\mathcal{Q}_{\varrho_2}$. The optimal success probability to distinguish between such two pure states in the worst case is given by

$$\begin{aligned} P_{\mathrm{WCD}} &= \inf_{\Pi_{|\varphi_i\rangle} \in \mathcal{Q}_{\varrho_i}} P_{\mathrm{MED}}(\Pi_{|\varphi_1\rangle}, \Pi_{|\varphi_2\rangle}) \\ &\equiv (1 + \mathcal{D}(\varrho_1, \varrho_2))/2 \,. \end{aligned} \tag{10}$$

Since no deterministic process can increase the trace distance between two states [2], a purifying process of ϱ_1 and ϱ_2 must not deterministically increase the distance between any pair of pure states $\Pi_{|\varphi_1\rangle} \in \mathcal{Q}_{\varrho_1}$ and $\Pi_{|\varphi_2\rangle} \in \mathcal{Q}_{\varrho_2}$. This may serve as a physical motivation for the inequality in (6).

Can the bound in (6) always be achieved by some purifying process Λ? The answer is affirmative, but in order to verify this to a satisfactory level there is no way to avoid the awkwardness of an explicit construction of a map which reaches equality in (6).

Let us first briefly recall the concept of Jordan bases and Jordan angles (sometimes also called canonical bases and canonical angles) [5, 8] of two subspaces $\mathcal{A}_1 \subset \mathcal{H}$ and $\mathcal{A}_2 \subset \mathcal{H}$. Orthonormal bases $|\psi_1^k\rangle$ of $\mathcal{A}_1$ and $|\psi_2^k\rangle$ of $\mathcal{A}_2$ are called Jordan bases, if

$$\langle \psi_1^k | \psi_2^l \rangle = 0 \qquad \text{for } k \neq l, \tag{11a}$$

$$\langle \psi_1^k | \psi_2^k \rangle = \cos \vartheta_k \qquad \text{for } k \leq \min_i \dim \mathcal{A}_i. \tag{11b}$$

Such bases always exist and ϑ_k are called the Jordan angles between $\mathcal{A}_1$ and $\mathcal{A}_2$.

The first step in the construction of the purifying map is to apply the distance-decreasing map Ω_φ defined in (5a)–(5d) on each pair of Jordan vectors $|\psi_1^k\rangle \in \operatorname{supp} \varrho_1$ and $|\psi_2^k\rangle \in \operatorname{supp} \varrho_2$, such that the distance is reduced to $\mathcal{D}(\varrho_1, \varrho_2)$: We define the Kraus operators A_1^k and A_2^k for $k \leq \min_i \operatorname{rank} \varrho_i$ analogously to (5b) and (5c) and choose $\sin \varphi_k = \mathcal{D}(\varrho_1 \varrho_2)$. In order to complete the set of Kraus operators, we in addition define the projector $A_3 = \mathbf{1} - \sum_k A_1^{k\dagger} A_1^k - \sum_k A_2^{k\dagger} A_2^k$ and write

$$\begin{aligned} \tilde{\Omega} \colon &\mathcal{B}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{B}(\mathcal{H}_{\mathrm{in}}) \\ \colon &\mathcal{X} \mapsto \sum_k A_1^k \mathcal{X} A_1^{k\dagger} + \sum_k A_2^k \mathcal{X} A_2^{k\dagger} + A_3 \mathcal{X} A_3^\dagger \,. \end{aligned} \tag{12}$$

Let $\Pi_{|\nu_i\rangle}$ be an arbitrary pure state in $\mathcal{Q}_{\varrho_i}$. One finds that

$$\begin{aligned} \tilde{\Omega}[\Pi_{|\nu_i\rangle}] = &\sum_k \operatorname{tr}(\Pi_{|\psi_i^k\rangle} \Pi_{|\nu_i\rangle}) \tilde{\Omega}[\Pi_{|\psi_i^k\rangle}] \\ &+ A_3 \Pi_{|\nu_i\rangle} A_3^\dagger \,. \end{aligned} \tag{13}$$

By construction, $\Pi_{|\tilde{\psi}_i^k\rangle} = \tilde{\Omega}[\Pi_{|\psi_i^k\rangle}]$ is again pure with $\langle \tilde{\psi}_i^k | \tilde{\psi}_j^l \rangle = 0$ for $k \neq l$ and $d(\Pi_{|\tilde{\psi}_1^k\rangle}, \Pi_{|\tilde{\psi}_2^k\rangle}) = \mathcal{D}(\varrho_1, \varrho_2)$. Furthermore $A_3 \Pi_{|\nu_i\rangle} A_3^\dagger \neq 0$ only if $\operatorname{rank} \varrho_i > \operatorname{rank} \varrho_j$ for $j \neq i$.

Using the above properties of $\tilde{\Omega}$, it is straightforward to find a CPTP map $\tilde{E} \colon \mathcal{H}_{\mathrm{in}} \to \mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{\mathrm{aux}}$, such that the vectors $|k\rangle|\varphi_1\rangle$ diagonalise $(\tilde{E} \circ \tilde{\Omega})[\varrho_1]$ and the vectors $|k\rangle|\varphi_2\rangle$ diagonalise $(\tilde{E} \circ \tilde{\Omega})[\varrho_2]$, where $\langle k|l\rangle = \delta_{kl}$ and $d(\Pi_{|\varphi_1\rangle}, \Pi_{|\varphi_2\rangle}) = \mathcal{D}(\varrho_1, \varrho_2)$. Thus the map $\operatorname{tr}_{\mathrm{in}} \circ \tilde{E} \circ \tilde{\Omega}$ is a map which reaches the bound in (6), i.e.,

$$\mathcal{D}(\varrho_1, \varrho_2) = \max_{\Lambda} d(\Lambda[\varrho_1], \Lambda[\varrho_2]) \,, \tag{14}$$

where the maximum is taken over all CPTP maps Λ satisfying $\{\varrho_1, \varrho_2\} \subset \operatorname{pur}(\Lambda)$. Furthermore, as already discussed in advance, due to (14), the maximizing map $\operatorname{tr}_{\mathrm{in}} \circ \tilde{E} \circ \tilde{\Omega}$ together with the distance-decreasing map Ω_φ allows mimicking of the action of any purifying map of the states ϱ_1 and ϱ_2.

This result characterizes the output of any process, which maps two input states to pure output states. For example one immediately finds that two states with overlapping support have a vanishing worst-case distinguishability and thus such states only can be mapped to identical pure states by a purifying process. In [4] the problem was investigated, how close the pure output states of a purifying map can get to a purification [9, 10] of the input states. The deviation from the optimal quality of such a purifying map was found to be limited by the difference $d(\varrho_1, \varrho_2) - \mathcal{D}(\varrho_1, \varrho_2)$. Furthermore the result in (14) turned out to be the key for the analysis of sets which can be mapped perfectly to their purifications [4].

In addition, the result in (14) can also be used as a general tool in quantum information theory, since results for pure states often are much simpler to obtain than results for mixed states. As an example, we provide a lower bound on the trace distance of any two product states $\varrho_1 \otimes \sigma_1$ and $\varrho_2 \otimes \sigma_2$:

$$\begin{aligned} &d(\varrho_1 \otimes \sigma_1, \varrho_2 \otimes \sigma_2)^2 \geq \\ &\qquad 1 - (1 - \mathcal{D}(\varrho_1, \varrho_2)^2)(1 - d(\sigma_1, \sigma_2)^2) \,. \end{aligned} \tag{15}$$

(From this inequality in particular $d(\varrho_1, \varrho_2) \geq \mathcal{D}(\varrho_1, \varrho_2)$ follows by setting $\sigma_1 = \sigma_2$.) This inequality follows by applying a map for which

$$\varrho_1 \otimes \sigma_1 \mapsto \Pi_{|\varphi_1\rangle} \otimes (q_1 \Pi_{|0\rangle} + (1 - q_1) \Pi_{|1\rangle}) \,, \tag{16}$$

$$\varrho_2 \otimes \sigma_2 \mapsto \Pi_{|\varphi_2\rangle} \otimes ((1 - q_2) \Pi_{|0\rangle} + q_2 \Pi_{|1\rangle}) \,. \tag{17}$$

Such a mapping can be implemented by a CPTP map for appropriate q_1, q_2 satisfying $q_1 + q_2 = 1 + d(\sigma_1, \sigma_2)$ and $\Pi_{|\varphi_i\rangle}$ satisfying $d(\Pi_{|\varphi_1\rangle}, \Pi_{|\varphi_2\rangle}) = \mathcal{D}(\varrho_1, \varrho_2)$, since then for the first system one applies the purifying map $\operatorname{tr}_{\mathrm{in}} \circ \tilde{E} \circ \tilde{\Omega}$ and for the second system one applies a minimum error discrimination of σ_1 and σ_2. Now using the fact that a CPTP map cannot increase the trace distance, it is straightforward to obtain (15).

3 Probabilistic purifying processes

Although we want to concentrate on deterministic processes, in this section we wish to briefly discuss the properties of probabilistic purifying processes. We exclude probabilistic processes $\bar{\Lambda}$ from our considerations, for which $\operatorname{tr} \bar{\Lambda}[\varrho] = 0$ for some $\varrho \in \mathcal{M}$, i.e., we call a process probabilistic on $\mathcal{M}$, only if for any state in $\mathcal{M}$ the success probability of the process is non-zero.

A simple example of a probabilistic purifying process is a process, which first performs an unambiguous state discrimination [8, 11] between the possible input states and then uses the unambiguous information to create a purification of the input state. In the language of probabilistic processes, unambiguous state discrimination of a set of states ϱ_i is a probabilistic map which maps ϱ_i to $p_i \Pi_{|i\rangle}$, where p_i is the success probability of unambiguously identifying ϱ_i and $\langle i | j \rangle = \delta_{ij}$. In [12] it was shown, that a probabilistic unambiguous state discrimination process for a set $\mathcal{M}$ exists, if and only if

$$\operatorname{supp} \varrho_i \not\subseteq \sum_{j \neq i} \operatorname{supp} \varrho_j , \quad \forall \varrho_i \in \mathcal{M} . \tag{18}$$

Hence, if one applies unambiguous state discrimination on such a set $\mathcal{M}$, in case of a successful discrimination one can map each state $\Pi_{|i\rangle}$ to a purification $\Pi_{|\psi_i\rangle}$ of ϱ_i. This map is purifying as well as reversible on $\mathcal{M}$ (with the reversible map being the partial trace over the purifying system) and it is successful, whenever the unambiguous state discrimination process succeeds.

However, there is a good reason not to deepen the analysis of probabilistic processes as a proper variant of purifying processes: physically, the information of a successful application of a probabilistic map is provided as a bit of classical information. Thus for a probabilistic purifying map $\bar{\Lambda}$ one can equivalently write the deterministic map

$$\Lambda \colon \mathcal{X} \mapsto \bar{\Lambda}[\mathcal{X}] + (\operatorname{tr}\mathcal{X} - \operatorname{tr}\bar{\Lambda}[\mathcal{X}]) \Pi_{|?\rangle} , \tag{19}$$

where $\Pi_{|?\rangle}$ is a state that is orthogonal to all output operators $\bar{\Lambda}[\mathcal{X}]$. However, the output of Λ is not pure, unless $\bar{\Lambda}$ is already deterministic and purifying.

4 Purifying and reversible processes

We now want to combine the purifying property of a deterministic process with the feature of reversibility. Since we already noticed that processes which are reversible on all states cannot increase the purity of any state (although it is possible to decrease the purity, e.g. using the map E_σ defined in (1)), in the fashion of Sect. 2 we demand reversibility only on a subset of states $\mathcal{M}$. We call a CPTP map Λ reversible on $\mathcal{M}$ if one can find a CPTP map Λ', such that $(\Lambda' \circ \Lambda)[\varrho_{\text{in}}] = \varrho_{\text{in}}$ for all $\varrho_{\text{in}} \in \mathcal{M}$. Let us again formalize this property. For a CPTP map $\Xi \colon \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{in}})$, let $ID(\Xi) = \{\varrho_{\text{in}} \in \mathcal{B}_+(\mathcal{H}_{\text{in}}) \mid \Xi[\varrho_{\text{in}}] = \varrho_{\text{in}}\}$ be the set of states that are unchanged by the action of Ξ. Thus for a reversible map Λ on $\mathcal{M}$, we demand that one can find a CPTP map Λ', such that $\mathcal{M} \subset ID(\Lambda' \circ \Lambda)$. Note, that Λ' does not need to be unique. Now a map Λ is purifying and reversible on $\mathcal{M}$, if and only if one can find a map Λ', such that $\mathcal{M} \subset \operatorname{pur}(\Lambda) \cap ID(\Lambda' \circ \Lambda)$. It is possible to completely characterize any such set $\mathcal{M}$:

Theorem 1. *A reversible and purifying process for a set of states $\mathcal{M} \subset \mathcal{B}_+(\mathcal{H}_{in})$ exists, if and only if for appropriate vector spaces $\mathcal{H}^i_C$, $\mathcal{H}^i_A$ and $\mathcal{H}^i_B$, satisfying $\mathcal{H}_{in} \otimes \mathcal{H}^i_C \cong \mathcal{H}^i_A \otimes \mathcal{H}^i_B$, one can find mixed states $\sigma^i_B \in \mathcal{B}_+(\mathcal{H}^i_B)$ and $\omega^i_C \in \mathcal{B}_+(\mathcal{H}^i_C)$, and unitary transformations U_i, such that $\mathcal{M} = \bigcup_i \mathcal{M}_i$ with $\mathcal{M}_i \perp \mathcal{M}_j$, $i \neq j$ and*

$$\{\varrho \otimes \omega^i_C \mid \varrho \in \mathcal{M}_i\} \subset \{U_i (\Pi_{|\varphi\rangle} \otimes \sigma^i_B) U_i^\dagger \mid |\varphi\rangle \in \mathcal{H}^i_A\} . \tag{20}$$

In Theorem 1, $\mathcal{M}_i \perp \mathcal{M}_j$ if $\operatorname{tr}(\varrho\sigma) = 0$ for all $\varrho \in \mathcal{M}_i$ and $\sigma \in \mathcal{M}_j$. Sets $\mathcal{M}_i$, which satisfy (20) are called essentially pure, i.e., a reversible and purifying process for $\mathcal{M}$ exists, if and only if $\mathcal{M}$ is an orthogonal union of essentially pure sets. Furthermore, note that basically, essentially pure sets are such sets which are generated by applying the map $E_U \circ E_\sigma$ on a set of pure states.

Proof (Theorem 1). In [4] it was shown, that $\mathcal{M}$ is an orthogonal union of essentially pure sets, if and only if a perfect purifier of $\mathcal{M}$ exists. A perfect purifier is a CPTP map, which maps any state in $\mathcal{M}$ to one of its purifications in $\mathcal{B}_+(\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{aux}})$. Hence a perfect purifier Λ of $\mathcal{M}$ in particular satisfies $\mathcal{M} \subset \operatorname{pur}(\Lambda)$ and $\mathcal{M} \subset ID(\operatorname{tr}_{\text{aux}} \circ \Lambda)$, i.e., it is purifying and reversible.

For the converse assume that a reversible and purifying map Λ for $\mathcal{M}$ exists. Then $\mathcal{M} \subset \operatorname{pur}(\Lambda)$ and one can find a CPTP map Λ', such that $\mathcal{M} \subset ID(\Lambda' \circ \Lambda)$. The map $\Gamma_{\Lambda',\Lambda}$ defined in (3) thus satisfies $\operatorname{pur}(\Gamma_{\Lambda',\Lambda}) = \operatorname{pur}(\Lambda) \supset \mathcal{M}$ and $ID(\operatorname{tr}_{\text{aux}} \circ \Gamma_{\Lambda',\Lambda}) = ID(\Lambda' \circ \Lambda) \supset \mathcal{M}$ and hence $\Gamma_{\Lambda',\Lambda}$ is a perfect purifier of $\mathcal{M}$. Using again the result in [4], it follows that $\mathcal{M}$ is an orthogonal union of essentially pure sets. □

Although Theorem 1 completely characterizes all sets for which a reversible and purifying process exists, it is in general not straightforward to test whether a set is of the structure as specified in (20). Only for the case where $\mathcal{M}$ consists of only two states, an operational necessary and sufficient criterion is known [4]: the set $\mathcal{M} = \{\varrho_1, \varrho_2\}$ is essentially pure or $\varrho_1 \perp \varrho_2$ if and only if $\mathcal{D}(\varrho_1, \varrho_2) = d(\varrho_1, \varrho_2)$. In the general case only some necessary operational conditions can be derived. The most obvious necessary criterion is, that in an essentially pure set all states must share the same spectrum. Another example of a necessary criterion is, that the Jordan angles between the support of any two states taken from an essentially pure set have to be completely degenerate. But these two properties are not sufficient for an essentially pure set, as the following simple counter-example demonstrates:

$$\varrho_1 = p\Pi_{|0\rangle} + (1-p)\Pi_{|1\rangle} \tag{21a}$$

$$\varrho_2 = p\Pi_{|\nu^+\rangle} + (1-p)\Pi_{|\nu^-\rangle} , \tag{21b}$$

where $|\nu^\pm\rangle = \frac{1}{2}(\pm|0\rangle + |1\rangle \pm |2\rangle + |3\rangle)$ and $0 < p < \frac{1}{2}$.

As a final remark let us note that it is possible to simplify the definition of essentially pure sets. A set of states $\mathcal{M} \subset \mathcal{B}_+(\mathcal{H}_{\text{in}})$ with $\varrho_0 \in \mathcal{M}$ is essentially pure if and only if one can find a unitary transformation U on $\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{aux}}$ and normalised vectors $|\varrho\rangle \in \mathcal{H}_{\text{aux}}$ corresponding to each $\varrho \in \mathcal{M}$, such that for each $\varrho \in \mathcal{M}$,

$$\varrho \otimes \Pi_{|\varrho_0\rangle} = U(\varrho_0 \otimes \Pi_{|\varrho\rangle})U^\dagger \tag{22}$$

holds. From the proof of Theorem 1 in [4] it is clear that in (20) one always can choose ω_C to be pure. Now the dimension of the kernel of each element on the left hand side of (20) is given by $\dim(\mathcal{H}_{\text{in}}) \dim(\mathcal{H}_C) - \operatorname{rank} \varrho$ while on the right hand side we find $\dim(\mathcal{H}_A) \dim(\mathcal{H}_B) - \operatorname{rank} \varrho$. One readily extends $\mathcal{H}_C$ and $\mathcal{H}_B$ such that $\dim \mathcal{H}_B$ is an integer multiple of $\dim(\mathcal{H}_{\text{in}})$. Then after a suitable rotation U' on $\mathcal{H}_A \otimes \mathcal{H}_B$, one has $U'(\Pi_{|\varphi\rangle} \otimes \sigma_B)U'^\dagger = (\varrho_0 \otimes \Pi_{|0\rangle}) \otimes \Pi_{|\varphi\rangle}$. Identifying $\Pi_{|0\rangle} \otimes \Pi_{|\varphi\rangle}$ with $\Pi_{|\varrho\rangle}$ finishes the proof of (22).

5 Conclusions

In summary we have analysed deterministic physical processes which are reversible or purifying, with particular focus on the combination of both properties. First we have shown that the properties of reversibility and purity of a physical processes are completely incompatible, as long as reversibility or purity is required to hold for any input state. For certain restricted sets, however, one can combine these properties. We investigated the case, where only two input states are mapped to pure states. It turned out that the trace distance of the output states of such a map is limited by the worst-case distinguishability of the input states. A map was provided, which always reaches this bound. Some applications of this result in quantum information theory were presented. For probabilistic processes we used unambiguous state discrimination to build a non-trivial example of a purifying and reversible process. We finally characterised all sets of states, for which a deterministic purifying and reversible process exists and it turned out that such sets have to be pure up to a common mixed contribution. Despite this result and the existence of an operational criterion for such essentially pure sets in the case, where the set consists of only two states, no operational necessary and sufficient condition for larger essentially pure sets was provided. Such criteria will be subject to further research. Furthermore, although some properties of reversible processes where presented here, another direction of future work will be to deepen the understanding of such processes.

ACKNOWLEDGEMENTS This work was partially supported by the European Commission (Integrated Projects SECOQC and SCALA).

REFERENCES

1 W. Forrest Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955)
2 M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (University Press, Cambridge, 2000)
3 K. Kraus, *States, Effects and Operations* (Springer, Berlin, 1983)
4 M. Kleinmann, H. Kampermann, T. Meyer, D. Bruß, Phys. Rev. A **73**, 062309 (2006)
5 G.W. Stewart, J.G. Sun, *Matrix Pertubation Theory* (Academic, San Diego, 1990)
6 C.W. Helstrøm, *Quantum Detection and Estimation Theory* (Academic, New York, 1976)
7 U. Herzog, J.A. Bergou, Phys. Rev. A **70**, 022302 (2004)
8 T. Rudolph, R.W. Spekkens, P.S. Turner, Phys. Rev. A **68**, 010301 (2003)
9 L.P. Hughston, R. Jozsa, W.K. Wootters, Phys. Lett. A **183**, 14 (1993)
10 A. Bassi, G.C. Ghirardi, Phys. Lett. A **309**, 24 (2003)
11 G. Jaeger, A. Shimony, Phys. Lett. A **197**, 83 (1995)
12 Y. Feng, R. Duan, Y. Mingsheng, Phys. Rev. A **70**, 012308 (2004)

Publication D

Commutator relations reveal solvable structures in unambiguous state discrimination
Journal of Physics A: Mathematical and General, Vol. 40, p. F871 (2007) *(8 pages)*
M. Kleinmann, H. Kampermann, Ph. Raynal, and D. Bruß

J. Phys. A: Math. Theor. **40** (2007) F871–F878 doi:10.1088/1751-8113/40/36/F04

FAST TRACK COMMUNICATION

Commutator relations reveal solvable structures in unambiguous state discrimination

M Kleinmann[1], H Kampermann[1], Ph Raynal[2] and D Bruß[1]

[1] Institut für Theoretische Physik III, Heinrich-Heine-Universität Düsseldorf,
D-40225 Düsseldorf, Germany
[2] Quantum Information Theory Group, Institut für Theoretische Physik I, and
Max-Plank-Forschungsgruppe, Institut für Optik, Information und Photonik,
Universität Erlangen-Nürnberg, D-91058 Erlangen, Germany

E-mail: kleinmann@thphy.uni-duesseldorf.de

Received 31 May 2007, in final form 17 July 2007
Published 21 August 2007
Online at stacks.iop.org/JPhysA/40/F871

Abstract

We present a criterion, based on three commutator relations, that allows us to decide whether two self-adjoint matrices with non-overlapping support are simultaneously unitarily similar to quasi-diagonal matrices, i.e., whether they can be simultaneously brought into a diagonal structure with (2×2)-dimensional blocks. Application of this criterion to unambiguous state discrimination provides a systematic test whether the given problem is reducible to a solvable structure. As an example, we discuss unambiguous state comparison.

PACS numbers: 03.67.−a, 03.65.−w, 02.10.Yn

1. Introduction

The commutator of two self-adjoint operators, which act on a Hilbert space, is a fundamental concept in quantum mechanics: two observables can be measured without uncertainty if and only if their commutator vanishes. This physical interpretation is connected to the mathematical fact that two Hermitian matrices can be diagonalized simultaneously if and only if their commutator is zero. A natural question to ask is when two Hermitian matrices can be simultaneously brought into a block-diagonal structure with blocks of the lowest non-trivial size, namely size 2×2. Such structures are known as quasi-diagonal form and criteria for existence have been studied in [1, 2]: Watters [1] showed that a family of normal matrices can be simultaneously brought into a quasi-diagonal form if and only if each member of the family commutes with the squared commutator of an element of the family with any element from the algebra generated by the family. (Thus, testing this criterion requires us to show that infinitely many commutators vanish.) Laffey [2] studied a family with two members only. He showed that when the matrices in the family are positive semi-definite, they are simultaneously unitarily similar to quasi-diagonal matrices if and only if six certain commutators vanish.

The question of simultaneous quasi-diagonalizability has a physical application in unambiguous discrimination of quantum states (see the next paragraph). In that context, it is sufficient to deal with positive semi-definite operators with non-overlapping supports (the support of an operator is the orthocomplement of its kernel). As we will show, this restriction leads to simpler commutator criteria. In this paper we will give a constructive proof that, given two self-adjoint operators with non-overlapping supports, they have a common block diagonal structure of dimension 2, if and only if a set of only three commutators vanishes. These commutators are also easier to calculate than those given in [2], as the latter are of maximal order 7, while the former are of maximal order 5.

Unambiguous state discrimination (USD) is a strategy for distinguishing non-orthogonal quantum states without being allowed to make an error. As it is impossible to discriminate non-orthogonal quantum states with unit probability, the measurement has to have inconclusive outcomes. The optimal USD strategy is the one that maximizes the success probability (i.e., minimizes the probability to get an inconclusive result). A different possibility to discriminate quantum states is called *minimum error discrimination*, where one minimizes the probability of making an error in the state identification.

In this contribution we want to focus onto the first strategy, namely unambiguous state discrimination. For two density operators, ρ_1 and ρ_2, acting on the Hilbert space $\mathcal{H}$ of finite dimension, this task is described by a positive operator-valued measure (POVM) on $\mathcal{H}$, consisting of three positive operators, E_1, E_2 and $E_?$, with $E_1 + E_2 + E_? = \mathbb{1}$. In order to make the discrimination unambiguous, the probability of wrong identification must vanish, i.e. $\text{tr}(E_1\rho_2) = 0$ and $\text{tr}(E_2\rho_1) = 0$. It is natural to allow ρ_1 and ρ_2 to have *a priori* probabilities p_1 and p_2, respectively, where $p_1 > 0$, $p_2 > 0$, and $p_1 + p_2 = 1$. The open problem in USD is to find a POVM $\{E_1, E_2, E_?\}$ which maximizes the success probability $p_{\text{succ}} = p_1 \,\text{tr}(E_1\rho_1) + p_2 \,\text{tr}(E_2\rho_2)$.

While the optimal solution for minimum error discrimination of two mixed states is already known for more than three decades [3], the optimal solution for unambiguous state discrimination has been found only for the pure state case [4] and certain special cases of mixed states [5–13]. A partial solution for unambiguous discrimination of mixed states is provided via the reductions of the density operators by the space where perfect and/or no USD is possible [8]. Otherwise, known optimal USD measurements for mixed states mainly belong to the class, where the problem can be decomposed into several pure state discrimination tasks [5, 9, 11]. A general representation of such states was recently discussed by Bergou *et al* [11].

It is not obvious how to decide whether the given density operators possess such a structure. In this contribution we present a method that allows us to systematically identify if the optimal USD of two mixed states can be simplified to the pure state task.

The paper is organized as follows. In section 2 we introduce the concept of common block-diagonal structures of two operators. We specifically consider the case of two-dimensional blocks, as the optimal measurement in two dimensions is well known. Simple commutator relations are presented to check for the existence of such a structure. In section 3 we discuss whether the block structures are preserved by the reductions. Finally, we study the example of unambiguous state comparison [7, 9, 14–16] to illustrate the power of the commutator test.

2. Block-diagonal structures

2.1. Independent orthogonal subspaces in USD

In [5] Bennett *et al* analyzed the parity check for a string of qubits, i.e., the question whether a sequence composed of states that are either $|\psi_0\rangle$ or $|\psi_1\rangle$, with $0 < |\langle\psi_0|\psi_1\rangle| < 1$, contains an even or odd number of occurrences of $|\psi_1\rangle$. This task is equivalent to the unambiguous discrimination of two certain mixed states. After a suitable (symmetric) choice of a basis these mixed states turned out to share the same block-diagonal shape, with each block $\blacksquare$ symbolizing a 2×2 matrix:

$$\rho_1 = \begin{pmatrix} \blacksquare & & \\ & \blacksquare & \\ & & \ddots \end{pmatrix}, \qquad \rho_2 = \begin{pmatrix} \blacksquare & & \\ & \blacksquare & \\ & & \ddots \end{pmatrix}. \tag{1}$$

The authors of [5] argued that due to this structure an optimal solution to the discrimination problem can be obtained by the simple composition of the optimal solutions in each block. The optimal solution in two dimensions is known, since only in the case of two pure states the solution is not obvious and this case was solved by Jaeger and Shimony [4].

Our aim is to provide a systematic method for finding such structures. We start with a formal definition of a block-diagonal structure: For a set of operators $\mathcal{O}$, a *common block-diagonal structure* (CBS) is a projection-valued measure $\{\Pi_k\}$ such that all operators in $\mathcal{O}$ commute with any Π_k. In other words, if the operators in $\mathcal{O}$ have a CBS, they can be simultaneously decomposed in orthogonal subspaces, and a von-Neumann measurement $\{\Pi_k\}$ projects onto these subspaces. Having the measurement outcome 'k', the support of the states is reduced to $\Pi_k\mathcal{H}$ (the image of Π_k). Thus one can focus on performing the optimal measurement in this subspace.

A common block-diagonal structure is *at most n-dimensional* if the rank of all Π_k is at most n. In particular, the existence of an at most one-dimensional CBS for a set $\mathcal{O}$ of normal operators (a normal operator is an operator that commutes with its adjoint) is equivalent to the existence of a common basis, in which all operators in $\mathcal{O}$ are diagonal. It is well known (cf, e.g., chapter IX, theorem 11 in [17]) that for normal operators this is possible if and only if all operators in $\mathcal{O}$ mutually commute. We will present a commutator criterion to verify whether two operators have an at most *two*-dimensional CBS (2D-CBS). This criterion, which is simpler (from an operational point of view) than the one introduced by Laffey [2], is valid in the case of non-overlapping support only, but is sufficiently general in order to detect any two-dimensional block structure in the case of USD.

2.2. Diagonalizing Jordan bases: definition and existence

Let us first relate the idea of a 2D-CBS to a concept that is widely used in the analysis of USD, namely the concept of *Jordan (or canonical) bases* of subspaces (cf, e.g., [18]): let P_A and P_B be self-adjoint projectors. Then by virtue of the singular value decomposition, one can find orthonormal bases $\{|\alpha_i\rangle\}$ of $P_A\mathcal{H}$ and $\{|\beta_j\rangle\}$ of $P_B\mathcal{H}$, such that

$$\langle\alpha_i|\beta_j\rangle \equiv \langle\alpha_i|P_A P_B|\beta_j\rangle = 0 \quad \text{for} \quad i \neq j, \tag{2a}$$

while for $i \leqslant \min\{\text{rank}\, P_A, \text{rank}\, P_B\}$,

$$\langle\alpha_i|\beta_i\rangle \equiv \langle\alpha_i|P_A P_B|\beta_i\rangle \equiv \cos\vartheta_i \geqslant 0 \tag{2b}$$

for some $0 \leqslant \vartheta_i \leqslant \pi/2$. The bases $\{|\alpha_i\rangle\}$ and $\{|\beta_j\rangle\}$ are called *Jordan bases* of the subspaces $P_A\mathcal{H}$ and $P_B\mathcal{H}$ and $\{\vartheta_i\}$ are the corresponding (unique) *Jordan angles*. The first equation expresses the bi-orthogonality of the Jordan bases. Note that in the case of degenerate Jordan angles (i.e., not all Jordan angles are different) or if $|\text{rank}\, P_A - \text{rank}\, P_B| \geqslant 2$, the Jordan bases are not unique.

For the analysis of USD, it turns out to be fruitful to consider density operators, which are diagonalized by a pair of Jordan bases [11]. For two normal operators A and B, *diagonalizing*

Jordan bases are Jordan bases of supp A and supp B, which diagonalize A and B, respectively. Of course, such diagonalizing Jordan bases do not always exist. As mentioned in [19], the existence of such bases implies the presence of a 2D-CBS, since the pairs $\{|\alpha_i\rangle, |\beta_i\rangle\}$ span mutually orthogonal two-dimensional subspaces. However, the converse is in general not true. It is possible that already in two dimensions no pair of diagonalizing Jordan bases exists. Consider the positive semi-definite matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{3}$$

Then up to some complex phases, the only orthonormal basis of supp A that diagonalizes A is the canonical basis $\{(1, 0), (0, 1)\}$ while supp B is spanned by $(1, 1)$. But $(1, 1)$ is orthogonal to neither $(1, 0)$ nor $(0, 1)$, i.e., no diagonalizing Jordan bases exist.

The exact relation between 2D-CBS and diagonalizing Jordan bases is given by the following

Lemma 1. *Let A and B be normal operators acting on $\mathcal{H}$. Then diagonalizing Jordan bases of A and B can be found if and only if a 2D-CBS of A and B exists and $[A, ABA] = 0$ and $[B, BAB] = 0$.*

Proof. Assume that diagonalizing Jordan bases of A and B exist. Then their structure readily provides an appropriate 2D-CBS. Furthermore, by writing A and B in diagonalizing Jordan bases, i.e., $A = \sum_i a_i|\alpha_i\rangle\langle\alpha_i|$ and $B = \sum_j b_j|\beta_j\rangle\langle\beta_j|$, and using equations (2), it is easy to verify that $[A, ABA] = 0$ and $[B, BAB] = 0$ holds.

For the contrary it is enough to prove the assertion in each subspace $\Pi_k\mathcal{H}$, where $\{\Pi_k\}$ is a 2D-CBS of A and B. Since A and B commute with all projectors Π_k, in each subspace the operators $A_k \equiv \Pi_k A \Pi_k$ and $B_k \equiv \Pi_k B \Pi_k$ are again normal. First suppose that A_k has a maximal rank, i.e., rank 2. Since A_k has full rank in $\Pi_k\mathcal{H}$, the condition $0 = \Pi_k[A, ABA]\Pi_k = A_k[A_k, B_k]A_k$ is equivalent to $\Pi_k[A_k, B_k]\Pi_k \equiv [A_k, B_k] = 0$, i.e., both operators can be diagonalized simultaneously and hence in particular diagonalizing Jordan bases exist. (An analogous argument holds if B_k has a maximal rank.) The remaining non-trivial case is that both operators have rank 1, in which case the diagonalizing Jordan bases are given by the vector spanning the support of each operator. □

Note that commutators of the form $[A, AXA]$ can always be rewritten as $A[A, X]A$, i.e., in the above lemma one could equivalently write the conditions $A[A, B]A = 0$ and $B[A, B]B = 0$.

2.3. Construction of diagonalizing Jordan bases

It is a simple observation that if diagonalizing Jordan bases for two normal operators A and B exist, then necessarily all commutators of the structure $[A, ABA]$, $[A, AB^2A]$ and so forth vanish (see the proof of lemma 1). In the following lemma we will state that certain of these commutators already suffice to explicitly construct a pair of diagonalizing Jordan bases.

Lemma 2. *Let A and B be self-adjoint operators on $\mathcal{H}$ with $[A, ABA] = 0$, $[A, AB^2A] = 0$ and $[B, BA^2B] = 0$. Furthermore, denote by $\{|k\rangle\}$ an orthogonal basis of* supp A *which simultaneously diagonalizes A, ABA and AB^2A.*

Then there exists vectors $\{|v\rangle\}$, such that (up to normalization) $\{A|k\rangle\}$ and $\{BA|k\rangle\} \cup \{|v\rangle\}$ are diagonalizing Jordan bases of A and B.

Proof. First note that all vectors $BA|k\rangle$ are mutually orthogonal (or trivial), since the basis $\{|k\rangle\}$ diagonalizes $ABBA$. Now consider the following expression:

$$\begin{aligned} w_k B(BA|k\rangle) &= BBA(ABA|k\rangle) \\ &= BA^2BBA|k\rangle \\ &= v_k BA|k\rangle, \end{aligned} \tag{4}$$

where w_k denotes the eigenvalue of ABA for $|k\rangle$ and v_k denotes the eigenvalue of AB^2A for $|k\rangle$. In the second step we used $[B, BA^2B] = 0$. The right-hand side can only vanish if $BA|k\rangle = 0$. Hence due to equation (4), $BA|k\rangle \in \operatorname{supp} B$ is either trivial or is an eigenvector of B. Furthermore, one readily finds eigenvectors $|\nu\rangle \in \operatorname{supp} B$ of B that complete the orthogonal basis of $\operatorname{supp} B$. These vectors are also orthogonal to all $A|k\rangle$, since by construction $b_\nu\langle\nu|A|k\rangle = \langle\nu|BA|k\rangle = 0$, where $b_\nu \neq 0$ is the eigenvalue of B for $|\nu\rangle$. It remains to verify that $\{A|k\rangle\}$ and $\{BA|k\rangle\}$ are bi-orthogonal. But this follows from the fact that $\{|k\rangle\}$ diagonalizes ABA. □

Note that it is straightforward to extend this lemma to normal operators. However, we are mainly interested in application for USD and hence specialize the results of this section in the following form:

Theorem. *For two self-adjoint A and B operators on a Hilbert space of finite dimension with* $\operatorname{supp} A \cap \operatorname{supp} B = \{0\}$ *the following statements are equivalent: (i) A and B have a 2D-CBS; (ii) diagonalizing Jordan bases of A and B exist; (iii)* $[A, ABA] = 0, [B, BA^2B] = 0$ *and* $[A, AB^2A] = 0$.

Proof. Remember that (ii) ⇒ (i) follows from the structure of Jordan bases (see lemma 1), and also (ii) ⇒ (iii) is a consequence of the properties of Jordan bases (i.e., that all commutators of the structure $[A, ABA]$, $[A, AB^2A]$ and so forth vanish). The implication (iii) ⇒ (ii) was proven in lemma 2. It remains to show that from (i) follows (ii). Due to lemma 1 this reduces to showing that $[A, ABA] = 0$ and $[B, BAB] = 0$ for the case where (i) holds and $\operatorname{supp} A \cap \operatorname{supp} B = \{0\}$. The condition of non-overlapping supports implies, together with (i), that $\operatorname{rank}(A_k) + \operatorname{rank}(B_k) \leqslant 2$, where $A_k = \Pi_k A \Pi_k$ and $B_k = \Pi_k B \Pi_k$, and $\{\Pi_k\}$ is a 2D-CBS of A and B. If either $\operatorname{rank}(A_k)$ or $\operatorname{rank}(B_k)$ is zero, the commutators $[A_k, A_k B_k A_k]$ and $[B_k, B_k A_k B_k]$ vanish trivially. They are also equal to zero for the remaining case of $\operatorname{rank}(A_k) = 1 = \operatorname{rank}(B_k)$. □

As soon as the supports of A and B overlap, in general, none of the commutators in the above theorem vanishes. But in such a situation one can make use of the fact that in two dimensions, the square of all commutators of the form $[A_k, B_k]$, $\left[A_k, B_k^2\right]$ and so forth is proportional to the identity operator. Laffey [2] showed that for positive operators the following set of commutators, given below, are already sufficient to prove the existence of a 2D-CBS.

Two positive semi-definite operators A and B have a 2D-CBS if and only if [2]

$$\begin{aligned} &[[A, B]^2, A] = 0, && [[B, A]^2, B] = 0, \\ &[[A, B^2]^2, A] = 0, && [[B, A^2]^2, B] = 0, \\ &[[A^2, B]^2, A] = 0, && [[B^2, A]^2, B] = 0. \end{aligned} \tag{5}$$

3. Application to USD

We now want to apply the above analysis to unambiguous discrimination of two mixed states ρ_1 and ρ_2. We denote the combination of the density operator and the according *a priori* probability by $\gamma_\mu = p_\mu \rho_\mu$, such that $\operatorname{tr} \gamma_\mu < 1$ ($\mu = 1, 2$). For technical reasons (see the map τ_0 below) we also allow that the *a priori* probabilities do not sum up to 1, $\operatorname{tr}(\gamma_1) + \operatorname{tr}(\gamma_2) \leqslant 1$.

3.1. Preservation of block structures under reduction of USD

In the above theorem the density operators need to satisfy the condition $\operatorname{supp} \gamma_1 \cap \operatorname{supp} \gamma_2 = \{0\}$, which in general is not the case. The first reduction theorem in [8], however, shows how to reduce any USD problem to that specific form. But one could imagine that this reduction might destroy an already present 2D-CBS, so that the combination of the first reduction theorem together with the above theorem would fail to detect certain block-diagonal structures. As we will see here, this is not the case and the application of any of the reductions in [8] preserves any CBS.

We repeat the reductions of [8] in the language of projectors. For a pair of positive semi-definite operators (γ_1, γ_2), let τ_0 be the (nonlinear) mapping

$$\tau_0 \colon (\gamma_1, \gamma_2) \mapsto \left(\gamma_1^0, \gamma_2^0\right), \tag{6}$$

where γ_μ^0 (with $\mu = 1, 2$) is the projection of γ_μ onto $(\ker \gamma_1 + \ker \gamma_2)$. In a similar fashion we define $\tau_\nu \colon (\gamma_1, \gamma_2) \mapsto \left(\gamma_1^\nu, \gamma_2^\nu\right)$ (with $\nu = 1, 2$) where

$$\gamma_\mu^\nu = P_\nu \gamma_\mu P_\nu + (\mathbb{1} - P_\nu)\gamma_\mu(\mathbb{1} - P_\nu). \tag{7}$$

Here, P_1 is the self-adjoint projector onto $(\ker \gamma_1 + \operatorname{supp} \gamma_2)$ and P_2 the projection onto $(\ker \gamma_2 + \operatorname{supp} \gamma_1)$. The reduction theorems in [8] now read as follows.

For $\tau \in \{\tau_0, \tau_1, \tau_2\}$, the pair (γ_1, γ_2) and the reduced pair $\tau(\gamma_1, \gamma_2)$ can be unambiguously discriminated with the same success probability [8].

What is relevant for our considerations is the fact that no reduction can destroy any CBS, i.e., a CBS $\{\Pi_k\}$ of (γ_1, γ_2) is also a CBS of $\tau(\gamma_1, \gamma_2)$ for all $\tau \in \{\tau_0, \tau_1, \tau_2\}$. In order to see this, it is enough to show that any of the projectors P_0, P_1 and P_2 (with P_0 denoting the projector onto $\ker \gamma_1 + \ker \gamma_2$) commutes with all Π_k. But this follows from the fact that the range of each of the projectors is the support of an operator that commutes with all Π_k (namely, $P_0\mathcal{H} = \operatorname{supp}(2\mathbb{1} - G_1 - G_2)$, $P_1\mathcal{H} = \operatorname{supp}(\mathbb{1} - G_1 + G_2)$ and $P_2\mathcal{H} = \operatorname{supp}(\mathbb{1} - G_2 + G_1)$, where G_μ is the projector onto $\operatorname{supp} \gamma_\mu$). Note, however, in contrast, that a CBS of $\tau(\gamma_1, \gamma_2)$ is not necessarily a CBS of (γ_1, γ_2), thus a reduction may give rise to new block-diagonal structures.

In order to check for a 2D-CBS it is necessary to first apply the reduction τ_0. If the reductions τ_1 and τ_2 are—from an operational point of view—feasible, then it is also worthwhile to apply those, since new 2D-CBS may arise.

3.2. Example: state comparison

We consider a special case of unambiguous state comparison 'two out of N' as defined in [16]. A source emits pure states $\{|\psi_1\rangle, \ldots, |\psi_N\rangle\}$, each of which appears with equal *a priori* probability $\frac{1}{N}$. We further assume that all states have the same (real) mutual overlap, $\langle\psi_i|\psi_j\rangle = \cos\vartheta$ for $i \neq j$. Given two of these pure states, the aim is to decide unambiguously whether the states are identical or not. This task is equivalent to the discrimination of

$$\gamma_1 = \frac{1}{N^2}\sum_{k=1}^{N} |\psi_k\psi_k\rangle\langle\psi_k\psi_k|, \tag{8}$$

$$\gamma_2 = \frac{1}{N^2}\sum_{k\neq l}^{N} |\psi_k\psi_l\rangle\langle\psi_k\psi_l|. \tag{9}$$

From the definition it follows that supp $\gamma_1 \cap$ supp $\gamma_2 = \{0\}$. Thus we can directly apply the theorem of section 2.3, i.e., we test whether it is true that $[\gamma_1, \gamma_1\gamma_2\gamma_1] = 0$, $\left[\gamma_1, \gamma_1\gamma_2^2\gamma_1\right] = 0$ and $\left[\gamma_2, \gamma_2\gamma_1^2\gamma_2\right] = 0$. For the first two commutators, it is sufficient to verify that $\omega_{kl} \equiv \langle\psi_k\psi_k|[\cdots]|\psi_l\psi_l\rangle = 0$ for any k and l. Here, $[\cdots]$ stands for any of the first two commutators. Obviously we have $\omega_{kl} = -(\omega_{lk})^*$ for all k and l, and since all overlaps are real, $\omega_{kk} = 0$. Due to the high symmetry, all ω_{kl} with $k \neq l$ must be equal. In particular, $\omega_{kl} = \omega_{lk} = -(\omega_{kl})^*$, and again due to reality of the overlaps, $\omega_{kl} = 0$ must hold.

It remains to test whether $\left[\gamma_2, \gamma_2\gamma_1^2\gamma_2\right] = 0$. This is equivalent to showing that $\gamma_2\left[\gamma_2 + \gamma_1, \gamma_1^2\right]\gamma_2 = 0$ or to showing that

$$\gamma_2(\gamma_2 + \gamma_1)\gamma_1^2\gamma_2 = \sum_{i,j;p,q} |\psi_i\psi_j\rangle\langle\psi_p\psi_q| A_{ij,pq} \tag{10}$$

is self-adjoint. For $i \neq j$ and also $p \neq q$, we have

$$A_{ij,pq} = \sum_{k,l,n,m} c_{ik}c_{jl}c_{kn}c_{ln}c_{nm}^2 c_{mp}c_{mq}, \tag{11}$$

with $c_{ij} \equiv \langle\psi_i|\psi_j\rangle = \cos\vartheta + (1 - \cos\vartheta)\delta_{ij}$. Otherwise, $A_{ij,pq} = 0$. First we find

$$\sum_k c_{ik}c_{kn} \propto \delta_{in} + \mu, \tag{12}$$

with some constant μ. Also, for $p \neq q$,

$$\sum_m c_{nm}^2 c_{mp}c_{mq} \propto \delta_{nq} + \delta_{np} + \sigma, \tag{13}$$

where σ is another constant. Hence for $i \neq j$ and $p \neq q$ we have

$$\begin{aligned} A_{ij,pq} &\propto \sum_n (\delta_{in} + \mu)(\delta_{jn} + \mu)(\delta_{np} + \delta_{nq} + \sigma) \\ &\propto \delta_{ip} + \delta_{iq} + \delta_{jp} + \delta_{jq} + \text{const.} \end{aligned} \tag{14}$$

In particular, $A_{ij,pq} = A_{pq,ij} \equiv (A_{pq,ij})^*$ holds, which demonstrates that $\gamma_2(\gamma_2 + \gamma_1)\gamma_1^2\gamma_2$ is self-adjoint and therefore $\gamma_2\left[\gamma_2 + \gamma_1, \gamma_1^2\right]\gamma_2 = 0$.

Thus we have shown that the symmetric state comparison 'two out of N' can be reduced to pure state discrimination. Note that this statement is in general not true for state comparison 'C out of N', with $C > 2$, i.e., the question whether C states taken from a set of N states (with equal overlaps) are identical or not. In this case the third commutator does not vanish before the reductions, and the corresponding state discrimination problem is not necessarily simplified to the pure state case.

4. Conclusions

In many practical situations of unambiguous state discrimination (USD) the pair of states that one wants to discriminate has a high symmetry which naturally gives rise to a two-dimensional common block-diagonal structure (2D-CBS) [5, 9, 11]. In this situation the optimal USD measurement has the very same 2D-CBS [11], where each block basically is given by the pure state solution of Jaeger and Shimony [4].

Here, we provided a tool to systematically identify whether a given USD task possesses such a structure. With the commutator relations presented in this paper it is easy to test whether a 2D-CBS for two self-adjoint operators with non-overlapping support exists. In order to derive these commutator relations, we studied the connection between the existence of a 2D-CBS and of diagonalizing Jordan bases. This also led to an explicit construction procedure for such bases.

We showed that the reduction method [8] for USD can only generate, but not destroy a 2D-CBS. Thus, applying the reductions as a first step ensures that the condition of non-overlapping support of the two operators is fulfilled.

We demonstrated the strength of the simple commutator relations by considering unambiguous state comparison [7, 9, 14–16], where it is easy to show that in completely symmetric situations for the specific case 'two out of N' a 2D-CBS exists.

Outlook. Note that the commutator relations in the theorem of section 2.3 are not symmetric in both operators (i.e., the missing commutator $[BAB, B]$ already vanishes). It would be interesting to understand the reason for this asymmetry. Furthermore, it would be useful to extend this concept to be applicable to more than two operators and also to the detection of larger block-diagonal structures (with respect to USD, e.g., four-dimensional structures would be interesting). In order to be operational, this would mean to extend the work by Watters [1] and Shapiro [20] (generalization to blocks of arbitrary dimension) and finding a *finite* set of commutators with possibly low order.

Acknowledgments

The authors would like to thank A Doherty, N Lütkenhaus, T Meyer, D Petz and R Unanyan for valuable discussions, and S Wehner for pointing out references [1, 20]. This work was partially supported by the EU Integrated Projects SECOQC and SCALA.

References

[1] Watters J F 1974 *Linear Algebra Appl.* **9** 103
[2] Laffey T J 1977 *Linear Algebra Appl.* **16** 189
[3] Helstrøm C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[4] Jaeger G and Shimony A 1995 *Phys. Lett.* A **197** 83
[5] Bennett C H, Mor T and Smolin J A 1996 *Phys. Rev.* A **54** 2675
[6] Sun Y, Bergou J A and Hillery M 2002 *Phys. Rev.* A **66** 032315
[7] Rudolph T, Spekkens R W and Turner P S 2003 *Phys. Rev.* A **68** 010301
[8] Raynal P, Lütkenhaus N and van Enk S J 2003 *Phys. Rev.* A **68** 022308
[9] Herzog U and Bergou J A 2005 *Phys. Rev.* A **71** 050301
[10] Raynal P and Lütkenhaus N 2005 *Phys. Rev.* A **72** 022342, 049909
[11] Bergou J A, Feldman E and Hillery M 2006 *Phys. Rev.* A **73** 032107
[12] Zhou X-F, Zhang Y-S and Guo G-C 2007 *Phys. Rev.* A **75** 052314
[13] Raynal P and Lütkenhaus N 2007 *Preprint* quant-ph/0702022
[14] Barnett S M, Chefles A and Jex I 2003 *Phys. Lett.* A **307** 189
[15] Chefles A, Andersson E and Jex I 2004 *J. Phys. A: Math. Gen.* **37** 7315
[16] Kleinmann M, Kampermann H and Bruß D 2005 *Phys. Rev.* A **72** 032308
[17] Gantmacher F 1959 *The Theory of Matrices* (New York: Chelsea Publication)
[18] Stewart G W and Sun J-g 1990 *Matrix Pertubation Theory* (San Diego, CA: Academic)
[19] Herzog U 2007 *Phys. Rev.* A **75** 052309
[20] Shapiro H 1979 *Linear Algebra Appl.* **25** 129

Publication E

Structural approach to unambiguous discrimination of two mixed quantum states
(submitted in March 2008) *(33 pages)*
M. Kleinmann, H. Kampermann, and D. Bruß

Structural approach to unambiguous discrimination of two mixed quantum states

M. Kleinmann[*], H. Kampermann, and D. Bruß

Institut für Theoretische Physik III, Heinrich-Heine-Universität Düsseldorf,
D-40225 Düsseldorf, Germany

Abstract

We analyze the optimal unambiguous discrimination of two arbitrary mixed quantum states. We show that the optimal measurement is unique and we present this optimal measurement for the case where the rank of the density operator of one of the states is at most 2 ("solution in 4 dimensions"). The solution is illustrated by some examples. The optimality conditions proven by Eldar *et al.* [Phys. Rev. A **69**, 062318 (2004)] are simplified to an operational form. As an application we present optimality conditions for the measurement, when only one of the two states is detected. The current status of optimal unambiguous state discrimination is summarized via a general strategy.

[*] electronic address: `kleinmann@thphy.uni-duesseldorf.de`

Contents

1 Introduction

Among the subtleties in quantum information processing and in quantum communication protocols are the properties that originate from the fact that in quantum mechanics non-orthogonal states cannot be discriminated perfectly. In the most naïve approach to quantum state discrimination – the minimum error discrimination (cf. Ref. [1, 2]) – this leads to the fact, that the identification of a state might be erroneous with some finite probability. Ivanovic [3] and Dieks [4] showed that one can avoid erroneous measurement results and that a measurement with a conclusive state identification is possible. In the case of non-orthogonal states, this strategy cannot work with a success probability of one. Peres showed in Ref. [5] how the optimum of this success probability can be achieved in the case of pure states, both having the same *a priori* probability. The discussion of the optimal unambiguous discrimination of two pure states was completed by Jaeger and Shimony in Ref. [6]. They derived the optimal solution for arbitrary *a priori* probabilities.

Although it was long ago stated to be an interesting problem [7], the unambiguous discrimination of *mixed* states did not attract much attention for a long time. This changed with an example introduced by Sun *et al.* in Ref. [8] and the first general analysis of the unambiguous discrimination of mixed states by Rudolph *et al.* in Ref. [9]. After that, several general results and special classes of optimal solutions were found, cf. Ref. [10, 11, 12, 13, 14, 15, 16]. While Bergou *et al.* derived in Ref. [10, 12] the optimal measurement for the unambiguous discrimination of a pure state and an arbitrary mixed state, no analysis so far did succeed to produce a general solution for the simplest instance of genuine mixed state discrimination, the discrimination of two mixed states where both density operators have a rank of 2. Also the simple question whether the optimal measurement in general is unique remained unanswered.

The answers to these two questions are among the central results of this contribution. The uniqueness of the optimal measurement is stated in Proposition 11 and the general solution for rank 2 density operators is presented in Sec. 6.

A valuable tool to approach both questions turned out to be a result by Eldar *et al.* in Ref. [17]. They showed necessary and sufficient conditions for a given measurement to be optimal. However, these conditions are difficult to verify, since the criterion implies the proof of the existence or non-existence of an operator with certain properties. In Corollary 9 we reformulate this criterion in such a way, that it can be directly applied to a given measurement. As a further immediate consequence of this Corollary we will be able to provide simple optimality conditions for a very special type of measurement: The measurement which only detects one out of the two states, cf. Sec. 5.1. It will also become possible to provide a simple proof and a deeper insight into the fidelity form measurement [13, 14], cf. Sec. 5.2.

Before we arrive at these results, we first provide an analysis of unambiguous state discrimination (USD), beginning in Sec 2, where we derive general results and continuing in Sec. 3, in which we specialize to the optimal case.

An analysis of the structure of the optimal measurement in particular yields Theorem 4. This Theorem is a cornerstone in order to prove the uniqueness of the optimal measurement and also provides a simple proof of the "second reduction" shown by Raynal *et al.* in Ref. [11]. We summarize and deepen the analysis carried out in Ref. [11] in Proposition 3, Proposition 6, and Lemma 7.

In Sec. 7 we will provide a generic scheme in order to approach a given optimization problem for USD. We conclude in Sec. 8.

2 Defining properties of USD

2.1 Main definitions

In quantum state discrimination of n quantum states it is usually assumed that the density operators $\rho_1, \dots, \rho_n$ of all possible input states are known, together with the probability $p_1, \dots, p_n$ of their occurrence. For $1 \leq \mu \leq n$, the *a priori* probability $p_\mu \geq 0$ and the corresponding density operator $\rho_\mu \geq 0$ with $\operatorname{tr}(\rho_\mu) = 1$ naturally combine to a *weighted density operator* $\gamma_\mu = p_\mu \rho_\mu$. Hence the trace of a weighted density operator γ_μ is the *a priori* probability of the state, $\operatorname{tr}(\gamma_\mu) = p_\mu$. Using this notation, the input states are represented by a family of positive semi-definite operators $\mathcal{S} = (\gamma_\mu)$. For a meaningful interpretation in terms of probability, we clearly need to have $\sum_\mu \operatorname{tr}(\gamma_\mu) = 1$. However, we will not require this normalization, as the subsequent definition and analysis is independent of it, and for certain statements (cf. e.g. Proposition 2) it will be useful to explicitly allow $\sum_\mu \operatorname{tr}(\gamma_\mu) < 1$.

In the following we will only consider the case of two input states, i.e., $\mu = 1, 2$. We restrict our analysis to finite-dimensional quantum systems, such that any possible quantum state of the system can be represented by a density operator which acts on a Hilbert space $\mathscr{H}$ of finite dimension. We will use the formalism of generalized measurements in which a physical measurement with M possible outcomes is described by a positive operator valued measure $\mathcal{E} = (E_1, \dots, E_M)$ on $\mathscr{H}$, i.e., by a family of M positive semi-definite operators which sum up to the identity, $\sum_k E_k = \mathbb{1}$.

Let us introduce our notation. We denote by $\ker A = \{|k\rangle \in \mathscr{H} \mid A|k\rangle = 0\}$ the kernel of an operator A, and we write $A\mathscr{H} = \{A|\phi\rangle \mid |\phi\rangle \in \mathscr{H}\}$ for its image. The support of a positive semi-definite operator ρ is written as $\operatorname{supp} \rho = \{|\phi\rangle \in \mathscr{H} \mid \exists \alpha > 0\colon \rho - \alpha|\phi\rangle\langle\phi| \geq 0\}$. Note, that the support of ρ is the orthocomplement of its kernel, $\operatorname{supp} \rho = (\ker \rho)^\perp$ and since ρ is self-adjoint, $\rho\mathscr{H} = \operatorname{supp} \rho$ holds.

By a *projector* we always mean an orthogonal projector, unless we explicitly state that the projector is oblique (cf. Lemma 18 in Appendix A). We use upper case Greek letters for orthogonal projectors, $\Sigma^\dagger = \Sigma = \Sigma^2$. The symbols "$\subset$" and "$\supset$" are used such that they also include equality, i.e., $\mathscr{A} = \mathscr{B}$ if and only if $\mathscr{A} \subset \mathscr{B}$ and $\mathscr{A} \supset \mathscr{B}$.

For a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$, we abbreviate

$$\operatorname{supp} \mathcal{S} \equiv \operatorname{supp}(\gamma_1 + \gamma_2) = \operatorname{supp} \gamma_1 + \operatorname{supp} \gamma_2, \tag{1}$$

for the collective support of $\mathcal{S}$, which is the physically relevant subspace for the discrimination task and $\ker \mathcal{S}$ for the common kernel of $\mathcal{S}$, which then is the trivial subspace,

$$\ker \mathcal{S} \equiv \ker(\gamma_1 + \gamma_2) = \ker \gamma_1 \cap \ker \gamma_2. \tag{2}$$

The task of optimal unambiguous discrimination of two mixed states is defined as follows.

Definition 1. *A positive operator valued measure $\mathcal{E} = (E_1, E_2, E_?)$ is called an* unambiguous state discrimination (USD) measurement *of a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$ if $\mathrm{tr}(E_2\gamma_1) = 0$ and $\mathrm{tr}(E_1\gamma_2) = 0$. The* success probability P_{succ} *of $\mathcal{E}$ of $\mathcal{S}$ is given by*

$$P_{\mathrm{succ}}(\mathcal{E}; \mathcal{S}) = \mathrm{tr}(E_1\gamma_1) + \mathrm{tr}(E_2\gamma_2). \tag{3}$$

A USD measurement $\mathcal{E}$ of $\mathcal{S}$ is optimal *if it has maximal success probability, i.e., if for any USD measurement $\mathcal{E}'$ of $\mathcal{S}$, $P_{\mathrm{succ}}(\mathcal{E}; \mathcal{S}) \geq P_{\mathrm{succ}}(\mathcal{E}'; \mathcal{S})$ holds. A USD measurement $\mathcal{E}$ of $\mathcal{S}$ is called* proper *if $\mathrm{supp}(E_1 + E_2) \subset \mathrm{supp}\,\mathcal{S}$.*

The condition $\mathrm{tr}(E_2\gamma_1) = 0$ is equivalent to $\mathrm{supp}\,E_2 \subset \ker\gamma_1$ and $\mathrm{tr}(E_1\gamma_2) = 0$ is equivalent to $\mathrm{supp}\,E_1 \subset \ker\gamma_2$. Thus if $(\ker\gamma_1 + \ker\gamma_2)$ is not trivial, it is simple to write down some non-trivial USD measurement for a given pair $\mathcal{S}$. In the next section we will see, that it is sufficient to consider proper USD measurements. But the set of proper USD measurements in particular is compact (cf. remark after Proposition 3) and hence there always exists at least one proper USD measurement, which maximizes the success probability.

2.2 Trivial subspaces

For any USD measurement $\mathcal{E} = (E_1, E_2, E_?)$ of $\mathcal{S} = (\gamma_1, \gamma_2)$ one readily constructs a proper USD measurement $\mathcal{E}' = (E_1', E_2', E_?')$ with the same marginal probabilities, i.e., $\mathrm{tr}(E_1\gamma_1) = \mathrm{tr}(E_1'\gamma_1)$ and $\mathrm{tr}(E_2\gamma_2) = \mathrm{tr}(E_2'\gamma_2)$. For that the most straightforward approach is to choose E_1' and E_2' to be the projection of E_1 and E_2 onto $\mathrm{supp}\,\mathcal{S}$ and to set $E_?' = \mathbb{1} - E_1' - E_2'$.

As an important feature of proper USD measurements we will show that the optimal proper USD measurement is unique (cf. Proposition 11). Such a statement of uniqueness clearly can only hold if we require that the measurement is proper. For illustrative reasons let us provide an example of an optimal USD measurement, which is not proper and where the measurement operators do not even commute with the projector onto $\mathrm{supp}\,\mathcal{S}$: We consider two non-orthogonal pure states with

$$\gamma_1 = \tfrac{1}{2}|1\rangle\langle 1|, \quad \gamma_2 = \tfrac{1}{2}|+\rangle\langle +|, \tag{4}$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and the measurement $\mathcal{E} = (E_1, E_2, \mathbb{1} - E_1 - E_2)$ with

$$E_\mu = (3 - 3/\sqrt{2})|e_\mu\rangle\langle e_\mu|, \tag{5}$$

where $|e_1\rangle = (|0\rangle - |1\rangle - |2\rangle)/\sqrt{3}$ and $|e_2\rangle = (\sqrt{2}|0\rangle + |2\rangle)/\sqrt{3}$. It is straightforward to verify, that this measurement is a USD measurement and has a success probability of $P_{\mathrm{succ}} = 1 - 1/\sqrt{2}$ as given by the optimal solution due to Peres [5].

The subspace $\ker\mathcal{S}$ cannot play any role in USD, since the support of γ_1 and γ_2 is orthogonal to this space. Similarly, the subspace $\mathrm{supp}\,\gamma_1 \cap \mathrm{supp}\,\gamma_2$ necessarily is orthogonal to the support of E_1 and E_2, since $\mathrm{supp}\,E_1 \subset \ker\gamma_2$ and $\mathrm{supp}\,E_2 \subset \ker\gamma_1$. The following proposition is a consequence of this observation:

Proposition 2 (cf. Theorem 1 in Ref. [11]). *Let $\mathcal{S} = (\gamma_1, \gamma_2)$ be a pair of weighted density operators. Denote by $\Pi_{\nparallel}$ the projector onto $(\ker\gamma_1 + \ker\gamma_2)$ and write $\mathcal{S}^{\nparallel} = (\Pi_{\nparallel}\gamma_1\Pi_{\nparallel}, \Pi_{\nparallel}\gamma_2\Pi_{\nparallel})$ for the projected pair. Let $q \geq 0$.*

Then $\mathcal{E}$ is a proper USD measurement for $\mathcal{S}$ with $P_{\mathrm{succ}}(\mathcal{E}; \mathcal{S}) = q$ if and only if $\mathcal{E}$ is a proper USD measurement for $\mathcal{S}^{\nparallel}$ with $P_{\mathrm{succ}}(\mathcal{E}; \mathcal{S}^{\nparallel}) = q$.

(Note, that $(\ker\gamma_1+\ker\gamma_2)$ is the orthocomplement of $(\operatorname{supp}\gamma_1\cap\operatorname{supp}\gamma_2)$, which can be considered to be the "parallel" part of the support of γ_1 and γ_2.)

Proof. If $\mathcal{E}$ is a USD measurement of $\mathcal{S}$, we have $\Pi_{\nparallel}E_\mu\Pi_{\nparallel}=E_\mu$ and so clearly $\operatorname{tr}(E_\mu\gamma_\nu)=\operatorname{tr}(E_\mu\Pi_{\nparallel}\gamma_\nu\Pi_{\nparallel})$ holds. We have that $\operatorname{supp}E_\mu\subset\operatorname{supp}\mathcal{S}$ and $\operatorname{supp}E_\mu\subset(\ker\gamma_1+\ker\gamma_2)$. Due to $\operatorname{supp}\mathcal{S}^{\nparallel}=\operatorname{supp}\mathcal{S}\cap(\ker\gamma_1+\ker\gamma_2)$, it follows that $\mathcal{E}$ is also proper for $\mathcal{S}^{\nparallel}$.

For the converse, since $\mathcal{E}$ is proper for $\mathcal{S}^{\nparallel}$ we have in particular $\Pi_{\nparallel}E_\mu\Pi_{\nparallel}=E_\mu$ and hence $\operatorname{tr}(E_\mu\Pi_{\nparallel}\gamma_\nu\Pi_{\nparallel})=\operatorname{tr}(E_\mu\gamma_\nu)$. Furthermore we have $\operatorname{supp}E_\mu\subset\operatorname{supp}\mathcal{S}^{\nparallel}=\operatorname{supp}\mathcal{S}\cap(\ker\gamma_1+\ker\gamma_2)$, i.e., $\mathcal{E}$ is proper for $\mathcal{S}$. □

2.3 The role of $E_?$

For the discussion of USD measurements it is useful to note that the measurement operator corresponding to the inconclusive result, $E_?$, already completely determines a proper USD measurement.

Proposition 3. *For an operator $E_?$ and a pair of weighted density operators $\mathcal{S}=(\gamma_1,\gamma_2)$ there exist operators E_1 and E_2, such that $\mathcal{E}=(E_1,E_2,E_?)$ is a proper USD measurement of $\mathcal{S}$, if and only if $E_?$ acts as identity on $\ker\mathcal{S}$, $E_?\geq 0$, $\mathbb{1}-E_?\geq 0$ and $\gamma_1(\mathbb{1}-E_?)\gamma_2=0$.*

Given $E_?$, the proper USD measurement $\mathcal{E}$ of $\mathcal{S}$ is unique.

Proof. It is straightforward to see that the conditions are necessary. The proof of sufficiency and uniqueness is constructive: Let us write Q_1 for the bijective oblique projector from $\ker\gamma_2\cap\operatorname{supp}\mathcal{S}$ to $\operatorname{supp}\gamma_1\cap(\ker\gamma_1+\ker\gamma_2)$ (for a brief introduction to bijective oblique projectors cf. Lemma 18 in Appendix A). Then we have for any proper USD measurement $E_2Q_1=0$ and $E_1=E_1Q_1$. Hence

$$E_1=Q_1^\dagger(E_1+E_2)Q_1=Q_1^\dagger(\mathbb{1}-E_?)Q_1 \tag{6}$$

is the only candidate for E_1, given $E_?$. Due to $\mathbb{1}-E_?\geq 0$, this construction ensures that $E_1\geq 0$. An analogous construction holds for E_2.

It remains to show that $E_1+E_2-(\mathbb{1}-E_?)=0$. We decompose the Hilbert space into the sum

$$\mathscr{H}=\ker\mathcal{S}\oplus(\operatorname{supp}\gamma_1+\operatorname{supp}\gamma_2). \tag{7}$$

With $\Pi_\perp$ the projector onto $\ker\mathcal{S}$, we have $E_\mu\Pi_\perp=0$ and since $E_?$ acts as identity on $\ker\mathcal{S}$, also $(\mathbb{1}-E_?)\Pi_\perp=0$ holds. Using, that by construction $\gamma_1E_\mu\gamma_2=0$, we furthermore have

$$\gamma_1[E_1+E_2-(\mathbb{1}-E_?)]\gamma_2=-\gamma_1(\mathbb{1}-E_?)\gamma_2=0. \tag{8}$$

From $\gamma_1(\mathbb{1}-E_?)\gamma_2=0$ and $\mathbb{1}-E_?\geq 0$ it follows that with $\Pi_\parallel$ the projector onto $\operatorname{supp}\gamma_1\cap\operatorname{supp}\gamma_2$, we have $(\mathbb{1}-E_?)\Pi_\parallel=0$. Furthermore, one verifies that $Q_1\gamma_1=(\mathbb{1}-\Pi_\parallel)\gamma_1$ and hence $\gamma_1(E_1+E_2)\gamma_1=\gamma_1E_1\gamma_1=\gamma_1(\mathbb{1}-E_?)\gamma_1$. A similar argument for γ_2 finishes the proof. □

Due to this Proposition 3 we sometimes refer to an operator $E_?$ as a proper USD measurement if it satisfies the conditions of the Proposition. From here it is also easy to prove, that the set of proper USD measurements is compact (e.g. using the operator norm induced by the inner product $\langle A, B\rangle = \mathrm{tr}(A^\dagger B)$): We only need to show the compactness for the set of operators which satisfy the conditions of Proposition 3. Obviously this set is bounded. It only remains to show that its complement (in the set of self-adjoint operators) is open. This follows from the necessary and sufficient properties provided by Proposition 3.

A proper USD measurement is already uniquely defined by $E_?(\gamma_2 - \gamma_1)E_?$ (as it will turn out below, in the optimal case this operator is in some sense much simpler than $E_?$ itself). Namely, with $\Pi_\perp$ the projector onto $\ker \mathcal{S}$ and $(\gamma_1 + \gamma_2)^-$ denoting the inverse of $(\gamma_1 + \gamma_2)$ on its support we have the identity

$$\begin{aligned} E_? = \Pi_\perp + (\gamma_1 + \gamma_2)^- \{ &\gamma_1\gamma_2 + \gamma_2\gamma_1 \\ &+ \sqrt{\gamma_1}\sqrt{\sqrt{\gamma_1}[\gamma_2 - E_?(\gamma_2 - \gamma_1)E_?]\sqrt{\gamma_1}}\sqrt{\gamma_1} \\ &+ \sqrt{\gamma_2}\sqrt{\sqrt{\gamma_2}[\gamma_1 - E_?(\gamma_1 - \gamma_2)E_?]\sqrt{\gamma_2}}\sqrt{\gamma_2} \\ &\}(\gamma_1 + \gamma_2)^-. \end{aligned} \tag{9}$$

In order to see this, first note that using $\sqrt{\gamma_1}(\mathbb{1} - E_?)\sqrt{\gamma_2} = 0$ and $E_? \geq 0$ the term in curly brackets can be rewritten as

$$(\gamma_1 + \gamma_2)^2 - \gamma_1^2 - \gamma_2^2 + \sqrt{\gamma_1}\sqrt{(\sqrt{\gamma_1}E_?\sqrt{\gamma_1})^2}\sqrt{\gamma_1} + \gamma_2 E_? \gamma_2. \tag{10}$$

Then due to $(\gamma_1 + \gamma_2)^-(\gamma_1 + \gamma_2) = \mathbb{1} - \Pi_\perp$ and once more $\gamma_1(\mathbb{1} - E_?)\gamma_2 = 0$ we see that the right hand side of Eq. (9) is given by

$$\begin{aligned} \Pi_\perp + (\gamma_1 + \gamma_2)^-(\gamma_1 + \gamma_2)[\mathbb{1} - (\mathbb{1} - E_?)](\gamma_1 + \gamma_2)(\gamma_1 + \gamma_2)^- \\ = \Pi_\perp + (\mathbb{1} - \Pi_\perp)E_?(\mathbb{1} - \Pi_\perp) \end{aligned} \tag{11}$$

This expression is equal to $E_?$, since for a proper measurement $E_?\Pi_\perp = \Pi_\perp$ holds.

Using the forthcoming Lemma 10, Eq. (9), and Proposition 3, it will become possible to reconstruct the optimal measurement given only the projective part of $E_?$. This projective part is given by $\ker(\mathbb{1} - E_?)$. It has a very specific structure, which originates in the condition $\gamma_1(\mathbb{1} - E_?)\gamma_2 = 0$. Let Π_μ denote the projector onto $\mathrm{supp}\,\gamma_\mu$ and $\Pi_\perp$ denote the projector onto $\ker \mathcal{S}$. For any proper measurement these projectors satisfy $\Pi_1(\mathbb{1} - E_?)\Pi_2 = 0$ and $(\mathbb{1} - E_?)\Pi_\perp = 0$, and hence Lemma 17 (Appendix A) applies, i.e., for any proper measurement,

$$\begin{aligned} \ker(\mathbb{1} - E_?) = \{\ker(\mathbb{1} - E_?) \cap \mathrm{supp}\,\gamma_1\} \\ + \{\ker(\mathbb{1} - E_?) \cap \mathrm{supp}\,\gamma_2\} + \ker \mathcal{S} \end{aligned} \tag{12}$$

holds. Although this result may seem to be quite technical, in certain situation it turns out to be a quite powerful tool.

3 Simple properties of optimal measurements

The following theorem makes a simple but fundamental statement about the structure of optimal measurements. It basically states that no vector, that is in the kernel of γ_1 or in the kernel of γ_2 must be in the support of $E_?$. This clearly gives an upper bound on the rank of $E_?$. On the other hand the condition $\gamma_1(\mathbb{1} - E_?)\gamma_2 = 0$ provides a lower bound on the rank of $E_?$. The second part of the Theorem states that in the optimal case these bounds coincide and fix the rank of $E_?$.

Theorem 4. *Let $\mathcal{E} = (E_1, E_2, E_?)$ be an optimal USD measurement for a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$. Then* $(\operatorname{supp} E_? \cap \ker \gamma_1) = (\operatorname{supp} E_? \cap \ker \gamma_2)$*.*

If $\mathcal{E}$ in addition is proper, then $\operatorname{supp} E_? \cap \ker \gamma_1 = \ker \mathcal{S}$ *and* $\operatorname{rank} E_? = \operatorname{rank} \gamma_1\gamma_2 + \dim \ker \mathcal{S}$*.*

(Remember, that the rank of an operator A is given by $\dim(A\mathscr{H}) \equiv \dim \mathscr{H} - \dim \ker A$, i.e., the number of strictly positive eigenvalues of $A^\dagger A$.)

Proof. Let $|\phi\rangle \in \operatorname{supp} E_? \cap \ker \gamma_1$. Then due to $|\phi\rangle \in \operatorname{supp} E_?$ there exists an $\alpha > 0$ such that $E_? - \alpha|\phi\rangle\langle\phi| \geq 0$. We define a new USD measurement by $\mathcal{E}' = (E_1, E_2 + \alpha|\phi\rangle\langle\phi|, E_? - \alpha|\phi\rangle\langle\phi|)$. From the optimality condition for $\mathcal{E}$, i.e., $P_{\text{succ}}(\mathcal{E}', \mathcal{S}) \leq P_{\text{succ}}(\mathcal{E}, \mathcal{S})$, we find $\alpha\langle\phi|\gamma_2|\phi\rangle \leq 0$ which only can hold if $\gamma_2|\phi\rangle = 0$. Since $|\phi\rangle \in \operatorname{supp} E_?$, $(\operatorname{supp} E_? \cap \ker \gamma_1) \subset (\operatorname{supp} E_? \cap \ker \gamma_2)$ follows. An analogous argument holds for the "$\supset$" part and finishes the proof of the first assertion.

From this result by intersection with $(\ker \gamma_1)$ one immediately finds $(\operatorname{supp} E_? \cap \ker \gamma_1) = (\operatorname{supp} E_? \cap \ker \mathcal{S})$. In the case of a proper measurement, however, $\operatorname{supp} E_? \supset \ker \mathcal{S}$ and hence $(\operatorname{supp} E_? \cap \ker \gamma_1) = \ker \mathcal{S}$ follows.

Let $E'_?$ denote $E_?$ projected onto $\operatorname{supp} \mathcal{S}$. Since the measurement is proper, $E_? - E'_?$ is the projector onto $\ker \mathcal{S}$ and $\operatorname{supp} E_? = \operatorname{supp} E'_? \oplus \ker \mathcal{S}$. From the previous results we have $E'_?\mathscr{H} \cap \ker \gamma_2 = \{0\}$ and $E'_?\gamma_2\mathscr{H} \cap \ker \gamma_1 = \{0\}$. Then due to Lemma 16 (Appendix A) if follows $\ker(\gamma_2 E'_?) = \ker E'_?$ and $\ker(\gamma_1 E'_? \gamma_2) = \ker(E'_?\gamma_2)$. Hence,

$$\begin{aligned}\dim \ker E'_? &= \dim \ker(\gamma_2 E'_?) = \dim \ker(E'_?\gamma_2) \\ &= \dim \ker(\gamma_1 E'_? \gamma_2) = \dim \ker(\gamma_1\gamma_2),\end{aligned} \tag{13}$$

where we used that $\dim \ker A = \dim \ker A^\dagger$ for any operator A and that $\gamma_1(\mathbb{1} - E'_?)\gamma_2 = \gamma_1(\mathbb{1} - E_?)\gamma_2 = 0$. □

3.1 Orthogonal subspaces

An important consequence of the first part of Theorem 4 is the following

Lemma 5. *Let $\mathcal{E} = (E_1, E_2, E_?)$ be an optimal USD measurement for a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$. Suppose that Π is a projector with $\Pi\mathscr{H} \subset (\ker \gamma_1 \cap \operatorname{supp} \mathcal{S})$.*

Then $E_1\Pi = 0$ if and only if $E_2\Pi = \Pi$.

Proof. The "if" part follows directly from $0 \leq \Pi E_? \Pi = -\Pi E_1 \Pi$. For the converse we have $\operatorname{supp} E_? \supset E_?\Pi\mathscr{H} = (\Pi - E_2\Pi)\mathscr{H} \subset \ker \gamma_1$ and thus due to Theorem 4, $E_?\Pi\mathscr{H} \subset \ker S$. But since $\ker S$ is orthogonal to $\Pi\mathscr{H}$, we have $\Pi E_? \Pi = 0$. Thus $0 = E_?\Pi = \Pi - E_2\Pi$. □

In particular let Σ_2 denote the projector onto $\ker \gamma_1 \cap \operatorname{supp} \gamma_2$. Then necessarily for any USD measurement $E_1\Sigma_2 = 0$ and hence by virtue of Lemma 5, for any optimal measurement $E_2\Sigma_2 = \Sigma_2$ holds. With Σ_1 denoting the projector onto $\ker\gamma_2 \cap \operatorname{supp}\gamma_1$ we obtain $E_1\Sigma_1 = \Sigma_1$ in an analogous way. These observations are at the core of the following

Proposition 6 (cf. Theorem 2 in Ref. [11]). *Let $\mathcal{S} = (\gamma_1, \gamma_2)$ be a pair of weighted density operators. Denote by $\Pi_{\rm skew}$ the projector onto $(\ker\gamma_1{+}\operatorname{supp}\gamma_2)\cap(\ker\gamma_2 + \operatorname{supp}\gamma_1)$ and write $\mathcal{S}^{\rm skew} = (\Pi_{\rm skew}\gamma_1\Pi_{\rm skew}, \Pi_{\rm skew}\gamma_2\Pi_{\rm skew})$ for the projected pair. Let $E_?$ and $E_?^{\rm skew}$ be two operators satisfying $E_?^{\rm skew} = E_? + (\mathbb{1} - \Pi_{\rm skew})$.*

Then $E_?$ is an optimal and proper USD measurement for $\mathcal{S}$, if and only if $E_?^{\rm skew}$ is an optimal and proper USD measurement for $\mathcal{S}^{\rm skew}$.

In this case, the failure probability *for $\mathcal{E}$ of $\mathcal{S}$ is the same as for $\mathcal{E}^{\rm skew}$ of $\mathcal{S}^{\rm skew}$,*

$$\operatorname{tr}(\gamma_1 + \gamma_2) - P_{\rm succ}(\mathcal{E}, \mathcal{S}) = \operatorname{tr}[\Pi_{\rm skew}(\gamma_1 + \gamma_2)] - P_{\rm succ}(\mathcal{E}^{\rm skew}, \mathcal{S}^{\rm skew}). \tag{14}$$

(Note, that $(\ker\gamma_1 + \operatorname{supp}\gamma_2) \cap (\ker\gamma_2 + \operatorname{supp}\gamma_1)$ is the orthocomplement of $(\Sigma_1 + \Sigma_2)\mathscr{H}$. For the projected pair $\mathcal{S}^{\rm skew}$, the spaces $\operatorname{supp}(\Pi_{\rm skew}\gamma_1\Pi_{\rm skew})$ and $\operatorname{supp}(\Pi_{\rm skew}\gamma_2\Pi_{\rm skew})$ are skew, where two spaces $\mathscr{A}$ and $\mathscr{B}$ are called *skew*, if $\mathscr{A} \cap \mathscr{B}^\perp = \{0\} = \mathscr{B} \cap \mathscr{A}^\perp$.)

Proof. Due to the discussion leading to the Proposition, for any optimal measurement $E_?$ we have $E_?\Pi_{\rm skew} = E_?$ and hence $\Pi_{\rm skew}(\mathbb{1} - E_?)\Pi_{\rm skew} = \mathbb{1} - E_?^{\rm skew}$. It follows that $\mathbb{1} - E_?^{\rm skew} \geq 0$ and that $\Pi_{\rm skew}\gamma_1\Pi_{\rm skew}(\mathbb{1} - E_?^{\rm skew})\Pi_{\rm skew}\gamma_2\Pi_{\rm skew} = 0$. Furthermore we find due to $\Pi_{\rm skew}\mathscr{H} \supset \ker\mathcal{S}$ that

$$\ker\mathcal{S}^{\rm skew} = \ker[(\gamma_1 + \gamma_2)\Pi_{\rm skew}] = \ker\Pi_{\rm skew} \oplus \ker\mathcal{S}, \tag{15}$$

where both terms in the direct sum are orthogonal. This shows that $E_?^{\rm skew}$ acts as a projector onto $\ker\mathcal{S}^{\rm skew}$. Since obviously $E_?^{\rm skew} \geq 0$ we have shown that $E_?^{\rm skew}$ is a proper USD measurement for $\mathcal{S}^{\rm skew}$. The converse, namely that $E_? = E_?^{\rm skew} - (\mathbb{1} - \Pi_{\rm skew})$ is a proper USD measurement of $\mathcal{S}$, in fact holds for any proper measurement $E_?^{\rm skew}$ of $\mathcal{S}^{\rm skew}$. This follows from Eq. (15) and by noticing that for $\mathcal{E}^{\rm skew} = (E_1^{\rm skew}, E_2^{\rm skew}, E_?^{\rm skew})$ the measurement defined by $E_?$ is given by $\mathcal{E} = (E_1^{\rm skew} + \Sigma_1, E_2^{\rm skew} + \Sigma_2, E_?)$.

In order to show that given $E_?$, the measurement $E_?^{\rm skew}$ is optimal, suppose, that $E_?^{\rm skew\prime}$ is proper and has a higher success probability than $E_?^{\rm skew}$. Then it is easy to see that $E_?' = E_?^{\rm skew\prime} - (\mathbb{1} - \Pi_{\rm skew})$ would yield a higher success probability for $\mathcal{S}$ than $E_?$, in contradiction to the assumption.

On the other hand, since $E_?\Pi_{\rm skew} = E_?$, any optimal and proper $E_?$ minimizes $\operatorname{tr}(E_?\Pi_{\rm skew}(\gamma_1 + \gamma_2)\Pi_{\rm skew})$. But this is minimal for optimal $E_?^{\rm skew}$, since $E_? = E_?^{\rm skew}\Pi_{\rm skew}$. □

Proposition 2 and Proposition 6 can be used independently from each other, in contrast to the original result in Ref. [11]. Proposition 2 and Proposition 6 provide a method to obtain all optimal measurements[1] for a given pair $\mathcal{S}$ by considering a different pair $\mathcal{S}'$ where $\dim\operatorname{supp}\mathcal{S} \geq \dim\operatorname{supp}\mathcal{S}'$. This is in particular useful, if $\dim\operatorname{supp}\mathcal{S}' \leq 4$, since in Section 6 we will provide an analytical solution for any such pair. If $\dim\operatorname{supp}\mathcal{S}' \leq 2$, then the general solution

[1] In the original work [11] it was only shown that one *may* choose the measurements in that specific way. Here we showed that all optimal measurements *must* have this structure.

can already be obtained due to the result by Jaeger and Shimony [6]. Also the pair $\mathcal{S}'$ might possess a two-dimensional common block diagonal structure which was not present in the original pair $\mathcal{S}$ and allows a solution of the problem (cf. Ref. [18]; for a simple criterion to detect such structures, cf. Ref. [19]). Apart from that, using both propositions all optimal measurements can be found by just considering pairs of states which do not possess any orthogonal (like $\operatorname{supp}\gamma_1 \cap \ker\gamma_2$) or parallel ($\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2$) components.

The following property simplifies actual calculations.

Lemma 7. *With the notations of Proposition 2 and Proposition 6 let $\tau_{\nparallel}$ denote the (non-linear) mapping from $\mathcal{S}$ to $\mathcal{S}^{\nparallel}$ and analogously τ_{skew} the mapping from $\mathcal{S}$ to $\mathcal{S}^{\mathsf{skew}}$.*

Then $\tau_{\nparallel} \circ \tau_{\nparallel} = \tau_{\nparallel}$, $\tau_{\mathsf{skew}} \circ \tau_{\mathsf{skew}} = \tau_{\mathsf{skew}}$ and $\tau_{\mathsf{skew}} \circ \tau_{\nparallel} = \tau_{\nparallel} \circ \tau_{\mathsf{skew}}$.

Sketch of Proof. We abbreviate $\tau[\mathcal{S}]_\mu$ for γ_μ after the application of τ, i.e., $(\tau[\mathcal{S}]_1, \tau[\mathcal{S}]_2) = \tau[\mathcal{S}]$. One verifies

$$\operatorname{supp}\tau_{\nparallel}[\mathcal{S}]_\mu = \operatorname{supp}\gamma_\mu \cap \operatorname{supp}\Pi_{\nparallel}, \tag{16}$$

and

$$\begin{aligned} \operatorname{supp}\tau_{\mathsf{skew}}[\mathcal{S}]_\mu &= \operatorname{supp}(\mathbb{1} - \Sigma_\mu)\gamma_\mu(\mathbb{1} - \Sigma_\mu) \\ &= \operatorname{supp}\gamma_\mu \cap \ker\Sigma_1. \end{aligned} \tag{17}$$

From the first equation we immediately get $\ker\tau_{\nparallel}[\mathcal{S}]_1 + \ker\tau_{\nparallel}[\mathcal{S}]_2 = \mathscr{H}$, i.e., $\tau_{\nparallel}$ is acts as identity on $\tau_{\nparallel}[\mathcal{S}]$. In order to show that τ_{skew} is idempotent one verifies that $\operatorname{supp}\tau_{\mathsf{skew}}[\mathcal{S}]_1 \cap \ker\tau_{\mathsf{skew}}[\mathcal{S}]_2 = \{0\}$.

Due to $\operatorname{supp}\tau_{\nparallel}[\mathcal{S}]_1 \cap \ker\tau_{\nparallel}[\mathcal{S}]_2 = \operatorname{supp}\gamma_1 \cap \ker\gamma_2$ it follows that

$$(\tau_{\mathsf{skew}} \circ \tau_{\nparallel})[\mathcal{S}] = (\Xi\gamma_1\Xi, \Xi\gamma_2\Xi), \tag{18}$$

where $\Xi = \Pi_{\mathsf{skew}}\Pi_{\nparallel} \equiv \Pi_{\nparallel}\Pi_{\mathsf{skew}}$. Analogously due to $\ker\tau_{\mathsf{skew}}[\mathcal{S}]_1 + \ker\tau_{\mathsf{skew}}[\mathcal{S}]_2 = \ker\gamma_1 + \ker\gamma_2$ we have

$$(\tau_{\nparallel} \circ \tau_{\mathsf{skew}})[\mathcal{S}] = (\Xi\gamma_1\Xi, \Xi\gamma_2\Xi), \tag{19}$$

and thus the third assertion holds. □

As an important consequence one can apply the mappings τ in any order and in particular due to $(\tau_{\mathsf{skew}} \circ \tau_{\nparallel})^{\circ 2} = \tau_{\mathsf{skew}} \circ \tau_{\nparallel}$, a second application of both mappings is never necessary.

The action of $\tau_{\nparallel}$ on $\mathcal{S} = (\gamma_1, \gamma_2)$ is non-trivial, if and only if $\operatorname{rank}(\gamma_1 + \gamma_2) < \operatorname{rank}\gamma_1 + \operatorname{rank}\gamma_2$. Similarly, the action of τ_{skew} is non-trivial if and only if $\operatorname{rank}\gamma_1 > \operatorname{rank}\gamma_1\gamma_2$ or $\operatorname{rank}\gamma_2 > \operatorname{rank}\gamma_1\gamma_2$. We call a pair of states $\mathcal{S}$ *strictly skew*, if $(\tau_{\mathsf{skew}} \circ \tau_{\nparallel})[\mathcal{S}] = \mathcal{S}$.

Let us briefly mention a convenient way to construct the mapping $\tau_{\mathsf{skew}} \circ \tau_{\nparallel}$. As shown in the proof of Lemma 7, we can write

$$\mathcal{S}' \equiv (\tau_{\mathsf{skew}} \circ \tau_{\nparallel})[\mathcal{S}] = (\Xi\gamma_1\Xi, \Xi\gamma_2\Xi), \tag{20}$$

with $\Xi = \mathbb{1} - \Pi_{\parallel} - \Sigma_1 - \Sigma_2$. Now let $(|s_{1i}\rangle)$ and $(|s_{2j}\rangle)$ be Jordan bases (cf. Appendix C) of $\operatorname{supp}\gamma_1$ and $\operatorname{supp}\gamma_2$, i.e., orthonormal bases of $\operatorname{supp}\gamma_1$ and $\operatorname{supp}\gamma_2$, respectively, such that $\langle s_{1k}|s_{2k}\rangle \geq 0$ and $\langle s_{1i}|s_{2j}\rangle = 0$ for $i \neq j$. Then

$\Pi_{\|} = \sum_{i\in\mathcal{X}} |s_{1i}\rangle\langle s_{1i}|$ and $\Sigma_\mu = \sum_{k\in\mathcal{Y}_\mu} |s_{\mu k}\rangle\langle s_{\mu k}|$, with $\mathcal{X} = \{k \mid \langle s_{1k}|s_{2k}\rangle = 1\}$, $\mathcal{Y}_1 = \{i \mid \forall j\colon \langle s_{1i}|s_{2j}\rangle = 0\}$, and $\mathcal{Y}_2 = \{j \mid \forall i\colon \langle s_{1i}|s_{2j}\rangle = 0\}$.

Summarizing Proposition 2 and Proposition 6, if $\mathcal{E}' = (E'_1, E'_2, E'_?)$ is an optimal and proper USD measurement of $\mathcal{S}' = (\tau_{\text{skew}} \circ \tau_{\nparallel})[\mathcal{S}]$, then $\mathcal{E} = (E'_1 + \Sigma_1, E'_2 + \Sigma_2, E'_? - \Sigma_1 - \Sigma_2)$ is an optimal and proper USD measurement of $\mathcal{S}$. The optimal success probability computes to

$$P_{\text{succ}}(\mathcal{E};\mathcal{S}) = P_{\text{succ}}(\mathcal{E}';\mathcal{S}') + \operatorname{tr}[(\Sigma_1 + \Sigma_2)(\gamma_1 + \gamma_2)]. \tag{21}$$

3.2 Classification of USD measurements

We want to introduce a classification of the different types of optimal measurements for USD. Given the dimension of $\operatorname{supp}\mathcal{S}$, the classification is according to the rank of the measurement operators. For a Hilbert space of dimension d, we consider the optimal and proper USD measurements $\mathcal{E} = (E_1, E_2, E_?)$ for pairs of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$. We restrict the analysis to the case, where τ_{skew} and $\tau_{\nparallel}$ act as identity on $\mathcal{S}$, i.e., to the case of *strictly skew* pairs. Then $\operatorname{rank}\gamma_1\gamma_2 = \operatorname{rank}\gamma_2 = \operatorname{rank}\gamma_1 \equiv r$ and $\dim\ker\mathcal{S} = d - 2r$ holds. All optimal measurements with $\operatorname{rank} E_1 = e_1$ and $\operatorname{rank} E_2 = e_2$ will be considered as one *type* of measurement, denoted by (e_1, e_2). As we will see in subsequent sections, the construction method of the known optimal measurement mainly depends on the type of the measurement. The symmetry of USD for exchanging the label of γ_1 and γ_2 makes it only necessary to develop a construction procedure for the case where e.g. $e_1 \leq e_2$. Thus a measurement *class* $[a, b]$ with $a \leq b$ denotes both measurement types (a, b) and (b, a). We now count the number of measurement types and measurement classes.

Since we consider proper measurements, we have $\operatorname{supp} E_1 \cap \operatorname{supp} E_2 = \{0\}$ and hence $e_1 + e_2 = \operatorname{rank}(E_1 + E_2)$ and $e_\mu \leq r$. Let us denote by δ the dimension of the projective part of $E_?$, i.e, $\delta = \dim\ker(\mathbb{1} - E_?)$. Then $e_1 + e_2 + \delta = d$ and $\delta \leq \operatorname{rank} E_?$. From Theorem 4 we have that $\operatorname{rank} E_? = r + (d - 2r)$. On the other hand, at least $\ker(\mathbb{1} - E_?) \supset \ker\mathcal{S}$, i.e., $\delta \geq d - 2r$. In summary we arrive at the constraints

$$e_1 \leq r, \quad e_2 \leq r, \quad \text{and} \quad r \leq e_1 + e_2 \leq 2r. \tag{22}$$

From the situation where γ_1 and γ_2 have a two-dimensional block diagonal structure, one can see that for any possible e_1 and e_2 which satisfy the constraints in Eq. (22), one can find a pair $\mathcal{S} = (\gamma_1, \gamma_2)$ such that an optimal measurement is of the type (e_1, e_2).

Counting the possible combinations to satisfy the conditions in Eq. (22), one finds

$$\#\text{types} = \tfrac{1}{2}(r+1)(r+2) \quad \text{and} \quad \#\text{classes} = \left\lfloor (\tfrac{r}{2} + 1)^2 \right\rfloor, \tag{23}$$

where #types denotes the number of measurement types and #classes the number of measurement classes. Here we used the floor function, $\lfloor x \rfloor = \max\{k \in \mathbb{Z} \mid k \leq x\}$.

Measurements of the type (e_1, e_2) with $e_1 + e_2 = r$ actually are von-Neumann measurements. (Obviously there are always r such measurement types.) This can be seen, since then $d - r = \operatorname{rank} E_? \geq \delta = d - e_1 - e_2 = d - r$, i.e., $\operatorname{rank} E_? = \dim\ker(\mathbb{1} - E_?)$ and hence $E_?$ is projective. But then $\operatorname{tr} E_1 + \operatorname{tr} E_2 = \operatorname{tr}(\mathbb{1} - E_?) = r = e_1 + e_2$ holds. Due to the positivity conditions $E_\mu \geq 0$ and

$\mathbb{1} - E_\mu \geq 0$, all eigenvalues of E_1 and E_2 are in the interval $[0, 1]$, and thus in our situation all eigenvalues are either 1 or 0. This proofs the assertion.

As we will see in Sec. 5.1 and Sec. 5.2, only two measurement classes are known in the general optimal case – the class $[r, r]$ and the special von-Neumann class $[0, r]$. These classes may occur for any $r \geq 1$ and thus in particular solve the two-dimensional case ($r = 1$) and "half" of the four-dimensional case ($r = 2$). The remaining two classes (one of which is von-Neumann) in four dimensions are solved in Sec. 6.1 and Sec. 6.2.

4 The optimality conditions by Eldar, Stojnic & Hassibi

Eldar, Stojnic, and Hassibi provided in Ref. [17] necessary and sufficient conditions for the optimality of a USD measurement[2]:

Theorem 8 (Eldar, Stojnic & Hassibi [17]). *Let $\mathcal{E} = (E_1, E_2, E_?)$ be a proper USD measurement for a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$. Denote by Λ_1 the projector onto $\ker \gamma_2 \cap \operatorname{supp} \mathcal{S}$ and by Λ_2 the projector onto $\ker \gamma_1 \cap \operatorname{supp} \mathcal{S}$. This measurement is optimal, if and only if one can find an operator Z such that for $\mu = 1, 2$,*

$$Z \geq 0, \quad Z E_? = 0, \tag{24a}$$

$$\Lambda_\mu (Z - \gamma_\mu) \Lambda_\mu \geq 0, \quad \textit{and} \quad \Lambda_\mu (Z - \gamma_\mu) E_\mu = 0. \tag{24b}$$

In Ref. [17], this statement was only proven for the case $\ker \mathcal{S} = \{0\}$. However, the generalization presented in Theorem 8 follows immediately from the original statement.

In Theorem 8 necessary and sufficient conditions for optimality where presented. However they are not operational, as the existence or non-existence of Z is difficult to prove. We show in Appendix B, that the unknown operator Z can be eliminated, and the above conditions can be re-expressed as follows:

Corollary 9. *With the preliminaries and notations as in Theorem 8, a proper measurement $\mathcal{E}$ of $\mathcal{S}$ is optimal if and only if*

$$(\Lambda_1 - \Lambda_2) E_? (\gamma_2 - \gamma_1) E_? (\Lambda_1 + \Lambda_2) \geq 0 \tag{25a}$$

$$(\Lambda_1 - \Lambda_2) E_? (\gamma_2 - \gamma_1) E_? (\mathbb{1} - E_?) = 0. \tag{25b}$$

The conditions for an optimal USD measurement are now expressed as a series of equations and positivity conditions on only $E_?$. Remember the fact that $E_?$ already completely determines a USD measurement (cf. Proposition 3).

The first condition in the above Corollary 9, Eq. (25a), relies on the fact, that a positive semi-definite operator in particular has to be self-adjoint. Thus, the condition in Eq. (25a) is only a compact notation for the three conditions

$$\Lambda_1 E_? (\gamma_2 - \gamma_1) E_? \Lambda_1 \geq 0, \tag{26a}$$

$$\Lambda_2 E_? (\gamma_1 - \gamma_2) E_? \Lambda_2 \geq 0, \tag{26b}$$

$$\Lambda_1 E_? (\gamma_2 - \gamma_1) E_? \Lambda_2 = 0. \tag{26c}$$

[2] Indeed Eldar *et al.* proved conditions for the optimality of a USD measurement for an arbitrary number of states.

(Obviously these conditions are sufficient for Eq. (25a). The necessity follows from multiplication of Eq. (25a) by $Q_\mu^\dagger$ from the left and Q_ν from the right. Here Q_μ are the bijective oblique projectors as defined in the proof of Proposition 3.)

The second equation, Eq. (25b), in Corollary 9 makes a statement about the projective part of $E_?$. This is the content of the following

Lemma 10. *Let $E_?$ be an optimal and proper USD measurement of a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$ with* $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$*. Denote by $\Pi_?$ the projector onto* $\operatorname{supp} E_?$ *and by Δ the projector onto* $\ker(\mathbb{1} - E_?)$*.*

Then $E_?(\gamma_2 - \gamma_1)E_? = \Pi_?(\gamma_2 - \gamma_1)\Pi_? = \Delta(\gamma_2 - \gamma_1)\Delta$.

Proof. Let $\Pi_\perp$ denote the projector onto $\ker \mathcal{S}$. Then due to $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$, we have $(\Lambda_1 - \Lambda_2)\mathscr{H} = (\mathbb{1} - \Pi_\perp)\mathscr{H}$. Due to $\Pi_\perp E_? \gamma_\mu = 0$, the optimality condition in Eq. (25b) hence reads $E_?(\gamma_2 - \gamma_1)E_?(\mathbb{1} - E_?) = 0$ or

$$E_?(\gamma_2 - \gamma_1)E_? = E_?(\gamma_2 - \gamma_1)E_?^2. \tag{27}$$

For the first equality we multiply this equation from the right by the inverse (on its support) of $E_?$ and in a second step from the left and obtain the equations

$$E_?(\gamma_2 - \gamma_1)\Pi_? = E_?(\gamma_2 - \gamma_1)E_?, \tag{28a}$$
$$\Pi_?(\gamma_2 - \gamma_1)\Pi_? = \Pi_?(\gamma_2 - \gamma_1)E_?. \tag{28b}$$

Since the right hand side of the first equation is self-adjoint, the assertion follows.

For the second equality we have $E_?\Delta = \Delta$ and $N \equiv E_?(\mathbb{1} - \Delta) = E_? - \Delta$ with $N\Delta = 0$. Thus $E_?(\mathbb{1} - E_?) = N(\mathbb{1} - N)$. But $\ker(\mathbb{1} - N) = \ker(\mathbb{1} - E_? + \Delta) = \{0\}$ and hence the optimality condition in Eq. (25b) reads $E_?(\gamma_2 - \gamma_1)N = 0$. Thus

$$E_?(\gamma_2 - \gamma_1)E_? = E_?(\gamma_2 - \gamma_1)\Delta \tag{29a}$$
$$\Delta(\gamma_2 - \gamma_1)E_? = \Delta(\gamma_2 - \gamma_1)\Delta \tag{29b}$$

holds, where in the second step we multiplied the first equation by Δ from the left. □

Lemma 10 is the key to prove the uniqueness of the optimal and proper USD measurement, since due to the identity in Eq. (9) we have seen that any USD measurement is solely defined by $E_?(\gamma_2 - \gamma_1)E_?$. Hence in the case of $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$, the optimal and proper USD measurement can be uniquely determined, given $\Pi_?$, the projector onto the support of $E_?$. But since the set of optimal and proper USD measurements of $\mathcal{S}$ is by virtue of Proposition 2 equal to the set of optimal and proper USD measurements of $\mathcal{S}^{\nparallel}$, having $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$ (cf. also Lemma 7), it remains to show that the support of $E_?$ is unique. This follows from the fact that the rank of $E_?$ is fixed by virtue of Theorem 4, together with the convexity of optimal and proper measurements. Namely, for any two optimal and proper USD measurements $E_?$ and $\tilde{E}_?$, also $\frac{1}{2}(E_? + \tilde{E}_?)$ is an optimal and proper USD measurement. But since $E_?$ and $\tilde{E}_?$ are positive semi-definite, $\operatorname{rank}(E_? + \tilde{E}_?) = \operatorname{rank} E_? = \operatorname{rank}\tilde{E}_?$ can only hold if $\operatorname{supp} E_? = \operatorname{supp}\tilde{E}_?$. Thus we have proven the following

Proposition 11. *For a given pair of weighted density operators, there exists exactly one optimal and proper USD measurement.*

5 Two special classes of optimal measurements

5.1 Single state detection

For certain pairs of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$ it may be advantageous to choose $\operatorname{tr}(E_1\gamma_1) = 0$, e.g. if $\operatorname{tr}(\gamma_1)$ is much smaller than $\operatorname{tr}(\gamma_2)$. We refer to this situation as *single state detection* of γ_2. In the classification scheme proposed in Sec. 3.2, the single state detection measurements can be identified with the class $[0, r]$, where $r = \dim \mathscr{H}/2$.

For a proper measurement, $\operatorname{tr}(E_1\gamma_1) = 0$ can only hold if already $E_1 = 0$. If the measurement is optimal then due to Lemma 5, $E_1 = 0$ implies $E_2 = \Lambda_2$. It follows that $E_? = \mathbb{1} - E_1 - E_2 = \mathbb{1} - \Lambda_2$ is a projector and hence satisfies the optimality condition in Eq. (25b). Thus the measurement is optimal if and only if Eq. (25a) holds, i.e.,

$$\Lambda_1 E_? (\gamma_2 - \gamma_1) E_? \Lambda_1 \geq 0. \tag{30}$$

Let us now assume that $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$. Then $(\mathbb{1} - \Lambda_2)\Lambda_1\mathscr{H} = \gamma_1\mathscr{H}$ and we arrive at the following

Proposition 12. *Let $\mathcal{E} = (E_1, E_2, E_?)$ be an optimal USD measurement for a pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$. Then $\operatorname{tr}(E_1\gamma_1) = 0$ if and only if $\gamma_1(\gamma_2 - \gamma_1)\gamma_1 \geq 0$.*

In this case the success probability is given by $P_{\text{succ}}(\mathcal{E}; \mathcal{S}) = \operatorname{tr}(\Lambda_2\gamma_2)$, and if $\mathcal{E}$ is proper, then $\mathcal{E} = (0, \Lambda_2, \mathbb{1} - \Lambda_2)$. ($\Lambda_2$ is the projector onto $\ker\gamma_1 \cap \operatorname{supp}\mathcal{S}$.)

Proof. Assume, that $\mathcal{E}$ is optimal and satisfies $\operatorname{tr}(E_1\gamma_1) = 0$. Then also for the corresponding proper measurement $\mathcal{E}' = (E_1', E_2', E_?')$ (cf. Sec. 2.2) we have $\operatorname{tr}(E_1'\gamma_1) = 0$ and hence $\gamma_1(\gamma_2 - \gamma_1)\gamma_1 \geq 0$ follows. For the contrary, we have already shown that the if $\gamma_1(\gamma_2 - \gamma_1)\gamma_1 \geq 0$ holds, then the proper measurement $\mathcal{E} = (0, \Lambda_2, \mathbb{1} - \Lambda_2)$ is an optimal measurement. But due to Proposition 11, this is the only optimal and proper measurement. Let now $\bar{\mathcal{E}} = (\bar{E}_1, \bar{E}_2, \bar{E}_?)$ be some optimal measurement, that is not proper. Then if the projection E_1 of $\bar{E}_1$ onto $\operatorname{supp}\mathcal{S}$ satisfies $\operatorname{tr}(E_1\gamma_1) = 0$, then necessarily also $\operatorname{tr}(\bar{E}_1\gamma_1) = 0$ holds. □

Let us consider the situation, where the success probability for the states ρ_1 and ρ_2 (both having unit trace) is analyzed in dependence of the *a priori* probability $0 < p_1 < 1$ of the state ρ_1, while the *a priori* probability of ρ_2 is $p_2 = 1 - p_1$. Then the optimality condition in Proposition 12 is satisfied, if and only if for any $|\varphi\rangle \in \operatorname{supp}\rho_1$,

$$(1 - p_1)\langle\varphi|\rho_2|\varphi\rangle \geq p_1\langle\varphi|\rho_1|\varphi\rangle. \tag{31}$$

If there exists a $|\varphi\rangle \in \operatorname{supp}\rho_1 \cap \ker\rho_2$ with $|\varphi\rangle \neq 0$, then this condition cannot be satisfied for any $p_1 > 0$. But if we assume $\operatorname{supp}\rho_1 \cap \ker\rho_2 = \{0\}$, single state detection of $p_2\rho_2$ is optimal if and only if $0 < p_1 \leq \ell_1$, where ℓ_1 is given by (with $|\varphi\rangle \in \operatorname{supp}\rho_1$ and $\langle\varphi|\varphi\rangle = 1$)

$$\ell_1 = \min_{|\varphi\rangle} \left\{ \frac{\langle\varphi|\rho_2|\varphi\rangle}{\langle\varphi|(\rho_1 + \rho_2)|\varphi\rangle} \right\} = \frac{\lambda_1}{1 + \lambda_1}, \tag{32}$$

where ($\sqrt{\rho_1}^{\,-}$ denotes the inverse of $\sqrt{\rho_1}$ on its support)

$$\lambda_1 = \min_{|\varphi\rangle} \langle\varphi|\sqrt{\rho_1}^{\,-}\rho_2\sqrt{\rho_1}^{\,-}|\varphi\rangle. \tag{33}$$

The minimum in the expression for λ_1 is given by the smallest non-vanishing eigenvalue of the operator $\sqrt{\rho_1}^- \rho_2 \sqrt{\rho_1}^-$ (remember, that we assumed $\operatorname{supp}\rho_1 \cap \ker \rho_2 = \{0\}$). Note that $\lambda_1 > 0$ and hence there always exists a finite parameter range for p_1, where single state detection of γ_2 is optimal.

An analogous construction yields ℓ_2, such that single state detection of γ_1 is optimal if and only if $0 < p_2 \leq \ell_2$.

5.2 Fidelity form measurement

An upper bound on the optimal success probability of USD was constructed by Rudolph, Spekkens and Turner in Ref. [9]. Let $|\gamma_\mu\rangle\langle\gamma_\mu|$ be a *purification* [20, 21] of γ_μ, i.e., a positive semi-definite operator of rank 1 acting on an extended Hilbert space $\mathscr{H} \otimes \mathscr{H}_{\text{aux}}$, such that the partial trace over $\mathscr{H}_{\text{aux}}$ yields back the original weighted density operator, $\operatorname{tr}_{\text{aux}} |\gamma_\mu\rangle\langle\gamma_\mu| = \gamma_\mu$. Since the partial trace can be implemented by physical means, the optimal unambiguous discrimination of $\mathcal{S} = (\gamma_1, \gamma_2)$ cannot have a higher success probability than $\mathcal{S}^{\text{pur}} = (|\gamma_1\rangle\langle\gamma_1|, |\gamma_2\rangle\langle\gamma_2|)$. But $\mathcal{S}^{\text{pur}}$ is a pair of pure states, for which the optimal success probability is known due to the result by Jaeger and Shimony [6]. The map from $\mathcal{S}$ to $\mathcal{S}^{\text{pur}}$, on the other hand, can only be performed physically in very special situations [22] and hence the success probability of $\mathcal{S}^{\text{pur}}$ in general only yields an upper bound. This bound is strongly related to the Uhlmann fidelity $\operatorname{tr}|\sqrt{\rho_1}\sqrt{\rho_2}|$ of $\rho_1 \equiv \gamma_1/\operatorname{tr}(\gamma_1)$ and $\rho_2 \equiv \gamma_2/\operatorname{tr}(\gamma_2)$ [23, 24]. The Uhlmann fidelity is the largest overlap between any purification of both states ρ_1 and ρ_2. Due to this relation the bound was named *fidelity bound* [13]. In Ref. [13, 14], necessary and sufficient conditions for the fidelity bound to be optimal where shown and the optimal measurement was constructed. In this section we summarize and extend these results.

We continue to assume $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$. Herzog and Bergou showed in Ref. [13] that the fidelity bound can be reached only if $E_?(\gamma_2 - \gamma_1)E_? = 0$. (From Corollary 9, it is obvious that any such measurement is optimal.) But due to Eq. (9) we find that any measurement with $E_?(\gamma_2 - \gamma_1)E_? = 0$ is given by

$$\begin{aligned} E_? &= \Pi_\perp + (\gamma_1 + \gamma_2)^- \{\gamma_1\gamma_2 + \gamma_2\gamma_1 + \sqrt{\gamma_1}F_1\sqrt{\gamma_1} + \sqrt{\gamma_2}F_2\sqrt{\gamma_2}\}(\gamma_1 + \gamma_2)^- \\ &= \mathbb{1} - (\gamma_1 + \gamma_2)^- \{\sqrt{\gamma_1}(\gamma_1 - F_1)\sqrt{\gamma_1} + \sqrt{\gamma_2}(\gamma_2 - F_2)\sqrt{\gamma_2}\}(\gamma_1 + \gamma_2)^-, \end{aligned} \tag{34}$$

where we abbreviated $F_1 = \sqrt{\sqrt{\gamma_1}\gamma_2\sqrt{\gamma_1}}$ and $F_2 = \sqrt{\sqrt{\gamma_2}\gamma_1\sqrt{\gamma_2}}$. The converse is also true:

Lemma 13. *Let $\mathcal{S} = (\gamma_1, \gamma_2)$ be a pair of weighted density operators with $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$ and let $\mathcal{E} = (E_1, E_2, E_?)$ be a proper USD measurement of $\mathcal{S}$. Then $E_?\gamma_2 E_? = E_?\gamma_1 E_?$ if and only if $E_?$ is given by Eq.* (34)*.*

Proof. It remains to show the "if" part. First we multiply the identity

$$(\gamma_1 + \gamma_2)(\gamma_1 + \gamma_2)^- \gamma_1 = \gamma_1 \tag{35}$$

from left by Q_1 as defined in the proof of Proposition 3, i.e., Q_1 is the bijective oblique projector from $\ker\gamma_2 \cap \operatorname{supp}\mathcal{S}$ to $\operatorname{supp}\gamma_1$. We then obtain due to Lemma 18 (cf. Appendix A) that $\gamma_1(\gamma_1 + \gamma_2)^-\gamma_1 = \gamma_1$ and thus $\gamma_2(\gamma_1 + \gamma_2)^-\gamma_1 =$

0. (One can show that $\gamma_1(\gamma_1+\gamma_2)^- = Q_1$.) From the polar decomposition of $\sqrt{\gamma_1}\sqrt{\gamma_2}$, it furthermore follows, that there exists a unitary transformation U, such that $\sqrt{\gamma_1}\sqrt{\gamma_2}U = F_1$ (and hence $U\sqrt{\gamma_1}\sqrt{\gamma_2} = F_2$, cf. also Ref. [14]). Thus

$$\begin{aligned} E_?\sqrt{\gamma_1} &= (\gamma_1+\gamma_2)^-\{0+\gamma_2\sqrt{\gamma_1}+\sqrt{\gamma_1}(\sqrt{\gamma_1}\sqrt{\gamma_2}U)+0\} \\ &= (\gamma_1+\gamma_2)^-\{\sqrt{\gamma_2}(U\sqrt{\gamma_1}\sqrt{\gamma_2})^\dagger+\gamma_1\sqrt{\gamma_2}\}U \\ &= E_?\sqrt{\gamma_2}U, \end{aligned} \tag{36}$$

i.e., we have $E_?\gamma_1E_? = E_?\gamma_2E_?$. □

We refer to the measurement characterized by Lemma 13 as *fidelity form measurement* due to the appearance of the operators F_1 and F_2, which satisfy $\operatorname{tr}|\sqrt{\gamma_1}\sqrt{\gamma_2}| = \operatorname{tr}F_1 = \operatorname{tr}F_2$. According to the classification in Sec. 3.2, the fidelity form measurements are a (strict) superset of the measurement class $[r,r]$, with $r = \dim\mathscr{H}/2$. (This can be seen from Lemma 10: in the class $[r,r]$ we have $\dim\ker(\mathbb{1}-E_?) = 0$ and hence in particular $E_?(\gamma_2-\gamma_1)E_? = 0$.)

Unfortunately, it is very rare that the operator given by Eq. (34) is part of a valid USD measurement. The following Proposition states necessary and sufficient criteria.

Proposition 14 (cf. Theorem 4 in Ref. [14]). *Let $\mathcal{S} = (\gamma_1,\gamma_2)$ be a pair of weighted density operators with $\operatorname{supp}\gamma_1\cap\operatorname{supp}\gamma_2 = \{0\}$. Then there exists a proper USD measurement $\mathcal{E} = (E_1,E_2,E_?)$ of $\mathcal{S}$ with $E_?$ given by Eq. (34) if and only if $\gamma_1-\sqrt{\sqrt{\gamma_1}\gamma_2\sqrt{\gamma_1}}\geq 0$ and $\gamma_2-\sqrt{\sqrt{\gamma_2}\gamma_1\sqrt{\gamma_2}}\geq 0$.*

If the measurement exists, it is optimal and the success probability is given by $P_{\text{succ}}(\mathcal{E},\mathcal{S}) = 1-2\operatorname{tr}|\sqrt{\gamma_1}\sqrt{\gamma_2}|$.

Proof. Due to the properties shown in the proof of Lemma 13, for $E_?$ given by Eq. (34) we have $\gamma_1(\mathbb{1}-E_?)\gamma_2 = 0$ and $\sqrt{\gamma_\mu}(\mathbb{1}-E_?)\sqrt{\gamma_\mu} = \gamma_\mu - F_\mu$. Furthermore, one can write $E_? = \Pi_\perp + AA^\dagger$, with (cf. Ref. [14])

$$A = (\gamma_1+\gamma_2)^-(\sqrt{\gamma_1}+\sqrt{\gamma_2}U)\sqrt{F_1}, \tag{37}$$

where U is a unitary transformation originating from the polar decomposition $\sqrt{\gamma_1}\sqrt{\gamma_2}U = F_1$ (cf. Lemma 13).

Due to these properties, the necessary and sufficient conditions on $E_?$ shown in Proposition 3 reduce to the assertion of the current Proposition. □

If the criterion in Proposition 14 is not satisfied, then the optimal measurement cannot be of the form as given by Eq. (34). Thus by virtue of Lemma 13, $E_?(\gamma_2-\gamma_1)E_? \neq 0$ holds and using Lemma 10 we have that $\ker(\mathbb{1}-E_?)\supsetneq\ker\mathcal{S}$. But due to Eq. (12) it follows that $\ker(\mathbb{1}-E_?)$ contains at least one vector either in $\operatorname{supp}\gamma_1$ or in $\operatorname{supp}\gamma_2$ (cf. Corollary 1 in Ref. [16]).

Similar to the discussion of the single state detection measurement in Sec 5.1, we ask for the values of the *a priori* probability $0 < p_1 < 1$ of ρ_1, for which the fidelity form measurement is optimal. The first condition in Proposition 14, $\gamma_1 - F_1 \geq 0$, is satisfied if and only if for any $|\varphi\rangle\in\operatorname{supp}\gamma_1$,

$$p_1\langle\varphi|\rho_1|\varphi\rangle \geq \sqrt{p_1(1-p_1)}\langle\varphi|R_1|\varphi\rangle \tag{38}$$

holds, where we abbreviated $R_1 = \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$. Thus $\gamma_1 - F_1 \geq 0$ if and only if $m_1 \leq p_1 < 1$, where m_1 is given by

$$m_1 = \frac{\mu_1^2}{1+\mu_1^2}, \tag{39}$$

with μ_1 the maximal eigenvalue of $\sqrt{\rho_1}^- R_1 \sqrt{\rho_1}^-$. With an analogous construction we get that $\gamma_2 - F_2 \geq 0$ if and only if $m_2 \leq p_2 \equiv 1 - p_1 < 1$. Then

$$m_1 \leq p_1 \leq 1 - m_2 \tag{40}$$

is the region where the fidelity form measurement is optimal. Note, that this region is empty when $m_1 + m_2 > 1$.

In summary, single state detection is optimal if and only if $\{(\gamma_1^2 - F_1^2 \leq 0)$ *or* $(\gamma_2^2 - F_2^2 \leq 0)\}$, while the fidelity form measurement is optimal if and only if $\{(\gamma_1 - F_1 \geq 0)$ *and* $(\gamma_2 - F_2 \geq 0\}$. The situations, where the optimal measurement is neither a single state detection measurement nor a fidelity form measurement seems to be related to the gap between "$A \geq B$" and "$A^2 \leq B^2$" for positive operators A and B. In the pure state case, however, A and B are of rank 1 and hence this gap does not exist. Indeed, in the pure state case, $\mu_1^2 = \mu_2^2 = \mathrm{tr}(\rho_1\rho_2) = \lambda_1 = \lambda_2$ (where λ_μ and ℓ_μ were defined at the end of Sec. 5.1). Thus $\ell_1 = m_1$ and $\ell_2 = m_2$ and hence either the single state detection or the fidelity form measurement is always optimal. This is exactly the solution for the pure state case, as given by Jaeger and Shimony in Ref. [6].

6 Solution in four dimensions

In this section we reduce the candidates for an optimal and proper USD measurement for the case where $\dim \operatorname{supp} \mathcal{S} = 4$ to a finite number. These candidates are obtained by finding the real roots of a high-order polynomial.

Due to Proposition 2 and Proposition 6, it is sufficient to discuss the case of strictly skew pairs $\mathcal{S} = (\gamma_1, \gamma_2)$ with $\ker \mathcal{S} = \{0\}$, i.e.,

$$\begin{aligned} \operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}, &\quad \ker\gamma_1 \cap \ker\gamma_2 = \{0\}, \\ \operatorname{supp}\gamma_1 \cap \ker\gamma_2 = \{0\}, &\quad \text{and} \quad \ker\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}. \end{aligned} \tag{41}$$

Then, following the discussion in Sec. 3.2, there are six types of optimal USD measurements which differ in the rank of the measurement operators E_1 and E_2. We list these possible types in Table 1.

The measurements of the class $[0,2]$ and the class $[2,2]$ where already extensively discussed in Sec. 5.1 and Sec. 5.2, respectively. For the class $[1,2]$, the kernel of $(\mathbb{1} - E_?)$ is one-dimensional. We will consider this class in Sec. 6.1. In the remaining case, where $\operatorname{rank} E_1 = 1 = \operatorname{rank} E_2$, the measurement is a von-Neumann measurement (cf. Sec 3.2). An example of this kind of measurement was first found in Ref. [16]. Sec. 6.2 is devoted for a general treatment of this class.

6.1 The measurement class $[1,2]$

In this class the dimension of $\ker(\mathbb{1} - E_?)$ is 1. According to Eq. (12), this kernel is either contained in $\operatorname{supp}\gamma_1$ or in $\operatorname{supp}\gamma_2$. Here we focus on $\ker(\mathbb{1} - E_?) \subset$

Table 1: Possible ranks of the measurement operators of the optimal USD measurement in four dimensions (cf. also Sec. 3.2).

$\operatorname{rank} E_1$	$\operatorname{rank} E_2$	$\operatorname{rank} E_?$	$\dim\ker(\mathbb{1}-E_?)$	class	cf. Sec.
0	2	2	2	$[0,2]$	5.1
1	2	2	1	$[1,2]$	6.1
2	2	2	0	$[2,2]$	5.2
1	1	2	2	$[1,1]$	6.2
2	1	2	1	$[1,2]$	6.1
2	0	2	2	$[0,2]$	5.1

$\operatorname{supp}\gamma_1$, i.e., to the measurement type (1,2); the case of $\ker(\mathbb{1}-E_?)\subset\operatorname{supp}\gamma_2$ follows along same lines. Thus there exists an orthonormal basis $(|\phi\rangle,|\phi^\perp\rangle)$ of $\operatorname{supp}\gamma_1$, such that

$$E_? = \nu|n\rangle\langle n| + |\phi\rangle\langle\phi|, \tag{42}$$

with $|n\rangle$ some normalized vector, orthogonal to $|\phi\rangle$ and $0<\nu<1$.

6.1.1 The necessary and sufficient conditions

Due to Proposition 3, the operator in Eq. (42) is a valid USD measurement, if and only if $\gamma_1(\mathbb{1}-E_?)\gamma_2=0$, i.e.,

$$\gamma_1|\phi^\perp\rangle\langle\phi^\perp|\gamma_2 = \nu\gamma_1|n\rangle\langle n|\gamma_2 \tag{43}$$

holds. This is equivalent to

$$\gamma_1|n\rangle = a\gamma_1|\phi^\perp\rangle \quad\text{and}\quad \gamma_2|n\rangle = b\gamma_2|\phi^\perp\rangle. \tag{44}$$

where $a=\langle\phi^\perp|n\rangle$ and $b=(\nu\langle n|\phi^\perp\rangle)^{-1}$. Remember, that due to Theorem 4 we have $|n\rangle\notin\ker\gamma_1$, i.e., $\langle n|\phi^\perp\rangle\neq 0$. On the other hand, from $\nu<1$ it follows $ab^*>1$ and thus $|n\rangle\notin\operatorname{span}\{|\phi^\perp\rangle\}$, i.e., $|n\rangle\notin\operatorname{supp}\gamma_1$. Given $|\phi^\perp\rangle$, we choose the phase of $|n\rangle$ such that $\langle n|\phi^\perp\rangle>0$.

The conditions in Corollary 9 can now be reduced to simple relations. By multiplying Eq. (25b) from the left with $|\phi^\perp\rangle\langle\phi^\perp|(\Lambda_1-\Lambda_2)^{-1}$, this first leads to

$$\langle n|\gamma_2-\gamma_1|n\rangle = 0. \tag{45}$$

Now it is straightforward to see that the conditions in Eq. (25) are equivalent to

$$E_?(\gamma_2-\gamma_1)|n\rangle = 0, \quad\text{and} \tag{46}$$

$$\langle\phi|\gamma_2-\gamma_1|\phi\rangle \geq 0. \tag{47}$$

By virtue of Eq. (44), we can re-express Eq. (46) in terms of $|\phi\rangle$ and $|\phi^\perp\rangle$, yielding

$$\sqrt{\langle\phi^\perp|\gamma_1|\phi^\perp\rangle}\langle\phi^\perp|\gamma_2|\phi\rangle = \sqrt{\langle\phi^\perp|\gamma_2|\phi^\perp\rangle}\langle\phi^\perp|\gamma_1|\phi\rangle, \quad\text{and} \tag{48}$$

$$\frac{a}{b} = \sqrt{\frac{\langle\phi^\perp|\gamma_2|\phi^\perp\rangle}{\langle\phi^\perp|\gamma_1|\phi^\perp\rangle}}. \tag{49}$$

The last equation enables us to construct $\nu|n\rangle\langle n|$ from

$$\begin{aligned}\sqrt{\nu}|n\rangle &= \sqrt{\nu}(\gamma_1+\gamma_2)^{-1}(\gamma_1+\gamma_2)|n\rangle \\ &= \sqrt{\nu}(\gamma_1+\gamma_2)^{-1}(a\,\gamma_1+b\,\gamma_2)|\phi^\perp\rangle \\ &= K|\phi^\perp\rangle,\end{aligned} \tag{50}$$

where in the last step we used $\sqrt{ab\nu}=1$ and we abbreviated

$$K=(\gamma_1+\gamma_2)^{-1}\left(\sqrt{a/b}\,\gamma_1+\sqrt{b/a}\,\gamma_2\right). \tag{51}$$

But due to Eq. (49), $\sqrt{a/b}$ is given in terms of $|\phi\rangle$ and $|\phi^\perp\rangle$. Note, that K has full rank and hence ensures $\nu\equiv\langle\phi^\perp|K^\dagger K|\phi^\perp\rangle>0$.

We summarize: *The optimal USD measurement is of type* $(1,2)$ *if and only if there exists an orthonormal basis* $(|\phi\rangle,|\phi^\perp\rangle)$ *of* $\operatorname{supp}\gamma_1$*, such that the conditions in Eq.* (47)*, in Eq.* (48) *and* $\langle\phi^\perp|K^\dagger K|\phi^\perp\rangle<1$ *are satisfied.*

6.1.2 Construction of a finite number of candidates for $E_?$

In the following we will show, that already Eq. (48) reduces the possible candidates of $\operatorname{span}\{|\phi\rangle\}$ to a finite number and hence the remaining positivity conditions can be easily checked. Eq. (48) is a complex equation. Thus the absolute value and the phase of the left hand side and the right hand side have to be identical. This leads to

$$\langle\phi|\gamma_2|\phi^\perp\rangle\langle\phi^\perp|\gamma_2|\phi\rangle\langle\phi^\perp|\gamma_1|\phi^\perp\rangle=\langle\phi|\gamma_1|\phi^\perp\rangle\langle\phi^\perp|\gamma_1|\phi\rangle\langle\phi^\perp|\gamma_2|\phi^\perp\rangle, \tag{52}$$

$$\langle\phi^\perp|\gamma_2|\phi\rangle\langle\phi|\gamma_1|\phi^\perp\rangle\geq 0. \tag{53}$$

Let $(|s_1\rangle,|s_2\rangle)$ be an orthonormal basis of $\operatorname{supp}\gamma_1$, such that $\gamma_1|s_i\rangle=g_{1i}|s_i\rangle$ with $g_{11}\geq g_{12}>0$. We abbreviate $g_{2i}=\langle s_i|\gamma_2|s_i\rangle>0$, $g_\mu=g_{\mu1}-g_{\mu2}$, and $g_{23}=\langle s_1|\gamma_2|s_2\rangle$. We ensure $g_{23}\geq 0$ by choosing a proper global phase of $|s_2\rangle$. In the case $g_1=0$ we use a basis where $g_{23}=0$.

First we consider, whether $|s_1\rangle$ or $|s_2\rangle$ is a candidate for $|\phi\rangle$. In either case, Eq. (48) can only be satisfied, if $g_{23}=0$. From Eq. (47) we find that $|s_1\rangle$ is a candidate only if $g_{21}\geq g_{11}$ and analogously, $|s_2\rangle$ only if $g_{22}\geq g_{12}$.

We now assume that neither of the above two cases is optimal. Then any of the remaining bases (apart from global phases) are parametrized by

$$|\phi\rangle=(1+x^2)^{-\frac{1}{2}}(|s_1\rangle+x\mathrm{e}^{i\vartheta}|s_2\rangle), \tag{54a}$$

$$|\phi^\perp\rangle=(1+x^2)^{-\frac{1}{2}}(x\mathrm{e}^{-i\vartheta}|s_1\rangle-|s_2\rangle), \tag{54b}$$

where $x\neq 0$ is real and $-\frac{\pi}{2}\leq\vartheta<\frac{\pi}{2}$.

Using these definitions, Eq. (53) can be written as

$$g_1g_{23}\sin\vartheta=0\quad\text{and}\quad xg_1(xg_2+g_{23}(x^2-1))\geq 0. \tag{55}$$

Let us first discuss the case where $g_{23}\neq 0$. In this situation we get[3] from Eq. (55) the phase $\vartheta=0$. The solutions of Eq. (52) are given by the real roots of the polynomial of degree six in x,

$$x^2g_1^2(x^2g_{21}+g_{22}-2xg_{23})-(xg_2+(x^2-1)g_{23})^2(x^2g_{11}+g_{12})=0. \tag{56}$$

[3] Remember, that we chose our basis such that $g_{23}\neq 0$ implies $g_1\neq 0$.

(Since $g_{23}^2 g_{12} > 0$, this polynomial cannot be trivial.)

It now remains to consider the special case, where $g_{23} = 0$, but neither $|\phi\rangle = |s_1\rangle$ nor $|\phi\rangle = |s_2\rangle$ is optimal. If $g_{23} = 0$, then Eq. (52) reads

$$x^2(g_1^2 g_{21} - g_2^2 g_{11}) = g_2^2 g_{12} - g_1^2 g_{22}. \tag{57}$$

Assume that this equation has some solutions where x is real (there might be infinitely many). None of these solutions leads to an optimal measurement, as we exclude by the following argumentation. Neither Eq. (47) nor Eq. (53) depend on ϑ. From the necessary and sufficient conditions for optimality (cf. end of Sec. 6.1.1), only the condition $\nu \equiv \langle\phi^\perp|K^\dagger K|\phi^\perp\rangle < 1$ remains to be satisfied. Since Eq. (49) does not depend on ϑ, also K is independent of ϑ.

Thus ν as a function of ϑ is of the form

$$\nu(\vartheta) = (1 + x^2)^{-1} \left(x^2 \kappa_{11} - 2\, x \,\mathrm{Re}(\mathrm{e}^{-i\vartheta} \kappa_{12}) + \kappa_{22}\right), \tag{58}$$

where we defined $\kappa_{ij} = \langle s_i|K^\dagger K|s_j\rangle$. In particular this function is continuous in ϑ. Assume, for a given $x \neq 0$ (satisfying Eq. (47) and Eq. (48)), there exists some $-\pi/2 \leq \vartheta < \pi/2$, such that $\nu(\vartheta) < 1$. Then there exists an $\epsilon > 0$, such that also $\vartheta + \epsilon < \pi/2$ and $\nu(\vartheta + \epsilon) < 1$. Hence for ϑ as well as $\vartheta + \epsilon$ all optimality conditions would be satisfied. But for both values, we get a different vector $|\phi\rangle$ and hence there would be two different operators $E_?$, both being optimal. This is a contradiction to Proposition 11. It follows that for $g_{23} = 0$, only $|\phi\rangle = |s_1\rangle$ and $|\phi\rangle = |s_2\rangle$ can yield an optimal solution.

6.1.3 Summary for measurement type $(1, 2)$

Let us briefly summarize the algorithm to obtain the optimal measurement $E_?$ for the case where $\operatorname{rank} E_1 = 1$ and $\operatorname{rank} E_2 = 2$.

We construct some basis $(|s_1\rangle, |s_2\rangle)$ of $\operatorname{supp} \gamma_1$ as described in the paragraph below Eq. (53). In a next step, we construct candidates for the basis $(|\phi\rangle, |\phi^\perp\rangle)$. There are two cases:

(i) $g_{23} = 0$. If $g_{21} \geq g_{11}$, then $(|\phi\rangle = |s_1\rangle, |\phi^\perp\rangle = |s_2\rangle)$ is a candidate. If $g_{22} \geq g_{12}$, then $(|\phi\rangle = |s_2\rangle, |\phi^\perp\rangle = |s_1\rangle)$ is a candidate.

(ii) $g_{23} \neq 0$. For any real root $x \neq 0$ of the polynomial in Eq. (56), the basis $(|\phi\rangle, |\phi^\perp\rangle)$ as defined in Eq. (54) is a candidate, where $\vartheta = 0$. A candidate in addition has to satisfy the second part of Eq. (55) and Eq. (47).

For any of the candidate bases (if any), we construct $\sqrt{\nu}|n\rangle$ using Eq. (50). At most one of the bases will satisfy $\nu \equiv \sqrt{\nu}\langle n|n\rangle\sqrt{\nu} < 1$. If such a basis exists, the optimal measurement is given by Eq. (42).

6.2 The measurement class $[1, 1]$

In Sec. 3.2 we have already seen, that if $\operatorname{rank} E_1 + \operatorname{rank} E_2 = \operatorname{rank} \gamma_1 (= \operatorname{rank} \gamma_2)$, then both E_1 and E_2 are projectors. Hence there are orthonormal bases $(|\psi_1\rangle, |\psi_1^\perp\rangle)$ of $\ker \gamma_2$ and $(|\psi_2\rangle, |\psi_2^\perp\rangle)$ of $\ker \gamma_1$ such that

$$E_1 = |\psi_1\rangle\langle\psi_1| \quad \text{and} \quad E_2 = |\psi_2\rangle\langle\psi_2|. \tag{59}$$

Since $\mathbb{1} - E_1 - E_2 \geq 0$, necessarily $\langle\psi_1|\psi_2\rangle = 0$ must hold. Using this notation, Eq. (26) is equivalent to

$$|\langle\psi_1^\perp|\psi_2\rangle|^2 \langle\psi_2|\gamma_2|\psi_2\rangle \geq \langle\psi_1^\perp|\gamma_1|\psi_1^\perp\rangle \tag{60a}$$

$$|\langle\psi_2^\perp|\psi_1\rangle|^2 \langle\psi_1|\gamma_1|\psi_1\rangle \geq \langle\psi_2^\perp|\gamma_2|\psi_2^\perp\rangle \tag{60b}$$

$$\langle\psi_2^\perp|\psi_1\rangle\langle\psi_1|\gamma_1|\psi_1^\perp\rangle = \langle\psi_2|\psi_1^\perp\rangle\langle\psi_2^\perp|\gamma_2|\psi_2\rangle, \tag{60c}$$

while Eq. (25b) is satisfied identically. (Note that these equations only follow if all vectors are normalized.)

6.2.1 Construction of a finite number of candidates for $E_?$

Let $(|k_{1i}\rangle)$ and $(|k_{2i}\rangle)$ be Jordan bases (cf. Appendix C) of $\ker\gamma_1$ and $\ker\gamma_2$, i.e., $(|k_{1i}\rangle)$ and $(|k_{2i}\rangle)$ are orthonormal bases of $\ker\gamma_1$ and $\ker\gamma_2$, respectively, such that $\langle k_{1l}|k_{2l}\rangle \geq 0$ and for $i \neq j$ one has $\langle k_{1i}|k_{2j}\rangle = 0$. Due to our assumptions in Eq. (41) we have $0 < \langle k_{1i}|k_{2i}\rangle < 1$. We choose $\langle k_{11}|k_{21}\rangle \geq \langle k_{12}|k_{22}\rangle$. In case of degenerate Jordan angles (i.e., $\langle k_{11}|k_{21}\rangle = \langle k_{12}|k_{22}\rangle$), these bases are not unique and we then choose bases, such that $\langle k_{21}|\gamma_1|k_{22}\rangle = 0$. We abbreviate

$$\begin{aligned} g_{1i} &= \langle k_{2i}|\gamma_1|k_{2i}\rangle, \quad g_{2i} = \langle k_{1i}|\gamma_2|k_{1i}\rangle, \\ g_{13} &= |\langle k_{21}|\gamma_1|k_{22}\rangle|, \quad g_{23} = |\langle k_{11}|\gamma_2|k_{12}\rangle|, \\ g_\mu &= g_{\mu 1} - g_{\mu 2}, \end{aligned} \tag{61}$$

and choose the global phases of the vectors $|k_{12}\rangle$ and $|k_{22}\rangle$ in such a way that still $\langle k_{12}|k_{22}\rangle > 0$ but also $g_{13}\mathrm{e}^{i\varphi} = \langle k_{21}|\gamma_1|k_{22}\rangle$ and $g_{23}\mathrm{e}^{-i\varphi} = \langle k_{11}|\gamma_2|k_{12}\rangle$, with $0 \leq \varphi < \pi$. We let $\varphi = 0$ if $g_{13} = 0$ or $g_{23} = 0$. Finally we define $c = \langle k_{12}|k_{22}\rangle / \langle k_{11}|k_{21}\rangle$.

After choosing the Jordan bases, let us first consider the case where $|\psi_1\rangle = |k_{21}\rangle$. Then one has (fixing the phases) $|\psi_2\rangle = |k_{12}\rangle$, $|\psi_1^\perp\rangle = |k_{22}\rangle$, and $|\psi_2^\perp\rangle = |k_{11}\rangle$. Eq. (60c) reduces to $cg_{23} = g_{13}$ and $\sin\varphi = 0$. Analogously, in the case $|\psi_1\rangle = |k_{22}\rangle$, we obtain $cg_{13} = g_{23}$ and $\sin\varphi = 0$.

We now consider the case, where $|\psi_1\rangle$ is not one of the basis vectors $|k_{2i}\rangle$. We fix the global phases of $|\psi_\mu\rangle$ and $|\psi_\mu^\perp\rangle$ and choose the parametrization

$$\begin{aligned} |\psi_1\rangle &= (1+x^2)^{-\frac{1}{2}}(|k_{21}\rangle + x\mathrm{e}^{i\vartheta}|k_{22}\rangle), \\ |\psi_1^\perp\rangle &= (1+x^2)^{-\frac{1}{2}}(x\mathrm{e}^{-i\vartheta}|k_{21}\rangle - |k_{22}\rangle), \\ |\psi_2\rangle &= (1+c^2x^2)^{-\frac{1}{2}}(-x\mathrm{e}^{-i\vartheta}c|k_{11}\rangle + |k_{12}\rangle), \\ |\psi_2^\perp\rangle &= (1+c^2x^2)^{-\frac{1}{2}}(-|k_{11}\rangle - x\mathrm{e}^{i\vartheta}c|k_{12}\rangle), \end{aligned} \tag{62}$$

with the real parameters $x \neq 0$ and $-\pi/2 \leq \vartheta < \pi/2$.

Using these definitions, Eq. (60c) reads

$$\begin{aligned} &(x^2c^2+1)^2\left(xg_1 - \mathrm{e}^{i(\vartheta+\varphi)}g_{13} + x^2\mathrm{e}^{-i(\vartheta+\varphi)}g_{13}\right) = \\ &\qquad c(x^2+1)^2\left(xcg_2 - \mathrm{e}^{i(\vartheta-\varphi)}g_{23} + x^2c^2\mathrm{e}^{-i(\vartheta-\varphi)}g_{23}\right) \end{aligned} \tag{63}$$

The imaginary part of this equation gives

$$\begin{aligned} &(x^2c^2+1)(1+x^2) \\ &\qquad \cdot \left[(x^2c^2+1)\sin(\vartheta+\varphi)g_{13} - (1+x^2)\sin(\vartheta-\varphi)cg_{23}\right] = 0, \end{aligned} \tag{64}$$

which can only hold if already the term in square brackets is zero. This is the case if and only if

$$A_1 \sin\vartheta = A_2 \cos\vartheta, \tag{65}$$

where

$$A_1 = (c(x^2+1)g_{23} - (x^2c^2+1)g_{13})\cos\varphi, \tag{66a}$$

$$A_2 = (c(x^2+1)g_{23} + (x^2c^2+1)g_{13})\sin\varphi. \tag{66b}$$

In order to get the solutions of Eq. (63), we consider now its real part,

$$\begin{aligned}(x^2c^2+1)^2(xg_1 + \cos(\vartheta+\varphi)g_{13}(x^2-1)) = \\ c(x^2+1)^2(xcg_2 + \cos(\vartheta-\varphi)g_{23}(x^2c^2-1)).\end{aligned} \tag{67}$$

Using the abbreviations

$$\begin{aligned}B_1 &= (x^2c^2+1)^2xg_1 - c(x^2+1)^2xcg_2,\\ B_2 &= \left[(x^2c^2+1)^2g_{13}(x^2-1) - c(x^2+1)^2g_{23}(x^2c^2-1)\right]\cos\varphi,\\ B_3 &= \left[(x^2c^2+1)^2g_{13}(x^2-1) + c(x^2+1)^2g_{23}(x^2c^2-1)\right]\sin\varphi,\end{aligned} \tag{68}$$

we get the equivalent expression

$$B_1 = B_3\sin\vartheta - B_2\cos\vartheta. \tag{69}$$

Taking the square of this equation, multiplied by $(A_1^2 + A_2^2)$, we obtain due to Eq. (65) the polynomial (with a degree of at most 8 in x^2)

$$B_1^2(A_1^2 + A_2^2) - (A_1B_2 - A_2B_3)^2 = 0. \tag{70}$$

This polynomial is trivial if and only if $g_{13} = g_{23} = 0$. (In order to see this, we consider the highest and lowest order term, which only can vanish, if $-(cg_{13} - g_{23})^2\cos^2\varphi = (cg_{13} + g_{23})^2\sin^2\varphi$ and if $-(cg_{23} - g_{13})^2\cos^2\varphi = (cg_{23} + g_{13})^2\sin^2\varphi$, respectively. But due to our particular choice of the bases, this can only hold if already $g_{13} = g_{23} = 0$.) It is straightforward to see, that in this case none of the conditions in Eq. (60) depend on ϑ. Now suppose there is a solution of these conditions with $x \neq 0$. Then any possible value of ϑ leads to a different, but optimal measurement, in contradiction to Proposition 11.

In any other situation we get from Eq. (70) a finite number of real solutions $x \neq 0$. The corresponding value for ϑ can be obtained as follows. If $A_1 \neq 0$, then from Eq. (65) we have $\vartheta = \arctan(A_2/A_1)$, while if $A_1 = 0$ and $A_2 \neq 0$, then $\vartheta = -\pi/2$. If $A_1 = A_2 = 0$, then $\sin\varphi = 0$ and

$$x^2 = \frac{1}{c}\frac{cg_{23} - g_{13}}{cg_{13} - g_{23}}, \tag{71}$$

where from $x^2 > 0$ it follows that $(cg_{23} - g_{13})(cg_{13} - g_{23}) > 0$. From Eq. (69) we have

$$2\cos\vartheta\, g_{13}g_{23}(cg_{23} - g_{13}) = xc(g_{23}^2g_1 - g_{13}^2g_2), \tag{72}$$

which can be used in order to obtain ϑ.

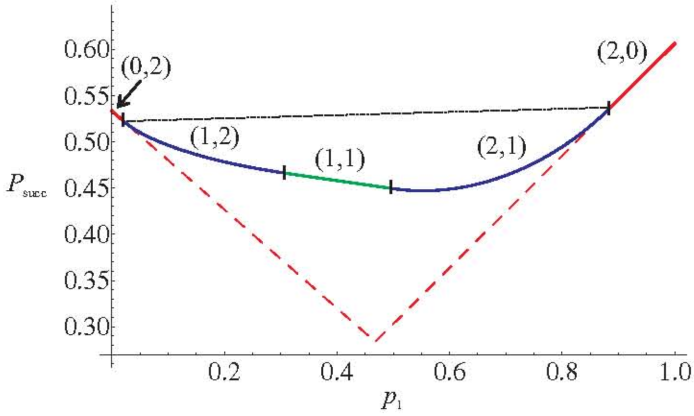

Figure 1: Optimal success probability of the states given in Eq. (73), depending on the relative probability p_1 of the occurrence of ρ_1 (solid line). The dashed lines denote the success probability of single state detection (lower bound) and the dotted line corresponds to a simple upper bound. In brackets we denote the measurement types as defined in Sec. 3.2.

6.2.2 Summary for measurement class $[1,1]$

In a first step we construct Jordan bases as described in the first paragraph of Sec. 6.2.1. Then we collect candidates for the bases $(|\psi_i\rangle, |\psi_i^\perp\rangle)$:

(i) If $\sin\varphi = 0$ and $cg_{23} = g_{13}$, then $(|\psi_1\rangle = |k_{21}\rangle, |\psi_1^\perp\rangle = |k_{22}\rangle)$ and $(|\psi_2\rangle = |k_{12}\rangle, |\psi_2^\perp\rangle = |k_{11}\rangle)$ is a candidate.

(ii) If $\sin\varphi = 0$ and $cg_{13} = g_{23}$, then $(|\psi_1\rangle = |k_{22}\rangle, |\psi_1^\perp\rangle = |k_{21}\rangle)$ and $(|\psi_2\rangle = |k_{11}\rangle, |\psi_2^\perp\rangle = |k_{12}\rangle)$ is a candidate.

(iii) For any root $x \neq 0$ of Eq. (70), we get a unique value for ϑ from Eq. (65) (if $A_1 \neq 0$ or $A_2 \neq 0$) or from Eq. (72) (if $A_1 = 0$ and $A_2 = 0$). If $A_1 \neq 0$ or $A_2 \neq 0$, then in addition Eq. (69) has to hold. For each (x, ϑ), we obtain the candidates from Eq. (62).

At most one of the candidates will satisfy Eq. (60a) and Eq. (60b). If such a candidate exists, the optimal measurement is provided by Eq. (59).

6.3 Examples

We want to discuss a few examples of USD in four dimensions, which demonstrate the structure of the previous results. We consider three examples which belong to case (vi) in the flowchart (Fig. 3), i.e., the considered states are strictly skew. In Fig. 1 the optimal success probability of the two states

$$\rho_1 = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \rho_2 = \frac{1}{45}\begin{pmatrix} 11 & 10 & 12 & 10 \\ 10 & 10 & 10 & 10 \\ 12 & 10 & 14 & 10 \\ 10 & 10 & 10 & 10 \end{pmatrix}, \tag{73}$$

is given in dependence of the *a priori* probability of p_1 of ρ_1 (solid line). Following the results of Sec. 5.1 and Sec. 5.2 we can directly calculate the probability

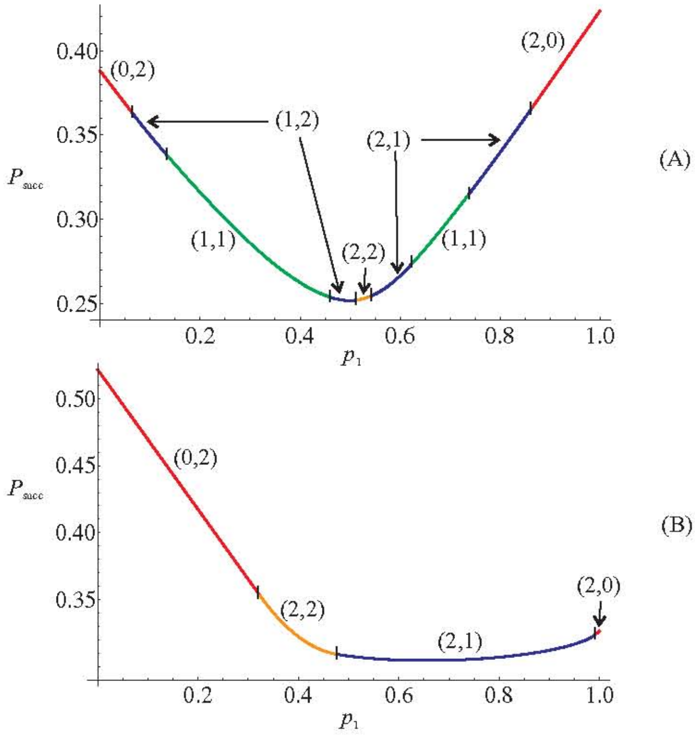

Figure 2: Optimal success probability depending on the relative probability p_1 of the occurrence of ρ_1'. (A): A "random" example (B): An asymmetric example, cf. Eq. (74). In brackets we denote the measurement types as defined in Sec. 3.2. For details see text in Sec. 6.3.

range where single state detection (class $[0,2]$) and the fidelity form measurement are optimal (roughly class $[2,2]$). The remaining optimal measurements belong to the classes $[1,2]$ or $[1,1]$ which are obtained by following Sec. 6.1 or Sec. 6.2, respectively. We also plotted the "bound triangle" which can be easily calculated for any pair of density operators. The lower bounds correspond to single state detection (dashed lines) and the upper bound connects the points where single state detection stops being optimal (dotted line). The latter one is an upper bound due to the convexity of the success probability function $P_{\text{succ}}(p_1)$.

The next example (Fig. 2 (A)) was found by generating "randomly" pairs of density operators. The data of the states is available as supplemental material. This example shows that the measurement types $(1,1)$, $(1,2)$, and $(2,1)$ can appear in more than one probability range. This is not possible for the other measurement types. It is also an example for a pair of states, where all possible measurement types can be optimal in some probability range.

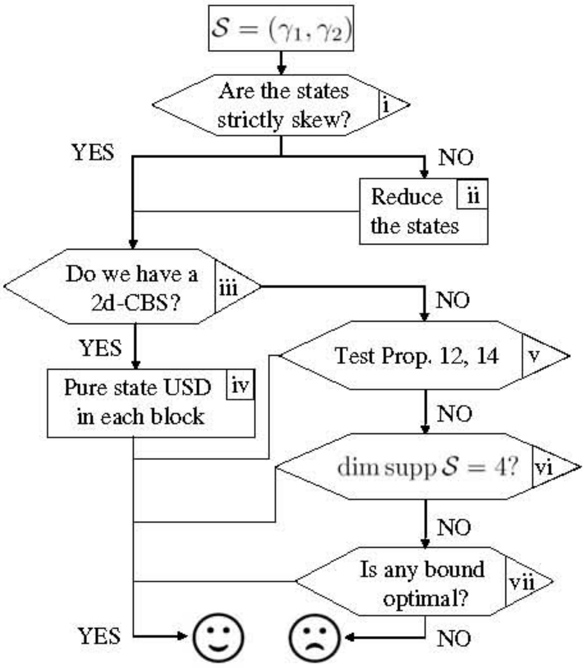
$\mathcal{S} = (\gamma_1, \gamma_2)$
Are the states strictly skew?
i
YES
NO
Reduce the states
ii
Do we have a 2d-CBS?
iii
NO
YES
Test Prop. 12, 14
v
Pure state USD in each block
iv
NO
$\dim \operatorname{supp} \mathcal{S} = 4$?
vi
NO
Is any bound optimal?
vii
YES
NO

Figure 3: Flow chart of a generic strategy to solve/reduce the USD of two mixed states. For details see text in Sec. 7.

The last example (Fig. 2 (B)), given by

$$\begin{aligned}
\rho_1' &= \left(\frac{1}{2} + \sqrt{\frac{5}{22}}\right)|0\rangle\langle 0| + \left(\frac{1}{2} - \sqrt{\frac{5}{22}}\right)|1\rangle\langle 1|, \\
\rho_2' &= \frac{5}{46}|v\rangle\langle v| + \frac{41}{46}|w\rangle\langle w|, \quad \text{with} \\
|w\rangle &= \frac{1}{2\sqrt{41}}\left((-1)^{1/7}\left((\sqrt{22}+2\sqrt{5})|0\rangle + (\sqrt{22}-2\sqrt{5})|1\rangle\right) + 2\sqrt{10}\,(|2\rangle + |3\rangle)\right), \\
|v\rangle &= \frac{1}{\sqrt{10}}\left(((-1)^{4/21})^*|0\rangle + (-1)^{17/21}|1\rangle + 2\sqrt{2}(-1)^{1/5}|2\rangle\right)
\end{aligned} \tag{74}$$

is devoted to show that there is no deeper general structure which optimal measurement classes can be connected with each other. Here, the fidelity form measurement directly follows the single state detection measurement. This example also shows a high asymmetry in the success probability function $P_{\rm succ}(p_1)$.

7 General strategy

In this section we want to summarize the known results for the unambiguous discrimination of two mixed states by suggesting a strategy in order to find the optimal success probability, cf. Fig. 3.

In step [i] we check whether the pair of weighted density operators $\mathcal{S} = (\gamma_1, \gamma_2)$ is *strictly skew*. A pair $\mathcal{S}$ is called strictly skew, if $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2 = \{0\}$

and $\operatorname{supp}\gamma_1 \cap \ker\gamma_2 = \{0\} = \ker\gamma_1 \cap \operatorname{supp}\gamma_2$. If the pair fails to be strictly skew, one can use the reduction method described in the discussion of Lemma 7 (Fig. 3 [ii]). After this reduction, it suffices to find the optimal measurement $\mathcal{E}'$ of the strictly skew pair $\mathcal{S}'$.

From now on we assume that the states are strictly skew. In the next step (Fig. 3 [iii]) we check with the commutator criteria shown in Ref. [19] whether the two states have an at most two-dimensional common block diagonal structure. If this is the case in Ref. [19] it was shown how to construct diagonalizing Jordan bases, which then can be used to compose the optimal measurement from the pure state case (Fig. 3 [iv]), as shown in Ref. [18, 19].

For states without an at most two-dimensional common block diagonal structure, optimality of one of the two generally solved measurement classes, i.e., single state detection (Proposition 12) or the fidelity form measurement (Proposition 14), can be checked (Fig. 3 [v]). Otherwise the optimal measurement can be calculated if the two states act on an effective four-dimensional Hilbert-space ($\dim\operatorname{supp}\mathcal{S} = 4$). Here the optimal measurement of the two remaining measurement classes $[1,2]$ and $[1,1]$ can be calculated and checked according to Sec. 6.1 and Sec. 6.2, respectively (Fig. 3 [vi]). In principle one could also find optimal measurements of states which have a four-dimensional common block diagonal structure, because the four-dimensional solution (Sec. 6) in each block would lead to the optimal measurement. Unfortunately, there is so far no known constructive method to identify such blocks. This question is left for further investigations.

The last possibility is to check optimality of upper and lower bounds on the optimal success probability (Fig. 3 [vii]). Examples of such bounds were presented e.g. in Ref [9, 13, 14, 25]. For some of these bounds, optimality conditions are known, while in the remaining cases one can use Corollary 9 in order to check for optimality.

If the procedure sketched here fails to deliver the optimal solution then there is up to now no systematic method known to find an analytic expression for the optimal USD measurement.

8 Conclusions

We analyzed the unambiguous discrimination of two mixed states ρ_1 and ρ_2 and the properties of an optimal measurement strategy. We first showed (cf. Proposition 3) that any unambiguous state discrimination measurement is completely determined by the measurement operator $E_?$, the operator which corresponds to the inconclusive measurement result. A further analysis for optimal measurements showed (cf. Theorem 4) that the rank of $E_?$ is determined by the structure of the input states, $\operatorname{rank} E_? = \operatorname{rank}\rho_1\rho_2$.

This fact leads to one of our main results, namely, that the optimal measurement for a given pair of mixed states and given *a priori* probabilities is unique (cf. Proposition 11). This uniqueness might have interesting consequences, e.g. in the analysis of the "complexity" of the optimal measurement. Interesting questions here are whether the optimal measurement is von-Neumann or whether the optimal measurement is non-local, as e.g. discussed in Ref. [26].

Eldar *et al.* in Ref. [17] provided necessary and sufficient conditions for a measurement to be optimal, but in many situations this result was not operational. We simplified these optimality conditions in Corollary 9, which now can be applied directly to a given measurement.

As an application of this result, we analyzed the single state detection case, where the measurement only detects one of the two states. Although this measurement may seem to be pathological, it turns out to be always optimal for a finite region of the *a priori* probabilities of the states. We derive an analytical expression for the bounds of this region.

Finally, we constructed the optimal measurement for the unambiguous discrimination of two mixed states having $\operatorname{rank}\rho_1 = 2 = \operatorname{rank}\rho_2$ (due to a results by Raynal *et al.* in Ref. [11], this construction can be extended to the case where one of the density operators has rank 2 and the other state has arbitrary rank). The solution splits into 6 different types. Although in principle this solution is analytical, in certain cases the roots of a high order (up to degree 8) polynomial are needed. Due to the complicated structure of this solution it may turn out to be quite difficult to analyze the next step, namely the discrimination of two mixed states with $\operatorname{rank}\rho_1 = 3 = \operatorname{rank}\rho_3$.

It would be interesting to find strategies in order to detect symmetries and solvable substructures in unambiguous state discrimination, e.g. to find 4-dimensional common block diagonal structures, analogously to the result for 2-dimensional common block diagonal structures in Ref. [19]. Also, there are only a few results on the optimal unambiguous discrimination on more than two states [27, 28, 29, 30, 17]. We think that several concepts presented in this contribution may also generalize to the discrimination of many states.

Acknowledgments

We would like to thank J. Bergou, T. Meyer, Ph. Raynal, and Z. Shadman for valuable discussions. This work was partially supported by the EU Integrated Projects SECOQC and SCALA.

A Appendix: Technical statements

The following two Lemmas are very basic. We put them here for completeness, but without a proof.

Lemma 15. *Let $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$ be subspaces of $\mathscr{H}$ with $\mathscr{A} \perp \mathscr{C}$ and $\mathscr{B} \perp \mathscr{C}$. Then $\mathscr{A} \cap (\mathscr{B} + \mathscr{C}) = \mathscr{A} \cap \mathscr{B}$.*

Lemma 16. *Let A and B be operators on $\mathscr{H}$ with $B\mathscr{H} \cap \ker A = \{0\}$. Then $\ker AB = \ker B$.*

The useful Eq. (12) is based on the following

Lemma 17. *Let X be an operator and (Π_k) be a family of projectors with $\ker(\sum_k \Pi_k) = \{0\}$. For $k \neq l$, assume that $\Pi_k X \Pi_l = 0$. (Note that $\Pi_k \Pi_l \neq 0$ is allowed.). Then $\ker X = \sum_k (\ker X \cap \Pi_k \mathscr{H})$.*

Proof. The "$\supset$" part is obvious. For the contrary let $|\Phi\rangle \in \ker X$. Then there exist vectors $|\varphi_k\rangle \in \Pi_k \mathscr{H}$ such that $|\Phi\rangle = \sum_k |\varphi_k\rangle$. We have

$$\sum_l \Pi_l X |\varphi_k\rangle = \Pi_k X |\varphi_k\rangle = \Pi_k X |\Phi\rangle = 0. \tag{75}$$

Since $\sum_k \Pi_k$ has a trivial kernel, it follows that $X|\varphi_k\rangle = 0$, i.e., $|\varphi_k\rangle \in \ker X$. □

The following Lemma lists properties of non-orthogonal projectors, i.e., idempotent operators which are not self-adjoint. Such operators where considered e.g. in Ref. [31, 32]. Our notion here differs slightly from the original definition, as we require the projector to be one-to-one. Each of the statements (ii) – (iv) is a valid definition for the bijective oblique projector, where (ii) is the formal definition, (iii) is an explicit construction and (iv) are the central properties for our purposes.

Lemma 18. *Let Λ and Π be two (self-adjoint) projectors on $\mathscr{H}$ with $\Lambda\mathscr{H} \cap (\Pi\mathscr{H})^\perp = \{0\}$ and $(\Lambda\mathscr{H})^\perp \cap \Pi\mathscr{H} = \{0\}$. For an operator Q the following statements are equivalent:*

(i) Q is the (unique) bijective oblique projector *from $\Lambda\mathscr{H}$ to $\Pi\mathscr{H}$.*

(ii) $Q^2 = Q$, $Q\mathscr{H} = \Pi\mathscr{H}$ and $\ker Q = \ker \Lambda$.

(iii) Q is the Moore-Penrose inverse of $\Lambda\Pi$.

(iv) $Q\Lambda = Q$, $\Pi Q = Q$, $\Lambda Q = \Lambda$ and $Q\Pi = \Pi$.

Proof. (ii) formally defines (i). In order to show that (iii) follows from (ii), we write Q in its singular value decomposition, $Q = \sum q_i |\pi_i\rangle\langle\lambda_i|$ with $q_i > 0$ and $(|\pi_i\rangle)$ and $(|\lambda_i\rangle)$ an appropriate pair of orthonormal and complete bases of $\Pi\mathscr{H}$ and $\Lambda\mathscr{H}$, respectively. Then $Q^2 = Q$ is equivalent to $\langle\lambda_k|\pi_k\rangle = q_k^{-1}$ and $\langle\lambda_i|\pi_j\rangle = 0$ for $i \neq j$. Since $\Lambda\Pi = \sum_i q_i^{-1} |\lambda_i\rangle\langle\pi_i|$, it immediately follows that Q is the Moore-Penrose inverse of $\Lambda\Pi$.

From the explicit form of the Moore-Penrose inverse, $Q = \sum_i \langle\lambda_i|\pi_i\rangle^{-1} |\pi_i\rangle\langle\lambda_i|$, one easily verifies the properties in (iv).

We have from (iv) that $Q\mathscr{H} = \Pi Q\mathscr{H} \subset \Pi\mathscr{H} = Q\Pi\mathscr{H} \subset Q\mathscr{H}$ and hence $Q\mathscr{H} = \Pi\mathscr{H}$ and analogously $\ker Q = (Q^\dagger\mathscr{H})^\perp = (\Lambda\mathscr{H})^\perp = \ker \Lambda$. Finally, $Q^2 = Q(\Pi Q) = (Q\Pi)Q = \Pi Q = Q$. □

B Appendix: Proof of Corollary 9

B.1 Necessity

Let us abbreviate $\bar{\mu} = 3 - \mu$, $Z_\mu = \Lambda_\mu Z \Lambda_\mu$ and $Y_{\bar{\mu}} = \Lambda_\mu Z \Lambda_{\bar{\mu}}$. From Theorem 8 it follows that for any optimal measurement we have $(\mu = 1, 2)$

$$Z_\mu - \Lambda_\mu \gamma_\mu \Lambda_\mu \geq 0, \tag{76a}$$

$$(Z_\mu - \Lambda_\mu \gamma_\mu \Lambda_\mu) E_\mu = 0, \tag{76b}$$

and

$$Z_\mu(\Lambda_\mu - E_\mu) = Y_{\bar{\mu}} E_{\bar{\mu}} \Lambda_\mu, \tag{77a}$$

$$Y_{\bar{\mu}}(\Lambda_{\bar{\mu}} - E_{\bar{\mu}}) = Z_\mu E_\mu \Lambda_{\bar{\mu}}, \tag{77b}$$

where the last two equations follow from $\Lambda_\mu Z E_? \Lambda_\nu = 0$ with $E_? = \mathbb{1} - E_\mu - E_{\bar\mu}$. From $Z \geq 0$ we have $Z_\mu \geq 0$ and $Y_{\bar\mu} = Y_\mu^\dagger$. We find

$$\begin{aligned} Z_\mu - \Lambda_\mu \gamma_\mu \Lambda_\mu &= (\Lambda_\mu - E_\mu)(Z_\mu - \gamma_\mu)(\Lambda_\mu - E_\mu) \\ &= \Lambda_\mu E_{\bar\mu} Y_{\bar\mu}^\dagger (\Lambda_\mu - E_\mu) - (\Lambda_\mu - E_\mu)\gamma_\mu(\Lambda_\mu - E_\mu) \\ &= \Lambda_\mu E_{\bar\mu} \gamma_{\bar\mu} E_{\bar\mu} \Lambda_\mu - (\Lambda_\mu - E_\mu)\gamma_\mu(\Lambda_\mu - E_\mu) \\ &= \Lambda_\mu E_? \gamma_{\bar\mu} E_? \Lambda_\mu - \Lambda_\mu E_? \gamma_\mu E_? \Lambda_\mu. \end{aligned} \tag{78}$$

Thus Eq. (26a), Eq. (26b) follow from Eq. (76a). From Eq. (76b) we have

$$\Lambda_\mu E_? (\gamma_{\bar\mu} - \gamma_\mu) E_? E_\mu = 0. \tag{79}$$

Combining Eq. (77a) and Eq. (77b) we obtain

$$Y_{\bar\mu} \Lambda_\mu = Z_\mu E_\mu \Lambda_{\bar\mu} \Lambda_\mu + Z_\mu (\Lambda_\mu - E_\mu) \tag{80}$$

and hence from Eq. (77a) and due to Eq. (76b),

$$\begin{aligned} \Lambda_{\bar\mu} Z_\mu (\Lambda_\mu - E_\mu) &= (Y_\mu \Lambda_{\bar\mu})^\dagger E_{\bar\mu} \Lambda_\mu \\ &= \Lambda_{\bar\mu} \Lambda_\mu E_{\bar\mu} \gamma_{\bar\mu} E_{\bar\mu} \Lambda_\mu + (\Lambda_{\bar\mu} - E_{\bar\mu}) \gamma_{\bar\mu} E_{\bar\mu} \Lambda_\mu \\ &= \Lambda_{\bar\mu} (\Lambda_\mu - E_\mu) Z_\mu (\Lambda_\mu - E_\mu) + (\Lambda_{\bar\mu} - E_{\bar\mu}) \gamma_{\bar\mu} E_{\bar\mu} \Lambda_\mu, \end{aligned} \tag{81}$$

where in the last step we used the result from Eq. (78). Thus we have found $\Lambda_{\bar\mu} E_\mu \gamma_\mu (\Lambda_\mu - E_\mu) = (\Lambda_{\bar\mu} - E_{\bar\mu}) \gamma_{\bar\mu} E_{\bar\mu} \Lambda_\mu$, i.e., Eq. (26c) follows.

This equation together with Eq. (79) for $\mu = 1$ and $\mu = 2$, finally yields Eq. (25b).

B.2 Sufficiency

We first get rid of the non-skew parts of $\mathcal{S}$:

Lemma 19. *With the definitions and preliminaries as in Proposition 6, if $E_?$ is a proper USD measurement of $\mathcal{S}$ and satisfies Eq.* (25) *for $\mathcal{S}$, then $E_?^{\text{skew}} = E_? + (\mathbb{1} - \Pi_{\text{skew}})$ is a proper USD measurement of $\mathcal{S}^{\text{skew}}$ and satisfies Eq.* (25) *for $\mathcal{S}^{\text{skew}}$.*

Hence, if we further show, that from the second part of the Lemma, it follows that $E_?^{\text{skew}}$ is optimal for $\mathcal{S}^{\text{skew}}$, then we have due to Proposition 6, that also $E_?$ is optimal for $\mathcal{S}$.

Proof of Lemma 19. We denote by Σ_1 the projector onto $\operatorname{supp}\gamma_1 \cap \ker\gamma_2$, by Σ_2 the projector onto $\operatorname{supp}\gamma_2 \cap \ker\gamma_1$ and by Π_μ the projector onto $\operatorname{supp}\gamma_\mu$. Then multiplication of Eq. (25a) by Σ_1 yields

$$\Sigma_1 E_? \gamma_2 E_? \Sigma_1 - \Sigma_1 E_? \gamma_1 E_? \Sigma_1 \geq 0. \tag{82}$$

Since for a USD measurement $\Pi_1 E_? \gamma_2 = \Pi_1 \gamma_2$, we have $\Sigma_1 \gamma_2 = 0$ and only the second term remains, which henceforth must vanish. This yields $\sqrt{\gamma_1} E_? \Sigma_1 = 0$ or equivalently $\Pi_1 E_? \Sigma_1 = 0$. A further multiplication from the left by Σ_1 together with the property $E_? \geq 0$ proofs that $E_? \Sigma_1 = 0$. An analogous argument can be used in order to show $E_?\Sigma_2 = 0$. Now, following the same lines of argument as in the proof of Proposition 6, it follows that $E_?^{\mathrm{skew}}$ is a proper USD measurement for $\mathcal{S}^{\mathrm{skew}}$.

From $E_?\Sigma_\mu = 0$ it in particular follows that $E_?(\gamma_2-\gamma_1)E_? = E_?^{\mathrm{skew}}\Pi_{\mathrm{skew}}(\gamma_2-\gamma_1)\Pi_{\mathrm{skew}}E_?^{\mathrm{skew}}$. Using

$$\ker \Pi_{\mathrm{skew}}\gamma_\mu\Pi_{\mathrm{skew}} = \ker \gamma_\mu(\mathbb{1} - \Sigma_\mu) = \Sigma_\mu\mathscr{H} \oplus \ker \gamma_\mu, \tag{83}$$

it is now straightforward to show that $E_?^{\mathrm{skew}}$ satisfies Eq. (25) for $\mathcal{S}^{\mathrm{skew}}$. □

It remains to consider the case where $\operatorname{supp}\gamma_1 \cap \ker\gamma_2 = \{0\} = \operatorname{supp}\gamma_2 \cap \ker\gamma_1$. We define R_μ to be the bijective oblique projector from $\ker\gamma_{\bar\mu}$ to $\ker\gamma_\mu$. Note that $R_\mu = R_{\bar\mu}^\dagger$. We furthermore denote by Q_μ the bijective oblique projector from $\ker\gamma_{\bar\mu} \cap \operatorname{supp}\mathcal{S}$ to $\operatorname{supp}\gamma_\mu \cap (\ker\gamma_1 + \ker\gamma_2)$. Then the multiplication of Eq. (25b) by $Q_\mu^\dagger$ from the left and by Q_μ from the right yields

$$\Lambda_\mu E_?(\gamma_{\bar\mu} - \gamma_\mu)E_?E_\mu = 0. \tag{84}$$

Let us define

$$V_\mu = \Lambda_\mu E_?(\gamma_{\bar\mu} - \gamma_\mu)E_?\Lambda_\mu + \Lambda_\mu\gamma_\mu\Lambda_\mu, \tag{85}$$
$$W_{\bar\mu} = (R_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu}) + \Lambda_\mu E_{\bar\mu})V_{\bar\mu}. \tag{86}$$

Then, using Eq. (26c),(84), we have

$$\begin{aligned}
V_\mu(\Lambda_\mu - E_\mu) &= \Lambda_\mu E_?(\gamma_{\bar\mu} - \gamma_\mu)E_?\Lambda_\mu + \Lambda_\mu\gamma_\mu E_?\Lambda_\mu \\
&= \Lambda_\mu E_?\gamma_{\bar\mu}E_?\Lambda_\mu + E_\mu\gamma_\mu E_?\Lambda_\mu \\
&= -\Lambda_\mu E_{\bar\mu}\gamma_{\bar\mu}E_?\Lambda_\mu - R_{\bar\mu}E_?\gamma_\mu E_?\Lambda_\mu \\
&= -(\Lambda_\mu E_{\bar\mu} + R_{\bar\mu}E_?)\gamma_{\bar\mu}E_?\Lambda_\mu \\
&= (\Lambda_\mu E_{\bar\mu} + R_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu}))\gamma_{\bar\mu}E_{\bar\mu}\Lambda_\mu \\
&= W_{\bar\mu}E_{\bar\mu}\Lambda_\mu.
\end{aligned} \tag{87}$$

And a similar equation holds for $W_{\bar\mu}$:

$$\begin{aligned}
W_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu}) =& R_{\bar\mu}E_?(\gamma_\mu - \gamma_{\bar\mu})E_?\Lambda_{\bar\mu} + R_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu})\gamma_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu}) \\
&+ \Lambda_\mu E_{\bar\mu}\gamma_{\bar\mu}(\Lambda_{\bar\mu} - E_{\bar\mu}) \\
=& R_{\bar\mu}E_?\gamma_\mu E_?\Lambda_{\bar\mu} - \Lambda_\mu E_?\gamma_\mu E_?\Lambda_{\bar\mu} \\
=& -(-E_\mu - (\Lambda_\mu - E_\mu))\gamma_\mu E_\mu\Lambda_{\bar\mu} \\
=& V_\mu E_\mu\Lambda_{\bar\mu}.
\end{aligned} \tag{88}$$

We combine these two equations and find $W_{\bar\mu} = V_\mu(E_\mu\Lambda_{\bar\mu} + (\Lambda_\mu - E_\mu)R_{\bar\mu})$, i.e., in comparison with the definition in Eq. (86), $W_{\bar\mu} = W_\mu^\dagger$. Furthermore, we have $W_1V_1^-V_1 = W_1$ and $V_2 = W_1V_1^-W_1^\dagger$, where V_1^- denotes the inverse of V_1 on its support. (The second identity follows from $W_1V_1^-W_1^\dagger = W_1(\Lambda_1 - E_1)R_2 + W_1E_1\Lambda_2$.)

We construct the operator Z as $Z = TV_1T^\dagger$ with $T = Q_1 + Q_2W_1V_1^-$. Since $V_1 \geq 0$ we have $Z \geq 0$. From our previous considerations, $\Lambda_\mu Z\Lambda_\mu = V_\mu$ and $\Lambda_\mu Z\Lambda_{\bar\mu} = W_{\bar\mu}$ follows. Then using Eq. (26a), Eq. (26b), and Eq. (84) it is now straightforward to verify that Z satisfies the conditions in Eq. (24b) of Theorem 8.

It remains to show that $ZE_? = 0$. First, with $\Pi_\|$ the projector onto $\operatorname{supp}\gamma_1 \cap \operatorname{supp}\gamma_2$, we have $ZE_?\Pi_\| = Z\Pi_\| = 0$ due to $\Pi_\| Q_\mu = 0$. Analogously with $\Pi_\perp$ the projector onto $\ker \mathcal{S}$, $ZE_?\Pi_\perp = Z\Pi_\perp = 0$. Thus we only need to show that $\Lambda_\mu Z E_? \Lambda_\nu = 0$. But this follows from Eq. (87) and Eq. (88).

C Appendix: Construction of Jordan bases

In this Appendix an explicit construction of Jordan bases (also sometimes called *canonical bases*) of two subspaces $\mathscr{A}$ and $\mathscr{B}$ of $\mathbb{C}^d$ is given (cf. also Ref. [33, 9, 19, 18]). Jordan bases are orthonormal bases $(|a_i\rangle)$ of $\mathscr{A}$ and $(|b_j\rangle)$ of $\mathscr{B}$, such that $\langle a_i|b_j\rangle = 0$ for all $i \neq j$ and $\langle a_k|b_k\rangle \geq 0$ for all k. With $S_\mathscr{A}$ we denote the $d \times n_\mathscr{A}$ dimensional matrix where the columns are given by the $n_\mathscr{A}$ basis vectors of some orthonormal basis of $\mathscr{A}$. Analogously we define $S_\mathscr{B}$ and $n_\mathscr{B}$ by using $\mathscr{B}$.

Consider a singular value decomposition of

$$S_\mathscr{A}^\dagger S_\mathscr{B} = U_\mathscr{A} D U_\mathscr{B}^\dagger, \quad \text{i.e.,} \quad (S_\mathscr{A} U_\mathscr{A})^\dagger (S_\mathscr{B} U_\mathscr{B}) = D, \tag{89}$$

where D is a $n_\mathscr{A} \times n_\mathscr{B}$ dimensional diagonal matrix. $U_\mathscr{A}$ and $U_\mathscr{B}$ are unitary matrices. Let us denote the ith column of $S_\mathscr{A} U_\mathscr{A}$ by $|a_i\rangle$ and the jth column of $S_\mathscr{B} U_\mathscr{B}$ by $|b_j\rangle$. Then $(|a_i\rangle)$ and $(|b_j\rangle)$ are Jordan bases of $\mathscr{A}$ and $\mathscr{B}$.

References

[1] C. W. Helstrøm, Quantum Detection and Estimation Theory, Acad. Press, New York, 1976.

[2] U. Herzog, J. A. Bergou, Distinguishing mixed quantum states: Minimum-error discrimination versus optimum unambiguous discrimination, Phys. Rev. A 70 (2004) 022302.

[3] I. Ivanovic, How to differentiate between non-orthogonal states, Phys. Lett. A 123 (1987) 257–259.

[4] D. Dieks, Overlap and distinguisability of quantum states, Phys. Lett. A 126 (1988) 303–306.

[5] A. Peres, How to differentiate between non-orthogonal states, Phys. Lett. A 128 (1988) 19.

[6] G. Jaeger, A. Shimony, Optimal distinction between two non-orthogonal quantum states, Phys. Lett. A 197 (1995) 83–87.

[7] C. H. Bennett, T. Mor, J. A. Smolin, Parity bit in quantum cryptography, Phys. Rev. A 54 (4) (1996) 2675–2684.

[8] Y. Sun, J. A. Bergou, M. Hillery, Optimum unambiguous discrimination between subsets of nonorthogonal quantum states, Phys. Rev. A 66 (3) (2002) 032315.

[9] T. Rudolph, R. W. Spekkens, P. S. Turner, Unambiguous discrimination of mixed states, Phys. Rev. A 68 (1) (2003) 010301(R).

[10] J. A. Bergou, U. Herzog, M. Hillery, Quantum filtering and discrimination between sets of boolean functions, Phys. Rev. Lett. 90 (25) (2003) 257901.

[11] P. Raynal, N. Lütkenhaus, S. J. van Enk, Reduction theorems for optimal unambiguous state discrimination of density matrices, Phys. Rev. A 68 (2) (2003) 022308.

[12] J. A. Bergou, U. Herzog, M. Hillery, Optimal unambiguous filtering of a quantum state: An instance in mixed state discrimination, Phys. Rev. A 71 (2005) 042314.

[13] U. Herzog, J. A. Bergou, Optimum unambiguous discrimination of two mixed quantum states, Phys. Rev. A 71 (5) (2005) 050301(R).

[14] P. Raynal, N. Lütkenhaus, Optimal unambiguous state discrimination of two density matrices: Lower bound and class of exact solutions, Phys. Rev. A 72 (2005) 022342, 049909(E).

[15] U. Herzog, Optimum unambiguous discrimination of two mixed states and application to a class of similar states, Phys. Rev. A 75 (5) (2007) 052309.

[16] P. Raynal, N. Lütkenhaus, Optimal unambiguous state discrimination of two density matrices: A second class of exact solutions, Phys. Rev. A 76 (2007) 052322.

[17] Y. C. Eldar, M. Stojnic, B. Hassibi, Optimal quantum detectors for unambiguous detection of mixed states, Phys. Rev. A 69 (2004) 062318.

[18] J. A. Bergou, E. Feldman, M. Hillery, Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination, Phys. Rev. A 73 (3) (2006) 032107.

[19] M. Kleinmann, H. Kampermann, P. Raynal, D. Bruß, Commutator relations reveal solvable structures in unambiguous state discrimination, Journal of Physics A: Mathematical and Theoretical 40 (36) (2007) F871–F878.

[20] L. P. Hughston, R. Jozsa, W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Phys. Lett. A 183 (1993) 14–18.

[21] A. Bassi, G. Ghirardi, A general scheme for ensemble purification, Phys. Lett. A 309 (2003) 24–28.

[22] M. Kleinmann, H. Kampermann, T. Meyer, D. Bruß, Physical purification of quantum states, Phys. Rev. A 73 (2006) 062309.

[23] A. Uhlmann, The "transition probability" in the state space of a *-algebra, Rep. Math. Phys. 9 (1976) 273–279.

[24] R. Jozsa, Fidelity for mixed quantum states, J. Mod. Opt. 41 (1994) 2315–2323.

[25] X.-F. Zhou, Y.-S. Zhang, G.-C. Guo, Unambiguous discrimination of mixed states: A description based on system-ancilla coupling, Phys. Rev. A 75 (5) (2007) 052314.

[26] M. Kleinmann, H. Kampermann, D. Bruß, Generalization of quantum-state comparison, Phys. Rev. A 72 (3) (2005) 032308.

[27] A. Chefles, Unambiguous discrimination between linearly independent quantum states, Phys. Lett. A 239 (1998) 339–347.

[28] A. Chefles, S. M. Barnett, Optimum unambiguous discrimination between linearly independent symmetric states, Phys. Lett. A 250 (1998) 223–229.

[29] A. Peres, D. R. Terno, Optimal distinction between non-orthogonal quantum states, J. Phys. A: Math. Gen. 31 (1998) 7105–7111.

[30] C. Zhang, Y. Feng, M. Ying, Unambiguous discrimination of mixed quantum states, Phys. Lett. A 353 (2006) 300–306.

[31] S. N. Afriat, Orthogonal and oblique projectors and the characteristics of pairs of vectors spaces, Proc. Camb. Philos. Soc. 53 (1957) 800–816.

[32] T. N. E. Greville, Solutions of the matrix equation $XAX = X$, and relations between oblique and orthogonal projectors, SIAM J. Appl. Math. 26 (4) (1974) 828–832.

[33] G. W. Stewart, J.-g. Sun, Matrix Pertubation Theory, Acad. Press, San Diego, 1990.

Danksagungen

An erster Stelle möchte ich Frau Prof. Dr. D. Bruß ganz herzlich danken, dass sie mir die Möglichkeit gegeben hat, in ihrer Gruppe zu arbeiten, für die große Freiheit die ich genossen habe, die vielen Diskussionen und Anregungen und für die konstruktive Kritik und aufmunternden Worte. Zu besonderem Dank bin ich auch Herrn Dr. H. Kampermann verpflichtet, der in unzähligen Diskussionen, mit seiner Hartnäckigkeit und mit seinem unzerstörbaren Optimismus ganz wesentlich zu der Entstehung dieser Arbeit beigetragen hat. Für eine konstruktive Zusammenarbeit möchte ich mich auch bei Herrn Dr. T. Meyer und Herrn Dr. Ph. Raynal bedanken.

Darüber hinaus möchte ich mich bei allen der gesamten Arbeitsgruppe bedanken, die mir immer wieder bei mehr oder weniger schlauen Fragen zu Seite gestanden haben und den Alltag im Institut so angenehm gemacht haben, neben den oben genannten also auch Frau Z. Shadman, Herrn Dr. R. Unanyan und Herrn P. Skwara.

Besonders lieber Dank gilt meinen Eltern und meinen Geschwistern, aber auch meinen guten Freunden und Bekannten für das Verständnis und den stetigen Rückhalt, den sie mir gegeben haben.